## Handbook of Biomedical Image Analysis

# Registration Models

Jasjit S. Suri David L. Wilson Swamy Laxminarayan



# Handbook of Biomedical Image Analysis

#### TOPICS IN BIOMEDICAL ENGINEERING INTERNATIONAL BOOK SERIES

Series Editor: Evangelia Micheli-Tzanakou Rutgers University Piscataway, New Jersey

Signals and Systems in Biomedical Engineering: Signal Processing and Physiological Systems Modeling Suresh R. Devasahayam

Models of the Visual System Edited by George K. Hung and Kenneth J. Ciuffreda

PDE and Level Sets: Algorithmic Approaches to Static and Motion Imagery Edited by Jasjit S. Suri and Swamy Laxminarayan

Frontiers in Biomedical Engineering: Proceedings of the World Congress for Chinese Biomedical Engineers Edited by Ned H.C. Hwang and Savio L-Y. Woo

Handbook of Biomedical Image Analysis: Volume I: Segmentation Models Part A Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

Handbook of Biomedical Image Analysis: Volume II: Segmentation Models Part B Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

Handbook of Biomedical Image Analysis: Volume III: Registration Models Edited by Jasjit S. Suri, David L. Wilson, and Swamy Laxminarayan

A Continuation Order Plan is available for this series. A continuation order will bring delivery of each new volume immediately upon publication. Volumes are billed only upon actual shipment. For further information please contact the publisher.

# Handbook of Biomedical Image Analysis

Volume III: Registration Models

Edited by

Jasjit S. Suri

Department of Biomedical Engineering Case Western Reserve University Cleveland, Ohio

## David L. Wilson

Department of Biomedical Engineering Case Western Reserve University Cleveland, Ohio

and

## Swamy Laxminarayan

Institute of Rural Health Idaho State University Pocatello, Idaho

Kluwer Academic/Plenum Publishers New York, Boston, Dordrecht, London, Moscow

ISBN 0-306-48607-5 eISBN 0-306-48608-3 set ISBN: 0-387-23126-9

©2005 Kluwer Academic/Plenum Publishers, New York 233 Spring Street, New York, New York 10013

http://www.wkap.nl/

 $10 \quad 9 \quad 8 \quad 7 \quad 6 \quad 5 \quad 4 \quad 3 \quad 2 \quad 1$ 

A C.I.P. record for this book is available from the Library of Congress

All rights reserved

No part of this book may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording, or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Permissions for books published in Europe: *permissions@wkap.nl* Permissions for books published in the United States of America: *permissions@wkap.com* 

Printed in the United States of America

Jasjit Suri dedicates this handbook to his youngest uncle Paramjeet Chadha and his immediate family: his late sister Sharan, his late brother Amarjeet, and his late parents Kulwant Kaur and Udam S. Chadha (Fellow of Royal Institute of London).

David Wilson dedicates this handbook to his family and students.

Swamy Laxminarayan dedicates this book in memory of his beloved parents who were a constant source of inspiration in his life and to his in-laws Corie and Derk Zwakman for their genuine sense of family attachments and friendship.

## Contributors

Jasjit S. Suri, Ph.D. Case Western Reserve University, Cleveland, OH, USA

**David L. Wilson, Ph.D.** Case Western Reserve University, Cleveland, OH, USA

**Baowei Fei, Ph.D** Case Western Reserve University, Cleveland, OH, USA

**Swamy Laxminarayan, DSc.** State University of Idaho, Pocatello, ID, USA

**Andrew Laine, Ph.D.** Columbia University, New York, NY, USA

**Elsa Angelini, Ph.D.** Columbia University, New York, NY, USA

**Yinpeng Jin, Ph.D.** Columbia University, New York, NY, USA

**Aly A. Farag, Ph.D.** University of Louisville, Louisville, KY, USA Sameh M. Yamany, Ph.D. University of Louisville, Louisville, KY, USA

Jeremy Nett, M.S. University of Louisville, Louisville, KY, USA

**Thomas Moriarty, M.D., Ph.D.** University of Louisville, Louisville, KY, USA

**Stephen Hushek, Ph.D.** University of Louisville, Louisville, KY, USA

**Robert Falk, M.D.** University of Louisville, Louisville, KY, USA

**A. Ardeshir Goshtasby, Ph.D.** Wright State University, Dayton, OH, USA

**Pierre Hellier, Ph.D.** IRISA-INRIA, Cedex, France

Michael Unser, Ph.D. Swiss Federal Institute of Technology, Lausanne Lausanne, Switzerland

**Contributors** 

**Jan Kybic, Ph.D.** Czech Technical University, Prague, Czech Republic

**Jeff Weiss, Ph.D.** University of Utah, Utah, UT, USA

Nikhil Phatak, Ph.D. University of Utah, Utah, UT, USA

Alex Veress, Ph.D. University of Utah, Utah, UT, USA

Xiao-Hong Zhu, M.S. University Name, City, State, USA

Yang Ming Zhu, Ph.D. Philips Medical Systems, Cleveland, OH, USA

**Steven M. Cochoff, M.S.** Philips Medical Systems, Cleveland, OH, USA

**Torsten Rohlfing, Ph.D.** Stanford University, Stanford, CA, USA

**Robert Brandt, Ph.D.** Visual Concepts GmbH, Berlin, Germany **Randolf Menzel, Ph.D.** Freie Universität Berlin, Berlin, Germany

**Daniel B. Russakoff, M.Sc.** Stanford University, Stanford, CA, USA

**Calvin R. Maurer, Jr., Ph.D.** Stanford University, Stanford, CA, USA

**Gary E Christensen, DSc.** The University of Iowa, Iowa, USA

**Jundong Liu, Ph.D.** Ohio University, Athens, OH

**Chi-Hsiang Lo, Ph.D.** National Ilan University, Ilan, Taiwan

Yujun Guo, M.S. Kent State University, Kent, OH, USA

**Cheng-Chang Lu, Ph.D.** Kent State University, Kent, OH, USA

**Chi-Hua Tung, Ph.D.** Philips Medical Systems, Cleveland, USA

viii

## Acknowledgments

This book is the result of collective endeavors from several noted engineering and computer scientists, mathematicans, medical doctors, physicists, and radiologists. The editors are indebted to all of their efforts and outstanding scientific contributions. The editors are particularly grateful to Drs. Petia Reveda, Alex Falco, Andrew Laine, David Breen, David Chopp, C. C. Lu, Gary E. Christensen, Dirk Vandermeulen, Aly A Farag, Alejandro Frangi, Gilson Antonio Giraldi, Gabor Szekely, Pierre Hellier, Gabor Herman, Ardeshir Goshtasby, Jan Kybic, Jeff Weiss, Jean-Claude Klein, Majid Mirmehdi, Maria Kallergi, Yang-ming Zhu, Sunanda Mitra, Sameer Singh, Alessandro Sarti, Xioping Shen, Calvin R. Maurer, Jr., Yoshinobu Sato, Koon-Pong Wong, Avdhesh Sharma, Rakesh Sharma, and Chun Yuan and their team members for working with us so closely in meeting all of the deadlines of the book. We would like to express our appreciation to Kluwer Publishers for helping create this invitational handbook. We are particularly thankful to Aaron Johnson, the acquisition editor, and Shoshana Sternlicht for their excellent coordination of the book at every stage.

Dr. Suri thanks Philips Medical Systems, Inc., for the MR datasets and encouragements during his experiments and research. Special thanks are due to Dr. Larry Kasuboski and Dr. Elaine Keeler from Philips Medical Systems, Inc., for their support and motivations. Thanks are also due to my past Ph.D. committee research professors, particularly Professors Linda Shapiro, Robert M. Haralick, Dean Lytle, and Arun Somani, for their encouragements.

We extend our appreciations to Drs. Ajit Singh, Siemens Medical Systems, George Thoma, Chief Imaging Science Division from National Institutes of Health, Dr. Sameer Singh, University of Exeter, UK for his motivations.

Acknowledgments

Special thanks go to the Book Series Editor, Professor Evangelia Micheli-Tzanakou, for advising us on all aspects of the book.

We thank the IEEE Press, Academic Press, Springer Verlag Publishers, and several medical and engineering journals for o permitting us to use some of the images previously published in these journals.

Finally, Jasjit Suri thanks his wife Malvika Suri for all the love and support she has showered over the years and to our baby Harman whose presence is always a constant source of pride and joy. I also express my gratitude to my father, a mathematician, who inspired me throughout my life and career, and to my late mother, who most unfortunately passed away a few days before my Ph.D. graduation, and who so much wanted to see me write this book. I love you, Mom. Special thanks to Pom Chadha and his family, who taught me life is not just books. He is my of my best friends. I would like to also thank my in-laws who have a special place for me in their hearts and have shown lots of love and care for me.

David Wilson acknowledges the support of the Department of Biomedical Engineering, Case Western Reserve University, in this endeavor. Special thanks are due to the many colleagues and students who make research in biomedical engineering an exciting, wondrous endeavor.

Swamy Laxminarayan expresses his loving acknowledgements to his wife Marijke and to his kids, Malini and Vinod, for always giving the strength of mind amidst all life frustrations. The book kindles fondest memories of my late parents who made many personal sacrifices that helped shape our careers and the support of my family members who were always there for me when I needed them most. I have shared many ideas and thoughts on the book with numerous of my friends and colleagues in the discipline. I acknowledge their friendship, feedbacks, and discussions with particular thanks to Prof. David Kristol of the New Jersey Institute of Technology, Peter Brett of Ashton University, Ewart Carson of the City University, London, Laura Roa of the University of Sevilla in Spain, and Jean Louis Coatrieux of the University of Rennes in France, for their constant support over the past two decades.

## Preface

Our goal is to develop automated methods for the segmentation of threedimensional biomedical images. Here, we describe the segmentation of confocal microscopy images of bee brains (20 individuals) by registration to one or several atlas images. Registration is performed by a highly parallel implementation of an entropy-based nonrigid registration algorithm using B-spline transformations. We present and evaluate different methods to solve the correspondence problem in atlas based registration. An image can be segmented by registering it to an individual atlas, an average atlas, or multiple atlases. When registering to multiple atlases, combining the individual segmentations into a final segmentation can be achieved by atlas selection, or multiclassifier decision fusion. We describe all these methods and evaluate the segmentation accuracies that they achieve by performing experiments with electronic phantoms as well as by comparing their outputs to a manual gold standard.

The present work is focused on the mathematical and computational theory behind a technique for deformable image registration termed Hyperelastic Warping, and demonstration of the technique via applications in image registration and strain measurement. The approach combines well-established principles of nonlinear continuum mechanics with forces derived directly from threedimensional image data to achieve registration. The general approach does not require the definition of landmarks, fiducials, or surfaces, although it can accommodate these if available. Representative problems demonstrate the robust and flexible nature of the approach.

Three-dimensional registration methods are introduced for registering MRI volumes of the pelvis and prostate. The chapter first reviews the applications,

challenges, and previous methods of image registration in the prostate. Then the chapter describes a three-dimensional mutual information rigid body registration algorithm with special features. The chapter also discusses the threedimensional nonrigid registration algorithm. Many interactively placed control points are independently optimized using mutual information and a thin plate spline transformation is established for the warping of image volumes. Nonrigid method works better than rigid body registration whenever the subject position or condition is greatly changed between acquisitions.

This chapter will cover 1D, 2D, and 3D registration approaches both rigid and elastic. Mathematical foundation for surface and volume registration approaches will be presented. Applications will include plastic surgery, lung cancer, and multiple sclerosis.

Flow-mediated dilation (FMD) offers a mechanism to characterize endothelial function and therefore may play a role in the diagnosis of cardiovascular diseases. Computerized analysis techniques are very desirable to give accuracy and objectivity to the measurements. Virtually all methods proposed up to now to measure FMD rely on accurate edge detection of the arterial wall, and they are not always robust in the presence of poor image quality or image artifacts. A novel method for automatic dilation assessment based on a global image analysis strategy is presented. We model interframe arterial dilation as a superposition of a rigid motion model and a scaling factor perpendicular to the artery. Rigid motion can be interpreted as a global compensation for patient and probe movements, an aspect that has not been sufficiently studied before. The scaling factor explains arterial dilation. The ultrasound (US) sequence is analyzed in two phases using image registration to recover both transformation models. Temporal continuity in the registration parameters along the sequence is enforced with a Kalman filter since the dilation process is known to be a gradual physiological phenomenon. Comparing automated and gold standard measurements we found a negligible bias (0.04%) and a small standard deviation of the differences (1.14%). These values are better than those obtained from manual measurements (bias = 0.47%, SD = 1.28%). The proposed method offers also a better reproducibility (CV = 0.46%) than the manual measurements (CV =1.40%).

This chapter will focus on nonrigid registration techniques. Nonrigid registration is needed to correct for deformations that occur in various contexts: respiration or organ motion, disease progression over time, tissue deformation

#### Preface

due to surgical procedure, intersubject comparison to build anatomical atlases, etc. Numerous registration techniques have been developed and can be broadly decomposed into intensity-based (photometric) and landmark-based (geometrical) techniques. This chapter will present up-to-date methods.

This chapter will then present how segmentation and registration methods can cooperate: accurate and fast segmentation can be obtained using nonrigid registration; nonrigid registration methods can be constrained by segmentation methods. Results of these cooperation schemes will be given.

This chapter will finally be concerned with validation of nonrigid registration methods. More specifically, an objective evaluation framework will be presented in the particular context of intersubject registration.

This chapter concerns elastic image registration for biomedical applications. We start with an overview and classification of existing registration techniques. We revisit the landmark interpolation and add some generalisations. We develop a general elastic image registration algorithm. It uses a grid of uniform B-splines to describe the deformation. It also uses B-splines for image interpolation. Multiresolution in both image and deformation model spaces yields robustness and speed. We show various applications of the algorithm on MRI, CT, SPECT and ultrasound data. A semiautomatic version of the registration algorithm is capable of accepting expert hints in the form of soft landmark constraints.

The chapter will include algorithms based on landmark and intensity-based image registration. It will compare traditional unidirectional registration algorithms to those that are bidirectional and minimize the inverse consistency error. It will discuss how small deformation models can nonrigidly be used for medical image registration in the brain, skull, and inner ear. It will also discuss how to extend the small deformation model to the large deformation model to accommodate locally large deformation image registration problems. We will provide examples using phantom images and brain images to demonstrate the large deformation case.

## Contents

1.	Medical Image Registration: Theory, Algorithm, and Case Studies in Surgical Simulation, Chest Cancer, and Multiple Sclerosis	1
	Aly A. Farag, Sameh M. Yamany, Jeremy Nett, Thomas Moriarty,	
	Ayman El-Baz, Stephen Hushek, and Robert Falk	
2.	State of the Art of Level Set Methods in Segmentation and Registration of Medical Imaging Modalities	47
	Elsa Angelini, Yinpeng Jin, and Andrew Laine	
3.	Three-Dimensional Rigid and Non-Rigid Image Registration for the Pelvis and Prostate	103
	Baowei Fei, Jasjit Suri, and David L. Wilson	
4.	Stereo and Temporal Retinal Image Registration by Mutual Information Maximization	151
	Xiao-Hong Zhu, Cheng-Chang Lu, and Yang-Ming Zhu	
5.	Quantification of Brain Aneurysm Dimensions from CTA for Surgical Planning of Coiling Interventions	185
	Monica Hemandez, Alejandro P. Prangi, and Gumernio Sapiro	• • •
6.	G. E. Christensen	219
7.	A Computer-Aided Design System for Segmentation of Volumetric Images	251
	Marcel Jackowski and Ardeshir Goshfasby	
8.	Inter-subject Non-rigid Registration: An Overview with Classification and the Romeo Algorithm	273
	Pierre Heiller	

#### Contents

9.	Elastic Registration for Biomedical Applications	339
	Jan Kybic and Michael Unser	
10.	Cross-Entropy, Reversed Cross-Entropy, and Symmetric Divergence Similarity Measures for 3D Image Registration: A Comparative Study Yang-Ming Zhu and Steven M. Cochoff	393
11.	Quo Vadis, Atlas-Based Segmentation?	435
	Torsten Rohlfing, Robert Brandt, Randolf Menzel, Daniel B. Russakoff, and Calvin R. Maurer, Jr.	
12.	Deformable Image Registration with Hyperelastic Warping	487
	Alexander I. Veress, Nikhil Phatak, and Jeffrey A. Weiss	
13.	Future of Image Registration	535
	Jasjit Suri, Baowei Fei, David Wilson, Swamy Laxminarayan,	
	Chi-Hsiang Lo, Yujun Guo, Cheng-Chang Lu, and Chi-Hua Tung	
The	Editors	555
Inde	ех	561

xvi

## **Chapter 1**

## Medical Image Registration: Theory, Algorithm, and Case Studies in Surgical Simulation, Chest Cancer, and Multiple Sclerosis

Aly A. Farag,<sup>1</sup> Sameh M. Yamany,<sup>2</sup> Jeremy Nett,<sup>1</sup> Thomas Moriarty,<sup>3</sup> Ayman El-Baz,<sup>1</sup> Stephen Hushek,<sup>4</sup> and Robert Falk<sup>5</sup>

## **1.1 Introduction**

Registration found its application in medical imaging due to the fact that physicians are frequently confronted with the practical problem of registering medical images. Medical registration techniques have recently been extended to relate multimodal images which makes it possible to superimpose features from different imaging studies. Registration techniques have been also used in stereotactic surgery and stereotactic radiosurgery that require images to be registered with the physical space occupied by the patient during surgery. New interactive, image-guided surgery techniques use image-to-physical space registration to track the changing surgical position on a display of the preoperative image sets of the patient. In such applications, the speed of convergence of the registration technique is of major importance.

<sup>&</sup>lt;sup>1</sup>Computer Vision and Image Processing Laboratory, Department of Electrical and Computer Engineering, University of Louisville, Louisville, KY 40292, USA

 $<sup>^2\</sup>mathrm{System}$  and Biomedical Engineering Department, Cairo University, Giza, Egypt

<sup>&</sup>lt;sup>3</sup>Department of Neurological Surgery, University of Louisville, KY 40292, USA

<sup>&</sup>lt;sup>4</sup>MRI Department, Norton Hospital, Louisville, KY, USA

<sup>&</sup>lt;sup>5</sup>Medical Imaging Division, Jewish Hospital, Louisville, KY, USA

	Vactor function denoting a point on the model surface
X	Vector function denoting a point on the model surface
y G	vector function denoting a point on the experimental surface
S	General surface
Т _	Transformation matrix
R	Rotation matrix
t	Translation vector
$d(\mathbf{y}_i, S)$	The distance of point $\mathbf{y}_i$ to shape S
$\mathcal{F}()$	Registration objective function
$\mathcal{C}($ )	The closet point operator
GCP()	Grid Closest Point transform
${\mathcal G}$	3D space subset
$ec{r}$	Displacement vector in the GCP grid
$\{R, C, H\}$	Coordinates of the GCP grid
δ	Grid resolution
$C_{ijk}$	Grid cell
$c_0{}^{ijk}$	Centroid of the cell $C_{ijk}$
$\alpha$ , $\beta$ , and $\gamma$	3D-angles of rotations
$\theta$	Simplex mesh angle
P	3D point on a free-form surface
H	Mean curvature of the surface
$ec{U}_P$	Normal vector at point $P$
A	Set of landmarks
λ	Curvature threshold
$E_n^2$	Matching value
0	Overlap ratio
s	Scale factor
F	A medical volume
h()	Entropy function
Rf	A reference medical volume.

 Table 1.1: Most important nomenclature used throughout the chapter

Another example of the use of medical image registration is in neurosurgery where it is useful to identify tumors with magnetic resonance images (MRI), yet the established stereotaxy technology uses computed tomography (CT) images. Being able to register these two modalities allows one to transfer the coordinates of tumors from the MR images into the CT stereotaxy. It is similarly useful to transfer functional information from SPECT or positron-emission tomography (PET) into MR or CT for anatomical reference, and for stereotactic exploitation.

The currently used imaging modalities can be generally divided into two main categories, one related to the anatomy being imaged and the other to the functionality represented in the image. The first one includes X-ray, CT (computed tomography), MRI (magnetic resonance imaging), US (ultrasound), portal images, and (video) sequences obtained by various catheter scopes, e.g., by laparoscopy or laryngoscopy. Other prominent derivative techniques include, MRA (magnetic resonance angiography), DSA (digital subtraction angiography, derived from X-ray), CTA (computed tomography angiography), and Doppler (derived from US, referring to the Doppler effect measured). Functional modalities include (planar) scintigraphy, SPECT (single photon emission computed tomography), PET (positron emission tomography), which together make up the nuclear medicine imaging modalities, and fMRI (functional MRI).

An eminent example of the use of registering different modalities can be found in the area of epilepsy surgery [1]. Patients may undergo various MR, CT, and DSA studies for anatomical reference; ictal and interictal SPECT studies; MEG and extra and/or intra-cranial (subdural or depth) EEG, as well as PET studies. Registration of the images from practically any combination will benefit the surgeon. A second example concerns radiotherapy treatment, where both CT and MR can be employed. The former is needed to accurately compute the radiation dose, while the latter is usually better suited for delineation of tumor tissue.

Yet, a more prominent example is the use of medical registration for the same modality, i.e., monomodale registration. For example, in the qualitative evaluation of multiple sclerosis (MS) studies, where multiple MRI scans of the same patient at different times must be compared with one another. Due to the largely arbitrary positioning of the anatomy in the scanner, in a slice-by-slice comparison between studies, quite different anatomy can by chance be located on the same slice numbers in different studies. The goal of registration, therefore, is to align the anatomy from one scan, to the anatomy from another.

Medical registration spans numerous applications and there exists a large score of different techniques reported in the literature. In what follows is an attempt to classify these different techniques and categorize them based on some criteria.

## **1.2 Medical Registration Classifications**

The classification of registration methods used in this chapter is based on the criteria formulated by van den Elsen, Pol and Viergever [2]. Maintz and Viergever

[1] provided a good survey on different classification criteria. In this section we will summarize seven basic classification criteria commonly used (for more details and further reading see the Maintz and Viergever review).

The seven criteria are:

- 1. Dimensionality
- 2. Nature of registration basis
- 3. Nature of transformation
- 4. Interaction
- 5. Optimization procedure
- 6. Modalities involved
- 7. Subject

#### 1.2.1 Dimensionality

The main division here is either the scope of the registration involves *spatial dimension only* or is *time series* also involved. For spacial registration, there are the (i) 3D/3D registration where two or more volumes of interest are to be aligned together, the (ii) 2D/2D registration where two medical images are to be aligned together. In general, 2D/2D registration is less complex than the 3D/3D registration. A more complex one is the (iii) 2D/3D registration which involves the direct alignment of spatial data to projective data (e.g., a preoperative CT image to an intraoperative X-ray image), or the alignment of a single tomographic slice to spatial data. Time can be another dimension involved when the patient's images and volumes are to be tracked with time for analysis or monitoring.

#### 1.2.2 Nature of Registration Basis

In this category registration can be divided into *extrinsic*, i.e., based on foreign objects introduced into the imaged space, and *intrinsic* methods, i.e., based on the image information as generated by the patient. Extrinsic methods rely on artificial objects attached to the patient, objects which are designed to be well visible and accurately detectable in all of the pertinent modalities. As such,

the registration of the acquired images is comparatively easy, fast, can usually be automated, and, since the registration parameters can often be computed explicitly, has no need for complex optimization algorithms. The main drawback of extrinsic registration is that, for good accuracy, *invasive* maker (e.g., stereotactic frame or screw markers) objects are used. *Non-invasive* markers (e.g., skin markers individualized foam moulds, head holder frames, or dental adapters) can be used, but as a rule are less accurate.

Intrinsic registration can rely on *landmarks* in the images or volumes to be aligned. These landmarks can be *anatomical* based on morphological points on some visible anatomical organ(s), or pure *geometrical* based. Intrinsic registration can also be based on segmentation results. Segmentation in this case can be *rigid* where anatomically the same structures (mostly surfaces) are extracted from both images to be registered, and used as sole input for the alignment procedure. They can also be *deformable model based* where an extracted structure (also mostly surfaces, and curves) from one image is elastically deformed to fit the second image. The rigid model based approaches are probably the most popular methods due to its easy implementation and fast results. A drawback of rigid segmentation based methods is that the registration accuracy is limited to the accuracy of the segmentation step. In theory, rigid segmentation based registration areas have largely been limited to neuroimaging and orthopedic imaging.

Another example of intrinsic registration are the *voxel based* registration methods that operate directly on the image gray values, without prior data reduction by the user or segmentation. There are two distinct approaches: the first is to immediately reduce the image gray value content to a representative set of scalars and orientations, the second is to use the full image content throughout the registration process. Principal axes and moments based methods are the prime examples of reductive registration methods. Within these methods the image center of gravity and its principal orientations (principal axes) are computed from the image zeroth and first order moments. Registration is then performed by aligning the center of gravity and the principal orientations. The result is usually not very accurate, and the method is not equipped to handle differences in scanned volume well. Despite its drawbacks, principal axes methods are widely used in registration problems that require no high accuracy, because of the automatic and very fast nature of its use, and the easy implementation. On the other hand, voxel based registration using full image content is more

flexible, yet more complex and the methods range from using cross correlation, variance minimization, histogram clustering and the famous maximization of mutual information (discussed later in details).

#### 1.2.3 Nature of Transformation

Since the registration process tries to recover the optimal transformation between two candidate subjects, the nature of such transformation categorize the registration procedure to be used. The most commonly used is the *rigid* registration where the transformation involves only translations and rotations. If the transformation maps parallel lines onto parallel lines it is called *affine*. If it maps lines onto lines, it is called *projective*. Finally, if it maps lines onto curves, it is called *curved* or *elastic*. Each type of transformation contains as special cases the ones described before it, e.g., the rigid transformation is a special kind of affine transformation. A composition of more than one transformation can be categorized as a single transformation of the most complex type in the composition, e.g., a composition of a projective and an affine transformation is a projective transformation. Also a transformation is called *global* if it applies to the entire image, and *local* if subsections of the image each have their own transformations defined.

Rigid and affine transformations are generally global, and curved transformations are local. This is due to the physical model underlying the curved transformation type. Affine transformations are typically used in instances of rigid body movement where the image scaling factors are unknown or suspected to be incorrect, such as in MRI images due to geometric distortions. The projective transformation type has no real physical basis in image registration except for 2D/3D registration, but is sometimes used as a constrained-elastic transformation when a fully elastic transformation behaves inadequately or has too many parameters to solve.

#### 1.2.4 Interaction

Three levels of interaction can be involved in registration procedures. *Automatic*, where the user only supplies the algorithm with the image data and possibly information on the image acquisition. *Interactive*, where the user does the registration himself, assisted by software supplying a visual or numerical impression of the current transformation, and possibly an initial transformation guess. *Semi-automatic*, where the interaction required can be of two different natures: the user needs to initialize the algorithm, e.g., by segmenting the data, or steer the algorithm, e.g., by rejecting or accepting suggested registration hypotheses.

#### **1.2.5 Optimization Procedure**

There exists two possible ways of finding the transformation parameters. Either they are *computed* directly from the available image information, or they are *looked for* based on a certain optimization criterion. Many applications use more than one optimization technique, frequently a fast but coarse technique followed by an accurate yet slow one (as shown later).

#### 1.2.6 Modalities Involved

Four classes of registration tasks can be recognized based on the modalities that are involved. In *monomodal* applications, the images to be registered belong to the same modality, as opposed to *multimodal* registration tasks, where the images to be registered stem from two different modalities. The other two are *modality to model* and *model to modality* registration where only one image is involved and the other modality is either a model or the patient himself. The model to modality is used frequently in intraoperative registration techniques. Monomodal tasks are well suited for growth monitoring, intervention verification, rest-stress comparisons, ictal-interictal comparisons, subtraction imaging (also DSA, CTA), and many other applications. The applications of multimodal registration are abundant and diverse, predominantly diagnostic in nature. A coarse division would be into anatomical-anatomical registration, where images showing different aspects of tissue morphology are combined, and functional-anatomical, where tissue metabolism and its spatial location relative to anatomical structures are related.

#### 1.2.7 Subject

There can be *intrasubject* registration involving images for the same patient, *intersubject* registration involving images for different patients and *atlas*  registration where one image is acquired from a single patient, and the other image is somehow constructed from an image information database obtained using imaging of many subjects.

## **1.3 General Registration Theory**

In general, registration is the process by which two or more data sets are brought into alignment. Registration can be defined as "the process of finding a set of transformation operations between two or more data sets such that the overlap between these sets in a certain common space minimizes a certain optimization criterion".

The registration problem can be mathematically represented as follows: A parametric shape *S*, either a curve segment or a surface, is a vector function,

$$\mathbf{x}:[a,b] \to \mathfrak{R}^3 \tag{1.1}$$

for curves where a and b are scalars and

$$\mathbf{x}:\mathfrak{R}^2\to\mathfrak{R}^3\tag{1.2}$$

for surfaces. Both curve and surface data sets are usually in the form of, or can be easily converted to, a set of 2D or 3D points, which represent the most general form of 2D or 3D curves and surfaces including free-from curves and surfaces. Let the points in the first, or model data set, *S*, be denoted by  $\{\mathbf{x}_i | i = 1, ..., m\}$ , and those in the second, or experimental data set, *S'*, be denoted by  $\{\mathbf{y}_j | j =$  $1, ..., n\}$ . We want to find a transformation matrix **T** such that when applied to *S'*, the distance from each point on the resulting surface and its corresponding point on the model surface *S* is zero in the noise free case.

For the case of rigid registration (without considering the scaling factor), the transformation matrix  $\mathbf{T}$  consists of two components: a rotation matrix  $\mathbf{R}$ , and a translation vector  $\mathbf{t}$ . The objective of registration is to determine  $\mathbf{R}$  and  $\mathbf{t}$  such that the following criterion is minimized

$$\mathcal{F}(\mathbf{R}, \mathbf{t}) = \sum_{i=1}^{n} d^2 (\mathbf{R} \mathbf{y}_i + \mathbf{t}, S).$$
(1.3)

where  $d(\mathbf{y}_i, S)$  denotes the distance of point  $\mathbf{y}_i$  to shape *S*.

If we add the scaling factor as a third component, then the matrix  $\mathbf{T}$  represents a matrix called the *similarity transformation* matrix. In this case the

new shape will be *similar* to the original shape but at a different scale. It should be noticed that we will use the term similarity transformation to represent a rotation, translation and scaling only, no shear or torsion or other deformable transformations are included.

The minimization of Eq. (1.3) is very difficult because  $d(\mathbf{R}\mathbf{y}_i + \mathbf{t}, S)$  is highly nonlinear since the correspondence between  $\mathbf{y}_i$  and S is not known beforehand.

To understand the challenges involved in solving the registration problem one needs to understand the following:

- 1. For two data sets, if the transformation from one to the other is precisely known, then the registration process would be trivial. But when the transformation is only approximately known, the problem becomes more difficult. It is here that most researchers have addressed this problem. However, few researchers have attempted to solve the problem when the transformation is totally unknown.
- 2. The search for an unknown, optimal transformation invariably assumes an initial transformation which is iteratively refined through the minimization of some evaluation function. Such search may lead to a local minimum and, unless a global optimizer is used [3], it is difficult to reach a global solution.
- 3. For many applications (e.g., intraoperative registration), the registration time can be very critical and near real-time registration process is still needed.

The registration process must compensate for three very important problems, which are translational offset, rotational misalignment, and partial data sets. Error due to translational offset occurs when the coordinate origins of the data sets are not the same point in *N*-dimensional space. This can be demonstrated by calculating the point by point error of two identical surfaces located at different locations in the *N*-dimensional space. Even though the data sets are identical, the average experimental data error will be equal to the distance of the offset between the two sets.

Registration must also correct for error due to rotational misalignment between the data sets. This can be visualized by viewing a non-symmetrical surface from two different angles. The two different views may appear very different even though they come from the same surface. Once the views are rotated into correct alignment, an accurate value for the error measure can be obtained. The last major problem that the registration process must address is aligning data sets that represents only a portion of the model data. A correspondence between the experimental data set and the corresponding portion of the model data set must be established before correcting for translation and rotation errors. Once this is accomplished, the error measure must be calculated for only the overlapping portions of the two sets. For example, consider scanning a tooth and attempting to calculate the error between the scanned tooth and a model of an entire human jaw. The registration process must be able to determine which region of the jaw coincides with the scanned tooth, assuming the tooth is distinct enough to distinguish between the other teeth, and then calculate the error measure using only the overlapping regions.

In terms of algorithmic implementations, all of the registration techniques fall under two global implementation categories: *distance-based* and *featurebased* approaches. In the distance-based approach, the goal is to calculate the transformation by minimizing a criterion relating the distance between the two data sets. In the feature-based approach some differential properties invariant to rigid transformation (such as gray level value, histogram, curvature, mutual information, entropy, etc.) are often used.

In the following sections we will discuss in some details examples of algorithms in both the approaches.

### 1.4 Distance-based Registration Techniques

Among the distance-based techniques, Besl and McKay [4] proposed the *Iterative Closest Point (ICP)* algorithm which establishes correspondences between data sets by matching points in one data set to the closest points in the other data set. ICP is an elegant way to register different data sets because it is intuitive and simple. Besl and McKay's algorithm requires an initial transformation that places the two data sets in approximate registration and operates under the condition that one data set be a proper subset of the other. Since their algorithm looks for a corresponding scene point for every model point, incorrect registration can occur when a model point does not have a corresponding scene point due to occlusion in the scene. Attempts at solving these problems have led to several variants of the original algorithm.

The ICP algorithm matches points in the experimental data set after applying the previously recovered transformation  $(\mathbf{R}, \mathbf{t})$ , where  $\mathbf{R}$  is a matrix represent-

ing the rotation transformation and **t** is a vector representing the translation transformation, with their closest points in the model data set. A least-squares estimation is then used to reduce the average distance between the matched points in the two data sets. The algorithm is relatively straightforward and can be summarized as follows:

- 1. Given a motion transformation that initially aligns two data sets to some degree, a set of correspondence is developed between the points in each set. This is done using a simple metric: for each point in the first data set, pick the point in the second set which is closest to it under the current transformation.
- 2. From this set of correspondence an incremental motion can be computed which further aligns these points to one another.
- 3. Those two steps are iterated until some convergence criterion is satisfied.

Figure 1.1 illustrates these steps. The ICP algorithm tries to find the optimal transformation matrix **T** between two shapes *S* and *S'* such that Eq. (1.3) is minimized using the closest point operator in distance calculations as follows:

$$d(\mathbf{y}_i, S) = ||\mathbf{y}_i - \mathcal{C}(\mathbf{y}_i, S)||$$
(1.4)

where  $C(\cdot)$  is defined as the closest point operator, i.e.,  $C(\cdot)$  finds the closest point in shape *S* to the point  $\mathbf{y}_i$ . At each step of the minimization process, a correspondent point on *S* has to be found for each point  $\mathbf{Ry}_i + \mathbf{t}$  on *S'*. This makes the operation of registration of order O(mn) and as a result ICP has many drawbacks:

- 1. One of the main disadvantages of the ICP algorithm is its computation complexity. This makes the algorithm not suitable for applications where near real-time performance is required.
- 2. The algorithm converges successfully to a local solution but there is no guarantee that it will reach a global solution.
- 3. Proper convergence only occurs if one of the data sets is a subset of the other. The presence of points in each set that are not in the other leads to incorrect correspondences, which subsequently generates nonoptimal transformations.



Figure 1.1: A diagram illustrating the distance-based registration algorithm steps which start by an initial alignment and then finding correspondence from which incremental motion is calculated and this process iterates until convergence.

Attempts at solving these problems have led to several variants of the original algorithm. In what follows, we provide a review of these improvements. Another good review can be found in [5].

#### **1.4.1 Improving Correspondence**

The first improvement to the basic algorithm changes the simple point-to-point correspondence used in many of the methods [4, 6, 7, 8, 9] to that between a point and a location on the *surface* represented by the other data set. This potentially increases the integration accuracy beyond that of the sampling resolution.

The first such effort was by Chen and Medioni [7]. They have improved the ICP algorithm by minimizing the distance from the sensed point to the nearest estimated plane approximating the model surface. They begin by finding the data point in the second set that is closest to a line through the point in the first set in the direction of its estimated surface normal. Then the tangent plane at this intersection point is used as the surface approximation. Yet finding the estimated plane involves another iterative procedure which further adds to the computation complexity. They reduced the complexity by selecting some important points on the smooth surfaces of the object and used these points for the registration. This works well if the smooth surfaces are uniformly distributed over the object, which is not the case for many free-form objects. More accurate but time-consuming estimates of the surface have also been used, such as octrees [10], triangular meshes [11], and parametric surfaces [12].

Most researchers have used, the simple Euclidean distance in determining the closest point [3, 4, 6, 7, 8, 9, 11, 12, 13, 14]. Fewer have used higher dimensional feature vectors, such as including the estimated surface normal [15].

#### 1.4.2 Thresholding Outliers

Most of the early algorithms were limited by the original assumption that one data set was a subset of the other [4, 7, 8, 11, 12, 14]. Proposals to bypass this limitation have involved imposing a heuristic threshold on either the distance allowed between points in a valid pairing [6, 9, 10, 13, 15] and the deviation of the surface normals of corresponding points [15]. Any point pairs which exceed these thresholds or constrains are assumed to be incorrect. These thresholds are usually predefined constants related to the estimated accuracy of the initial transformations and can be difficult to choose robustly. Dynamically adjustable thresholds have been based on both the distribution of the distance errors at each iteration [6] and a fuzzy set classification of inlier and outliers correspondences [16].

#### **1.4.3 Computational Requirements**

In all of the techniques, computing potential correspondences is generally the most time-consuming step. In a brute-force approach [4, 15, 12], an  $O(N^2)$  number of comparisons is performed to find N pairings. One way to reduce

the actual time, with a potential loss of accuracy [11], is to subsample the original data sets. Criteria for sub-sampling include taking a simple fraction of the original number [13], using multiple scales of increasing resolution [3], or taking points in areas away from surface discontinuities [7], in areas of fine detail [8], and in small random sets for robust transform estimation [14]. A more accurate and slower alternative is to use the full original data sets, but organize the closest point search using efficient data structures such as the octree [10] and k-d tree [6]. The k-d tree is even more efficient,  $O(N\log N)$ , when higher order features of the points are incorporated in the distance metric.

#### **1.4.4 Computing Intermediate Motions**

Once a set of correspondences has been determined, a motion transform must be computed that best aligns the points. The most common approach is to use one of several least squares techniques to minimize the distances between corresponding points [4, 6, 8, 11, 14, 7]. In certain cases [6], individual point contributions are weighted based on the suspected noise of different portions of the data sets. More robust estimation using the least median squares technique (clustering many transforms computed from smaller sets of points) has been tried by Masuda et al. [14]. Alternatively, a Kalman filter has been used to track the intermediate motion at each iteration as new correspondences are computed [6]. More involved techniques compute the motion transform via some form of search in the space of possible transforms, trying to minimize a cost function such as the sum of distance errors across all corresponding points. Movements in transform parameter space are computed based on the changing nature of the function. Such standard search strategies as Levenberg-Marquardt [10, 12], simulated annealing [13] and genetic algorithms [9] have been used. Correspondences must be periodically updated during the search to keep the error function current. Updating too frequently can drastically increase the amount of computation, while too few updates can lead to an incorrect minimization.

#### **1.4.5** Initialization and Convergence of Searches

As mentioned earlier, an ICP-based refinement occurs after some initial set of transformations has been determined. Some researchers assume that this

estimate is determined by a previous process [6, 7, 11, 12, 14], possibly calculated using feature sets. Other prior estimates can be given by a rotary table [10], a robot arm [13], or even the user. Most such estimates are assumed to be quite accurate so that using one of the various distance thresholds during matching will prune outliers correctly. Other researchers do their own featurebased alignment prior to refinement using such characteristics as principal moments [4] or similar triangles on a mesh representation of the data [8]. If these distinguishing features are absent, a uniform distribution of starting points can always be processed [4]. All of the ICP algorithms must use some set of criteria to detect convergence of the final transformation. For those techniques that compute intermediate motions using least squares methods, convergence is achieved when the transform implies a sufficiently small amount of motion [6, 8, 14] or the distance between corresponding points becomes suitably close [4, 7, 11, 14]. The iterative searches of parameter space typically converge based on small changes in the parameters or error value, or if the shape of the cost function at the current value indicates a function minimum. Any method can be terminated if convergence is not detected after some maximal number of iterations [3].

#### 1.4.6 A Genetic Distance-based Technique

Another enhancements on the ICP algorithm for fast registration of two sets of 3D curves or surfaces was done by applying a distance transform to the model surface [3, 9]. The distance transform essentially converts the 3D space surrounding the two data sets into a field in which every point stores the magnitude and direction of a displacement vector from this point to the nearest surface element. Thus the cost function is largely precomputed. Such a transform is called the *grid closest point* (GCP) [9]. A genetic algorithm (GA) is then used to minimize the cost function.

Genetic algorithms (GAs) [17] provide a powerful and domain-independent search method for complex, poorly understood search spaces. Genetic algorithms have been applied to a wide diversity of problems such as combinatorial optimization [18], VLSI layout [19], machine learning [20], scene recognition [21], and adaptive image segmentation [22].

As mentioned in the previous section, to perform registration most of the computation time is spent in finding the closest point in the model set to every point in the experimental data set. This time can be significantly reduced by applying the following grid closest point (GCP) transform.

The GCP transform **GCP** :  $\Re^3 \to \Re^3$  maps each point in the 3D space  $\mathcal{G} \subset \Re^3$  that encloses the two surfaces *S* and *S'* to a displacement vector,  $\vec{r}$ , which represents the displacement from the closest point in the model set *S*. Thus for all  $\mathbf{z} \in \mathcal{G}$ 

$$\mathbf{GCP}(\mathbf{z}) = \vec{r} = \mathbf{x}_m - \mathbf{z} \tag{1.5}$$

such that

$$d(\mathbf{z}, \mathbf{x}_m) = \min_{\mathbf{x}_i \in S} \{ d(\mathbf{z}, \mathbf{x}_i)$$
(1.6)

where  $d(\cdot)$  is the Euclidean distance. For each point in  $\mathcal{G}$ , the transform calculates a displacement vector to the closest point in the model data set which can be used subsequently to find matching points between S and S' during the minimization process.

In the discrete case, assume that  $\mathcal{G}$  consists of a rectangular box that encloses the two surfaces. Furthermore, assume that  $\mathcal{G}$  is quantized into a set of  $L \times W \times H$  cells of size  $\delta^3$ 

$$\{C_{ijk} \mid 0 \le i \le L, \quad 0 \le j \le W, \quad 0 \le k \le H\}$$
(1.7)

such that

$$W = (W_{\rm max} - W_{\rm min})/\delta \tag{1.8}$$

$$L = (L_{\max} - L_{\min})/\delta \tag{1.9}$$

and

$$H = (H_{\text{max}} - H_{\text{min}})/\delta \tag{1.10}$$

Figure 1.2 shows a 2D illustration of such a grid.

Each cell  $C_{ijk}$  will hold a displacement vector  $\vec{r}_{ijk}$  which is a vector from its centroid, denoted by  $\mathbf{c}_0^{ijk}$ , to its closest point in the model set.

The GCP transform is applied only once at the beginning of the registration process. After its application, each cell in  $\mathcal{G}$  has a displacement vector to its closet point in the model set. During the minimization, to calculate the closest point  $\tilde{\mathbf{x}} = C(\mathbf{y}_v, S)$  to a point  $\mathbf{y}_v \in S'$ , you first have to find the intersection of  $\mathbf{y}_v$ and  $\mathcal{G}$ . Assuming uniform quantization, the indices of the cell  $C_{ijk}$  in which the



Figure 1.2: (Left) superimposing a uniform grid  $\mathcal{G}$  of size  $\delta$  onto the space that encloses the model and experimental data sets. (Right) for each cell  $C_{ijk} \in \mathcal{G}$ , calculate a displacement  $\vec{r}_{ijk}$ , from the cell centroid,  $\mathbf{c}_0^{ijk}$ , to its closet point on S.

point  $\mathbf{y}_v = \{u, v, w\}$  lies can be found by

$$i = \frac{u - L_{\min}}{\delta}, \quad j = \frac{v - W_{\min}}{\delta}, \quad k = \frac{w - H_{\min}}{\delta}$$
 (1.11)

If the content of the cell  $C_{ijk}$  is  $\vec{r}_{ijk}$  then

$$\tilde{\mathbf{x}} = \mathcal{C}(\mathbf{y}_v, S) = \vec{r}_{ijk} + \mathbf{c}_0^{\ ijk} \tag{1.12}$$

An approximation of the closest point can be obtained by using the point itself instead of the centroid of the cell in which it lies

$$\tilde{\mathbf{x}} = \mathcal{C}(\mathbf{y}_v, S) \approx \vec{r}_{ijk} + \mathbf{y}_v \tag{1.13}$$

Equation (1.13) introduces an error which is a function of  $\delta$ , the quantization step. This error can be reduced to some extent by using a non-uniform quantization. It should be noted that the GCP transform is spatially quantized and its accuracy depends largely on the selection of  $\delta$ . The error in the displacement vector is  $\leq \frac{3}{4}\delta^2$ . Therefore, smaller values for  $\delta$  will give higher accuracy but on the extent of larger memory requirements and larger number of computations of the GCP for each cell. To solve this problem, you can select a small value for  $\delta$  in the region that directly surrounds the model set, and a slightly larger value (in this work  $2\delta$ ) for the rest of the space  $\mathcal{G}$ . This enables a coarse matching process at the beginning of the registration and fine matching toward the end of the minimization.

The next step is to search for the transformation parameters using genetic algorithms (GAs). GAs, pioneered by Holland [23], are adaptive, domain independent search procedures derived from the principles of natural population genetics. GAs borrow its name from the natural genetic system. In natural genetic system whether a living cell will perform a specific and useful task in a predictable and controlled way is determined by its genetic makeup, i.e., by the instructions contained in a collection of chemical messages called genes [24]. Genetic algorithms are briefly characterized by three main concepts: a Darwinian notion of fitness or strength which determines an individual's likelihood of affecting future generations through reproduction; a reproduction operation which selects individuals for recombination to their fitness or strength; and a recombination operation which creates new offspring based on the genetic structure of their parents.

Genetic algorithms work with a coding of a parameter set, not the parameters themselves and search from a population of points, not a single point. Also genetic algorithms use payoff (objective function) information, not derivatives or other auxiliary knowledge and use probabilistic transition rules, not deterministic ones. These four differences contribute to genetic algorithms' robustness and resulting advantage over other more commonly used search and optimization techniques (see Fig. 1.3).

Since the genetic algorithm works by maximizing an objective function, the fitness function,  $\mathcal{F}_r(\mathbf{R}, \mathbf{t})$ , can be defined as in Eq. (1.3).

Combining (1.3) with the GCP transform to find matching points between the two data sets, Eq. (1.3) can be rewritten as

$$\mathcal{F}_r(\mathbf{R}, \mathbf{t}) = -\sum_{i=1}^n d^2(\mathbf{y}', \mathbf{GCP}(\mathbf{y}'_i) + \mathbf{y}'_i)$$
(1.14)

where  $\mathbf{y}'_i = \mathbf{R}\mathbf{y}_i + \mathbf{t}$ . By maximizing (1.14) we effectively minimize (1.3).

The objective of the registration process is to obtain the rotation matrix **R** and the translation vector **t**. Thus in 3D space, there are six parameters that need to be evaluated; the three angles of rotation around the three principal axes  $\alpha$ ,  $\beta$ , and  $\gamma$ , and the displacement  $\Delta x$ ,  $\Delta y$ , and  $\Delta z$  in the x, y, z directions, respectively.



Figure 1.3: An example of a non-uniform grid  $\mathcal{G}$ .

In the 2D space, the parameter set is reduced to  $\Delta x$ ,  $\Delta y$ , and the angle of rotation  $\theta$ . These parameters are represented by binary notation to minimize as much as possible the length of the schemata d(H) and the order of the schemata O(H). The number of bits  $n_p$ , assigned to each parameter, p, depends on the type of application and the required degree of accuracy. The number of bits should be chosen as small as possible to minimize the time of convergence of the genetic algorithm. For example, you can assign 8 bits each, thus allowing a displacement of  $\pm 127$  units. A range of  $\pm 31$  can also enforced over the angles of rotation. Therefore 6 bits are assigned for each angle of rotation. As shown in Fig. 1.4, the genes are formed from the concatenation of the binary coded parameters.

The selection operator chooses the highest fitted genes for mating using a Roulette wheel selection [24]. The crossover and mutation operators are implemented by choosing a random crossover and mutation point with probabilities  $P_c^p$  and  $P_m^p$ , respectively, for each coded parameter p. The generated strings are concatenated together to form one string from which the populations are formed (see Fig. 1.5).

More details and error analysis of the GCP technique can be found in [9]. The major drawback of the GCP/GA algorithm is that the range of the transformation



Figure 1.4: Different gene structures.



Figure 1.5: An illustration of the GA use in calculating the transformation parameters. The process starts by coding the transformation parameters into one string. An initial population of strings is randomly generated. Applying the selection, crossover and the mutation operations, new generations are obtained. Notice how the registration fitness function average increases with the generations. Maximum fitness can be reached eventually.

parameters needed to be known beforehand to efficiently code them into a string. Also if the range is large, the GA convergence can be slow.

## 1.5 Feature-Based Registration Techniques

Feature-based registration techniques rely on extracting and matching similar features vector between two or more data sets in order to find corresponding data points. So the two critical procedures involved in feature-based registration is *feature extraction* and *feature correspondence*. Feature representations, which are invariant to scaling, rotation, and translation, are more desirable in the matching process.

One of the most successful feature-based registration techniques, especially for multimodal registration, is by *maximization of mutual information* (MI) [25]. This technique works well for both MR and CT since they are informative of the same underlying anatomy and there will be mutual information between the MR image and the CT image. Such a technique would attempt to find the registration by maximizing the information that one volumetric image provides about the other. It requires no a priori model of the relationship between the modalities, it only assumes that one volume provides the most information about the other one when they are correctly registered.

Unfortunately, if initial transformation between the two modalities is unknown, the MI will converge slowly. So, we will demonstrate how to use a composite registration procedure that integrates another feature-based registration technique, mainly a *surface-based* [26] registration technique, to estimate the initial transformation. Surface-based registration techniques use features available on the data set surface mesh such as density or curvature. The surfacebased registration techniques work better for free-form surfaces, such as the skin contours, while MI works better for voxel-based volumes. Such composite registration procedures have become recently the state-of-the-art in most registration applications due to the fact that most feature-based techniques are complementary in nature.

#### 1.5.1 Surface-Based Registration Algorithm

In order for any surface-based registration algorithm to perform accurately and efficiently, an appropriate representation scheme for the surface is needed.
Most of the surface representation schemes found in literature have adopted some form of shape parameterization especially for the purpose of object recognition. However, free-form surfaces, in general (e.g., CT/MRI skin contours), may not have simple volumetric shapes that can be expressed in terms of parametric primitives. Some representation schemes for free-form surfaces found in literature include the "splash" representation proposed by Stein and Medioni [27] in which the surface curvature along the intersection of the surface and a sphere centered at the point of interest is calculated for different sphere diameters and a signature curve is obtained for this point. Also the work of Chua and Jarvis [28], who proposed the "point signature" representation which describes the local underlying surface structure in the neighborhood of a point. This is obtained by plotting the distance profile of a circle of points to a plane defined by that circle of points. Dorai and Jain [29] proposed another representation called "COSMOS" for free-form objects in which an object is described in terms of maximal surface patches of constant shape index from which properties such as surface area, curvedness and connectivity are built into the representation. Johnson and Hebert [30] recently introduced a new representation scheme called the "spin image". This image represents the histogram of the surface points relative to a certain oriented point. This image is generated for each oriented point on the surface and matching between two surfaces is done by matching the spin images of the points in the two surfaces. Yamany and Farag [26] introduced another technique based on surface signatures. Surface signatures are 2D images formed by coding the 3D curvature information seen from a local point. These images are invariant to most transformation. In what follows are some details for some of these algorithms.

#### 1.5.1.1 The "Splash" Surface Registration

Geometric indexing have been one of the most used surface indexing techniques because it used the geometrical relationships between invariant features. However, another form of indexing that uses local shape information has become more popular. As it is based on structural information local to the neighborhood of a point, this indexing method is called "Structural indexing".

Circle of points was first used to describe the underlying surface structure around a given point. This can be done by decomposing the local surface patch



Figure 1.6: An illustration of the "splash" representation scheme. At specific points on the surface, the intersections of the surface patches and the spheres of prefixed radii are obtained. For each intersection a curve representing the average normal of the points in the intersection and the point in study is obtained. These curves are further used for matching.

around a specific point into a series of contours, each of which is the locus of all points at a certain distance from the specific point.

Stein and Medioni [27] extended this idea further. Instead of decomposing a surface patch into a series of contours of different radii, a few contours at prefixed radii are extracted as shown in Fig. 1.6. On each contour of points, surface normals are computed. This contour is called a "splash". A 3D curve is constructed from the relationship between the splash and the normal at the center point. This curve is converted into piecewise linear segments. Curvature angles between these segments and torsion angles between their binormals are computed. These two features are used to encode the contour. Matching is performed using the contour codes of points on the two surfaces. Fast indexing was achieved by hashing the codes for all models in the library into an index table.

At runtime recognition, the splashes of highly structured regions are computed and encoded using the same encoding scheme. Models which contain similar codes as the splashes appearing in the scene are extracted. Verification is then performed for each combination of three correspondences.

#### 1.5.1.2 The "Point Signature" Surface Registration

This approach was introduced by Chua and Jarvis [28] for fast recognition. They establish a "signature" for each of the given 3D data points rather than just depending on the 3D coordinates of the point. This is similar to the "splash" representation but instead of using the relationship between the surface normals of the center points and its surrounding neighbors, they used the point set itself.

For a given point p, they place a sphere of radius r, centered at p (similar to the method used in splash as depicted in Fig. 1.6). The intersection of the sphere with the object surface is a 3D space curve whose orientation can be defined by a directional frame formed of the normal to the plane fitted to the curve, a reference direction and their cross product. The next step is to sample the points on this curve starting from the reference direction. For each sampled point there exist two information, the distance from itself to the fitted plane and the clockwise angle about the normal from the reference direction. Figure 1.7 shows some typical examples of point signatures for different surface types.

Due to this simple representation, the 3D surface matching is transformed into 1D signature matching. In their paper they analyze this signature matching and estimate the accepted error tolerance in the matched signature. Prior to recognition, the model library is built by computing the point signatures for each point in the model and for every model. They also used hashing to index the signatures in a table. At runtime, the surface under study is sampled at a fixed interval and the sampled points are used in the matching process.

#### 1.5.1.3 The "COSMOS" Surface Registration

The goals of the COSMOS (Curvedness-Orientation-Shape Map On Sphere) representation scheme were, first, to be a general representation scheme that can



Figure 1.7: Examples of point signatures: (a) peak, (b) ridge, (c) saddle, (d) pit, (e) valley, (f) roof edge.

be used to describe sculpted objects, as well as objects composed of simple analytical surface primitives. Second, to be as compact and as expressive as possible for accurate recognition of objects from single range image.

The representation uses the "shape index" to represent complex objects. An object is characterized by a set of maximally sized surface patches of constant shape index and their orientation dependent mapping onto the unit sphere. The patches that get assigned to the same point on the sphere are aggregated according to the shape categories of the surface components.

The concept of "shape spectrum" features is also included in the COSMOS framework. This allows free-form object views to be grouped in terms of the shape categories of the visible surfaces and their surface areas.

For the recognition purpose, COSMOS adapted a feature representation consisting of the moments computed from the shape spectrum of an object view. This eliminated unlikely object view matches from a model database of views. Once a small subset of likely candidate views has been isolated from the database, a detailed matching scheme that exploits the various components of the COSMOS representation is performed to derive a matching score and to establish view surface patch correspondence.

#### 1.5.1.4 The "Spin Image" Surface Registration

Johnson and Hebert [30] presented an approach for recognition of complex objects in cluttered 3D scenes. Their surface representation, the "spin" image, comprises descriptive images associated with oriented points on the surface. Using a single point basis, the positions of the other points on the surface are described by two parameters. These parameters are accumulated for many points on the surface and result in an image at each oriented point which is invariant to rigid transformation.

Through correlation of images, point correspondences between a model and scene data are established. Geometric consistency is used to group the correspondences from which plausible rigid transformations that align the model with the scene are calculated. The transformations are then refined and verified using a modified ICP algorithm.

The spin image is generated by first considering an oriented point (a 3D point with a normal direction) and defining its basis. The basis is defined using the tangent plane perpendicular to the point direction and the point direction itself. A spin-map is then defined using the point basis. In this spin map, any other point on the surface is related to the oriented point by two parameters, one is the perpendicular distance to the oriented point line direction and the other is the signed perpendicular distance to the plane passing through the oriented point and perpendicular to the point direction.



Figure 1.8: The spin image generation process can be visualized as a sheet spinning around the oriented point basis, accumulating points as it sweeps space.

The term "spin-map" comes from the cylindrical symmetry of the oriented point basis; the basis can spin about its axis with no effect on the coordinates with respect to the basis. To create a spin image, first an oriented point on the surface is selected. Then for each other point on the surface, the spin-map parameters are computed. These parameters are then accumulated in a 2D array. Once all the points on the surface have been processed, the 2D array is converted into a gray image. Figure 1.8 shows the spin-image generation process and visualizes it as a sheet spinning around the oriented point basis, accumulating points as it sweeps space. Figure 1.9 shows examples of spin images generated for the surface of a statue. The darker the pixel, the higher the number of points projected into this location.

They compared spin images using linear correlation coefficients. They used the magnitude of the correlation coefficients as well as the confidence in the correlation results which is measured by the variance of the correlation coefficient. They also modeled the effect of clutter and occlusion to predict a lower bound on the correlation coefficient.



Figure 1.9: Examples of spin images generated for the surface of a statue. The darker the pixel, the higher the number of points projected into this location.

### 1.5.1.5 The "Surface Signature" Surface Registration

The surface signature algorithm captures the 3D curvature information of any free-form surface and encodes it into a 2D image corresponding to a certain point on the surface. This image is unique for this point and is independent from the object translation or orientation in space.

The process starts by identifying special points on the surface. These points are called *Important* points due to the information they carry. Then an image capturing the surface curvature information seen from each *important* point, is formed. This image is called the Surface Signature at this point because it is almost unique for each point on the surface. Surface registration is then performed by matching signature images of different surfaces and hence finding

corresponding points in each surface. A point P on a surface/curve S, is called *important* point,  $P_I$ , if and only if the absolute value of the curvature at this point is larger than a certain positive value (a threshold).

$$A = \{P_I\} = \{P \in S | |\operatorname{Curv}(P)| > \epsilon, \epsilon > 0\}$$

$$(1.15)$$

As the *important* points are landmarks, one may expect that they are stable for the same object. However, due to scanning artifacts, their number and locations may vary. By adjusting the curvature threshold, a common subset can be found. Otherwise, the object has either suffered from non-rigid transformations or its visible surface has no landmarks.

The signature, computed at each important point, encodes the surface curvature seen from this point, thus giving it discriminating power. As shown in Fig. 1.10 (a), for each important point  $P \in A$  defined by its 3D coordinates and the normal  $U_P$  at the patch where P is the center of gravity, each other point  $P_i$  on the surface can be related to P by two parameters: The distance

$$d_i = ||P - P_i|| \tag{1.16}$$

and the angle

$$\alpha_{i} = \cos^{-1} \left( \frac{U_{P} \cdot (P - P_{i})}{||P - P_{i}||} \right)$$
(1.17)

Also you can notice that there is a missing degree of freedom in this representation which is the cylindrical angular parameter. At each location in the image, the gray value encodes the angle

$$\beta_i = \cos^{-1}(U_P \cdot U_{P_i})$$
(1.18)

This represents the change in the normal at the surface point  $P_i$  relative to the normal at P. Figure 1.10 (b) shows some signature images taken at different important points on the skin model of a patient's head.

The next step in the registration process is to match corresponding signature images of two surfaces. The ultimate goal of the matching process is to find at least a three-points correspondence in order to calculate the transformation parameters. The benefit of using the signature images to find the correspondence is that we can now use image processing tools in the matching, hence reducing the time taken to find accurate transformation. One such tool is *Template Matching* in which a value defines how well a portion of an image matched a template.





Figure 1.10: (a) For each important point on the surface a signature image is obtained. This image encodes the angle difference between the normal at the point in focus and the normal at each other point on the surface. (b) Examples of signature images taken at different important points locations. Notice how the image provides features the curvature information. The dark intensity in the image represents a high curvature seen from the point while the light intensity represents a low curvature. Also notice how different the image corresponding to a location is from images of other locations.

The end result of the matching process is a list of groups of likely threepoints correspondences that satisfies the geometric consistency constraint. The list is sorted such that correspondences that are far apart are at the top of the list. A rigid transformation is calculated for each group of correspondences and a verification stage [9] is performed to obtain the best group. Detailed discussion concerning the surface signature sensitivity and robustness can be found in [26].

# 1.5.2 Maximization of Mutual Information (MI) Algorithm

MI is a basic concept from information theory, measuring the statistical dependence between two random variables or the amount of information that one variable contains about the other. The MI registration criterion used states that the MI of corresponding voxel pairs is maximal if the two volumes are geometrically aligned [31]. No assumptions are made regarding the nature of the relation between the image intensities in either modality.

Consider the two medical volumes to be registered as the reference volume R and the floating volume F. A voxel of the reference volume is denoted R(x), where x is the coordinates vector of the voxel. A voxel of the floating volume is denoted similarly as F(x). Given that T is a transformation matrix from the coordinate space of the reference volume to the floating volume, F(T(x)) is the floating volume voxel associated with the reference volume voxel R(x).

MI seeks an estimate of the transformation matrix that registers the reference volume R and floating volume F by maximizing their mutual information. The vector x is treated as a random variable over coordinate locations in the reference volume. Mutual information is defined in terms of entropy in the following way [25]:

$$I(R(x), F(T(x))) \equiv h(R(x)) + h(F(T(x))) - h(R(x), F(T(x))).$$
(1.19)

where h(R(x)) and h(F(T(x))) are the entropy of R and F, respectively. h(R(x), F(T(x))) is the joint entropy. Entropy can be interpreted as a measure of uncertainty, variability, or complexity. The mutual information defined in Eq. (1.19) has three components. The first term on the right is the entropy in the reference volume, and is not a function of T. The second term is the entropy of the part of the floating volume into which the reference volume projects. It encourages transformations that project R into complex parts of F. The third term, the (negative) joint entropy of R and F, contributes when R and F are functionally related. Maximizing MI tends to find as much of the complexity that is in the separate volumes (maximizing the first two terms) as possible so that at the same time they explain each other well (minimizing the third term) [25, 31].

#### **1.5.2.1** Computation of MI Metric

Based on the definition of relative entropy, also known as Kullbak Leibler distance, between two probability mass functions, Eq. (1.19) can be written in terms of probability distribution functions as follows

$$I(R(x), F(T(x))) = \sum_{R(x), F(T(x))} p(r(x), F(T(x))) \cdot \log_2 \frac{p(r(x), F(T(x)))}{p_{R(x)}(r(x) \cdot p_{F(T(x))}(F(T(x)))}.$$
 (1.20)

where p(x, y) is the joint distribution function and  $p_x(x)$  is the marginal probability mass functions. The marginals can be obtained directly from the joint probability function. The joint probability mass function p(r(x), F(T(x))) will be approximated by the normalized joint histogram H(r(x), F(T(x))). Here, normalization refers to scaling of the histogram, such that the sum of approximated probabilities equals 1.0. The marginals are then approximated from H()by summation over the rows, and then the columns. Computation of H() involves a complete iteration over each sample in the floating volume. For each sample, the transformation T is applied, to arrive at a coordinate set in the image coordinate system of the reference volume. If the transformed coordinate is outside the measured reference volume, then the remaining operations are not executed, and the process starts again with the next sample in the floating volume. Otherwise, a sample in the reference volume at the transformed coordinates is approximated using trilinear interpolation, and discretized. The two samples, one from the floating volume, and one from the reference volume, are then binned in the joint histogram.

Computation of the joint histogram involves the processing of each sample in the floating volume, application of a transformation to the coordinate of the sample in the floating volume to obtain a coordinate in the reference volume, interpolation in the reference volume, and binning in the joint histogram. For a typical 256 by 256 by 20 MRI volume, there are thus 256 by 256 by 20 = 1,310,720

samples to process. Following computation of the joint histogram, normalization and computation of the marginal histograms must be performed. This involves one pass over the joint histogram, therefore processing 256 by 256 = 65,536elements from the joint histogram. This processing consists of normalization, and summation to compute the marginal histograms. Following this operation, the mutual information metric itself may be computed. This processing involves computation of the sum given in Eq. (1.19), and involves one pass over the joint histogram, therefore processing 256 by 256 = 65,536 elements that compose the sum given in Eq. (1.19). Therefore, computation of the joint histogram is by far the most computationally costly component in computation of the mutual information metric. Therefore, performance may be best increased by decreasing the execution time of the computation of the joint histogram. Computation of the joint histogram is an amenable problem for parallel execution, as computation of a part of the joint histogram does not depend on the computational results of any other part of the joint histogram, allowing individual bins, or entire regions of the joint histogram to be computed independently, and then merged to form the total joint histogram. Figure 1.11 illustrates such parallel architecture.



Figure 1.11: Illustration of the parallel computation of the joint histogram necessary for computation of the mutual information metric.

#### 1.5.2.2 Optimization Methods

The mutual information metric provides a quantitative measure of spatial alignment between two volumes, given a choice of registration parameters. To obtain the best alignment, it is necessary to maximize the metric. Maximization of the metric, which is parameterized in terms of the registration parameters, is numerically accomplished with the use of an optimization algorithm. In the original formulation of registration by maximization of mutual information in [31], Powell's multidimensional optimization method, with Brent line minimizations was used for maximization of the mutual information metric [32]. Subsequently, [33] compares different classical optimization methods maximizing the mutual information metric. One such method included was the use of the classic Nelder and Mead or simplex algorithm for maximization. This method solely uses the objective function directly for optimization, and therefore does not require the expensive computation of derivatives. This method is a geometry-based method, using the geometric operations of contraction, expansion, and reflection to manipulate a simplex to a maximum of the objective function. GA was also used as an optimization criteria in maximizing the mutual information metric as demonstrated in [34].

### **1.6 Practical Examples**

### 1.6.1 Composite Approach for Mutlimodal Registration

A composite registration approach can perform a fast, six degrees of freedom registration and accurately locate the position and orientation of medical volumes (obtained from CT/MRI scans for the same patient) with respect to each other. The technique uses surface registration technique followed by a volume registration approach. The advantage of this combination is to have an accurate alignment and to reduce the time needed for registration. The surface registration uses the surface signature approach and for the volume registration, the maximization of mutual information (MI) is used as a matching criterion and is enhanced by a genetic based optimization technique. Figure 1.12 shows a block diagram of the composite registration. Figure 1.13 shows some results and Table 1.2 shows approximate registration time for both single and composition registration.



Figure 1.12: Two different volumes are fed into the system. Surface registration provides an estimate transformation between them. This result is further enhanced by maximizing the mutual information. A Genetic Algorithms technique is used in the optimization process.

### 1.6.2 Multiple Sclerosis (MS) Application

Multiple sclerosis is a disease of the human central nervous system, affecting approximately 250,000 to 350,000 people in the United States alone [35]. MS results



Figure 1.13: Two different signature samples. The upper two images are from the MR dataset; while the lower two are from the CT dataset. This figure shows similarity of the signature images of each matched pair, although they are produced from two different datasets (CT and MR) with a different 3D volume size (number of slices).

		Composite				
Case	No. of slices	Surface Matching		Volume	Total time	Single-Stage MI-GA
		Features Extraction	Matching	Matching	(sec)	(sec)
CT/CT CT/MR	236 33/19	20 sec 8 sec	30 sec 20 sec	100 sec 80 sec	150 108	300 250

 Table 1.2:
 Comparison between approximate times for single and composite

 volume registration
 Image: Comparison between approximate times for single and composite

in a variety of clinically observable deficiencies, such as speech difficulties, pain, impairment of senses, loss of muscle control, and cognitive abnormalities [36]. In the brain, MS results in the inflammation and destruction of myelin, a fatty covering insulating nerve cells [36]. This damage results in decreased ability of the nervous system to control the body, leading to the clinically-observable symptoms of the disease. The causes of MS are not clearly accepted in the medical community. Geographic, genetic, and environmental factors all seem to be present [35]. Many researchers have proposed that MS is an autoimmune disease [35]. Though there are promising research developments, at this time, no cure is known for this disease. Several treatment options are available, and allow for management of the disease. Despite this, most patients progress in disability over the course of their life [35]. Though not usually a fatal disease in and of itself, the resulting disabilities may contribute to accidental mishaps [35].

When patients with MS are imaged using MRI modalities, lesions (also referred to as plaques or deficits) can be contrasted against surrounding, normal brain tissue, by choice of appropriate scan parameters, and depending on the state of the lesion [37]. Figure 1.14 shows several slices of different FLAIR MR images of patients with MS. MS lesions appear as hyper intense regions in the patient's brain.

The use of MR imagery in the evaluation of MS involves the identification of abnormal brain tissue (MS lesions), and normal, non-diseased brain tissue (gray and white matter). MRI has been found to be a sensitive marker to changes in disease progression, and evaluation of MRI studies of MS patients can be useful as an outcome measurement in MS studies [4]. Despite these statements, evaluation of the disease cannot be completely based upon MRI findings. MS lesion activity observed in MRI studies of the brain does not always correspond



Figure 1.14: Sample slices from FLAIR MRI studies of patients with MS. Hyperintense regions in the brain are indicative of plaques caused by MS.

to clinically observed deficits [37]. As well, quantification of MRI studies of MS have yet to be perfected, and most studies of MS do not generate enough MRI data to evaluate the nuances in the application of MRI to study MS.

It is desirable to develop computer tools to assist experts in the study of MS using MR imaging. Allowing a computer to automatically identify normal and abnormal brain tissue would free an expert from the arduous task of manually examining each slice of a study, while generally increasing the reproducibility of the identification by removing the subjectivity of the human observer. Medical registration tools would also be useful for automatic, retrospective alignment of patient studies, taken at different points in time, to allow for qualitative comparison of different studies of a patient of the course of his or her treatment. This type of analysis would aid the expert in deciding if the disease is responding well to the present treatment, or if a change in the treatment is warranted.

Typically in an MRI study, a patient is placed in the scanner with little regard for positioning of the anatomy of interest. The only constraints are that the anatomy of interest falls within the scanning volume, and that the patient is generally placed in some orientation such that gross anatomical features are placed in some direction (for example, the patient's nose points upward in the scanner). In the context of qualitative evaluation of MS studies however, this arrangement leads to hindrances in evaluation, due to the fact that multiple scans must be compared with one another. Due to the largely arbitrary positioning of the anatomy in the scanner, in a slice-by-slice comparison between studies, quite different anatomy can by chance be located on the same slice numbers in different studies.

It is desirable to be able to perform registration using the intrinsic approaches, rather than imposing limitations in the scanning procedure, or affixing artificial fiducial markers on the patient's head. For best accuracy, artificial markers would likely be affixed to the skull, and therefore would be inconvenient and potentially painful for the patient. Additionally, this procedure would also introduce a risk of infection. Furthermore, using intrinsic registration techniques, it is also desirable to be able to apply registration retroactively, allowing for current data sets to be aligned with data sets taken previously in a patient's history, or perhaps with an imaging modality that prevents the use of artificial markers. In the study of MS using MRI, for comparison of scans taken at different points of time in a clinical study, of the same patient, a registration technique is necessary. Such a tool would allow for alignment of the patient's anatomy in different scans. When this alignment is accomplished, qualitative comparison of scans becomes easier to an expert viewer, as image slices will now contain the same anatomy, and quantitative comparison between studies is enabled in the same manner. Registration can also be used to assist in segmentation. For example, if a model of the patient's anatomy is known, then a study can be registered to that model, allowing for segmentation of certain classes of problems to be made trivial, as the segmentation of the data set is then known a priori from the model. In the MS research, this approach is used for segmentation of a patient's brain from his head. In this context, registration is performed by maximization of mutual information. This technique has generally been found to perform well, and is useful in the MS clinical settings. This technique was studied, implemented, and tested using MS patient studies. Additionally, the performance of this method was enhanced by application of parallel programming techniques.

Figure 1.15 shows sample slices from the registration studies. Each row of the figure corresponds to a single study. The first column of the figure shows a sample slice from the floating volume used for each study. The second column of



Figure 1.15: Sample registration results from each samples of MS data sets. The first column is a sample slice from the floating volume used. The second column is the corresponding slice from the re-sampled reference volume. The third column is the checkerboard composite image of the two corresponding slices from the floating and resampled reference volumes. The floating volume and reference volumes used in each trial were from the same patient.

the figure shows the corresponding slice from the re-sampled reference volume. Finally, the third column of the figure shows the checkerboard composite image formed by fusion of the corresponding floating and re-sampled reference volume slices.

### 1.6.3 Lung Cancer Application

Another practical application for monomodal registration in the lung cancer diagnosis. Lung cancer remains the leading cause of mortality cancer. In 1999, there were approximately 170,000 new cases of lung cancer [38]. The five-year survival rate from the diseases is 14% and has increased only slightly since the early 1970s despite extensive and costly research efforts to find effective therapy. The disparity in survival between early and late-stage lung cancer is substantial, with a five-year survival rate of approximately 70% in stage 1A disease compared to less than 5% in stage IV disease according to the recently revised lung cancer



Figure 1.16: (a) CT scan June 2000 (b) CT scan June 2001 (c) CT scan June 2002.



Figure 1.17: MI fitness function at different GA iterations.

staging criteria [38]. The disproportionately high prevalence and mortality of lung cancer has encouraged attempts to detect early lung cancer with screening programs aimed at smokers. Smokers have an incidence rate of lung cancer that is 10 times that of nonsmokers and accounts for more than 80% of lung cancer cases in the United States [38]. One in every 18 woman and every 12 men develop lung cancer, making it the leading cause of cancer deaths. Early detection of lung tumors (visible on chest film as nodules) may increase the patient's chance of survival.



Figure 1.18: (a) Result of registration of data in Figs. 1.16.a and 1.16.b. (b) Result of registration of data in Figs. 1.16.a and 1.16.c.

The Jewish Hospital Foundation in Louisville, KY., proposed a cancer screening and early detection study in a randomized trial with the following specific aims: (1) to determine whether the use of spiral CT scanning of the chest detects early lung abnormalities that lead to cancer, which are not visible on chest X-rays in patients at high-risk for developing lung cancer; and (2) to determine whether annual spiral chest CT scans of the chest in high-risk patients result in an improvement in survival. To test and prove these hypothesis, data was collected from 1000 symptomatic patients above 60 years of age with positive smoking history will undergo screening with low dose spiral CT (LDCT) and chest radiography. Screening was performed every three months on the selected 1000 symptomatic patients. The role of the image registration process was to help in studying the development of abnormalities.

Again, maximization of mutual information enhanced by using genetic algorithms was used in this application. Figure 1.16 shows different CT images for the same patient taken at various periods. Registration of these images using the mutual information criterion was performed fully automatically on a PC computer with microprocessor 2.4 GHz. The algorithm took less than 8 min. to register all data. The recovered rotational transformation parameters were generally smaller than 15 degrees, while the translational parameters varied up to 20 mm. Figure 1.17 shows how the GA iterations improve the MI measure overtime until maximum fitness is reached. Figure 1.18 shows the results of registration for the data shown in Fig. 1.16. The registration process was able to show that new parts of abnormality appeared in later scans between the period June 2001 to June 2002.

# Questions

- 1. Based on the classification defined in this chapter, define all the possible classifications of the following registration techniques:
  - (a) Registration by maximization of Mutual Information.
  - (b) Surface Signature registration.
  - (c) Grid Closest Point registration.
- 2. Using the formulation for the Surface Signature images, derive and draw the signature image for the following parametric shapes. (choose points of interest on each shape)
  - (a) a sphere of radius r

$$Sph(u, v) = \begin{bmatrix} r \cos(2\pi u) \\ r \sin(2\pi u)\cos(2\pi v) \\ r \sin(2\pi u)\sin(2\pi v) \end{bmatrix}$$
(1.21)

(b) a cylinder of radius r and height h

$$Cyl(u, v) = \begin{bmatrix} r \cos(2\pi u) \\ r \sin(2\pi u) \\ v/h \end{bmatrix}$$
(1.22)

- 3. Describe how the Genetic Algorithm technique can be used as an optimizer for the MI registration technique.
- 4. Derive Eq. (1.19) from Eq. (1.20).

# Bibliography

- Maintz, J. B. A. and Viergever, M. A., A survey of medical image registration, Medical Imaging Analysis, Vol. 2, no. 1, pp. 1–36, 1998.
- [2] van den Elsen, P. A. and Viergever, M. A., Automated CT and MR brain image registration using geometrical feature correlation., Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 1827–1830, 1993.
- [3] Yamany, S. M., Ahmed, M. N., and Farag, A. A., Novel surface registration using the grid closest point (GCP) transform, Proc. IEEE International Conference on Image Processing, Chicago, October 1998.
- [4] Besl, B. J. and McKay, N. D., A method for registration of 3D shapes, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, no. 2, pp. 239–256, 1992.
- [5] Eggert, D. W., Fitzgibbon, A. W., and Fisher, R. B., Simultaneous registration of multiple range views for use in reverse engineering of CAD models, Computer Vision and Image Understanding, Vol. 69, no. 3, pp. 253– 272, March 1998.
- [6] Zhang, Z., Iterative point matching for registration of free-form curves and surfaces, International Journal of Computer Vision, Vol. 13, no. 2, pp. 119–152, 1994.
- [7] Chen, Y. and Medioni, G., Object modelling by registration of multiple range images, Image Vision Computing, Vol. 10, no. 3, pp. 145–155, 1992.
- [8] Bergevin, R., Laurendeau, D., and Poussart, D., Registering range views of multipart objects, Computer Vision and Image Understanding, Vol. 61, no. 1, pp. 1–16, January 1995.
- [9] Yamany, S. M., Ahmed, M. N., and Farag, A. A., A New genetic-based technique for matching 3D curves and surfaces, Pattern Recognition, Vol. 32, no. 10, p. 1817–1820, 1999.
- [10] Champleboux, G., Lavallee, S., Szeliski, R., and Burnie, L., From accurate range imaging sensor calibration to accurate model-based 3D

object recognition, Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Champaign, pp. 83–89, 1992.

- [11] Brujic, D. and Ristic, M., Analysis of free form surface registration, Proc IEEE International Conference on Image Processing (ICIP), Lausanne, Vol. 17, p. 2–5, 1996.
- [12] Menq, C. H., Yau, H. T., and Lai, G. Y., Automated precision measurement of surface profile in CAD-directed inspection, IEEE Transactions on Robotics Automation, Vol. 8, pp. 268–278, 1992.
- [13] Blais, G. and Levine, M. D., Registering multiview range data to create 3D computer objects, IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol. 17, pp. 820–824, 1995.
- [14] Masuda, T. and Yokoya, N., A robust method for registration and segmentation of multiple range images, Computer Vision and Image Understanding, Vol. 61, pp. 295–307, 1995.
- [15] Lu, F. and Milios, E. E., Robot pose estimation in unknown environments by matching 2D range scans, Proc. IEEE Computer Vision and Pattern Recognition (CVPR), Seattle, pp. 935–938, 1994.
- [16] Tarel, J.-P. and Boujemaa, N., A coarse to fine 3D registration method based on robust fuzzy clustering, Computer Vision and Image Understanding, Vol. 73, no. 1, pp. 14–28, 1999.
- [17] Neri, F. and Saitta, L., Exploring the power of genetic search in learning symbolic classifiers, IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol. 18, no. 11, pp. 1135–1141, 1991.
- [18] Grefenstette, J. J., Optimization of control parameters for genetic algorithms classifiers, IEEE Transactions on Systems, Man, and Cybernatics, Vol. 16, no. 1, pp. 122–128, January 1996.
- [19] Fourman, M. P., Compaction of symbolic layout using genetic algorithms, Proc. of International Conference of Genetic Algorithms and their Applications, pp. 141–153, 1985.
- [20] DeJong, K., Learning with genetic algorithms: An overview, Machine Learning, Vol. 3, pp. 121–138, 1988.

- [21] Ankenbrandt, C. A., Buckles, B. P., and Petry, F. E., Scene recognition using genetic algorithms with semantic nets, Pattern Recognition Letters, Vol. 11, pp. 285–293, 1990.
- [22] Bhanu, B. and Lee, S., Genetic Learning for Adaptive Image Segmentation, Kluwer Academic Publishers, Dordrecht, 1994.
- [23] Holland, J. H., Adaptation in Natural and Artificial Systems, University of Michigan Press, Michigan, 1975.
- [24] Goldberg, D. E., Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Welsey, Reading, MA, 1989.
- [25] Wells, W. M., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R., Multimodal volume registration by maximization of mutual information, Medical Imaging Analysis, Vol. 1, no. 1, pp. 35–54, March 1996.
- [26] Yamany, S. M. and Farag, A. A., Surface signatures: An orientation independent free-form surface representation scheme for the purpose of objects registration and matching, IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol. 24, no. 8, pp. 1105–1120, 2002.
- [27] Stein, F. and Medioni, G., Structural indexing: Efficient 3D object recognition, IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol. 14, no. 2, pp. 125–145, 1992.
- [28] Chua, C. S. and Jarvis, R., Point signatures: A new representation for 3D object recognition, International Journal of Computer Vision, Vol. 25, no. 1, pp. 63–85, 1997.
- [29] Dorai, C. and Jain, A. K., COSMOS; A representation scheme for 3D free-form objects, IEEE Transactions on Pattern Analysis and Machine Intellegence, Vol. 19, no. 10, pp. 1115–1130, October 1997.
- [30] Johnson, A. and Hebert, M., Surface matching for object recognition in complex three-dimensional scenes, Image and Vision Computing, Vol. 16, pp. 635–651, 1998.
- [31] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximization of mutual

information, IEEE Transaction on Medical Imaging, Vol. 6, no. 2, pp. 187–198, 1997.

- [32] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., Numerical Recipes in C: The Art of Scientific Computing, 2nd edn., Cambridge University Press, Cambridge, 1992.
- [33] Maes, F., Vandermeulen, D., and Suetens, P., Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information, Medical Imaging Analysis, Vol. 3, no. 4, pp. 373–386, 1999.
- [34] Eldeib, A., Yamany, S. M., and Farag, A. A., Multimodal medical volumes fusion by surface matching, Proc. Second International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'99), Cambridge, England, Sept 1999.
- [35] Noseworthy, J. H., Lucchinetti, C., Rodriguez, M., and Weinshenker, B. G., Multiple sclerosis, The New England Journal of Medicine, Vol. 343, no. 13, pp. 938–952, Sept 2000.
- [36] National Institute of Neurological Disorders and Stroke, Multiple Sclerosis: Hope through Research, Publication 96-75, NIH, July 2001.
- [37] Miller, D. H., Albert, P. S., Barkhof, F., Francis, G., Frank, J. A., Hodgkinson, S., Liblin, F. D., D. W. Paty, S. C. R., and Simon, J., Guidelines for the use of magnetic resonance techniques in monitoring the treatment of multiple sclerosis, Annals of Neurology, Vol. 39, no. 1, Jan 1996.
- [38] Boiselle, P. M. and White, C. S., New Technique in Thoracic Imaging, Marcel Dekker, New York, 2002.

# **Chapter 2**

# State of the Art of Level Set Methods in Segmentation and Registration of Medical Imaging Modalities

Elsa Angelini,<sup>1</sup> Yinpeng Jin,<sup>1</sup> and Andrew Laine<sup>1</sup>

### **2.1 Introduction**

Segmentation of medical images is an important step in various applications such as visualization, quantitative analysis and image-guided surgery. Numerous segmentation methods have been developed in the past two decades for extraction of organ contours on medical images. Low-level segmentation methods, such as pixel-based clustering, region growing, and filter-based edge detection, require additional pre-processing and post-processing as well as considerable amounts of expert intervention or information of the objects of interest. Furthermore, the subsequent analysis of segmented objects is hampered by the primitive, pixel or voxel level representations from those region-based segmentation [1].

Deformable models, on the other hand, provide an explicit representation of the boundary and the shape of the object. They combine several desirable features such as inherent connectivity and smoothness, which counteract noise and boundary irregularities, as well as the ability to incorporate knowledge about the object of interest [2, 3, 4]. However, parametric deformable models have two main limitations. First, in situations where the initial model and desired object boundary differ greatly in size and shape, the model must be reparameterized dynamically to faithfully recover the object boundary. The second limitation

<sup>&</sup>lt;sup>1</sup>Columbia University, New York, NY, USA

is that it has difficulty dealing with topological adaptation such as splitting or merging model parts, a useful property for recovering either multiple objects or objects with unknown topology. This difficulty is caused by the fact that a new parameterization must be constructed whenever topology change occurs, which requires sophisticated schemes [5, 6]. Level set deformable models [7, 8], also referred to as geometric deformable models, provide an elegant solution to address the primary limitations of parametric deformable models. These methods have drawn a great deal of attention since their introduction in 1988. Advantages of the contour implicit formulation of the deformable model over parametric formulation include: (1) no parameterization of the contour, (2) topological flexibility, (3) good numerical stability, (4) straightforward extension of the 2D formulation to n-D. Recent reviews on the subject include papers from Suri [9, 10].

In this chapter we give a general overview of the level set segmentation methods with emphasis on new frameworks recently introduced in the context of medical imaging problems. We then introduce novel approaches that aim at combining segmentation and registration in a level set formulation. Finally, we review a selective set of clinical works with detailed validation of the level set methods for several clinical applications.

# 2.2 Level Set Methods for Segmentation

A recent paper from Montagnat, Delingette, and Ayache [11] reviews the general family of deformable models and surfaces with a classification based on their representation. This classification has been reproduced to some extent in Fig. 2.1. Level set deformable models appear in this classification diagram as part of continuous deformable models with implicit representation.

### 2.2.1 Level Set Framework

Segmentation of an image *I* via active contours, also referred to as snakes [2], operates through an energy functional controlling the deformation of an initial contour curve C(p),  $p \in [0, 1]$  under the influence of internal and external forces achieving a minimum energy state at high-gradient locations. The generic energy



Figure 2.1: Geometric representations of deformable surfaces.

functional for active contour models is expressed as:

$$E(C) = \alpha \int_0^1 |C'(s)|^2 ds + \beta \int_0^1 |C''(s)| ds - \lambda \int_0^1 |\nabla I(C(s))|^2 ds \qquad (2.1)$$

where  $(\alpha, \beta, \lambda)$  are positive parameters. The first two terms control the rigidity and elasticity of the contour (defining the internal energy of the deformable object) while the last term attracts the model to high-gradient locations in the image *I* (defining the external energy of the model).

Active contour segmentation via minimization of the energy functional in Eq. (2.1) is typically implemented with a parametric framework in which the deformable model is explicitly formulated as a parameterized contour on a regular spatial grid, tracking its point positions in a Lagrangian framework [11].

In their original paper from 1988 [7], Osher and Sethian introduced the concept of geometric deformable models, which provide an implicit formulation of the deformable contour in a level set framework. To introduce the concept of the level set framework we focus on the boundary-value problem of a close contour C deforming with a speed V along its normal direction:

$$|\nabla C| V = 1, \quad V > 0 \tag{2.2}$$

Their fundamental idea is, instead of tracking in time the positions of the front C(x, y) on a regular grid as:

$$\Gamma(t) = \{(x, y) | C(x, y) = t\}$$
(2.3)

to embed the curve into a higher dimension function  $\phi(x, y, t)$  such that:

1. at time zero the initial contour  $C_0$  corresponds to the level zero of the function  $\phi$ :

$$C_0 = \{(x, y) / \phi(x, y, 0) = 0\}.$$
(2.4)

2. the function  $\phi$  evolves with the dynamic equation:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| V. \tag{2.5}$$

In this framework, at any time *t*, the front implicitly defined by:

$$\Gamma(t) = \{(x, y)/\phi(x, y, t) = 0\}$$
(2.6)

corresponds to the solution of the initial boundary value problem defined parametrically in Eq. (2.3). This result is illustrated in Fig. 2.2.



Figure 2.2: Correspondence between a parametric and implicit level-set formulation of the deformation of a contour with a speed term oriented along the normal direction. In their pioneer paper, Osher and Sethian focused on motion under mean curvature flow where the speed term is expressed as:

$$V = div\left(\frac{\nabla\phi}{|\nabla\phi|}\right). \tag{2.7}$$

Since its introduction, the concept of deformable models for image segmentation defined in a level set framework has motivated the development of several families of method that include: geometric active contours based on mean curvature flow, gradient-based implicit active contours and geodesic active contours.

### 2.2.2 Geometric Active Contours

In their work introducing geometric active contours, Caselles *et al.* [12] proposed the following functional to segment a given image *I*:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| g(|\nabla I|) \left( div \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + \nu \right), \tag{2.8}$$

with

$$g(|\nabla I(x, y)|) = \frac{1}{1 + |\nabla G_{\sigma}(x, y)^* I(x, y)|^2},$$
(2.9)

where  $\nu \geq 0$  and  $G_{\sigma}$  is a Gaussian convolution filter of standard deviation  $\sigma$ . The idea defining geometric deformable models is to modify the initial mean curvature flow of Eq. (2.7) by adding a constant inflation force term  $\nu$  and multiplying the speed by a term inversely proportional to the smooth gradient of the image. In this context the model is forced to inflate on smooth areas and to stop at high-gradient locations as the speed decreases towards zero.

### 2.2.3 Gradient-Based Level Set Active Contours

In their initial work on applications of the level set framework for segmentation of medical images, Malladi *et al.* [8] presented a gradient-based speed function for the general Hamilton-Jacobi type equation of motion in Eq. (2.5).

Their general framework decomposed the speed term into two components:

$$V = V_a + V_G, (2.10)$$

where  $V_a$  is an advection term, independent of the geometry of the front and  $V_G$  is a remainder term that depends on the front geometry.

The authors studied the design of the speed term to stop the front propagation at high-gradient locations depending on the value of  $V_G$ .

In the first case, for  $V_G = 0$  they proposed the following speed term:

$$V = \left(-V_a + \frac{V_a}{(M_1 - M_2)} (|\nabla G_{\sigma}^* I| - M_2)\right), \qquad (2.11)$$

where  $(M_1, M_2)$  are the maximum and minimum values of the smooth gradient image  $|\nabla G_{\sigma}^* I|$ .

In the case where  $V_G \neq 0$ , the speed term needs to be multiplied by a gradientbased term to stop the front evolution, as follows:

$$V = \left(\frac{1}{1 + \left|\nabla G_{\sigma}^* I\right|}\right) \times \left(V_a + V_G\right).$$
(2.12)

Numerical schemes for approximation of spatial derivatives with theses speed terms must respect the appropriate entropy condition for propagating fronts as discussed in detail in [13] and [14]. This entropy condition ensures that the propagating front corresponds to the boundary of an expanding region. An analogy invoked by Sethian to illustrate the entropy principle is to consider the moving front as a source for a burning flame and expand the flame thus ensuring that once a point in the domain is ignited, it stays burnt. The entropy principle puts some constraints in the choice of particular numerical schemes for temporal and spatial derivatives of the level set function. In their work, Malladi *et al.* [8] used a forward difference in time, upwind scheme for the constant inflation term and central differences for the remainder term (see Fig. 2.3).

A second issue with this framework arises from the fact the image-based speed terms are only defined on the zero-level of the moving front, as it was designed to stop the evolution of this level at high-gradient locations. On the other hand, the energy functional is defined over the entire domain and the speed term must have a consistent definition over all values of the level set function. This is done by extending the speed term from its values defined only on the level zero. There are several methods available to perform the extension. One of the most popular methods assigns the values of the closest point on the level zero to a given point of the domain.



Figure 2.3: Illustration of extension for gradient-based speed terms. (a) Region of interest from a chest MRI scan with a level set curve initialized inside the spine.(b) Gradient map of the MRI image with prior smoothing with a Gaussian filter.(c) Multiplying term for speed function, proportional to the inverse edge map from (b). (d) Extension of the speed term in (c) from values under the zero level contour.

### 2.2.4 Geodesic Active Contours

Geodesic active contours were introduced simultaneously by Kichenassamy *et al.* [15] and Caselles *et al.* [16] as a segmentation framework, derived from energy-based snakes active contours, performing contour extraction via the computation of geodesics, i.e., minimal distance curves in a Riemannian space derived from the image. Given an image *I* and for a given differentiable curve  $C(p), p \in [0, 1]$  they define the following energy:

$$E(C) = \int_0^1 g(|\nabla I(C(p))|) |C(p)| dp, \qquad (2.13)$$

where *g* is a positive decreasing function. Segmentation is achieved via minimization of this energy functional equivalent to the computation of geodesics in a Riemannian space according to a metric that weights the Euclidian length of the curve with the term  $g(|\nabla I(C(p))|)$ .

Minimization of the functional is performed via derivation of the Euler-Lagrange system:

$$\frac{\partial C}{\partial t} = g\left(|\nabla I|\right)\kappa\vec{N} - (\nabla g(|\nabla I|).\vec{N})\vec{N},\tag{2.14}$$

where  $\kappa$  is the Euclidian curvature of the curve *C* and  $\vec{N}$  is the unit normal vector to the curve. Implementation with a level set framework is performed by embedding the curve *C* into a level set function  $\phi$ . Using the following property on the curvature term:

$$\kappa = div \left(\frac{\nabla \phi}{|\nabla \phi|}\right),\tag{2.15}$$

and the following equivalence of relationships between a curve *C* and its associated level set function  $\phi$ :

$$\frac{\partial C}{\partial t} = \alpha \vec{N}$$

$$\frac{\partial \phi}{\partial t} = \alpha \left| \nabla \phi \right|,$$
(2.16)

the level set formulation is expressed as:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \left( g\left( |\nabla I| \right) div \left( \frac{\nabla \phi}{|\nabla \phi|} \right) + \nabla g\left( |\nabla I| \right) \frac{\nabla \phi}{|\nabla \phi|} \right) 
= |\nabla \phi| div \left( g\left( |\nabla I| \right) \frac{\nabla \phi}{|\nabla \phi|} \right)$$
(2.17)

To improve convergence speed and allow the detection of non-convex objects, the authors also introduced a modification of the initial formulation with the introduction of a constant inflation term  $\nu g(|\nabla I|) |\nabla \phi|$  leading to the following functional:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \left( div \left( g \left( |\nabla I| \right) \frac{\nabla \phi}{|\nabla \phi|} \right) + \nu g \left( |\nabla I| \right) \right) 
= g \left( |\nabla I| \right) (\kappa + \nu) |\nabla \phi| + \nabla g \left( |\nabla I| \right) \nabla \phi.$$
(2.18)

Applications of the geodesic deformable model to medical imaging have been tested by both groups of pioneering authors. Yezzi *et al.* tested their geodesic deformable model in [17] on 2D images for cardiac MRI, breast ultrasound with

a cyst and bone CT. No clinical validation was performed. Caselles with Malladi *et al.* in [18] compared 2D geometric and 3D geodesic deformable models with applications on 3D CT of human thighs and 3D cardiac MRI data sets. Measurement of soft tissue and ventricular cavity volumes are reported but no clinical validation was performed. A recent review of the use of geodesic deformable models for medical image analysis is provided in [19] with comparison of performance between geometric and parametric deformable models.

# 2.2.5 Tuning Level Set Speed Functions for Segmentation

The main problem of the boundary-based level set segmentation methods is related to contour leakage at locations of weak or missing boundary data information. An illustration of the phenomenon is provided in Fig. 2.4 for segmentation of a high-resolution abdominal MRI slice.

Several efforts have been performed to add stopping criteria on the entire front [20, 21] and local pixel freezing rules [21], or combine gradient with region



Figure 2.4: Leakage of level set deformable model at location of weak edges with gradient-based speed terms. (a) T2-weighted abdominal MRI with region of interest selected to contain subcutaneous fat. (b) Edge map derived from gradient computation to define speed term. (c) Leakage of front outside the fat compartment at two locations due to interstices with poor edge contrast. The level zero curve used to initialize the segmentation is displayed with a thick line.

information [22, 23] to make the segmentation process more robust to poor edge definition. When dealing with weak boundaries the most radical solution to leaking problems is to remove the expansion term at the cost of requiring an initialization close to the final solution [24]. An alternative to this approach was proposed by Jin et al. [25] initially keeping the expansion term for pushing the model and turning it off as it approaches the object boundary. Detection of the boundary location was performed using a homogeneity map derived from scalebased fuzzy connectivity [26]. A more recent effort to address the problem of segmentation of an object with missing boundaries was presented by Sarti et al. in [27] introducing a new geometric model for subjective surfaces. Starting from a reference point inside the object to segment, the "point of view", a geometric deformable model is evolved with mean curvature flow and image-derived speed terms until a piecewise constant solution is reached. This piecewise constant solution is the subjective surface defined by the segmentation process that is flat inside the object and has boundaries defined by geodesic curves. The authors also introduced the notion of "modal" contours which are contours that are perceived in the visual context and "amodal" contours which are associated with partially occluded objects. Segmentation of amodal contours can be performed with their subjective surface framework through iterations of edge-map computation and contour extraction. The authors produced very nice illustrations of the performance of their subjective surface segmentation on threedimensional ultrasound data with a fetal echogram, recovering the shape of the fetus.

All the level set segmentation methods presented above are based on image gradient intensity making them prone to leaking problems in areas with low contrast. A second problem related to the use of the image gradient as the only image-derived speed term is that it makes the segmentation process very sensitive to the initial position of the level set function as the model is prone to converge to false edges that correspond to local minima of the functional. Medical images typically suffer from insufficient and spurious edges inherent to physics of acquisition and machine noise from different modalities.

Two approaches can be followed to address these limitations. The first approach is to fuse regularizer terms in the speed function as reviewed in [9]. A second approach is to reformulate the problem in terms of region-based segmentation methods derived from the Mumford-Shah functional implemented in

a level set framework. We give an overview of these two families of methods in the next section.

### 2.2.6 Level Set Speed Functions with Regularizers

Suri *et al.* review in [9] recent works on the fusion of classical geometric and geodesic deformable models speed terms with regularizers, i.e., regional statistics information from the image. Regularization of the level set speed term is desirable to add prior information on the object to segment and prevent segmentation errors when using only gradient-based information in the definition of the speed. Four main types of regularizers were identified by the authors of the review:

- 1. Clustering-based regularizers
- 2. Bayesian-based regularizers
- 3. Shape-based regularizers
- 4. Coupling-surfaces regularizers

We give in the next section a brief overview of each method.

(1) *Clustering-based Regularizers*: Suri proposed in [28] the following energy functional for level set segmentation:

$$\frac{\partial \phi}{\partial t} = (\varepsilon \kappa + V_p) \left| \nabla \phi \right| - V_{\text{ext}} \nabla \phi, \qquad (2.19)$$

where  $V_p$  is a regional force term expressed as a combination of the inside and outside regional area of the propagating curve. This term is proportional to a region indicator taking value between 0 and 1, derived from a fuzzy membership measure as described in [29].

(2) *Bayesian-based Regularizers*: Recent work from Baillard *et al.* [30] proposed an approach similar to the previous one where the level set energy functional expressed as:

$$\frac{\partial \phi}{\partial t} = g\left(|\nabla I|\right)(\kappa + V_0)\left|\nabla \phi\right| \tag{2.20}$$

uses a modified propagation term  $V_0$  as a local force term. This term was derived from the probability density functions inside and outside the structure to
segment. The authors also modified the data consistency term  $g(|\nabla I|)$  as expressed in Eq. (2.9) using a transitional probability from going inside to outside the object to be segmented.

(3) Shape-based Regularizers: Leventon *et al.* [31] introduced shape-based regularizers where curvature profiles act as boundary regularization terms more specific to the shape to extract than standard curvature terms. A shape model is built from a set of segmented exemplars using principle component analysis applied to the signed-distance level set functions derived from the training shapes. The principal modes of variation around a mean shape are computed. Projection coefficients of a shape on the identified principal vectors are referred to as shape parameters. Rigid transformation parameters aligning the evolving curve and the shape model are referred to as pose parameters. To be able to include a global shape constraint in the level set speed term, shape and pose parameters of the final curve  $\phi^*(t)$  are estimated using maximum *a posteriori* estimation. The new functional is derived with a geodesic formulation as in Eq. (2.18) with solution for the evolving surface expressed as:

$$\phi(t+1) = \phi(t) + \lambda_1 \left( g\left( |\nabla I| \right) (c+\kappa) |\nabla \phi(t)| + \nabla g\left( |\nabla I| \right) . \nabla \phi(t) \right) + \lambda_2 (\phi^*(t) - \phi(t)),$$
(2.21)

where  $(\lambda_1, \lambda_2)$  are two parameters that balance the influence of the gradientcurvature term and the shape-model term. In more recent work, Leventon *et al.* [32] introduced further refinements of their method by introducing prior intensity and curvature models using statistical image-surface relationships in the regularizer terms. Limited clinical validation have been reported using this method but some illustrations on various applications including segmentation of the femur bone, the corpus callosum and vertebral bodies of the spine showed efficient and robust performance of the method.

(4) *Coupling-surfaces Regularizers*: Segmentation of embedded organs such as the cortical gray matter in the brain have motivated the introduction of a level set segmentation framework to perform simultaneous segmentation of the inner and outer organ surfaces with coupled level set functions. Such method was proposed by Zeng *et al.* in [33]. In this framework, segmentation is performed with the following system of equations:

$$\begin{cases} \phi_{\rm in} + V_{\rm in} |\nabla \phi_{\rm in}| = 0\\ \phi_{\rm out} + V_{\rm out} |\nabla \phi_{\rm out}| = 0 \end{cases}$$
(2.22)

where the speed terms ( $V_{in}$ ,  $V_{out}$ ) are functions of the surface normal direction (e.g., curvature of the surface), image-derived information and the distance between the two surfaces. When this distance is within the desired range, the two surfaces propagate according to the first two terms of the speed term. When the distance is out of the desired range, the speed term based on the distance controls the deformation as to correct the surface positions.

When defining the initial level set function as the signed distance function to its level zero, and ensuring that the distance function is preserved during the deformation process of the front through reinitialization, the distance of any point on the inner surface to the outer surface is directly read as the value of the outer level set function and vice versa.

Defining the speed terms as:

$$\begin{cases} V_{\rm in} = F_{\rm in} h\left(\phi_{\rm out}\right) \\ V_{\rm out} = F_{\rm out} h\left(\phi_{\rm in}\right) \end{cases}$$
(2.23)

with  $(F_{in}, F_{out})$  speed terms derived from image and curvature properties and h() a smooth approximating the windowing step function defined for a range of distance  $[d_1 \ d_2]$  that is equal to one inside this interval and 0 outside.

Zeng *et al.* [33] applied this framework for the segmentation of brain cortical gray matter (GM) surfaces. In this application, the speed terms were defined as:

$$\begin{cases} V_{\text{inside}} = S^{-} \left( I - I_{\text{in}} \right) + S^{+} \left( \phi_{\text{out}} + \varepsilon \right) \\ V_{\text{outside}} = S^{-} \left( I - I_{\text{out}} \right) + S^{+} \left( \phi_{\text{in}} - \varepsilon \right), \end{cases}$$
(2.24)

where *I* is the intensity of the MRI,  $I_{in}$  is a threshold value corresponding to the white matter and  $I_{out}$  a threshold value corresponding to the gray matter,  $\varepsilon$  is the desired thickness of the gray matter layer,  $(S^-, S^+)$  are two sigmoid functions, respectively, decreasing and increasing with bounded value between [-1, 1]. If the curve evolution is implemented with Eq. (2.5), the magnitude of the gradients  $(|\nabla \phi_{in}|, |\nabla \phi_{out}|)$  will increase and the estimation of the distance between the zero-levels of the two functions will be overestimated, leading  $(\phi_{in}, \phi_{out})$  to get closer as they evolve and eventually collide until the level set functions are reinitialized. Results are illustrated on three regions of interest from three MRI slices and show very interesting results but no quantitative evaluation of the accuracy of the method was performed.

#### 2.2.7 Reconciling Level Set and Distance Function

In a recent paper [34], Gomes and Faugeras introduced a reformulation of the Hamilton-Jacobi equation of Eq. (2.5) underlying the level set initial formulation from Osher and Sethian [7] to eliminate problems related to reinitialization of the distance function and the need to extend the velocity field away from the level zero.

The fact that the solution to Hamilton-Jacobi equations of the form in Eq. (2.5) are not distance functions has been demonstrated formally in [35]. In [34] the authors provide two simple examples illustrating this result. There are both theoretical and practical reasons pointed out by the authors to motivate the preservation of the signed distance function during the segmentation process. Theoretically, the signed distance function gives a unique equivalence to the implicit description of the moving front. From a practical point of view, the use of a signed distance function enables to directly extract from the level set function geometrical properties of the front and guarantees bounded values of the level set function gradient, ensuring numerical stability of the segmentation iterative process.

To derive the new dynamic equation, the authors initialize the level set function  $\phi_0 = \phi(x, 0)$  at t = 0 as the signed distance function from the initial front. The goal is to redefine a speed function F such that  $\frac{\partial \phi}{\partial t} = F$  which (1) preserves  $\phi$ as the signed distance function from the level zero, and (2) ensures that the level zero of  $\phi$  evolves as in Eq. (2.2). These constraints are expressed mathematically as:

$$\begin{cases}
F_{/\phi=0} = V \\
\frac{\partial \phi}{\partial t} = F \\
|\nabla \phi| = 1
\end{cases}$$
(2.25)

where  $F_{\phi=0}$  denotes the restriction of F to the zero-level of  $\phi$ . The authors derived the following dynamic equation as the solution to this system:

$$\frac{\partial \phi}{\partial t} = V \left( x - \phi \nabla \phi \right) \tag{2.26}$$

for any point  $x \in \mathbb{R}^3$ , which is not a Hamilton-Jacobi equation.

Implementation of the equation is proposed with a narrow-band framework, shock-detecting gradient computation and as described in [14].

The authors provide a very nice application for segmentation of cortical gray matter surfaces from MRIs derived from the initial work of Zeng *et al.* [33]. With this method, if the curve evolution in Eq. (2.24) is implemented with model in Eq. (2.5), the magnitude of the gradients  $(|\nabla \phi_{in}|, |\nabla \phi_{out}|)$  will increase and the estimation of the distance between the zero-levels of the two functions will be overestimated, leading  $(\phi_{in}, \phi_{out})$  to get closer as they evolve and eventually collide until the level set functions are reinitialized. Results are illustrated on three regions of interest from three MRI slices and show very interesting results but no quantitative evaluation of the accuracy of the method was performed.

#### 2.2.8 Region-based Level Set Active Contours

Region-based active contour were derived from the Mumford-Shah segmentation framework initially proposed in [36]. In their initial work, Mumford and Shah defined a new segmentation framework performing segmentation of a given image I into a set of contours S and a smooth approximation f of the image via minimization of the following framework:

$$E(S, f) = \alpha \int_{\Omega} (f - I)^2 dx + \beta \int_{\Omega \setminus S} |\nabla f| \, dx + H^{n-1}(S), \qquad (2.27)$$

where  $H^{n-1}(S)$  is the (n-1) dimensional Hausdorff measure, and  $(\alpha, \beta)$  are positive real parameters. In this functional, the first term ensures that f is a good approximation of the original image I, the second term ensures that f is smooth and the last term minimizes the length of the set of contours of the segmentation. This type of region-based segmentation method relies on the homogeneity of the object to segment. This assumption is often violated with medical images due to motion of the organ, presence of corrupting machine noise or acquisition artifacts that introduce flat field inhomogeneities.

Based on the Mumford-Shah segmentation framework, Chan and Vese introduced in a series of papers a new type of active contour models without gradient information [37–41]. In the simplest case, assume that an image I defined on  $\Omega$  is composed of two regions (e.g., an object and a background) with homogeneous intensities around values  $c_0$  and  $c_1$ . Given a curve C that defines the boundary of a region inside the image I, they introduce the following homogeneity-based functional:

$$E(C) = \int_{\text{inside}C} |I - C_0|^2 d\Omega + \int_{\text{outside}C} |I - C_1|^2 d\Omega, \qquad (2.28)$$

where  $(C_0, C_1)$  are the average intensity values of *I* inside and outside the curve *C*. With this functional, the boundary between the two regions is defined by its minimum state. They further combined this homogeneity-based fitting term with regularizing terms that put constraints on the length and the area of the curve with the following functional:

$$E(C_0, C_1, C) = \lambda_0 \int_{\text{inside}C} |I - C_0|^2 d\Omega + \lambda 1 \int_{\text{outside}C} |I - C_1|^2 d\Omega + \mu \operatorname{length}(C) + \mu \operatorname{Area}(C).$$

$$(2.29)$$

Details for the mathematical definitions of the length and the area of the boundary curve C can be found in [40].

In a level-set framework implementation, the functional (2.29) is expressed as:

$$E(C_0, C_1, \phi) = \lambda_0 \int |I - C_0|^2 H(\phi) d\Omega + \lambda_1 \int |I - C_1|^2 (1 - H(\phi)) d\Omega + \mu \int \delta(\phi) |\nabla \phi| d\Omega + \nu \int H(\phi) d\Omega$$
(2.30)

Advantages of this method include the possibility of segmenting objects with discontinuous edges and robustness of the method to arbitrary initialization, avoiding the problem of local minima at spurious edge locations or leakage of the model at missing edge locations. The initial work from these authors have generated many applicative research works for segmentation of medical images, starting with works from the authors themselves in [37] with illustration of their method on brain MRI, three-dimensional ultrasound.

A simultaneous and parallel effort to the work of Chan and Vese, from Tsai *et al.* [42] proposed a reformulation of the Mumford-Shah functional from a curve evolution perspective using a gradient flow formulation and a level set framework implementation. Recent works applying this segmentation method to three-dimensional cardiac ultrasound include Angelini *et al.* [43], and Lin *et al.* [44].

We note two powerful extensions of this region-based implicit deformable model for applications to medical images:

- (1) This method is easily extended to segmentation of vectorial images with integration of the multiple channels information in the homogeneity measure. This property, described in detail in [45], has potential applications in segmentation of multiprotocols MRI brain data sets or any co-registered multimodality data sets where combination of spatial information can assist the definition of a particular organ contours.
- (2) This method is extensible to multiphases segmentation using a system of n coupled dynamic PDEs with {φ<sub>1</sub>,..., φ<sub>n</sub>} level set functions defining 2<sup>n</sup> phases in the segmented data. Extensive description of the multiphase method is provided in [41]. Potential applications of the multiphase formulation include segmentation of brain MRIs into multiple tissue types. An illustration of this application is provided in Fig. 2.5.

# 2.3 Joint Image Registration and Segmentation

#### 2.3.1 Motivations

Combining registration and segmentation has been motivated by the need to incorporate prior information to guide and constrain the segmentation process. The quality of the images acquired by the various medical screening modalities is often poor due to the presence of multiple noise sources in the acquisition system, degradation of data content during reconstruction processes (e.g., tomographic reconstruction with Radon transform), motion and respiratory artifacts introduced by motion of the patient, and inherent limitations of system acquisition accuracy. The combination of these factors degrade the signal to noise ratio of the data, limit the spatial resolution, introduce inhomogeneities in the tissue appearance across volumetric slices, and deteriorate boundary definitions between specific organs and their surrounding tissues. These issues are encountered with other medical imaging modalities such as ultrasound, MRI, PET and SPECT and CT.

In the context of brain MRI segmentation for example, incorporation of atlas information to assist the segmentation task of a particular data set has been a very successful and popular approach for many years as reviewed in [46]. For organs with very characteristics shapes such as cardiac ventricles, the corpus



Figure 2.5: Multiphase segmentation of brain MRI with region-based implicit deformable model. (a) Diagram illustrating the definition of two phases with one level set function. (b) Diagram illustrating the definition of 4 phases with 2 level set functions. (c) Original axial slice from 3D brain MRI data set with initialization of the two level set functions. (d) Results of implicit deformable model for from each phase: WM (1st line), GM (2nd line), CSF (3rd line) and background (4th line). (e) Manually labeled data. (f) TP/FP error maps.

callosum in the brain, or cartilages of the knee, shape priors (including active shape models, active appearance models and statistical shape descriptors) have been used with great success in the context of constrained segmentation [47–50].

The use of an atlas (or a shape model) to assist the segmentation process requires that the target image data and the atlas (or the model) are being aligned via either preregistration or via a new concept of combined registration and segmentation. When considering registration as a pre-processing step, common atlas-based segmentation methods use warping of the atlas to the target data via minimization of mean square errors of image pair intensities at control points. Alternative popular techniques use robust estimators, optimization of correlation ratios, optical flow, fluid-flow non-rigid deformation models and mutual information methods to construct statistical deformation models. An extensive review of registration methods applied to medical imaging can be found in [51]. Among recent work in this area we mention here the method of Vemuri *et al.* [52] who derived a novel curve evolution approach in a level set framework for image intensity morphing and non-linear associated PDE for the corresponding coordinate registration between an atlas and an image. Applications of the method included a clinical study on segmentation of the corpus callosum via morphing of a shape model defined in the atlas space, after registration of the data with the proposed method.

In this chapter we focus on methods that explicitly combine segmentation and registration in a variational framework. By combining registration and segmentation, one can recover the image region that corresponds to the organ of interest, given a model of this structure. Level set deformable models offer a very flexible framework to propagate a moving front with segmentation-driven constraints while registering the segmentation result (i.e., the level zero curve) to a given model. Distance transforms have been successfully applied in the past to registration problems [53-55]. In a level set framework, Paragios has published several papers recently focusing on matching geometric shapes in a variational framework for global as well as local registration [56-58]. The first attempt at combining segmentation and registration in a single geometric deformable model framework might be attributed to Yezzi et al. in [59]. Their key observation is that multiple images may be segmented by evolving a single contour as well as the mappings of that contour into each image. In the context of level set framework, multiple recent works can be referenced that incorporate shape priors in the segmentation process as reviewed in [60]. The main trend of the reported efforts uses a shape model and incorporates a constraint in the energy of the geometric deformable model that forces the evolving contour to fit to the shape model [56, 61, 62]. In an effort to derive a rigorous and complete scheme, Paragios and Rousson [56] focused on the integration of a shape model, defined directly in a level set space, to derive a shape prior in an energetic form and integrate it with a data-driven variational segmentation framework. Applications of their combined registration and segmentation framework focused on the segmentation of physically corrupted or incomplete natural images.

In this chapter we selected to focus on recent works applied to segmentation and registration of medical images as this application typically involves tuning of a general framework to the specificity of the task at hand. We describe in details two different approaches in the next sections.

# 2.3.2 Shape Priors into a Variational Segmentation Framework

Several applications in medical imaging can benefit from the introduction of shape priors in the segmentation process using deformable models [49; 63–65]. Only few works on segmentation of medical imaging with level set framework attempted to perform simultaneous registration and segmentation into a single energy functional and we review three of them in this section.

We first review the work of Chen *et al.* [60, 66, 67] that proposes a Mumford-Shah type energy functional plus a parameterized registration term embedded in a level set formulation for segmentation of brain MRI. Their approach consists of constraining the segmentation process with a level set framework by incorporating an explicit registration term between the detected shape and a prior shape model. They proposed two approaches either with a geodesic, gradient-based active contour or with a Mumford-Shah region-based functional.

The geodesic active contour minimizes the following functional:

$$E(C, s, R, T) = \int_0^1 \left\{ g(|\nabla I|) C(p) + \frac{\lambda}{2} d^2 (sRC(p) + T) \right\} |C(p)| dp \qquad (2.31)$$

with C(p) a differentiable curve parameterized with  $(p \in [0, 1])$  defined on image I, g a positive decreasing function, (s, R, T) are rigid transformation parameters for scale, rotation and translation and  $d(C(p)) = d(C^*, C(p))$  is the distance between a point C(p) on the curve C and the curve  $C^*$  representing the shape prior for the segmentation task. A level set formulation is derived by embedding the curve C into a level set function  $\phi$  positive inside the curve. Let's introduce the Heaviside function H(z) = 1 if  $z \ge 0$ , H(z) = 0 otherwise, and the Dirac measure  $\delta(z) = H(z)$  (with derivative in the distribution sense), the energy functional in Eq. (2.31) is reformulated as:

$$E(\phi, \mu, R, T) = \int_{\Omega} \delta(\phi) \left( g(|\nabla I|) + \frac{\lambda}{2} d^2 (\mu R x + T) \right) |\nabla \phi|.$$
(2.32)

Four evolution equations are derived for the Euler-Lagrange system for temporal derivatives of the level set function  $\phi$  and the rigid registration parameters ( $\mu$ , R, T) with detailed numerical implementation described in [68]. In this paper the authors also report on experiments performed with this method to segment the endocardial borders of the left ventricle on an ultrasound image and segment the corpus callosum on misaligned functional MRI images in a time series.

In [60], the authors proposed a second functional for combining registration and segmentation in an implicit deformable model framework where the image gradient term is replaced by an homogeneity measure. Their approach is derived from the Mumford–Shah functional [36] in a similar fashion as described in the previous section of region-based level set methods. They proposed the following functional:

$$E(\phi, c_{0}, c_{1}, \mu, R, T) = \alpha \int_{\Omega} H(\phi) \left(I - S^{+}\right)^{2} + \alpha \left(1 - H(\phi)\right) \left(I - S^{-}\right)^{2} + \beta H(\phi) \left|\nabla S^{+}\right| + \beta \left(1 - H(\phi)\right) \left|\nabla S^{-}\right| + \delta(\phi) d^{2} \left(\mu Rx + T\right) \left|\nabla \phi\right| d\Omega$$
(2.33)

with  $(S^+, S^-)$  smooth approximations of the image *I* on, respectively,  $\{x/\phi(x) > 0\}$  and  $\{x/\phi(x) < 0\}$  and  $(\alpha, \beta)$  a set of positive parameters.

In a similar effort to combine registration and segmentation Paragios *et al.* [69] proposed a level-set approach for knowledge-based registration and segmentation of the left ventricle. In their method a level set framework was used to perform simultaneous segmentation of the epicardial and the endocardial surfaces of the myocardium muscle via coupling two level set functions ( $\phi_0, \phi_1$ ). The proposed functional is generalized as:

$$E(\phi_0, \phi_1, A_0, A_1) = \alpha E_G(\phi_0, \phi_1) + \beta E_R(\phi_0, \phi_1) + \gamma E_A(\phi_0, \phi_1) + \delta E_S(\phi_0, \phi_1, A_0, A_1).$$
(2.34)

This functional integrates four components:

- 1.  $E_G$  is a constraint on the regularity of the contour via minimization of its length.
- 2.  $E_R$  is an intensity-based region component. This component identifies the partition of the image into regions that maximizes the posterior segmentation probability given a priori gray level histogram distribution for the endocardium, the epicardium and the background;

- 3.  $E_A$  is an anatomy-driven constraint. It preserves the distance between the myocardium border surfaces within an admissible range of values.
- 4.  $E_S$  is a shape-driven global consistency constraint. This knowledge-based term performs a registration of the evolving contour to a prior shape model  $(A_0, A_1)$  via rigid deformation. The prior shape models were defined in a pixel-wise stochastic level set representation [56].

The authors reported some experiments on segmentation of the endocardium from 2D cardiac MRI images. These experiments revealed that the anatomical constraint played a minor role in controlling the deformation of the segmenting surface and that the regularity term was overwritten by the shape prior term. Final segmentation results showed a reliable performance of the method but no quantitative validation was performed. It was pointed out by the authors that the integration of the different modules was difficult and that future refinements of the approach were considered such as the use of a single level set function for segmentation of the myocardium, and tracking contours in time by replacing the prior shape model with the segmentation from the previous frame in the context of consecutive time frames segmentation over a cardiac cycle.

In two related papers, Paragios [63, 70] proposed modified versions of the method.

In [70] Paragios had proposed a version of the method where the regularity term consisted of a boundary component derived from gradient vector flow [71] to detect cardiac boundaries and curvature constraints on the segmented shape. No shape-driven constraint was proposed in this early work.

In a posterior work [63], Paragios modified the method for segmentation of the endocardial surface on ultrasound. The model was first modified to replace the regularity term  $E_R$  by a boundary constraint  $E_B$  derived from a geodesic active contour formulation [16]. The model was further modified to integrate temporal tracking of the segmented contours between consecutive time frames. A time-tracking constraint, in the form of a bounded error function using a robust norm  $\rho$  was introduced as:

$$E_T(\phi^t, \phi^{t+1}, T) = \int H(\phi^t) \rho(I^t - I^{t+1}(T)) d\Omega + \int H(\phi^{t+1}) \rho(I^t(T^{-1}) - I^{t+1}) d\Omega$$
(2.35)

where *H* refers to the Heaviside function (equal to 1 for negative values and 0 on positive values) and *T* is an optimal transformation to track the targeted structure of interest between to consecutive time-frames images  $I^t$  and  $I^{t+1}$  satisfying the visual consistency constraint:

$$I^{t}(x, y) \approx I^{t+1}(x, y), \ \forall (x, y) / H(\phi^{t}(x, y)) \ge 0$$
(2.36)

with  $\phi$  defined with negative values inside the object to segment (i.e., the ventricle blood cavity in this case).

This work uses a shape-model defined in a level set framework. Several interesting recent efforts have focused on the use of level set framework for shape modeling and registration toward model-based shape-driven object extraction as reviewed in [58].

# 2.3.3 Registering Contours for Multimodalities Segmentation

In a recent paper Yezzi *et al.* [59] introduced a new variational deformable model framework that interleaves segmentation and feature-based registration for combined segmentation of a single organ from multiple screening modalities (e.g., skin surface from head CT and MRI).

They defined their problem as follows: They want to find closed surfaces S and  $\hat{S}$  to segment an object in images I and  $\hat{I}$  so that the curves, segmenting the same organ, are related through a geometrical mapping:  $\hat{S} = g(S)$ . The authors used rigid registration for the mapping (i.e., combination of rotation and translation) and defined the following coupled functionals for the surface S and the registration parameters  $g = [g_1, g_2, \ldots, g_n]$ :

$$\frac{\partial S}{\partial t} = \left( f(x) + \hat{f}(g(x)) \right) N - \kappa N$$

$$\frac{dg_i}{dt} = \int_S \hat{f}(g(x)) \left\langle \frac{\partial g(x)}{\partial g_i}, \hat{N} \right\rangle dA$$
(2.37)

where  $\kappa$  and dA denote the mean curvature and area element of the surface S,  $(N, \hat{N})$  are the unit normals of  $(S, \hat{S})$ . The registration vector is modelized as:

$$g\left(x\right) = Rx + T \tag{2.38}$$

with  $R = R_X R_Y R_Z$  is the combination of rotations around the three orthogonal axis (X, Y, Z) defining the 3D domain, and  $T = [T_X, T_Y, T_Z]$  is the translation vector in each axis direction.

The function f(x) is defined in a homogeneity-based framework [40] as:

$$f(x) = f_{\rm in}(x) - f_{\rm out}(x).$$
(2.39)

The functions  $f_{in}$  and  $f_{out}$  are defined as:

$$f_{\rm in} = (I - u)^2 f_{\rm out} = (I - v)^2$$
(2.40)

where u and v denote the mean values of the image I inside and outside the surface S. Analogous definitions of  $\hat{f}$  with statistics on  $\hat{I}$  are also derived.

The authors reported three experiments on simultaneous segmentation and registration of MRI/CT images of the head and the spine both in 2D and 3D. Validation via visual inspection showed accurate contour extraction for these limited experiments.

# 2.4 Review of Clinical Validations

In this section we review in detail several recent papers that apply level set segmentation and registration methods to medical images and provide a detailed validation of their method through a clinical study for qualitative and quantitative assessment of the accuracy of the method in assisting or performing a particular clinical diagnosis task.

### 2.4.1 Important Clinical Segmentation Problems

We introduce in some details the two major applications in the domain of segmentation of clinical images: Segmentation of the brain and segmentation of the left ventricular cardiac cavity.

#### 2.4.1.1 Segmentation of Brain Images

The two major modalities used for brain screening are MRI and SPECT/PET.

Regarding MRI brain imaging, the paper in [72] gives a very nice review of the potential applications of MRI in quantification of brain disease. The brain cortex is a highly convoluted layer of gray matter that lies between the white matter and the cerebrospinal fluid (CSF). Clinical applications require reconstruction of the cerebral cortex from MRI data for:

- 1. brain visualization
- 2. analysis of brain geometry
- 3. registration with other data sets (multimodality or repetitive scans of a single patient)
- 4. surgical planning
- 5. cortex mapping

The interface between WM and GM is clearly visible on T1-weighted MRIs. Difficulties of MRI segmentation arise from imaging noise, inhomogeneities, partial volume effects and the highly convoluted geometry of the cortex.

Regarding quantitative measurements of the brain anatomy [33] using MRI, whole brain volume, cortical gray matter volume, white matter volume, cortical surface area, cortical shape characteristics and cortical thickness map are among the most interesting to study brain anatomy and function. Such measurements can typically assist in characterizing, predicting or assessing neurological and psychiatric disorders via correlation to abnormality in the measurements. These measurements are all easily derived from the final level set function in a distance-preserving framework (typically ensured by reinitialization of the level set function during the iterative deformation process).

Two sources of MRI brain data for testing segmentation algorithms are available as open source databases on the web:

• The Internet Brain Segmentation Repository (IBSR) available at <u>http://www.cma.mgh.harvard.edu/ibsr/</u>. This repository provides manually-guided expert segmentation results along with magnetic resonance brain image data. Its purpose is to encourage the evaluation and development of segmentation methods. The IBSR is supported by the National Institute of Neurological Disorders and Stroke at the NIH part of a grant that funds research in MR brain segmentation by researchers at Boston University, Draper Laboratory, Northeastern University, Massachusetts Institute of Technology, and Massachusetts General Hospital/Harvard Medical School. The IBSR is a World Wide Web resource providing access to magnetic resonance brain image data and segmentation results contributed and utilized by researchers from all over the world. Its purpose is to encourage the development and evaluation of segmentation methods by providing raw test and image data, human expert segmentation results, and methods for comparing segmentation results.

• The **Brain Web:** Simulated Brain Database (SBD) from the McConnell Brain Imaging Centre Montréal Neurological Institute, at McGill University (http://www.bic.mni.mcgill.ca/brainweb/). This database contains a set of realistic MRI data volumes produced by an MRI simulator [73–76]. These data can be used by the neuro-imaging community to evaluate the performance of various image analysis methods in a setting where the truth is known. Currently, the SBD contains simulated brain MRI data based on two anatomical models: normal and multiple sclerosis (MS). Full threedimensional data volumes have been simulated using three sequences (T1-, T2-, and proton-density- (PD-) weighted) and a variety of slice thicknesses, noise levels, and levels of intensity non-uniformity. These data are available for viewing in three orthogonal views (transversal, sagittal, and coronal), and for downloading. Customization of the MRI simulations is also available allowing the user to run his own custom MRI simulation with any of several pulse sequences and source digital phantoms, and arbitrary values of the acquisition artifacts.

Regarding PET imaging, this modality uses small amounts of tracer drugs chemically attached to glucose or other compounds injected into the patient. As the tracer travels through the body, it emits signals and eventually collects in the organs targeted for examination. If an area in an organ is cancerous, the signals will be stronger than in the surrounding tissue. A scanner records these signals and transforms them into pictures of chemistry and function.

PET modality is used for brain screening for:

- diagnosis of Alzheimer's disease,
- location of tumors and scar tissue in the brain,

- location of seizures focus for patients with epilepsy, and
- accurate assessment of tumor sites in the brain for surgery planning.

In practice, segmentation of PET brain data is difficult to perform in an automatic fashion because of the poor contrast and high noise level in the images. A standard procedure to delineate structures from PET brain images is to segment structures from the corresponding anatomical magnetic resonance images and then to superimpose them on the PET images. This method relies on an accurate registration between the two imaging modalities, handling the incongruity of structures and functions. Thresholding methods can provide a direct way to segment PET images with locally uniform radioactivity concentration and consistent structures. For more noisy data, advanced methods such as deformable models have been studied for segmentation of cortical structure [77]. These methods must adapt to changes in individual radioactivity concentrations.

#### 2.4.1.2 Segmentation of Cardiac Images

A reliable noninvasive imaging modality is essential for evaluating and monitoring patients with cardiac disease. Traditional screening techniques for quantitative assessment of cardiac function include the following modalities:

- *Multigated Angiography (MUGA)*: This is a slow screening modality that requires the injection of a radiopharmaceutical agent by a clinician. The purpose of MUGA screening is to examine the pumping function of the heart. After injection of a radioactive agent that labels red blood cells, a nuclear camera creates an image of the heart's chambers by counting these cells over numerous cardiac cycles. The information obtained can quantify ejection fraction but not ventricular volumes.
- *Magnetic Resonance Imaging (MRI)*: Because of its complexity and even though MRI machines abound in the United States, cardiac MRI has largely been limited to university hospitals where there is a strong interest in research. This screening modality has proven very useful in evaluating patients' cardiac anatomy prior to surgery, in locating and characterizing cardiac tumors and in identifying and treating cardiac abnormalities for children with complex congenital heart disease. These clinical situations

are relatively rare and cardiac MRI has yet to become a commonly used tool in clinical medicine.

- *Computerized Tomography (CT)*: Multidetector technology has made cardiac CT possible enabling angiography, perfusion and function studies. The main limitation of this screening modality remains the acquisition time with multirow detectors, which may be solved with upcoming spiral CT, electron beam CT and ultrafast CT technologies.
- *Single Photon Emission Computed Tomography (SPECT)*: Commonly referred to as myocardial perfusion imaging, this technique is used to visualize myocardial blood flow distribution using intravenous injection of a radionuclide detected by single crystal gamma camera rotating around the patient's body. This modality can be used to assess ejection fraction and regional wall motion but cannot provide detailed views of anatomical structures.
- *Positive Emission Tomography (PET)*: Similar to SPECT, this technique visualizes myocardial blood flow using intravenous injection of positronemitting tracers detected by multiple rings of stationary detectors encircling the patient's body to produce a series of multiple tomographic images encompassing the heart. Specific tracers have been developed for the evaluation and quantification of numerous physiological processes, including regional myocardial blood flow, metabolic processes, oxygen consumption, receptor activity, and membrane function. When compared to SPECT, PET images have been shown to be more accurate in clinical studies but PET scanners remain costly and therefore less widely available than standard SPECT systems.
- *Two-Dimensional Echocardiography (2DE)*: Two-dimensional echocardiography is the fastest, least expensive, and least invasive screening modality for imaging the heart. Because of the three-dimensional structure and deformation of the heart muscle during the cardiac cycle, analysis of irregularly shaped cardiac chambers or description of valve morphology using 2D images is inherently limited. A second existing problem with 2DE, constrained to planar views, is that highly trained clinicians are required to perform the studies. But, despite its limited image quality and its limitation to planar acquisition, 2DE is one of the most

popular cardiac screening modalities, available at hospitals, medical centers and cardiologists' offices. Critical information for assessment of cardiac pathology such as ventricular shape, wall deformation, valve motion, and blood flow (via Doppler acquisition mode) can be quickly assessed with this non-invasive, portable and relatively inexpensive screening modality.

- *Three-dimensional Echocardiography (3DUS)*: Three-dimensional ultrasound was introduced in the late 1980s with offline 3D medical ultrasound imaging systems. Many review articles have been published over the past decade, assessing the progress and limitations of 3D ultrasound technology for clinical screening [78–81]. These articles reflect the diversity of 3D systems that were developed for both image acquisition and reconstruction. The evolution of 3D ultrasound acquisition systems can be divided into three generations:
  - Freehand scanning. With freehand probes, planar images are acquired at arbitrary spatial positions and orientation with a 2D transducer. A positioning device, attached to the transducer, tracks its location in space and time during the acquisition. A three-dimensional volume can then be reconstructed by associating each acquired image with its 3D spatial position and integrating in 3D space. This method offers a great scanning flexibility and provides high quality images. The principal limitations of this method are related to the precision of the positioning device (either mechanical, acoustic or magnetic) and the experience of the clinician in positioning the device to acquire sufficient amount of information for an accurate 3D reconstruction. These issues are most critical when scanning small moving structures such as valves and myocardium wall defect in cardiac applications.
  - Mechanical scanning. With mechanical scanning, a 2D transducer is moved at regular intervals along a specified path, ensuring an accurate sampling of the volume to reconstruct with a probe whose position is controlled in space and time. The three most common scanning paradigms use linear, fan, and rotational sweep. This technology provides a more accurate three-dimensional reconstruction of the anatomy than freehand scanning at the cost of a slower acquisition

rate, and a smaller field of view. When screening the beating heart, gating is necessary for both modalities in order to acquire data corresponding to similar instants in the cardiac cycle, so that the anatomy is approximately the same, assuming a periodic movement, for each planar view.

- Phased arrays. Real-time three-dimensional (RT3D) volumetric imaging is the only true three-dimensional ultrasound modality [82–87]. This technology, pioneered by Dr. Olaf Von Ramm and Dr. Stephen Smith at Duke University, is fundamentally different from the former generations of 3D systems as a volume is acquired with a 2D array of pulse transmitter and receiver elements, enabling the cardiologist to view moving cardiac structures from any given plane in real-time [82, 87–91]. Electronic scanning controls the acquisition depth allowing real-time signal acquisition through a pyramidal shape volume. This existing technique is still limited by hardware component size and speed. The low spatial resolution and high noise level have prevented this technology from meeting its initial expectation and reaching its full potential. It remains, nevertheless, the only true 3D ultrasound modality that can enable accurate temporal visualization of cardiac deformation during a single cardiac cycle.

In current clinical practice, cardiologists use anatomical images from CT, US and MRI modalities to quantify cardiac function through measurements of ventricular volumes at end diastole (ED) which corresponds to the end of the blood filling phase in the cardiac cycle, and end systole (ES) which corresponds to the end of the blood ejection phase, stroke volume (SV) which is equal to the difference of blood volumes and ED and ES and ejection fraction (EF) which is equal to SV over ED volume. These measurements are performed via simple visual inspection or manual tracing of 2D slices extracted from the 3D data. A second phase of analysis of the images aims at analyzing myocardium wall deformation and localization of abnormalities. Physiological images provided by MUGA, SPECT and PET/SPECT modalities aim at quantifying myocardium tissue blood perfusion for localization of ischemic tissue.

All these diagnostic tasks require the intervention of an expert cardiologist familiar with the modality for quick visual inspection. A more detailed processing of the data always require segmentation of the myocardium tissue versus the blood pool and sometimes the outside tissues. Few automated segmentation algorithms, with minimal manual intervention, are available on clinical consoles to assist the segmentation task with a significant saving of time. Segmentation of cardiac images is still a very active research area and level set segmentation methods have proved in the recent years to offer a very flexible threedimensional tool that can handle the volumetric and dynamic nature of the data.

# 2.4.2 Open Source Software Tools for Level Set Segmentation

#### 2.4.2.1 Snake Automatic Partitioning (SNAP)

This software was developed by the Medical Image Display and Analysis Group at the University of North Carolina and is available for download at www. midag.cs.unc.edu. SNAP is a segmentation tool for volumetric image data using 3D level set methods with either a region-probability deformable model or a gradient-based deformable model framework. Some interaction with parameter settings of the segmentation method and prior-filtering is available. Interactive visualization of the deformation process in provided.

#### 2.4.2.2 Insight Segmentation and Registration Toolkit (ITK)

The National Library of Medicine Insight Segmentation and Registration Toolkit (ITK) is an open-source software system to support the Visible Human Project. The toolkit is available for free download at www.itk.org. Under active development, ITK employs leading-edge segmentation and registration algorithms in multiple dimensions. The Insight Toolkit was developed by six principal organizations, three commercial (Kitware, GE Corporate R&D, and Insightful) and three academic (UNC Chapel Hill, University of Utah, and University of Pennsylvania). Additional team members include Harvard Brigham & Women's Hospital, University of Pittsburgh, and Columbia University. The funding for the project is from the National Library of Medicine at the National Institutes of Health. NLM in turn was supported by member institutions of NIH (see sponsors). Several level set segmentation methods are implemented in this toolkit including: fast marching methods, shape detection segmentation, geodesic

active contours, threshold level set, canny-edge level set and Laplacian level set methods.

#### 2.4.3 Applications to Clinical Studies

### 2.4.3.1 Robust Adaptive Segmentation of 3D Medical Images with Level Sets

This work was published by Baillard et al. in [30].

**Method:** The proposed method uses a 3D level set algorithm with the introduction of an adaptive adjustment of the time step and the external propagation force at each iteration. A region-based force is derived from intensity probability density functions over the data. Assumptions are made on the input data which is modeled as a mixture of distributions. Mixture of Gaussian distributions for MRI and Gaussian and Rayleigh distributions for ultrasound data are validated through two experiments. Each distribution defines a class  $c_k$  through a parameter vector that contains the distribution parameters and the probability  $p_k$  that a voxel belongs to class  $c_k$ . The parameters vector is estimated from the data using the stochastic expectation-maximization (SEM) algorithm [92], which is a stochastic version of the EM algorithm that utilizes probabilistic learning stage. Advantages of the SEM over the EM algorithm include: (1) Only an overestimation of the number of classes is required, (2) it is less dependent on the initialization. The stopping criterion for the deformation process is based on the stabilization of the average segmented volume size.

**Experiments:** Experiments were performed on brain MRI volumes. The statistical model was initialized with seven classes.

**1.** A first experiment used simulated brain MRIs from the MNI group [93]. Brain MRI volumes of size  $(181 \times 217 \times 181)$  simulating WM, GM and CSF were generated under noiseless conditions and three different combinations of noise and inhomogeneities. The segmentation method was applied to extract together GM and WM volumes. Initialization was performed by defining a large cube of size  $(100 \times 70 \times 70)$  inside the data volume. Gaussian distribution parameters for WM + GM were automatically estimated prior to segmentation. Quantitative validation was performed using overlapping measurements [94] between the result and the known ground truth on these phantom data sets. The measures included estimation of the number of true-positive (TP) true-negative (TN),

false-positive (FP) and false negative (FN) voxels and the definition of the following measures:

sensitivity = 
$$TP/(TP + FN)$$
  
specificity =  $TN/(FP + TN)$  (2.41)  
total performance =  $(TP + TN)/(TP + FP + TN + FN)$ 

These measures are very helpful to assess the global performance of a segmentation method such as under-segmentation characterized by a low sensitivity of over-segmentation characterized by a low specificity. The total performance of the proposed algorithm stabilized around 98.3% under all noise conditions. The authors further compared their segmentation performance to morphological operators performance and reported an improvement of sensitivity performance with the level set method.

**2.** A second set of experiments with a database of 18 real brain MRIs of size  $(256 \times 256 \times 176)$  was performed. Results reported a 94% success ratio of segmentation convergence (one case failed), requiring on an average 1,000 iterations. Segmentation of individual tissue classes (WM, GM and CSF) required a coarse approximation of tissue segmentation for class definition and computation of a priori statistical models.

**Limitations:** The proposed segmentation method has a performance limited by the fact that the SEM algorithm does not guarantee an optimal solution. In practice, an initial partitioning roughly representative of the inside and outside distributions of the organs to segment lead to a correct solution. This means that tissue classes need to be initialized with relatively accurate average intensity values.

### 2.4.3.2 Topology Preserving Geometric Deformable Models for Brain Reconstruction

This research work was published by Han et al. in [72].

**Method:** The authors proposed a 3D level set segmentation method with a speed term based on binary flow forces, mean curvature flow and gradient vector flow. The originality of the method was to focus on the topology of the evolving front and use the notion of simple points and update the front deformation only at their locations. Given a set of points defining a 3D surface, a point is simple if its addition or removal from the object does not change the topology of either the object or the background. The topology of an object is defined through its number of connected components, cavities and handles. The algorithm was implemented with a narrow-band update and 3D level set fast marching propagation scheme for computational efficiency. The final object surface, which corresponds to the zero-level of the level set function is extracted with a *connectivity consistent* marching cubes algorithm (CCMC) [95, 96]. This algorithm is a modification of the standard marching cubes algorithm where the resolution of ambiguous cubes depends on predefined digital connectivity rules.

**Experiments:** The authors performed two sets of experiments on brain MRIs.

**1.** The first experiment compared visual quality of segmented data with the topology preserving deformable model to a standard geometric deformable model and a parametric deformable model using the same initialization scheme. Results showed very similar looking surfaces for the three methods but close inspection revealed critical differences:

- the parametric deformable model surface had self-intersection points,
- the number of handles with the simple geometric deformable model was 40 versus 0 for the two other methods (corresponding to the correct manifold).

In this experiment the authors also provided an example from a brain MRI data set where part of the WM seemed to display a handle when viewed in 3D corresponding to an incorrect topology. This type of errors, mostly due to MRI noise, can only be corrected with a topology preserving segmentation method such as the proposed level set framework.

**2.** A second set of experiments employed 21 T1-weighted MRI volumes with voxel size  $(0.9375 \times 0.9375 \times 1.5 \text{ mm})$  from the public database of the Baltimore study on aging [97]. Volumes were preprocessed to remove extracranial tissues, cerebellum and brain stem. The experiments were performed with digital connectivity defined as: 18-connectivity for the object (WM) and 6-connectivity for the background. The experiments focused on the extraction of central cortical surfaces. Prior to segmentation, the volumes were processed with fuzzy connectedness [98] for labeling into memberships to different tissue types. The result of this labeling was used as an initial segmentation of the WM and used to fill the ventricles and the sub cortical structures (including the thalamus,

hypothalamus, caudate nucleus and putamen). The filled WM volume was then binarized via thresholding of the fuzzy values at 0.5. This binary volume was further processed for topology correction with a multiscale graph-based algorithm [99]. The CCMC was then used to extract the WM surface of the volume. At this point, the WM surface was used as the initial level zero of the level set segmentation to extract three cortical surfaces: WM/GM surface, central cortical surface, and pial (CM/CSF) surface.

Segmentation accuracy was assessed through error measurements at 10 landmark points manually selected on major sulci and gyri on six MRI cases. Landmark error was measured as the minimum distance between the landmark points and the segmented surfaces. The overall average error was 0.87 mm (std 0.5 mm) outperforming a previous method from the same group based on a parametric deformable model that produced an average error of 1.22 mm (std 1.01 mm) [100]. Visual inspection of the segmented data did not reveal any self-intersection on the extracted surfaces. The algorithm computational time was about 40 minutes on a SGI O2 workstation for reconstruction of the three surfaces. This performance compares favorably to typical deformable model algorithms with arbitrary initialization as claimed by the authors.

**Limitations:** The main limitation of this algorithm is the involvement of the preprocessing for initialization of the WM that make the process difficult to reproduce.

### 2.4.3.3 Segmentation and Measurement of the Cortex from 3D MR Images using Coupled-Surfaces Propagation

This research was published by Zeng et al. in [33].

**Method:** The authors proposed the segmentation and measurement of the cortical GM thickness from brain MRI data with a level set method using coupled-surfaces propagation. As stated by the authors, coupling surfaces can prevent two problems:

- The inner cortical surface can collapse with the CSF due to higher contrast at the CSF/GM interface than at the WM/GM interface.
- The presence of eye sockets with no CSF signal can drive the outer cortical surface to expend outward from the brain.

Constraining the cortical thickness during the segmentation process prevent the collapse of leakage of the surfaces.

The level set segmentation method with surface coupling is described in the first section of this chapter. The traditional gradient features in the speed term were replaced with tissue interface probability measurements based on statistical priors summarized here. The statistical models are based on the assumption of Gaussian independent distribution functions of voxel intensities in MRI volume data for WM, GM, and CSF. Let us assume the presence of two tissue types A and B in the data with independent Gaussian probabilities  $G(\mu_A, \sigma_A)$  and  $G(\mu_B, \sigma_B)$ . For each voxel s, a set of 26 immediate 3D-neighborhood voxels can be defined. For each neighbor voxel, a normal direction  $\theta$  along the line passing through the center voxel and the neighbor voxel is computed which defines a plane that separates the neighborhood into two regions  $(R_1, R_2)$ . The probability of the center voxel belonging to an interface between the two tissue types (A, B) is then computed as:

$$p_{AB}(\theta) = \prod_{r \in R_1} \frac{1}{\sqrt{2\pi}\sigma_A} \exp\left(\frac{-(I(r) - \mu_A)^2}{\sigma_A^2}\right) \times \prod_{r \in R_2} \frac{1}{\sqrt{2\pi}\sigma_B} \exp\left(\frac{-(I(r) - \mu_B)^2}{\sigma_B^2}\right)$$
(2.42)

where I(r) is the intensity value of the data at neighbor voxel r. The final density probability at voxel s is set to the highest value of  $p_{AB}(\theta)$  over all the 26 directions. An illustration of a feature map based on this tissue interface probability measure is provided on a single brain MRI slice computed with  $(R_1, R_2)$  containing only one voxel. The example illustrates well the better performance of the interface probability feature at extracting locations of tissue transitions for WM, GM, and CSF when compared to standard gradient maps.

**Experiments:** Validation was performed on T1-weighted MRIs. The segmentation process was initialized with several pairs of concentric spheres with a constraint on starting inside the WM for robust behavior.

1. The first experiment used simulated MRI data from the McConnell Brain Imaging Center at the Montreal Neurological Institute [93]. The authors simulated T1-weighted brain MRIs with 3% noise and 1 mm<sup>3</sup> voxel size. The distance range between two surfaces was set to [1.5 mm 5.5 mm] leading to bandwidth ranges for the inner and outer surfaces of [-3 mm 6 mm] and [-6 mm 3 mm], respectively. Segmentation was validated by comparing the binary segmented

volumes to the thresholded membership values of the corresponding voxels (with reference to WM, GM, and CSF) above 0.5 from the database ground truth data. The authors computed TP, FP rates and total volume ratios (VR) to compare segmented volumes and thresholded ground truth volumes. Results with [TP FP VR] were: [92.3% 2.0% 96.3%] for (WM + GM), [92.8% 6.0% 103.2%] for cortical GM and [92.4% 3.3% 98.1%] for WM showing good performance of the algorithm in isolating brain tissue and segmenting the cortex.

2. A second experiment was performed using 20 T1-weighted spoiled gradient MRIs of normal brain subject from the Internet Brain Segmentation Repository (IBSR) of the Center for Morphometric Analysis at the Massachusetts General Hospital [101]. Cases were acquired with two different scanners and all registered to a standard 3D brain coordinate system. Manual segmentation from medical experts was available. An overlap metric was defined by the IBSR to evaluate the performance of automatic segmentation methods, measuring the ratio TP/(1 + FP). Such ratio ranges from 0 for no agreement to 1 for perfect agreement with the manual segmentation considered as the ground truth. MRI data was interpolated from 3 mm thick coronal slices (as provided) to 1 mm thickness, achieving isotropic voxels. The GM overlap metric on the whole brain was 0.657. It outperformed other reported segmentation performances on the same data sets from [102] with overlap ranking from 0.47 to 0.56. Overlap metric specifically computed on the cerebral cortex (excluding brain stem and cerebellum) was further improved to 0.701. The authors pointed out that if applied to the phantom data, the overlap metric was 0.875 that compared to the manual segmentation variability of 0.876 reported in the IBSR database.

**3.** A third experiment was performed for a study of the frontal lobe anatomy on 7 high-resolution MRI data sets acquired with SPGR with isotropic voxel size of 1.2 mm<sup>3</sup>. The patient population of this study included young autistic and control subjects for comparison of frontal lobe volumes. The MRI volumes were preprocessed for inhomogeneity correction. Segmentation was performed with coupled level set functions and the frontal lobe was manually isolated with anatomical landmarks. Segmentation accuracy was compared to expert manual tracing. TP and FP on entire frontal lobe averaged 94.1% and 2.1%. TP and FP for cortical GM on frontal lobe averaged 86.7% and 20.8%. The authors further evaluated the reliability statistics on the volume measurements obtained on the segmented frontal lobe volumes using the method proposed by Schultz and Chakraborty [103]. The agreement between the expert tracing and the level set

segmentation method was very strong for both the entire frontal lobe and only the GM but the level set algorithm systematically underestimated the frontal lobe volume with a mean difference of 4%.

4. In a fourth experiment, regional cortical thickness was quantitatively analyzed on seven high-resolution MRI data sets acquired with SPGR and isotropic voxel size of  $1.2 \text{ mm}^3$ . Cortical measurements were performed in four lobes to compare level set segmentation and expert manual tracing. The authors first compared the mean thickness of each lobe to the results from the study of 63 postmortem males by Pakkenberg and Gundersen [104]. The new segmentation produced similar results with a frontal cortex thicker than the occipital cortex. On overall, postmortem data was 5–14% thinner than the new results while variability of thickness was equal to 1.5 mm for both studies. Statistical tests also showed that frontal and temporal lobes were thicker than parietal and occipital lobes.

This method relies on prior statistics for WM, GM, and CSF for the construction of interface probability maps. The method also requires manual initialization by clicking center points of concentric spheres on slices. The authors demonstrated on one example that the number of spheres or their localization does not have a critical influence on the accuracy or reproducibility of the method. On the other hand, major advantages of the method include the fact that it does not require stripping of non-brain data, its performance was demonstrated on a wide range of MRI image quality and it offers fast computation times compared to existing methods.

### 2.4.3.4 Segmentation of RT3D Ultrasound with Geodesic Deformable Models

This research was published by Corsi et al. in [105].

**Method:** The proposed method uses a geodesic deformable model as defined in Eq. (2.17), with weighting parameters associated with the two terms and a function g() defined as in [8]. To handle the noisy nature of the ultrasound data and the poor definition of the myocardium wall borders in some frames, the authors did not use any inflationary force in the speed term. This type of model requires an initialization close to the final endocardial surface to ensure that the moving front is attracted to local high data gradient locations. The algorithm was implemented with the narrow-band technique for computation efficiency. **Experiments:** The authors performed 3D segmentation of echocardiographic real-time three dimensional (RT3D) ultrasound data for extraction of ventricular volumes. Prior to segmentation, a clinician roughly defined endocardial contours on a limited set of short-axis views. The polygonal surface defined by this manual tracing was then used to initialize the level set segmentation process. Deformation of the moving front required about 40s for a single volume on a PC and initial manual tracing less than 1 min. The authors performed three experiments for assessment of the method accuracy and robustness to the initialization:

**1.** In the first experiment the authors performed multiple segmentation of the same RT3D volume using manual initialization from six different users. Comparing volume measurements to precise manual tracing they report a mean square error of 3.8% and a maximum error of 4.38%. They concluded from these results that the segmentation is rather sensitive to the manual initialization.

2. In the second experiment, the authors segmented invitro phantom data of 18 balloons filled with water and immersed in a tank of water. To take into account the sensitivity of the segmentation technique to the initialization, two operators performed two separate tracings and the segmentation was run with these different initializations. A linear regression coefficient of 0.99 was reported between true volume values and measurements from 70 level set segmentations. The standard error of estimate was equal to 9.35 ml, the average error of measurement was -2.63 ml (std 10.81 ml). Intraobserver variability was estimated for each operator as: 1.66% and -1.36%. Interobserver variability was estimated as 1.63%. This experiment reported a maximum error of measurement of 40 ml for large volumes (above 200 ml). Errors of measurements decreased significantly with balloons' true volumes suggesting that ventricular volumes can be accurately measured with this technique for physiological volume ranges.

**3.** An invivo study was also performed on about 18 RT3D cases (exact number not specified in the paper) using again two different manual tracings from two operators. The ventricular volume range for this study was [151–467 ml] which suggest that it included dilated cardiomyopathy patients with enlarged ventricular cavity. Volume measurements were compared to measurement from manual tracing on MRI. Linear regression was performed with a correlation coefficient of 0.97. The standard error of estimate was 20.13 ml, average error was -15.58 ml (std 20.55 ml). Intraobserver variability was 0.16% and -2.04%, and inter-observer variability was -2.16%. Ejection fraction measurements were

performed on nine RT3D cases. The correlation coefficient was 0.87 when compared to MRI measurement. The main limitation of this study is the absence of testing on normal physiological ventricular volume in the range [40–150 ml] for which the behavior of the level set segmentation can be significantly different as these volumes are much smaller.

### 2.4.3.5 Segmentation of RT3D Ultrasound with Implicit Deformable Models Without Gradients

This research was published by Angelini, Holmes, Laine and Homma in [43].

**Method:** This study focused on the same clinical problem as the previous study for segmentation of echocardiographic RT3D ultrasound data. The proposed method uses the homogeneity-based implicit deformable model proposed by Chan and Vese in [40] as an extension of the Mumford-Shah segmentation functional. Motivations for selection of this method include robustness with arbitrary initialization of the object anywhere in the image, topology adaptation for multiobject segmentation (for potential cosegmentation of ventricles and atria, for example), self-adaptation of the deformation flow to inward and outward flows. Minor modifications of the method were performed to adapt the design to the specificity of the 3D ultrasound data. The homogeneity terms from Eq. (2.28) were weighted by the mean intensity value as:

$$E(C) = \int_{\text{inside}C} \left(\frac{I - c_0}{c_0}\right)^2 d\Omega + \int_{\text{outside}C} \left(\frac{I - c_1}{c_1}\right)^2 d\Omega.$$
(2.43)

A similar approach was followed by Lin *et al.* [44] for segmentation of 3D echocardiographic data where they normalized the homogeneity term by the variance of the data inside and outside the object segmented, after pre-processing with multiscale Gaussian filtering.

Parameters were set to v = 0 (no constant inflation force was used),  $\mu = 1$ ,  $\lambda_1 = 0.25$ ,  $\lambda_2 = 0$  (no homogeneity constraint on the outside of the ventricle to reduce the effect of the noisy myocardium texture),  $\Delta x = \Delta y = \Delta z = 1$ ,  $\Delta t = \min(\Delta x, \Delta y, \Delta z)/|v + \mu + \lambda_1 + \lambda_2|$  (to respect CFL condition with explicit numerical scheme). The system was let to deform over 20 iterations.

**Experiments**: A clinical study was performed on 10 patients with pulmonary hypertension to segment both right and left ventricular volumes. A 2D parametric deformable model and a 3D level set deformable model illustrated in

Figure 2.6 were used for segmentation of the ultrasound data after denoising with a spatio-temporal brushlet expansion [106]. The model was initialized with a cone in which dimensions were defined manually on slices at the base and apex. Manual tracing on ultrasound data was performed by an expert clinician. MRI data were also acquired on the patients and manually traced by a second expert. Absolute errors of measures were computed for RV and LV ejection-fraction. Mean-error values and standard deviation over the 10 cases for the two ventricles were equal to [Mean Std Max Min]: [8.6% 5.7% 17.8% 0.3%] for manual tracing on ultrasound vs. MRI, [4.9% 4.1% 12.21% 0.2%] for 2D parametric deformable model vs. MRI, [4.6% 4.2% 13.9% 0.8%] for 3D level set deformable model vs. MRI. Improvement of correlation measurements with deformable models (with good statistical significance) was reported when compared to MRI as well as better accuracy with a Bland-Altman analysis. The study concluded that errors of EF measurements using deformable models were within



Figure 2.6: Segmentation of right and left ventricular volumes with a 3D implicit deformable model on RT3D ultrasound data. (a) Initialization of the segmentation with a cone shape surface (dashed line) and final position of the contour (continuous line) on the endocardial surface. (b) Illustration of diversity of right and left ventricular shapes and sizes extracted for the clinical study reported in [43].

the range of inter- and intraobserver variability for both ventricles and compared favorably to similar studies performed by other groups using RT3D ultrasound for quantification of cardiac function. Manual tracing measures were significantly less reliable with large standard deviation of errors and low correlation coefficients. Finally, the 3D level set deformable model achieved the highest degree of accuracy, which can be explained by a more accurate segmentation of small and distorted ventricular shapes when integrating the third spatial dimension.

## 2.5 Conclusion

Level set methods for segmentation and registration of medical images have been the focus of intense research for the past decade producing very promising results. Major advantages of the method include its robustness to noisy conditions, its aptitude in extracting curved objects with complex topology and its clean numerical framework of multidimensional implementation. Despite their success, these methods still need to be refined to address two limitations:

- 1. Computation time needs to be further reduced, for viability of the method in clinical application where interactivity (and therefore close to real-time computation) is critical. This optimization will have to handle the constant increase in data size observed in medical imaging applications with improvements of spatial resolution, temporal resolution and now the introduction of combo scanners such as PET/CT machines.
- 2. Robustness to variation in image quality and organ anatomy needs to be studied. Unfortunately, the methods described in this chapter were only rarely validated in clinical studies. On the other hand, it is well known that these methods require tuning of their parameters to adapt to the nature of the image data to segment. In that optic, it is therefore critical to evaluate robustness of the performance on a set of data that covers the range of quality encountered in clinical practice for a particular examination. For methods based on shape models, it is also critical to test the method on a variety of abnormal (e.g., disease) cases that differ from the average anatomy that they typically represent. Such validation for medical application should always clearly specify the context of the problem at

hand in terms of anatomy of interest (e.g., endocardial surface of myocardium muscle), imaging modality (e.g., three-dimensional real-time ultrasound) and clinical application targeted (e.g., quantification of volume). Only in this context can a segmentation method be really tuned, tested, and validated for clinical application [107].

# Questions

- 1. Define the principle idea of the level set framework for segmentation of an image given an initial contour C and a speed function V > 0. What are the advantages of using an implicit formulation of the problem?
- 2. What are the limitations associated with gradient-based speed terms? Why is it especially problematic with medical images?
- 3. What is a regularization term? What is its main functionality? Give some examples.
- 4. What is the entropy principle for implementation of a level set deformable model with finite difference? In what case does it apply?
- 5. Explain the concept of speed extension for image-based speed terms? Why is it necessary? Propose a simple algorithm to implement it.
- 6. Is the standard level set framework preserving the distance function? Why is this an important concept for segmentation applications?
- 7. Why is there a need for reinitialization of the distance function?
- 8. Outline in a flowchart the structure of an iterative level set segmentation algorithm using a gradient-based speed term. Use a convergence criteria (without detailing it) to stop the iterations.
- 9. Design a level set segmentation algorithm for extraction of the endocardial and epicardial surfaces of the left ventricle from an MRI volume? What are the properties of the data that can be used to define the speed function? Is there a way to perform simultaneous segmentation of the two surfaces?

# Bibliography

- McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A survey, Medical Image Analysis, Vol. 1, pp. 91–108, 1996.
- [2] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, International Journal of Computer Vision, Vol. 1, pp. 321–331, 1987.
- [3] Metaxas, D. N., Physics-based deformable models. Applications to computer vision, graphics and medical imaging, 1997.
- [4] Xu, C., Pham, D., and Prince, J., Image segmentation using deformable models., In: Handbook of Medical Imaging, Vol. 2: SPIE, 2000, pp. 129– 174.
- [5] McInerney, T. and Terzopoulos, D., Topologically adaptable snakes, In: Proceedings of 5th International Conference on Computer Vision, pp. 840–845, 1995.
- [6] Durikovic, R., Kaneda, K., and Yamashita, H., Dynamic contour: a texture approach and contour operations, The Visual Computer, Vol. 11, pp. 277–289, 1995.
- [7] Osher, S. and Sethian, J. A., Fronts propagating with curvaturedependent speed: Algorithms based on Hamilton-Jacobi formulations, Journal of Computational Physics, Vol. 79, pp. 12–49, 1988.
- [8] Malladi, R., Sethian, J. A., and Vemuri, B. C., Shape modeling with front propagation: A level set approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, pp. 158–175, 1995.
- [9] Suri, J. S., Liu, K., Singh, S., Laxminarayan, S. N., Zeng, X., and Reden, L., Shape recovery algorithms using level sets in 2D/3D medical imagery: a state-of-the-art review, IEEE Transactions on Information Technology in Biomedicine, Vol. 6, pp. 8 -28, 2002.
- [10] Suri, J. S. and Liu, K., Level set regularizers for shape recovery in medical images, In: Proceedings of IEEE Symposium on Computer-Based Medical Systems, pp. 369–374, Bethesda, MD, USA, 2001.

- [11] Montagnat, J., Delingette, H., and Ayache, N., A review of deformable surfaces: topology, geometry and deformation, Image and Vision Computing, Vol. 19, pp. 1023–1040, 2001.
- [12] Caselles, V., Catte, F., Coll, T., and Dibos, F., A geometric model for active contours, Numerische Mathematik, Vol. 66, pp. 1–31, 1993.
- [13] Sethian, J., Numerical algorithms for propagating interfaces: Hamilton-Jacobi equation and conservation laws, Journal of Differential Geometry, Vol. 31, pp. 131–161, 1990.
- [14] Sethian, J. A., Level Set Methods and Fast Marching Methods: Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science, Cambridge University Press, 1999.
- [15] Kichenassamy, S., Kumar, A., Olver, P., Tannenbaum, A., and Yezzi, A., Gradient flows and geometric active contour models, In: Proceedings of IEEE International Conference in Computer Vision, pp. 810–815, Cambridge, MA, USA, 1995.
- [16] Caselles, V., Kimmel, R., and Sapiro, G., Geodesic active contours, In: Proceedings of International Conference on Computer Vision, pp. 694– 699, Cambridge, MA, USA, 1995.
- [17] Yezzi-Jr, A., Kichenassamy, S., Kumar, A., Olver, P., and Tannenbaum, A., A geometric snake model for segmentation of medical imagery, IEEE Transactions on Medical Imaging, Vol. 16, pp. 199–209, 1997.
- [18] Malladi, R., Kimmel, R., Adalsteinsson, D., Sapiro, G., Caselles, V., and Sethian, J. A., A geometric approach to segmentation and analysis of 3D medical images, In: Proceedings of Workshop on Mathematical Methods in Biomedical Image Analysis, pp. 244–252, 1996.
- [19] Niessen, W. J., Romeny, B. M. t. H., and Viergever, M. A., Geodesic deformable models for medical image analysis, IEEE Transactions on Medical Imaging, Vol. 17, pp. 634 - 641, 1998.
- [20] Siddiqi, K., Lauziere, Y. B., Tannenbaum, A., and Zucker, S. W., Area and Length Minimizing Flows for Shape Segmentation, IEEE Transactions on Image Processing, Vol. 7, pp. 433–443, 1998.

- [21] Deschamps, T., Curve and Shape Extraction with Minimal Path and Level-Sets techniques. Applications to 3D Medical Imaging, In: Mathématiques Appliquées et Traitements d'Images. Paris, France: Université de Paris Dauphine, 2001.
- [22] Yezzi-Jr., A., Tsai, A., and Willsky, A., A statistical approach to snakes for bimodal and trimodal imagery, In: Proceedings of IEEE International Conference on Computer Vision, pp. 898–903, Kerkyra, Greece, 1999.
- [23] Paragios, N. and Deriche, R., Geodesic active contours and level sets for the detection and tracking of moving objects, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, pp. 266–280, 2000.
- [24] Sarti, A. and Malladi, R., A Geometric Level Set Model for Ultrasounds Analysis, University of California, Berkeley, 1999.
- [25] Jin, Y., Laine, A., and Imielinska, C., An adaptive speed term based on homogeneity for level-set segmentation, In: Proceedings of SPIE Medical Imaging Conference, pp. 383–390, San Diego, CA, USA, 2002.
- [26] Saha, P. K., Udupa, J. K., and Odhner, D., Scale-based fuzzy connected image segmentation: theory, algorithms, and validation, Computer Vision and Image Understanding, Vol. 77, pp. 145–174, 2000.
- [27] Sarti, A., Malladi, R., and Sethian, J. A., Subjective surfaces: a geometric model for boundary completion, International Journal of Computer Vision, Vol. 46, pp. 201–221, 2002.
- [28] Suri, J. S., Leaking prevention in fast level sets using fuzzy models: an application in MR brain, In: Proceedings of Information Technology Applications in Biomedicine, pp. 220–225, Arlington, VA, USA, 2000.
- [29] Suri, J. S., Leaking prevention in fast level sets using fuzzy models: an application in MR brain, In: Proceedings of IEEE EMBS International Conference on Information Technology Applications in Biomedicine, pp. 220–225, Arlington, VA, USA, 2000.
- [30] Baillard, C., Barillot, C., and Bouthemy, P., Robust Adaptive Segmentation of 3D Medical Images with Level Sets, INRIA, Rennes, France, Research Report Nov. 2000.

- [31] Leventon, M. E., Grimson, W. E. L., and Faugeras, O., Statistical shape influence in geodesic active contours, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 316–325, 2000.
- [32] Leventon, M., Grimson, E., Faugeras, O., Wells, S., and Kikinis, R., Knowledge-based segmentation of medical images, In: Geometric Level Set Methods in Imaging, Vision and Graphics, Osher, N. P. S., ed., New York: Springer Verlag, 2003.
- [33] Zeng, X., Staib, L. H., Schultz, R. T., and Duncan, J. S., Segmentation and measurement of the cortex from 3D MR images using coupledsurfaces propagation, IEEE Transactions on Medical Imaging, Vol. 18, pp. 927–937, 1999.
- [34] Gomes, J. and Faugeras, O., Segmentation of the inner and outer surfaces of the human cortex: an approach based on partial differential equations, In: Proceedings of 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1764–1774, Chicago, IL, USA, 2000.
- [35] Barles, G., Soner, H. M., and Souganidis, P. E., Front propagation and phase field theory, SIAM Journal of Control and Optimization, Vol. 31, pp. 439–469, 1993.
- [36] Mumford, D. and Shah, J., Boundary detection by minimizing functional, In: Proceedings of International Conference on Computer Vision and Pattern Recognition, pp. 22–26, San Francisco, CA, USA, 1985.
- [37] Chan, T. F. and Vese, L. A., Active contour and segmentation models using geometric PDE's for medical imaging, In Geometric Methods in Bio-Medical Image Processing, Mathematics and Visualization, R. Malladi, ed., Springer, 2002.
- [38] Chan, T. F. and Vese, L. A., A level set algorithm for minimizing the Mumford-Shah functional in image processing, In: Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision, pp. 161–168, Vancouver, BC, Canada, 2001.
- [39] Chan, T. F. and Vese, L. A., An efficient variational multiphase motion for the Mumford-Shah segmentation model, In: Proceedings of 34th
Asilomar Conference on Signals, Systems and Computers, pp. 490–494, Pacific Grove, CA, USA, 2000.

- [40] Chan, T. F. and Vese, L. A., Active contours without edges, IEEE Transactions on Image Processing, Vol. 10, pp. 266–277, 2001.
- [41] Vese, L. A. and Chan, T. F., A multiphase level set framework for image segmentation using the Mumford and Shah model, University of California, Los Angeles, CA, USA, Computational and Applied Mathematics Report 01-25, 2001.
- [42] Tsai, A., Yezzi-Jr, A., and Willsky, A. S., Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification, IEEE Transactions on Image Processing, Vol. 10, pp. 1169–1186, 2001.
- [43] Angelini, E. D., Holmes, J., Laine, A. F., and Homma, S., Segmentation of real-time 3D cardiac ultrasound with implicit deformable models without gradients, In: Proceedings of International Symposium on Image and Signal Processing and Analysis, pp. 711–716, Rome, Italy, 2003.
- [44] Lin, N., Yu, W., and Duncan, J. S., Combinative multiscale level set framework for echocardiographic image segmentation, Medical Image Analysis, Vol. 7, pp. 529–537, 2003.
- [45] Chan, T. F., Sandberg, B. Y., and Vese, L. A., Active contours without edges for vector-valued images, Journal of Visual Communication and Image Representation, Vol. 11, pp. 130–141, 2000.
- [46] Maintz, J. B. and Viergever, M. A., A survey of medical image registration, Medical Image Analysis, Vol. 2, pp. 1–36, 1998.
- [47] Ginneken, B. v., Frangi, A. F., Staal, J. J., Romeny, B. M. t. H., and Viergever, M. A., Active shape model segmentation with optimal features, IEEE Transactions on Medical Imaging, Vol. 21, pp. 924–933, 2002.
- [48] Frangi, A. F., Rueckert, D., Schnabel, J. A., and Niessen, W. J., Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling, IEEE Transactions on Medical Imaging, Vol. 21, pp. 1151–1166, 2002.

- [49] Taylor, C. J., Cootes, T. F., Hill, A., and Haslam, J., Medical image segmentation using active shape models, In: Medical Imaging, M. H. K. L. Beolchi, ed., IOS Press, Amsterdam, 1994.
- [50] Angelini, E. D. and Ciaccio, E. J., Optimized region finding and edge detection of knee cartilage surfaces from magnetic resonance images, Annals of Biomedical Engineering, Vol. 31, pp. 336–345, 2003.
- [51] Medical Image Registration: CRC Press, 2001.
- [52] Vemuri, B. C., Ye, J., Chen, Y., and Leonard, C. M., Image registration via level-set motion: Applications to atlas-based segmentation, Medical Image Analysis, Vol. 7, pp. 1–20, 2003.
- [53] Yahia, H. M., Huot, E. G., Herlin, I. L., and Cohen, I., Geodesic distance evolution of surfaces: a new method for matching surfaces, In: Proceedings of IEEE CVPR, pp. 663–668, Hilton Head Island, SC, USA, 2000.
- [54] Cohen, I. and Herlin, I., Curve matching using geodesic paths, In: Proceedings of IEEE CVPR, pp. 741–746, Santa Barbara, CA, USA, 1998.
- [55] Lavallee, S. and Szeliski, R., Recovering the position and orientation of free-form objects from image contours using 3D distance maps, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, pp. 378–390, 1995.
- [56] Rousson, M. and Paragios, N., Shape priors for level set representations, In: Proceedings of European Conference on Computer Vision, pp. 78–93, Copenhagen, Denmark, 2002.
- [57] Paragios, N., Rousson, M., and Ramesh, V., Matching distance functions: A shape-to-area variational approach for global-to-local registration, In: Proceedings of European Conference in Computer Vision, Copenhagen, Denmark, 2002.
- [58] Paragios, N. and Rousson, M., Shape analysis towards model-based segmentation, In: Geometric Level Set Methods in Imaging, Vision and Graphics, N. P. S. Osher, ed., Springer Verlag, New York, 2003.
- [59] Yezzi, A., Zollei, L., and Kapur, T., A variational framework for joint segmentation and registration, In: Proceedings of IEEE Workshop on

Mathematical Methods in Biomedical Image Analysis, pp. 44–51, Kauai, HI, USA, 2001.

- [60] Vemuri, B. and Chen, Y., Joint image registration and segmentation, In: Geometric Level Set Methods in Imaging, Vision and Graphics, N. P. S. Osher, eds., Springer Verlag, New York, pp. 251–269, 2003.
- [61] Tsai, A., Yezzi-Jr., A., Wells, W., III, Tempany, C., Tucker, D., Fan, A., Grimson, W. E., and Willsky, A., Model-based curve evolution technique for image segmentation, In: Proceedings of Conference on Computer Vision and Pattern Recognition, pp. 463–468, 2001.
- [62] Cremers, D., Schnorr, C., and Weickert, J., Diffusion-snakes: combining statistical shape knowledge and image information in a variational framework, In: Proceedings of Workshop on Variational and Level Set Methods in Computer Vision, pp. 137–144, Vancouver, BC, Canada, 2001.
- [63] Paragios, N., A level set approach for shape-driven segmentation and tracking of the left ventricle, IEEE Transactions on Medical Imaging, Vol. 22, pp. 773–776, 2003.
- [64] Wang, Y. and Staib, L., Boundary finding with correspondence using statistical shape models, In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Santa Barbara, CA, USA, 1998.
- [65] Staib, L. H. and Duncan, J. S., Model-based deformable surface finding for medical images, IEEE Transactions on Medical Imaging, Vol. 15, pp. 720–731, 1996.
- [66] Chen, Y., Thiruvenkadam, S., Tagare, H. D., Huang, F., Wilson, D., and Geiser, E. A., On the incorporation of shape priors into geometric active contours, In: IEEE Workshop on Variational and Level Set Methods in Computer Vision, Vancouver, BC, Canada, 2001.
- [67] Chen, Y., Huang, F., Tagare, H. D., Murali, R., Wilson, D., and Geiser, E. A., Using prior shape and intensity profile in medical image segmentation, In: Proceedings of IEEE ICCV, pp. 1117–1124, 2003.
- [68] Chen, Y., Thiruvenkadam, S., Tagare, H. D., Huang, F., Wilson, D., and Geiser, E. A., On the incorporation of shape priors into geometric

active contours, In: Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision, Vancouver, BC, Canada, 2001.

- [69] Paragios, N., Rousson, M., and Ramesh, V., Knowledge-based registration & segmentation of the left ventricle: a level set approach, In: Proceedings of IEEE Workshop on Applications of Computer Vision, 2002.
- [70] Paragios, N., A variational approach for the segmentation of the left ventricle in MR cardiac images, In: Proceedings of IEEE Workshop on Variational and Level Set Methods in Computer Vision, Vancouver, BC, Canada, 2001.
- [71] Xu, C. and Prince, J. L., Snakes, shapes and gradient vector Flow, IEEE Transactions on Image Processing, pp. 359–369, 1998.
- [72] Han, X., Xu, C., and Prince, J., Topology preserving geometric deformable models for brain reconstruction, In: Geometric Level Set Methods in Imaging, Vision and Graphics, N. P. S., Osher, ed., New York: Springer Verlag, 2003.
- [73] Kwan, R. K.-S., Evans, A. C., and Pike, G. B., An extensible MRI simulator for post-processing evaluation, In: Proceedings of Visualization in Biomedical Computing, Hamburg, Germany, 1996.
- [74] Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J., and Evans, A. C., Design and construction of a realistic digital brain phantom, IEEE Transactions on Medical Imaging, Vol. 17, pp. 463–468, 1998.
- [75] Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., Evans, A. C. and v., no. 4, part 2/4, S425, May 1997., BrainWeb: online interface to a 3d MRI simulated brain database, in the Proceedings of III International Conference on Functional Mapping of the Human Brain, pp. S425, Copenhagen, Denmark, 1997.
- [76] Kwan, R. K.-S., Evans, A. C., and Pike, G. B., MRI simulation-based evaluation of image-processing and classification methods, IEEE Transactions on Medical Imaging, Vol. 18, pp. 1085–1097, 1999.

- [77] Mykkanen, J., Tohka, J., and Ruotsalainen, U., Automated delineation of brain stuctures with snakes in PET, In Physiological Imaging of the Brain with PET, A. Press, ed., 2001, pp. 39–43.
- [78] Ofili, E. O. and Nanda, N. C., Three-dimensional and four-dimensional echocardiography, Ultrasound Medical Biology, Vol. 20, 1994.
- [79] Fenster, A. and Downey, D. B., Three-Dimensional Ultrasound Imaging, In: Handbook of Medical Imaging. Volume1. Physics and Psychophysics, Vol. 1, H. L. K. Jacob Beutel, Richard L. Metter, ed., Bellingham, WA, USA.: SPIE- The International Society of Optical Engineering, 2000, pp. 463–510.
- [80] Rankin, R. N., Fenster, A., Downey, D. B., Munk, P. L., Levin, M. F., and Vellet, A. D., Three-dimensional sonographic reconstruction: technique and diagnostic applications, American Journal of Radiology, Vol. 161, pp. 695–702, 1993.
- [81] Belohlavek, M., Foley, D. A., Gerber, T. C., Kinter, T. M., Greenleaf, J. F., and Seward, J. B., Three- and four-dimensional cardiovascular ultrasound imaging: a new era for echocardiography, In: Mayo Clinic Proceedings, Vol. 68, pp. 221–240, 1993.
- [82] Stetten, G., Ota, T., Ohazama, C., Fleishman, C., Castelucci, J., Oxaal, J., Ryan, T., Kisslo, J., and Von Ramm, O. T., Real-time 3D ultrasound: a new look at the heart, Journal of Cardiovascular Diagnosis Procedures, Vol. 15, pp. 73–84, 1998.
- [83] Smith, S. W., Pavy, J. H. G., Miglin, J. A., and Von Ramm, O. T., Improved real-time volumetric ultrasonic imaging, Ultrasonic Imaging, Vol. 14, pp. 186–211, 1992.
- [84] Von Ramm, O. T., Pavy, J. H. G., Smith, S. W., and Kisslo, J., Realtime, three-dimensional echocardiography: the first human images, Circulation, Vol. 84, pp. 685, 1991.
- [85] Ramm, O. T. Von and Smith, S. W., Real time volumetric ultrasound imaging system, Journal of Digital Imaging, Vol. 3, pp. 261–266, 1990.

- [86] Ota, T., Fleishman, C. E., Strub, M., Stetten, G., Ohazama, C. J., Ramm, O. T. Von, and Kisslo, J., Real-time, three-dimensional echocardiography: Feasibility of dynamic right ventricular volume measurement with saline contrast., American Heart Journal, Vol. 137, pp. 958–66, 1999.
- [87] Kisslo, J., Firek, B., Ota, T., Kang, D. H., Fleishman, C. E., Stetten, G., Li, J., Ohazama, C. J., Adams, D., Landolfo, C., Ryan, T., and Ramm, O. T. Von, Real-time volumetric echocardiography: the technology and the possibilities, Echocardiography, Vol. 17, pp. 773–779, 2000.
- [88] Morsy, A. A., Stetten, G. D., and Ramm, O. T. Von, Detection and quantification of true 3D motion components of the myocardium using 3D speckle tracking in volumetric ultrasound scans: simulations and initial experimental results, In: Proceedings of Medical Imaging 1997: Physiology and Function from Multidimensional Images, pp. 346–353, 1997.
- [89] Ota, T., Fleishman, C. E., Ohazama, C. J., Stetten, G., Lewis, C. W., Glower, D. D., Li, J., Ryan, T., Kisslo, J., and Ramm, O. T. Von, Measurement of left ventricular volume by real-time, three-dimensional echocardiography in dogs, Circulation, Vol. 94, pp. 379, 1990.
- [90] Shattuck, D. P., Weinshenker, M. D., Smith, S. W., and Ramm, O. T. Von, Exploroscan: A parallel processing technique for high speed ultrasound imaging with linear phased arrays, The Journal of the Acoustical Society of America, Vol. 75, pp. 1273–1282, 1984.
- [91] Shiota, T., Jones, M., Chikada, M., Fleishman, C., Castellucci, J., Cotter, B., DeMaria, A., Ramm, O. T. Von, Kisslo, J., Ryan, T., and Sahn, D., Real-time three dimensional echocardiography for determining right ventricular stroke volume in an animal model of chronic right ventricular volume overload, Circulation, Vol. 97, pp. 1897–1900, 1998.
- [92] Masson, P. and Pieczynski, W., SEM algorithm and unsupervised statistical segmentation of satellite images, IEEE Transactions on Geoscience and Remote Sensing, Vol. 31, pp. 618–633, 1993.
- [93] M. B. I. Centre: Montréal Neurological Institute, McGill University, Montréal, Canada.

- [94] Handbook of Medical Informatics. Heidelberg: Springer-Verlag, 2002.
- [95] Han, X., Xu, C., Rettmann, M., and Prince, J., Automatic segmentation editing for cortical surface reconstruction, In: Proceedings of SPIE Medical Imaging, pp. 194–203, San Diego, CA, USA, 2001.
- [96] Lorensen, W. E. and Cline, H. E., Marching cubes: A high resolution 3D surface construction algorithm, Computer Graphics, Vol. 21, pp. 163–169, 1987.
- [97] Resnick, S., Goldszal, A., Davatzikos, C., Golski, S., Kraut, M., Metter, E., Bryan, R., and Zonderman, A., One-year changes in MRI brain volumes in older adults, Cerebral Cortex, Vol. 10, pp. 464–472, 2000.
- [98] Pham, D. L. and Prince, J. L., Adaptive fuzzy segmentation of magnetic resonance images, IEEE Transactions on Medical Imaging, Vol. 18, pp. 737–752, 1999.
- [99] Han, X., Xu, C., Braga-Neto, U., and Prince, J., Topology correction in brain cortex segmentation using a multiscale graph-based algorithm, IEEE Transactions on Medical Imaging, Vol. 21, pp. 109–121, 2002.
- [100] Xu, C., Pham, D. L., Rettmann, M. E., Yu, D. N., and Prince, J. L., Reconstruction of the human cerebral cortex from magnetic resonance images, IEEE Transactions on Medical Imaging, Vol. 18, pp. 467–480, 1999.
- [101] M. G. H. Center for Morphometric Analysis, Brain Segmentation Repository.
- [102] Rajapakse, J. C. and Fruggel, F., Segmentation of MR images with intensity inhomogeneities, Image and Vision Computing, Vol. 16, pp. 165–180, 1998.
- [103] Schultz, R. T. and Chakraborty, A., Magnetic resonance image analysis, In: Handbook of Human Brain Function: Neuroimaging, E. Bigler, ed., Plenum, New York, 1996.
- [104] Pakkernenberg, B. and Gundersen, H. J., Neocortical neuron number in humans: effect of sex and age, Journal of Comparative Anatomy, Vol. 384, 1997.

- [105] Corsi, C., Saracino, G., Sarti, A., and Lamberti, C., Left ventricular volume estimation for real-time three-dimensional echocardiography, IEEE Transactions on Medical Imaging, Vol. 21, pp. 1202–1208, 2002.
- [106] Angelini, E., Laine, A., Takuma, S., Holmes, J., and Homma, S., LV volume quantification via spatio-temporal analysis of real-time 3D echocardiography, IEEE Transactions on Medical Imaging, Vol. 20, pp. 457–469, 2001.
- [107] Udupa, J., LeBlanc, V., Schmidt, H., Imielinska, C., Saha, P., Grevera, G., Zhuge, Y., Molholt, P., Jin, Y., and Currie, L., A methodology for evaluating image segmentation algorithm, In: Proceedings of SPIE Conference on Medical Imaging, pp. 266–277, San Diego CA, USA, 2002.

# **Chapter 3**

# Three-Dimensional Rigid and Non-Rigid Image Registration for the Pelvis and Prostate

Baowei Fei,<sup>1</sup> Jasjit Suri,<sup>1</sup> and David L Wilson<sup>1,2</sup>

# 3.1 Introduction

Several important applications require registration of images of the pelvis and prostate [1, 2]. First, comparison of registered MR images acquired before and immediately after therapies can be used to determine whether a tumor is adequately treated. This is particularly helpful in instances where the edematous response to treatment can be confused with a highly perfused tumor. Second, registration of serial examinations can be used to follow regression/progression of tumor. Third, registration of functional, biochemical images such as single photon emission computed tomography (SPECT), positron emission tomography (PET), and MR spectroscopy with anatomical MR images is useful for detecting and localizing cancer. Fourth, incorporating the functional, biochemical images into the interventional magnetic resonance imaging (iMRI) paradigm will aid image-guided treatments. Fifth, on a low-field magnet system during iMRI treatments where fast imaging is important, it might be highly desirable to register high quality MR image from a conventional MR scanner to the live-time iMRI images [3, 5].

There are challenges for registration in the pelvis and prostate that might reduce the effectiveness of automatic voxel-based registration. First, the

<sup>&</sup>lt;sup>1</sup>Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, 44106, USA

 $<sup>^2 \</sup>rm Department$  of Radiology, University Hospitals of Cleveland & Case Western Reserve University, Cleveland, OH 44106, USA

abdomen has irregular boundaries, unlike the head to which registration has been most often applied. Second, the normal prostate is a small organ that when healthy measures only about 38.0 mm in its widest dimension transversely across the base [6]. Third, different patient positions such as legs flat and raised significantly change the legs in lower portions of image volumes as well as cause movement and deformation of internal organs in the pelvis. Fourth, the prostate might move relative to the pelvic bones due to changes in bladder and rectal filling [7, 8]. The alignment of the pelvic bones, a most prominent anatomical feature in MR grayscale images, does not necessarily mean that the prostate registration because no external markers are available.

Many reports describe methods and evaluations for registration in the brain [9]. A few describe results for the pelvis or prostate. For example, manual registration has been used where an operator cues on segmented vascular structures [10] or other anatomical landmarks [11, 13]. Others have used automated 3D schemes that match contours of bones and sometimes other structures that are extracted using manual or interactive segmentation [14, 16]. Manual segmentation has also been used to create surfaces for automatic registration [17, 18]. All of these methods require either segmentation or visual identification of structures. Voxel based methods, particularly those based upon mutual information, are robust, require no segmentation that can be prone to error, are suitable for multimodality registration, are highly accurate for brain registration [19], and are suitable for abdominal registration [20]. For registration of brain and other organs, registration accuracy has been assessed using fiducial markers [21, 22] and anatomical landmarks [23, 25].

The next section will describe a three-dimensional mutual information rigid body registration algorithm with special features for MRI volumes of the pelvis and prostate. Section 3.3 describes a three-dimensional nonrigid registration algorithm that is based upon independent optimization of many interactively placed control points using mutual information and a thin plate spline transformation. Detailed implementation, comparisons with rigid body method, and discussions are reported at the end of the section.

# 3.2 Three-Dimensional Rigid Body Registration Algorithm with Special Features

## **3.2.1 Similarity Measurements**

Two similarity measures, mutual information and correlation coefficient (CC), are used in the registration. Suppose one volume R is the *reference*, and the other F is *floating*. Their mutual information MI(R,F) is given below [19]:

$$MI(R, F) = \sum_{r, f} p_{RF}(r, f) \log \frac{p_{RF}(r, f)}{p_R(r) \cdot p_F(f)}$$

The joint probability  $p_{RF}(r, f)$  and the marginal probabilities  $p_R(r)$  of the reference image and  $p_F(f)$  of the floating image, can be estimated from the normalized joint and marginal intensity histogram, respectively. The correlation coefficient CC(R, F) is given below [26]:

$$CC(R,F) = \frac{\sum (R(r) - \overline{R}(r))(F(f) - \overline{F}(f))}{\sqrt{\sum (R(r) - \overline{R}(r))^2 \sum (F(f) - \overline{F}(f))^2}}$$

Here  $\overline{R}(r)$ ,  $\overline{F}(f)$  denote the average intensities of the reference and floating volumes and the summation includes all voxels within the overlap of both volumes.

In Fig. 3.1, we compare the two similarity measures at different resolutions. Plotted are MI and CC values as a function of translation along the transverse axis where the origin is the optimal transformation. For images at a resolution of 1/4 voxels along a linear dimension, the CC curves are much smoother than MI, which is noisy and contains many local maximums as shown in Fig. 3.1a. In addition, there is a false global maximum in Fig. 3.1a at 18 voxels. At full resolution, Fig. 3.1c shows that MI is much more peaked than CC, but there is high frequency noise in the MI curves far from the optimum that give rise to local maximums that must be avoided. From these figures, we infer that CC is better at low resolution and that MI is better at full resolution when one is close



Figure 3.1: MI and CC similarity functions are plotted to show their relative advantages for registration at different resolutions. Two high-resolution MRI volumes were registered to obtain the optimal parameters. We then computed similarity values as a function of translation along the transverse axis. MI is plotted in (a) and (c); CC is plotted in (b) and (d). Graphs on the top, (a) and (b), are at a resolution of 1/4 voxels along a linear dimension, giving a distance between voxel centers of  $\approx$ 5.5 mm. MI gives a noisy plot having many local maximums, and a false global maximum occurs at 18 voxels. Graphs on the bottom are obtained at full resolution. MI has a much sharper peak than CC, which is relatively flat. The voxel size is 1.4 mm. Images are from volunteer V2 in the diagnostic and reference conditions.

to the optimum value. As described in Section 3.2.2, our registration algorithm makes use of these features.

## 3.2.2 Registration Algorithm with Special Features

The algorithm shown in Fig. 3.2 include special features to improve robustness for registration of MR prostate images. We use a multiresolution approach and

Set initial CC\_duresholds

DO registration FROM lowest resolution TO highest resolution BEGIN Resample both volumes to <sup>14</sup>, <sup>14</sup> or full number of voxels along linear dimension Initialize registration\_results and number\_restarts to zero for restarting registrations REPEAT

Optimize similarity (CC" or MI") between the reference and floating volumes

1. Transform the floating volume and interpolate to get a new reformatted volume

2. Calculate similarity between the reference volume and the new reformatted volume

3. Repeat 1 and 2 until meeting function tolerance or maximum iteration number

Calculate CC between the reference volume and the optimal reformatted volume

Record CC, MI<sup>RE</sup> values and transformation parameters to registration\_results

Perturb the initial transformation parameters:

initial\_transformation\_parameters = initial\_transformation\_parameters - random • factor Trace the number of restarts: number\_restarts - number\_restarts + 1

UNTIL (CC > CC\_threshold) OR (number\_restarts > maximum\_restarts)

Select the final transformation parameters based on CC\* or MT\* among registration results

Scale the parameters and assign to initial\_transformation\_parameters of next higher resolution END

Figure 3.2: Registration algorithm. Capital bold words are computer language. The outer loop from DO to END gives the multiresolution approach. The interloop from REPEAT to UNTIL is for restarting registration. Registration\_results and number\_restarts are used to store temporary values in the program. See text for details. \* CC is used at lower resolutions, 1/4 or 1/2 number of voxels. \*\* MI is used only at high resolution, full number of voxels.

perform registration from low to high resolution. At low resolution, we resample both images at 1/4 or 1/2 number of voxels along each linear dimension, respectively. Iterative optimization of the similarity is used to vary the six rigid body transformation parameters (three translations and three angles). We use the correlation coefficient at the two lower resolutions because of two superiorities: (a) it gives fewer local maximums and (b) it can be calculated faster than MI. We use MI at full resolution because the peaked similarity function gives a more precise solution than CC.

We create a method to avoid local minima by restarting the registration with randomly perturbed parameters obtained from a uniform distribution about the very first initial guess at each resolution. The distribution is centered on the initial guess because we want to use the best solution from the lower resolution. The algorithm restarts until the absolute correlation coefficient between the reference and registered volumes is above a threshold or the maximum number of restarts is reached. The perturbation range is  $\pm 5^{\circ}$  and  $\pm 5$  voxels corresponding to  $\pm 27.3$  mm,  $\pm 13.7$  mm or  $\pm 6.8$  mm for resolutions 1/4, 1/2, or full voxels, respectively. Absolute CC is used for the restart test rather than MI because of three reasons: (a) CC has a well-defined range between 0 and 1, (b) CC provides an independent check of the MI result, and (c) CC has fewer problems with local and incorrect global maximums for registrations at low resolution far from the optimum value.

We record all important results following an optimization cycle including the CC and/or MI values, the number of restarts, and the transformation parameters. At the end of processing at a lower resolution, we always select the transformation parameters having the maximum CC value. We then scale the translation parameters appropriately and assign the new parameters to be initial values at the next higher resolution. At the highest resolution, we select the final transformation parameters to be those with the maximum MI value.

There are several implementation details. We use rigid body transformation (three translations and three angles) and trilinear interpolation. For optimization, we use the downhill simplex method of Nelder and Mead [27]. Optimization of alignment ends either when the maximum number of MI calculations is reached (typically 500) or the fractional change in MI is smaller than a tolerance (typically 0.001). The very first initial guess at the lowest resolution is all zeros for the three displacements and three angles. The CC thresholds can be 0.65, 0.70, and 0.75, and the maximum numbers of restarts be 20, 10, and 5, from low to high resolutions, respectively.

There are several preprocessing details. Isotropic voxels are created using 3D linear interpolation or higher order interpolation methods [20]. From the top and bottom of the volume, we optionally crop transverse slices that are over 35 mm away from the prostate rim. Cropping is done to remove slices having reduced brightness due to sensitivity fall off from the receiver coils, artifacts from a small field of view, displacement of the legs in the treatment position, and/or bladder deformation.



Figure 3.3: MR prostate image with labeled features used to analyze registration error. This transverse image is from the reference volume of prostate cancer patient P3. The prostate boundary was manually segmented near the image center. The four vertical dash lines from left to right indicate the rim of the right acetabular socket, the right and left rims of the prostate, and the rim of the left acetabular socket, respectively. The five crosses from bottom to top indicate the coccyx, the prostate posterior rim, the 2D centroid automatically calculated from the segmented prostate area, the anterior rim, and the pubic symphysis. The image also shows the bottom of the bladder, the rectum, the pubic symphysis, and hip joints.

## 3.2.3 Evaluation of Registration

## 3.2.3.1 Registration Accuracy Based on Bony Landmarks

We evaluate the registration of the pelvis by measuring the displacement of bony landmarks following registration. There are six easily found bony landmarks consisting of two great sciatic notches, two lesser sciatic notches, the pubic symphysis, and the coccyx, some of which are illustrated in Fig. 3.3. Previously, sciatic notches and the pubic symphysis were used to register CT and MRI images for prostate conformal radiation [28]. To measure landmark displacements, we use visualization software such as *RegViz*, a program written

in IDL (Interactive Data Language, Research System Inc., USA) and created in the authors' laboratory for visualizing and analyzing registered image volumes. We navigate transverse, coronal, and sagittal MR images slice-by-slice to search the landmarks. The same unique features such as corners and intersections are identified with a cursor on magnified images. A single person repeats this several times over a few weeks, and results are averaged to give a 3D location for each landmark. A radiologist confirms the landmark selection. Following registration, we calculate the root-mean-squared (RMS) distance over the six landmarks [21].

Although this method provides an independent means for evaluating skeletal registration accuracy, there is error in localizing the bony landmarks. To determine the effect of localization error, we perform least-squares point-to-point registration [22] and compare results to MI registration. The rationale is that if we could identify point landmarks without error on the bony pelvis, point-to-point registration would be perfect. Hence, any displacement left after registration is introduced by localization error. We determine the optimal transformation for matching the six corresponding landmarks. Points are transformed, and distances between corresponding points are determined. RMS values are computed and compared to the RMS values from MI registration.

#### **3.2.3.2 Registration Consistency**

We use the registration consistency as proposed by [29] for registration evaluation. For each of the three volunteers, we use three volumes: reference, diagnosis, and empty bladder, all of which are obtained with the subject in the similar position. We call these three volumes A, B, and C, respectively. They give three pairs of registrations (A–B, B–C, and C–A) and three sets of transformation parameters ( $T_{ab}$ ,  $T_{bc}$ ,  $T_{ca}$ ). Using the transformation parameters, we transform voxel positions in A to B, and then to C, and then back to A. The distance between the original location and the final position is calculated. Since this is introduced by three transformations, we estimate the error for a single transformation, by multiplying by  $3^{-1/2}$  [29].

### 3.2.3.3 Voxel Displacements

To test the dependency of registration on algorithmic features such as image cropping, one can compare transformation parameters. However, we choose a more meaningful approach that consists of finding the average displacement of voxels in a region-of-interest (ROI) [20]. The 3D distances between transformed voxels are calculated in millimeters and averaged over a cubic ROI just covering the prostate.

# 3.2.3.4 Other Evaluation Methods Including Displacement of Prostate Centroids

A variety of other methods are used to evaluate the registration of the pelvis and prostate. First, we measure potential displacements of the 3D centroid of manually segmented prostates. Second, we use multiple visualization and analysis methods found in RegViz such as contour overlap and color overlay. Third, we calculate the intensity difference between the reference and registered volumes on a voxel-by-voxel basis and compute statistics. Fourth, we calculate the correlation coefficient (CC) between corresponding voxels to measure the quality of registration of two MR volumes acquired with identical parameters.

# 3.2.4 Examples and Results

## 3.2.4.1 Data Acquisition

We acquired MRI volumes from a 1.5 T Siemens MRI system (Magnetom Symphony, Siemens Medical Systems, Erlangen, Germany). An 8-element phased array body coil was used to ensure coverage of the prostate with a uniform sensitivity. Typically two anterior and two posterior elements were enabled for signal acquisition. We used two different MR sequences. First, we used a 3D RF spoiled gradient echo steady state pulse sequence (FLASH) with TR/TE/flip parameters of 12/5.0/60 which give  $256 \times 256 \times 128$  voxels over a  $330 \times 330 \times 256$ -mm field of view (FOV) to yield  $1.29 \times 1.29 \times 2.0$ -mm voxels oriented to give the highest resolution for transverse slices. The acquisition time was 5.63 min. This sequence was good for pelvic imaging but was not ideal for the prostate. Second, we used a 3D rapid gradient echo sequence (PSIF) designed to acquire the spin-echo component of the steady state response, rather than the free induction decay. The spin echo component formed immediately prior to the RF pulse, and it was shifted toward the prior RF pulse through appropriate gradient waveform design. The sequence with 9.4/5.0/60 (TR/TE/flip) yielded  $160 \times 256 \times 128$  voxels

over a  $219 \times 350 \times 192$ -mm rectangular FOV and  $1.37 \times 1.37 \times 1.5$ -mm voxels oriented to give the highest resolution for transverse slices. There was oversampling at 31% in the slice direction to reduce aliasing artifacts. The acquisition time was 4.25 minutes. Most often, we used the second sequence, which gave excellent image contrast for the prostate and its surroundings.

#### 3.2.4.2 Image Volumes for Registration

We acquired 3D MRI volume images from three prostate cancer patients and three normal volunteers under four conditions simulating anticipated situations in diagnostic and treatment applications. They are *diagnostic position*, treatment position, empty bladder, and diagnosis 1 week. In the diagnostic position, the subject laid supine throughout MR scanning. The reference volume was always obtained in the diagnostic position. In the treatment position, the subject was supine, and his legs were supported at 30°-60° relative to the horizontal and separated in a "V" with an angle of  $60^{\circ}$ -90° between the legs. This is similar to the lithotomy position used in some prostate therapies, and it should provide access for needle insertion in brachytherapy or RF thermal ablation. In some experiments, the subject micturated to create an empty bladder prior to imaging. For each subject, volumes were typically obtained within an imaging session of 1-2 hours. We imaged one subject (V2) a week before the standard imaging session, and we refer to this volume as diagnosis 1 week. Between volume acquisitions, subjects got off the MRI table, stretched, and walked around to ensure that they would assume a different position on the table. The coil array was centered on the prostate. All images of a subject were acquired using the same pulse sequence and acquisition parameters so as to ensure very similar gray values. In total, we registered 22 volume pairs consisting of one pair for each patient, six pairs for each volunteer, and one additional pair for volunteer V2.

#### 3.2.4.3 Assessments of Pelvic Registration

Following registration, we determined displacements between the six bony landmarks. For each subject, there was no consistent displacement of landmarks in one direction versus another. Hence, we measured 3D distances and determined RMS values over the six landmarks. Registration results are plotted in Fig. 3.4. The smallest errors are obtained when subjects are in the diagnostic position for



Figure 3.4: Registration error as determined from bony landmarks. Plotted are RMS distances as well as maximums and minimums that show the spread of the data. Conditions along the x-axis such as the "treatment" position are described in Methods. P's and V's refer to patients and volunteers, respectively. Averaging data across all subjects for the best case (diagnosis-reference) gives  $1.6 \pm 0.2$  mm. Averages are  $2.9 \pm 0.7$  mm and  $2.0 \pm 0.1$  mm for treatment-reference and empty bladder-reference, respectively.

both imaging sessions, labeled diagnosis-reference. The average error across the three patients and three volunteers is only  $1.6 \pm 0.2$  mm. Consistently larger errors are obtained when we compare volumes acquired in the treatment position with those in the reference position. Even though the MR acquisition technique used for the patients gave inferior image quality as compared to that for the volunteers, errors were small.

Additional error analyses are performed on the volunteer images to assess the accuracy of point landmark localization. We used images obtained with the rapid gradient echo sequence, which have improved contrast of the prostate and bony landmarks and which give us more confidence in measurements. The isotropic voxels are 1.4 mm on a side, almost as large as the  $1.7 \pm 0.5$  mm error obtained for the volunteer diagnosis-reference data. We assess the error in localizing the bony landmarks by performing point-based registration on 4–6 points per volume pair. The RMS distances after registration averaged across the three volunteers was



Figure 3.5: Image intensity difference between registered volume pairs. Plotted are means and standard deviations calculated over the entire volumes. Other details are given in the legend of Fig. 3.4. Eight out of 10 registrations have a mean absolute gray level difference less than 1.5 gray levels. Average results are  $0.9 \pm 1.8$ ,  $1.5 \pm 3.9$ , and  $0.4 \pm 2.6$  gray levels, or 1.0%, 1.7%, and 0.4% of typical mean values of 90 gray levels, for diagnosis-reference, treatment-reference, and empty bladder-reference, respectively.

 $1.5 \pm 0.2$  mm, very nearly the value obtained with MI registration. Hence, the "error" reported for MI is probably overestimated due to landmark location error. This analysis was prompted by ideas in a previous report [22] that numerically demonstrated the relationship between point localization uncertainty and point-based registration uncertainty.

Figure 3.5 shows image intensity differences between reference and registered volumes. The means are quite small with 8 out of 10 registrations giving a mean absolute value <1.5 gray levels, or only 1.7% of typical mean values of 90 gray levels for these 3D MR acquisitions. Again, the only consistent outliers occur when we compare the treatment position to the reference. For the case diagnosis-reference, extremely small image differences are found with V1 giving  $0.1 \pm 1.6$  gray values, a standard deviation that compares favorably to the expectation from image noise alone, or 1.5 gray values. For this volume pair, the subtracted images have very little structure except at the skin surface, indicating

excellent registration (not shown). We know that 3D alignment is achieved because all slices across the entire pelvis are well aligned and because rendered images show that the prostate matches well.

Registration consistency, as described in Section 3.2.3.2, provides yet another means of evaluating the quality of registration. Values were 0.4, 0.8, and 0.7 mm for volunteers V1, V2, and V3, respectively. The average is  $0.6 \pm 0.2$  mm, a value less than half the dimension of a voxel indicating excellent registration consistency.

### 3.2.4.4 Assessment of Prostate Registration

We determined the quality of prostate registration by visually examining nearly all of the roughly 800 registered image slices using one or more of the methods found in RegViz. A typical example for the case of diagnosis-reference is shown in Fig. 3.6 where the boundary overlap is excellent and probably within the manual segmentation error. In some other cases such as treatment-reference, small displacements of the prostate were observed. In a typical volume pair, the prostate is displaced to the posterior direction by  $\approx$ 3.0 mm when the legs are raised. There are no obvious displacements in other directions.

Centroid vector displacements can also be analyzed following registrations. For the case of diagnosis-reference, centroid displacements are only  $1.4 \pm 0.2$  mm. In the case of treatment-reference, there is a consistent displacement ( $\approx$ 3 mm) in the posterior direction with relatively little change in the two orthogonal directions. In the case of empty bladder-reference, two of the three volume pairs show a displacement in the posterior direction while the other is displaced in the anterior direction. Finally, in the case of a diagnostic volume obtained one week before the reference, there was a 4 mm displacement in the caudal direction due mostly to changes in rectal and bladder filling. Because the 3D centroid of the prostate averages over a large region, we believe these measurements to be relatively insensitive to segmentation error. Even so, we consider the uncertainty to be at least 1 mm, and displacements less than this should be disregarded. All significant results above can be visually confirmed.

Prostate volumes were measured for each subject. The typical difference between volumes in an imaging session was <1.5%, indicating that segmentation errors were small and that prostate volumes did not change. The average prostate volume for the healthy volunteers was  $23.9 \pm 3.2$  cm<sup>3</sup>. Volume measurements are particularly useful for clinicians when assessing the response



Figure 3.6: The prostate overlap between reference and registered images. Following registration, the prostate was manually segmented in reference (a) and diagnosis (b) images. The rectangular region in (b) is zoomed in (c) with both boundaries superimposed. Images are from volunteer V2.

of prostate cancer treatments such as brachytherapy, chemo- or radiation therapy.

# 3.2.4.5 Effects of Image Cropping

In Figure 3.7, we plot registration error as determined from bony landmarks with and without the cropping operation described in Methods. For the case treatment-reference, cropping always improved registration accuracy, and for V3, error reduced greatly from 12.6 mm to 3.4 mm. For all other cases, subjects were always in the supine position with legs flat on the table, and there was no consistent effect of cropping. If anything, cropping tends to increase error in these cases, with an increase in 5 of 7 volume pairs. Correlation coefficient always improves with image cropping.



Figure 3.7: The effect of image cropping on registration accuracy. The light and dark bars are RMS distances between bony landmarks with and without image cropping, respectively, as defined in Methods. Conditions on the *x*-axis are described in Methods.

Displacements of bony landmarks might significantly overestimate the change near the prostate. Hence, as described in section 3.3, we investigated the displacement of voxels in a ROI surrounding the prostate between registrations with and without cropping. For 9 of 10 analyzed volume pairs, the average voxel displacement was <0.5 mm indicating that prostate registration is fairly insensitive to cropping. However, for V3 treatment-reference, a much larger voxel displacement of 7.4 mm was obtained indicating that cropping is critical for this volume pair.

## 3.2.4.6 Implementation Issues

The algorithm was quite robust and always gave very nearly the same transformation parameters (less than 0.01 voxels and 0.01°) for the 22 volume pairs in this study using a wide variety of initial guesses. The restarting and multiresolution features are important, and we report some results for a typical volume pair registration. The multiresolution approach enabled the program to get close to the final value quickly because of the reduced number of calculations. That is, the time for reformatting at the lowest resolution (1/4) was 9.8 sec, which was less than 1/59 times that at the highest resolution, a value nearly equal to the 1/64 expected from the change in the number of voxels. The number of restarts was 5, 1, and 1 for resolutions at 1/4, 1/2, and the full number of voxels. Each call to the simplex optimization resulted in 55 to 94 MI evaluations before the tolerance (0.001) was reached. The simplex optimization method worked about 1.5–2.0 times faster than the Powell method in our implementation. The time for registration using Simplex, typically 5 minutes on a Pentium IV, 1.8GHz CPU, with 1GB of memory, could probably be greatly improved with optimized C code rather than IDL.

## 3.2.5 Discussion

#### **3.2.5.1 Registration Accuracy**

Our results suggest that MI can be used to accurately register, with an error on the order of a voxel, MR pelvic images obtained under similar conditions. Because it gives an independent, true 3D measurement, we like to use the method of point bony landmarks to assess accuracy. However, as argued in section 3.2.4.3, the true MI registration accuracy might be better than our ability to measure it with point bony landmarks. That is, following point landmark registration, the distance between registered, corresponding landmarks was on the order of that following MI registration. Very possibly, MI is more accurate than point registration using bony landmarks. Additional, independent evidence of excellent MI accuracy comes from the very low error value from the registration consistency measurement ( $0.6 \pm 0.2$  mm). Interestingly, this is obtained even though the interpolation artifact present in MI similarity surfaces should reduce the likelihood of subvoxel accuracy [30]. Our results for the pelvis with image volumes obtained under the same conditions compare favorably with those for the brain, where MI registers images very accurately giving errors as small as 0.7–0.8 mm for CT-MR [31].

Visual and quantitative evaluation of prostate organ movement showed good registration even when we acquired images under conditions that greatly stressed the ability to register the images. The small prostate displacements in our study are consistent with earlier reports on respiration-induced prostate movement of  $\leq 1$  mm for most patients in supine position with "quiet" respiration [32]. The difference between the treatment and diagnostic positions resulted in the most consistent and largest displacement of the prostate. When images were acquired in the diagnostic position one week apart, there was significant displacement of the prostate due to a change in rectal filling. This is consistent with previously reported results [8, 14], which found rectal filling to be a significant factor in prostate displacement.

There are ways to limit the small displacements of the prostate. One obvious remedy is to acquire images in the same position. That is, if we want to register an image with one obtained in the treatment position, we should obtain it in the treatment position. Although it is unknown how accurately one must repeat the treatment position, a device to support and constrain the legs is probably required. In addition, there is a dependence of registration error on bladder and rectum content. One solution is use clinical preparations often employed to void the bladder and rectum prior to prostate imaging or therapy. We anticipate that this might even lessen prostate displacements between the diagnostic and treatment positions.

We must consider our results with regard to potential applications such as those described in section 3.1. First, registered images acquired before and immediately after treatment can be used to determine whether a tumor is adequately treated. Second, serial examinations can be registered to determine tumor progression or regression. Third, registration of functional images from other modalities such as nuclear medicine or from MR spectroscopy can give molecular markers for prostate cancer [33, 34]. Fourth, we want to register high quality MR images with a few live-time interventional MR images to aid treatment decisions [4, 5]. Our results indicate that registering images from the treatment and diagnostic positions can lead to errors, and potential steps are described above to limit this error. With images acquired in the same position, our results place a lower limit on registration error of about 1 voxel.

#### 3.2.5.2 Assessment of Registration

We are involved in a long-term effort to use registration for detection, assessment, and therapy of prostate cancer. Hence, we have developed and used several methods to assess pelvic and prostate registration. It is highly desirable to have an automatic method for evaluating the quality of a registration so that a poor one can be flagged before it is used clinically. The correlation coefficient would be applicable whenever one uses MR images obtained with identical pulse sequences. It compares favorably with the bony landmark results. Registration consistency provides an additional means to evaluate registration accuracy that does not rely on operator interaction.

Other evaluation methods are applicable for clinical or research applications. RegViz provided visual inspection tools for quick evaluation of the quality of registration and potential prostate displacement. Such methods can be used to verify the quality of registration and possibly account for small displacements in some applications. Boundary overlays provide a good means to evaluate organ deformation as well as displacement. Point anatomical landmarks provide a useful, independent test, but it is time consuming to identify them and MI might be more accurate than the point landmarks. Centroids are obtained reliably because small segmentation errors are removed by integrating over the entire prostate volume. Centroids provide a good means of quantifying prostate displacements.

#### 3.2.5.3 Algorithm with Combined Similarity Measures

Using both CC and MI at different resolutions was an important feature that increased robustness. When only mutual information was used, registrations at low resolution sometimes gave false solutions that mislead registration at the next higher resolution. However, CC performed well and gave many fewer local maximums at the lower resolutions (Figs. 3.1a and 3.1b). But MI gave a more accurate solution at the full resolution due to the peaked MI surface (Figs. 3.1c and 3.1d). Our registration algorithm combined advantages from the two similarity measures.

There are probably several reasons why mutual information does not work well at low resolution. First, the similarity curve is noisy with periodic oscillations from the so-called interpolation artifact [30] that is accentuated at reduced resolutions [35]. This results in the many local maximums in Fig. 3.1a that can trap the optimization. A similar result was reported for brain registration [19, 36]. Second, when images are of low resolution and there is only a small region of overlap, the mutual information function can even contain incorrect global maximums [35]. Such a result was found in Fig. 3.1a where the global maximum was obtained at very large displacements where the overlap was reduced. This occurs because MI is not only a function of how well the images match in the overlap, but also by how much information is provided by the two images in the overlap [37].

#### **3.2.5.4** Computer Implementation

Accuracy is an important issue for automatic registration, but there are others such as robustness, speed, and requirements for operator interaction. With the multiresolution and restarting features, our modified MI algorithm is quite robust. For a wide range of initial guesses, it worked well for all 22 volume pairs reported here. Three of the volume pairs were from patients, and we are confident that routinely acquired clinical images will have sufficient quality for registration. Because good starting values are unimportant, operator interaction is minimal. In one instance, cropping of the legs was important for registering an image volume obtained in the treatment position with that in the diagnostic position. It is not surprising that legs in a very different position have to be cropped. Although this is easy to do manually, we can probably determine an automated method if it is deemed desirable.

The mutual information similarity measure is quite robust. Even though our images are very similar, we had less success with some other measures such as the sum of the squared image difference. An advantage of MI is that it can be used with images from different modalities, a feature that we are starting to use.

# 3.3 Three-Dimensional Non-Rigid Body Registration Algorithm

# 3.3.1 Why Non-Rigid Registration

In the previous section, we discussed rigid body registration of the prostate. For volume pairs acquired over a short time span from a supine subject with legs flat on the table, registration accuracy of both prostate centroids (typically <1 mm) and bony landmarks (average 1.6 mm) was on the order of a voxel ( $\approx$ 1.4 mm).

We obtained somewhat larger prostate registration errors of about 3.0 mm when volume pairs were obtained under very different conditions, e.g., legs flat and legs raised, or with and without bladder or rectal filling. Rigid body registration of the pelvis cannot follow prostate movements due to changes in the postures of legs and deformation of the bladder and rectum [8]. In this section, we discuss the ability of non-rigid registration to express this deformation.

Non rigid registration studies are reported for the brain [38, 39], for the breast [40, 41, 41, 42, 42], for a variety of other organs [23, 43, 45], and for excised tissue [46]. Far few reports described results of the pelvis and prostate. Bharaha *et al.* reported a method using manually segmented prostate for rigid body registration followed by finite element-based warping in the application of prostate brachytherapy [47]. Voxel based methods, particularly those based upon mutual information, are robust, require no segmentation that can be prone to error, are highly accurate for brain registration [31], and are suitable for abdominal registration where there can be deformation [20]. We are discussing voxel-based non-rigid registration for the particular application in the pelvis and prostate.

In this section, we perform experiments to compare non-rigid and rigid body registration for the prostate and pelvis. By using high-resolution MR images giving distinctive anatomic detail, we test the ability of a non-rigid algorithm to correct anatomical variations throughout the pelvic region. We include conditions with very significant changes in posture possible in interventional applications, that is, we attempt to register image volumes from a diagnostic scan with legs flat to those from a treatment acquisition with legs raised. We qualitatively and quantitatively evaluated registration results using 17 volume pairs from three volunteers.

## 3.3.2 Non-Rigid Registration Algorithm

Figure 3.8 outlines the non-rigid registration algorithm that includes three major steps: control point selection, control point optimization, and thin plate spline warping. Prior to non-rigid registration, we perform rigid body registration as reported in Section 3.2. Again, the unchanging volume is the *reference*, and the one to be warped is *floating*.

The manual selection of CP's is an important step. We used *RegViz* for visualizing and analyzing image volumes. Following rigid body registration, the

Register two volumes using rigid body registration Create a new floating volume from rigid body registration parameters Select N control points (CP) in the reference and new floating volumes Record 3D coordinates of the CP's in both volumes Initialize the size of volume of interest (VOI) centered at CP's Bin the reference and floating volumes to 256 gray levels FOR CP FROM 0 to N-1 DO BEGIN Optimize mutual information (MI) between the reference and floating VOI's 1. Transform the floating VOI using three translation parameters 2. Interpolate to get a reformatted VOI 3. Calculate MI between the reference and reformatted VOI's 4. Vary the three translation parameters 5. Repeat the above steps 1-4 until meeting function tolerance or maximum iteration number Record the optimized CP coordinates in the floating volume END Calculate thin plate spline transformation using the reference and optimized CP's Interpolate the floating volume and get a warped volume

Figure 3.8: Flow chart of the warping registration algorithm. Following rigid body registration, N control points are selected in both the reference and floating volumes. A small cubic volume of interest (VOI) is centered on each control point. Optimization is performed by varying the x, y, and z locations of the floating VOI until the mutual information between corresponding voxels is maximized. Each control point is optimized independently, and then the optimized control points are used to establish a three-dimensional thin plate spline transformation for the entire volume.

aligned two volumes are displayed in two rows slice-by-slice. Images can be transverse, coronal, or sagittal slices. It is quite straightforward to find corresponding features at the pelvis, prostate, bladder, and rectum. We normally select control points (CPs) using recognizable organ features such as corners and intersections of edges because of their unique positions. Corresponding CPs in the two volumes are placed using a cursor, and sometimes they are in different image slices. The 3D coordinates are automatically stored in a file. Because of the optimization that occurs later, the correspondence can be up to 15 mm or  $\approx 10$  voxels in error. Experiences with CP selection are described in Section 3.3.4.4. Typically, we used 180 CPs for a volume with  $256 \times 256 \times 140$  isotropic voxels.

The next step of the non-rigid algorithm (Fig. 3.8) is the CP optimization. We define a small cubic volume of interest (VOI) centered at each CP. The VOI can be 16, 32, 48 or 64 voxels on a side. As reported later, the selection of the VOI size depends on the amount of warping required. A simplex optimization algorithm varies the x, y, and z transformation parameters of the floating VOI until the mutual information with the reference VOI is optimized. Each control point is optimized independently and the 3D coordinates of the optimal CPs are recorded.

The final major step is to warp the floating volume using the corresponding optimal CPs coordinates to establish a three-dimensional thin-plate spline (TPS) transformation [48, 49]. We now briefly go through the three computing steps for the TPS transformation.

First, let  $P_1 = (x_1, y_1, z_1)$ ,  $P_2 = (x_2, y_2, z_2)$ ,...,  $P_n = (x_n, y_n, z_n)$  be *n* control points in the image coordinate of the *reference* volume. Write  $r_{ij} = |P_i - P_j|$  for the distance between point *i* and *j*. We define matrices:

$$P = \begin{bmatrix} 1 & x_1 & y_1 & z_1 \\ 1 & x_2 & y_2 & z_2 \\ \dots & \dots & \dots \\ 1 & x_n & y_n & z_n \end{bmatrix}, \quad n \times 4;$$
$$K = \begin{bmatrix} 0 & r_{12} & r_{13} & \dots & r_{1n} \\ r_{21} & 0 & r_{23} & \dots & r_{2n} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & r_{n3} & \dots & 0 \end{bmatrix}, \quad n \times n;$$

and

$$L = \begin{bmatrix} K & P \\ P^T & O \end{bmatrix}, \quad (n+4) \times (n+4);$$

where T is the matrix transpose operator and O is a  $4 \times 4$  matrix of zero.

Second, let  $Q_1 = (u_1, v_1, w_1)$ ,  $Q_2 = (u_2, v_2, w_2)$ , ...,  $Q_n = (u_n, v_n, w_n)$  be *n* corresponding control points in the image coordinate of the *floating* volume. We get matrices:

$$V = \begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \\ w_1 & w_2 & \cdots & w_n \end{bmatrix}, \quad 3 \times n,$$
$$Y = \left( V | 0 \quad 0 \quad 0 \quad 0 \right)^T, \quad 3 \times (n+4),$$

and define the vector  $W = (w_1, w_2, ..., w_n)$  and the coefficients  $\alpha_1, \alpha_x, \alpha_y$ , and  $\alpha_z$  by the equation

$$L^{-1}Y = (W|\alpha_1 \quad \alpha_u \quad \alpha_v \quad \alpha_w)^T.$$

Third, use the elements of  $L^{-1}Y$  to define a function f(u', v', w') everywhere in the entire volume:

$$f(u', v', w') = \alpha_1 + \alpha_u u + \alpha_v v + \alpha_w w + \sum_{i=0}^n w_i |P_i - (u, v, w)|$$

Thus, any voxel  $(u_i, v_i, w_i)$  in the floating volume is transformed to a new coordinate  $(u'_i, v'_i, w'_i)$  and a warped volume can be obtained by trilinear interpolation.

Additional algorithm details are now described. For both VOI optimization and rigid body registration, we use trilinear interpolation. Optimization of similarity ends either when the maximum number of calculations is reached (typically 500) or the fractional change in the similarity function is smaller than a tolerance (typically 0.001). We use IDL as the programming language.

## 3.3.3 Registration Evaluation

We used the multiple visualization features of RegViz to visually evaluate registration results. First, we manually segmented prostate boundaries in image slices and copied them to corresponding slices from the other volume. This enabled visual determination of the overlap of prostate boundaries over the entire volume. We applied the same method to evaluate pelvic registration. Second, color overlay displays were used to evaluate overlap of structures. One image was rendered in gray and the other in the "hot-iron" color scheme available in IDL. To visualize potential differences, it was quite useful to interactively change the contribution of each image using the transparency scale. Third, we used a sector display, which divided the reference and registered images into rectangular sectors and created an output image by alternating sectors from the two input images. Even subtle shifts of edges could be clearly seen [1].

Voxel gray value measures were calculated as indicators of registration quality. Mutual information and correlation coefficient between registered volumes were computed. Since volumes to be registered were acquired using the same acquisition parameters, high absolute CC values were obtained when registration was good [41]. Because voxel intensities were comparable, we created difference images and calculated statistics such as the voxel mean and standard deviation following registration.

Finally, we used a variety of tools in RegViz to evaluate registration quality. We used contour overlap and color overlay to assess the prostate registration. We manually segmented the prostate across all slices and calculated the potential displacements of the prostate 3D centroid.

# 3.3.4 Examples and Results

### 3.3.4.1 Image Acquisition

All MRI volumes were acquired using a 1.5 T Siemens MRI system (Magnetom Symphony, Siemens Medical Systems, Erlangen, Germany). As described in Section 3.2.4.1, we used two MR sequences. First, a 3D FLASH sequence with TR/TE/flip parameters of 12/5.0/60 gave  $256 \times 256 \times 128$  voxels over a  $330 \times 330 \times 256$ -mm field of view (FOV) to yield  $1.29 \times 1.29 \times 2.0$ -mm voxels oriented to give the highest resolution for transverse slices. This sequence was used for volunteer S1. Second, a 3D PSIF sequence with 9.4/5.0/60 (TR/TE/flip) yielded  $160 \times 256 \times 128$  voxels over a  $219 \times 350 \times 192$ -mm rectangular FOV and  $1.4 \times 1.4 \times 1.5$ -mm voxels oriented to give the highest resolution for transverse slices. The second sequence was used for volunteers S2 and S3.

#### 3.3.4.2 Imaging Experiments

We acquired 3D MRI volume images from three normal volunteers under a variety of conditions simulating anticipated conditions in diagnostic and treatment applications. Before image acquisition, each volunteer drank water and had a relatively *full bladder*. In the *diagnostic position*, the subject laid supine throughout MR scanning. In the *treatment position*, the subject was supine, and his legs were supported at  $30^{\circ}$ – $60^{\circ}$  relative to the horizon. In some experiments, the subject micturated to create an *empty bladder* prior to imaging. We imaged volunteers a week before the standard imaging session, and we refer to these volumes as *diagnosis 1 week*. Between volume acquisitions, volunteers got off the MRI table, stretched, and walked around to ensure that they would assume a different position when they laid back on the table. All images of a volunteer were acquired with the same MRI acquisition parameters so as to ensure very

similar gray values. In total, there are 4, 4, and 8 volumes for volunteer S1, S2, and S3, respectively. The permutation of the volumes gives many possible volume pairs for registration experiments.

#### **3.3.4.3** Volumes for Registration Experiments

We registered 17 volume pairs under five different conditions as defined above. Five pairs are *treatment-diagnosis*; seven pairs are *full bladder-empty bladder*; two pairs are *diagnosis 1 week-diagnosis*; and three pairs are *diagnosisdiagnosis*. For each case, other conditions were controlled. For example, for the case of diagnosis 1 week-diagnosis, both volumes were acquired with empty bladder and comparable conditions. Rigid body and warping registration were applied to each of the volume pairs. Results were evaluated as described next.

# 3.3.4.4 Effect of Control Point Selection on Registration Quality

In well over 100 registration experiments using different numbers and placement of CPs, we investigated effects on non-rigid registration quality. For each of the three volunteers, we selected one typical volume pair from the diagnostictreatment positions for systematic experiments. We progressively increased the number of CPs from 15 to 250. We found that less than 120 CPs did not produce good visual matching of our high-resolution MR images showing great anatomical detail. More than 220 CPs did not give significantly improved results but required more time for manual selection and optimization. When we used  $\approx 180$ CPs placed strategically using rules described later, we obtained excellent results over the entire pelvis and internal organs. As a result of our experience, we modified the registration method to be suitable for many CPs.

Some rules follow for strategic placement of CPs. For registration of treatment and diagnostic image volumes, most CPs were selected using transverse slices because they best showed the pelvic displacement when moving the legs to the treatment position (Fig. 3.9). About 25 CP pairs were placed near the edge and point features having recognizable correspondence on each of 5–8 transverse slices with a *z* interval of  $\approx$ 8 mm, covering the entire pelvic region. Additionally, we placed about 25 CPs from sagittal slices



Figure 3.9: Control point selection when images are acquired in the treatment and diagnostic positions. Image (a) is from the reference volume acquired in the treatment position with legs raised. Image (b) is to be warped and is from the volume acquired in the diagnostic position with the subject supine on the table. Transverse slices best show the deformations, especially at the legs. As described in the text, control points indicated by the white dots are selected around the pelvic surface and the prostate. Each control point is located at one voxel but displayed much bigger for better visualization. Volumes are from volunteer S2.

because they provided other structures that can be missed in the transverse images. It was also important to include CPs from organs other than the prostate because they constrained warps. We always placed CPs at critical regions such as the prostate center, pelvic surface, bladder border, and rectal walls.

For registration of image volumes with full and empty bladder, most CPs were placed from sagittal slices because they best showed the deformation of the bladder and rectum (Fig. 3.10). About 10–20 CPs were placed at the borders of the bladder and rectum on each of 8–10 sagittal slices with an equal interval of  $\approx$ 8 mm, covering the entire pelvic region including the prostate, bladder, and rectum.



Figure 3.10: Control point selection when images are acquired with a week's interval between them. Image (a) is from the reference volume acquired one week later with an empty bladder. Image (b) is to be warped and is from the volume acquired earlier with a full bladder. Sagittal slices best show the deformations at the bladder (vertical arrow) and rectum (horizontal arrow) where most control points are placed. Volumes are from volunteer S3.

# 3.3.4.5 Registration Quality of Non-Rigid and Rigid Body Registration

In Fig. 3.11 we compare non-rigid and rigid body registration for a typical volume pair in the treatment and diagnostic positions. Following non-rigid registration, the prostate boundary overlap is excellent (Fig. 3.11e) and probably within the manual segmentation error. Similar results were obtained in other transverse slices throughout the prostate. The prostate 3D centroid calculated from segmented images displaced by only 0.6 mm, or 0.4 voxels, following warping. Following rigid body registration, the prostate was misaligned with a displacement to the posterior of  $\approx$ 3.4 mm when in the treatment position (Fig. 3.11d), as previously reported by us [1]. Using rigid body registration, there is significant misalignment throughout large regions in the pelvis (Fig. 3.11f) that is greatly reduced with warping (Fig. 3.11g). Note that warping even allows the outer surfaces to match well. Other visualization methods such as two-color overlays and difference images, quickly show matching of structures without segmentation but do not reproduce well on a printed page.



Figure 3.11: Comparison of warping and rigid body registration for volumes acquired in the treatment and diagnostic positions. Image (a) is from the reference volume acquired in the treatment position, and the prostate is manually segmented. Images in the left and right columns are from the floating volume acquired in the diagnostic position following rigid body and warping registration, respectively. To show potential mismatch, the prostate contour from the reference in (a) is copied to (b) and (c) and magnified as the dashed contours in (d) and (e). The 3 mm movement of the prostate to the posterior is corrected with warping (e) but not rigid body registration (d). Pelvic boundaries manually segmented from the reference show significant misalignment with rigid body (f) that is greatly improved with warping (g). Images are transverse slices from subject S2.

We next examine the effect of conditions such as bladder and rectal filling that might change from one imaging session to the next. In Fig. 3.12 we compare non-rigid and rigid body registration for a volume pair with one-week between imaging sessions. One volume is with an empty bladder and the other is with a relatively full bladder. There is also a difference in rectal filling. Non-rigid registration closely aligns the prostate (Fig. 3.12e) while rigid body does not


Figure 3.12: Comparison of rigid body and warping registration for volumes acquired with an interval of one week between imaging sessions. The reference image (a) with a manually segmented prostate was acquired later with an empty bladder (vertical arrow) and partial rectal filling (horizontal arrow). Images in the left and right columns are from the floating volume acquired earlier following rigid body and warping registration, respectively. To show potential mismatch, contours from the reference are shown on images following registration, as described in Fig. 3.11. The full bladder in (d) has pushed the prostate, shown by the continuous curve, in the caudal direction. After warping, prostate contours match closely (e). The bladder, rectum, and other organs closely align following warping (g). With rigid body (f), proceeding from left to right, the front of the pelvis, the bladder (arrow), and the rectum are all misaligned. Images are sagittal slices from S3.

(Fig. 3.12d). In addition, rigid body registration does not align the bladder and parts of the rectum (Fig. 3.12f). With warping, the bladder closely matches the reference, and the rectum is better aligned (Fig. 3.12g). Other visualization methods showed excellent alignment of internal and surface edges. Difference images show that warping greatly improves alignment of internal structures as compared to rigid body registration (Fig. 3.13). The difference image following rigid body registration shows bright regions indicating misalignments (Fig. 3.13d) that are removed with warping (Fig. 3.13e).

We also examined volume pairs with both volumes acquired in the diagnostic position under comparable conditions. In the current data set, five volume



Figure 3.13: Comparison of registration quality for rigid body and warping registration. The reference image (a) was acquired with a relatively empty bladder (arrow). Images (b) and (c) are from the floating volume acquired with a full bladder following rigid body and warping registration, respectively. Images (d) and (e) are the absolute difference images between the reference and registered images, respectively. Bright regions following rigid body indicate misalignments (d) that are removed with warping (e). Images (d) and (e) are displayed using the same grayscale window and level values. Images are coronal slices from S3 volumes shown in Fig. 3.12.

pairs fit these criteria. In all such cases, rigid body registration worked as well as warping. There were no noticeable deformations in the pelvis, and prostate centroids typically displaced less than 1.0 mm between the two registered volumes. Note that this was obtained even though subjects always got up from the table and moved around before being imaged again.

### 3.3.4.6 Quantitative Evaluation of Non-Rigid Registration

Figure 3.14 shows the correlation coefficient and mutual information values between registered volumes. Warping increased CC and MI values in every case, and a paired two-tailed *t*-test indicated a significant effect of warping at p < 0.5%. The most significant improvement was in the case of treatment-diagnosis where improvements in CC and MI were as high as 102.7% and 87.8%, respectively.

Statistics of image differences following rigid body and non-rigid registration are shown in Fig. 3.15. Warping reduces the absolute intensity difference between corresponding voxels (Fig. 3.15a), and the mean across all image volumes is only 4.2 gray levels, a value corresponding to only 4.7% of the mean image value of 90. We used the absolute intensity difference because signed



Mutual Information

1.0 0.8

0.6

0.4 0.2 0.0

Diagnosis

Treatment

Figure 3.14: Voxel similarity measures for rigid body and warping registration. Correlation coefficient (CC) (a) and mutual information (MI) (b) following registration with (light bars) and without (dark bars) warping are plotted. Conditions described in Experimental Methods are listed on the *x*-axis. Warping increase CC and MI in all cases. The most significant increases occurred in the case of the treatment-diagnosis volume pairs where maximum increase in CC and MI are 101.7% and 87.8%, respectively. For volumes acquired within the same diagnostic position and comparable conditions (two right-most groups), warping did not have significant improvement over rigid body method.

Full Bladder ~

Empty Bladder

Diagnosis

1 Week

Diagnosis

Diagnosis

Diagnosis



Figure 3.15: Image statistics of absolute intensity difference images for rigid body and warping registration. The mean (a) and standard deviation (b) are plotted. See the legend of Fig. 3.14 for other details. Warping decreased the mean and standard deviation in each case, but the most significant decreases occurred in the case of the treatment-diagnosis volume pairs. After warping, the intensity averaged over all data is  $4.2 \pm 1.9$  gray levels, a value corresponding to only  $\approx 4.7\%$  of the mean image value of  $\approx 90$  gray levels.



Figure 3.16: Optimization time and mutual information as a function of VOI size. The left vertical axis is mutual information, and the right vertical axis is the total VOI's optimization time. The horizontal axis is the size of the VOI on a side. In each case, the VOI is centered on the CP, but since even numbers of voxels are used, the CP is displaced consistently to the upper left-hand corner by one voxel. With increasing VOI size, time increases linearly with the number of voxels within the VOI. The peak mutual information value is at a VOI size of 64 on a side. A treatment-diagnosis volume pair is used from S2 with 180 CPs.

values canceled when averaged over the entire image. The standard deviation of absolute difference is also reduced (Fig. 3.15b).

These quantitative measures match observation from visual inspection. For example, the third pair of the first group (diagnosis-treatment) in Figs. 15 and 16 correspond with the images in Fig. 3.11. After warping, registration greatly improved. Another interesting example is the difference images in Figs. 3.13d and 3.13e that correspond to the last pair of the second group (full-empty bladder) in Fig. 3.15. Once again, the statistical measures reflect the great change in visual quality.

### 3.3.4.7 Algorithmic Implementation

In rigid body registration, the multiresolution approach and restarting algorithm were important modifications. First, these two features improved robustness.

The algorithm always gave very nearly the same transformation parameters  $(<0.01 \text{ voxels and } 0.01^{\circ})$  for the 17 volume pairs in this study using a wide variety of initial guesses. We also found that MI was more accurate than CC at the highest resolution [1]. Second, the multiresolution approach enabled the program to get close to the final value quickly because of the reduced number of calculations. That is, the time for reformatting at the lowest resolution of 1/4 number of voxels in a linear dimension was 0.16 minute, less than 1/63 times that at the highest resolution, a value nearly equal to the 1/64 expected from the change in the number of voxels. In a typical example, the number of restarts was 5, 1, and 1 for resolutions at 1/4, 1/2, and the full number of voxels in a linear dimension, respectively. When we checked the restarts at the resolution of 1/4 number of voxels, we determined that none of the five restarts converged to the same transformation. It has been our experience that more restarts are desirable at the lower resolutions, and the algorithm includes this feature. Each call to the Simplex optimization resulted in 50 to 100 MI evaluations before the tolerance (0.001) was reached. In some experiments on multiple volumes, we reduced the tolerance value but found little difference in registration quality, probably because of the restarting and multiresolution features. The time for rigid body registration, typically 5-10 min on a Pentium IV, 1.8 GHz CPU, with 1.0 GB of memory, could possibly be reduced to within one minute with optimized C code rather than the high level language IDL.

Some technical aspects of non-rigid registration are of interest. Fig. 3.16 shows the optimization time and MI values between registered volumes as a function of VOI size. The optimization time for 180 CPs increases roughly linearly with the number of voxels within a VOI, about 0.5 minutes for VOIs with 16 voxels on one side and 30 minutes for VOIs with 64 voxels on a side. In Fig. 3.16, the MI curve saturates at the VOI size of 64 voxels on a side which means the size of 64 gave better MI value. These curves are for the case of treatment-diagnosis for volunteer S2. When we examined the cases of full-empty bladder and volumes acquired over one week time interval, we found that the VOI size of 16 voxels on a side worked best. Using the same computer above, for a volume with  $256 \times 256 \times 140$  voxels and 180 CPs, the non-rigid registration typically takes about 15–45 minutes depending on the VOI size.

We report some details on VOI optimization for a typical treatmentdiagnosis volume pair from subject S2. Following rigid body registration, the mean distance between the manually selected reference and floating CPs was  $15.5 \pm 0.7$  mm, where the latter number is the standard deviation. The maximum distance was 53.2 mm. After VOI optimization, the algorithm moved the floating CPs an average of  $9.0 \pm 6.5$  mm. This value shows that one does not have to be very careful in marking corresponding CPs.

### 3.3.5 Discussions

### 3.3.5.1 Applicability of Non-Rigid Registration

For MR images of the pelvis and prostate, non-rigid registration is desirable whenever images are acquired in different positions or with different conditions of bladder and rectal filling. Local deformations throughout the pelvis can be corrected, and, more importantly, the prostate can be accurately registered. However, when images are acquired in the same position under comparable conditions such as our case called diagnosis–diagnosis, rigid body registration worked satisfactorily as previously reported by us [1]. Similarly, if one were to reproduce the treatment position with reasonable accuracy, we believe that prostate registration would be very good.

Our goal is to get good matching throughout the entire pelvic region not just at the prostate because proper localization of other organs is important for interpretation of some functional images and because anatomical spatial integrity is important for treatment planning. Hence, we used high-resolution MR images that provide a very stringent test for warping. Many anatomical details are evident, and even a small mismatch can clearly be seen. As a result, we found that  $\approx$ 180 control points were required to get excellent quality registration. When we applied the method to register CT images with PET images of the lung having much less resolution, many fewer points ( $\approx 50$ ) were required [50]. With a sufficient number of control points, the TPS transformation excellently approximated the deformations of the pelvis and internal structures of our MR images. Even when we warped the volume in the diagnostic position to one in the treatment position, most organs were closely aligned, despite very significant movements. The method performed equally well for correcting the deformation and organ displacement arising from changes in bladder and rectal filling.

With our graphical user interface, interactive control point selection is quite easy after training. It usually took an experienced user about 15 minutes to select 180 CPs. Based on our experience, we think that it is possible to create an automatic or semiautomatic method for selection of appropriate CPs in the pelvis. For example, one might use a grayscale threshold to detect the pelvic outer boundaries and apply edge enhancement to extract feature of internal structures. CPs would be placed on such structures automatically. We are investigating this and other methods for CP selection followed by automatic non-rigid registration.

One way to adjust the movement of control points is to change the size of the VOI. In the case of treatment-diagnosis volume pairs, a large VOI size of 64 on a side worked better than smaller ones because displacements were large, because larger VOIs tend to give a more robust optimization, and because no small local deformations were required. However, a size of 16 on a side worked better for the case of full-empty bladder volume pairs because small VOIs better capture the small, local deformations. VOIs with a size of 64 on a side covered most of the bladder and could not generate small local deformations. For volumes with both large and small-scale deformations, we suggest using different VOI sizes for different CPs.

With non-rigid registration, we have to be concerned about potential warping errors affecting the application of interest. For the prostate, we used 3–5 CPs near the prostate center because we desired to maintain the spatial integrity of the organ and to preserve the tissue volume. We placed many CPs around the pelvic surface to produce reasonable warping.

### 3.3.5.2 Evaluation of Non-Rigid Registration

Since there is no gold standard for non-rigid registration of anatomical images, we used a variety of methods to evaluate registration quality. First, for routine evaluation, a color overlay is simple, fast, and intuitive. To better visualize the two data sets, we interactively adjust the transparency scale of each image. Second, for illustration of subtle difference along an edge, we recommend a sector display because it best shows small shifts. Third, for visual evaluation of a specific organ such as the prostate, we like to superimpose manually marked contours from one image onto another as shown in Figs. 3.12 and 3.13. This clearly shows any displacement or deformation even in a printed figure. Fourth, a more quantitative approach is obtained by calculating the displacement in

millimeters from the 3D centroid of a segmented organ such as the prostate. Finally, when images have comparable gray levels, a difference image can provide a visual evaluation or a quantitative evaluation from image statistics. A downside with MR difference images is that the inhomogeneity of the signal response and interpolation can introduce artifacts in difference images. Since MR image intensity can vary with different MR sequence parameters and the signal response of MR coil, gray value statistic may have some limitations when image acquisitions are not carefully repeated.

### 3.3.5.3 Algorithmic Robustness and Efficiency

The rigid body algorithm is robust for a global registration. Because of two principal design features, the algorithm is quite robust and accurate for volume pairs acquired in the same positions and with comparable conditions [1]. First, using both CC and MI at different resolutions was an important feature that increased robustness. CC gave fewer local minimums at low resolutions and MI was more accurate at high resolution [1, 5]. Second, the restarting mechanism was also quite important. Without restarting, we found that registrations sometimes failed in cases of volumes with large mismatches and significant deformation. Even these cases resulted in a proper solution when restarting was employed.

Based upon our initial experiments with interactive CP selection, we determined that many CPs were required for good matching throughout the pelvis. As a result, we designed algorithm features to be computationally efficient for TPS warping with hundreds of CPs. First, the optimization of small VOIs is very fast. Second, we optimized each CP separately because the optimization of three parameters (x, y, and z) is simple and fast. Conversely, as previously reported by others [43, 44], the simultaneous optimization of many CPs leads to a much more complicated error surface and local maximums. If one were to use 180 CPs and optimize the 540 free parameters simultaneously, the optimization process would become extraordinarily complex. Third, we applied the TPS transformation once to the final, optimal CPs, which saved considerable time. If TPS was applied in each iteration, the registration time would be unacceptable for our application. If we were to use optimized C code, the total time for rigid body and non-rigid registration should reduce to within 5 minutes.

### 3.3.5.4 Applications

The flexibility introduced with manual selection of CPs makes the current software suitable for non-rigid registration in many applications in addition to the clinical procedures described in the Introduction. We have successfully applied it to human MR-MR prostate images as shown here, rat CT-CT images, and CT-PET lung images [50]. We believe that the registration method can be applied to many organs other than the pelvis and prostate, multimodality images, and intersubject images. In addition, we think it applicable to a variety of animal experiments in which we are involved, including iMRI-guided thermal ablation in pig and rabbit, prostate imaging studies in dog, and controlled drug release studies in rat.

We conclude that our mutual information non-rigid registration is fast and can be applied to a variety of applications. For prostate and pelvic imaging, it works better than rigid body registration whenever the subject position or condition is greatly changed between acquisitions. It will probably be a useful tool for many applications in prostate diagnosis, staging, and therapy.

### 3.3.6 Summary

Many applications in prostate cancer management such as tumor localization, possibly tumor staging, tumor targeting during therapy, assessment of adequate treatment, and treatment follow up, require image registration of MRI volumes and/or volumes from other imaging modalities. With regard to interventional MRI guided radiofrequency thermal ablation for the minimally invasive treatment of prostate cancer, registration applications include the comparison of registered MR images acquired before and immediately after RF ablation to determine whether a tumor is adequately treated. When images are acquired in different patient positions and/or different conditions, the pelvis, prostate, bladder, and rectum can deform and displace. Non-rigid registration is desired to correct for such deformations.

We created a two-step, three-dimensional registration algorithm using mutual information and thin plate spline warping for the prostate MR images. First, automatic rigid body registration was used to capture the global transformation. Features included a multiresolution approach, two similarity measures, and automatic restarting to avoid local minimums. Second, local non-rigid

registration was applied. Interactively placed control points were automatically optimized by maximizing the mutual information of corresponding voxels in small volumes of interest and by using a three-dimensional thin plate spline to express the deformation throughout the image volume. More than 100 registration experiments with 17 MR volume pairs determined the quality of registration under conditions simulating potential interventional MRI-guided treatments of prostate cancer. Evaluations included visual inspection; voxel gray value measures such as mutual information, correlation coefficient, and intensity difference; and displacement of the centroids of segmented prostates. For image pairs that stress rigid body registration (e.g., supine, the diagnostic position, versus legs raised, the treatment position), both visual and numerical evaluation methods showed that warping consistently worked better than rigid body. Non-rigid registration rectified the misalignment in the pelvis following rigid body registration. The prostate centroid displacement for a typical volume pair was reduced from 3.4 mm to 0.6 mm when warping was added. Experiments showed that  $\approx 180$  strategically placed control points were sufficiently expressive to capture important features of the deformation. When only 120 control points were used, warping throughout the pelvis was visually less satisfactory but the prostate was aligned reasonably well. For volume pairs with images acquired in the same position (diagnosis-diagnosis) and comparable conditions, the rigid body method worked sufficiently well, and the prostate centroid displacements were <1.0 mm. In conclusion, the non-rigid registration method works better than rigid body registration whenever patient position or condition is greatly changed between acquisitions. It is very computational efficient for hundreds of control points and can very well approximate the deformation of the pelvis and internal organs. It will probably be a useful tool for many applications.

# Questions

- 1. Describe the concept of image registration.
- 2. Describe the steps for voxel-based image registration.
- 3. Given two images A and B, describe how to compute their mutual information value.

- 4. Describe three visual inspection methods for the evaluation of prostate registration.
- 5. Describe the quantitative evaluation methods for the prostate and pelvic registration.
- 6. Given two sets of corresponding control points, describe how to compute the thin plate transformation.

# Bibliography

- [1] Fei, B. W., Wheaton, A., Lee, Z., Duerk, J. L., and Wilson, D. L., Automatic MR volume registration and its evaluation for the pelvis and prostate, Physics in Medicine and Biology, Vol. 47, pp. 823–838, 2002, (http://www.iop.org/EJ/journal/PMB/).
- [2] Fei, B. W., Lee, Z., Duerk, J. L., Lewin, J.S., Sodee, D. B., and Wilson, D. L., Registration and fusion of SPECT, high resolution MRI, and interventional MRI for thermal ablation of the prostate cancer, IEEE Transactions on Nuclear Science, Vol. 51, No. 1, pp. 177–183, 2004.
- [3] Fei, B. W., Duerk, J. L., and Wilson, D. L., Automatic 3D registration for interventional MRI-guided treatment of prostate cancer, Computer Aided Surgery, Vol. 7, No. 5, pp. 257–267, 2002.
- [4] Fei, B. W., Duerk, J. L., Boll, D. T., Lewin, J. S., and Wilson, D. L., Slice to volume registration and its potential application to interventional MRI guided radiofrequency thermal ablation of prostate cancer, IEEE Transactions on Medical Imaging, Vol. 22, No. 4, pp. 515–525, 2003.
- [5] Fei, B. W., Wheaton, A., Lee, Z., Nagano, K., Duerk, J. L., and Wilson, D. L., Robust registration algorithm for interventional MRI guidance for thermal ablation of prostate cancer, Ki Mun, S., In: Proceedings of SPIE Medical Imaging on Visualization, Display, and Image-Guided Procedures, Vol. 4319, pp. 53–60, Bellingham, WA, SPIE. 2001.
- [6] Gray, H., Anatomy, Descriptive and Surgical (The classic collector's edition), pp. 1010, Gramercy Books, New York, 1977.
- [7] vanHerk, M., Bruce, A., Kroes, A. P. G., Shouman, T., Touw, A., and Lebesque, J. V., Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration, International Journal of Radiation Oncology Biology Physics, Vol. 33, No. 5, pp. 1311–1320, 1995.
- [8] Tenhaken, R. K., Forman, J. D., Heimburger, D. K., Gerhardsson, A., Mcshan, D. L., Pereztamayo, C., Schoeppel, S. L., and Lichter, A. S., Treatment planning issues related to prostate movement in response to

differential filling of the rectum and bladder, International Journal of Radiation Oncology Biology Physics, Vol. 20, No. 6, pp. 1317–1324, 1991.

- [9] Hill, D. L. G., Batchelor, P. G., Holden, M., and Hawkes, D. J., Medical image registration, Physics in Medicine and Biology, Vol. 46, No. 3, pp. 1–45, 2001.
- [10] Hamilton, R. J., Blend, M. J., Pelizzari, C. A., Milliken, B. D., and Vijayakumar, S., Using vascular structure for CT-SPECT registration in the pelvis, Journal of Nuclear Medicine, Vol. 40, No. 2, pp. 347–351, 1999.
- [11] Liehn, J. C., Loboguerrero, A., Perault, X., and Demange, L., Superimposition of computed-tomography and single photon emission tomography immunoscinti-graphic images in the pelvis—validation in patients with colorectal or ovarian carcinoma recurrence, European Journal of Nuclear Medicine, Vol. 19, No. 3, pp. 186–194, 1992.
- [12] Balter, J. M., Sandler, H. M., Lam, K., Bree, R. L., Lichter, A. S., and Ten Haken, R. K., Measurement of prostate movement over the course of routine radiotherapy using implanted markers, International Journal of Radiation Oncology Biology Physics, Vol. 31, No. 1, pp. 113–118, 1995.
- [13] Narayana, V., Roberson, P. L., Winfield, R. J., and McLaughlin, P. W., Impact of ultrasound and computed tomography prostate volume registration on evaluation of permanent prostate implants, International Journal of Radiation Oncology Biology Physics, Vol. 39, No. 2, pp. 341–346, 1997.
- [14] Herk, M. V., de Munck, J. C., Lebesque, J. V., Muller, S., Rasch, C., and Touw, A., Automatic registration of pelvic computed tomography data and magnetic resonance scans including a full circle method for quantitative accuracy evaluation, Medical Physics, Vol. 25, No. 10, pp. 2054–2067, 1998.
- [15] Antolak, J. A., Rosen, I. I., Childress, C. H., Zagars, G. K., and Pollack, A., Prostate target volume variations during a course of radiotherapy, International Journal of Radiation Oncology Biology Physics, Vol. 42, No. 3, pp. 661–672, 1998.

- [16] Remeijer, P., Geerlof, E., Ploeger, L., Gilhuijs, K., van Herk, M., and Lebesque, J. V., 3-D portal image analysis in clinical practice: An evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients, International Journal of Radiation Oncology Biology Physics, Vol. 46, No. 5, pp. 1281–1290, 2000.
- [17] Roeske, J. C., Forman, J. D., Mesina, C. F., He, T., Pelizzari, C. A., Fontenla, E., Vijayakumar, S., and Chen, G. T. Y., Evaluation of changes in the size and location of the prostate, seminal vesicles, bladder, and rectum during a course of external beam radiation therapy, International Journal of Radiation Oncology Biology Physics, Vol. 33, No. 5, pp. 1321–1329, 1995.
- [18] Scott, A. M., Macapinlac, H. A., Divgi, C. R., Zhang, J. J., Kalaigian, H., Pentlow, K., Hilton, S., Graham, M. C., Sgouros, G., Pelizzari, C., Chen, G., Schlom, J., Goldsmith, S. J., and Larson, S. M., Clinical validation of SPECT and CT/MRI image registration in radiolabeled monoclonal-antibody studies of colorectal carcinoma, Journal of Nuclear Medicine, Vol. 35, No. 12, pp. 1976–1984, 1994.
- [19] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, 1997.
- [20] Carrillo, A., Duerk, J. L., Lewin, J. S., and Wilson, D. L., Semiautomatic 3-D image registration as applied to interventional MRI liver cancer treatment, IEEE Transactions on Medical Imaging, Vol. 19, No. 3, pp. 175–185, 2000.
- [21] Wang, M. Y., Maurer, C. R., Fitzpatrick, J. M., and Maciunas, R. J., An automatic technique for finding and localizing externally attached markers in CT and MR volume images of the head, IEEE Transactions on Biomedical Engineering, Vol. 43, No. 6, pp. 627–637, 1996.
- [22] Maurer, C. R., Fitzpatrick, J. M., Wang, M. Y., Galloway, R. L., Maciunas, R. J., and Allen, G. S., Registration of head volume images using implantable fiducial markers, IEEE Transactions on Medical Imaging, Vol. 16, No. 4, pp. 447–462, 1997.

- [23] Wilson, D. L., Carrillo, A., Zheng, L., Genc, A., Duerk, J. L., and Lewin, J. S., Evaluation of 3D image registration as applied to MR-guided thermal treatment of liver cancer, Journal of Magnetic Resonance Imaging, Vol. 8, No. 1, pp. 77–84, 1998.
- [24] Fitzpatrick, J. M., Hill, D. L. G., Shyr, Y., West, J., Studholme, C., and Maurer, C. R., Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain, IEEE Transactions on Medical Imaging, Vol. 17, No. 4, pp. 571–585, 1998.
- [25] Peters, A. R., Muller, S. H., de Munck, J. C., and van Herk, M., The accuracy of image registration for the brain and the nasopharynx using external anatomical landmarks, Physics in Medicine and Biology, Vol. 45, No. 8, pp. 2403–2416, 2000.
- [26] Press, W. H., Teukolsky, S. A., Vellerling, W. T., and Flannery, B. P., Numerical Recipes in C: The Art of Scientific Computing, 2nd edn, Vol. 2, p. 636., The Press Syndicate of the Cambridge University, New York, 1992.
- [27] Nelder, J. and Mead, R. A., A simplex method for function minimization, Computer Journal, Vol. 7, pp. 308–313, 1965.
- [28] Kagawa, K., Lee, W. R., Schultheiss, T. E., Hunt, M. A., Shaer, A. H., and Hanks, G. E., Initial clinical assessment of CT-MRI image fusion software in localization of the prostate for 3D conformal radiation therapy, International Journal of Radiation Oncology Biology Physics, Vol. 38, No. 2, pp. 319–325, 1997.
- [29] Freeborough, P. A., Woods, R. P., and Fox, N. C., Accurate registration of serial 3D MR brain images and its application to visualizing change in neurodegenerative disorders, Journal of Computer Assisted Tomography, Vol. 20, No. 6, pp. 1012–1022, 1996.
- [30] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A., Interpolation artefacts in mutual information-based image registration, Computer Vision and Image Understanding, Vol. 77, No. 2, pp. 211–232, 2000.
- [31] West, J., Fitzpatrick, J. M., Wang, M. Y., Dawant, B. M., Maurer, C. R., Kessler, R. M., Maciunas, R. J., Barillot, C., Lemoine, D., Collignon, A.,

Maes, F., Suetens, P., Vandermeulen, D., vandenElsen, P. A., Napel, S., Sumanaweera, T. S., Harkness, B., Hemler, P. F., Hill, D. L. G., Hawkes, D. J., Studholme, C., Maintz, J. B. A., Viergever, M. A., Malandain, G., Pennec, X., Noz, M. E., Maguire, G. Q., Pollack, M., Pelizzari, C. A., Robb, R. A., Hanson, D., and Woods, R. P., Comparison and evaluation of retrospective intermodality brain image registration techniques, Journal of Computer Assisted Tomography, Vol. 21, No. 4, pp. 554–566, 1997.

- [32] Malone, S., Crook, J. M., Kendal, W. S., and Szanto, J., Respiratoryinduced prostate motion: Quantification and characterization, International Journal of Radiation Oncology Biology Physics, Vol. 48, No. 1, pp. 105–109, 2000.
- [33] Lee, Z., Sodee, D. B., Duerk, J. L., Nelson, A. D., and Berridge, M. S., Automatic registration of SPECT-MRI in the pelvis, Journal of Nuclear Medicine, Vol. 41, No. 5, pp. 232–232, 2000.
- [34] Lee, Z., Sodee, D. B., Faulhaber, P. F., Lancaster, T. L., MacLennan, G. T., and Wilson, D. L., Comparison of SPECT and PET imaging for prostate cancer with histological correlation, Journal of Nuclear Medicine, Vol. 42, No. 5, pp. 1226–1226, 2001.
- [35] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A., Image registration by maximization of combined mutual information and gradient information, IEEE Transactions on Medical Imaging, Vol. 19, No. 8, pp. 809–814, 2000.
- [36] Lau, Y. H., Braun, M., and Hutton, B. F., Non-rigid image registration using a median-filtered coarse-to-fine displacement field and a symmetric correlation ratio, Physics in Medicine and Biology, Vol. 46, No. 4, pp. 1297–1319, 2001.
- [37] Studholme, C., Hill, D. L. G., and Hawkes, D. J., Automated threedimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Medical Physics, Vol. 24, No. 1, pp. 25–35, 1997.
- [38] Rohr, K., Stiehl, H. S., Sprengel, R., Buzug, T. M., Weese, J., and Kuhn, M. H., Landmark-based elastic registration using approximating

thin-plate splines, IEEE Transactions on Medical Imaging, Vol. 20, No. 6, pp. 526–534, 2001.

- [39] Maurer, C. R., Hill, D. L. G., Martin, A. J., Liu, H. Y., McCue, M., Rueckert, D., Lloret, D., Hall, W. A., Maxwell, R. E., Hawkes, D. J., and Truwit, C. L., Investigation of intraoperative brain deformation using a 1.5-t interventional MR system: Preliminary results, IEEE Transactions on Medical Imaging, Vol. 17, No. 5, pp. 817–825, 1998.
- [40] Davis, M. H., Khotanzad, A., Flamig, D. P., and Harms, S. E., A physicsbased coordinate transformation for 3-D image matching, IEEE Transactions on Medical Imaging, Vol. 16, No. 3, pp. 317–328, 1997.
- [41] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., and Hawkes, D. J., Nonrigid registration using free-form deformations: Application to breast MR images, IEEE Transactions on Medical Imaging, Vol. 18, No. 8, pp. 712–721, 1999.
- [42] Denton, E. R. E., Sonoda, L. I., Rueckert, D., Rankin, S. C., Hayes, C., Leach, M. O., Hill, D. L. G., and Hawkes, D. J., Comparison and evaluation of rigid, affine, and nonrigid registration of breast MR images, Journal of Computer Assisted Tomography, Vol. 23, No. 5, pp. 800–805, 1999.
- [43] Meyer, C. R., Boes, J. L., Kim, B., Bland, P. H., Zasadny, K. R., Kison, P. V., Koral, 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, Vol. 1, No. 3, pp. 195–206, 1996.
- [44] Krucker, J. F., Meyer, C. R., LeCarpentier, G. L., Fowlkes, J. B., and Carson, P. L., 3D spatial compounding of ultrasound images using image-based nonrigid registration, Ultrasound in Medicine and Biology, Vol. 26, No. 9, pp. 1475–1488, 2000.
- [45] Kimiaei, S., Noz, M., Jonsson, E., Crafoord, J., and Maguire, G. Q., Evaluation of polynomial image deformation using anatomical landmarks for matching of 3D-abdominal MR-images and for atlas construction, IEEE Transactions on Nuclear Science, Vol. 46, No. 4, pp. 1110–1113, 1999.

- [46] Kim, B., Boes, J. L., Frey, K. A., and Meyer, C. R., Mutual information for automated unwarping of rat brain autoradiographs, Neuroimage, Vol. 5, No. 1, pp. 31–40, 1997.
- [47] Bharatha, A., Hirose, M., Hata, N., Warfield, S. K., Ferrant, M., and Zou, K. H., Three-dimensional finite element-based deformable registration of pre- and intraoperative prostate imaging, Radiology, Vol. 221, pp. 224–224, 2001.
- [48] Goshtasby, A., Registration Of Images With Geometric Distortions, IEEE Transactions on Geoscience and Remote Sensing, Vol. 26, No. 1, pp. 60–64, 1988.
- [49] Bookstein, F. L., Thin-plate splines and the atlas problem for biomedical images, Lecture Notes in Computer Science, Vol. 511, pp. 326–342, 1991.
- [50] Lee, Z., Kemper, C., uzic, R. F., erridge, M. S., and ilson, D. L., Quantitative pulmonary imaging based on PET-CT co-registration with warping, Journal of Nuclear Medicine, Vol. 42, No. Suppl, pp. 10–11, 2001.
- [51] Thacker, N. A., Jackson, A., Moriarty, D., and Vokurka, E., Improved quality of re-sliced MR images using re-normalized sinc interpolation, Journal of Magnetic Resonance Imaging, Vol. 10, No. 4, pp. 582–588, 1999.

# **Chapter 4**

# Stereo and Temporal Retinal Image Registration by Mutual Information Maximization

Xiao-Hong Zhu,<sup>1</sup> Cheng-Chang Lu,<sup>2</sup> and Yang-Ming Zhu<sup>3</sup>

## 4.1 Introduction

Image registration is the process of determining a one-to-one mapping between the coordinates in one space and those in another such that points in the two spaces that correspond to the same point are mapped to each other. The mappings, which are also called transformation, are two dimensional (2D) for 2D spaces and three dimensional (3D) for 3D spaces. The simplest examples are the rigid-body transformations, which are transformations in which the distances among all points are preserved. The rigid-body transformations are typically used to compensate for different imaging orientations of rigid objects. Other possible transformations include affine where parallel lines remain parallel, projective that is from 3D to 2D, and warping which is nonlinear in general.

The most prominent application of image registration is multimodality medical image fusion. Different image modalities can provide different information about the imaged organs, and most of the time this information is complementary in nature. For example, computer tomography (CT) image reveals the anatomical structure of the organ, magnetic resonance (MR) image is capable of

<sup>&</sup>lt;sup>1</sup>X.H. Zhu, Zircoa, Inc., 31501 Solon Rd., Solon, OH 44139

<sup>&</sup>lt;sup>2</sup>C.C Lu, Dept of Computer Science, Kent State Univ., Kent, OH 44242

<sup>&</sup>lt;sup>3</sup>Y.M. Zhu, PET Engineering, Philips Medical Systems, Cleveland, OH 44143

depicting soft-tissue anatomy, and photon emission tomography (PET) and single photon emission computed tomography (SPECT) show the functional property of the organ under study. Before different types of images are fused, these images have to be registered. After registration, the skeletal structures and areas of contrast enhancement seen in CT images can be overlaid on MR images, and likewise, functional lesions detected with PET or SPECT can be viewed in the context of anatomy imaged with CT or MR. Image registration can also be applied to multiple data sets obtained with the same modality at different times for the purpose of quantitative comparison, which increases the precision of treatment monitoring with serial images.

Another major biomedical application of image registration is to retinal image matching. Retinal or fundus photographs are standard diagnostic tools in ophthalmology. In the follow-up of age-related macular degeneration, drusen deposits need to be tracked and compared (see Sbeh et al. [1], Rapantzikos et al. [2]). Screening of diabetic retinopathy can involve a follow-up over many years (see [3, 4]). To determine the progression of glaucoma, a series of opticnerve-head topographies are assessed and compared [5]. Serial photographs of the flow of fluorescein dye are also used to determine areas of ischemia, hemorrhaging, neovascularization, and occlusions in diseases such as diabetic retinopathy (see [5]). A noise reduction technique is reported for laser scanning ophthalmoscope using image registration [6]. Multimodality registration is also performed in retinal imaging. In glaucoma diagnosis, for example, the optic-nerve-head is assessed from color stereo images and the nerve fiber layer is assessed from red-free images [7]. In retinal analysis, two types of images, fluorescein images (angiographic images taken under ultra violet light after injection of fluorescein dye) and green images (taken under natural light with a green filter), are often used for the diagnosis of the gravity of diabetic retinopathy [3]. Physicians often use more than one image to identify a lesion and assess its seriousness, or base their diagnosis on detection of various image features in different modality images. To make this comparison and assessment objective, it is necessary that all images be registered.

The research of image registration has a relatively short history. Due to diverse applications, many registration algorithms have been developed from different perspectives. Brown summarized the research work before 1992, mostly for 2D-2D registration [8]. Since the most important and fruitful application of image registration is in medical imaging, several authors reviewed registration algorithms from a medical imaging perspective. Gerlot-Chiron and Bizais have presented a unified description of existing registration methods [9]. Maurer and Fitzpatrick later adopted a similar scheme when they reviewed the registration algorithms within the neurosurgery context [9]. Van den Elsen *et al.* reviewed and classified medical image registration algorithms [10]. Their classification criteria have been augmented and detailed by Maintz and Viergever recently [11]. In addition to various survey articles and book chapters (e.g. Fitzpatrick *et al.* [12]), a monograph on image registration has also been published [13]. For further elaboration, the reader is advised to refer to the original surveys, book chapters, and monograph.

The increasingly complex schemes for classification reflect the sheer amount of literature on image registration methodologies. It is impossible for us to detail these algorithms here. However, we do want to point out a recent trend in image registration research and practice, i.e., the voxel property-based registration methods have become increasingly popular. Compared to other registration algorithms, the voxel property-based methods using the full image content offer several advantages: they work on the image gray-value without any prior data reduction; they can be automated and the results are objective; they require no segmentation and involve little or no user interaction.

To this category, various paradigms have been reported, including crosscorrelation in spatial or (Fourier) transformed domain, minimization of variance of intensity ratios, minimization of variance of gray values within segments, histogram clustering and minimization of histogram dispersion, minimization of the joint histogram entropy of different images, and maximization of mutual information, among many others. Studholme *et al.* [14] compared five similaritybased algorithms and Fitzpatrick *et al.* [15] compared 16 of these algorithms. The reports from various independent groups confirm that the mutual information maximization approach to image registration is one of the most robust and has superior performance.

Mutual information image registration was independently proposed by Maes *et al.* [16] and Wells *et al.* [17]. However, their development is a natural consequence of early effect on the analysis of voxel value joint histogram (see [18]). Hill *et al.* [19] used third-order moments of the joint histogram as well as other measures to characterize the clustering of the joint histogram at registration. Collignon *et al.* [20, 21] used joint entropy as a criterion for registration, but reported that it had a small capture range, i.e., only when the initial

position of two images is sufficiently close to the optimal alignment can they be registered well. To increase the capture range, naturally the next step is to exploit mutual information. Due to its super performance, mutual information becomes the first choice for automatic image registration. The advancement on mutual information registration has been reviewed recently by Pluim *et al.* [22].

### 4.1.1 Review of Retinal Image Registration

Retinal image registration is the main focus of this chapter. This registration generally involves large x translation, due to changes between sittings and smaller y translation from changes in position of the chin cup. Rotation occurs due to tilting of the head and through ocular torsion, and scaling is caused by changes in the distance between the camera and the head, due to equipment changes or differing head positions (see [5, 23]). This section reviews some registration methods as applied to retinal images. By no means is this review complete. The interested readers may refer to [3, 5, 23] and references therein for more related work.

Peli *et al.* [24] reported on a correlation method that preprocesses the images using an adaptive threshold procedure to select vessel points. The normalized sum of differences is then calculated with these vessel points. Via an exhaustive search, this method produces pixel-level registration for x- and y-translation only. It is not robust toward large changes in image intensity and white noise. The absolute value of difference of pixel intensities was also used as a comparison measure for retinal image registration. The images can be processed twice, using optic discs as features for coarse alignment and the blood vessels as features for fine alignment [25].

If one image is a scaled, rotated, and translated version of another image, then the Fourier transform of that image is a scaled and rotated version of the Fourier transform of the other image. Thus, image registration can also be done in Fourier transformed space. Cideciyan *et al.* [26] computed the scaling and rotation differences of the Fourier transformed images by cross-correlation. These results are then used to transform one image in spatial domain. The final translation differences in spatial domain are then found via cross-correlation. When the images are taken at different times, where the translation difference and the image intensity difference may be large, this approach is problematic. This approach is not applicable to multimodality image registration. Hart *et al.* [27] exploited a blood vessel filter to detect the ends of blood vessel segments and then used them as control points to register retinal images. A specific scheme was designed to eliminate erroneous point pairs until either there are four control points left or the mean square error given by the least square fitting drops below five pixels. This method has some problems as pointed out by Ritter *et al.* [5].

Ritter *et al.* [5] recently applied the mutual information maximization to retinal image registration. To find a global optimum, simulated annealing is used in the multiresolution optimization. Although this method can successfully find the global optimum registration with translation, rotation, and scaling, it is time consuming. To assess the accuracy of registration, Ritter *et al.* compared the registration results against the solutions obtained by an exhaustive search. This comparison has an intrinsic drawback. What they studied is how the simulated annealing behaves which is an implementation artifact, not how the mutual information maximization behaves as a registration criterion.

Matsopoulos *et al.* [28] used matched filters (see [29]) to segment the vessel trees and registered the segmented trees automatically. To ensure that a global optimal registration is found, simulated annealing and genetic algorithms were employed. They also studied the suitability and efficiency of different image transformation models. The criterion used in the optimization is a correlation function defined on the segmented, binary images.

Zana *et al.* [3] reported on a multimodal retinal registration scheme based on vessels detection and Hough transform. The vascular tree is segmented first, and then the bifurcation points are detected. Those tree and points are features used to register the images. Although their algorithm is attractive, it involves a fair amount of user interaction in the preprocessing and in the final registration selection (the solution given by Bayesian selection is not necessarily the best). Laliberte *et al.* developed a similar technique that also used the blood vessel bifurcation points, but did not need the assumption of a Gaussian shape vessel intensity profile which is inappropriate for low resolution optical images (see [23]). In spite of about 10 adjustable parameters in the algorithm, it seems that the success rate for this latter method is low (36 out of 61 pairs). Can *et al.* [30] developed a hierarchical scheme to match the feature points in two images, using a progressively complex transformation model and a reduced set of matching points. This algorithm is attractive when one builds the retinal map since warping is generally required. This chapter reports on our development of an object-oriented software system for automatic retinal image registration by mutual information maximization. For maximum portability the software is written in Java which "is written once, but runs everywhere", using MVC (model-view-control) framework. We use the simplex downhill method (see [31]) as the optimization algorithm which is easy to implement, and is quick in practice. We demonstrate that this algorithm registers temporal and stereo retinal image pairs of four patients with a very high success rate (86%), a satisfactory registration accuracy compared to point matching results, and within a clinically acceptable time  $(12 \pm 3 \text{ sec.})$ .

# 4.2 Registration by Mutual Information Maximization

### 4.2.1 Mutual Information

For two random variables A and B, the mutual information is

$$I(A, B) = \sum_{A} \sum_{B} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_{A}(a)p_{B}(b)},$$

where  $p_{AB}(a, b)$  is the joint probability density function (pdf), and  $p_A(a)$  and  $p_B(b)$  are marginal pdfs, Gonzalez *et al.* [32]. I(A, B) is related to entropy [H(A), H(B)], conditional entropy [H(A|B), H(B|A)], and joint entropy [H(A, B)] by

$$I(A, B) = H(A) + H(B) - H(A, B)$$
  
=  $H(A) - H(A|B)$   
=  $H(B) - H(B|A).$ 

Mutual information measures the interdependence of two random variables. If two variables are independent, then their joint pdf is the product of their marginal pdfs, i.e.,  $p_{AB}(a, b) = p_A(a)p_B(b)$ . Substituting this into the definition of mutual information one gets zero. That is to say, the mutual information is minimal. On the other hand, if two random variables are related by a one-to-one mapping, the mutual information is maximal. In fact, in the latter case, H(A) = H(B).

Two images involved in registration are called reference and floating images. The floating image has undergone scaling, rotation, and translation to match the reference image. In the mutual information image registration context, we treat the voxel values a and b at corresponding points in two images that are to be registered as random variables A and B. Mutual information measures the interdependency between the reference image and floating image. We assume that the mutual information of random variables A and B has a maximum value when two images are registered (i.e., maximum interdependency). That is, the uncertainty of one image given another image is minimized, and we have more confidence in using one image to interpret another. Note that, the voxel values a and b are related by the registration transformation T. The mutual information registration states that the images are registered under transformation  $T^*$  for which I(A, B) is maximal.

If both marginal pdfs are independent of the registration parameters (i.e., no matter what T is, they would not change much), then mutual information maximization is reduced to minimization of joint entropy. If either pdf is independent of the registration parameters, which is the case when one image geometrically contains another image, maximization of mutual information is reduced to minimization of conditional entropy. However, if both images only partially overlap, which is very likely during the optimization as we will see later, the overlap will change as the transformation changes and both marginal entropies generally depend on the transformation. Mutual information takes overlapping explicitly into account.

Sometimes retinal images are not grayscale images. They have RGB color channels. One has three options to handle this multichannel registration. (1) Define the mutual information for each channel and maximize the sum of these mutual information values. (2) Convert the color images to luminance images and then register the converted images. (3) Pick a color channel and register this channel and other two channels are presumably registered once that selected channel has been registered. Since the green channel has the highest contrast, it can be registered first if one chooses the third option.

In Fig. 4.1 we illustrate the procedures involved in the mutual information maximization approach to image registration. In the following sections we discuss what kind of transformation we pose, how to compute the mutual information, how to update the transform and what is the criterion for optimal transform, etc.



Figure 4.1: Flow-chart of image registration by mutual information maximization.

Our implementation is based on Java, using a lot of Java 2D graphics and image processing capabilities. Java 2D API includes a set of classes that can be used to create high quality graphics. It has features like geometric transformation, alpha composition, and image processing (see [33, 34]).

## 4.2.2 Transformation

In image registration different transformations can be considered: rigid-body transformation, affine, and nonlinear warping, with increasingly technical difficulty. In this chapter we consider affine transformation with a scaling factor, rotation angle, and two translational offsets. It is reported that the affine transform is sufficient for most retinal registration [23].

In homogeneous coordinate systems, an isotropic scaling transformation matrix is

$$S = \begin{bmatrix} s & 0 & 0 \\ 0 & s & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

where s is the scaling factor. The rotation matrix is

$$R = \begin{bmatrix} \cos\theta & \sin\theta & 0\\ -\sin\theta & \cos\theta & 0\\ 0 & 0 & 1 \end{bmatrix},$$

where  $\theta$  is the rotation angle. The transformation matrix corresponding to the translation is

$$T = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix},$$

where  $t_x$  and  $t_y$  are translation offsets.

A successive transformation amounts to multiplication of corresponding matrices. We enforce the order of operations as scaling, rotation, and then translation, in matrix form,  $T \cdot R \cdot S$ . We seek  $(s, \theta, t_x, t_y)$  parameters and these parameters are applied to the floating image in the order discussed above.

An affine transform can be easily composed in Java:

```
AffineTransform at = new AffineTransform();
at.translate(tx, ty);
at.scale(scale, scale);
at.rotate(Math.PI/180.0*angle);
```

To create a transformed image, one can invoke the filter method of the AffineTransformOp, which can be constructed from the rendering hints and affine transform. The Java code is similar to

```
RenderingHints rh = new RenderingHints(/* specify here */);
AffineTransform at = new AffineTransform();
// define transform here
AffineTransformOp atop = new AffineTransformOp(at, rh);
BufferedImageOp biop = (BufferedImageOp) atop;
BufferedImage bi = biop.filter(bi, null);
```

For retinal image registration, the maximum horizontal translation  $t_x$  is 100 pixels, the vertical translation  $t_y$  is in the order of 10 pixels, the maximum rotation angle  $\theta$  is 5° and a typical scaling factor falls in [0.95, 1.05]. In the optimization process, we clamp the scaling factor to a certain range. A negative scaling factor is relatively impossible. A zero scaling is also impractical and Java 2D would generate a runtime exception.

## 4.2.3 Interpolation

After a transformation is applied, the grid point in the floating image will typically not coincide with another grid point in the transformed space. For a simple example, consider translation  $t_x = 0.25$ . Grids (0, 0), (1, 0) become (0.25, 0), and (1.25, 0). Since the pixel values of the reference image are known on grid, and under certain transformations, only the pixel value of the transformed floating image on grid are of interest, the pixel values of the transformed floating image on grid have to be estimated. The technique is called interpolation, which is to estimate a pixel value based on the pixel values of its surrounding pixels.

There are different kinds of interpolation methods: nearest neighbor (0-order), bilinear (first-order), etc. Tsao evaluated the interpolation effects on registration performance [31]. Nearest neighbor uses the pixel value at the nearest grid pixel to approximate the pixel value at the new position. That is, the pixel value at (x, y) is approximated by that at ((int)(x + 0.5), (int)(y + 0.5)). The nearest neighbor interpolation is insensitive to the magnitude of translation up to one pixel. Therefore, it is not sufficient in order to achieve subpixel registration accuracy.

Bilinear interpolation assumes the pixel value in each x and y direction changes in a linear fashion. To get the pixel value at point (x, y), one can (a) interpolate the value at (x, j) from (i, j) and (i + 1, j), (b) interpolate the value at (x, j + 1) from (i, j + 1) and (i + 1, j + 1), and (c) interpolate the value at (x, y) from (x, j) and (x, j + 1), where integer i and j satisfy  $i \le x < i + 1$  and  $j \le y < j + 1$ .

In Java 2D, the interpolation method is governed by the rendering hints. To specify a nearest neighbor interpolation, the rendering hints can be defined by

Similarly, to use a bilinear interpolation, the rendering hints can be specified as

RenderingHints rh=new RenderingHints ( RenderingHints.KEY\_INTERPOLATION, RenderingHints.VALUE\_INTERPOLATION\_BILINEAR);

Both interpolation methods were used. If a nearest neighbor interpolation is used, the optimization process tends to be quicker, but the success rate of registration tends to be lower. Thus, all results reported in this chapter are obtained using bilinear interpolation.

## 4.2.4 Estimation of Joint and Marginal Distributions

To calculate the mutual information, one has to have the marginal and joint pdfs. Since they are unknown, they have to be estimated from the data under study. We estimate these pdfs from their corresponding histograms.

To compute the histograms, the pixel pairs at the same grid are binned. Assume the bin size is  $B_r \cdot B_f$ , where  $B_r$  and  $B_f$  are the bin size for the pixels in reference image and floating image, respectively.

A fixed bin size (32) is used. We will not discuss bin size effect since our preliminary results indicate that there are no apparent effects on the registration performance. For the byte image as the retinal GIF files normally are, there are 256 values in each color channel. However, when images are digitized and then translated from color to grayscale, there is some error. The image does not have very bright and very dark regions. Moreover, the lossy JPEG image compression is frequently used. These facts lead to a sparsely filled histogram if a  $256 \times 256$  bin is used. Therefore, before we bin the pixel pairs their values are processed and then binned.

Assume the maximum pixel value is  $I_{\text{max}}$  and the minimum pixel value is  $I_{\text{min}}$ . Since we only have B bins, each bin needs to accommodate step =  $\lceil (I_{\text{max}} - I_{\text{min}})/B \rceil$  different pixel values. Given a pixel value v, obviously it will be binned to bin  $(v - I_{\text{min}})$ /step. Notice that we did not distinguish reference and floating images. The formula and schemes are applicable to both images. Also notice that interpolation never increases or decreases the max and min pixel values (except it may introduce zero-valued pixels due to transformation). This assertion is clear for the nearest neighbor interpolation. It is also true for the bilinear interpolation since, based on the formula, the new pixel value is a convex combination of old ones. This property is used in the implementation and steps are only calculated once and never changed thereafter.

The overlapping pixels are scanned and the joint histogram is then computed. Let the joint histogram be H(i, j). The marginal histograms are then

$$H_r(i) = \sum_j H(i, j),$$
$$H_f(j) = \sum_i H(i, j).$$

Here the subscripts r and f stand for reference and floating, respectively. Let N be the total number of pixel pairs examined in the overlapping region then the joint and marginal pdfs are estimated as

$$p(i, j) = \frac{H(i, j)}{N},$$
$$p_r(i) = \frac{H_r(i)}{N},$$
$$p_f(j) = \frac{H_f(j)}{N}.$$

Substituting those estimated joint and marginal pdfs into the definition of mutual information, its value can be readily computed. Since the iterative optimization routine is for minimization problem, the negated mutual information is actually computed.

### 4.2.5 Optimization

The sole goal of mutual information approach to image registration is to find a transformation under which the mutual information between the reference image and the transformed floating image is maximized. This is a typical optimization problem and many optimization methods have been employed, including exhaustive search, gradient descent, simplex downhill, Powell's method, simulated annealing, and genetic algorithms [36]. This chapter studies the simplex downhill optimization since it is reported that it is faster than other algorithms with similar accuracy [36]. Simplex is also the easiest method to understand and does not require any derivative calculation.

The simplex downhill method may be trapped at local minimum. Simulated annealing method can avoid the problem, but it is computationally expensive. To avoid that we incorporate multiresolution strategies into the iterative process.

#### 4.2.5.1 Exhaustive Search

The exhaustive search can be used to find the true global optimum solution. This brute-force search is very expensive. If we want 0.1 pixels, 0.1 rotations, and 0.001 scaling accuracy, we would have  $10^{10}$  iterations.

### 4.2.5.2 Simplex

The downhill simplex method implements an entirely self-contained strategy and does not make use of any 1D optimization algorithm [31]. It requires only function evaluation, not derivatives.

A simplex is a geometrical figure, consisting of N + 1 vertices in N dimensions and all their interconnecting line segments, polygonal faces, etc. In the optimization process, a simplex reflexes, expands, and contracts, around a vertex, trying to enclose the optimum point within its interior.

The downhill simplex method must be started with N + 1 points, defining an initial simplex. For our image registration problem, there are four unknown registration parameters. Given an initial guess of the registration parameters vector ( $t_x$ ,  $t_y$ ,  $\theta$ , s), a non-degenerate simplex can be formed by points ( $t_x + 1$ ,  $t_y$ ,  $\theta$ , s), ( $t_x$ ,  $t_y + 1$ ,  $\theta$ , s), ( $t_x$ ,  $t_y$ ,  $\theta + 1$ , s), and ( $t_x$ ,  $t_y$ ,  $\theta$ , s + 0.1). In our implementation, the initial registration parameter vector is (0, 0, 0, 1), i.e., all translation offsets are 0, rotation angle 0, and the scaling factor 1.

Note that we use different guesses for the problem's characteristic length scale in different directions since we don't expect that the optimized scaling factor will be significantly deviated from 1. We use degrees rather than radians such that the characteristic length scale of rotation angle is comparable with those for translations (1 radian amounts to  $57.3^{\circ}$ ). In Java a zero scaling factor will generate an exception. In reality, a non-positive scaling factor does not make any sense. From the characteristic of the retinal image registration problem, one knows that the scaling factor falls in the range of 0.95–1.05. The simplex optimization routine cannot guarantee that. Thus, at each iteration the proper range of the scaling factor is checked. A relaxed lower and upper bounds is used. We clamp the scaling factor at a lower bound of 0.9 and an upper bound of 1.10.

The termination condition in any multidimensional optimization routine can be delicate. One can terminate the iterative process when the vector distance moved in a step is fractionally smaller in magnitude than some preset tolerance. Alternatively, one can require that the decrease in the function value in the terminating step be fractionally smaller than a preset tolerance. In the simplex optimization, the relative difference between the highest and lowest vertices is compared against a tolerance (0.0001).

## 4.2.5.3 Multiresolution Strategies

The above optimization process can be attracted to a local minimum (thus local maximum of mutual information). To avoid that multiresolution or subsampling optimization scheme can be used. The idea of multiresolution strategy is simple: Find the optimal registration on coarse images first. Then using the found solution as the starting point, find the optimal registration for the fine images. The coarse images are derived from the original images, by averaging several pixels in the original images. As a variation of this multiresolution, subsampling is also used sometimes. Instead of taking the average, a single pixel is picked to form a coarse image are subsampled periodically, with a gradually increasing sampling frequency. The optimization result at the lower sampling frequency is used as the starting point for the higher sampling. Since the images are subsampled, some information is lost.

We propose a multiresolution subsampling scheme. The idea is to combine the advantages of multiresolution and subsampling. Our experimental results indicate that this scheme can increase the convergence speed considerably. The pseudo code for this scheme follows:

```
For resolution r1 > r2 > ... > rn
    Rescale the translation offsets (divided by the
    folding number)
    Prepare the coarse images
    For sampling s1 > s2 ... > sm
        Do simplex optimization
Rescale translation offsets to correct different
        resolution effect
```

In the implementation, all these resolutions r and sampling frequencies s can be adjusted. In the experimental results reported below, the resolutions are 27, 9, 3, 1 and the sampling frequencies are 3 and 1.

The simplex optimization in the inner loop is identical to the one described earlier, except the images are scaled down by some folding factor which is equal to the resolution. To carry the registration parameters to different resolution, the parameters are also scaled down before entering the optimization loop and scaled up after leaving the loop. Only translation offsets need scaling. When the sampling is not 1, s - 1 pixels are skipped over when the pixel pairs are binned. In the inner simplex optimization, the "reference image" and "original floating" image should be understood as the down-scaled reference image and down-folded floating image at each resolution.

# 4.2.6 Object-Oriented Software Implementation and Architecture

The MVC framework is adopted to facilitate the software design [37]. Models represent the information, views present information, and controllers interpret user manipulation. Since the algorithm is automatic and there are not many user interactions, the control and view can be combined, which is similar to the document-view pattern used in Microsoft Foundation Class Doc-View Framework.

In our case the model has the information on the two images to be registered, i.e., the reference image and the floating image and the current registration parameters. Both images are stored as Buffered Image. The view object gets two images as well as the registration parameters from this model and displays the images properly. The registration object also gets two images from the model object and sets the registration parameters after a solution is found.

The iterative optimization is implemented with two classes. Simplex is an abstract base class and implements the logic of simplex optimization. The method to calculate the objective function (mutual information in this case) is not implemented in the class and thus is abstract. Its subclass should implement the objective function calculation method. The subclass, MIMax, extends the Simplex abstract class and provides all necessary methods to compute the mutual information.

# 4.3 Registration Results and Discussion

## 4.3.1 Description of the Image Files

The retinal images used in this study were provided by Dr. Nicola Ritter of the Lions Eye Institute of Perth and the Glaucoma Foundation of Perth. The images were taken using a Nidek stereo fundus camera. The original images were taken with color slide film and then digitized with a Polaroid slide digitizer. The images used for registration had 256-gray levels and comprise about 5% of the total surface of the retina centered on the optic nerve head.

Following is a brief description of the image files. All image files are labeled based on the patient identification, l/r, l/r, and a number. The first "l" or "r" indicates that those are the images of the patient's left or right eye. The second "l" or "r" designates the left or right stereo images. These images were taken simultaneously onto a single photographic slide and were separated into two images manually after digitization. The last digits, 5, 6, or 7, indicate whether the images were taken in 1995, 1996, or 1997.

*Image set 1*. There are four images in this patient file: ll5, ll7, lr5, and lr7. They are the same images as in Fig. 4.2 [5]. These images of left eyes were taken in 1995 and 1997, 18 months apart, during which time the patient had an operation to relieve pressure related to glaucoma. They are displayed in Fig. 4.2 here. This image set gives us two stereo image pairs (ll5/lr5, ll7/lr7) and two temporal pairs (ll5/lr7, lr5/lr7).

*Image set 2 (B853)*. There are six images in this set: ll6, ll7, lr6, lr7, rl7 and rr7, which gives us three stereo pairs (ll6/lr6, ll7/lr7, and rl7/rr7) and two temporal pairs (ll6/ll7 and lr6/lr7).

*Image set 3 (H3397)*. There are four images in this set: ll6, ll7, lr6, and lr7, which gives us two stereo pairs (ll6/lr6 and ll7/lr7) and two temporal pairs (ll6/ll7 and lr6/lr7).

*Image set 4 (P374)*. There are eight images in this set: ll5, ll6, lr5, lr6, rl5, rl6, rr5, and rr6, which gives us four stereo pairs (ll5/lr5, ll6/lr6, rl5/rr5, and rl6/rr6) and four temporal pairs (ll5/ll6, lr5/lr6, rl5/rl6, and rr5/rr6).

All together we have 11 stereo image pairs and 10 temporal image pairs, which is a subset of the images used by Ritter *et al.* [5].

### 4.3.2 Mutual Information as a Measure

In this section we study mutual information as a measure for retinal image registration. Look at an extreme case first: an image registers to itself. At registration, the pixel values in two images have an exact one-to-one (identical) relation. In their joint histogram, there would be a straight, diagonal line. For any two



(a)



(b)

Figure 4.2: Typical temporal and stereo retinal images. (a) ll5, (b) lr5, (c) ll7, and (d) lr7 in image set 1.

images, we would not expect to see any structure in the joint histogram when they are not registered.

For two images of the same modality, one would expect a similar situation. However, due to noise, change of imaging condition, or change of the imaged


(c)



<sup>(</sup>d) Figure 4.2: (*Cont.*)

scene itself, the pixel values for the same point at registration would not be identical. It is safe to expect that the majority of them will follow some simple and most likely linear relation. When the images are not registered, it is difficult to justify the existence of a simple structure in the joint histogram.

For two images of different modality, one would expect the pixel values at registration to still be related since they are the images of the same scene.



Figure 4.3: Joint histogram of images (c) and (d) in Fig. 4.2. (a) Before registration; (b) after registration by mutual information maximization (registration parameters: x and y translations are -4.62 and -2.21 pixels, rotation angle is  $0.50^{\circ}$ , and scaling factor is 0.9979).

However, the relationship becomes convoluted. Again, the joint histogram at registration is more structured than at misregistration.

As a demonstration of this joint histogram aggregation concept, we display some joint histograms at registration and at misregistration in the following figures. Figure 4.3 depicts the joint histograms of images (c) and (d) (stereo registration) in Fig. 4.2 before and after registration. The registration parameters are shown in the figure caption. The horizontal direction corresponds to the pixel values in image (c) and the vertical direction corresponds to the pixel values in image (d). The histogram size is  $256 \times 256$ . Since the minimum pixel value is not 0 and the maximum pixel value is not 255, the low end and the high end are virtually black, i.e., there are no pixel value pair aggregations there. Along the diagonal, Fig. 4.3b is brighter than Fig. 4.3a, indicating the histogram in (b) is more aggregated.

Figure 4.4 shows the joint histograms of images (b) and (d) (temporal registration) in Fig. 4.2 before and after registration. Again we put the registration



(a)



(b)

Figure 4.4: Joint histogram of images (b) and (d) in Fig. 4.2. (a) Before registration; (b) after registration by mutual information maximization (registration parameters: x and y translations are 65.00 and -4.56 pixels, rotation angle is  $0.31^{\circ}$ , and scaling factor is 0.9858).

parameters in the figure caption. The histogram after registration (Fig. 4.4b) is more aggregated than that before registration.

Those joint histograms were normalized before being displayed. The maximum histogram value may be larger or smaller than 255, depending on the image contents and the image size. The maximum histogram value is always normalized to 255 with a nonlinear transformation. The nonlinear transformation is relatively simple. First a linear transformation  $(x/\max)$  is used to normalize the histogram such that the maximum histogram value is 1.0. A nonlinear transform  $x^{0.25}$  is used to change the histogram such that the small values are enhanced. The resultant histogram values are multiplied by 255 afterwards. Alternatively one can use a logarithmic operation to rescale the dynamic range of the histogram, as employed to display the Fourier transform of an image.

The joint histogram aggregation at registration can be studied and characterized by entropy. Entropy is a measure of randomness. A higher disordered system has larger entropy. If the histogram is well structured, then the entropy of the joint pdf of the pixel values has smaller entropy. In fact, entropy minimization was exploited as a measure for image registration. However, it is too sensitive to the overlapping size of two images. To overcome that mutual information is now used instead. In practice, mutual information maximization proves to be a robust measure for image registration.

For retinal image registration there are four registration parameters. It is difficult to visualize how the mutual information as a registration measure behaves. We use mutual information maximization to register images (b) and (d) (temporal registration) and calculate the mutual information in the vicinity of the optimized solution. For this particular image pair the registration parameters are: x translation of 65.00 pixels, y translation of -4.56 pixels, rotation angle of  $0.31^{\circ}$  and scaling factor of 0.9858. To visualize the mutual information surface in the hyperspace, we fix three registration parameters and change the other one in the neighborhood of the optimal.

Figure 4.5 shows the mutual information values in the vicinity of the optimal x translation while other three registration parameters are fixed. Figure 4.6 shows the results when the y translation is varied around the optimal value. The rotation angle dependent behavior is illustrated in Fig. 4.7. Figure 4.8 displays how the mutual information changes when the scaling factor varies. It can be seen that the mutual information indeed has a maximum value around a good registration. Note that the mutual information is very sensitive



Figure 4.5: Mutual information as a function of x-translation offset from a good registration.



Figure 4.6: Mutual information as a function of y-translation offset from a good registration.



Figure 4.7: Mutual information as a function of rotation angle offset from a good registration.



Figure 4.8: Mutual information as a function of scaling offset from a good registration.

to the change in the rotation angle as revealed by the sharp peak in Fig. 4.7. This is expected since a slight change in rotation angle can amount to a very large offset in pixel positions if the distance from the rotation center is large. It is very important to have an accurate estimation of the rotation angle in registration.

### 4.3.3 Success Rate, Speed, and Accuracy

Our implementation of mutual information maximization for retinal image registration works relatively well. Figure 4.9 shows the registration results of Figs. 4.2a and 4.2b (temporal registration). The registration images can be displayed side by side (not shown here). Figure 4.9a shows the two registered images in a checkerboard format, where the size of each checkerboard can be adjusted. Figure 4.9b shows these two images in a moving curtain format, where the vertical line can be moved left or right to check the continuity of the image features (vessels) across the dividing lines. The lines can also be horizontal. To the left of the line is the reference image. The part of the reference image to the right of the line is clipped out by the line and not displayed. To the right of the line is the registered floating image. The left side of the matched floating image is also clipped out by the line and is not displayed. Figure 4.9c shows these two images in an overlay format, where the alpha is 0.5 so one can see one image through the other image. The maximum, minimum, average, absolute difference, color composition, and other fusion methods are also implemented in our software. The color composition selectively extracts the color channels and assigns them to the composite image which is very useful when one inspects the registration and presents the fusion results of color images (for example, red-free and angiograph).

In this section we compare our results against Ritter *et al.* [5] in terms of success rate, registration speed, and registration accuracy. They used the simulated annealing as the optimization routine. Their program was written in C and their results were obtained on a Pentium Pro 200 running Linux. Our program was written in Java (JDK 1.4.0) and the results were obtained on a Pentium IV running Windows XP. We also ran the program on a Pentium 233 running Windows 95. Their success rate is 100%. On average it takes 1240 seconds. In their implementation they used the nearest neighbor interpolation first and then bilinear interpolation in the last layer of iteration. As one would expect, if



(a)



(b)



Figure 4.9: Two registered retinal images displayed in (a) checkerboard, (b) moving curtain, and (c) overlay formats.

	Ritter et al.	Our implementations	
Operating system	Linux	Windows XP	
CPU	Pentium Pro 200	Pentium IV, 2.4 GHz	
Language	С	Java (JDK 1.4.0)	
Interpolation	Nearest neighbor + bilinear	Bilinear only	
Optimization	Simulated annealing	Simplex	
Multiresolution	Yes	Yes	
Subsampling	No	Yes	

Table 4.1: Comparison of Ritten *et al.* and our implementations of mutual information maximization for retinal image registration

bilinear interpolation were used in all iterations, it would take longer. Table 4.1 summarizes the differences.

Ritten *et al.* reported the root mean square error (RMSE) of the registration parameters. For the stereo registration, they are 0.59, 0.52, 0.31, 0.0058 for x and y translations, rotation angle, and the scaling factor, respectively; for the temporal registration, they are 0.29, 0.29, 0.15, 0.0038, respectively.

For this kind of registration the ground truth is unknown. Therefore, it is hard to evaluate its accuracy. Ritter *et al.* compared their results against the results obtained from the exhaustive search in a limited area [5]. What they assessed is not the mutual information maximization as a registration method, but the simulated annealing as a global optimization scheme. It is well known that the simulated annealing has a large chance to find the global optimal solution. Rather than comparing the results against those of an exhaustive search, we compare our results against the point matching results. In CT/MR/PET/SPECT image registration, point matching is regarded as the gold standard [15], as long as one can identify the corresponding points correctly.

It is possible to identify the corresponding points in retinal images if there are enough vessel trees. Our images are about 5% of the total surface of the retina centered on the optic nerve head. Some retinas have severe degradation. Some images are blurred. The vessels are usually very thick. Thus the corresponding points are not always easy to identify.

The point matching results of temporal image pairs and stereo image pairs were obtained as the average of 4 or 5 independent registrations performed by two operators. As we mentioned earlier, it is difficult to pair the corresponding points in some image pairs. Thus we have less confidence in their results. For example, in temporal registration, it is difficult to register the images in Set 1. In general, the stereo images are difficult to register by point matching. Nevertheless, the average registrations seem reasonable.

### 4.3.3.1 Registration of Temporal Images

Visual inspection of the registration results reveals that mutual information maximization succeeded 9 times out of 10 and it failed on lr6/lr7 of set 3. Thus the success rate is 90%. The average registration time was 12.5 seconds with a standard deviation of 3.3.

We can compute the speedup of our implementation against Ritter *et al.* [5]. Considering the CPU difference and ignoring all other differences, we define the speedup as

$$speedup = \frac{RittersTime \cdot TheirCPUClockRate}{OurTime \cdot OurCPUClockRate} - 1$$

Substituting all numbers in, our speedup is 7.27. It is worth pointing out that Pentium Pro and Pentium IV processors have different architectures, thus, this comparison only has indicative meaning.

Define the misregistration as mutual information registration—"true" registration. The mean and standard deviation of these temporal misregistration parameters are  $(-0.18 \pm 1.17, 0.33 \pm 0.56, 0.01 \pm 0.15, 0.0000 \pm 0.0040)$ . Overall, the estimation of the rotation angle and the scaling factors are very accurate. The large errors in the translation offsets are primarily due to those errors related to the first image set.

We also compute the RMSE and they are (1.17, 0.65, 0.15, 0.0040). Those RMSE numbers shall be very much close to the standard deviation reported above for a large dataset. The translation RMSE numbers are larger than what Ritter *et al.* reported, but the rotation and scaling are better than their results. As we mentioned earlier, it is difficult to register image set 1. Excluding set 1, the RMSE numbers are (0.34, 0.40, 0.06, 0.0020). The RMSE numbers of the translations are comparable to those of Ritter *et al.*, while the rotation and the scaling factor are much better than theirs.

#### 4.3.3.2 Registration of Stereo Images

Visual inspection of the registration results reveals that mutual information maximization registration succeeded 9 times out of 11 and it failed on ll5/lr5 of set 1 and ll7/lr7 of set 3. The success rate is 82%. The average registration time is 11.0 seconds with a standard deviation of 1.5, which amounts to a speedup of 8.39. Again, the speedup value is only indicative.

For the stereo registration, the mean and standard deviation of misregistration parameters are  $(-0.46 \pm 1.69, -0.67 \pm 1.72, -0.02 \pm 0.24, 0.0043 \pm 0.0082)$ . The estimation of the rotation angle and the scaling factors are very accurate.

The RMSE numbers are (1.75, 1.85, 0.25, 0.0090). The translation RMSE numbers are larger than what Ritter *et al.* reported, but the rotation and scaling are comparable to their results. As we mentioned earlier, it is hard to pair the points in two stereo images.

#### 4.3.3.3 Registration for Temporal and Stereo Images

If we combine all results discussed above, then the success rate is 86%. The average time for registration is 11.7 seconds with a standard deviation of 2.6, which is about 7.83 speedup against Ritter *et al.* [5].

The average misregistrations are (-0.32, -0.17, -0.01, 0.0021) for *x* and *y* translations, rotation angle, and scaling factor, respectively. The RMSE are (1.49, 1.39, 0.21, 0.0070).

## 4.4 Summary

Retinal image registration and fusion is an important tool in ophthalmology for diagnosis, lesion progression assessment, and treatment monitoring. Registration by mutual information maximization is one of the most popular automatic algorithms. We implemented such a software system for automatic retinal image registration by mutual information maximization. The software was written in Java and Java 2D for maximum portability. The robust software architecture (MVC framework) was employed to ensure software maintainability and extensibility. The simplex downhill method was exploited to optimize the mutual information function as a function of the registration parameters. To increase the search range and to avoid being trapped by local optimal, multiresolution and subsampling optimization schemes were explicitly employed. Other implementation strategies were also tested, including nearest neighbor and bilinear interpolation and histogram estimation. Our implementation has an 86% success rate. Compared to the published results, our implementation is about 7.83 times quicker with comparable accuracy (mean of misregistration parameters) and precision (standard deviation of misregistration parameters).

Besides the improved registration performance, our system also provided new representation tools to visualize the registration results. The registered images can be displayed side by side to allow direct comparison. Moreover, the registered images can be displayed in a moving curtain fashion and in a checkerboard format, where parts of two images are displayed together. Furthermore, two images can be overlaid to allow one to see one image through the other. These tools were integrated seamlessly to allow the user to check and interpret the registration findings.

We attribute our fast registration speed to the multiresolution and subsampling scheme. To our knowledge this is the first time anyone has taken advantage of the benefits of both in a single implementation. Depending on the image size, the coarse image can be very coarse, thus the registration can be very fast.

While Ritter *et al.* reported a 100% success rate, we only achieved an 86% rate. We found that the resolutions and subsampling frequencies in our multiresolution and subsampling scheme can be adjusted so that the failed registrations can be registered successfully. Our implementation provides a facility to allow the user to change them at runtime. Considering that the registration can be done in less than 15 seconds, it is practical and acceptable to register the failed image pairs with a different set of parameters. Another approach to achieve a higher success rate is to bring the images close to the optimal alignment before starting the automatic alignment. This prealignment proves to be useful in 3D registration and our preliminary results indicate that it is helpful in 2D retinal registration too.

# Acknowledgement

We are grateful to Dr. Nicola Ritter for providing us the retinal images used in this study. Thanks are also due to Dan Kovalik for his proofreading of the manuscript.

# Questions

- 1. What is image registration?
- 2. How is the image registration used in retinal imaging?
- 3. What are the characteristics of retinal image registration?
- 4. What is mutual information?
- 5. What is mutual information based registration?
- 6. What is image interpolation? Why is it important in image registration process?
- 7. Why is the exhaustive optimization not feasible in the mutual information registration process?
- 8. What is a multiresolution optimization? Why is it needed? How is it applied in the registration process?
- 9. Why is the mutual information registration algorithm written in Java?
- 10. What is the MVC framework?

# Bibliography

- Sbeh, Z.B., Cohen, L.D., Momoun, G. and Coscas, G., A new approach of geodesic reconstruction for drusen segmentation in eye fundus images, IEEE Transactions on Medical Imaging, Vol. 20, No. 12, pp. 1321–1333, 2001.
- [2] Rapantzikos, K., Zervakis, M. and Balas, K., Detection and segmentation of drusen deposits on human retina: Potential in the diagnosis of agerelated macular degeneration, Medical Image Analysis, Vol. 7, No. 1, pp. 95–108, 2003.
- [3] Zana, F. and Klein, J.C., A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform, IEEE Transactions on Medical Imaging, Vol. 18, No. 5, pp. 419–428, 1999.
- [4] Osareh, A., Mirmehdi, M., Thomas, B. and Markham, R., Classification and localization of diabetic-related eye disease, Lecture Notes in Computer Science, Vol. 2353, pp. 502–516, 2002.
- [5] Ritter, N., Owens, R., Cooper, J., Eikelboom, R.H. and van Saarloons, P., Registration of stereo and temporal images of the retina, IEEE Transactions on Medical Imaging, Vol. 18, No.5, pp. 404–418, 1999.
- [6] Wade A.R. and Fitzke, F.W., A fast, robust pattern recognition system for low light level image registration and its application to retinal imaging, Optics Express, Vol. 3, No. 5, pp. 190–197, 1998.
- [7] Wienreb, R.N., Assessment of the optic nerve and nerve fiber layer in glaucoma, J. Glaucoma, Vol. 2, pp. 135–137, 1993.
- [8] Brown, L.G., A survey of image registration techniques, ACM Computing Survey, Vol. 24, No. 4, pp. 325–376, 1992.
- [9] Maurer, C.R. and Fitzpatrick, J.M., A review of medical image registration, in Interactive Image-Guided Neurosurgery, Maciunas, R.J., ed., Park Ridge, IL: American Association of Neurological Surgeons, pp. 17–44, 1993.

- [10] Van den Elsen, P.A., Pol, E.D. and Viergever, M.A., Medical image matching—A review with classification, IEEE Engineering in Medicine and Biology, pp. 26–39, 1993.
- [11] Maintz, J.B.A. and Viergever, M.A., A survey of medical image registration, Medical Image Analysis, Vol. 2, No. 1, pp. 1–36, 1998.
- [12] Fitzpatrick, J.M., Hill, D.L.G. and Maurer, C.R., Image Registration, in Handbook of Medical Imaging: Vol. 2, Medical Image Processing and Analysis, Sonka, M. and Fitzpatrick, J.M., eds, SPIE Press, Bellingham, 2000.
- [13] Hajnal, J.V., Hill, D.L.G. and Hawkes, D.J., eds., Medical Image Registration, Boca Raton, FL: CRC, 2001.
- [14] Studholme, C., Hill, D.L.G. and Hawkes, D.J., Automatic threedimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Medical Physics, Vol. 24, No. 1, pp. 25–35, 1997.
- [15] Fitzpatrick, J.M., Hill, D.L.G., Shyr, Y., West, J., Studholme, C. and Maurer, C.R. Jr., Visual Assessment of the accuracy of retrospective registration of MR and Ct images of the brain, IEEE Transactions on Medical Imaging, Vol. 17, No. 4, pp. 571–585, 1998.
- [16] Maes, F., Collognon, A., Vandermeulen, D., Marchal, G. and Suetens, P., Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187– 198, 1997.
- [17] Wells, W.M. III, Viola, P., Atsumi, H., Nakajima, S. and Kikinis, R., Multimodal volume registration by maximization of mutual information, Medical Image Analysis, Vol. 1, No. 1, pp. 35–51, 1996.
- [18] Hawkes D.J. and Hill, D.L.G., Medical imaging at Guy's Hospital, King's College London, IEEE Transactions on Medical Imaging, Vol. 22, No. 9, pp. 1033–1041, 2003.

- [19] Hill, D., Studholme, C. and Hawkes, D., Voxel similarity measures for automated image registration. In: Proceedings of Third Conference Visualization in Biomedical Computing, SPIE, pp. 205–216, 1994.
- [20] Collignon, A., Automated multi-modality image registration based on information theory, in Bizais, Y. ed., In: Proceedings of Information Processing in Medical Imaging Conference, Kluwer Academic Publishers, Dordrecht, pp. 263–274, 1995.
- [21] Collignon, A., Vandermuelen, D., Suetens, P., and Marchal, G., 3d multimodality medical image registration using feature space clustering, Ayache, N. ed., In: Computer Vision, Virtual Reality and Robotics in Medicine, Springer Verlag, Berlin, pp. 195–204, 1995.
- [22] Pluim, J.P., Maintz, J.B.A. and Viergever, M.A., Mutual informationbased registration of medical images: A survey, IEEE Transactions on Medical Imaging, Vol. 22, No. 8, pp. 986–1004, 2003.
- [23] Laliberte, F., Gagnon, L. and Sheng, Y., Registration and fusion of retinal images—An evaluation study, IEEE Transactions on Medical Imaging, Vol. 22, No. 5, pp. 661–673, 2003.
- [24] Peli, E., Augliere, R.A. and Timberlake, G.T., Feature-based registration of retinal images, IEEE Transactions on Medical Imaging, Vol. 6, No. 2, pp. 272–278, 1987.
- [25] Yu, J.J.H., Hung, B.N. and Liou, C.L., Fast algorithm for digital retinal alignment, Proc. IEEE Ann. Conf. Engineering Medicine Biology Society, Vol 2, pp. 374–375, 1989.
- [26] Cideciyan, A.V., Jacobson, S.G., Kemp, C.M., Knighton, R.W. and Nagel, J.H., Registration of high resolution images of the retina, Proc. SPIE, Vol. 1652, pp. 310–322, 1992.
- [27] Hart, W.E. and Goldbaum, M.H., Registration of retinal images using automatically selected control points, Proc. IEEE Int. Conf. Image Processing, pp. 576–580, 1994.
- [28] Matsopoulos, G.K., Mouravliansky, N.A., Delibasis, K.K. and Nikita, K.S., Automatic retinal image registration scheme using global

optimization techniques, IEEE Transactions on Information Technology in Biomedicine, Vol. 3, No. 1, pp. 47–60, 1999.

- [29] Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M. and Goldbaum, M., Detection of blood vessels in retinal images using two-dimensional matched filters, IEEE Transactions on Medical Imaging, Vol. 8, No. 3, pp. 263–269, 1989.
- [30] Can, A., Stewart, C.V., Roysam, B. and Tanenbaum, H.L., A featurebased, robust, hierarchical algorithm for registering pairs of images of the curved human retina, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 3, pp. 347–364, 2002.
- [31] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., Numerical Recipes in C, Cambridge University Press, Cambridge, Chap. 10, 1992.
- [32] Gonzalez, R.C. and Woods, R.E., Digital Image Processing, Addison-Wesley, Reading, MA, Chap. 6, 1992.
- [33] Arnold, K. and Gosling, J., The Java Programming Language, 2nd edn, Addison-Wesley, Reading, MA, 1998.
- [34] Knudsen, J., Java 2D graphics, O'Reilly & Associates. Inc., 1999.
- [35] Tsao, J., 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.
- [36] Maes, F., Vandermeulen, D. and Suetens, P., Comparative evaluation of multi-resolution optimization strategies for multimodality image registration by maximization of mutual information, Medical Image Analysis, Vol. 3, No. 4, pp. 373–386, 1999.
- [37] Szyperski, C., Component Software: Beyond Object-Oriented Programming, Addison-Wesley, 1998.

# **Chapter 5**

# Quantification of Brain Aneurysm Dimensions from CTA for Surgical Planning of Coiling Interventions

Mónica Hernández,<sup>1</sup> Alejandro F. Frangi,<sup>1,2</sup> and Guillermo Sapiro<sup>3</sup>

# Abbreviations

nD	n dimensional, $n \in \{2, 3\}$
3DRA	Three dimensional Rotational Angiography
ACA	Anterior Cerebral Artery
ACoA	Anterior Communicating Artery
ANOVA	Analysis Of Variance
BA	Basilar Artery
CTA	Computed Tomography Angiography
DSA	Digital Substraction Angiography
GAC	Geodesic Active Contours
GAR	Geodesic Active Regions
GDC	Guglielmi Detachable Coil
ICA	Internal Carotid Artery
kNN	k-Nearest Neighbor
MAP	Maximum A Posteriori

<sup>&</sup>lt;sup>1</sup>Computer Vision Lab, Aragon Institute of Engineering Research, University of Zaragoza, Zaragoza, Spain

<sup>&</sup>lt;sup>2</sup>Department of Technology, Pompeu Fabra University, Barcelona, Spain

<sup>&</sup>lt;sup>3</sup>Electrical Engineering and Computer Sciences Department, University of Minnesota, Minneapolis, USA

MCA	Middle Cerebral Artery
MIP	Maximum Intensity Projection
MRA	Magnetic Resonance Angiography
NA	Not Applicable
PCA	Posterior Cerebral Artery
PCoA	Posterior Communicating Artery
PDF	Probability Density Function
SAH	Subarachnoid hemorrhage
SD	Standard Deviation

# 5.1 Introduction

### 5.1.1 Brain Aneurysms

Brain aneurysms are pathological dilatations of cerebral arteries. These dilatations consist in a progressive enlargement and deformation of the vessel wall produced by blood flow pressure. Brain aneurysms tend to occur at or near arterial bifurcations, mostly at the Circle of Willis, the vascular system that irrigates the basis of the brain.

The Circle of Willis is made up of several vascular segments (Fig. 5.1). The precommunicating segments (A1, A2) of the left and right Anterior Cerebral Arteries (ACA) and the Anterior Communicating Artery (ACoA) form the anterior part of the circle. The postcommunicating segments (P1, P2) of the left and right Posterior Cerebral Arteries (PCA) with the Posterior Communicating Arteries (PCoA) form the posterior part of the circle. The left and right PCoAs emerge from the left and right Internal Carotid Arteries (ICA). The Basilar Artery (BA) and the Middle Cerebral Arteries (MCA), complete the description of the Circle of Willis.

Brain aneurysms are classified into saccular and non-saccular types according to their shape (Fig. 5.2). Non-saccular aneurysms include atherosclerotic, fusiform, traumatic, and mycotic types. Saccular, or berry, aneurysms typically arise at a bifurcation or along turns of the parent vessel, or they point in the direction in which the blood flow would proceed if the turn were not present. Brain aneurysms are named according to the artery or segment of origin. For example, anterior communicating aneurysms are located at the ACoA, and



Figure 5.1: Three-dimensional model of the Circle of Willis. Image courtesy of Dr. Juan R. Cebral, George Mason University. (Color Slide)

posterior communicating aneurysms are usually located at the ICA, near the origin of the PCoA.

The prevalence of unruptured cerebral aneurysms is unknown, but it is estimated to be as high as 5% of the population [1]. The most serious complication of a brain aneurysm is its rupture and the consequent aneurysmal subarachnoid hemorrhage (SAH) with an incidence of sudden death of the 12.4% and rates



Figure 5.2: Models of aneurysm. (a) Saccular aneurysm. (b) Non-saccular (fusiform) aneurysm.

of fatality from 32% to 67% after the hemorrhage [2, 3]. Morbidity rates reach 10.9% due to intra cranial bruise, subsequent recurrent bleeding, hydrocephaly and spasms in the surrounding brain vessels [4].

### 5.1.2 Planning Endovascular Interventions

The treatment of cerebral aneurysms aims at their complete elimination from the circulation. The traditional surgical technique consists in clipping the aneurysm. In the case of subarachnoid hemorrhage (SAH), early management prevents rebleeding and future rupture. More concretely, it has been observed that an early treatment of the aneurysm reduces the risk of SAH, mortality and morbidity [5]. Nevertheless, the risk of lesions during the intervention is still high in surgically unfavorable locations.

In the past years, there has been a growing trend to practise minimally invasive endovascular procedures to treat cerebral aneurysms [6]. In cases of potential aneurysm rupture, these techniques stabilize the patient and facilitate further aggressive treatments with the purpose of preventing the spasm after intra-cerebral bleeding. Aneurysm coiling with Guglielmi detachable coils (GDC) is probably the most widespread technique for permanent aneurysm embolization. The placement of coils inside the aneurysm promotes blood clotting by electrothrombosis, thus avoiding blood flow and pressure, and hampering its rupture [7]. Figure 5.3 shows a picture of a GDC, and two digital substraction angiographies of a brain aneurysm before and after patient embolization.

A correct size selection and placement of the GDC inside the aneurysm is crucial for the success of the treatment, for which it is highly desirable to have pre-surgical knowledge of the dimensions as well as from the three-dimensional morphology of the aneurysm and connected vessels. For example, the amount of aneurysm filling following the coiling is strongly dependent on the geometry and dimensions of the aneurysm. In particular, it has been shown that knowledge about the maximum neck diameter and the main axes dimensions of the aneurysm, dome width and depth, play an important role in the selection of patients and materials for an appropriate treatment with GDCs [8]. Figure 5.4 shows the traditional measurements used for the planning of the endovascular procedures.



Figure 5.3: Process of endovascular embolization with GDCs. (a) An example of GDC. (b) Digital substraction angiography of an embolized aneurysm using GDC coils. Introduction of the GDC inside the aneurysm by catheterization. (c) Final result of the embolization. The placement of the coil promotes blood clotting inside the aneurysm. The clot avoids blood flow inside the aneurysm as is appreciated in the DSA image. The progressive reduction of intracranial blood flow pressure avoids its rupture.



Figure 5.4: Maximum Intensity Projection (MIP) of a brain aneurysm reconstructed from a CTA image. The measurements of the neck, dome width and depth needed to carry out the selection of the coil size in the endovascular treatment are indicated over the MIP.

Two-dimensional digital substraction angiography (DSA) is considered the gold standard technique for the detection and quantification of brain aneurysms. However, other less invasive acquisition techniques like Computed Tomography Angiography (CTA) or Magnetic Resonance Angiography (MRA) are also used as complementary methods for these aims [9, 10]. In order to estimate the dimensions of an aneurysm from 3D CTA or MRA, it is customary to perform lineal measurements on Maximum Intensity Projections (MIP) of the original volumetric scan. The MIP provides a 2D image of the 3D data in an angle considered optimal by the neuroradiologist. The optimal viewpoint is chosen so that the magnitudes of interest are maximum. Manual measurements are then carried out on the basis of this 2D image using electronic callipers.

The selection of the optimal view angle, introduces a high degree of subjectivity to the quantification of the aneurysm. Window levelling is often used to enhance the arterial structures of interest thus increasing the subjectivity in the quantification of the aneurysm morphology. In some images, the presence of nearby bony structures imposes restrictions to the selection of the viewpoint. This forces, in some cases, the selection of suboptimal views to perform the measurements. In MIP images, it is often difficult to determine depth relationships between the aneurysm and the surrounding and sometimes overlapping vessels. Therefore, the use of computerized 3D segmentation techniques is crucial for accurate quantification of the aneurysm dimensions as well as for a correct interpretation of the 3D morphology.

### 5.1.3 Implicit Deformable Models

Many of the recent approaches used to segment vascular structures from medical images use deformable models. The use of implicit deformable models combined with level set methods in arterial structures has become very popular over the last few years. The ability to capture the topology of complex structures makes it a very suitable technique for the extraction of the shape of arterial structures. Loncaric *et al.* [11] use classic Geodesic Active Contours (GAC) model combined with level set methods in the segmentation of aortic abdominal aneurysms (AAA). However, classic GAC methods are not able to deal with narrow and twisted vessels. To this end, some improvements to the traditional method have been proposed in the literature. For example, Lorigo *et al.* [12, 13] use codimension two GACs for the extraction of the cerebral vasculature. This method successfully obtained a segmentation of the whole tree. However, the final segmentation shows relatively thinner vessels compared to MIP images and thresholding schemes, and the model seems to be not able to capture abnormal vessel lumina.

Other authors propose to combine implicit deformable models with a smart initialization of the evolving surface inside the vessels of interest. Van Bemmel *et al.* [14, 15], for instance, compute the central axis of the artery and use it as initialization for the GAC model in the segmentation of carotid arteries and the aorta. Deschamps *et al.* [16] use the output of a vessel enhancement filter [17] as speed for the Fast Marching method for a fast and rough initialization of the model. The evolved surface is used as initialization of the implicit deformable model to obtain a more accurate segmentation.

Classical deformable models depend on the gradient of the image as stopping force. The front in evolution usually suffers from leakage or bleeding in places with weak or inhomogeneous image gradient, and does not provide good results in brain vessels. Since the method of region competition proposed by Zhu and Yuille [18], there have been several efforts to include statistical region-based information in the process of segmentation. Paragios made a wide study about the inclusion of this information in the GAC model [19]. In places with weak gradient, regional information drives the evolution of the front, thus avoiding leakage in the edges of the object of interest. Similar work including statistical information on the implicit model was done by Yezzi *et al.* [20]. Deschamps successfully used Paragios' Geodesic Active Regions (GAR) model in the segmentation of brain aneurysms [16].

Although a number of algorithms based on implicit deformable models have addressed the problem of cerebral vessels segmentation [13, 16], these do not produce satisfactory results when confronted with images of standard quality in average radiology departments. For example, the work reported by Deschamps [16] deals with rotational angiography (3DRA) where the background and bone tissues have a well differentiated contrast with respect to vessels (Fig. 5.5a). On the other hand, the ranges of vessel and bone intensities in CTA ususally overlap (Fig. 5.5b). Most of the previous attempts to solve this problem have presented segmentation results only on few selected images. Unfortunately, there is a general lack of larger evaluation studies on image databases acquired under routine clinical conditions.



**Figure 5.5:** (a) Histogram from a 3DRA image of a brain aneurysm. (b) Histogram from a CTA image of a brain aneurysm. In the first case, the distributions from background-bone (intensity values from -20,000 to -10,000 approximately) and vessels (intensity values from -10,000 to 0) are well differentiated. In the second case, the limit of the distribution from background (intensity values from 1000 to 1200 approximately) is not differentiated from the distributions of vessel (intensity values from 1100 to 1600) and bone (intensity values from 1100 to 2400) that are totally overlapped.

### 5.1.4 Chapter Outlook

This chapter presents a method to address the problem of brain aneurysm segmentation in CTA images. Due to the limited resolution of CTA brain images, the front is usually not able to evolve through narrow and twisted objects as, for example, the thinnest brain vessels of the Circle of Willis. For this reason, a two-stage algorithm is devised. In the first stage, a fast and rough segmentation of all the tissues present in the image is obtained using a Maximum a posteriori (MAP) classifier. In the second stage, a GAR method is used to obtain the final segmentation with sub-voxel accuracy. The novelty of the method consists in the use of differential image descriptors of high order, and the inclusion of non-parametric information in the GAR model. This is done using a k-Nearest Neighbor (kNN) classifier to estimate the underlying probability density functions of the main tissue types that are present in the CTA images. The result is an algorithm that provides accurate segmentations with very little user intervention in the selection of its parameters.

The method has been evaluated on a database of 39 brain aneurysms placed within the Circle of Willis. The technique is compared against manual measurements of three geometrical descriptors of the aneurysm morphology, which are standard in assessing the viability of surgical treatment with GDCs.

The chapter is organized as follows. Section 5.2 introduces the proposed method to address the segmentation of brain aneurysms. Section 5.3 describes the materials and methods used in clinical practise to obtain the measurements for the planning of the endovascular intervention. The procedures for manual and computerized measurements are also described. The evaluation of the computerized technique is reported in section 5.4. Finally, in section 5.5 the results of the evaluation study are discussed and some concluding remarks are made in section 5.6.

# 5.2 Aneurysm Segmentation

The GAR model includes region-based information into the classical GAC model. This is done using region descriptors, which are defined in terms of the negative logarithm of a Probability Density Function (PDF) associated with the region and the image inside that region. The generation of the PDF for each region, involves the definition of the features that characterize the image inside the region and a suitable PDF estimation technique. Classically, region descriptors were defined using the intensity values of the image inside the region as features, and a Gaussian PDF model [18, 19]. However, higher order information was not used while we argue that it could provide a more complete description of regional properties.

#### 5.2.1 Image Features

A common approach to analyze the local behavior of an image,  $I(\mathbf{x})$ , is to consider its Taylor expansion in the neighborhood of a point  $\mathbf{x}_0$  at scale  $\sigma$ 

$$I_{\sigma}(\mathbf{x}_{0} + \delta \mathbf{x}_{0}) \approx I_{\sigma}(\mathbf{x}_{0}) + \delta \mathbf{x}_{0}^{T} \nabla_{\sigma} I(\mathbf{x}_{0}) + \delta \mathbf{x}_{0}^{T} H_{\sigma}(\mathbf{x}_{0}) \delta \mathbf{x}_{0}$$
(5.1)

where  $I_{\sigma}$ ,  $\nabla_{\sigma} I$ , and  $H_{\sigma}$  are the image intensity function, the gradient vector, and the Hessian matrix of the image computed after convolution with a Gaussian kernel of scale  $\sigma$ . While the intensity provides information about the tissue properties (i.e., X-ray absorption in CTA), and the norm of the gradient indicates the amount of contrast between neighboring tissues, the second order structure provided by the Hessian matrix indicates local curvature and bending [17, 21]. We argue that all this information can be relevant in characterizing a region.

### 5.2.2 Non-parametric Tissue Probability Estimation

Pattern recognition techniques can be used to obtain non-parametric PDF estimators. In this framework, the PDF is computed using classifiers. The estimation is usually done using a sparse set of neighboring samples. For example, the Parzen density estimation method computes the probability for a point  $\mathbf{x}$  to belong to a region by computing the number of samples belonging to that region in a fixed neighborhood of radius k. The k-Nearest Neighbor (kNN) estimation technique can be interpreted as a Parzen approach with a neighborhood size adjusted automatically depending on the location of the point. It is reasonable to assume that points with similar local image structure belong to the same tissue class. For this reason, we use kNN rule for the generation of the PDF in our algorithm.

The nearest neighbor principle is one of the best studied techniques for pattern classification [22]. A classifier based on this principle uses a training set of vectors as a collection of labelled cases. For a given pattern, it searches for the nearest vector in this learning set according to a given metric. The traditional classification consists in assigning to the pattern the label of the most voted class. Combinations of number of votes (kNN rule) and distances can be also used for classification. A PDF estimation can be derived from these voting systems.

#### 5.2.2.1 Training Set Construction

We use the kNN rule to estimate the probability density function for each tissue class as follows. For the construction of the training set, some representative CTA images are selected from the whole data base. Then, *N* points are manually picked from these images, and labelled with one of these three tissue classes: vessel, background or bone. A label and a feature vector is associated with each point in the training set. The feature vector is derived from the local differential structure of the image at a small scale  $\sigma$  (Eq. (5.1)). For a point  $\hat{\mathbf{x}}$  in the training set, we associate the feature vector

$$\mathbf{f}(\hat{\mathbf{x}}) = (I_{\sigma}, |\nabla I_{\sigma}|, \lambda_{1_{\sigma}}, \lambda_{2_{\sigma}}, \lambda_{3_{\sigma}})$$
(5.2)

where  $I_{\sigma}$  represents the convolution of the image with a Gaussian kernel and  $\nabla I_{\sigma}$  its gradient. The parameters  $\lambda_{i_{\sigma}}$  represent the eigenvalues of the Hessian matrix of the image  $I_{\sigma}$ , ordered by increasing magnitude.

### 5.2.2.2 Training Set Normalization

As is classical in pattern recognition theory, the feature vectors of the training set are normalized. We applied the normalization

$$\widehat{f}_n^m = \frac{f_n^m - \mu_n}{\sigma_n} \tag{5.3}$$

where  $f_n^m$ ,  $\mu_n$  and  $\sigma_n$  are the value of the *n*-th component of the *m*th feature vector, and the mean and the standard deviation of the *n*th component over the training set, respectively [23].

#### 5.2.2.3 Probability Density Function Estimation

The kNN rule is used to estimate the underlying PDF as follows. For a given voxel  $\mathbf{x}$ , the feature vector  $\mathbf{f}(\mathbf{x})$  is defined as in Eq. (5.2) and normalized as in Eq. (5.3). Then, the *k* nearest feature vectors are found in the training set according to the Euclidean distance. The probability for a voxel of intensity *i* to belong to a tissue class  $C_j$ , is computed from the formula

$$P(I(\mathbf{x}) = i | C_j) = \frac{\sum_{\hat{\mathbf{x}} \in \mathcal{L}_j \cap \mathcal{N}_k(\mathbf{x})} d(\mathbf{f}(\mathbf{x}), \mathbf{f}(\hat{\mathbf{x}}))}{\sum_{\hat{\mathbf{x}} \in \mathcal{N}_k(\mathbf{x})} d(\mathbf{f}(\mathbf{x}), \mathbf{f}(\hat{\mathbf{x}}))}$$
(5.4)

where  $\mathcal{L}_j$  represents the set of points in the training set that belongs to the class  $C_j$ ,  $\mathcal{N}_k(\mathbf{x})$  is the set of the *k* nearest neighbors and *d* represents the Euclidean distance. Figure 5.6 shows an example of the probability density functions estimated by the kNN rule. In the sequel,  $C_0$ ,  $C_1$  and  $C_2$  will stand for vessel, background and bone class, respectively.

### 5.2.3 Maximum A Posteriori Tissue Classification

A MAP tissue classifier is used to obtain a partition of the image domain into regions matching with vessel, background and bone. The probabilities estimated from the kNN rule provide a learned prior probability for a particular voxel to



Figure 5.6: Cross-section of the PDF images estimated by the kNN rule. Brighter areas correspond to higher probabilities. (a) Gray level image. (b–d) Probability for vessel, background and bone, respectively.

belong to a certain class,  $P(I(\mathbf{x}) = i | \mathbf{x} \in C_j)$ . All tissue classes are assumed to be equiprobable.

The Bayes rule is then applied to calculate the posterior probability for a given voxel to belong to a particular class given its intensity,  $P(C_j = c_j | I(\mathbf{x}) = i)$ . The MAP classifier uses the maximum a posteriori probability estimate after anisotropic smoothing [24] to obtain a classification of the voxels of the image

$$C_{j}^{*} = \arg \max_{c_{j} \in \{C_{0}, C_{1}, C_{2}\}} P^{*}(C_{j} = c_{j} | I(\mathbf{x}) = i)$$
(5.5)

where  $P^*$  corresponds to the posterior probabilities after diffusion driven by the equation

$$\frac{\partial P}{\partial t} = \operatorname{div}\left(\frac{\nabla P}{|\nabla P|}\right)^{1/3} |\nabla P|$$
(5.6)



Figure 5.7: Maximum a posteriori classification. Cross-section with the MAP labels. Black corresponds to the vessel tissue, white to bone and gray to the background.

This Partial Differential Equation verifies the Maximum Principle [25]. Therefore, the posteriors remain being probability functions after diffusion. Applying anisotropic diffusion introduces spatial coherence before the MAP decision thus improving the classification results [26]. Figure 5.7 shows one example of the MAP classification. Voxels labelled as vessel are used as initialization of the GAR method introduced in the next section.

## 5.2.4 Geodesic Active Regions

### 5.2.4.1 Geodesic Active Contours

Caselles *et al.* proposed an implicit deformable model known in the literature as Geodesic Active Contours [27]. It conciles Parametric Models and the Level Set Theory. It is based on the idea from geodesic snakes [28] of evolving an initial curve to a local minimum of an energy functional.

The GAC energy functional is defined for surfaces as

$$E(S) = \iint g(S) \, da \tag{5.7}$$

where S represents a parametrization of the evolving surface, g is an inverse edge detector function, and da is the surface area element. This energy functional defines a Riemannian Space, with metric associated to the edge detector function g. The geodesic surfaces on this space, are defined by the edges of the image.

To compute the differential equation that drives the evolution of the surface, traditional variational techniques are used. The Euler-Lagrange equation associated to the functional is

$$(\kappa g - \langle \nabla g, \vec{n} \rangle)\vec{n} = 0 \tag{5.8}$$

where  $\kappa$  is the mean curvature of the evolving surface and  $\vec{n}$  is its outer unitary normal vector. So, the evolution of the surface is driven by the associated gradient descent equation

$$S_t = (\kappa g - \langle \nabla g, \vec{n} \rangle) \vec{n}$$
(5.9)

Following this equation, the surface evolves toward the minimum of the functional, which is achieved at the edges of the image. The resulting steadystate surface is a model of the object of interest. The level set method [29] is used to track its motion, allowing topological changes in the surface and avoiding numerical instabilities. Basically, the level set method consists in embedding the evolving surface in a manifold one dimension higher than S, and implicitly represented by a function  $\phi$ . The surface S can be reconstructed as the level set zero of  $\phi$ . If the manifold evolves following the equation

$$\phi_t = g \cdot div \left(\frac{\nabla \phi}{|\nabla \phi|}\right) |\nabla \phi| - \langle \nabla g, \nabla \phi \rangle \tag{5.10}$$

then the evolution of the zero level set of  $\phi$  is equivalent to the evolution of S driven by Eq. (5.9).

### 5.2.4.2 Introduction of Region-Based Information in the GAC Model

The GAR model [19] combines the classical GAC model [27] with region-based statistical information incorporated into the classical energy functional. Therefore, in places where the gradient is weak, regional information drives the evolution of the surface thus being more robust than GAC.

A surface *S* provides a partition of the space in three regions: inside  $(\Omega_{in})$ , outside  $(\Omega_{out})$ , and the boundary itself (*S*). Considering the surface evolving in the domain of a three-dimensional image, the surface provides the following partition at each time:

$$\Omega(t) = \Omega_{\rm in}(t) \cup \Omega_{\rm out}(t) \cup S(t).$$
(5.11)

Given a partition of the image domain defined by S(t), the descriptor of the inner region is defined as

$$k_{\rm in}(\mathbf{x};t) = -\log(P_{\rm in}(\mathbf{x};t)) \tag{5.12}$$

where  $P_{in}$  is the probability for a voxel **x** to belong to  $\Omega_{in}$ . An analogous definition holds for the outer region.

Following the ideas of the theory of GAC, the energy functional associated to the region-based model is defined as

$$E(t) = \zeta \int_{\Omega_{\rm in}(t)} k_{\rm in}(\mathbf{x}; t) \, d\mathbf{x} + \zeta \int_{\Omega_{\rm out}(t)} k_{\rm out}(\mathbf{x}; t) \, d\mathbf{x} + \eta \int_{S(t)} g(\mathbf{x}) da \quad (5.13)$$

where  $\zeta$  and  $\eta$  control the contribution of the regional and boundary information, respectively.

The evolution of the regions in which the domain is divided can be simplified by expressing it in terms of the evolution of the boundary S(t). So the evolution can be expressed by the partial differential equation

$$\frac{\partial S(\mathbf{x};t)}{\partial t} = F(\mathbf{x},t)\overrightarrow{n}$$
(5.14)

where  $F(\mathbf{x}; t)$  is the evolution speed.

The gradient descent flow associated to the minimization of the functional E(t) is<sup>4</sup>

$$\frac{\partial S(\mathbf{x};t)}{\partial t} = \zeta (k_{\text{out}} - k_{\text{in}}) \overrightarrow{n} - \eta (g\kappa + \langle \nabla g, \overrightarrow{n} \rangle) \overrightarrow{n}$$
(5.15)

So the associated level set equation will be

$$\phi_t + \zeta (k_{\text{out}} - k_{\text{in}}) |\nabla \phi| - \eta (g\kappa |\nabla \phi| + \nabla g \nabla \phi) = 0.$$
(5.16)

As region descriptors, we propose to use the negative logarithm of the probabilities learned from the kNN rule. When using multiple features, like in our approach, a non-linear PDF estimation technique can better adapt to the distributions of the underlying tissues than the traditionally used Gaussian PDFs. So, the probability of the inner region is computed as

$$P_{\rm in} = P(I(\mathbf{x}) = i|C_0) \tag{5.17}$$

<sup>&</sup>lt;sup>4</sup>In the case of time-dependent region descriptors, other additive terms extracted from  $\int_{\Omega_{in}(t)} \frac{\partial k_{in}}{\partial t} d\mathbf{x}$  and  $\int_{\Omega_{out}(t)} \frac{\partial k_{out}}{\partial t} d\mathbf{x}$  appear in the Euler Lagrange equation. In the case of Gaussian and kNN descriptors, it can be shown that these terms have zero contribution (see [16] for the case of Gaussian descriptors).

The outer probability combines the region information of the bone and background tissues. If we assume this distributions to be independent (i.e.,  $C_1 \cap C_2 = \emptyset$ ), then

$$P_{\text{out}} = P(I(\mathbf{x}) = i | C_1 \cup C_2) = P(I(\mathbf{x}) = i | C_1) + P(I(\mathbf{x}) = i | C_2) \quad (5.18)$$

#### 5.2.4.3 Considerations of Implementation

*kNN Algorithm.* For the construction of the classifier, six images were selected, and a total of N = 1830 points were manually picked and labelled. Image features were computed with a Gaussian aperture of  $\sigma = 0.8$  mm. In our approach, we used k = 10 as size of the kNN neighborhood.

For the-extraction of the nearest neighbors, a brute-force computation is highly time-consuming, specially if the size of the training set or the dimension of the feature vector are large. For this reason, we used the approximate nearest neighbors (ANN) algorithm [30], that computes efficiently the knearest neighbors from a given vector. The ANN library implements a number of different data structures, based on kd-trees and box-decomposition trees, and employs different search strategies for speeding-up the search process.

Numerical Schemes for GAR. The level set equation  $\phi_t = F |\nabla \phi|$  is an hyperbolic equation that belongs to the family of Hamilton-Jacobi equations. Only in one dimension, the level set equation is an hyperbolic conservation law. Nevertheless, the numerical methodology developed to solve conservation laws can be generalized to obtain numerical schemes that solve Hamilton-Jacobi equations and, in particular, the level set Equation [31]. The Courant-Friedrichs-Levy (CFL) condition is used to compute the optimal time step for the GAR level set equation [32]. The term  $\zeta (k_{\text{out}} - k_{\text{in}}) |\nabla \phi|$  represents a motion in the normal direction. An upwind numerical scheme is used to select the correct solution. The term  $g\kappa |\nabla \phi|$  represents a motion involving mean curvature. This term acts as a viscous force so a scheme with central differences is sufficient to select the correct solution. The term  $\nabla g \nabla \phi$  represents a motion in an externally generated velocity field. This term is also called advection term. Upwind numerical differences are also used to select the correct solution in this case.

Initialization and Reinitialization with the Fast Marching Method. The Fast Marching algorithm is an efficient method to compute the solution of the Eikonal equation  $F \cdot |\nabla T| = 1$  with initial condition  $T_0 = 0$ . If F = 1, the solution

of the equation is the distance transform to the surface represented by the initial condition,  $T_0 = 0$ . The initial surface is computed as the boundary vessel voxels in the MAP classification. The Fast Marching algorithm is used to compute the distance transform to this surface. This function is used as initialization to the level set algorithm.

The level set equation, Eq. (5.16), does not preserve the distance transform through the iterations. The numerical approximations described below may cause a numerical deterioration of the solutions if the evolving front  $\phi$  is not smooth enough. To avoid this phenomenon,  $\phi$  is replaced in each iteration by a distance transform  $\hat{\phi}$  with the same level set zero as  $\phi$ . To do this, the Fast Marching algorithm is used. The values of the voxels where there is a change of sign define the evolving surface. These voxels are introduced in the Fast Marching algorithm as initial condition instead of explicitly computing the zero level set and then use it as initial condition. The result is also a distance transform to the evolving surface.

*Curvature Constraints.* The mean curvature of a surface is defined in differential geometry as the inverse of the radius of the osculant sphere. If the minimum grid size is  $\Delta x$ , it makes no sense considering osculant spheres of sizes less than  $\Delta x$ . So, the discrete values of the curvature are limited to the interval  $\left[-\frac{1}{\Delta x}, \frac{1}{\Delta x}\right]$ . The numerical approximation introduces singularities in the calculus of the curvature function. To avoid their propagation in the GAR algorithm, a Gaussian smoothing with  $\sigma = 0.8$  mm was applied to the curvature.

*Parameter Selection*. Parameters  $\zeta$  and  $\eta$  control the influence of the region and boundary based forces in the motion of the surface. The choice of these parameters depends on the confidence of the user in the different descriptors. For this application,  $\zeta$  and  $\eta$  were chosen equal to 1.0.

Regarding the selection of the edge detector function g, there is an interesting property that relates the parameters involved in the level set equation to a bound on the curvature of the front in evolution. The front evolution is driven by Eq. (5.15). At the steady-state,  $S_t = 0$ , so  $\zeta(k_{\text{out}} - k_{\text{in}}) - \eta(g\kappa + \nabla g \vec{n}) = 0$ . This formula gives upper and lower bounds for the curvature

$$|\kappa| \le \left|\frac{\zeta}{g\eta}\right| \cdot |k_{\text{out}} - k_{\text{in}}| + \left|\frac{\nabla g}{g}\right|$$
(5.19)

In the case of segmentation of narrow vessels, the curvature of the model is high at certain locations. Therefore, the selection of the parameters has to be such that the curvature is not limited by the right-hand term of inequality (5.19). Remember that the curvature was also bounded by  $\frac{1}{\Delta x}$ , so the resolution of the image grid provides another constraint to the algorithm in the capture of narrow structures.

Traditionally, two functions have been used as contour detectors

$$g(\mathbf{x}) = \frac{1}{(1 + |\nabla_{\sigma} I(\mathbf{x})|)^p}, \ p = 1, 2$$
(5.20)

$$g(\mathbf{x}) = e^{-\epsilon |\nabla_{\sigma} I(\mathbf{x})|}, \ \epsilon > 0$$
(5.21)

where  $\nabla_{\sigma}I = \nabla(I * G_{\sigma})$  and  $G_{\sigma}$  is a Gaussian kernel. In this work, the Leclerc function (Eq. (5.21)) was used with  $\sigma = 0.8$  mm, as edge detector function. For the selection of  $\epsilon$ , we noted that in places where  $k_{out} \approx k_{in}$ , Inequality (5.19) leads to  $|\kappa| \leq \epsilon ||\nabla_{\sigma}I||_{\infty}$ . So if  $\epsilon_1 < \epsilon_2$ , the steady-state model associated with  $\epsilon_2$  will be able to capture narrower structures than the steady-state model associated to  $\epsilon_1$ . Therefore, in all the experiments, the Leclerc contour detector function g was selected with  $\epsilon = 10$ . Less than 20 iterations were then enough to assure convergence of the algorithm.

# 5.3 Evaluation: Materials and Methods

### 5.3.1 Clinical Protocol

The evaluation study was performed on a data base of 39 brain aneurysms located mostly at the Circle of Willis; more specifically, at the Anterior Communicating (16), Posterior Communicating (10) and Middle Cerebral (10) Arteries. There were also three aneurysms located at the Pericallosal (1) and Internal Carotid Arteries (2). Aneurysm sizes (dome depth size) varied from 2.9 to 16.3 mm. Table 5.1 shows the distribution of sizes, manually measured, according to the location of the aneurysm.

Image acquisition was performed using an Helical Elscint CT Twin scanner (Marconi; Haifa, Israel) with 120 kV/300 mA for the amplifier tube, 1.2-mm collimation with an helical pitch of 1 and slice spacing of 0.65 mm. Images were reconstructed on a  $512 \times 512$  matrix with a square Field Of View (FOV) of 20.8 cm yielding an in-plane resolution of 0.4 mm. A total of 140 ml of non-ionic contrast fluid was intravenously administrated (Omnitrast 300 mg; Schering,

	n	Neck (mm)	Width (mm)	Depth (mm)
ACoA	16	$2.67\pm0.98$	$5.73 \pm 4.18$	$7.27 \pm 4.22$
ACoP	10	$3.03\pm0.88$	$4.45 \pm 1.79$	$5.50 \pm 1.99$
MC	10	$2.71 \pm 1.11$	$5.40 \pm 2.01$	$6.07 \pm 1.91$
Pericallosal	1	$3.50 \pm NA$	$8.27 \pm NA$	$9.52 \pm NA$
ICA	2	$3.13 \pm 1.46$	$6.01\pm2.20$	$6.63 \pm 4.47$

Table 5.1: Overview of aneurysm dimensions present in our database assessed via manual measurements done by a neuroradiologist. The table shows mean  $\pm$  SD of the mean of the manual measurements by two observers. Aneurysms were stratified according to their location

Berling, Germany) at a rate of 3 ml/s, starting the scanning 20 seconds after the onset of contrast administration.

The acquired images were transferred to a SGI Indigo2 workstation (Silicon Graphics, Mountain View, CA) for viewing and postprocessing. The manual quantification of the aneurysms was performed using 2D MIP images and measuring tools provided by the console software Omnipro (Marconi; Haifa, Israel). The clinical parameters needed for the planning of the endovascular intervention were the maximum neck diameter, the maximum dome width, and maximum dome depth of the aneurysm. As it is customary in clinical routine, the measurements were carried out along several projection angles and, from those, the neuroradiologist chose the view-angle producing maximal measurements.

### 5.3.2 Computerized Protocol

Using the marching cubes algorithm, a 3D model of the aneurysm was reconstructed from the zero-level set of  $\phi$ . To make the computerized measurements comparable to the manual gold standard, the models were rendered with a viewpoint selected according to the criterion of maximality used to generate the MIP images. Two points were manually pinpointed in the 3D scene, corresponding to the end-point of the measured magnitude from that angle. Measurements are then performed by projecting this points into the camera plane to emulate MIP-based measurements.
## 5.4 Results

#### 5.4.1 Examples

In Fig. 5.8 and 5.9 we show some examples of the segmentations of some representative aneurysms from our data base.

#### 5.4.2 Evaluation

Two experts carried out the manual measurements twice per aneurysm with enough delay between sessions to consider them independent. The average of the manual measurements is considered as gold-standard and compared to the measurements obtained by the model based approach. Analysis of variance is used to estimate the reproducibility of the manual method. Bland–Altman analysis [33] is used to obtain the repeatability of the manual method for each of the two observers, the agreement between the observers, and the agreement between the manual and the computerized method.

#### 5.4.2.1 Analysis of Variance

In the ANalysis Of VAriance (ANOVA) model, measurements are considered samples from a random variable that follows a linear statistical model. The associated variance is factored out into the components of the variance according to

$$\sigma^2 = \sigma_S^2 + \sigma_O^2 + \sigma_H^2 + \sigma_W^2 \tag{5.22}$$

where  $\sigma_S^2$  is the variance due to inter-subject differences,  $\sigma_O^2$  is the variability introduced by the observers,  $\sigma_H^2$  is the variability introduced by the observersubject interaction, and  $\sigma_W^2$  is the within-subject random variation. The analysis of variance allows to estimate the significance that each component has in their contribution to the total variance of the measurements and the intra- and interobserver variation.

A first analysis of variance reported no statistically significant interaction between the observer and the subject ( $\sigma_H$ ) with p > 0.55 in all cases. The results of the analysis of variance with the suppression of this term are presented in Tables 5.2–5.4. Table 5.5 presents the results of the intra- and interobserver variation.

#### Quantification of Brain Aneurysm Dimensions



(a)



(b)



(c)

Figure 5.8: Some representative examples of the models obtained by the algorithm. (a), ACoA, (b), PCoA, and (c) MCA aneurysms. (Color slide).



(a)



(b)





Figure 5.9: More representative examples of the models obtained by the algorithm. (a), ACoA, (b), PCoA, and (c) MCA aneurysms. (Color slide).

Source of variation	SSq	DF	MSq	F	р
Observer	19.462	1	19.462	9.51	0.0026
Subject	1462.977	38	38.499	18.81	< 0.0001
Session	237.431	116	2.047		
Total	1719.869	155			

Table 5.2: Two-way ANOVA of manual measurements for the neck<sup>a</sup>

 $^a$  SSq (Sum of Squares), DF (Degrees of Freedom), MSq (Mean Squares), F (F of Snedecor) and p (Snedecor test significance).

Table 5.3: Two-way ANOVA of manual measurements for the dome width<sup>a</sup>

Source of variation	SSq	DF	MSq	F	р
Observer	4.743	1	4.743	3.47	0.065
Subject	1331.455	38	35.038	25.64	< 0.0001
Session	158.542	116	1.367		
Total	1494.740	155			

 $^a$  SSq (Sum of Squares), DF (Degrees of Freedom), MSq (Mean Squares), F (F of Snedecor) and p (Snedecor test significance).

Table 5.4:	Two-way ANOVA of manual	measurements for	the dome
depth <sup>a</sup>			

Source of variation	SSq	DF	MSq	F	р
Observer	19.462	1	19.462	9.51	0.0026
Subject	1462.977	38	38.499	18.81	< 0.0001
Session	237.431	116	2.047		
Total	1719.869	155			

 $^a$  SSq (Sum of Squares), DF (Degrees of Freedom), MSq (Mean Squares), F (F of Snedecor) and p (Snedecor test significance).

Table 5.5:	Intra- and	inter-obs	server va	ariability	of manua	l
measureme	ents					

	Neck (mm)	Width (mm)	Depth (mm)	
Intraobserver	1.43	1.16		
Interobserver	1.50	1.18	1.50	

Table 5.6: Repeatability and agreement study within observers and methods, respectively. ObsI and ObsII stand for each observer, and MB indicates the model-based technique. The table shows  $\mu \pm$  SD of the difference of the measurements in mm. The standard deviation has been corrected for repeated measurements in the agreement study [33]

	Neck (mm)	Width (mm)	Depth (mm)
ObsI	$-0.07\pm1.09$	$0.94 \pm 1.87$	$-0.65 \pm 2.41$
ObsII	$-0.51\pm0.86$	$-0.34 \pm 1.35$	$0.18 \pm 1.50$
ObsI vs ObsII	$-0.03\pm1.22$	$0.34 \pm 1.91$	$-0.70\pm2.45$
<b>ObsI vs MB</b>	$-0.47 \pm 1.05$	$0.23 \pm 1.86$	$-0.69\pm2.12$
<b>ObsII vs MB</b>	$-0.44\pm0.91$	$-0.11\pm1.43$	$0.00 \pm 1.55$

#### 5.4.2.2 Bland–Altman Study

The Bland–Altman plot is a statistical method of comparison of two clinical measurement techniques. The agreement between the two techniques can be quantified using the standard deviation of the differences between observations made on the same subjects. Bland–Altman graphs show the distribution of the differences by plotting the mean against the differences of paired measurements. The information of the Bland–Altman graphs can be summarized by providing the bias ( $\mu$ ) and standard deviation (SD) of the differences of the measurements. The limits of agreement, defined as  $\mu \pm 1.96 \cdot$  SD, provide an interval within which the 95% of the differences between measurements are expected to lie. When repeated measurements from two techniques are available, a corrected standard deviation is computed [33]. A very similar analysis to the limits of agreement approach can be applied to quantify the repeatability of a method from replicated measurements obtained from the same measurement technique.

The results of the Bland-Altman study are shown in Table 5.6. Figure 5.10 shows the Bland-Altman graphs.

## 5.5 Discussion

Classic GAC approaches were unsatisfactory for segmenting the cerebral vasculature from CTA and more sophisticated speed functions introducing statistical



Figure 5.10: In the first column, Bland–Altman graphs comparing the two manual measurements of each observer. Symbols  $\bigcirc$  and  $\triangle$  stand for ObsI and ObsII, respectively. In the second column, Bland–Altman graphs comparing the intersession measurement for manual and computerized methods. Symbol  $\bigcirc$  stands for ObsI vs ObsII study,  $\nabla$  for ObsI vs MB and  $\triangle$  for ObsII vs MB. In all plots the horizontal axis of the plot indicates the average and the vertical axis indicates the difference between measurements. The bias and the limits of agreement are indicated.



(a)

(b)



Figure 5.11: Cross-section with the probability density function estimated from the Gaussian model. Brighter areas correspond to higher probabilities. (a) Original gray level image. (b–d) Probability density functions for vessel, background and bone, respectively.

information from the image were required to improve them. Most of the approaches found in the literature use a Gaussian model for the intensities of each region. Figure 5.11 shows an example of the tissue probability density functions modelled by Gaussian distributions. Compared to Fig. 5.6, it can be appreciated that the probability of vessel is higher in the transition between bone and background. The probability of bone in the interior of the aneurysm is also higher. Background tissue inside the bone has high probability for vessel tissue. The introduction of these features in the region-based term makes the model less robust and very sensitive to the parameter settings of the algorithm, which have to be tuned for each patient to compensate the effect of the misclassification.

The use of non-parametric statistical information provides more accurate segmentations with minimal sensitivity to the selection of the parameters. In the case of Gaussian distributions, the region descriptors are time and front dependent. In contrast, the non-parametric approach presented in this work does not impose these two constraints.

The features used by the kNN rule are computed at a single scale. It would seem that, due to the nature of the object to be segmented, a multiscale approach should provide better results. However, it was observed that results were worse than when using a single scale. This could be explained by the fact that as the number of scales increases, the dimensionality of the feature space also increases. This may deteriorate the performance of the classifier owing to the *peaking phenomenon* [34]. We are currently working on improvements on the PDF estimation technique using a multiscale approach and dimensionality reduction strategies.

The aneurysms involved in the study had a mean size of 2.81 mm for the neck diameter, 5.40 mm for the width, and 6.44 mm for the depth with standard deviations of 0.84, 2.95 and 3.10 mm, respectively. The ANOVA reported an intraand inter- observer standard deviation of less than 1.50 mm in all the cases. Results obtained in the Bland-Altman study with the manual method showed that both observers have a similar performance in independent sessions. The repeatability study shows a bias less than 0.94 mm in all the cases. The standard deviation is larger in the measurements of the aneurysm width and depth than in the neck diameter.

The agreement study indicated a bias less than 0.70 mm in all cases. The standard deviation is larger in the measurement of the aneurysm width and depth than in the neck diameter as happened in the repeatability study. This is logical as minimal variations in the selection of the view angle can yield large variations in the saccular dimensions of the aneurysm when measured on the 2D projection images. These variations are less significative at the neck due to its smaller size and symmetry.

When comparing manual and computerized measurements, it can be observed that the bias is, in the worst case, approximately of the order of a voxel (-0.69 mm). The standard deviations are lower than in the agreement study between observers in all cases. Therefore, the computerized method has a higher agreement with each observer separately than the agreement achieved between the observers themselves.

In some of the patients, MIP images tended to induce misinterpretation of the overlapping vessels leading to wrong estimates of the neck size. Therefore, particularly in these cases, aneurysm quantification from 3D models may help in a more accurate determination of coil dimensions for the surgical intervention.

## 5.6 Conclusion

We have presented a method for three dimensional quantification of brain aneurysms for the purpose of surgical planning and the corresponding evaluation study. This study demonstrates the feasibility of using implicit deformable models combined with non-parametric statistical information to quantify aneurysm morphology and to obtain clinically relevant parameters. In summary, the technique presented in this work will contribute to the computerized surgical planning of coiling procedures by allowing more accurate and truly 3D quantification of brain aneurysms.

## Acknowledgments

The authors would like to acknowledge Dr. R. Barrena and Dr. G. Hernández from the Aragon Institute of Health Sciences (Zaragoza, Spain) for providing the data, measurements and clinical background. We also would like to acknowledge F. Mémoli for interesting conversations during this work. MHG is supported by the grant AP2001-1678 from the Spanish Ministry of Education, Culture and Sport. AFF is supported by a Ramon y Cajal Research Fellowship from the Spanish Ministry of Science and Technology (MCyT). GS is supported by ONI28NSF. This research was partially supported by the grants TIC2002-04495-C02 and TIC2000-1635-C04-01 from the MCyT, and G03/185 from ISCIII.

## Questions

1. Why are normalized the feature vectors? In the normalization of the sample feature vectors, the mean and standard deviation extracted from the training set are used. Explain why. 2. Infer the gradient descent flow of the surface S associated with the kNN region-based energy functional

$$E_{\text{region}}(t) = \int_{\Omega_{\text{in}}(t)} k_{\text{in}}(\mathbf{x}; t) \, d\mathbf{x} + \int_{\Omega_{\text{out}}(t)} k_{\text{out}}(\mathbf{x}; t) \, d\mathbf{x}$$
(5.23)

3. Explain the influence of each of the terms involved in the evolution of the surface driven by

$$\frac{\partial S(\mathbf{x};t)}{\partial t} = \zeta (k_{\text{out}} - k_{\text{in}}) \overrightarrow{n} - \eta (g\kappa + \langle \nabla g, \overrightarrow{n} \rangle) \overrightarrow{n}$$
(5.24)

- 4. How could we determine the convergence of the level set algorithm?
- 5. Explain what could be the advantages of using 3D models for quantification against 2D MIP images. Regarding the evaluation study presented in this chapter, was it taking real profit of the 3D measurement capability?

## Bibliography

- Winn, H., Jane, J., Taylor, J., Kaiser, D., and Britz, G., Prevalence of asymptomatic incidental aneurysms: review of 4568 arteriograms, J. Neurosurg., Vol. 97, No. 1, p. 240, July 2002.
- [2] Huang, J. and van Gelder, J., The probability of sudden death from rupture of intracranial aneurysms: a meta-analysis, Neurosurgery, Vol. 51, No. 5, pp. 1101–5, Nov. 2002.
- [3] Hop, J., Rinkel, G., Algra, A., and van Gijn, J., Case-fatality rates and functional outcome after subarachnoid hemorrhage: a systematic review, Stroke, Vol. 28, pp. 660–664, 1997.
- [4] Raaymakers, T., Rinkel, G., Limburg, M., and Algra, A., Mortality and morbidity of surgery for unruptured intracranial aneurysms: a metaanalysis., Stroke, Vol. 29, No. 8, pp. 1531–8, Aug. 1998.
- [5] Spetzger, U., Rohde, V., and Gilsbach, J., Clips and coils in aneurysm treatment, Wien Med Wochenschr., Vol. 147, Nos. 7–8, pp. 172–7, 1997.
- [6] Murayama, Y., Nien, Y., Duckwiler, G., Gobin, Y., Jahan, R., Frazee, J., Martin, N., and Vinuela, F., Guglielmi detachable coil embolization of cerebral aneurysms: 11 years' experience, J Neurosurg, Vol. 98, No. 5, pp. 959–66, May 2003.
- [7] Guglielmi, G., Vinuela, F., Spetka, I., and Macellari, V., Electrothrombosis of saccular aneurysms via endovascular approach, Neruosurgery, Vol. 75, No. 1, pp. 1–7, July 1991.
- [8] Debrun, G., Aletich, V., Kehrli, P., Misra, M., Ausman, J., Charbel, F., and Shownkeen, H., Aneurysm geometry: an important criterion in selecting patients for Guglielmi detachable coiling, Neuro. Med. Chir., Vol. 38, pp. 1–20, April 1998.
- [9] Anderson, G., Steinke, D., Petruk, K., Ashforth, R., and Findlay, J., Computed Tomographic Angiography versus Digital Subtraction Angiography for the diagnosis and early treatment of ruptured intracranial aneurysms, Neurosurgery, Vol. 45, No. 6, pp. 1315–20, Dec. 1999.

- [10] White, P., Teasdale, E., Wardlaw, J., and Easton, V., Intracranial aneurysms: CT angiography and MR angiography for detection prospective blinded comparison in a large patient cohort., Radiology, Vol. 219, No. 3, pp. 739–49, June 2001.
- [11] Loncaric, S. and Sorantin, M. S. E., 3-D deformable model for aortic aneurysm segmentation from CT images, Engineering in Medicine and Biology Society, 2000. In: Proceedings of the 22nd Annual International Conference of the IEEE, Vol. 1, pp. 398–401, July 2000.
- [12] Lorigo, L., Faugeras, O., Grimson, W., Keriven, R., Kikinis, R., and Westin, C., Co-dimension 2 Geodesic Active Contours for MRA Segmentation, Lecture Notes in Computer Science, Vol. 1613, pp. 126, 1999.
- [13] Lorigo, L., Faugeras, O., Grimson, W., Keriven, R., Kikinis, R., Nabavi, A., and Westin., C., CURVES: Curve evolution for vessel segmentation, Med. Image. Analysis, Vol. 5, No. 3, pp. 195–206, 2001.
- [14] van Bemmel, C., Spreeuwers, L., Viergever, M., and Niessen, W., Level-Set Based Carotid Artery Segmentation for Stenosis Grading, Medical Image Computing and Computer-Assisted Intervention (MICCAI 2002), Lecture Notes in Computer Science, Vol. 2489, pp. 36–43, 2002.
- [15] van Bemmel, C., Spreeuwers, L., Viergever, M., and Niessen, W., Level-Set-Based Artery-Vein Separation in Blood Pool Agent CE-MR angiograms, IEEE Transactions on Medical Imaging, Vol. 22, No. 10, pp. 1224–1234, Oct. 2003.
- [16] Deschamps, T., Extraction de courbes et surfaces par méthodes de chemins minimaux et ensembles de niveaux. Applications en imagerie medicale 3D, Ph.D. Thesis, University of Paris-Dauphine, France, 2001.
- [17] Frangi, A., Niessen, W., Vincken, K., and Viergever, M., Multiscale Vessel Enhancement Filtering, Lecture Notes in Computer Science, Vol. 1496, pp. 130–137, 1998.
- [18] Zhu, S. and Yuille, A., Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. In: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 9, pp. 884–900, 1996.

- [19] Paragios, N., Geodesic Active Regions and Level Set Methods: Contributions and Applications in Artificial Vision, Ph.D. Thesis, University of Nice Sophia-Antipolis, France, 2000.
- [20] Yezzi, A., Tsai, A., and Willsky, A., A statistical approach to snakes for bimodal and trimodal imagery, In: Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol. 2, pp. 898–903, 1999.
- [21] Sato, Y., Nakajima, S., Shiraga, N., Atsumi, H., Yoshida, S., Koller, T., Gerig, G., and Kikinis, R., Three-dimensional multiscale line filter for segmentation and visualization of curvilinear structures in medical images, Med. Image Anal., Vol. 2, No. 2, pp. 143–168, 1998.
- [22] Dasarathy, B., Nearest Neighbor Pattern Classification Techniques, IEEE Computer Society Press, Los Alamitos, CA, 1990.
- [23] Duda, R., Stork, D., and Hart, P., Pattern Classification and Scene Analysis: Pattern Classification, John Wiley and Sons Inc., Chester, UK 2000.
- [24] Perona, P. and J.Malik, Scale-space and edge detection using anisotropic diffusion, IEEE Trans. Pattern Anal. Machine Intel., Vol. 12, No. 7, pp. 629–639, 1990.
- [25] Sapiro, G., Geometric Partial Differential Equations and Image Analysis, Cambridge University Press, Cambridge, 2001.
- [26] Teo, P., Sapiro, G., and Wandell, B., Segmenting cortical gray matter for functional MRI visualization, In: 6th Int. Conf. Comput. Vision, pp. 292– 297, 1998
- [27] Caselles, V., Kimmel, R., and Sapiro, G., Geodesic Active Contours, Int. J. Comput. Vision, Vol. 22, No. 1, pp. 61–79, 1997.
- [28] Kass, M., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, Int. J. Comput. Vision, Vol. 1, No. 4, pp. 321–331, 1988.
- [29] Osher, S. and Sethian, J., Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations, J. Comp. Phy, Vol. 79, pp. 12–49, 1988.
- [30] ANN library : http://www.cs.umd.edu/~mount/ANN.

- [31] Sethian, J., Level Set Methods and Fast Marching Methods, Cambridge University Press, Cambridge, 2000.
- [32] Osher, S. and Fedkiw, R., Level Set Methods and Dynamic Implicit Surfaces, Springer-Verlag, Berlin 2003.
- [33] Bland, J. and Altman, D., Statistical methods for assessing agreement between two methods of clinical measurement, Lancet, Vol. 8476, pp. 307–10, Feb. 1986.
- [34] Jain, A. and Chandrasekaran, B., Dimensionality and sample size considerations in pattern recognition practice, Handbook of Statistics, P. R. Krishnaiah and L. N. Kanal (Eds) Vol. 2, North Holland, pp. 835–855, 1982.

## **Chapter 6**

## **Inverse Consistent Image Registration**

 $G. E. Christensen^1$ 

## 6.1 Introduction

Image registration has many uses in medicine such as multimodality fusion, image segmentation, deformable atlas registration, functional brain mapping, image guided surgery, and characterization of normal vs. abnormal anatomical shape and variation. The fundamental assumption in each of these applications is that image registration can be used to define a meaningful correspondence mapping between anatomical images collected from imaging devices such as CT, MRI, cryosectioning, etc. It is often assumed that this correspondence mapping or transformation is one-to-one, i.e., each point in image T is mapped to only one point in image S and vice versa. A fundamental problem with a large class of image registration techniques is that the estimated transformation from image T to S does not equal the inverse of the estimated transform from S to T. This inconsistency is a result of the matching criteria's inability to uniquely describe the correspondences between two images. Inverse consistent registration seeks to overcome this limitation by jointly estimating the transformation from T to S and from S to T while minimizing the correspondence inconsistencies between the forward and reverse transformations. Forward and reverse transformations that are inverses of each other are defined to be inverse consistent with each other. The inverse consistency error is a measure of the difference between the forward transformation and the inverse of the reverse transformation, and vica versa.

<sup>&</sup>lt;sup>1</sup>Department of Electrical and Computer Engineering, The University of Iowa, Iowa City, IA 52242 USA



Figure 6.1: Consistent image registration is based on the principle that the mappings *h* from *T* to *S* and *g* from *S* to *T* define a point by point correspondence between *T* and *S* that are consistent with each other. This consistency is enforced mathematically by jointly estimating *h* and *g* while minimizing the inverse consistency error  $(||h - g^{-1}|| + ||g - h^{-1}||)$ . The inverse consistency error is minimized when *h* and *g* are inverse mappings of one another.

The forward transformation h from image T to S and the reverse transformation g from S to T are pictured in Fig. 6.1. Ideally, the transformations h and g should be uniquely determined and should be inverses of one another provided the images only differ in shape and not structure. Estimating h and g independently very rarely results in a consistent set of transformations due to a large number of local minima. To overcome this deficiency in current registration systems, we jointly estimate h and g while minimizing the inverse consistency error defined as  $||h - g^{-1}|| + ||g - h^{-1}||$ . Notice that the inverse consistency error is minimized when the transformations h and g are inverse mappings of one another. Jointly estimating the forward and reverse transformations provides additional correspondence information and helps ensure that these transformations define a consistent correspondence between the images. Although uniqueness is very difficult to achieve in medical image registration, the joint estimation should lead to more consistent and biologically meaningful results since information from one registration direction minimizes registration ambiguities in the other direction.

Image registration algorithms use landmarks [1–4], contours [5–7], surfaces [8–11], volumes [6, 12–21], or a combination of these features [22] to manually, semi-automatically or automatically define correspondences between two images. The need to impose the invertibility consistency constraint depends on the particular application and on the correspondence model used for registration. In general, registration techniques that do not uniquely determine the correspondence between image volumes should benefit from the consistency constraint.

This is because such techniques often rely on minimize/maximize a similarity measure which has a large number of local minima/maxima due to the correspondence ambiguity. Examples include methods that minimizing/maximizing similarity measures between features in the source and target images such as image intensities, object boundaries/surfaces, etc. In theory, the higher the dimension of the transformation the more local minima these similarity measures have. Methods that use specified correspondences for registration will benefit less or not at all from the invertibility consistency constraint. For example, landmark based registration methods implicitly impose an invertibility constraint at the landmarks because the correspondence defined between landmarks is the same for estimating the forward and reverse transformations. However, the drawbacks of specifying correspondences can not always be specified, and such methods usually only provide coarse registration due to the small number of correspondences specified.

In this chapter, we will restrict our analysis to the class of applications that can be solved using diffeomorphic transformations. A diffeomorphic transformation is defined to be continuous, one-to-one, onto, and differentiable. The diffeomorphic restriction is valid for a large number of problems in which the two images have the same structures and neighborhood relationships but have different shapes.

Diffeomorphic transformations maintain the topology and guarantee that connected subregions of an image remain connected, neighborhood relationships between structures are preserved, and surfaces are mapped to surfaces. Preserving topology is important for synthesizing individualized electronic atlases the knowledge base of the atlas maybe transferred to the target anatomy through the topology preserving transformation providing automatic labeling and segmentation. If the total volume of a nucleus, ventricle, or cortical subregion are an important statistic it can be generated automatically. Topology preserving transformations that map the template to the target can also be used to study the physical properties of the target anatomy such as mean shape and variation. Likewise, preserving topology allows data from multiple individuals to be mapped to a standard atlas coordinate space [23]. Registration to an atlas removes individual anatomical variation and allows information from many experiments to be combined and associated with a single canonical anatomy. Other investigators have proposed methods for enforcing pairwise consistent transformations. For example, Woods *et al.* [24] computes all pairwise registrations of a population of image volumes using a linear transformation model, i.e., a  $3 \times 3$  matrix transformation. They then average the transformation from T to S with all the transformations from T to X to S. The original transformation from T to S is replaced with average transformation. The procedure is repeated for all the image pairs until convergence. This technique is limited by the fact that it can not be applied to two data sets. Also, there is no guarantee that the generated set of consistent transformations are valid. For example, a poorly registered pair of images can adversely affect all of the pairwise transformations.

The method described in this chapter is most similar to the approach described by Thirion [6]. Thirion's idea was to iteratively estimate the forward h, reverse g, and residual  $r = h \circ g$  transformations in order to register the images T and S. At each iteration, half of the residual r is added to h and half of the residual r is mapped through h and added to g. After performing this operation,  $h \circ g$  is close to the identity transformation. The advantage of Thirion's method is that it enforces the inverse consistency constraint without having to explicitly compute the inverse transformations as in Eq. (6.2). The residual method is an approximation to the inverse consistency method in that the residual method approximates the correspondences between the forward and reverse transformations while the inverse consistency method computes those correspondences. Thus, the residual approach only works under a small deformation assumption since the residual is computed between points that do not correspond to one another. This drawback limits the residual approach to small deformations and it therefore can not be extended to nonlinear transformation models. On the other hand, the approach presented in this paper can be extended to the nonlinear case by modifying the procedure used to calculate the inverse transformation to include nonlinear transformations.

## 6.2 Inverse Consitent Image Registration

#### 6.2.1 Problem Statement

Assume that T and S correspond to two continuous images defined on the coordinate system  $\Omega = [0, 1)^3$ . Traditionally, the image registration problem has been stated as: Find the transformation  $h: \Omega \to \Omega$  that maps the template image volume *T* into correspondence with the target image volume *S*. Alternatively, the problem can be stated as: Find the transformation  $g: \Omega \to \Omega$  that transforms *S* into correspondence with *T*. For inverse consistent registration, the previous two statements are combined into a single problem and restated as:

**Problem Statement:** Jointly estimate the transformations *h* and *g* such that *h* maps *T* to *S* and *g* maps *S* to *T* such that the inverse consistency constraint  $||h - g^{-1}|| + ||g - h^{-1}||$  is minimized.

The image volumes T and S can be of any dimension such as 1D, 2D, 3D, 4D, or higher dimensional and in general can be multivalued. Image data sets may represent information such as anatomical structures like the brain, heart, lungs, etc., or could represent symbolic information such as structure names, object features, curvature, brain function, etc., or could represent image frames in movies that need to be matched for morphing, interpolating transitional frames, etc., or images of a battlefield with tanks, artillery, etc., or images collected from satellites or robots that need to be fused into a composite image, etc.

The transformations are vector-valued functions that map the image coordinate system  $\Omega$  to itself, i.e.,  $h: \Omega \mapsto \Omega$  and  $g: \Omega \mapsto \Omega$ . Regularization constraints are placed on h and g so that they preserve topology. Throughout it is assumed that h(x) = x + u(x),  $h^{-1}(x) = x + \tilde{u}(x)$ , g(x) = x + w(x) and  $g^{-1}(x) = x + \tilde{w}(x)$  where  $h(h^{-1}(x)) = x$  and  $g(g^{-1}(x)) = x$ . The vector-valued functions  $u, w, \tilde{u}$ , and  $\tilde{w}$  are called displacement fields since they define the transformation in terms of a displacement from a location x. All of the functions  $h, g, h^{-1}, g^{-1}, u, \tilde{u}, w$ , and  $\tilde{w}$  are  $(3 \times 1)$  vector-valued functions defined on the  $\Omega$ .

Registration is defined using a symmetric similarity cost function that describes the distance between the transformed template  $T \circ h$  and target S, and the distance between the transformed target  $S \circ g$  and template T. To ensure the desired properties, the transformations h and g are jointly estimated by minimizing the similarity cost function while satisfying regularization constraints and inverse transformation consistency constraints. The regularization constraints can be enforced on the transformations by constraining them to satisfy the laws of continuum mechanics [25]. The image registration problem can be stated mathematically as finding the transformations h and g that minimize the cost function

$$\begin{split} C(\mu,\eta) &= \int_{\Omega} |T(x+u(x)) - S(x)|^2 + |S(x+w(x)) - T(x)|^2 dx \\ &+ \chi \int_{\Omega} ||h(x) - g^{-1}(x)||^2 + ||g(x) - h^{-1}(x)||^2 dx \\ &+ \rho \int_{\Omega} ||Lu(x)||^2 dx + ||Lw(x)||^2 dx \end{split}$$

The constants  $\sigma$ ,  $\chi$  and  $\rho$  are used to enforce/balance the constraints (see [26] for complete details on how to minimize this cost function).

#### 6.2.2 Symmetric Similarity Cost Function

A problem with many image registration techniques is that the image similarity function does not uniquely determine the correspondence between two image volumes. In general, similarity cost functions have many local minima due to the complexity of the images being matched and the dimensionality of the transformation. It is these local minima (ambiguities) that cause the estimated transformation from image T to S to be different from the inverse of the estimated transformation from S to T. In general, this becomes more of a problem as the dimensionality of the transformation increases.

To overcome correspondence ambiguities, the transformations from image T to S and from S to T are jointly estimated. This is accomplished by defining a cost function to measure the shape differences between the deformed image  $T \circ h$  and image S and the differences between the deformed image  $S \circ g$  and image T. Ideally, the transformations h and g should be inverses of one another, i.e.,  $h = g^{-1}$ . In this work, the transformations h and g are estimated by minimizing a cost function

$$C_{\text{SIM}}(T \circ h, S) + C_{\text{SIM}}(S \circ g, T) = \int_{\Omega} |T(h(x)) - S(x)|^2 dx + \int_{\Omega} |S(g(x)) - T(x)|^2 dx$$
(6.1)

where the intensities of T and S are scaled between 0 and 1. To use this similarity function, the images T and S must correspond to the same imaging modality and they may require preprocessing to equalize the intensities of the image. In practice, MRI images require intensity equalization while CT images do not. A

simple but effective method for intensity equalizing MRI data is to compute the histograms of the two images, scale the axis of one histogram so that the grayand white-matter maximums match, and then apply the intensity scaling to the image.

This joint estimation approach applies to both linear and non-linear transformations. In general, the squared-error similarity functions in Eq. (6.1) can be replaced by any suitable similarity function—mutual information [27, 28], demons [6], an intensity variance cost function [24], etc.—where the choice is dependent on the particular registration application.

#### 6.2.3 Inverse Consistency Constraint

Minimizing a symmetric cost function like Eq. (6.1) is not sufficient to guarantee that h and g are inverses of each other because the contributions of h and g to the cost function are independent. In order to couple the estimation of h with that of g, an inverse consistency constraint is imposed that is minimized when  $h = g^{-1}$ . The inverse consistency constraint is given by

$$C_{\rm ICC}(u,\tilde{w}) + C_{\rm ICC}(w,\tilde{u}) = \int_{\Omega} ||h(x) - g^{-1}(x)||^2 dx + \int_{\Omega} ||g(x) - h^{-1}(x)||^2 dx$$
$$= \int_{\Omega} ||u(x) - \tilde{w}(x)||^2 dx + \int_{\Omega} ||w(x) - \tilde{u}(x)||^2 dx. \quad (6.2)$$

Notice that the inverse consistency constraint is written in a symmetric form like the symmetric cost function for similar reasons.

#### 6.2.4 Computation of the Inverse Transformation

The procedure used to compute the inverse transformation of a transformation with minimum Jacobian greater than zero is as follows. Assume that h(x) is a continuously differentiable transformation that maps  $\Omega$  onto  $\Omega$  and has a positive Jacobian for all  $x \in \Omega$ . The fact that the Jacobian is positive at a point  $x \in \Omega$  implies that it is locally one-to-one and therefore has a local inverse. It is therefore possible to select a point  $y \in \Omega$  and iteratively search for a point  $x \in \Omega$  such that ||y - h(x)|| is less than some threshold provided that the initial guess of x is close to the final value of x.

The inverse transformation is computed in the following way [26]. First, note that all images including transformed images are discrete. Therefore, it is

only necessary to compute the transformations and the inverse of the transformations at the discrete voxel locations. Let  $\Omega_d$  denote the discrete center locations of the voxels in the coordinate system  $\Omega$ . The discrete inverse transformation is computed using the following procedure only at the discrete voxel points.

```
For each y \in \Omega_d do {

Set \delta = [1, 1, 1]^T, x = y, iteration = 0.

While (||\delta|| > threshold) do {

\delta = y - h(x)

x = x + \frac{\delta}{2}

iteration = iteration + 1

if (iteration > max_iteration) then

Report algorithm failed to converge and exit.

}

h^{-1}(y) = x
```

The threshold is typically set between  $10^{-2}$  and  $10^{-4}$  and the maximum number of iterations is set to 1000. In practice, the algorithm converges when the minimum Jacobian of h is greater than zero although we have not proved this mathematically. Reducing the value of the threshold gives a more accurate inverse but increases the iteration time. This algorithm normally converges quickly and is computationally efficient. However, this algorithm has a tendency to get stuck in osciations and is detected by the if statement. The inverse at these oscillatory points can be estimated using gradient descent to solve the minimization ||y - h(x)|| for x keeping y fixed. Alternatively, the failure of the algorithm to converge at a point can be ignored since it will not have a significant effect on the registration and will be corrected at the next iteration.

#### 6.2.5 Regularization Constraint

Minimizing the cost function in Eq.(6.2) does not ensure that the transformations h and g are diffeomorphic transformations except for when  $C_{\rm ICC} = 0$ . Continuum mechanical models such as linear elasticity [22, 29] and viscous fluid [15, 22] can be used to regularize the transformations. For example, a linear-elastic constraint has the form

$$C_{\text{REG}}(u) + C_{\text{REG}}(w) = \int_{\Omega} ||Lu(x)||^2 dx + \int_{\Omega} ||Lw(x)||^2 dx$$
(6.3)

and can be used to regularize the transformations. The linear elasticity operator *L* has the form  $Lu(x) = -\alpha \nabla^2 u(x) - \beta \nabla (\nabla \cdot u(x)) + \gamma u(x)$  where  $\nabla = \left[\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right]$  and  $\nabla^2 = \nabla \cdot \nabla = \left[\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}\right]$ . In general, *L* can be any nonsingular linear differential operator [30]. The limitation of using linear differential operators is that they can't prevent the transformation from folding onto itself, i.e., destroying the topology of the images under transformation [31]. This includes the linear elasticity and thin-plate spline models. The linear elasticity operator is used in this work to help prevent the Jacobian of the transformation from going negative. At each iteration the Jacobian of the transformation is checked to make sure that it is positive for all points in  $\Omega_d$  which implies that the transformation preserves topology when transforming images.

The purpose of the regularization constraint is to ensure that the transformations maintain the topology of the images T and S. Thus, the elasticity constraint can be replaced by or combined with other regularization constraints that maintain desirable properties of the template (source) and target when deformed. An example would be a constraint that prevented the Jacobian of both the forward and reverse transformations from going to zero or infinity. A constraint that penalizes small and large Jacobian values is given by  $C_{\text{Jac}}(h) + C_{\text{Jac}}(g) =$  $\int_{\Omega} (J(h(x)))^2 + (\frac{1}{J(h(x))})^2 + (J(g(x)))^2 + (\frac{1}{J(g(x))})^2 dx$  where J denotes the Jacobian operator. Further examples of regularization constraints that penalize large and small Jacobians can be found in Ashburner *et al.* [21].

#### 6.2.6 Transformation Parameterization

Until now, the forward and reverse transformations have been described as general functions. In order to estimate the transformations, they must be parameterized. Examples of transformation parameterizations that are used in practice include the 3D Fourier series [26], polinomials [32], b-splines [19,24], wavelets [17], and vector displacements [14, 15]. We will concentrate on a 3D Fourier series parameterization in this chapter. In a 3D Fourier series parameterization, each basis coefficient is interpreted as the weight of a harmonic component in a single coordinate direction. The discretized displacement fields

are given by

$$u_d[n] = \sum_{k \in G} \mu[k] e^{j\langle n, \theta[k] \rangle} \quad \text{and} \quad w_d[n] = \sum_{k \in G} \eta[k] e^{j\langle n, \theta[k] \rangle} \tag{6.4}$$

for  $n \in G$  and  $G = \{(n_1, n_2, n_3) : 0 \le n_1 < N_1, 0 \le n_2 < N_2, 0 \le n_3 < N_3\}$ . The displacement fields associated with the inverse of the forward and reverse transformations are given by replacing u, w,  $\mu$ , and  $\eta$  in Eq. (6.4) with  $\tilde{u}$ ,  $\tilde{w}$ ,  $\tilde{\mu}$ , and  $\tilde{\eta}$ , respectively. The Fourier series parameterization is periodic and therefore imposes cyclic boundary conditions on the boundary of  $\Omega$ .

#### 6.2.7 Multiresolution Registration

Multiresolution formulation of the registration problem has the benefit of minimizing computation time and helps to avoid local registration minima. The Fourier series parameterization in Eq. (6.4) is an example of a multiresolution decomposition of the displacement fields in parameter space. Let  $G[r] = G \setminus G[r]$ represent a family of subsets of  $\Omega_d$  where  $G[r] = \{n \in G | r_1 < n_1 < N_1 - r_1; r_2 < n_2 < N_2 - r_2; r_3 < n_3 < N_3 - r_3\}$  and the set subtraction notation  $A \setminus B$  is defined as all elements of A not in B. In practice, the low frequency basis coefficients are estimated before the higher ones allowing the global image features to be registered before the local features. This is accomplished by replacing Eq. (6.4) by

$$u_d[n,r] = \sum_{k \in \Omega_d[r]} \mu[k] e^{j\langle n,\theta[k] \rangle} \quad \text{and} \quad w_d[n,r] = \sum_{k \in \Omega_d[r]} \eta[k] e^{j\langle n,\theta[k] \rangle}.$$
(6.5)

where  $r \in G$  determines the number of harmonics used to represent the displacement fields. The components of r are initially set small and are periodically increased throughout the iterative minimization procedure. The set G[r]can be replaced by G when all of the components of r are greater than or equal to (N-1)/2 since the set G[r] is empty. The constants  $r_1$ ,  $r_2$ , and  $r_3$  represent the largest  $x_1$ ,  $x_2$ , and  $x_3$  harmonic components of the displacement fields. Each displacement field in Eq. (6.5) is efficiently computed using three  $N_1 \times N_2 \times N_3$ FFTs, i.e., each component of the  $3 \times 1$  vectors  $u_d$  and  $w_d$  are computed with a FFT after zeroing out the coefficients not present in the summations.

The approach of spatial multiresolution is to register two images at a course spatial resolution initially and then to refine the registration at a higher spatial resolution later to get the final transformation. The advantage of solving the problem on a course grid is that the algorithm requires fewer computations per iteration than a finer grid. This results in reduced computation time at low resolution. Each time the resolution of the grid is increased by a factor of two in each dimension, the computation time increases by a factor of eight. The drawback of solving the problem at low resolution is that there can be significant registration errors due to the loss of detail in the down sampling procedure. The trade-off between quicker execution times at low resolution and more accurate registration at higher resolution can be exploited by solving the registration problem at low spatial resolution during the initial iterations to approximate the result and then increasing the spatial resolution to get a more accurate result at the later iterations.

The spatial multiresolution approach works well with the frequency multiresolution approach provided by increasing the number of harmonics used to represent the displacement fields. The number of harmonics used to represent the displacement fields is initially set small and then increased as the number of iterations are increased. A low-frequency registration result is an approximation of the desired high-frequency registration result. Computing the gradient descent for a low-frequency basis coefficient at low spatial resolution gives approximately the same answer as using high spatial resolution but the computational burden is much less.

# 6.2.8 Tracking the Jacobian During the Estimation Procedure

It is important to track both the minimum and maximum values of the Jacobian during the estimation procedure. The Jacobian measures the differential volume change of a point being mapped through the transformation. At the start of the estimation, the transformation is the identity mapping and therefore has a Jacobian of one. If the minimum Jacobian goes negative, the transformation is no longer a one-to-one mapping and as a result folds the domain inside out [31]. Conversely, the reciprocal of the maximum value of the Jacobian corresponds to the minimum value of the Jacobian goes to infinity, the minimum value of the Jacobian of the inverse mapping. Thus, as the maximum value of the inverse mapping goes to zero.



Figure 6.2: Notation used to describe transformations from one coordinate system to another.

## 6.3 Transformation Properties

The following definitions will be used throughout this section. Let  $T_i$  for  $i \in Q = \{A, B, C, \ldots\}$  denote a set of homogeneous, topologically-equivalent anatomical images defined on the coordinate system or domain  $\Omega = [0, 1]^3$ . For example,  $T_A, T_B$ , etc., may correspond to a set of 3D MRI brain images collected from age and sex matched normals or abnormals or some other suitable classification criteria. Let  $h_{AB}$  represent the transformation from the coordinate system of image  $T_A$  to that of image  $T_B$  in terms of the coordinate system of image  $T_B$  as shown in Fig. 6.2. Let the linear transformation  $x = h_{AB}(y)$  deform image  $T_A(x)$  into a new image  $\tilde{T}(y) = T_A(h_{AB}(y))$  that resembles the shape of image  $T_B(y)$  by transforming the coordinate system of image  $T_A$  to that of image  $T_B$ . Define H as the set of all transformations  $h_{AB}(x)$  for  $A, B \in Q$  and  $x \in \Omega$ . Let  $||x|| = \sqrt{x_1^2 + x_2^2 + x_3^2}$  denote the standard 2-norm.

#### 6.3.1 Invertibility Property

Many nonlinear image registration algorithms have difficulty producing inverse consistent transforms because they use a finite set of basis functions (eigenfunctions of an operator, polynomials, etc.) that are not always closed under composition. This observation is a major reason why it is important to measure the inverse consistency error produced by different registration algorithms. This fact motivates the minimization of the inverse consistency constraint error that is used by the consistent linear elastic registration error since it is not possible to reduce the inverse consistency error to zero when using the a finite set of complex exponential basis functions.

Another reason why it is difficult to produce inverse consistent transformations is that numerical optimization techniques used to find the optimal image transformation often take a long time to converge or get stuck in local minima. The large number of parameters estimated and the nonlinearity introduced by the images being mapped makes it very difficult to find the optimal transformation. Placing a limit on the acceptable inverse consistency error may be one way of specifying the stopping criteria for a particular optimization technique.

Formally, transformations  $h_{AB}: \Omega \to \Omega$  and  $h_{BA}: \Omega \to \Omega$  are said to be inverses of one another if the transformation  $h_{BA}$  exists and satisfies  $h_{AB}(h_{BA}(x)) = x$  and  $h_{BA}(h_{AB}(x)) = x$  for all  $x \in \Omega$ . A set of linear transformations *H* is said to have the invertibility property if  $h_{AB}(h_{BA}(x)) = x$  for all  $A, B \in Q$  and  $x \in \Omega$ .

The average Inverse Consistency (IC) error within a region of interest (ROI) M is defined as

$$E_{AIC}(h_{AB}, h_{BA}, M) = \frac{1}{M} \int_{M} ||h_{AB}(h_{BA}(x)) - x|| dx$$
(6.6)

and the maximum IC error is defined as

$$E_{MIC}(h_{AB}, h_{BA}, M) = \max_{x \in M} ||h_{AB}(h_{BA}(x)) - x||.$$
(6.7)

Eqs. (6.6 and 6.7) are discretized for implementation.

It is important to define a ROI because the amount of padding applied to the image data can have a significant effect on the average error calculation. The ROI restricts the error measurements to areas of interest preventing the situation where the largest error occurs in the background of the image.

#### 6.3.2 Transitivity Property

Image registration algorithms that have a difficult time producing inverse consistent transformations have an even harder time producing transformations that satisfy the transitivity property. In the paper we investigate how an algorithm that reduces the inverse consistency error compared to another also reduces the transitivity error.

A set of image transformations *H* is said to have the transitivity property if  $h_{CB}(h_{BA}(x)) = h_{CA}(x)$  or equivalently if  $h_{AC}(h_{CB}(h_{BA}(x))) = x$  for all *A*, *B*, *C*  $\in$  *Q* and  $x \in \Omega$ .

These transitivity relationships are illustrated in Fig. 6.2. Assume that the points x, y, and z correspond to the same landmark in images A, B, and C, respectively. Assume that the set of transformations  $H = \{h_{AB}, h_{BA}, h_{BC}, h_{CB}, h_{AC}, h_{CA}\}$  has the invertibility and transitivity properties such that

$$y = h_{BA}(x), \quad z = h_{CB}(y), \quad x = h_{AC}(z).$$

Substituting the first equation into the second and the second into the third equation gives the result

$$x = h_{AC}(h_{CB}(h_{BA}(x)))$$

The average transitivity error is defined as

$$E_{ATRAN}(h_{AB}, h_{BC}, h_{CA}, M) = \frac{1}{M} \int_{M} ||h_{AB}(h_{BC}(h_{CA}(x))) - x|| dx$$
(6.8)

and the maximum transitivity error is defined as

$$E_{MTRAN}(h_{AB}, h_{BC}, h_{CA}, M) = \max_{x \in M} ||h_{AB}(h_{BC}(h_{CA}(x))) - x||.$$
(6.9)

Equations (6.8) (6.9) are discretized for implementation.

Figure 6.3 demonstrates an advantage of producing transformations that satisfy the transitivity property. The left panels show that the minimum number of invertible transformations required to map information from one coordinate system to another is N - 1 where N is the number of image volumes. The



Figure 6.3: The left panel shows the minimum number of pairwise transformations needed to map a point from one brain to its corresponding location in another. The right panel shows all of the pairwise mappings between the brains.

correspondence between any two coordinate systems is determined explicitly by one of the displayed transformations or indirectly by concatenating two of the transformations. For example, a point x in coordinate system B is mapped to y in coordinate system C by the mapping  $y = h_{BA}(h_{AC}(x))$ , etc.

Figure 6.3 demonstrates that it is advantageous to design pairwise registration algorithms rather than *N*-wise registration algorithms that satisfy the transitivity property. The first advantage is that a pairwise algorithm only needs to compute N - 1 pairwise transformations as opposed to (N - 1)! pairwise transformations. This reduces computation time and computer storage requirements by a factor of (N - 2) factorial. Another advantage is that a pairwise algorithm only requires one additional set of pairwise transformations to be computed to add a new data set to the population. An *N*-wise registration algorithm requires that all of the transformations to be recomputed to produce a set of transformations with the transitivity property.

In general, pairwise image registration algorithms do not produce transformations that have the transitivity property. The degree of transitivity can be evaluated by measuring the difference between the identity mapping and the composition the transformations from image *A*-to-*B*, *B*-to-*C*, and *C*to-*A*.

## 6.4 Inverse Consistent Registration Algorithms

## 6.4.1 Intensity-Based Small Deformation Inverse Consistent Linear Elastic (I-SICLE) Image Registration

The intensity-based small deformation inverse consistent linear elastic (I-SICLE) image registration algorithm [26,33–35] jointly estimates the forward and reverse transformations h and g by minimizing Eq. (6.1). The image intensity alone is used to register the images with this algorithm. A small deformation linear elastic continuum mechanical model is used to regularize the transformations. A generalization of this technic allows for multiple template image modalities or subvolumes to be registered to corresponding target modalities. The cost

function for the multimodality I-SICLE algorithm is given by

$$C = \sum_{i=1}^{l} \sigma_{i} \int_{\Omega} |T_{i}(h(x)) - S_{i}(x)|^{2} + |S_{i}(g(x)) - T_{i}(x)|^{2} dx + \rho \int_{\Omega} ||Lu(x)||^{2} + ||Lw(x)||^{2} dx$$
(6.10)  
+  $\chi \int_{\Omega} ||u(x) - \tilde{w}(x)||^{2} + ||w(x) - \tilde{u}(x)||^{2} dx$ 

where  $\sigma_i$  are relative weighting factors for each of imaging modalities and  $\rho$  and  $\chi$  define the relative importance of the bending energy minimization and the inverse consistency terms. The constants  $\sigma_i$ , define the relative importance of each modality with respect to the regularization terms of the cost function. The I-SICLE algorithm has been applied to registration of brain images [35], skull images [26] lung images [36, 37], and inner ear images [38].

## 6.4.2 Inverse Consistent Landmark-Based Thin-Plate Spline (CL-TPS) Image Registration

Before describing the inverse consistent landmark-based, thin-plate spline (CL-TPS) image registration algorithm, we discuss the traditional unidirection thin-plate spline registration algorithm. The unidirectional landmark-based, thin-plate spline (UL-TPS) image registration algorithm [1,2,39] registers a template image T with a target image S by matching corresponding landmarks identified in both images. Registration at non-landmark points is accomplished by interpolation such that the overall transformation smoothly maps the template into the shape of the target image.

The unidirectional landmark image registration problem can be thought of as a Dirichlet problem [40] and can be stated mathematically as finding the displacement field u that minimizes the cost function

$$C = \int_{\Omega} ||Lu(x)||^2 dx \tag{6.11}$$

subject to the constraints that  $u(p_i) = q_i - p_i$  for i = 1, ..., M where  $p_i$  and  $q_i$  are corresponding landmarks in the target and template images, respectively. The operator *L* denotes a symmetric linear differential operator [41] and is used to interpolate the displacement field *u* between the corresponding landmarks.

When  $\mathcal{L} = \nabla^2$ , the problem reduces to the thin-plate spline image registration problem given by

$$C = \int_{\Omega} ||\nabla^2 u(x)||^2 dx = \sum_{i=1}^2 \int_{\Omega} \left(\frac{\partial^2 u_i(x)}{\partial^2 x_1}\right)^2 + 2\left(\frac{\partial^2 u_i(x)}{\partial x_1 \partial x_2}\right) + \left(\frac{\partial^2 u_i(x)}{\partial^2 x_2}\right)^2 dx_1 dx_2$$
(6.12)

subject to the constraints that  $u(p_i) = q_i - p_i$  for i = 1, ..., M.

It is well known [1, 2, 39] that the thin-plate spline displacement field u(x) that minimizes the bending energy defined by Eq. (6.12) has the form

$$u(x) = \sum_{i=1}^{M} \xi_i \phi(x - p_i) + Ax + b$$
(6.13)

where  $\phi(r) = r^2 \log r$  and  $\xi_i$  are  $2 \times 1$  weighting vectors. The  $2 \times 2$  matrix  $A = [a_1, a_2]$  and the  $2 \times 1$  vector b define the affine transformation where  $a_1$  and  $a_2$  are  $2 \times 1$  vectors.

The thin-plate spline interpolant  $\phi(r) = r^2 \log r$  is derived assuming infinite boundary conditions, i.e.,  $\Omega$  is assumed to be the whole plane  $R^2$ . The thin-plate spline transformation is truncated at the image boundary when it is applied to an image. This presents a mismatch in boundary conditions at the image edges when comparing forward and reverse transformations between two images. It also implies that a thin-plate spline transformation is not a one-to-one and onto mapping between two image spaces. To overcome this problem and to match the periodic boundary conditions assumed by the intensity-based consistent image registration algorithm, approximate periodic boundary conditions are imposed on the registration problem (see [34] for details).

The inverse consistent landmark-based, thin-plate spline (CL-TPS) image registration algorithm is solved by minimizing the cost function given by

$$C = \rho \int_{\Omega} ||\mathcal{L}u(x)||^{2} + ||\mathcal{L}w(x)||^{2} dx + \chi \int_{\Omega} ||u(x) - \tilde{w}(x)||^{2} + ||w(x) - \tilde{u}(x)||^{2} dx \text{subject to } p_{i} + u(p_{i}) = q_{i} \text{ and } q_{i} + w(q_{i}) = p_{i} \text{ for } i = 1, \dots, M$$
(6.14)

The first integral of the cost function defines the bending energy of the thinplate spline for the displacement fields u and w associated with the forward and reverse transformations, respectively. This term penalizes large derivatives of the displacement fields and provides the smooth interpolation away from the landmarks. The second integral is the inverse consistency constraint (ICC) and is minimized when the forward and reverse transformations are inverses of one another. This integral couples the estimation of the forward and reverse transformations together and penalizes transformations that are not inverses of one another. The constants  $\rho$  and  $\chi$  define the relative importance of the bending energy minimization and the inverse consistency terms of the cost function.

The cost function in Eq. (6.14) is iteratively minimized until the landmark error and the inverse consistency error fall below problem specific thresholds or until a specified number of iterations are reached. In practice, this algorithm converges to an acceptable solution within five to 10 iterations and therefore we use a maximum number of iterations as our stopping criteria. See [34] for more details of this algorithm.

## 6.4.3 Combined Intensity and Feature-Based Inverse Consistent Registration

Landmark based registration algorithms provide good registration at landmark points where correspondence is known, but use interpolation away from the landmarks to define correspondence. The correspondence defined by the interplolation function does not always give acceptable correspondence away from the landmarks. On the other hand, intensity based registration provides good registration of intensity features contained in the images. However, intensity based correspondence funtions provide correspondence without regard to the structure of the objects being matched causing noncorresponding structures to be registered. Combining landmark and information from other features such as contours, surfaces, and subvolumes with intensity information helps avoid registration errors in uncertain or ambiguous areas of the respective cost functions. For example, landmarks are good at getting corresponding points registered and the intensity based cost function is good at registering points in between landmarks. In general, the more information that the registration algorithm has to define correspondences, the better the registration result will be. Examples of inverse consistent image registration combining landmark, subvolume, and intensity information can be found in [34-38].

## 6.5 Results

#### 6.5.1 Landmark Registration

The first experiment compares the inverse consistency error associated with the traditional unidirectional landmark thin-plate spline (UL-TPS) algorithm to that of the consistent landmark thin-plate spline (CL-TPS) algorithm. This simple experiment is designed to show that the UL-TPS algorithm can have significant inverse consistency error while this error is minimized using the CL-TPS algorithm. The experiment shown in Fig. 6.4 consisted of matching eight landmarks in one image to their corresponding landmarks in a second image using both



Figure 6.4: The location of local displacements at the landmarks points for the forward, and reverse transformations of images with  $100 \times 100$  pixels. Application of the thin-plate spline deformation fields to uniformly spaced grids for the forward and reverse transformations.

the UL-TPS and the CL-TPS algorithm. The arrows in the first and second panels show the displacement between the corresponding landmarks in the forward and reverse directions, respectively. The four landmarks in the corners of the images were fixed. The forward transformation h maps the four inner points to the four outer points and the reverse transformation g maps the outer points to the inner points. Applying the CL-TPS transformations to a rectangular grid shows that the forward transformation—defined with respect to a Eulerian frame of reference—causes the center of the image to expand (third panel of Fig. 6.4) while the reverse transformation causes a contraction of the central portion of the image (fourth panel of Fig. 6.4).

The top row of Fig. 6.5 shows the spatial locations and magnitudes of the inverse consistency errors of the forward and reverse transformations generated by the UL-TPS algorithm. The images in the left column were computed by taking the Euclidean norm of the difference between the forward transformation h and the inverse of the reverse transformation  $g^{-1}$ . The images in the center column were computed in a similar fashion with g and  $h^{-1}$ . The CL-TPS result



Figure 6.5: The left and center panels show the inverse consistency errors of the forward and reverse transformations, respectively. The tables in the right columns list the landmark errors associated with selected image points. The top and bottom rows are the inverse consistency errors associated with the unidirectional (UL-TPS) and consistent (CL-TPS) landmark thin-plate spline algorithms, respectively.

was created using AUL-TPS initialization and minimizing for 100 iterations with  $\alpha = 0.5$  and  $\beta = 0.012$ .

The tables in Fig. 6.5 tabulate the inverse consistency error at four representative points in the images. The points A and C are located at points away from landmarks while the points B and D are located at landmark locations. The inverse consistency error at the landmark points is small for both algorithms. However, the landmark error is quite large away from the landmark locations in the UL-TPS algorithm. The range of intensities on the color bar for each method shows that the range of inverse consistency errors for the UL-TPS algorithm was in the range of 0.002 to 4.9 pixels while this same error for the CL-TPS algorithm ranged from 0.00 to 0.009. This shows that the CL-TPS algorithm reduced the inverse consistency error by over 500 times that of the UL-TPS algorithm for this example.

A pair of transformations are point-wise consistent if the composite function h(g(x)) maps a point x to itself. Spatial deviations from the identity mapping can be visualized by applying the composite mapping to a uniformly spaced grid. The grid is deformed by the composite transformation in regions where the forward and reverse transformations have inverse consistency errors. The composite transformation does not deform the grid for a perfectly inverse consistent set of forward and reverse transformations. Fig. 6.6 shows the composite mapping



Figure 6.6: Deformed grids showing the error between the forward and reverse transformations estimated with the landmark-based thin-plate spline algorithm (left panel) and the CL-TPS algorithm (right panel). The grids were deformed by the transformation constructed by composing the forward and reverse transformations together, i.e., g(h(x)). Ideally, the composition of the forward and reverse transformations is the identity mapping which produces no distortion of the grid as in the right panel. The fuzziness associated with the grids are due to the bilinear interpolation.

produced by the UL-TPS (left) and the CL-TPS (right) applied to a rectangular grid for this experiment. Notice that there is a considerable amount of inverse consistency error in the UL-TPS algorithm while there is no visually detectable inverse consistency error produced by the CL-TPS algorithm. The blurring of the grid is due to bilinear interpolation used to deform the grid images with the error displacements. Both images are created with the same technique, but the inverse consistent image needs very little interpolation since there is nearly zero displacement error.

The minimum and maximum Jacobian values of the forward (reverse) transformation specify the maximum expansion and contraction of the transformation, respectively. The Jacobian error, calculated as  $\frac{1}{2}|\min\{Jac(h)\} - 1/\max\{Jac(g)\}| + \frac{1}{2}|\min\{Jac(g)\}| - 1/\max\{Jac(h)\}|$ , provides an indirect measure of the inconsistency between the forward and reverse transformations. The Jacobian error is zero if the forward and reverse transformations are inverses of one another, but the converse is not true. Table 6.1 shows that the Jacobian error was 1000 times smaller for the CL-TPS algorithm compared to the UL-TPS algorithm.

#### 6.5.2 Landmark and Intensity Registration

The five 2D transverse MRI data sets shown in Fig. 6.7 were used to compare the performance of the unidirectional landmark (UL-TPS); consistent landmark (CL-TPS); consistent intensity (CI-TPS); and consistent landmark and intensity

Table 6.1: Comparison between the unidirectional (UL-TPS) and consistent (CL-TPS) thin-plate spline image registration algorithms. The table columns are the Experiment, (ICC), transformation Direction (TD), average landmark error (ALE) in pixels, maximum landmark error (MLE), maximum inverse error (MIE) in pixels, average inverse error (AIE) in pixels, minimum Jacobian value (MJ), inverse of the maximum Jacobian value (IJ), and the Jacobian error (JE)

Experiment	ICC	TD	ALE	MLE	AIE	MIE	MJ	IJ	JE
UL-TPS	No	Forward Reverse	0.010 0.0056	0.016 0.010	2.2 2.0	4.1 4.9	2.4 2.9	4.8 3.2	1.4
CL-TPS (100 iter.)	Yes	Forward Reverse	0.00055 0.00046	0.0011 0.00094	$0.0028 \\ 0.0024$	$0.0078 \\ 0.0088$	0.28 0.48	0.48 0.28	0.0012


Figure 6.7: Five corresponding image slices from MRI acquired brains with manually identified points of correspondence.

(CLI-TPS) thin-plate spline algorithms. These  $256 \times 320$  pixel images with 1 mm isotropic pixel dimension were extracted from 3D MRI data sets such that they roughly corresponded to one another. Each data set was registered with the other four data sets for each of the four algorithms producing 10 forward and reverse transformations for each algorithm. For brevity of presentation, we only present some of the results of the experiments that are representative of all of the results. A set 39 of the corresponding landmarks were manually defined in data sets *B*2 and *B*4 and a subset of the 39 landmarks were manually defined in the additional three datasets (see Fig. 6.7). Only data sets B2 and B4 had all 39 landmarks identified on them since it was not possible to locate the corresponding locations for all the landmarks on the other data sets due to missing or different shaped sulci. Only corresponding landmarks between two images were used for registration and calculating the landmark error, i.e., if one image set was missing landmark 15, then landmark 15 was not used for registration or for calculating the landmark error.

The result of transforming MRI data set *B*5 in to the shape of *B*2 using each of the four registration algorithms is shown in Fig. 6.8. These results are typical of the other pairwise registration combinations. The images are arranged left to right from the worst to the best similarity match as shown by the corresponding difference images shown below the transformed images. The UL-TPS and CL-TPS algorithms perform almost identically with respect to similarity matching. The CI-TPS and CLI-TPS intensity based registrations produce better similarity match than the two landmark only methods. In particular, the intensity based methods match the border locations and non-landmark locations better than the landmark thin-plate spline or CL-TPS algorithms. The difference between the CI-TPS and CLI-TPS methods is that the CLI-TPS method produces much

#### Christensen



Figure 6.8: Intensity matching results for registering dataset B5 to dataset B2 with the four registration algorithms. The top row shows the data set B5 transformed into the shape of B2 using each algorithm and the bottom row shows the absolute difference image between the transformed B5 image and the target B2 image. Note that the intensity difference images of the CI-TPS and CLI-TPS are very similar since both algorithms minimize the intensity differences between the deformed template and target images. However, the difference between these two results is that the CLI-TPS also produces much smaller landmark errors which cannot be seen in the intensity difference images.

smaller landmark errors than the CI-TPS method which cannot be seen in the intensity difference images.

The images in Fig. 6.9 show the Jacobian of the forward and reverse transformations between images *B*2 and *B*1 produced by the CL-TPS (left two panels) and CLI-TPS (right two panels) algorithms, respectively. The value of the Jacobian at a point is encoded such that bright pixels represent expansion, and dark pixels represent contractions. Notice that the intensity pattern of the forward and reverse Jacobian images appear nearly opposite of one another since expansion in one domain corresponds to contraction in the other domain. These images show the advantage of using both landmark and intensity information together as opposed to just using landmark information alone. Notice that the CL-TPS algorithm has very smooth Jacobian images compared to the CLI-TPS



Figure 6.9: This figure shows the Jacobians of the forward and reverse transformations for the registration of data sets *B*2 and *B*1 for the CL-TPS (left two panels) and CLI-TPS (right two panels) algorithms. The bright pixels of the Jacobian images represent regions of expansion, and dark pixels represent regions of contraction.

algorithm. This is because the CL-TPS algorithm matches the images at the corresponding landmarks and smoothly interpolates the transformation between the landmarks. Conversely, the patterning of the local distortions in the CLI-TPS registration resemble the underlying intensity patterning. This indicates that combining the intensity information with the landmark information provides additional local deformation as compared to using the landmark information alone. This improved registration between landmarks produces more distortion of the template image and therefore there is a larger range of Jacobian values for the CLI-TPS algorithm than the CL-TPS algorithm as shown by the color bar scales.

Inverse consistency error images are computed by taking the Euclidean norm of the difference between the forward and the inverse of the reverse transformations at each voxel location in the image domain. Figure 6.10 shows the inverse consistency error images for the registration of data sets *B*2 and *B*5 using the UL-TPS, CL-TPS, CI-TPS, and and CLI-TPS algorithms. Note that each image is on its own color-scale and that the UL-TPS algorithm has 10 to 200 times more maximum inverse consistency error than the consistent registration algorithms. The UL-TPS algorithm had 50 to 500 times more average inverse consistency error than the consistent registrations algorithms. This can be seen by comparing large regions of bright pixels in the UL-TPS image to the small regions of



Figure 6.10: Images that display the magnitude and location of forward transformation inverse consistency errors for matching data sets *B*2 and *B*5 with UL-TPS, CL-TPS, CI-TPS, and CLI-TPS registration algorithms.

bright pixels in the other images. This figure shows that consistent registration algorithms produced forward and reverse transformations that had sub-voxel inverse consistency errors at all voxel locations. The inverse consistent errors in the UL-TPS and CL-TPS algorithms are greatest away from the landmark driving forces because the landmark driving forces are implicitly inverse consistent. The largest inverse consistency errors in the CI-TPS and CLI-TPS algorithms occur near edges where there is a correspondence ambiguity associated with the intensity matching solution.

Table 6.2 summarizes the representative statistics collected from the experiments. Comparing the results of the UL-TPS and CL-TPS algorithms shows that the addition of inverse consistency constraint (ICC) improved the inverse consistency of the transformations with no degradation of the landmark matching. Note that for the UL-TPS algorithm, the inverse consistency error tends to be larger as one moves away from landmarks and that the inverse consistency error can be decreased by defining more corresponding landmarks.

Table 6.2 also demonstrates that the CI-TPS and CLI-TPS registrations have a smaller average intensity difference but larger landmark errors. The CLI-TPS has smaller average intensity difference and smaller landmark errors than the CI-TPS registration algorithm. The CLI-TPS algorithm produces a better similarity match because the landmark driving force pulls the intensity driving function out of local minima. It should be noted that the large number of landmarks used in the CLI-TPS registration limits the effect of the intensity driving force in neighborhoods of the landmarks. In practice, when the

Table 6.2: Experimental results produced by mapping MRI brain image 2 into images 1, 3, 4, and 5 (see Fig. 6.7). The thin-plate spline algorithms compared in this table are the unidirectional landmark (UL-TPS), consistent landmark (CL-TPS), consistent intensity (CI-TPS), and consistent landmark and intensity (CLI-TPS) algorithms. The statistics computed for these experiments were the average landmark error (ALE) in pixels, maximum landmark error (MLE), maximum inverse error (MIE) in pixels, average inverse error (AIE) in pixels, maxked average intensity difference (MAID), minimum Jacobian value (MJ), inverse of the maximum Jacobian value (LJ) and the Jacobian error (JE)

Algorithm	Exp.	ALE	MLE	AIE	MIE	MAID	MJ	IJ	JE
None	b2b1	6.9	12			0.23			
	b2b3	4.9	13			0.19			
	b2b4	8.8	21			0.22			
	b2b5	8.7	19			0.26			
UL-TPS	b2b1	0.066	0.087	0.90	2.7	0.16	0.56	0.75	0.053
	b2b3	0.073	0.098	0.78	3.1	0.18	0.50	0.57	0.092
	b2b4	0.062	0.088	0.94	3.4	0.13	0.51	0.66	0.090
	b2b5	0.030	0.061	1.2	3.8	0.16	0.56	0.67	0.050
CL-TPS	b2b1	0.000030	0.00011	0.0012	0.028	0.16	0.59	0.73	0.0011
20 iter.	b2b3	0.000034	0.00014	0.0016	0.022	0.18	0.55	0.53	0.0014
	b2b4	0.0083	0.083	0.079	0.42	0.13	0.54	0.62	0.0011
	b2b5	0.000006	0.00037	0.0024	0.015	0.16	0.56	0.62	0.00021
CI-TPS	b2b1	1.5	3.1	0.0045	0.048	0.097	0.26	0.47	0.011
1000 iter.	b2b3	1.6	2.9	0.0043	0.052	0.11	0.25	0.29	0.017
	b2b4	1.0	2.2	0.0040	0.063	0.084	0.26	0.44	0.0075
	b2b5	1.4	3.4	0.0044	0.099	0.092	0.18	0.32	0.0091
CLI-TPS	b2b1	1.1	2.0	0.020	0.40	0.091	0.19	0.37	0.036
300 iter.	b2b3	1.1	2.0	0.021	0.62	0.10	0.13	0.23	0.030
	b2b4	0.75	1.6	0.017	0.61	0.080	0.12	0.39	0.025
	b2b5	1.1	2.8	0.021	0.96	0.088	0.10	0.17	0.034

landmark points are more sparse the intensity driving force plays a more important role.

### 6.6 Acknowledgments

This work was supported in part by the NIH under grants NS35368, CA75371, and DC03590, and a grant from the Whitaker Foundation.

## Questions

- 1. What is the main benefit of using inverse consistent image registration compared uni-directional registration?
- 2. How much computation time is added by converting a uni-directional registration algorithm into a inverse consistent image registration algorithm?
- 3. How do you compute in the inverse of a transformation?

# Bibliography

- Bookstein, F. L., The Measurement of Biological Shape and Shape Change, Vol. 24, Lecture Notes in Biomathematics, Springer-Verlag: New York, 1978.
- [2] Bookstein, F. L., Morphometric Tools for Landmark Data. Cambridge University Press, New York, 1991.
- [3] Evans, A. C., Beil, C., Marret, S., Thompson, C. J., and Hakim, A., Anatomical-functional correlation using an adjustable MRI-based region of interest atlas with positron emission tomography. Journal of Cerebral Blood Flow and Metabolism, Vol. 8, pp. 513–530, 1988.
- [4] Rohr, K., Fornefett, M., and Stiehl, H. S., Approximating thin-plate splines for elastic registration: Integration of landmark errors and orientation attributes. In: Information Processing in Medical Imaging, Kuba, A. and Samal, M. eds, LCNS 1613, Springer-Verlag, Berlin, pp. 252–265, 1999.
- [5] Davatzikos, C. A., Prince, J. L., and Bryan, R. N., Image registration based on boundary mapping. IEEE Transactions on Medical Imaging, Vol. 15, No. 1, pp. 112–115, Feb. 1996.
- [6] Thirion, J. P., Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis, Vol. 2, pp. 243–260, 1998.

- [7] Subsol, G., Crest lines for curve-based warping, In Toga [42], pp. 241– 262, 1999.
- [8] Pelizzari, C. A., Chen, G. T. Y., Spelbring, D. R., Weichselbaum, R. R., and Chen, C. T., Accurate three-dimensional registration of CT, PET, and/or MR images of the brain. Journal of Computer Assisted Tomography, Vol. 13, No. 1, pp. 20–26, 1989.
- [9] Thompson, P. M. and Toga, A. W., A surface-based technique for warping three-dimensional images of the brain. IEEE Transactions on Medical Imaging, Vol. 15, No. 4, pp. 1–16, 1996.
- [10] Krahnstover, N. and Lorenz, C., Development of a point-based shape representation of arbitrary three-dimensional medical objects suitable for statistical shape modeling. In: Medical Imaging 1999: Image Processing, Hanson, K. M. ed. Proceedings of SPIE Vol. 3661, pp. 620–631, Feb. 1999.
- [11] Downs, J. H., Lancaster, J. L., and Fox, P. T., Surface-based spatial normalization using convex hulls, In Toga [42], pp. 263–282, 1999.
- [12] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching. Computer Vision, Graphics, and Image Processing, Vol. 46, pp. 1–21, 1989.
- [13] Woods, R. P., Mazziotta, J. C., and Cherry, S. R., MRI–PET registration with automated algorithm. Journal of Computer Assisted Tomography, Vol. 17, No. 4, pp. 536–546, July/August 1993.
- [14] Miller, M. I., Christensen, G. E., Amit, Y., and Grenander, U., Mathematical textbook of deformable neuroanatomies. In: Proceedings of the National Academy of Sciences, Vol. 20, No. 24, pp. 11944–48, Dec. 1993.
- [15] Christensen, G. E., Rabbitt, R. D., and Miller, M. I., Deformable templates using large deformation kinematics. IEEE Transactions on Image Processing, Vol. 5, No. 10, pp. 1435–1447, Oct. 1996.
- [16] Collins, D. L., Neelin, P., Peters, T. M., and Evans, A. C., Automatic 3D intersubject registration on MR volumetric data in standardized talairach space. Journal of Computer Assisted Tomography, Vol. 18, No. 2, pp. 192–205, March/April 1994.

- [17] Amit, Y., A non-linear variational problem for image matching. SIAM Journal on Scientific Computation, Vol. 15, No. 1, pp. 207–224, Jan. 1994.
- [18] Gee, J. C., Briquer, L., Le, Barillot, C., and Haynor, D. R., Probabilistic matching of brain images. In: Information Processing in Medical Imaging, Y. Bizais, C. Braillot, and R. Di Paola, eds, Vol. 3, pp. 113–125. Kluwer Academic Publishers, Boston, June 1995.
- [19] Woods, R. P., Grafton, S. T., Holmes, C. J., Cherry, S. R., and Mazziotta, J. C., Automated image registration: I. general methods and intrasubject, intramodality validation. Journal of Computer Assisted Tomography, Vol. 22, No. 1, pp. 139–152, 1998.
- [20] Ashburner, J., and Friston, K. J., Spatial normalization. In Toga [42], pp. 27–44, 1999.
- [21] Ashburner, J., Andersson, J. L. R., and Friston, K. J., High-dimensional image registration using symmetric priors. NeuroImage, Vol. 9, pp. 619– 628, 1999.
- [22] Christensen, G. E., Joshi, S. C., and Miller, M. I., Volumetric transformation of brain anatomy. IEEE Transactions on Med. Imaging, Vol. 16, No. 6, pp. 864–877, Dec. 1997.
- [23] Talairach, J. and Tournoux, P., Co-Planar Stereotactic Atlas of the Human Brain. Beorg Thieme Verlag, Stuttgart, 1988.
- [24] Woods, R. P., Grafton, S. T., Watson, J. D., Sicotte, N. L., and Mazziotta, J. C., Automated image registration: II. intersubject validation of linear and nonlinear models. Journal of Computer Assisted Tomography, Vol. 22, No. 1, pp. 153–165, 1998.
- [25] Segel, L. A., Mathematics Applied to Continuum Mechanics. Dover Publications, New York, 1987.
- [26] Christensen, G. E. and Johnson, H. J., Consistent image registration. IEEE Transactions on Medical Imaging, Vol. 20, No. 7, pp. 568–582, July 2001.
- [27] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximization of mutual

information. IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, April 1997.

- [28] Wells III, W. M., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R., Multi-modal volume registration by maximization of mutual information. Medical Image Analysis, Vol. 1, No. 1, pp. 35–51, 1996.
- [29] Christensen, G. E., Rabbitt, R. D., and Miller, M. I., 3D brain mapping using a deformable neuroanatomy. Physics in Medicine and Biology, Vol. 39, pp. 609–618, 1994.
- [30] Miller, M. I., Banerjee, A., Christensen, G. E., Joshi, S. C., Khaneja, N., Grenander, U., and Matejic, L., Statistical methods in computational anatomy. Statistical Methods in Medical Research, Vol. 6, pp. 267–299, 1997.
- [31] Christensen, G. E., Rabbitt, R. D., Miller, M. I., Joshi, S. C., Grenander, U., Coogan, T. A., and Essen, D. C. Van., Topological properties of smooth anatomic maps. In: Information Processing in Medical Imaging, Bizais, Y., Braillot, C., and Di Paola, R., eds, Vol. 3, pp. 101–112. Kluwer Academic Publishers, Boston, June 1995.
- [32] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O., and Hawkes, D. J., Nonrigid registration using free-form deformations: application to breast MR images. IEEE Transactions on Medical Imaging, Vol. 18, No. 8, pp. 712–721, 1999.
- [33] Christensen, G. E., Consistent linear-elastic transformations for image matching. In: Information Processing in Medical Imaging, Kuba, A. and Samal, M., eds, LCNS 1613, Springer-Verlag, Berlin, pp. 224–237, 1999.
- [34] Johnson, H. J. and Christensen, G. E., Consistent landmark and intensity-based image registration. IEEE Transactions on Medical Imaging, Vol. 21, No. 5, pp. 450–461, 2002.
- [35] Magnotta, V. A., Bockholt, H. J., Johnson, H. J., Christensen, G. E., and Andreasen, N. C., Subcortical, cerebellar and MR based consistent brain image registration. NeuroImage, Vol.19, No. 2, pp. 233–245, 2003.

- [36] Baojun, Li, Christensen, Gary E., Dill, John, Hoffman, Eric A., and Reinhardt, Joseph M., 3D inter-subject warping and registration of pulmonary CT images for a human lung model. In: Proceedings of Medical Imaging 2002: Physiology and Function from Multidimensional Images, Clough, A. V. and Chen, C. T., eds, SPIE San Diego, CA, Vol. 4683, pp. 324– 335, 2002.
- [37] Li, Baojun, Christensen, Gary E., McLennan, Geoffrey, Hoffman, Eric A., and Reinhardt, Joseph M., Establishing a normative atlas of the human lung: Inter-subject warping and registration of volumetric CT. Academic Radiology, Vol. 10, No. 3, pp. 255–265, March 2003.
- [38] Christensen, G. E., He, J., Dill, J. A., Rubinstein, J. T., Vannier, M. W., and Wang, G., Automatic measurement of the labyrinth using image registration and a deformable inner ear atlas. Academic Radiology, Vol. 10, No. 9, 2003.
- [39] Bookstein, F. L., Linear methods for nonlinear maps: Procrustes fits, thin-plate splines, and the biometric analysis of shape variability. In Toga [42], pp. 157–181, 1999.
- [40] Joshi, S. C., Miller, M. I., Christensen, G. E., Banerjee, A., Coogan, T. A., and Grenander, U., Hierarchical brain mapping via a generalized Dirichlet solution for mapping brain manifolds. In: Proceedings of SPIE Vol. 2573, Melter, R. A., Wu, A. Y., Bookstein, F. L., and Green, W. D., eds, Vision Geometry IV, pp. 278–289, 1995.
- [41] Grenander, U. and Miller, M. I., Computational anatomy: An emerging discipline. Quarterly of Applied Mathematics, Vol. 56, pp. 617–694, 1998.
- [42] Brain, Warping., A., Toga, ed. Academic Press, San Diego, 1999.

## **Chapter 7**

# A Computer-Aided Design System for Segmentation of Volumetric Images

 $Marcel Jackowski^1 and Ardeshir Goshtasby^2$ 

## 7.1 Introduction

Image segmentation is the process of partitioning an image into meaningful regions. For the regions to be meaningful, they should represent objects or their parts. Difficulties arise when properties within objects vary or boundaries of objects become blurred. The problem is worsened when sensor inaccuracies exist and noise is present in the image. These variations, which are often unpredictable, make it impossible to develop an automatic method that can segment all images correctly. Because a high accuracy is demanded in the segmentation of medical images, the user has a critical role in examining the results and correcting the possible errors.

Image segmentation is perhaps the most studied area in image analysis. A large number of papers on this topic is published annually in image analysis journals and conference proceedings. The developed methods often take into consideration various properties of images or objects, and when such properties deviate from the anticipated ones, errors occur. Even for a limited class of images, for instance MR brain images, various methods have been developed, none of which is guaranteed to work correctly on a new image. This may be because there are sensor variations; variations in the brain's shape, size, and

<sup>&</sup>lt;sup>1</sup>Department of Computer Science and Engineering, Wright State University, Dayton, OH 45435, USA

 $<sup>^2\</sup>mathrm{PET}$  Section, Kettering Medical Center, Kettering, OH45429, USA

intensity distribution; and variations in intensities of tissues surrounding the brain. Since an error-proof image segmentation method cannot be developed, user assistance is needed to correct the obtained errors. At present, the best one can hope for is to have a segmentation method that can correctly find most areas of an object of interest, and in areas where it makes a mistake, allow the user to correct them.

We have developed a computer-aided design system that allows a user to revise the result of an automatically determined segmentation. We assume the region obtained by an automatic method has a spherical topology. We also assume the region represents voxels forming the bounding surface of an object of interest in a volumetric image. The developed system fits a parametric surface to the voxels and overlays the surface with the volumetric image. By viewing both the image and the surface, the surface is edited until the desired shape is obtained. The idea behind the proposed method is depicted in Fig. 7.1.

Various user-guided and interactive segmentation methods have been developed. Barrett, Falcão, Udupa, Mortensen, and others [1, 8, 21, 22, 26] describe



Figure 7.1: The computer-aided design system used in region editing. The system starts with a region obtained from an automatic segmentation method. It then represents the region by a free-form parametric surface and overlays the surface with the volumetric image. The user then revises the surface while viewing both the volumetric image and the surface. The final result is generated parametrically or in digital form.

a method known as "live-wire" with which a user roughly draws the boundary of a region of interest. An automatic process then takes over and revises the boundary by optimizing a cost function. An alternative method is introduced that allows the user to select a number of points on the region boundary, and the program then automatically finds boundary segments between consecutive points, again by minimizing the related cost functions. These methods have been optimized for speed [9]. They have also been extended to 3D [7]. In 3D, the program receives boundary contours in a few strategically placed slices and produces contours in other slices.

Cabral *et al.* [4] describe editing tools that are associated with a regiongrowing method, enabling a user to add or remove image voxels in a region to revise the region. Hinshaw and Brinkley [14] developed a 3D shape model that uses prior knowledge of an object's structure to guide the search for the object. Object structure is interactively specified with a graphical user interface.

Höhne and Hanson [15] developed low-level segmentation functions based on morphological operators that interactively delineate regions of interest. Pizer *et al.* [24] describe a method that segments a volumetric image into regions at a hierarchy of resolutions. Then, the user, by pointing to an object in a crosssectional image at a certain resolution, selects and revises a region. Welte *et al.* [27] describe an interactive method for separating vessels from each other and from the background in MR angiographic images. To reduce the complexity of the displayed structures during the interactive segmentation, a capability to select substructures of interest is provided.

Energy-minimizing models or "snakes" are another set of tools that can be used to guide a segmentation and revise the obtained results [18, 20]. With an energy-minimizing model, a contour or a wireframe is initiated approximately where an object of interest is believed to exist. An optimization process is then activated to iteratively revise the contour or the wireframe to minimize a local cost function that defines the energy of the snake. Since some points in a snake may trap in local minima, the globally optimal solution may be missed. To avoid this, often the user is allowed to intervene and either move some of the snake's points that are thought to have converged to local minima, or guide the snake to the optimal position by interactively controlling the external forces.

An interactive segmentation method based on a genetic algorithm is described by Cagnoni *et al.* [5]. In this method, the boundary contour of a region of interest is manually drawn in one of the slices. The boundary contour is then considered an initial contour in the subsequent slice and the contour is refined by a genetic algorithm using image information. The refined boundary is then considered an initial contour in the next slice and the process is repeated until all slices in a volumetric image are segmented. Interactive segmentation methods provide varying levels of user control. The control can be as little as selecting a contour among many [19] or as much as manually drawing a complete region boundary. Methods that require a lot of user interaction are highly reliable, but they also have a high interuser variability. On the other hand, methods that require very little user interaction are not as reliable, but they have a low interuser variability. A survey of interactive segmentation methods providing different levels of user control is given by Olabarriaga and Smeulders [23].

The new idea introduced in this paper is to use the capabilities of a computeraided design system to quickly and effectively refine the result of a 3D segmentation, just like editing a 3D geometric model. By having a mental picture of an object of interest and viewing the information present in a volumetric image, the user interactively modifies the result of an automatically obtained segmentation until the desired shape is sculpted. This is achieved by representing the region by a parametric surface and overlaying the surface with the volumetric image. Then, the user views both the image and the surface together and modifies the surface until the satisfactory region is obtained.

We assume an automatic segmentation method that correctly finds most parts of a region of interest is available. The capability introduced in this paper enables the user to revise parts of the region that are believed to be inaccurate. This revision is achieved through a mechanism that sculpts a desired shape from a rough initial one. The proposed method is not the same as a dynamic snake model that creates a desired shape by interactively changing the external forces that guide the snake [20]. Rather, it is based on a parametric surface fitting and editing model.

#### 7.2 The Computer-Aided Design System

We assume a volumetric image has been segmented and a region of interest has been extracted. We also assume the given region is composed of connected voxels that represent the bounding surface of an object of interest. We will call such a region a *digital volumetric shape*, or a *digital shape*. In the following, a method that approximates a digital shape by a parametric surface is described. Since voxels belonging to a digital shape do not usually form a regular grid, we choose the rational Gaussian (RaG) formulation [11, 12], which does not require a regular grid of control points to represent a free-form shape. We will show how to parametrize voxels in a digital shape and how to determine the control points of a RaG surface that approximate the digital shape by the least-squares method. The obtained RaG surface is then overlaid with the volumetric image and the user is allowed to revise the surface by moving its control points.

#### 7.2.1 Surface Approximation

Given a set of (control) points  $\{V_i : i = 1, ..., n\}$ , the RaG surface that approximates the points is given by [11, 12]

$$\mathbf{P}(u,v) = \sum_{i=1}^{n} \mathbf{V}_{i} g_{i}(u,v), \qquad u,v \in [0,1],$$
(7.1)

where  $g_i(u, v)$  is the *i*th blending function of the surface defined by

$$g_i(u, v) = \frac{G_i(u, v)}{\sum_{j=1}^n G_j(u, v)},$$
(7.2)

and  $G_i(u, v)$  is a 2D Gaussian of height 1 centered at  $(u_i, v_i)$ :

$$G_i(u, v) = \exp\{-[(u - u_i)^2 + (v - v_i)^2]/2\sigma^2\}.$$
(7.3)

 $\{(u_i, v_i) : i = 1, ..., n\}$  are the parameter coordinates associated with the points. The parameter coordinates determine the adjacency relation between the points. In the subsequent section, we will see how to estimate them. Formulas (7.1)–(7.3) are for an open surface. If a surface is required to close from one side, like a generalized cylinder, formula (7.3) should be replaced with

$$G_i(u,v) = \sum_{k=-\infty}^{\infty} \exp\{-[(u-u_i)^2 + (v-v_i+k)^2]/2\sigma^2\}.$$
 (7.4)

If the opening at each end of a generalized cylinder converges to a point, a closed surface will be obtained. In a cylindrical surface, a 2D Gaussian wraps around the closed side of the surface infinitely. However, since a Gaussian approaches zero exponentially, its effect vanishes after a few cycles. Therefore, in practice, the  $\infty$  in formula (7.4) is replaced with a small number such as 1 or 2 [11]. An alternative method for obtaining a closed surface is to use the

formulation of a torus, which is closed along both u and v. Staib and Duncan [25] make a torus that closes at two points and separate the segment between the points by selecting proper weights in the formulation of the torus. An alternative method [13] is to transform a torus to a sphere by giving the exterior and interior circles that define the torus the same center and the same radius, allowing parametrization of an object with spherical topology using parameter coordinates of the torus.

The standard deviation of Gaussians in formulas (7.3) and (7.4) determines the smoothness of a generated surface. A surface with a smaller standard deviation represents local details better than a surface with a larger standard deviation. The larger the standard deviation, the smoother the obtained surface.

When the control points are the voxels representing a closed 3D region, the region can be represented by a parametric surface by mapping the voxels to a sphere. RaG surfaces described by Eqs. (7.1), (7.2) and (7.4) represent surfaces with a spherical topology. Assuming parameters  $\phi \in [-\pi/2, \pi/2]$  and  $\theta \in [0, 2\pi]$  represent spherical coordinates of voxels defining an object, we will need to set  $u = (\phi + \frac{\pi}{2})/\pi$  and  $v = \theta/2\pi$  in the equations of a half-closed RaG surface. In the following section, we will show how to spherically parametrize voxels in a closed digital shape, and in the subsequent section, we will show how to find the control points of a RaG surface in order to approximate a digital shape while minimizing the sum of squared errors.

#### 7.2.2 Parametrizing the Shape Voxels

Brechbühler *et al.* [2, 3] describe a method for mapping simply connected shapes to a sphere through an optimization process. Although this method can find parameters of voxels in various shapes, the process is very time consuming. We use the coarse-to-fine method described in [17] to parametrize a digital shape. In this method, first, a digital shape is approximated by an octahedron and at the same time a sphere is approximated by an octahedron. Then, correspondence is establishes between triangles in the shape approximation and triangles in the sphere approximation. By knowing parameters of octahedral vertices in the sphere approximation, parameters of octahedral vertices in the shape approximation are determined. This coarse approximation step is depicted in Fig. 7.2. The process involves placing a regular octahedron inside the shape and extending its axes until they intersect the shape and replacing the octahedral vertices



Figure 7.2: (a) Approximation of a digital shape by an octahedron. (b) Approximation of a sphere by an octahedron. Parameter coordinates of octahedral vertices in the sphere are assigned to the octahedral vertices in the shape.

with the obtained intersection points. The center of the octahedron is placed at the center of gravity of the shape and its axes are aligned with the axes of the shape [10]. If the shape is very irregular so that the center of gravity of the shape falls outside the shape, the intersection of the major axis of the shape with the shape is found and the midpoint of the longest segment of the axis falling inside the shape is taken as the center of the octahedron.

Next, the voxels associated with each triangle in the octahedral approximation are determined. This is achieved by finding the bisecting plane of each octahedral edge and determining the shape voxels that lie in that plane. In Fig. 7.3b, the bisecting planes passing through the edges of a triangle and intersecting the shape are shown. The bisecting plane passing through each octahedral edge and intersecting the shape will be an edge contour. A triangle, therefore, produces three edge contours that start and end at the vertices of the triangle and enclose the shape voxels that belong to that triangular face in the octahedral approximation. The triangles obtained in the octahedral subdivision are entered into a list. After this initial step, a triangle is removed from the list and is subdivided into smaller triangles and the triangles whose distances to the associating triangular patches are larger than a given tolerance are again entered into the list. In this manner, the triangles are removed from the list, one at a time, and processed until the list becomes empty.



Figure 7.3: (a) Octahedral approximation of a shape. (b) Edge contours delimiting a triangular patch.

Subdivision of a triangle is achieved as follows. If distances of voxels in an edge contour to the associating edge are all within the required tolerance, that edge is not subdivided. Otherwise, the farthest voxel in the contour to the edge is used to segment the contour, producing two smaller contours. The farthest contour point is then connected to the end points of the contour to produce two new edges. In this manner, a triangular face is subdivided into 2, 3, or 4 smaller triangles depending on whether 1, 2, or 3 edges of the triangle are replaced with smaller edges. This is depicted in Figs. 7.4a–7.4c. If distances of voxels in all edge contours to corresponding edges in a triangle are below the required tolerance, a test is performed to determine whether or not distances of voxels associated with the triangle are within a required tolerance to that triangle. If all distances



Figure 7.4: (a)–(c) Subdividing one, two, or three of the triangular edges, respectively. (d) When no more triangular edges can be subdivided, error between the triangular patch and the associating triangle is determined and, if that error is above the given tolerance, the farthest voxel in the patch to the triangle is determined and used to subdivide the triangle.

are below the required tolerance, the triangle is not subdivided. Otherwise, the voxel that is farthest from the triangle is connected to the three vertices of the triangle to obtain three smaller triangles. This is depicted in Fig. 7.4d. The reason for subdividing the edge contours first is to avoid very long and narrow triangles in the final subdivision.

By subdividing a triangle, finer triangles are obtained. The process is repeated until distances of all shape voxels to the associating triangles become smaller than the required tolerance. Note that this subdivision is performed in parallel in the sphere as well. Therefore, whenever a triangle in the shape approximation is subdivided, the corresponding triangle in the sphere approximation is also subdivided. For that reason, there always exists a one-to-one correspondence between triangles in the shape approximation and triangles in the sphere approximation. By knowing the parameters of mesh vertices in the sphere, we will know the parameters of corresponding mesh vertices in the shape. This process assigns spherical parameters to the mesh vertices approximating the shape. The process is graphically shown in Fig. 7.5. The only requirement of the described parametrization is for the given shape to have spherical topology.

By knowing the parameters at vertices of a triangle, parameters at points inside the triangle can be computed from barycentric coordinates [16] of parameters at the vertices. Parameters of voxels in a triangular patch are obtained by projecting the voxels to the associating triangle and assigning parameters of the triangle points to the voxels. If a triangular patch does not fold over,



Figure 7.5: (a) A sphere. (b–e) Subdivision of the sphere. (f) A digital shape. (g–j) Subdivision of the shape. The shape and the sphere are subdivided in parallel.

this mapping will be unique. If fold overs occur, the subdivision process should be continued until all fold overs disappear. When the maximum distance between a triangle and the associating patch is less than two voxels, fold overs cannot occur. More details about this parametrization algorithm and its characteristics can be found in [17].

The parameters obtained by this algorithm uniquely map shape points to triangular faces. The mapping is continuous but not smooth. To obtain a smooth parametrization, the parameters obtained here should be used as initial values to the nonlinear optimization described by Brechbühler *et al.* [3]. The surface-fitting method used in this work, however, does not require a smooth parametrization of the points. It only requires that the parameters vary continuously.

If the vertices of a triangular mesh approximating a digital shape are used as the control points of a RaG surface and the parameters at mesh vertices are used as the nodes of the surface, a smooth parametric surface can be obtained that approximates the shape. The surface obtained in this manner only approximates the mesh vertices. We can improve this shape recovery process by making the surface interpolate the mesh vertices. In the following section, a least-squares method that determines the control points of a RaG surface interpolating the mesh vertices is described.

#### 7.2.3 Least-Squares Computation of the Control Points

Suppose a digital shape is available and the shape voxels are parametrized according to the procedure outlined in the preceding section. Also, suppose the shape is composed of N voxels: { $\mathbf{P}_j : j = 1, ..., N$ } with parameter coordinates { $(u_j, v_j) : j = 1, ..., N$ }. We would like to determine a RaG surface with control points { $\mathbf{V}_i : i = 1, ..., n$ } that can approximate the shape points optimally in the least-squares sense. Let's suppose  $\mathbf{P}_j = (X_j, Y_j, Z_j)$ ,  $\mathbf{P}(u, v) = [x(u, v), y(u, v), z(u, v)]$ , and  $\mathbf{V}_i = (x_i, y_i, z_i)$ . Then the sum of squared distances between the voxels and the approximating surface can be written as

$$E^{2} = \sum_{j=1}^{N} \{ [x(u_{j}, v_{j}) - X_{j}]^{2} + [y(u_{j}, v_{j}) - Y_{j}]^{2} + [z(u_{j}, v_{j}) - Z_{j}]^{2} \}$$
(7.5)

$$= \left\{ \sum_{j=1}^{N} [x(u_j, v_j) - X_j]^2 + \sum_{j=1}^{N} [y(u_j, v_j) - Y_j]^2 + \sum_{j=1}^{N} [z(u_j, v_j) - Z_j]^2 \right\}$$
(7.6)

$$= E_x^2 + E_y^2 + E_z^2. ag{7.7}$$

Since the three components of the surface are independently defined, to minimize  $E^2$ , we minimize  $E^2_x$ ,  $E^2_y$ , and  $E^2_z$ , separately. To minimize

$$E_x^2 = \sum_{j=1}^N [x(u_j, v_j) - X_j]^2,$$
(7.8)

since

$$x(u_j, v_j) = \sum_{i=1}^n x_i g_i(u_j, v_j),$$
(7.9)

we minimize

$$E_x^2 = \sum_{j=1}^N \left[ \sum_{i=1}^n x_i g_i(u_j, v_j) - X_j \right]^2.$$
(7.10)

This involves determining the partial derivatives of  $E_x^2$  with respect to the  $x_i$ 's, setting the partial derivatives to zero and solving the obtained system of equations. This results in

$$\sum_{j=1}^{N} g_k(u_j, v_j) \sum_{i=1}^{n} [x_i g_i(u_j, v_j) - X_j] = 0; \quad k = 1, \dots, n.$$
(7.11)

This represents a system of n linear equations, which can be solved for  $\{x_i : i = 1, ..., n\}$ . Since RaG basis functions monotonically decrease from a center point, if  $\sigma$  is not very large, Eq. (7.11) will have a diagonally dominant matrix of coefficients, ensuring a solution. In the same manner,  $\{y_i : i = 1, ..., n\}$  and  $\{z_i : i = 1, ..., n\}$  can be determined by minimizing  $E_y^2$  and  $E_z^2$ , respectively. Note that the above process positions the n control points of a RaG surface so that the surface will approximate the N image voxels with the least sum of squared errors. n depends on the size and complexity of the shape being approximated. n is typically a few hundred.

Since shape voxels are mapped to a sphere, spherical parameters are obtained for them. Assuming the approximating surface is represented by  $\mathbf{P}(u, v)$ , the distance of voxel  $\mathbf{V}_i = (x_i, y_i, z_i)$  to the surface is estimated from  $E(u_i, v_i) =$  $||\mathbf{V}_i - \mathbf{P}(u_i, v_i)||$ . The adjacency information between the control points is provided in the *u* and *v* parameter coordinates. Therefore, index *i* is arbitrary and the control points with their associated nodes can be rearranged in Eq. (7.1) without having any effect in the obtained surface.

When the standard deviation in a RaG surface is very small, the surface follows individual voxels. The selected standard deviation should be large enough to smooth digital and image noise in segmentation. As the standard deviation of Gaussians is increased, a smoother surface will be obtained approximating the same set of voxels. Usually, the standard deviation of Gaussians in a RaG surface should be made proportional to the average distance between adjacent nodes. The denser the control points, the smaller the distance between their nodes, and thus, a smaller standard deviation should be used. Experimental results show that standard deviations from the average distance between adjacent nodes to five times that are appropriate for surface fitting. We will select this parameter interactively during shape editing.

Figure 7.6a shows a region representing the bounding surface of a brain tumor. Subdivision of this region to a triangular mesh with a tolerance of 3 voxels is shown in Fig. 7.6b. The tolerance shows the maximum distance between the given digital shape and the approximating triangular mesh. Approximation of the tumor with a RaG surface of standard deviation 0.002 is shown in Fig. 7.6c. Increasing the standard deviation to 0.0025, 0.003, and 0.004, we obtain the



Figure 7.6: (a) A segmented brain tumor in an MR image. (b) Approximation of the tumor by a triangular mesh with a tolerance of 3 voxels. (c)–(f) RaG surfaces approximating the tumor with standard deviations 0.002, 0.0025, 0.003, and 0.004, resulting in RMSE of 1.9725, 1.9711, 2.1439, and 2.2497 pixels, respectively.

results shown in Figs. 7.6d, 7.6e, and 7.6f, respectively. Root-mean-squared-error (RMSE) for Figs. 7.6c–7.6f are 1.9725, 1.9711, 2.1439, and 2.2497 voxels, respectively. The RMSE obtained in a RaG approximation of a digital region is usually much smaller than the tolerance used to approximate a digital shape by a triangular mesh. Figures 7.6c–7.6f appear very similar, but from the RMSE obtained, we see that their local geometries are somewhat different. The shape in Fig. 7.6f is much smoother and rounder than the shape in Fig. 7.6c.

At the standard deviation that matches the level of detail in a shape, the smallest surface-fitting error is obtained. This minimum error can be determined by a steepest-descent algorithm. However, since the given region is known to contain errors, finding the surface that is very close to the region may not be of particular interest. Currently, after the control points of an approximating surface are determined, the user interactively varies the smoothness (standard deviation) of the surface and views the obtained surface as well as the associating RMSE. In this manner, the standard deviation of Gaussians can be interactively selected to reproduce a desired level of details in a constructed shape.

#### 7.2.4 Shape Editing

Once the result of an automatic segmentation is represented by a free-form parametric surface, the surface can be revised to a desired geometry by appropriately moving its control points. In the system we have developed, an obtained surface is overlaid with the original volumetric image. Then, by going through different image slices along one of the three orthogonal directions, the user visually observes the intersection of the surface with the image slices and verifies the correctness of the segmentation. When an error is observed, one or more of the control points are appropriately moved to correct the error. As the control points are moved, the user will observe changes in the surface immediately.

An example of shape editing by the proposed method is shown in Fig. 7.7. Figure 7.7a shows the surface approximating a brain tumor within the original volumetric image. The user selects a number of control points using a small sphere that is attached to the cursor and whose center lies in the image slice being reviewed. By placing the cursor near the area where an error has occurred in one of the slices and pressing the mouse button, the sphere is activated and the control points falling in the sphere are selected. By changing the radius of sphere, the number of control points selected for movement are changed. Control points



Figure 7.7: (a) Overlaying of the approximated tumor surface and the volumetric image. The blue dots show the selected control points during surface editing. The upper-left window shows the 3D view of the image volume with all three orthogonal slices. The other three windows show the individual orthogonal views in axial, sagittal, and coronal directions. (b) Another 3D viewing mode showing the surface in wireframe form to enable viewing of image information inside and behind the surface. The red dots show the control points of the surface.

selected by the sphere are then moved with the motion of the mouse. Control points inside the sphere are not all moved by the same amount and in the same direction. A point is moved in the appropriate direction by connecting the point to the center of the sphere and by using the amount proportional to the cosine of the angle between that direction and the direction of the motion of the mouse. Only those control points falling inside the hemisphere with positive cosines are moved. This avoids motion of control points with negative cosines in the opposing direction. It also ensures that discontinuities will not occur between points that are moved and points that are not. Intermediate results in surface modification are shown in Fig. 7.7a. Surface revision can be performed gradually and repeatedly while observing the image information. The sensitivity of the surface to the motion of the mouse can be changed by increasing or decreasing the weights assigned to the control points. To better view the intersection of the surface with the image planes, the surface can be shown in wireframe form as depicted in Fig. 7.7b. The edited surface can be digitized as shown in Fig. 7.8b to create the final segmentation in digital form.



Figure 7.8: (a) The tumor after necessary modifications. This is the final result in parametric form. (b) Digitization of the tumor. This is the final result in digital form.

## 7.3 Results

A few examples of image segmentation by the proposed method are shown in Fig. 7.9. The first column shows the original images, the second column shows the initial segmentation results, and the third column shows the results after the necessary revisions. The images represent a short-axis cardiac MR image (first row), an MR brain image containing a tumor (second row), an MR image containing only the brain (third row), and a PET image of the head (fourth row). The ventricular blood pool and the brain tumor were initially obtained by a smoothing operation and an optimal intensity thresholding method. In the thresholding method, first a subvolume including the object of interest is selected. Then the intensity threshold value that produces the minimum change in the region of interest as a result of change in threshold value by 1 is determined. The optimal threshold value is considered to be the intensity where the most stable segmentation is obtained. At the optimal threshold value, a small change in threshold value will change the segmentation result minimally. This threshold value corresponds to the intensity at object boundaries where intensities change sharply. Therefore, a slight error in estimation of the threshold value will not change the segmentation result drastically.

The brain image was roughly segmented slice by slice by hand, and the PET image was segmented by our 3D implementation of the Canny edge detector



Figure 7.9: First row: A short-axis cardiac MR image and segmentation of the left ventricular cavity. Second row: An MR brain image and segmentation of the tumor. Third row: An MR brain image and segmentation of the brain. Fourth row: A PET image and extraction of the surface of the head. The first column shows the original images, the middle column shows the initial segmentation results, and the right column shows the results after the necessary modifications.



Figure 7.10: An abdominal CT image.

[6]. The Canny edge detector produces a large number of edges. First, weak edges were removed by interactively varying the gradient threshold value and observing the obtained edges. Then, an edge surface of interest was selected by pointing to the surface with the mouse and extracting it from the image. In these figures, results of the initial segmentation are shown after RaG surface fitting by the least-squares method. RaG surfaces were then interactively revised as needed while viewing the overlaid surface and volumetric image. Final segmentation results are shown in the third column of Fig. 7.9.

Another set of examples is shown in Figs. 7.10 and 7.11. Figure 7.10 is an abdominal CT image. Segmentation of different regions via intensity thresholding or edge detection, subdivision of obtained regions to triangular meshes, and fitting of RaG surfaces to the mesh vertices are shown in Fig. 7.11. Regions corresponding to the liver, one of the kidneys, and the spleen were selected one at a time and, after representing each by a RaG surface, were edited to remove inaccuracies in segmentation. Final segmentation results are shown in the column on the right in Fig. 7.11.

The time needed to obtain an initial segmentation and the time needed to modify the initial segmentation to obtain the final result vary from image to image. In the image shown in Figs. 7.9 and 7.11, approximation of the initial



Figure 7.11: First row: Polygon mesh and RaG surface approximation of the liver. Second row: Polygon mesh and RaG surface approximation of the kidney. Third row: Polygon mesh and RaG surface approximation of the spleen.

regions by triangular meshes took from 10 to 30 seconds and approximation of the regions with RaG surfaces by the least-squares method took from 40 to 60 seconds. Interactive revision of the initial surfaces to obtain the final surfaces took from 1 to 2 minutes. All these times are measured on an SGI Octane computer with R10000 processor and 128 MB RAM. Although the time to subdivide a region into a triangular mesh and the time to fit a RaG surface to a volumetric region are fixed for a given region, the time needed to revise an initial surface to a desired one depends on the speed of the user and the severity of errors in the initial segmentation.

The final result of a segmentation obtained by the proposed system is user dependent. In a typical image, the user judges what a correct segmentation is based on his/her past experiences while taking into consideration the information present in the image. Since users have different experiences in image interpretation, results obtained by different users will be different. Even the same user may segment an image differently at different times. The intrauser variability and interuser variability are not the characteristics of the proposed system, but rather those of the users. The proposed system provides tools with which a user can modify the result of a segmentation in any way desired. There are no limitations in shape, size, or complexity of a region under consideration. The only requirement is that the given region have a spherical topology.

### 7.4 Conclusions

Image segmentation is an important component of any image analysis system. In medical imaging, it is essential that an image is accurately segmented so that different measurements about the region are accurately determined. In this paper, the idea of using a computer-aided design system to effectively revise the result of an automatically determined segmentation was introduced. In the proposed system, a RaG surface is fitted to voxels representing a 3D region by the least-squares method. The surface and the original volumetric image are then overlaid and the surface is interactively revised until the desired segmentation is achieved.

The system provides the option of using the output of an automatically obtained segmentation as the input or manually creating an initial segmentation by selecting a number of 3D points in the given image volume. In the latter case, an initial surface is created from the points and overlaid with the image. The user can then observe the image data and revise the surface to a desired shape. Because a region of interest is represented by a parametric surface, the surface may be sent to a computer-aided manufacturing system for construction of an actual 3D model of the region.

## 7.5 Acknowledgements

The image used in Fig. 7.10 is from Georg Gotschuli and Erich Sorantin, University of Garz, Austria; the cardiac MR image in Fig. 7.9 is from David Turner, Presbyterian-St. Luke's Hospital, Chicago, IL; and the MR brain image containing only the brain in Fig. 7.9 is from Terry Oaks, University of Wisconsin, Milwaukee. We appreciate all these contributions.

# Bibliography

- Barrett, W. and Mortensen, E., Fast, accurate, and reproducible live-wire boundary extraction, In: Proc. Visualization in Biomedical Computing, Hamburg, 1996.
- [2] Brechbühler, Ch., Gerig, G., and Kübler, O., Surface parametrization and shape description, In: SPIE Workshop on Visualization in Biomedical Computing, Vol. 1808, pp. 80–89, 1992.
- [3] Brechbühler, Ch., Gerig, G., and Kübler, O., Parametrization of closed surfaces for 3D shape description, Computer Vision and Image Understanding, Vol. 61, No. 2, pp. 154–170, 1995.
- [4] Cabral, Jr., J. E., White, K. S., Kim, Y., and Effmann, E. L., Interactive segmentation of brain tumors in MR images using 3D region growing, SPIE Medical Imaging Conference, Vol. 1989, pp. 171–179, 1993.
- [5] Cagnoni, S., Dobrzeniecki, A. B., Ploi, R., and Yanch, J. C., Genetic algorithm-based interactive segmentation of 3D medical images, Image and Vision Computing, Vol. 17, pp. 881–895, 1999.
- [6] Canny, J., A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, pp. 679–698, 1986.
- [7] Falcão, A. X. and Udupa, J. K, A 3D generalization of user-steered live-wire segmentation, Medical Image Analysis, Vol. 4, pp. 389–402, 2000.
- [8] Falcão, A. X., Udupa, J. K., Samarasekera, S., Sharma, S., Hirsch, B. E., and Lotufo, R. A., User-steered image segmentation paradigms: live-wire and live-lane, Graphical Models and Image Processing, Vol. 60, No. 4, pp. 233–260, 1998.
- [9] Falcão, A. X., Udupa, J. K., and Miyazawa, F. K., An ultra-fast usersteered image segmentation paradigm: live-wire-on-the-fly, In: SPIE Medical Imaging Conference, San Diego, CA, Vol. 3661, pp. 184–191, 1999.

- [10] Galvez, J. M. and Canton, M., Normalization and shape recognition of three-dimensional objects by 3D moments, Pattern Recognition, Vol. 26, No. 5, pp. 667–682, 1993.
- [11] Goshtasby, A., Design and recovery of 2D and 3D shapes using rational Gaussian curves and surfaces, International Journal of Computer Vision, Vol. 10, No. 3, pp. 233–256, 1993.
- [12] Goshtasby, A., Geometric modeling using rational Gaussian curves and surfaces, Computer-Aided Design, pp. 363–375, May 1995.
- [13] Goshtasby, A., Parametric circles and spheres, Computer-Aided Design, Vol. 35, May 2003, pp. 487–494.
- [14] Hinshaw, K. P. and Brinkley, J. F., Shape-based interactive threedimensional medical image segmentation, SPIE Medical Imaging Conference, Vol. 3034, pp. 236–242, 1997.
- [15] Höhne, K. H. and Hanson, W. A., Interactive 3D segmentation of MRI and CT volumes using morphological operations, Journal of Computer Assisted Tomography, Vol. 16, No. 2, pp. 258–294, 1992.
- [16] Hoschek, J. and Lasser, D. Computer Aided Geometric Design, A. K. Peters, pp. 289–291, 1989.
- [17] Jackowski, M., Satter, M., and Goshtasby, A., Approximating digital 3D shapes by rational Gaussian surfaces, IEEE Trans. Visualization and Computer Graphics, Vol. 9, No. 1, Jan–March 2003, pp. 56–69.
- [18] Kass, A., Witkin, A., and Terzopoulos, D., Snakes: Active contour models, International Journal of Computer Vision, Vol. 1, pp. 321–331, 1987.
- [19] Krivanek, A. and Sonka, M., Ovarian ultrasound image analysis: Follicle segmentation, IEEE Trans. Medical Imaging, Vol. 17, No. 6, pp. 935–944, 1998.
- [20] McInerney, T. and Terzopoulos, D., A dynamic finite element surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis, Computerized Medical Imaging and Graphics, Vol. 19, No. 1, pp. 69–83, 1995.

- [21] Mortensen, E. N. and Barrett, W. A., Interactive segmentation with intelligent scissors, Graphical Models and Image Processing, Vol. 60, No. 5, pp. 349–384, 1998.
- [22] Mortensen, E. N., Morse, B. S., Barrett, W. A., and Udupa, J. K., Adaptive boundary detection using live-wire two-dimensional dynamic programming, In: IEEE Proc. Computer in Cardiology, pp. 635–638, 1992.
- [23] Olabarriaga, S. D. and Smeulders, A. W. M., Interaction in the segmentation of medical images: A survey, Medical Image Analyis, Vol. 5, pp. 127–142, 2001.
- [24] Pizer, S. M., Cullip, T. J., and Fredericksen, R. E., Toward interactive object definition in 3D scalar images. In: 3D Imaging in Medicine, K. H. Höhne *et al.*, eds., Springer, Berlin: pp. 83–105, 1990.
- [25] Staib, L. H. and Duncan, J. D., Model-based deformable surface finding for medical images, IEEE Trans. Medical Imaging, Vol. 15, No. 5, pp. 720–731, 1996.
- [26] Udupa, J. K., Samarasekera, S., and Berrett, W. A., Boundary detection via dynamic programming, Proc. SPIE, Vol. 1808, pp. 33–39, 1992.
- [27] Welte, D., Grunert, T., Klose, U., Petersen, D., and Becker, E., Interactive 3D segmentation and visualization of vessels, Proceedings of Computer Assisted Radiology, pp. 329–335, 1996.

## **Chapter 8**

# Inter-Subject Non-Rigid Registration: An Overview with Classification and the Romeo Algorithm

Pierre Hellier<sup>1</sup>

#### 8.1 Introduction

Registration is basically the process of estimating the spatial transformation that matches two images. Registration is especially important when analyzing motion and deformations of natural phenomena. Registration is a very active field of research. Various techniques have been proposed so far, concerning rigid and non-rigid methods.

Rigid registration is adapted in many applications except in some situations: complex deformations such as soft tissue deformation, evolution of lesions over time, matter appearance or dissipation and so on. Therefore, non-rigid registration methods (also called deformable registration) have been developed. This chapter will be mostly concerned with non-rigid registration methods and more particularly with a specific application: registration of brains of different subjects.

During the last few years, the development of electronic brain atlases has emerged by overcoming some limitations of traditional paper-based atlases [49, 62, 81, 92, 97, 139]. To do so, non-rigid intersubject registration methods have been developed in order to account for the intersubject variability [92].

<sup>&</sup>lt;sup>1</sup>IRISA-INRIA, Campus de Beaulieu, 35042 Rennes, France

Electronic atlases have two main purposes:

- Automatic segmentation of a given subject by matching the segmentation and labeling of anatomical structures of a template. Labels of the template can be deformed into another subject, under the assumption that there is a total relation between the points of the atlas (source) and the points of the studied subject (target). This objective has been pursued for a long time in medicine and was traditionally treated by paper atlases with generally rather simple transformations. The most known example is the atlas of Talairach with its famous AC-PC referential and its related proportional squaring [131].
- Understanding of brain functions. Many techniques have been developed to record brain activity (SPECT, PET, MEG/EEG, fMRI). However, the links between anatomy and functional organization are often not well known: the superimposition of multiindividual neurofunctional recordings on the same anatomy is useful to better understand the human brain functional organization. In this case, inherent anatomical variability between individuals may disturb this interpretation. Therefore, spatial normalization, which is the goal of non-rigid registration methods, makes it possible to study the functional variability. A better knowledge of this anatomy-function relationship is of great interest for the researcher in cognitive neuroscience, as well as for the surgeon and the neurologist who intend to delineate relevant functional areas before surgery.

This chapter is divided into two sections: an overview with classification of non-rigid registration techniques will be presented first. The Romeo algorithm (robust multigrid elastic registration based on optical flow) will then be described.

# 8.2 Overview of Non-Rigid Registration Methods

#### 8.2.1 Introduction

Non-rigid registration is a very active field of research and numerous methods have been proposed. This section does not intend to propose an exhaustive list of methods but to present generic and up-to-date methods. The interested reader will refer to [19, 60, 61, 87, 90, 91, 140, 143, 151] for a complete survey on this subject. This section will therefore be restricted to an overview with classification of non-rigid registration methods, more particularly applied to non-rigid registration of brains of different subjects.

Methods can generally be classified according to the following criteria:

- Features that will be matched. This includes both the dimension of the data (classically from 2D to 4D) as well as the homologous structures that are chosen for matching.
- Transformation type. This includes the transformation domain: local or global. A transformation is called "global" when the modification of one parameter affects the entire image. This also includes the transformation type (rigid, affine, projective and so on).
- The similarity measure. The similarity models the interaction between the data (features used for matching defined above) and the variables to be estimated (parameters of the transformation for instance).
- The regularization. The regularization can be implicit (regularized transformation model for instance) or explicit (first-order regularization for instance).
- The optimization method. Once the registration problem has been formalized, the optimization plays a crucial role in estimating the registration variables.

We have chosen to divide non-rigid registration methods into two classes: geometrical methods that are based on the extraction and matching of sparse features; and photometric (or intensity-based) methods that exploit luminance information directly.

#### 8.2.2 Geometric Methods

The amount of data in a 3D MR image is enormous: it contains more than 10 million voxels. The computation of a dense deformation field is a tough problem: more than  $40.10^6$  variables have to be estimated. This complexity has motivated geometric methods: sparse anatomical features reduce the dimension of the

problem. Methods that are presented here extract geometrical features from images and compute a transformation that matches these features while interpolating smoothly the deformation throughout the image.

#### 8.2.2.1 Points

Earliest methods rely on points. The most famous one, which is still a reference in the field of neuroscience, is the Talairach stereotaxic space [130]. It has then been extended to the Talairach proportional squaring system [132]. Both methods rely on the identification of the anterior comissure AC and posterior comissure PC, as well as five brain extrema which makes it possible to specify a partition of the volume into 12 subvolumes. The transformation associated with the Talairach proportional squaring system is a piecewise linear one that makes it possible to embed the brain into a "box" centered at AC and whose anatomical axes are known. This framework is known to be quite accurate in the central region but less accurate for cortical areas.

Other authors have proposed methods based on anatomical points to register brains of different subjects [16, 26, 50, 116]. However, the number of points that can be reproducibly identified among a population of subjects is limited. It has been evaluated as 36 [38] or 26 [50]. This number of points seems limited to understand the intersubject variability; in addition to this, the extraction step might be erroneous. To limit the dependency toward extraction, some authors have proposed differential geometry operators to automate the process [135, 114, 115].

#### 8.2.2.2 Curves

Guéziec [66], Subsol [126] and Declerck [41] describe methods to register two volumes thanks to curves: smoothing and curve matching in [66], application to the registration of brains in [41], building of skull atlases in [126]. Crest lines, introduced by Monga *et al.* [100], are defined as maximal curvature points and can be automatically extracted using the marching lines algorithm [137].

Gueziec *et al.* [66] approximate curves using B-splines. This enables the direct computation of features such as position, curvature and so on. Curves are then registered using an iterative approach like the Kalman filter. Subsol [126] and Declerck [41] have extended the ICP algorithm (Iterative closest points
proposed by Zhang [150] and Besl [12]). The iterative approach matches each point with the closest point on the target curve. [41] models the transformation as a B-spline while [126] estimates successively a rigid, affine, polynomial and spline transformation.

## 8.2.2.3 Surfaces

In 3D medical imaging, anatomical structures are likely to be closed surfaces than points or curves. Several authors [51, 56, 138, 128] have therefore proposed methods to register brain surfaces (like ventricles, central nuclei, brain surface). In a nutshell, deformable models have been extensively used for segmentation of medical images [8, 15, 94]. They can lead either to an explicit representation of contours (e.g., snakes) or to an implicit representation of contours (e.g., level sets [124]).

Thompson and Toga [138] have proposed an original method based on the extraction and matching of the cortical surface. Surfaces are first modeled by a superquadric [25] than refined by the "balloons" method [32]. Feldmar *et al.* [52] have extended the curve matching methods (see Section 8.2.2 to the registration of free-form surfaces. The iterative approach matches points of similar curvature.

## 8.2.2.4 From "Contour" to Volume

We have previously presented registration methods based on the extraction and matching of sparse features. The next step is to extrapolate smoothly the deformation to the entire volume. To do so, two kinds of approaches can be used: thin plate spline methods and free form deformations.

The thin plate spline (TPS) approach consists in minimizing a functional under constraints. Initiated by Duchon [47] and Meinguet [95], this method is now widespread, thanks to the work of Bookstein [16]. The problem can be formulated as the minimization of:

$$\min_{u} \int |\nabla^{m} u(x)| dx, \text{ under the constraint } \forall i \in E, u(a_{i}) = \alpha_{i},$$

where u is the deformation field, E is the set of contour points,  $a_i$  being matched with  $\alpha_i$ . Bookstein describes the TPS method that minimizes the energy of a thin plate under constraint. The solution can be expressed as local

solutions of the biharmonic equation  $\triangle^2 U = 0$  at different scales. Solutions are expressed as  $z(x, y) = -U(r) = -r^2 \log(r^2)$  in dimension 2 (with  $r = \sqrt{x^2 + y^2}$ ) and z(x, y) = |r| in dimension 3 (with  $r = \sqrt{x^2 + y^2 + z^2}$ ). This spline transformation ensures the matching of landmarks as well as a smooth interpolation of the deformation. Chui and Rangarajan have proposed the TPS-RPM algorithm [30] where they address both the correspondence and the transformation problem. They propose the softassign algorithm to solve the correspondence problem and the TPS for the transformation.

Another approach is the use of free-form deformations [123]. Initially introduced to model and deform objects [6, 148], they have also been used to model deformations [43, 64, 70, 96, 107, 117, 129, 134]. Splines models are quite powerful to extrapolate deformations indeed.

## 8.2.3 Photometric Methods

The number of features that can be extracted reproductively among a population of subjects is rather low. Therefore, photometric (also called "intensity-based" or iconic) methods have been developed to take into account the entire information of the volume. Photometric methods rely on a similarity (or dissimilarity) that measures the dependency between two volumes. We have chosen to present the registration methods according to the following classification: methods that derive from the laws of continuum mechanics; methods that use cross-correlation; the demon's method; methods based on optical flow, and finally methods that estimate jointly an intensity correction and a geometrical transformation.

## 8.2.3.1 Models Based on Continuum Mechanics

Considering two MR images of two different subjects, the estimation of a "plausible" transformation must be sought. The notion of a "plausible" transformation in this context being particularly difficult to state, some authors have proposed to comply with the laws of continuum mechanics, either elastic or fluid (Section 8.2.3).

#### 8.2.3.2 Elastic Models

Elastic models have been introduced by Broit [18] and extended by Bajcsy and Kovacic [4, 5]. These models are nowadays used by various authors [39, 38, 54,

59, 104, 105, 118, 121, 146]. The estimated deformation field should basically obey the rule of the Navier equation:

$$\mu \nabla^2 u + (\lambda + \mu) \nabla (div(u)) + F = 0,$$

where u is the deformation field to estimate,  $\lambda$  and  $\mu$  are the Lame coefficients and F is the sum of forces that are applied on the system. The problem is to specify the forces F that will lead to a correct registration. Bajcsy proposes to compute these forces so as to match the contours [5]. Davatzikos [39] and Peckar [104] do not compute any forces but segment the brain surface and the ventricles using two different methods. The matching of these surfaces provide boundary conditions that make it possible to solve the problem. These two approaches are therefore sensitive to segmentation errors.

The use of elastic methods raises the following questions:

- What should be the values of Lame coefficients? The choice of these coefficients influence the deformation. Earliest work proposed that  $\lambda = 0$  but it appears nowadays to be a limitation.
- This modeling cannot handle large deformations. As a matter of fact, the equation of Navier is only valid for small displacements. To solve this problem, two kind of approaches can be used. A rigid registration can provide a good initialization (Bajcsy [5] uses principal inertia axes and Davatzikos [38] uses the stereotaxic space). Another way [104] is to solve the problem iteratively using a multiresolution approach.
- The topology of present structures will be preserved. This may be interesting in some applications but more questionable when matching brains of different subjects. Ono [103] has shown that cortical structures are not topologically equivalent among subjects indeed.

## 8.2.3.3 Fluid Models

Following the same inspiration as elastic models, Christensen and Miller [27] propose to compute a deformation that obeys the rule of fluid mechanics (equation of Navier–Stokes). The major difference with the elastic modeling is the fact that the fluid continuously "forgets" about its initial position. Large displacements and complex motions are therefore much easier to handle. The equation

of Navier-Stokes can be written as

$$\frac{\partial \vec{u}}{\partial t} - \nu \bigtriangleup \vec{u} + (\vec{u} \cdot \vec{\nabla}) \, \vec{u} + \vec{\nabla p} = 0$$

where  $\nu$  is the fluid viscosity,  $\vec{u}$  its speed and  $\vec{p}$  its pressure. This equation is highly non-linear (cross-terms) and its resolution is complex, leading to large computation times. Christensen imposes the constraint that the Jacobian be positive [27], leading to an homeomorphic transformation.

Christensen and Johnson [28] have extended the registration approach to introduce the reversibility constraint. Given two subjects A and B, the method jointly estimates transformation from A to B and from B to A. The inverse consistency error is zero when the forward and reverse transformations are inverses of one another. Furthermore, the transformations obey the rules of continuum mechanics and are parameterized by Fourier series.

Bro-Nielsen [17] has proposed an improvement to solve the following partial differential equation:

$$\mathcal{L}v = \mu \nabla v(x) + (\lambda + \mu) \, div(v) = f(x, u(x))$$

where u is the displacement and v the instantaneous speed. For a small time change, internal forces are constant and the equation is linear. While Christensen uses a finite element scheme, Bro-Nielsen considers the impulse response associated with operator  $\mathcal{L}$ . The solution is then expressed as linear combinations of eigenvectors of operator  $\mathcal{L}$ . This significantly decreases the computation time.

Wang and Staib [146] have also proposed two methods that obey the rule of continuum mechanics. The methods respect the properties of elastic solids or viscous fluids. A statistical shape information (sparse set of forces) is mixed with a luminance information (dense set of forces within a Bayesian framework).

## 8.2.3.4 Correlation

Cross-correlation is a widespread similarity measure. It has been used by popular methods such as ANIMAL [35] and Gee *et al.* [59]. ANIMAL uses a multiresolution strategy to estimate local linear transformations that maximizes cross-correlation. At a resolution level  $\sigma$ , the regularization is based on the statement that the norm of displacement vectors should not exceed  $\sigma$ . Colllins *et al.* [36] has extended ANIMAL so that sulcal constraints can be taken into account in the registration process.

Gee, first interested in mechanical models [59], adopted a statistical Bayesian framework [58]. Let us note  $I_R$  the reference volume,  $I_T$  the target volume,  $z = \{I_R, I_T\}$  the data and u the deformation field. The problem is then to minimize the cost functional:

$$P(z|u) \propto \exp -\{\int_{x\in \Omega_T} S\left(I_T(x), I_R(x+u\left(x
ight))
ight) dx\},$$

where *S* is the similarity measure that has been chosen to be cross-correlation. The regularization follows either a membrane model  $P(u) \propto \lambda \int (u_x^2 + u_y^2) dx$  or a thin-plate model  $P(u) \propto \lambda \int (u_{xx}^2 + 2u_{xy}^2 + u_{yy}^2) dx$ . Gee also made it possible to incorporate landmark points in the registration process. If the transformation *X* matches  $p_i$  with  $p'_i$ , the associated potential is:  $P(Z = (p_i, p'_i)|\theta = X) \propto \exp{-\frac{1}{2\sigma_i^2}}||X(p_i) - p'_i||^2$ . This probabilistic approach is useful to mix mechanical regularization, photometric similarity and landmark matching. It also make it possible to experiment and compare different kinds of regularization [58].

Cachier *et al.* [21] have proposed the Pasha algorithm where the local correlation coefficient is used. This coefficient can be efficiently computed using convolutions with a Gaussian window function. The regularization is a mixture of competitive and incremental regularization using quadratic energies.

## 8.2.3.5 Demons

Thirion has proposed a method well known as the Demon's algorithm [136]. At each demon's location, force is computed so as to repulse the model toward the data. The force depends on the polarity of the point (inside or outside the model), the image difference and gradients. For small displacements, it has been shown that the demon's method and optical flow are equivalent. The method is alternated: computation of forces and regularization of the deformation field by a Gaussian smoothing. The choice of the smoothing parameter is therefore important. The Demon's algorithm has been successfully used by Dawant *et al.* [40].

Cachier and Pennec [106] have shown that the Demon's method can be viewed as a second-order gradient descent of the SSD (Sum of Square Differences). This amounts to a minmax problem: maximization of similarity and regularization of solution.

## 8.2.3.6 Displaced Frame Difference and Optical Flow

The displaced frame difference (DFD) measures the difference between voxel intensities. It can be used either directly [1, 101, 144] or linearized (known as optical flow) [46, 76, 125]. The DFD is known to be highly non-linear whereas optical flow is linear. However, optical flow is only valid for small displacements and can estimate motion only in the direction of the image gradient (aperture problem). In both cases, this similarity will not be valid if luminance is not conserved (this may happen because of image acquisition, acquisition systems or parameters, MR inhomogeneities and so on).

Close to mechanical approaches, Song and Leahy [125] and Devlaminck [46] have proposed to estimate the optical flow with a mechanical regularization. More specifically, when images are density images (the luminance is directly related to a physical quantity), the mass conservation hypothesis may be introduced to constraint the estimation in a plausible way [37, 125].

In the field of cardiac imaging, Reissmann *et al.* [112] have proposed to use the neuractive pyramid to register images using the optical flow. The elastic grid that is the kernel of the deformation deforms so as to reject the discontinuities at boundaries of the grid. The minimization is therefore alternated between the deformation and the optimal shape of the grid.

The SPM spatial normalization approach [2] estimates warps by matching each skull-stripped image to the skull-stripped reference. Registration involves minimizing the mean squared difference between the images, which had been previously smoothed by convolving with an isotropic 8 mm FWHM Gaussian kernel. The non-rigid deformation is modeled by a linear combination of lowfrequency cosine transform basis functions [2]. Displacements in each direction are parameterized by 392 basis function coefficients, making a total of 1176 parameters in total. Regularization is obtained by minimizing the membrane energy of the warps.

Vemuri [144] also uses the optical flow but models the deformation as a combination of splines similarly to [127]. Finally, Musse *et al.* [101] describe a hierarchical method to estimate the deformation using the SSD criterion. The solution is sought as a combination of the spine's functions that ensure the regularity of the solution.

# 8.2.3.7 Joint estimation of Intensity and Geometric Transformations

Many artifacts can modify the luminance of an MR image. One of them is the inhomogeneity of the magnetic field for instance [80]. As a consequence, the hypothesis of luminance conservation might not be valid anywhere. One solution consists in using robust estimators to get rid of inconsistent data. Another solution consists in estimating jointly an intensity correction and a spatial transformation [53, 55, 65].

Gupta and Prince [65] propose an affine correction model for tagged MR:  $f(\mathbf{r} + \mathbf{dr}, t + dt) = m(\mathbf{r}, \mathbf{dr}, t, dt) f(\mathbf{r}, t) + c(\mathbf{r}, \mathbf{dr}, t, dt)$ . The optical flow equation then becomes:

$$f(\mathbf{r},t) + \nabla f(\mathbf{r},t) \cdot U(\mathbf{r},t) - f(\mathbf{r},t) \frac{\partial m(\mathbf{r},t)}{\partial t} - \frac{\partial c(\mathbf{r},t)}{\partial t} = 0$$

The equation is solved in a variational framework using a first-order regularization.

Friston [55] and Feldmar [53] propose to embed the intensity correction and the spatial transformation in the same cost functional:

$$\mathcal{C}(\mathbf{f},g) = \sum_{M_i \in i_1} (I_2(f(M_i)) - g(I_1(M_i), M_i))^2,$$

where f is the 3D transformation and g is the intensity correction. Feldmar generalizes this approach and considers 3D images as 4D surfaces. The criterion becomes:

$$\mathcal{C}(\mathbf{f},g) = \sum_{(x_j,i_j)} d \left( (\mathbf{f}(x_j), g(x_j, i_j)), CP_{4D}(\mathbf{f}(x_j), g(x_j, i_j)) \right)^2,$$

where  $x_j$  is the point of intensity  $i_j$  and  $CP_{4D}$  is the function that renders the closest point. In this sense, this method is a generalization of the ICP (iterative closest point) algorithm. Functions f and g can be modeled according to the application. For instance, for a intra-subject monomodal registration, f is rigid and g is the identity. For inter-subject registration, f can be a combination of radial basis functions and f should correct acquisition artifacts.

## 8.2.3.8 Non-Rigid Multimodal Registration

Although many efforts have been made to perform rigid multimodal registration, as far as we know, there has been few research concerning non-rigid multimodal registration. As a matter of fact, this is quite a challenging problem, since the number of variables to be estimated can be very large (intensity mapping, and geometrical transformation, the two being dependent). Two different approaches have been developed.

One option is to estimate the geometrical transformation with the original intensities of the two images to be registered. In this category, Maintz *et al.* [89] and Gaens *et al.* [57] proposed an algorithm that seek a non-rigid transformation by maximization of mutual information. They use a "block-matching" minimization scheme with a Gaussian filtering of the estimated deformation field to avoid blocky effects. On local windows, the estimation does not take into account the spatial context of the deformation field and only a translation is estimated. Furthermore, these methods are only performed in 2D. Rueckert *et al.* [117] and Kybic *et al.* [84] proposed an approach based on cubic B-splines and mutual information. The spline deformation model intrinsically contains the regularization and provides a smooth interpolation of the field. Displacement of the nodes are computed such as to maximize the similarity measure (mutual information, or normalized mutual information).

Another appealing option has been proposed by Guimond *et al.* [63]. This method considers the multimodal registration problem as a monomodal registration problem, and therefore estimates alternatively an intensity correction and a monomodal registration. The originality of the method resides in the decomposition of the problem into two "easier" ones: a polynomial intensity mapping and a monomodal registration problem based on the demon's algorithm [136].

## 8.2.4 Discussion

This section has presented a brief overview of non-rigid registration techniques. Methods have been arbitrarily classified into two groups: geometric methods that rely on the extraction and matching of geometrical features; and photometric methods (or intensity-based) that rely on the luminance information directly.

Geometric methods are attractive because they rely on anatomical features. The deformation is expected to be consistent in the vicinity of features that are used. In addition, the complexity is significantly reduced compared to the method that uses the entire data. Despite these advantages, these methods appear limited in the context of inter-subject registration. As a matter of fact, the number of features that can be reproductively identified among a population of subjects is limited compared to the inter-subject variability. Furthermore, a lot of information present in the data are not used by geometric methods while photometric take advantage of all information available. This rapid comparison may explain the popularity of photometric methods, which has been proved in the particular context of rigid multimodal fusion [147].

Photometric methods differ by numerous aspects. Among them, two important ones are the similarity measure and the regularization.

The choice of the similarity is crucial since this models the interaction between the data and the estimated variables. Roche *et al.* have shown [113] that the choice of a similarity can be guided by the a priori knowledge that we have about the data. Regularization is also crucial since it expresses the a priori knowledge that we have about the deformation. The choice of a correct regularization in the context of inter-subject normalization is difficult and still debated since we do not know what should be the "ideal" deformation field between two brains of two different subjects. Regularization often conserves the topology of brain structures. While valid for internal structures such as ventricles, the conservation of topology is a strong hypothesis when dealing with cortical structures. Anatomists have indeed shown that cortical sulci have different shapes and topology among individuals [103].

Recently there has been an increasing number of promising methods [22, 29, 36, 68, 73, 79, 141] that combine the benefits of photometric and geometric approaches to register brains of different subjects. In these methods, landmarks are used to drive the registration process so that the deformation field is consistent with the matching of sparse anatomical structures.

# 8.3 Romeo: Robust Multigrid Elastic Registration Based on Optical Flow

## 8.3.1 Introduction

We consider the registration problem as a motion estimation problem, which has been studied by different authors [7, 9, 10, 11, 13, 31, 76, 82, 102, 120]. Our

3D method performs a non-linear multimodality registration of MRI acquisition of different subjects. The similarity measure that we use incorporates robust estimators whose utility is twofold: on the one hand we want to limit the influence of the acquisition noise, on the other hand, we want to cope with possible modifications of structures' topology [75].

Since the luminance of MR images might not be directly comparable, we propose an intensity correction scheme that is anatomically consistent [71]. This correction method will be described in Section 8.3.2. Then volumes to be registered are rigidly aligned by maximizing mutual information (described in Section 8.3.2).

Many tasks in computer vision may be expressed as the minimization of a cost function. The optimization is often difficult to achieve, because the cost function is non-convex and because the optimization involves a very large number of variables. Therefore efficient iterative multigrid (or multilevel) approaches have been developed [67, 93] and applied in motion estimation [48] and in early vision [133].

To take into account large deformations, we use a multiresolution optimization scheme. Besides, at each resolution level, we use a multigrid minimization to accelerate the algorithm and improve the quality of the estimation. Within this hierarchical approach, we designed an adaptive partition of the volume to refine the estimation on the regions of interest and avoid useless efforts elsewhere. An anatomical segmentation of the cortex is introduced and used in two ways: at each resolution level, we initialize the partition as an octree subdivision based on the segmentation, and the segmentation mask is used in the subdivision criterion which controls the refinement of the estimation.

The method will first be extensively presented in Section 8.3.2. We will also present an extension of this method to multimodal data [72] in Section 8.3.2. Results on synthetic and real data will then be presented in Section 8.3.3.

## 8.3.2 Method

## 8.3.2.1 General Formulation

The optical flow hypothesis, or brightness constancy constraint, introduced by Horn and Schunck [76], assumes that the luminance of a physical point does not vary much between the two volumes to register. It amounts to zeroing the so-called DFD (displaced frame difference):

$$f(s + w_s, t_1) - f(s, t_2) = 0,$$

where *s* is a voxel of the volume,  $t_1$  and  $t_2$  are the indexes of the volumes (temporal indexes for a dynamic acquisition, indexes in a database for multisubject registration), *f* is the luminance function and *w* the expected 3D displacement field. The DFD may not be valid everywhere, because of noise and intensity inhomogeneities of MR acquisition. The robustness of the registration process with respect to acquisition artifacts will be discussed later on, the sections 8.3.2 and 8.3.3.

Generally, a linear expansion of this equation is preferred :  $\nabla f(s, t) \cdot \boldsymbol{w}_s + f_t(s, t) = 0$  where  $\nabla f(s, t)$  stands for the spatial gradient of luminance and  $f_t(s, t)$  is the voxelwise difference between the two volumes. The resulting set of undetermined equations has to be complemented with some prior on the deformation field. Using an energy-based framework (which can be viewed either from the Bayesian point of view, or from the one of the regularization theory), the registration problem may be formulated as the minimization of the following cost function:

$$U(\boldsymbol{w};f) = \sum_{s \in S} [\boldsymbol{\nabla} f(s,t) \cdot \boldsymbol{w}_s + f_t(s,t)]^2 + \alpha \sum_{\langle s,r \rangle \in \mathcal{C}} ||\boldsymbol{w}_s - \boldsymbol{w}_r||^2, \quad (8.1)$$

where *S* is the voxel lattice, *C* is the set of neighboring pairs w.r.t. a given neighborhood system  $\mathcal{V}$  on *S* ( $\langle s,r \rangle \in \mathcal{C} \Leftrightarrow s \in \mathcal{V}(r)$ ), and  $\alpha$  controls the balance between the two energy terms. The first term captures the brightness constancy constraint, thus modeling the interaction between the field (unknown variables) and the data (given variables), whereas the second term captures a simple smoothness prior. The weaknesses of this formulation are known:

- (a) Due to the linearization, the optical flow constraint (OFC) is not valid in case of large displacements.
- (b) The OFC might not be valid in all the regions of the volume, because of the acquisition noise, intensity non-uniformity in MRI data, and occlusions.
- (c) The "real" field is not globally smooth and it probably contains discontinuities that might not be preserved because of the quadratic smoothing.

To cope with the (b) and (c) limitations, we replace the quadratic cost by robust functions. To face the problem (a), we use a multiresolution plan and a multigrid strategy to improve the minimization at each resolution level.

We have here introduced a simple regularization term that makes almost no assumption on the estimated deformation field. One could imagine choosing different regularizations for the different brain tissues, but that involves specific assumptions on the "real" deformation that we do not address in that paper. However, the introduction of a robust estimator on the regularization term makes it possible to take into account possible discontinuities on the border of structures having different physical properties.

#### 8.3.2.2 Rigid Registration Step

Given two images with potentially large displacement, it first seems reasonable to estimate a rigid transformation. This step is performed by estimating a rigid transformation that maximizes mutual information [33, 145]. Given two images A and B, considered as discrete random variables, let us note  $p_A(a)$  and  $p_B(b)$  their respective marginal probability distribution, and  $p_{A,B}(a, b)$  the joint distribution. Mutual information I(A, B) is then defined as [33, 145]:

$$I(A, B) = \sum_{a,b} p_{A,B}(a, b) \log_2 \frac{p_{A,B}(a, b)}{p_A(a)p_B(b)} = H(A) + H(B) - H(A, B),$$

with

$$H(A) = -\sum_{a} p_A(a) \log_2(p_A(a))$$
 and  
 $H(A, B) = -\sum_{a,b} p_{A,B}(a, b) \log_2(p_{A,B}(a, b))$ 

In some particular cases, such as brain images for instance, it is possible to define a reference coordinate system that takes into account some information about the scene (such as resolution of pixels/voxels, orientation of axes, etc.). In such cases, the two volumes to be registered are mapped in this reference coordinate system and the rigid transformation is expressed in this coordinate system. If this a priori information is not available, the rigid transformation is estimated in the coordinate system attached to the data.

The registration is performed through a multiresolution optimization scheme (construction of a pyramid of volumes by successive isotropic Gaussian filtering and subsampling in each direction) [69, 109]. At each resolution level, the similarity I(A, T(B)) is maximized w.r.t. the parameters of the transformation using a Powell's algorithm [110]. We calculate the joint histogram on the overlapping part of A with T(B) by partial volume interpolation, the latter being known to provide a smoother cost function.

#### 8.3.2.3 Intensity Correction

The hypothesis of luminance conservation is strong and cannot stand when considering a large database. Actually, studies nowadays involve distributed databases. Since the MR acquisition can come from different systems, the intensity difference of MR images of different subjects needs to be corrected prior to registration. Let us formulate the problem as:

Given two 3D images  $I_1$  and  $I_2$ , and their histograms  $h_1$  and  $h_2$ , the problem is to estimate a correction function g such that corresponding anatomical tissues of  $g(I_1)$  and  $I_2$  have the same intensity, without registering volumes  $I_1$  and  $I_2$ .

**Estimation of Mixture Model.** The intensity correction f should be anatomically consistent, i.e., the intensity of gray matter (resp. white matter) of  $g(I_1)$  should match the intensity of gray matter (resp. white matter) of  $I_2$ . To ensure this coherence, we estimate a mixture of n Gaussian distributions [3, 83, 86, 122, 149] that models the two histograms  $h_1$  and  $h_2$  using the expectation-maximization (EM) algorithm [44] or a stochastic version, the stochastic expectation maximization (SEM) algorithm [23].

Basically, the EM algorithm consists of two steps: Step E where conditional probabilities are computed, and step M where mixtures parameters are estimated so as to maximize the likelihood. Contrary to the EM algorithm, the SEM algorithm consists in adding a stochastic "perturbation" between the E and M step. The labels are then randomly chosen from their current conditional distribution. The SEM algorithm is supposed to be less sensitive to initialization but also to converge more slowly than the EM algorithm.

It is well known that the MR histogram can be roughly modeled as the mixture of five Gaussian laws modeling the main tissues: background, cerebrospinal fluid (CSF), gray matter (GM), white matter (WM) and a mixture of fat and muscle. The Gaussian mixture has proved to be relevant for fitting MR-T1 histograms [83]. It has also been shown that mixture tissues (interface gray-CSF and Gray-White) can also be modeled by additional Gaussian laws to model partialvolume effects. To do so, a mixture of seven models can be used instead.

In every case (EM or SEM algorithm, five or seven Gaussian models), we model each class  $\kappa$  by a Gaussian distribution of mean  $\mu_{\kappa}$  (respectively,  $\nu_{\kappa}$ ) for image  $I_1$  (respectively, image  $I_2$ ).

**Parametric Correction.** To align the intensities of the anatomical tissues and to interpolate smoothly the correction, we choose a polynomial correction function of order p (see [63] for a similar modeling of intensity correction) such that  $g^{p}(x) = \sum_{i=0}^{i=p} \theta^{i} x^{i}$ . The coefficients  $\theta^{i}$  are estimated such as to minimize the following cost:

$$\sum_{l=1}^{l=n} \left( g^p(\mu_j) - \nu_j \right)^2.$$

The intensity correction aims at aligning the mean values of each classes while interpolating smoothly between the samples. This least-square problem amounts to inverting a linear system of order p. The resulting correction can then be applied to the voxel intensities of volume  $I_1$ .

#### 8.3.2.4 Robust Estimators

Cost function Eq. (8.1) does not make any difference between relevant data and inconsistent data, nor between neighboring pairs where the field is smooth and neighboring pairs where the field is discontinuous. Therefore, we introduce robust functions [77] and more precisely two robust M-estimators [14], the first one on the data term and the second one on the regularization term. We do not describe in details the properties of robust M-estimators, referring the reader to [14, 98] for further explanations. The cost function (8.1) can then be modified as:

$$U(\boldsymbol{w};f) = \sum_{s \in S} \rho_1 \left( \nabla f(s,t) \cdot \boldsymbol{w}_s + f_t(s,t) \right) + \alpha \sum_{\langle s,r \rangle \in \mathcal{C}} \rho_2 \left( ||\boldsymbol{w}_s - \boldsymbol{w}_r|| \right). \quad (8.2)$$

According to some properties of robust M-estimators [14, 24], it can be shown that the minimization of U (Eq. 8.1) is equivalent to the minimization of an augmented function, noted U:

$$\overset{*}{U}(\boldsymbol{w},\delta,\beta;f) = \sum_{s\in S} \delta_s \left(\boldsymbol{\nabla} f(s,t) \cdot \boldsymbol{w}_s + f_t(s,t)\right)^2 + \psi_1(\delta_s) + \alpha$$
$$\times \sum_{\langle s,r \rangle \in \mathcal{C}} \beta_{sr} ||\boldsymbol{w}_s - \boldsymbol{w}_r||^2 + \psi_2(\beta_{sr}), \tag{8.3}$$

where  $\delta_s$  and  $\beta_{sr}$  are auxiliary variables (acting as "weights") to be estimated. This cost function has the advantage to be quadratic with respect to  $\boldsymbol{w}$ . It also shows clearly that, when a discontinuity gets larger, the contribution of the pair of neighbors is limited by the reduction of the associated weight  $\beta_{sr}$ . The minimizers of  $\overset{*}{U}$  with respect to the auxiliary variables are obtained in closed form [14, 24]. The overall minimization of such function consists in an alternated weights computation and quadratic minimizations (with respect to  $\boldsymbol{w}$ ).

# 8.3.2.5 Multiresolution Incremental Computation of the Optical Flow

In cases of large displacements, we use a classical incremental multiresolution procedure [11, 48] (see Fig. 8.1). We construct a pyramid of volumes  $\{f^k\}$  with successive Gaussian smoothing and subsampling in each direction [20]. For each direction  $i = x, y, z, d_i$  is the spatial resolution of a voxel (the spatial resolution of MR acquisition is around 1 mm, depending on the system). We perform a Gaussian filtering using the recursive implementation proposed in [45] with a standard deviation of  $2d_i$  in direction *i*, in order to satisfy Nyquist's criterion. This implementation allows to perform infinite impulse response filtering at a constant computation cost.

At the coarsest level, displacements are reduced, and cost function (8.3) can be used because the linearization hypothesis becomes valid. For the next resolution levels, only an increment  $dw^k$  is estimated to refine the estimate  $\hat{w}^k$  obtained



Figure 8.1: Incremental estimation of the optical flow.

from the previous level. We perform the registration from resolution  $k_c$  until resolution  $k_f$  (in general  $k_f = 0$ ). This is done using cost function (8.2) but with  $\nabla \tilde{f}^k(s,t) \triangleq \nabla f^k(s+\hat{w}_s^k,t_2)$  and  $\tilde{f}^k_t(s,t) \triangleq f^k(s+\hat{w}_s^k,t_2) - f^k(s,t_1)$  instead of  $\nabla f^k(s,t)$  and  $f^k_t(s,t)$ .

To compute the spatial and temporal gradients, we construct the warped volume  $f^k(s + \hat{w}_s^k, t_2)$  from volume  $f^k(s, t_2)$  and the deformation field  $\hat{w}_s^k$ , using trilinear interpolation. The spatial gradient is hence calculated using the recursive implementation of the derivatives of the Gaussian [45]. At each voxel, we calculate the difference between the source volume and the reconstructed volume, and the result is filtered with a Gaussian to construct the temporal gradient. As previously, these quantities come from the linearization of the constancy assumption expressed for the whole displacement  $\hat{w}_s^k + dw_s^k$ . The regularization term becomes  $\sum_{< s, r > \in \mathcal{C}} \rho_2(||\hat{w}_s^k + dw_s^k - \hat{w}_r^k - dw_r^k||)$ .

#### 8.3.2.6 Multigrid Minimization Scheme

**Motivations.** The direct minimization of Eq. (8.3) is intractable. Some iterative procedure has to be designed. Unfortunately, the propagation of information through local interaction is often very slow, leading to an extremely time-consuming algorithm. To overcome this difficulty (which is classical in computer vision when minimizing a cost function involving a large number of variables), multigrid approaches have been designed and used in the field of computer vision [48, 98, 133]. Multigrid minimization consists in performing the estimation through a set of nested subspaces. As the algorithm goes further, the dimension of these subspaces increases, thus leading to a more accurate estimation. In practice, the multigrid minimization usually consists in choosing a set of basis functions and estimating the projection of the "real" solution on the space spanned by these basis functions.

**Description.** At each level of resolution, we use a multigrid minimization (see Fig. 8.2) based on successive partitions of the initial volume [98]. At each resolution level k, and at each grid level  $\ell$ , corresponding to a partition of cubes, we estimate an incremental deformation field  $dw^{k,\ell}$  that refines the estimate  $\hat{w}^k$ , obtained from the previous resolution levels. This minimization strategy, where the starting point is provided by the previous result—which we hope to be a rough estimate of the desired solution—improves the quality and the



Figure 8.2: Example of multiresolution/multigrid minimization. For each resolution level (on the left), a multigrid strategy (on the right) is performed. For legibility reasons, the figure is a 2D illustration of a 3D algorithm with volumetric data.

convergence rate as compared to the standard iterative solvers (such as Gauss– Seidel).

At grid level  $\ell$ ,  $\Xi_{\ell} = \{\Xi_n, n = 1, ..., N_{\ell}\}$  is the partition of the volume *B* into  $N_{\ell}$  cubes  $\Xi_n$ . At each grid level  $\ell$  corresponds a deformation increment  $T_{k,\ell}$  that is defined as follows: A 12-dimensional parametric increment deformation field is estimated on each cube  $\Xi_n$ , hence the total increment deformation field  $d\boldsymbol{w}^{k,\ell}$  is piecewise affine. At the beginning of each grid level, we construct a reconstructed volume with the target volume  $f^k(s, t_2)$  and the field estimated previously (see section 8.3.2). We compute the spatial and temporal gradients at the beginning of each grid level and the increment deformation field  $d\boldsymbol{w}^{k,\ell}$  is initialized to zero. The final deformation field is hence the sum of all the increments estimated at each grid level, thus expressing the hierarchical decomposition of the field.

Contrary to block-matching algorithms, we model the interaction between the cubes (see Section 8.3.2) of the partition, so that there is no "block-effects" in the estimation. At each resolution level k, we perform the registration from grid level  $\ell_c$  until grid level  $\ell_f$ . Depending on the application, it may be useless to compute the estimation until the finest grid level, i.e.,  $\ell_f = 0$ . We will evaluate this fact later on (see section 8.3.3).

Adaptive Partition. To initialize the partition at the coarsest grid level  $\ell_c$ , we consider a segmentation of the brain obtained by morphological operators. After a threshold and an erosion of the initial volume, a region growing process is performed from a starting point that is manually chosen. A dilatation

operation allows us to end up with a binary segmentation. At grid level  $\ell_c$ , the partition is initialized by a single cube of the volume size. We iteratively divide each cube as long as it intersects the segmentation mask and as long as its size is superior to  $2^{3\ell_c}$ . We finally get an octree partition which is anatomically relevant.

When we change from grid level, each cube is adaptively divided. The subdivision criterion depends first on the segmentation mask (we want a maximum precision on the cortex), but it also depends on the local distribution of the variables  $\delta_s$  (see Eq. (8.3)). More precisely, a cube is divided if it intersects the segmentation mask or if the mean of  $\delta_s$  on the cube is below a given threshold. As a matter of fact,  $\delta_s$  indicates the adequation between the data and the estimated deformation field at voxel *s*. Therefore, this criterion mixes an indicator of the confidence about the estimation with a relevant anatomical information.

#### 8.3.2.7 Parametric Model

We now introduce the deformation model that is used. We chose to consider an affine 12-parameter model on each cube of the partition. That kind of model is quite usual in the field of computer vision but rarely used in medical imaging. If a cube contains less than 12 voxels, we only estimate a rigid 6-parameter model, and for cubes that contain less than 6 voxels, we estimate a translational displacement field. As we have an adaptive partition, all the cubes of a given grid level might not have the same size. Therefore, we may have different parametric models, adapted to the partition.

At a given resolution level k and grid level  $\ell$ ,  $\Xi_{k,\ell} = \{\Xi_n, n = 1 \cdots N_{k,\ell}\}$  is the partition of the volume into  $N_{k,\ell}$  cubes  $\Xi_n$ . On each cube  $\Xi_n$ , we estimate an affine displacement defined by the parametric vector  $\Theta_n^{k,\ell}$ :  $\forall s = (x, y, z) \in$  $\Xi_n$ ,  $\mathrm{d} \boldsymbol{w}_s = P_s \Theta_n^{k,\ell}$ , with

A neighborhood system  $V^{k,\ell}$  on the partition  $\Xi_{k,\ell}$  derives naturally from  $\mathcal{V}$  (see section 8.3.2):

 $\forall n, m \in \{1 \cdots N_{k,\ell}\}, m \in V^{k,\ell}(n) \Leftrightarrow \exists s \in \Xi_n, \exists r \in \Xi_m \setminus r \in \mathcal{V}(s). \mathcal{C}$  being the set of neighboring pairs on  $S^k$ , we must now distinguish between two types of such pairs: the pairs inside one cube and the pairs between two cubes:

$$\forall n \in \{1 \dots N_{k,\ell}\}, < s, r > \in \mathcal{C}_n^{\ell} \Leftrightarrow s \in \Xi_n, r \in \Xi_n \text{ and } r \in \mathcal{V}(s).$$
  
$$\forall n \in \{1 \dots N_{k,\ell}\}, \forall m \in V^{\ell}(n), < s, r > \in \mathcal{C}_{nm}^{\ell} \Leftrightarrow m \in V^{l}(n), s \in \Xi_n, r \in \Xi_m$$
  
and  $r \in \mathcal{V}(s).$ 

For the sake of concision, we will now drop the resolution index k. With these notations, the cost function (8.3) becomes

$$\overset{*^{\ell}}{U} (\Theta^{\ell}, \delta^{\ell}, \beta^{\ell}; \boldsymbol{w}, f^{\ell}) = \sum_{n=1}^{N_{\ell}} \sum_{s \in \Xi_{n}} \delta^{\ell}_{s} [\boldsymbol{\nabla} \tilde{f}^{T}_{s} P_{s} \Theta^{\ell}_{n} + \tilde{f}_{t}(s, t)]^{2} + \psi_{1}(\delta^{\ell}_{s}) 
+ \alpha \sum_{n=1}^{N_{\ell}} \left[ \sum_{m \in V^{\ell}(n)} \sum_{\langle s, r \rangle \in \mathcal{C}^{\ell}_{nm}} \beta^{\ell}_{sr} || (\boldsymbol{w}_{s} + P_{s} \Theta^{\ell}_{n}) - (\boldsymbol{w}_{r} + P_{r} \Theta^{\ell}_{m}) ||^{2} + \psi_{2}(\beta^{\ell}_{sr}) \right] 
+ \alpha \sum_{n=1}^{N_{\ell}} \left[ \sum_{\langle s, r \rangle \in \mathcal{C}^{\ell}_{n}} \beta^{\ell}_{sr} || (\boldsymbol{w}_{s} + P_{s} \Theta^{\ell}_{n}) - (\boldsymbol{w}_{r} + P_{r} \Theta^{\ell}_{n}) ||^{2} + \psi_{2}(\beta^{\ell}_{sr}) \right]. \quad (8.4)$$

Considering the auxiliary variables of the robust estimators as fixed, one can easily differentiate the cost function (8.4) with respect to any  $\Theta_n^{\ell}$  and get a linear system to be solved. We use a Gauss-Seidel method to solve it for its implementation simplicity. However, any iterative solver could be used (solvers such as conjugate gradient with an adapted preconditioning would also be efficient). In turn, when the deformation field is "frozen", the weights are obtained in a closed form [14, 24]. The minimization may therefore be naturally handled as an alternated minimization (estimation of  $\Theta_n^{\ell}$  and computation of the auxiliary variables). Contrary to other methods (minmax problem like the demon's algorithm for instance), that kind of minimization strategy is guaranteed to converge [24, 42, 99] (i.e., to converge toward a local minimum from any initialization).

Moreover, the multigrid minimization makes the method invariant to intensity inhomogeneities that are piecewise constant. As a matter of fact, if the intensity inhomogeneity is constant on a cube, the restriction of the DFD on that cube is modified by adding a constant. As a consequence, minimizing the cost function 8.3.2 gives the same estimate, whenever the cost at the optimum is zero or a constant (see section 8.3.3 for an illustration on that issue).

#### 8.3.2.8 Multimodal Non-Rigid Registration

We have proposed a multimodal version of Romeo [72] where the optical flow is replaced by a more adapted and general similarity measure: mutual information. Mutual information has been presented in the section dedicated to rigid registration (8.3.2).

Let us note  $T_{\boldsymbol{w}}$  as the transformation associated with the deformation field  $\boldsymbol{w}$ . The total transformation  $T_{\boldsymbol{w}} \circ T_0$  maps the floating volume B onto the reference volume A. The field  $\boldsymbol{w}$  is defined on  $S_B$ , where  $S_B$  denotes the lattice of volume B(pixel lattice or voxel lattice). The cost function to be minimized then becomes:

$$U(\boldsymbol{w}; A, B, \mathsf{T}_0) = -I(A, (\mathsf{T}_{\boldsymbol{w}} \circ \mathsf{T}_0)(B)) + \alpha \sum_{\langle s, r \rangle \in \mathcal{C}_B} ||\boldsymbol{w}_s - \boldsymbol{w}_r||^2,$$

where  $C_B$  is the set of neighboring pairs of volume B (if we note  $\mathcal{V}$  a neighborhood system on  $S_B$ , we have:  $\langle s, r \rangle \in C_B \Leftrightarrow s \in \mathcal{V}(r)$ ).

A multiresolution and multigrid minimization are also used in this context. At grid level  $\ell$  and on each cube  $\Xi_n$ , we estimate an affine displacement increment defined by the parametric vector  $\Theta_n^{\ell}$ :  $\forall s \in \Xi_n$ ,  $d\boldsymbol{w}_s = P_s \Theta_n^{\ell}$ , with  $P_s = \mathbb{I}_2 \otimes [1x_s y_s]$  for 2D images, and  $P_s = \mathbb{I}_3 \otimes [1x_s y_s z_s]$  for 3D images (operator  $\otimes$  denotes the Kronecker product).

To be more explicit, in 3D we have:

Let us note  $T_{\Theta_n^{\ell}}$ , as the transformation associated with the parametric field  $\Theta_n^{\ell}$ . We have  $T_{\boldsymbol{\ell}} = T_{d\boldsymbol{U}^{\ell}}$  and  $T_{\Theta_n^{\ell}} = T_{d\boldsymbol{U}^{\ell}|\Xi_n}$ , where  $T_{d\boldsymbol{U}^{\ell}|\Xi_n}$  denotes the restriction of  $T_{\Theta_n^{\ell}}$  to the cube  $\Xi_n$ .

A neighborhood system  $V^{\ell}$  on the partition  $\Xi_{\ell}$  derives naturally from  $\mathcal{V}$ :

 $\forall n, m \in \{1 \cdots N_{\ell}\}, m \in V^{\ell}(n) \Leftrightarrow \exists s \in \Xi_n, \exists r \in \Xi_m/r \in \mathcal{V}(s). \mathcal{C}$  being the set of neighboring pairs on  $S^k$ , we must now distinguish between two types of such pairs: the pairs inside one cube and the pairs between two cubes:

$$\forall n \in \{1 \dots N_{\ell}\}, < s, r > \in \mathcal{C}_n^{\ell} \Leftrightarrow s \in \Xi_n, r \in \Xi_n \text{ and } r \in \mathcal{V}(s).$$
  
$$\forall n \in \{1 \dots N_{\ell}\}, \forall m \in V^{\ell}(n), < s, r > \in \mathcal{C}_{nm}^{\ell} \Leftrightarrow m \in V^{l}(n), s \in \Xi_n, r \in \Xi_m$$
  
and  $r \in \mathcal{V}(s).$ 

With these notations, at grid level  $\ell$ , the cost function can be modified as:

$$\overset{*}{U} (\Theta^{\ell}; A, B, \mathsf{T}_{0}, \boldsymbol{w}^{\ell}) = -\sum_{n=1}^{N_{\ell}} I(A, (\mathsf{T}_{\Theta_{n}^{\ell}} \circ \mathsf{T}_{\boldsymbol{w}^{\ell}} \circ \mathsf{T}_{0})(B_{|\Xi_{n}}))$$

$$+ \alpha \sum_{n=1}^{N_{\ell}} \left[ \sum_{m \in V^{\ell}(n) < s, r > \in \mathcal{C}_{m}^{\ell}} ||(\boldsymbol{w}_{s}^{\ell} + P_{s}\Theta_{n}^{\ell}) - (\boldsymbol{w}_{r}^{\ell} + P_{r}\Theta_{m}^{\ell})||^{2} \right]$$

$$+ \alpha \sum_{n=1}^{N_{\ell}} \left[ \sum_{< s, r > \in \mathcal{C}_{n}^{\ell}} ||(\boldsymbol{w}_{s}^{\ell} + P_{s}\Theta_{n}^{\ell}) - (\boldsymbol{w}_{r}^{\ell} + P_{r}\Theta_{n}^{\ell})||^{2} \right],$$

$$(8.5)$$

where  $B_{|\Xi_n}$  denotes the restriction of volume *B* to the cube  $\Xi_n$ . The minimization is performed with Gauss-Seidel iterative solver (each cube is iteratively updated while its neighbors are "frozen"). On each cube, Powell's algorithm [110, 111] is used to estimate the parametric affine increment.

#### 8.3.2.9 Implementation

The algorithm has been implemented in C + + using a template class for volumetric images.<sup>2</sup> A synopsis of the algorithm is presented in Fig. 8.3.

## 8.3.3 Results

## 8.3.3.1 Intensity Correction

We have evaluated the approach on various MR acquisitions. We present results on real data of the intensity correction, comparing the EM and SEM approaches and comparing the number of Gaussian laws used to model the histogram.

We have tested the approach on various T1-MR images and the algorithm has proved to be robust and reliable. Furthermore, it does not require any spatial alignment between the images to be corrected and can therefore be applied in various contexts: MR time series or MR of different subjects. Figure 8.4 presents cut-planes images of volumetric MR.

Figure 8.5 presents the effect of the correction using a EM algorithm and Fig. 8.6 the correction using a SEM algorithm. For each estimation scheme, we test a mixture of five (left) and seven Gaussian distributions to model the histogram. In each case, a fourth order parametric correction has been estimated.

<sup>&</sup>lt;sup>2</sup>http://www.irisa.fr/vista/Themes/Logiciel/VIsTAL/VIsTAL.html



Figure 8.3: Overview of the multiresolution and multigrid minimization. The convergence of the multigrid minimization is based on a percentage of cubes for which the parametric incremental field has been updated (percentage of the total number of cubes of the image partition).



Figure 8.4: MR images of different subjects. The intensity of tissue classes is different for source (top) and target (bottom) volume.



Figure 8.5: Intensity correction using the expectation maximization (EM) algorithm. The corrected source volume is presented, as well as the parametric intensity correction (to be compared with the identity function). The histogram has been modeled by five Gaussian distributions (top) and seven Gaussian distributions (bottom). Points represent the mean of Gaussian laws that model the histogram.



Figure 8.6: Intensity correction using the stochastic expectation maximization (SEM) algorithm. The corrected source volume is presented, as well as the parametric intensity correction (to be compared with the identity function). The histogram has been modeled by five Gaussian distributions (top) and seven Gaussian distributions (bottom). Points represent the mean of Gaussian laws that model the histogram.

Corrected source volumes and parametric correction functions are presented. The corrected volume seems visually more similar to the target volume (when comparing intensities of corresponding tissues). Modeling the histogram with seven classes seems more adequate in this context. This is actually more relevant from an anatomical point of view and provides more sample to estimate the correction function. The experiments we have conducted so far do not



**Before correction** 

After correction

Figure 8.7: Joint histogram before and after intensity correction. To compute the joint histogram, MR volumes have been previously rigidly registered by maximizing mutual information [88].

favor the SEM or the EM algorithm. There may be an indication that the SEM is more adapted in presence of field inhomogeneity and should be investigated further.

The relevance of this intensity correction can be assessed using the joint histogram (Fig. 8.7). To compute the joint histogram, a spatial alignment of the volumes needs to be performed. To do so, we estimate a rigid displacement that maximizes mutual information [88]. Figure 8.7 shows the joint histogram before and after intensity correction (using EM and seven Gaussian laws to model the histogram). It must be noted that the same displacement has been applied to the corrected and uncorrected volume (in other words, the effect of a possible misalignment is equal for both histograms). The joint histogram shows the relevance of the intensity correction.

## 8.3.3.2 Experiments on Simulated Data

**Evaluation on the MNI Phantom.** To evaluate the global registration method, we use the simulated data provided by the MNI<sup>3</sup> [34]. Data have been collected with three levels of noise and inhomogeneity. We design a synthetic deformation field made up of a global affine field with large deformations combined with local stochastic perturbations. We do not try to build a "realistic"

<sup>&</sup>lt;sup>3</sup>Brainweb: http://www.bic.mni.mcgill.ca/brainweb

field, but rather a field with the following properties: large deformations and local perturbations that modify the topology of the structures, in order to validate the basic hypothesis of our work. The "local" field is generated from 2,000 voxels which are randomly picked in the volume. For each voxel, each of the three components of the deformation is the realization of a Gaussian random variable of standard deviation 120 mm. We then perform a Gaussian smoothing with a small average deviation in order to propagate this perturbation to a local neighborhood while preserving discontinuities. The volumes and the results are shown on Fig. 8.8. We compare the multigrid method with a global affine registration method, in which a 12-parameter deformation is estimated for the entire volume.

To asses the quality of the registration, we compute the mean square error (MSE)<sup>4</sup> which is an indicator of the quality of the registration. However, it would be unfair to evaluate the registration only with a measure that is the underlying driving force of the estimation. Therefore, as we have the binary classification of the phantom, we can also assess the quality of the registration based on the overlap of two volumes: the first volume is the initial classification, i.e., a gold standard (gray matter/white matter), the second volume is the deformed classification, registered with the estimated deformation field. We then measure out overlapping ratios like the sensitivity, the specificity, and the total performance [142]. Results are presented in Table 8.1. Despite the use of binary classes, the resulting measures that we obtain are very satisfactory. Particularly, the robustness of the method is demonstrated in critical conditions (9% noise and 40% inhomogeneity), which are far tougher than in any realistic acquisition.

The numerical evaluation also allows to study the sensitivity of the algorithm with respect to the parameters of the algorithm, i.e., parameters of the robust estimators. We have two parameters to fix,  $\sigma_1$  and  $\sigma_2$ .  $\sigma_1$  corresponds to the hyperparameter of robust function  $\rho_1$ , associated with the similarity term, while  $\sigma_2$  corresponds to the hyperparameter of robust function  $\rho_2$  associated with the regularization term. We made the parameters  $\sigma_1$  and  $\sigma_2$  vary in a cube of size  $[1.0e^4, 1.0e^5] \times [1, 20]$  with step, respectively, of  $1.0e^4$  and 1 (which means that we performed the registration with 200 different sets of parameters), and we observe that the final result (the mean square error between the source volume

 $<sup>{}^{4}</sup>MSE = \frac{1}{N}\sum_{i=1}^{i=N}(I_{1}(i) - I_{2}(i))^{2}$ , where  $I_{1}$  and  $I_{2}$  are the volumes to compare, and N is the number of voxels.



original phantom deformed with the synthetic field Reconstructed volumes





Global affine registration Non-linear robust registration Difference volumes



Global affine registration



Multigrid robust registration

Figure 8.8: Results of the registration process on simulated data. The 3D MRI phantom has been deformed on the top of the figure. In the middle, the reconstructed volumes are shown and must be compared with the initial volume to evaluate the quality of the registration. On the bottom, the difference volumes show the benefits of non-linear registration.

and the reconstructed volume) varies less than 5% of the nominal MSE. This indicates that the sensitivity of algorithm with respect to these two parameters is very low.

For simulated data, mean square error (MSE) is a direct measure of the quality of the registration. Therefore we can also evaluate the influence of  $\ell_f$  (see section 8.3.2) on the computation time and on the accuracy of the registration.

performs	ance measures ar	e given for thre	e levels o	f noise an	d two registra	tion meth	ods. The r	egistration pro	ocesses ar	e
performe	ed until resolutio	n 0 (voxel size	1 mm <sup>3</sup> ). W	e manage	to recover up	to 93% of	the defor	mation even in	n presence	of
importar	nt noise (9%) and	image intensity	/ inhomog	eneity (40	)%). The CPU t	times are g	given for a	m Ultra Sparc	at 333 Mh	z
		Noise	%0		Noise	3%		Noise	6%	
		Inhomogeneity	%0		Inhomogeneity	20%		Inhomogeneity	40%	
		Target	Grey	White	Target	Grey	White	Target	Grey	White
		volume	matter	matter	volume	matter	matter	volume	matter	matter
	Computation time	10′			10′			10′		
	MSE	964.63	2679.49	1751.41	1104.22	3305.50	2171.13	2005.14	6933.05	5031.49
Global	sensitivity		93.78%	91.19%		93.26%	89.01%		83.21%	77.33%
affine	specificity		93.16%	93.72%		91.69%	92.48%		83.19%	85.42%
	total performance		93.27%	93.41%		91.97%	92.06%		83.19%	85.42%
	Computation time	55'			61'			76′		
	MSE	138.57	1383.46	886.53	233.23	1534.48	970.42	667.88	3186.49	1463.87
Robust	sensitivity		97.83%	97.35%		97.09%	96.36%		95.50%	93.27%
multigrid	specificity		94.28%	94.35%		94.76%	94.90%		90.73%	93.67%
	total performance		94.91%	94.71%		95.35%	95.03%		91.50%	93.80%

304

Table 8.1: Objective measures of the quality of the registration on simulated data. Specificity, sensitivity and total



Figure 8.9: Evolution of the MSE with respect to the grid level (at finest resolution 1 mm) and computation time needed to perform the registration until a given grid level. We observe that the MSE decreases significantly at the coarsest grid level, whereas at the finest grid level it continues to decrease, but less rapidly. At the same time, the computation time increases continuously. If we look at the difference between grid level 2 (the smallest cubes are of size  $2^2 \times 2^2 \times 2^2$  and the incremental deformation field is affine on each cube) and grid level 0 (the smallest cubes are reduced to a voxel and the incremental deformation field is translational for the smallest cubes), the computation time increases by 100%, whereas the MSE variation is only 5.3%. That suggests that, depending on the application, the user can make a compromise between the accuracy of the registration and the computation time if the resources are limited.

Figure 8.9 shows the evolution of the MSE with respect to the grid level (at finest resolution 1 mm) and also shows the computation time needed to perform the registration until a given grid level. We observe that the MSE decreases significantly at coarsest grid level, whereas at the finest grid level it continues to decrease, but less rapidly. At the same time, the computation time increases continuously. If we look at the difference between grid level<sup>5</sup> 2 and grid level<sup>6</sup> 0, the computation time increases by 100%, whereas the MSE variation is only 5.3%. That suggests that, depending on the application, the user can make a compromise between the accuracy of the registration and the computation time

<sup>&</sup>lt;sup>5</sup>The smallest cubes are of size  $2^2 \times 2^2 \times 2^2$  and the increment deformation field is affine on each cube.

<sup>&</sup>lt;sup>6</sup>The smallest cubes are reduced to a voxel and the increment deformation field is translational for the smallest cubes.



Figure 8.10: Synthetic data to validate the link between robust estimator on the regularization term and local changes of topology.

if its resources are limited. In our case, we find that  $\ell_f = 1$  (the smallest cubes are of size  $2 \times 2 \times 2$  and the allowed deformation is rigid on the smallest cube) is generally a good compromise.

**Importance of Robust Estimator.** We have introduced robust estimators in the registration process, in order to let local discontinuities of the deformation field occur. We now want to verify on simulated data the direct link between the introduction of a robust function and the possibility to locally change the topology of the structures. Therefore, we construct two volumes (see Fig. 8.10) to be registered, with a local modification of the topology. The volumes are composed of two homogenous classes, each one being defined by a unique gray level. With these two volumes, we obviously face the aperture problem, which is classical in the optical flow literature.

We first register the two volumes without any robust estimator. Results are presented in Fig. 8.11. The reconstructed volumes are computed with the target volume and the estimated deformation field with trilinear interpolation. One must therefore compare the reconstructed volume and the source volume to assess the quality of the registration. The different volumes shown in Fig. 8.11 correspond to different values of the parameter  $\alpha$ . This parameter balances the importance of the similarity term and the regularization term. When this parameter is high, the solution is smooth but the topology is not modified. When  $\alpha$  decreases, the solution is not smooth, the aperture problem is obvious, whereas the topology is not correctly modified.



Figure 8.11: Results of the registration without robust estimator. The different volumes correspond to different values of the parameter  $\alpha$ , and must be compared to the source volume.

We then perform the robust multigrid registration process, with a robust function only on the regularization term. Results are presented in Fig. 8.12, with two "extreme" values of the parameter  $\alpha$ . In that case, the modification of the topology is possible, while preserving the global smoothness of the solution. However, the aperture problem is still present in the tubular structure on the right. This experiment makes it possible to verify the link between the introduction of a robust estimator on the regularization term and the possibility to handle local change of topology. In addition, the robust registration process appears to be also more robust with respect to the parameter  $\alpha$ , because the results of the registration are very similar, when  $\alpha$  varies in a range of [100, 3,000].

#### 8.3.3.3 Experiment on Two Subjects

**Importance of Intensity Correction.** We first want to present cases where the registration method cannot work properly without a prior intensity



Figure 8.12: Results of the registration with a robust estimator on the regularization term. The reconstructed volumes must be compared to the source volume. We can handle with local topology changes, while preserving the global smoothness of the solution.

correction. Figure 8.13 shows the impact of the intensity correction step on the registration of two volumes. The source image is first corrected using the EM algorithm, seven classes to model the histogram and a fourth order parametric correction. Figure 8.13 presents the source image deformed toward the target image, as well as the difference image. While the registration has failed without intensity correction due to a very large intensity difference, it has performed successfully with an intensity correction step. It must be noted that the set of parameters is the same for both registration processes. That demonstrated the usefulness of such correction for a non-rigid registration task.

**Extensive Results for Two Subjects.** Results of the 3D method are presented in Figs. 8.14, 8.15, and 8.16. Two 3D MRI-T1 volumes of two different subjects are registered. The source volume, the target volume and the reconstructed volume are presented in Fig. 8.14. The reconstructed volume  $f_2(s + \hat{w}_s)$  is computed with the target volume  $f_2$  and the final displacement field  $\hat{w}$  by the way of a trilinear interpolation. To assess the quality of the registration, one must compare the source volume with the reconstructed volume.

We also present the volumes of difference, before and after registration in Fig. 8.15. In the same figure, the adaptive partition at grid level 3 is also presented (we do not present further grid levels for readability reasons). The difference volumes must be interpreted carefully, since we get the superposition of two



Figure 8.13: Impact of the intensity correction for the non-rigid registration. Top row: the images to be registered. Middle row: The source image is registered toward the target image, without intensity correction (left) and with correction (right). Bottom row: the differences images show the relevance of the intensity correction on the non-rigid registration.

errors: the first one is the registration error which comes from the anatomical variability that we could not apprehend. The second error is due to the difference of acquisition of the two volumes, which makes the histograms of the source and target volumes different.



Target volume  $f_2$ 

Figure 8.14: Final 3D results of the registration on real data. The volumes are T1-MRI acquisitions of two different subjects. The reconstructed volume is computed by trilinear interpolation with the target volume and the final dense displacement field. In order to evaluate the quality of the registration, we must therefore compare the source volume and the reconstructed volume.

In Fig. 8.16, the outliers are drawn, i.e., the data outliers map (variable  $\delta_s$ ) and the spatial outlier map (for each point *s*, we compute the mean of variable  $\beta_{sr}$  with respect to  $r \in \mathcal{V}(s)$ ). Looking at the data outliers map, the dark points represent areas where the optical flow hypothesis is inadequate, because of occlusions for instance (see the jaw in Fig. 8.16). For these points, the regularization term overwhelms the similarity term. Looking at the spatial outlier term, we observe that dark regions are located in the cortex. At that locations, the importance of the regularization term is reduced, and discontinuities can appear. The fact that discontinuities appear in the cortex is significant because we know that inter-subject variability is very high on the cortex.



Adaptive partition at grid level 3

Figure 8.15: Final 3D results of the registration on real data. Top: difference before registration. Middle: difference after registration. Bottom: adaptive partition at grid level 3. The difference volumes must be interpreted carefully, since we get the superposition of two errors: the first one is the registration error which comes from the anatomical variability that we could not apprehend. The second error is due to the difference of acquisition of the two volumes, which makes the two original histograms of the two volumes different.

The 3D deformation field is presented in Fig. 8.17. The vector field is subsampled in order to be easier to look at, and we also show the three components of the field on the sagittal view. Although discontinuities are visible, the general spatial coherence of the final deformation field is visible, due to the regularization. The field also confirms that there is no "block-effect" in the registration process.

The computation takes about 1 : 30 hour on an Ultra Sparc 30 (300 MHz). The volumes are  $256 \times 256 \times 200$ . We use three levels of resolution (k = 0, 1, 2) because the displacement amplitude may reach 30 voxels, and at each resolution level we perform the registration from grid level 4 until grid level 0.



Spatial outliers

Figure 8.16: Final 3D results of the registration on real data. Top: data outlier map. Middle: spatial outlier map. Looking at the spatial outlier term, we observe that dark regions are located in the cortex. Dark areas shows that the importance of the regularization term is reduced, and discontinuities can appear. The fact that discontinuities appear in the cortex is significant because we know that inter-subject variability is very high on the cortex.



3 components of the field on the sagittal view

Figure 8.17: Top: deformation field. Bottom: Images of the three components of the field on the sagittal view. The 3D deformation field is subsampled in order to be easier to look at. Although discontinuities are visible, the general spatial coherence of the final deformation field is visible, due to the regularization. The field also confirms that there is no "block-effect" in the registration process.
#### 8.3.3.4 Experiments on a Dataset of 18 Subjects

In order to validate the registration method on a larger database, we acquire MRI-T1 volumetric data of 18 patients. One subject is chosen as the reference subject. We then perform the registration between the reference volume (source) and each of the other subjects (target) using *always* the same set of parameters for the algorithm. Finally, we get 17 reconstructed volumes that can be compared to the reference volume. We average all the reconstructed volume in order to have a global overview of the quality of the method.

**Average Deformed Volume.** Figures 8.18 and 8.19 present the averaging between 17 patients after a global affine registration (top), after a quadratic multigrid registration, i.e., the method without robust estimators (middle), and the average volume after a robust multigrid registration (bottom). After global affine registration and averaging, we notice that the internal anatomical structures are blurred, because the registration is not precise enough. However, after a robust multigrid registration, we may distinguish precisely the contours of anatomical structures, such as ventricles, deep nuclei, white matter tracks, and even cortical regions (sylvian fissure and parietal region for instance).

The comparison between the quadratic registration and the robust registration shows the benefit of robust functions, because cortical regions are better registered. The MSE between the reference volume and the averaged volume is 892 for quadratic registration, and drops to 584 for robust registration. We must note that, considering two subjects, the MSE is not a good absolute measure of the quality of the registration because of the acquisition (a simple translation between the two histogram can lead to large MSE). However, the MSE is a good relative measure to compare two registration processes over a large database.

These experiments clearly show the significant impact of robust estimators. All the more, it validates the assumption that it is necessary to let discontinuities appear in the deformation field to register brains correctly. These experiments also demonstrate the robustness of the method (robustness with respect to the acquisitions and also with respect to the algorithm parameters) over a realistic database of subjects.

**Overlapping of Brain Tissues.** The evaluation must not be based only on a measure that is more or less related to the image similarity. Therefore, as in



Global affine registration



Quadratic multigrid registration



Robust multigrid registration

Figure 8.18: Results of experiments on a database of 18 subjects. One subject was chosen as the reference subject (see Fig. 8.19), and we averaged all the reconstructed volumes after global affine registration (top), after quadratic multigrid registration (middle) and after robust multigrid registration (bottom). We kept the same set of parameters for all the subjects. This demonstrates the robustness of the method, and the importance of robust estimators (the quadratic registration is less accurate on the cortex).

section 8.3.3, we evaluate in this section the registration process by computing the overlap between the tissues (gray matter and white matter) of the reference volumes and the tissues of each studied volume after registration. We measure the overlap with the total performance, which has already been presented in section 8.3.3.

The extraction of gray matter and white matter is performed using a technique presented in [85]. It consists in a 3D texture analysis to compute statistical attributes of each voxel. A clustering procedure is used to find the initial discrimination of the data, and a bayesian relaxation refines the primary decision.





Reference subject

Figure 8.19: Results of experiments on a database of 18 subjects. One subject was chosen as the reference subject (bottom), and we averaged all the reconstructed volumes after robust multigrid registration (top). We kept the same set of parameters for all the subjects, which demonstrates the robustness of the method. Results of the averaging after registration show the accuracy of the registration (after averaging we can distinguish precisely anatomical structures such as ventricles, deep nuclei, white matter tracks and even cortical regions).

For gray matter tissue, the average overlap after registration is 93.9% (mean of total performance). For white matter, the average overlap is 94.9%. If we perform a rigid registration by maximization of mutual information, we obtain 88.3% and 87.1% of average overlap, for gray matter and white matter, respectively. These measures must be interpreted carefully for two reasons. We use binary classes (and not fuzzy classes) and a simple trilinear interpolation scheme, which may introduce some error. Furthermore, the classification algorithm introduces errors that disturb the overlapping measure. In the last 5% to recover, it is difficult to distinguish what is due to the registration process and what is due to interpolation and segmentation errors. However, these overlapping measures show the benefit of non-rigid registration.

#### 8.3.3.5 Experiments on Multimodal Datasets

The extension of Romeo to multimodal dataset was motivated by a particular problem: the correction of distortions in echo-planar images. Among the functional images of the brain, fMRI is an appealing technique because it offers a good trade-off between spatial and temporal resolution. To increase its temporal resolution, echo-planar imaging (EPI) is used because it makes possible to collect at least five slices per second at a reduced spatial resolution. The drawback of this impressive acquisition rate is that it may introduce artifacts and distortions in the data. More details about these distortions can be found in [78].

If the distortions do not vary during the time series, they will not affect much the detection of subtle signal changes, but they will perturb the localization of the functional activity once being overlapped to the anatomical volume. It becomes necessary to correct these geometrical distortions in order to accurately identify activated areas.

**Simulated Data.** To evaluate the multimodal registration method, we use the simulated database of the MNI (Brainweb)<sup>7</sup> [34]. The T1-weighted MR volume is the reference volume (3% noise and 9% inhomogeneity), whereas the T2-weighted MR volume is the floating volume.

From the T2-weighted MR volume, we extract a subvolume and we apply a rigid transformation (three rotations and three translations). To simulate local geometrical distortions, we apply a thin-plate spline [16] deformation to the volume. The thin-plate deformation is computed by choosing one point in the volume and a displacement for this point. We choose a displacement of magnitude 5 voxels, with no privileged direction. Furthermore, the thin-plate deformation field is constrained to be naught at the border of the volume.

After rigid registration (see Fig. 8.20), distortions are clearly visible. On the axial view, ventricles are not well registered; on the sagittal and coronal view, the sagittal mid-plane is not well aligned. We then perform the multigrid non-rigid registration from grid level 7 until grid level 5 to avoid useless computational efforts. In this case we do not need to estimate a dense transformation, since the distortions are rather smooth and regular. Furthermore, the statistical similarity measure is only meaningful for a large number of voxels, i.e., for large cubes. After non-rigid registration, the internal structures are accurately registered (see ventricles on the axial view, and sagittal mid-plane on the coronal view for instance).

<sup>&</sup>lt;sup>7</sup>http://www.bic.mni.mcgill.ca/brainweb



Figure 8.20: Results of the registration on simulated data. Top: results after rigid registration. Distortions are visible on axial view (ventricles) and on coronal view (sagittal mid-plane). Bottom: results after performing a 3D multimodal non-rigid registration. Misregistration is significantly reduced (see ventricles on axial view for instance).

In Fig. 8.21, we present the evolution of the similarity measure, computed for all the volume. As the hierarchical estimation is performed, the similarity increases, which means that the dependence between the volumes to be registered also increases, leading to a more and more accurate registration. This is a numerical confirmation of visual assessment.

At this stage, the evaluation is not completely fair, since the criterion is related to the similarity used to drive the registration process. As we have the segmentation of the phantom (gray matter and white matter classes), we can evaluate objectively the registration process. We deform the segmentation volumes as described at the beginning of section (8.3.3). We can assess the quality of the registration by computing overlapping measures (specificity, sensitivity and total performance, see [142] for tutorial) between the initial classes and the deformed classes, once registered with the estimated deformation field. Let us recall that:

sensitivity 
$$= TP/(TP + FN)$$
  
specificity  $= TN/(FP + TN)$  (8.6)  
total performance  $= (TP + TN)/(TP + FP + TN + FN)$ ,



Figure 8.21: Evolution of the similarity measure. As the hierarchical estimation is performed, the similarity increases, which means that the volumes are more accurately registered.

where *TP*, *TN*, *FP* and *FN*, respectively, denote the number of true positive, true negative, false positive, and false negative points.

These numerical results are shown in Table (8.2). At the end of grid level 5, we manage to recover more than 95% of the segmentation. This result is satisfactory, due to the use of binary classes, and due to a simple trilinear interpolation scheme that causes artefacts [108].

**Real Data.** We have performed the algorithm on real data (see Fig. 8.22). The patient has a cyst and a bone tumor, therefore the multiple interfaces (air/cyst/bone) introduce large distortions that are visible after rigid registration. For instance, on the left hemisphere, distortions are clearly visible on the posterior part of the cyst.

There are many artefacts in this fMRI acquisition: there has been signal saturation and signal drops (visible in the cyst and in the border of the skull). This illustrates the difficulty of registering real clinical data. Although quality of the results is quite difficult to quantify, we can see that the cyst (on the axial view) and the ventricles (on the sagittal view) are better aligned after non-rigid registration.

Table 8.2: Numerical evaluation of the multimodal registration method on simulated data. The overlapping measures (specificity, sensitivity, and total performance) are computed after rigid registration and at each grid level of the non-rigid registration process

Registration	Overlap measure	Grey matter	White matter
Rigid	sensibility	74.7%	76.6%
	specificity	93.0%	92.8%
	Total performance	87.0%	87.6%
Non-rigid	sensibility	84.7%	86.0%
grid level 7	specificity	97.2%	96.2%
	Total performance	93.2%	92.9%
Non-rigid	sensibility	86.6%	86.8%
grid level 6	specificity	98.5%	97.3%
	Total performance	94.6%	93.9%
Non-rigid	sensibility	87.5%	87.0%
grid level 5	specificity	98.9%	98.0%
	Total performance	95.8%	95.3%



Figure 8.22: Results of clinical data. Top: results of the rigid registration. The multiple artefacts are visible: distortions, signal saturation, signal drops. Bottom: Results of the non-rigid registration. The registration is more accurate, in particular, for the ventricles and for the cyst. The data are courtesy of "laboratoire IDM, Hopital de Pontchaillou, Rennes".

# 8.4 Conclusion

This chapter has presented an overview of the classification of non-rigid registration methods with particular focus on non-rigid registration of brains of different subjects. Methods can be broadly classified into two groups: geometric methods that rely on the extraction and matching of sparse features (points, curves, surfaces), and photometric (or intensity-based) methods that rely on image luminance directly.

Geometric methods reduce the dimensionality of the problem and are consistent in the vicinity of features used for registration. However, the registration might be incorrect far from used features. Photometric methods use all the available information present in the volume but lead to a complex problem involving a very large number of variables.

We have presented here the Romeo algorithm (Robust Multigrid Elastic registration based on optical flow) that refers to photometric methods. Romeo uses the optical flow as a similarity measure and relies on an efficient multiresolution and multigrid optimization scheme. Robust estimators are introduced to limit the effect of erroneous data and to preserve discontinuities of the deformation field when needed. Prior to the non-rigid registration step, two preprocessing steps are performed: a rigid registration by maximization of mutual information and an intensity correction so that the luminance of the volumes to be registered are comparable. An extension to multimodal data has been presented. The multiresolution and multigrid framework is flexible enough to be adapted to multimodal similarity measures such as mutual information for instance.

It has been shown that photometric methods fail in matching cortical structures such as cortical sulci, for instance [74]. This can be explained by the high variability of cortical structures among subjects. Anatomists have pointed out [103] that cortical sulci of different subjects are very different in shapes. This has motivated mixed approaches where a photometric registration method incorporates sparse anatomical structures so as to drive the registration process [22, 29, 36, 68, 73, 79, 141].

In this context, it must be noted that validation is difficult and should be investigated further. Validation of non-rigid registration methods on anatomical structures have been conducted [74, 119]. However, the impact of non-rigid registration methods on functional data is still unknown. As a matter of fact, since these methods are dedicated to anatomical and functional normalization, it would be interesting to know how much of the intersubject functional variability can be understood and compensated by anatomical non-rigid registration. This is a challenging research subject that requires a better knowledge about the relationship between the anatomy of the brain and its functional organization.

# Questions

- 1. What are photometric and geometric registration methods? How can these methods be compared?
- 2. What is optical-flow?
- 3. What is the aperture problem? How can it be solved?
- 4. What are the advantages and drawbacks of optical-flow based registration?
- 5. What are robust estimators?
- 6. What are the advantages and drawbacks of robust M-estimators?
- 7. How useful is a multiresolution scheme?
- 8. How should the Gaussian standard deviation be chosen for building the multiresolution pyramid?
- 9. What is a multigrid optimization scheme?
- 10. What are the different options to regularize the deformation field?

# Bibliography

- Ashburner, J., Andersson, J., and Friston, K. J., Image registration using a symetric prior in three dimensions. Human Brain Mapping, Vol. 9, pp. 212–225, 1998.
- [2] Ashburner, J. and Friston, K., Nonlinear Spatial Normalization using basis functions. Human Brain Mapping, Vol. 7, No. 4, pp. 254–266, 1999.
- [3] Cuadra, M. Bach, Platel, B., Solanas, E., Butz, T., and Thiran, J. P., Validation of tissue modelization and classification techniques in T1-weighted MR brain images, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Dohi, T. and Kikinis, R., ed., No. 2488 In Lecture Notes in Computer Science, pp. 290–297, Springer, Tokyo, September 2002.
- [4] Bajcsy, R. and Broit, C., Matching of deformed images, In: Proc. International Conference on Pattern Recognition, Vol. 1, pp. 351–353, IEEE, New York, Munich, West Germany, Oct. 1982.
- [5] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching, Computer Vision, Graphics, Image Processing, Vol. 46, pp. 1–21, 1989.
- [6] Barr, A., Global and local deformations of solid primitives, Computer Graphics (SIGGRAPH), Vol. 18, no. 3, pp. 21–30, 1984.
- [7] Barron, J., Fleet, D., Beauchemin, S., and Burkitt, T., Performance of optical flow techniques, In: Proc. Conference on Computer Vision Pattern Recognition, pp. 236–242, Champaign, Illinois, Jun. 1992.
- [8] Bascle, B., Contributions et applications des modèles déformables en vision par ordinateur, Ph.D. thesis, Université de Nice Sophia-Antipolis, Sep. 1994.
- [9] Battiti, R., Amaldi, E., and Koch, C., Computing optical flow over multiple scales: an adaptive coarse-to-fine strategy, International Journal of Computer Vision, Vol. 6, No. 2, pp. 133–146, 1991.
- [10] Beauchemin, S. and Barron, J., The computation of optical flow, ACM Computing Surveys, Vol. 27, No. 3, pp. 433–467, 1995.

- [11] Bergen, J., Anadan, P., Hanna, K., and Hingorani, R., Hierarchical modelbased motion estimation, In: Proc. European Conference on Computer Vision, pp. 5–10, 1991.
- [12] Besl, P. and McKay, N., A method for registration of 3D shapes, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 14, No. 2, pp. 239– 255, 1992.
- [13] Black, M. and Anandan, P., The robust estimation of multiple motions: parametric and piecewise-smooth flow fields, Computer Vision and Image Understanding, Vol. 63, No. 1, pp. 75–104, 1996.
- [14] Black, M. and Rangarajan, A., On the unification of line processes, outlier rejection, and robust statistics with application in early vision, International Journal of Computer Vision, Vol. 19, No. 1, pp. 57–91, 1996.
- [15] Blake, A. and Isard, M., Active contours. Springer, 1998.
- [16] Bookstein, F., Principal warps: Thin-plate splines and the decomposition of deformations, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 6, pp. 567–585, 1989.
- [17] Bro-Nielsen, M. and Gramkow, C., Fast fluid registration of medical images, In: Proc. Visualization in Biomedical Computing, Hohne, K-. H. and Kikinis, R., ed., No. 1131 In Lecture Notes in Computer Science, pp. 267–276, Springer, Sep. 1996.
- [18] Broit, C., Optimal registration of deformed images, Ph.D. Dissertaition, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, 1981.
- [19] Brown, L. F., A survey of image registration techniques, ACM: Computer Surveys, Vol. 24, No. 4, pp. 325–376, 1992.
- [20] Burt, J. P., Multiresolution image processing and analysis, Chapter 2 : The pyramide as a structure for efficient computation, pp. 6–38, No. 12 In Springer series in Information Science, Springer-Verlag, 1984.
- [21] Cachier, P., Bardinet, E., Dormont, D., Pennec, X., and Ayache, N., Iconic feature based nonrigid registration: the PASHA algorithm, Computer Vision and Image Understanding, Vol. 89, No. 2/3, pp. 272–299, 2003.

- [22] Cachier, P., Mangin, J. F., Pennec, X., Rivière, D., Papadopoulos-Orfanos, D., Régis, J., and Ayache, N., Multisubject non-rigid registration of brain MRI using intensity and geometric features, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Niessen, W. and Viergever, M., eds., No. 2208 in Lecture Notes in Computer Science, pp. 734–742, 2001.
- [23] Celeux, G., Chauveau, D., and Diebolt, J., Stochastic version of the EM algorithm: an experimental study in the mixture context, Journal of Computationnal Statistical Computation and Simulation, Vol. 55, pp. 287–314, 1996.
- [24] Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M., Deterministic edge preserving regularization in computed imaging, IEEE Transactions Image Processing, Vol. 6, No. 2, pp. 298–311, 1997.
- [25] Chen, C., Huang, T., and Arrott, M., Modeling, analysis and vizualisation of left-ventricule shape and motion by hierarchical decomposition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 4, 1994.
- [26] Chen, G., Kessler, M., and Pitluck, S., Structure transfer between sets of three dimensionnal medical imaging data, In: VI Conference on Computer Graphics, Vol. 6, pp. 171–177, 1985.
- [27] Christensen, G., Rabbit, R., and Miller, M. I., Deformable templates using large deformation kinematics, IEEE Transactions on Image Processing, Vol. 5, No. 10, pp. 1435–1447, 1996.
- [28] Christensen, G. and Johnson, H. J., Consistent image registration, IEEE Transactions Medical Imaging, Vol. 20, No. 7, pp. 568–582, 2001.
- [29] Chui, H., Rambo, J., Duncan, J., Schultz, R., and Rangarajan, A., Registration of cortical anatomical structures via robust 3D point matching, In: Proc. Information Processing in Medical Imaging, Cuba, *et al.*, ed., No. 1613 In: Lecture Notes in Computer Science, pp. 168–181, Springer Verlag, 1999.

- [30] Chui, H. and Rangarajan, A., A new point matching algorithm for nonrigid registration, Computer Vision and Image Understanding, Vol. 89, No. 2/3, pp. 114–142, 2003.
- [31] Cohen, I. and Herlin, I., Optical flow and phase portrait methods for environmental satellite image sequences, In: Proc. European Conference on Computer Vision, pp. II:141–150, Cambridge, UK, April 1996.
- [32] Cohen, L. and Cohen, I., Finite element method for active contour models and ballons for 2D and 3D images, IEEE Transactions of Pattern Analysis and Machine Intelligence, Vol. 15, No. 11, pp. 1131– 1147, 1993.
- [33] Collignon, A., Vanderneulen, D., Suetens, P., and Marchal, G., 3D multimodality medical image registration using feature space clustering, In: Proc. of Computer Vision, Virtual Reality and Robotics in Medecine, pp. 195–204, Nice, France, 1995.
- [34] Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J., and Evans, A. O., Design and construction of a realistic digital brain phantom, IEEE Transactions on Medical Imaging, Vol. 17, No. 3, pp. 463–468, June 1998.
- [35] Collins, L. and Evans, A., Animal : validation and applications of nonlinear registration-based segmentation, International Journal of Pattern Recognition Artificial Intelligence, Vol. 8, No. 11, pp. 1271–1294, 1997.
- [36] Collins, L., Goualher, Le, G., and Evans, A., Non-linear cerebral registration with sulcal constraints, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Colchester, A. and Delp, S., eds., No. 1496 in Lecture Notes in Computer Science, pp. 974–985, Springer, October 1998.
- [37] Corpetti, T., Mémin, E., and Pérez, P., Dense estimation of fluid flows, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 3, pp. 365–380, 2002.
- [38] Davatzikos, C., Spatial transformation and registration of brain images using elastically deformable models, Computer Vision and Image Understanding, Vol. 66, No. 2, pp. 207–222, 1997.

- [39] Davatzikos, C., Prince, L., and Bryan, N., Image registration based on boundary mapping, IEEE Transactions on Medical Imaging, Vol. 15, No. 1, pp. 111–115, 1996.
- [40] Dawant, B., Hartmann, S., Thirion, J. P., Maes, F., Vandermeulen, D., and Demaerel, P., Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations: Part I, Methodology and validation on normal subjects, IEEE Transactions on Medical Imaging, Vol. 18, No. 10, pp. 909–916, 1999.
- [41] Declerck, J., Subsol, G., Thirion, J. P., and Ayache, N., Automatic retrieval of anatomical structures in 3D medical images, In: Proc. of Computer Vision, Virtual Reality and Robotics in Medicine, Ayache, N., ed., No. 905 in LNCS, pp. 153–162, Springer, Nice, France, April 1995, Electronic version: ftp://ftp.inria.fr/INRIA/publication/RR/RR-2485.ps.gz.
- [42] Delaney, A. and Bresler, Y., Globally convergent edge-preserving regularized reconstruction: an application to limited-angle tomography, IEEE Transanction Image Processing, Vol. 7, No. 2, pp. 204–221, 1998.
- [43] Delingette, H., Hebert, M., and Ikeuchi, K., Shape representation and image segmentation using deformable models, Image and Vision Computing, Vol. 10, No. 3, pp. 132–144, 1992.
- [44] Dempster, A. P., Laird, N. M., and Rubin, D. B., Maximum likelihood for incomplete data via the EM algorithm, Journal of the Royal Statistical Society, Vol. 39, pp. 1–38, 1977.
- [45] Deriche, R., Recursively implementing the gaussian and its derivatives, Technical Report 1893, INRIA, http://www.inria.fr/RRRT/RR-1893.html, April 1993.
- [46] Devlaminck, V., A functional for compressive and incompressive elastic deformation estimation, IEEE Signal Processing Letters, Vol. 6, No. 7, pp. 162–164, 1999.
- [47] Duchon, J., Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, RAIRO Analyse Numérique, Vol. 10, pp. 5–12, 1976.

- [48] Enkelmann, W., Investigations of multigrid algorithms for the estimation of optical flow fields in image sequences, Computer Vision, Graphics, Image Processing, Vol. 43, No. 2, pp. 150–177, 1988.
- [49] Evans, A., Collins, L., and Milner, B., An MRI-based stereotaxic atlas from 250 young normal subjects, Society for Neuroscience Abstract, Vol. 18, pp. 408, 1992.
- [50] Evans, A. C., Dai, W., Collins, D. L., Neelin, P., and Marrett, T., Warping of computerized 3D atlas to match brain image volumes for quantitative neuroanatomical and functionnal analysis, In: Proc. of the International Society of Optical Engineering: Medical Imaging V, SPIE, 1991.
- [51] Feldmar, J. and Ayache, N., Locally affine registration of free-form surfaces, In: Proc. Conference on Computer Vision Pattern Recognition, pp. 496–500, Seattle, June 1994.
- [52] Feldmar, J. and Ayache, N., Rigid, affine and locally affine registration of free-form surfaces, International Journal of Computer Vision, Vol. 18, No. 2, pp. 99–119, 1996.
- [53] Feldmar, J., Declerck, J., Malandain, G., and Ayache, N., Extension of the ICP algorithm to nonrigid intensity based registration of 3D images, Computer Vision and Image Understanding, Vol. 66, No. 2, pp. 193–206, 1997.
- [54] Ferrant, M., Warfield, S., Guttmann, C., Mulkern, R., Jolesz, F., and Kikinis, R., 3D image matching using a finite element based elastic deformation model, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Taylor, C. and Colchester, A., eds., No. 1679 in Lecture Notes in Computer Science, pp. 202–209, Cambridge, UK, September 1999.
- [55] Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., Heather, J. D., and Frackowiak, R. SJ., Spatial registration and normalisation of images, Human Brain Mapping, Vol. 2, pp. 165–189, 1995.
- [56] Gabrani, M. and Tretiak, O., Surface-based matching using elastic transformation, Pattern Recognition, Vol. 32, pp. 87–97, 1999.

- [57] Gaens, T., Maes, F., Vandermeulen, D., and Suetens, S., Non-rigid multimodal image registration using mutual information, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, No. 1496 in Lecture Notes in Computer Science, pp. 1099–1106, October 1998. Springer Verlag, Boston.
- [58] Gee, J., Briquer, Le, L., Barillot, C., and Haynor, D., Probabilistic matching of brain images, In: Proc. Information Processing in Medical Imaging, Bizais, *et al.*, ed., Kluwer Academic Publisher, Brest, Jun. 1995.
- [59] Gee, J. C., Reivicj, M., and Bajcsy, R., Elastically deforming 3D atlas to match anatomical brain images, Journal of Computer Assisted Tomography, Vol. 17, No. 2, pp. 225–236, 1993.
- [60] Gerlot, P. and Bizais, Y., Image registration: a review and a strategy for medical application, In: Proc. Information Processing in Medical Imaging, De Graaf, C. and Viergever, M., eds., pp. 81–89, Plenum Press, New York, 1988.
- [61] Goshtasby, A., Staib, L., Studholme, C., and Terzopoulos, D., eds., Computer Vision and Image Understanding: Special issue on nonrigid image registration, Vol. 89 No. 2/3, Academic Press, February/March 2003.
- [62] Keyserlingk, Graf von, D., Niemann, K., and Wasel, J., A quantitative approach to spatial variation of human cerebral sulci, Acta Anatomica, Vol. 131, pp. 127–131, 1988.
- [63] Guimond, A., Roche, A., Ayache, N., and Meunier, J., Multimodal brain warping using the demons algorithm and adaptative intensity corrections, IEEE Transactions on Medical Imaging, Vol. 20, No. 1, pp. 58–69, 2001.
- [64] Gupta, S. and Bajcsy, R., Volumetric segmentation of range images of 3D objects using superquadrics models, CVGIP: Image Understanding, Vol. 58, No. 3, pp. 302–326, 1993.
- [65] Gupta, S. and Prince, J., On variable brightness optical flow for tagged MRI, In: Proc. Information Processing in Medical Imaging, Bizais *et al.*, ed, pp. 323–334, Kluwer Academic Publisher, Brest, June 1995.

- [66] Guéziec, A. and Ayache, N., Smoothing and matching of 3D space curves, International Journal of Computer Vision, Vol. 12, No. 1, pp. 79–104, 1994.
- [67] Hackbusch, W., Multigrid Methods and Applications, Springer Verlag, 1985.
- [68] Hartkens, T., Hill, D., Castellano-Smith, A., Hawkes, D., Maurer, C., Martin, A., Hall, W., Liu, H., and Truwit, C., Using points and surfaces to improve voxel-based non-rigid registration, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Dohi, T. and Kikinis, R., eds., No. 2488 in Lecture Notes in Computer Science, pp. 565–572, Springer, Tokyo, September 2002.
- [69] He, Renjie and Narayana, Ponnada A. Global optimization of mutual information: application to three dimensional retrospective registration of magnetic resonance images. Computerized Medical Imaging and Graphics, Vol. 26, No. 4, pp.277–292, 2002.
- [70] Hebert, M. and Ikeuchi, H., K.and Delingette, A spherical representation for recognition of free form surfaces, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 7, pp. 681–690, 1995.
- [71] Hellier, P., Consistent intensity correction of MR images, In: IEEE conference on Image Processing, ICIP 2003, Barcelona, September 2003.
- [72] Hellier, P. and Barillot, C., Multimodal non-rigid warping for correction of distortions in functional MRI, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, MICCAI'00, Delp, S. *et al.*, ed., No. 1935 in Lecture Notes in Computer Science, pp. 512–521, Springer, Pittsburgh, Pennsylvania, October 2000.
- [73] Hellier, P. and Barillot, C., Coupling dense and landmark-based approaches for non-rigid registration, IEEE Transactions on Medical Imaging, Vol. 22, No. 2, pp. 217–227, 2003.
- [74] Hellier, P., Barillot, C., Corouge, I., Gibaud, B., Goualher, Le G., Collins, D. L., Evans, A., Malandain, G., Ayache, N., Christensen, G. E., and Johnson, H. J., Retrospective evaluation of inter-subject brain registration,

IEEE Transactions on Medical Imaging, Vol. 22, No. 9, pp. 1120–1130, 2003.

- [75] Hellier, P., Barillot, C., Mémin, E., and Pérez, P., Hierarchical estimation of a dense deformation field for 3D robust registration, IEEE Transactions on Medical Imaging, Vol. 20, No. 5, pp. 388–402, 2001.
- [76] Horn, B. and Schunck, B., Determining optical flow, Artificial Intelligence, Vol. 17, pp. 185–203, August 1981.
- [77] Huber, P., Robust Statistics, Wiley, 1981.
- [78] Jezzard, P. and Clare, S., Sources of distortions in functional MRI data, Human Brain Mapping, Vol. 8, pp. 80–85, 1999.
- [79] Johnson, H. J., and Christensen, G. E., Consistent Landmark and intensity-based Image registration, IEEE Transactions on Medical Imaging, Vol. 21, No. 5, pp. 450–461, 2002.
- [80] Jones, R. and Witte, D., Signal intensity artefacts in clinical MR imaging, Radiographics, Vol. 20, pp. 893–901, 2000.
- [81] Kikinis, R., Shenton, M., Iosifescu, D., McCarley, R., Saiviroonporn, P., Hokama, H., Robatino, A., Metcalf, D., Wible, C., Portas, C., Donnino, R., and Jolesz, F., A digital brain atlas for surgical planning, model driven segmentation and teaching, IEEE Transactions on Visualization and Computer Graphics, Vol. 2, No. 3, pp. 232–241, 1996.
- [82] Konrad, J. and Dubois, E., Bayesian estimation of motion vector fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, No. 9, pp. 910–927, 1992.
- [83] Kovacevic, N., Lobaugh, N. J., Bronskill, M. J., Levine, B., Feinstein, A., and Black, S. E., A robust method for extraction and automatic segmentation of brain images, Neuroimage, Vol. 17, pp. 1087–1100, 2002.
- [84] Kybic, J., Thevenaz, P., Nirkko, A., and Unser, M., Unwarping of unidirectionally distorted EPI images, IEEE Transactions on Medical Imaging, Vol. 19, No. 2, pp. 80–93, 2000.

- [85] Lachmann, F. and Barillot, C., Brain tissue classification from mri data by means of texture analysis, In: Medical Imaging VI: Image Processing, SPIE Press, ed., Vol. 1652, pp. 72–83, 1992.
- [86] Leemput, K. V., Maes, F., Vandermeulen, D., and Suetens, P., Automated model-based tissue classification of MR images of the brain, IEEE Transactions on Medical Imaging, Vol. 18, No. 10, pp. 897–908, 1999.
- [87] Lester, H. and Arridge, S., A survey of hierarchical non-linear medical image registration, Pattern Recognition, Vol. 32, pp. 129–149, 1999.
- [88] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximisation of mutual information, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, April 1997.
- [89] Maintz, J., Meijering, E., and Viergever, M., General multimodal elastic registration based on mutual information, In: Proc. of Medical Imaging: Image Processing, Hanson, M., ed., No. 3338 in SPIE Proceedings, pp. 144–154, San Diego, USA, April 1998.
- [90] Maintz, J. and Viergever, M. A., A survey of medical image registration, Medical Image Analysis, Vol. 2, No. 1, pp. 1–36, 1998.
- [91] Maurer, C. R., and Fitzpatrick, J. M., A review of medical image registration, In: Interactive Image Guided Neurosurgery, pp. 17–44, American Association of Neurological Surgeons, 1993.
- [92] Mazziotta, J., Toga, A., Evans, A., Fox, P., and Lancaster, J., A probabilistic atlas of the human brain: theory and rationale for its development, Neuroimage, Vol. 2, pp. 89–101, 1995.
- [93] McCormick, S., Multilevel Adaptive Methods for Partial Differential Equations, SIAM, Philadelphia, Pennsylvania, 1989.
- [94] McInerney, T. and Terzopoulos, D., Deformable models in medical image analysis: A Survey, Medical Image Analysis, Vol. 1, No. 2, pp. 91–108, 1996.
- [95] Meinguet, J., Multivariate interpolation at arbitrary points made simple, Journal of Applied Mathematics, Vol. 30, 1979.

- [96] Metaxas, D. and Terzopoulos, D., Shape and nonrigid motion estimation through physics-based synthesis, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 6, pp. 580–591, 1993.
- [97] Miller, M. I., Christensen, G. E., Amit, Y., and Grenander, U., Mathematical textbook of deformable neuroanatomies, In: Proceedings of the National Academy of Sciences, Vol. 90, No. 24, pp. 11944–48, December 1993.
- [98] Mémin, E. and Pérez, P., Dense estimation and object-based segmentation of the optical flow with robust techniques, IEEE Transactions on Image Processing, Vol. 7, No. 5, pp. 703–719, 1998.
- [99] Mémin, E. and Risset, T., Vlsi design methodology for edgepreserving image reconstruction, Real-Time Imaging, 2000, Available at: http://www.irisa.fr/EXTERNE/bibli/pi/1220/1220.html.
- [100] Monga, O., Ayache, N., and Sander, P., From voxel to intrinsic surface feature, Computer Vision and Image Understanding, Vol. 10, No. 6, pp. 403–417, 1992.
- [101] Musse, O., Heitz, F., and Armspach, J. P., 3D deformable image matching using multiscale minimization of global energy functions, In: Proc. Conference Computer Vision Pattern Recognition, Vol. 2, pp. 478–485, Fort Collins, Colorado, June 1999.
- [102] Odobez, J. M., and Bouthemy, P., Robust multiresolution estimation of parametric motion models, Journal of Visual Communication and Image Representation, Vol. 6, No. 4, pp. 348–365, 1995.
- [103] Ono, M., Kubik, S., and Abernathey, C., Atlas of the Cerebral Sulci, Verlag, 1990.
- [104] Peckar, W., Schnorr, C., Rohr, K., and Stiehl, S., Parameter-free elastic registration of brain images using prescribed displacements, Technical report, Universitat Hamburg Hachbereich Informatik, 1997.
- [105] Peckar, W., Schnorr, C., Rohr, K., and Stiehl, S., Parameter-free elastic deformation approach for 2D and 3D registration using prescribed

displacement, Journal of Mathematical Imaging and Vision, Vol. 10, No. 2, pp. 143–162, 1999.

- [106] Pennec, X., Cachier, P., and Ayache, N., Understanding the demon's algorithm : 3D non-rigid registration by gradient descent, In: Proc. of Medical Image Computing and Computer-Assisted Intervention, Taylor, C. and Colchester, A., eds., No. 1679 in Lecture Notes in Computer Science, pp. 597–605, Springer, Cambridge, UK, September 1999.
- [107] Pentland, A. and Sclaroff, S., Closed-form solutions for physicallybased shape modeling and recognition, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 7, pp. 715–729, 1991.
- [108] Pluim, J., Maintz, J., and Viergever, M., Interpolation artefacts in mutual information-based image registration, Computer Vision and Image Understanding, Vol. 77, pp. 211–232, 2000.
- [109] Pluim, J., Maintz, J., and Viergever, M., Mutual information matching in multiresolution context, Image and Vision Computing, Vol. 19, No. 1–2, pp. 45–52, 2001.
- [110] Powell, M., An efficient method for finding the minimum of a function of several variables without calculating derivatives, The Computer Journal, pp. 155–162, 1964.
- [111] Press, W., Teukolsky, S., Vetterling, V., and Flannery, B., Numerical Recipies in C, 2nd edn, Cambridge University Press, 1992.
- [112] Reissman, P. J., and Magnin, I. E., The active pyramid : a new tool for modelling 3D deformable object, International Journal Pattern Recognition on Artificial Intelligence, Vol. 11, No. 7, pp. 1129–1139, 1997.
- [113] Roche, A., Malandain, G., and Ayache, N., Unifying maximum likelihood approaches in medical image registration, International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging, Vol. 11, No. 1, pp. 71–80, 2000.
- [114] Rohr, K., On 3D differential operators for detecting point landmarks, Image and Vision Computing, Vol. 15, No. 3, pp. 219–233, 1997.

- [115] Rohr, K., Extraction of 3D anatomical point landmarks based on invariance principles, Pattern Recognition, Vol. 32, pp. 3–15, 1999.
- [116] Rohr, K., Stiehl, H., Sprengel, R., Beil, W., Buzug, T., Weese, J., and Kuhn, M., Point-based elastic registration of medical image using approximating thin-plate splines, In: Proc. Visualization in Biomedical Computing, Hohne, H. and Kikinis, R., eds., No. 1131 in Lecture Notes in Computer Science, pp. 297–306, Springer, September 1996, Hamburg.
- [117] Rueckert, D., Somoda, I., Hayes, C., Hill, D., Leach, M., and Hawkes, D., Non-girid registration using free-form deformations: application to breast MR images, IEEE Transactions on Medical Imaging, Vol. 18, No. 8, pp. 712–721, 1999.
- [118] Sandor, S. and Leahy, R., Surface-based labeling of cortical anatomy using a deformable atlas, IEEE Transactions on Medical Imaging, Vol. 16, No. 1, pp. 41–54, 1997.
- [119] Schnabel, J. A., Tanner, C., Castellano-Smith, A. D., Degenhard, A., Leach, M. O., Hose, D. R., Hill, D. L. G, and Hawkes, D. J., Validation of non-rigid image registration using finite element methods: Application to breast MR images, IEEE Transactions on Medical Imaging, Vol. 22, No. 2, pp. 238–247, 2003.
- [120] Schnörr, C., A study of convex variational diffusion approach for image segmentation and feature extraction, Journal of Mathematical Imaging and Vision, Vol. 8, No. 3, pp. 271–292, 1998.
- [121] Schormann, T., Henn, S., and Zilles, K., A new approach to fast elastic alignement with application to human brains, In: Proc. Visualization in Biomedical Computing, Hohne, K. H., and Kikinis, R., eds., No. 1131 in Lecture Notes in Computer Science, pp. 337–342, Springer, September 1996.
- [122] Schroeter, P., Vesin, J., Langenberger, T., and Meuli, R., Robust parameter estimation on intensity distribution for brain magnetic resonance images, IEEE Transactions on Medical Imaging, Vol. 17, No. 2, pp. 172– 186, 1998.

- [123] Sederberg, T. and Parry, S., Free-form deformations of solid geometric models, Computer Graphics (SIGGRAPH), Vol. 20, No. 4, pp. 151–160, 1986.
- [124] Sethian, J. A., Level Set Methods, Cambridge University Press, 1996.
- [125] Song, S. and Leahy, M., Computation of 3D velocity fields from 3D cine CT images of a human heart, IEEE Transactions on Medical Imaging, Vol. 10, No. 3, pp. 295–306, 1991.
- [126] Subsol, G., Thirion, J. P., and Ayache, N., A general scheme for automatically building 3D morphometric anatomical atlases: application to a skull atlas, Medical Image Analysis, Vol. 2, No. 1, pp. 37–60, 1998.
- [127] Szeliski, R. and Coughlan, J., Hierarchical spline-based image registration, In: Proc. Conference Computer Vision Pattern Recognition, pp. 194–201, Seattle, June 1994.
- [128] Szeliski, R. and Lavallée, S., Matching 3D anatomical surfaces with non-rigid deformations using octree-splines, In: SPIE Geometric Methods in Computer Vision, Vol. 2031, pp. 306–315, 1993.
- [129] Szeliski, R. and Lavallée, S., Matching 3D anatomical surfaces with non-rigid deformations using octree-splines, Inteternational Journal of Computer Vision, Vol. 18, No. 2, pp. 171–186, 1996.
- [130] Talairach, J., Szikla, G., Tournoux, P., Prosalentis, A., and Bornas-Ferrier, M., Atlas d'anatomie stéréotaxique du télencéphale, Masson, Paris, 1967.
- [131] Talairach, J. and Tournoux, P., Co-planar Stereotaxic Atlas of the Human Brain, Georg Thieme Verlag, Stuttgart, 1988.
- [132] Talairach, J. and Tournoux, P., Referentially Oriented Cerebral MRI Anatomy, Georg Thieme Verlag, New-York, 1993.
- [133] Terzopoulos, D., Image analysis using multigrid relaxation methods, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 2, pp. 129–139, 1986.

- [134] Terzopoulos, D. and Metaxas, D., Dynamic 3D models with local and global deformations: deformable superquadrics, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No. 7, pp. 703–714, 1991.
- [135] Thirion, J. P., Extremal points, definition and application to 3D image registration, In: Proc. Conference Computer Vision Pattern Recognition, pp. 587–592, Seattle-Washington, USA, 1994.
- [136] Thirion, J. P., Image matching as a diffusion process: an analogy with Maxwell's demons. Medical Image Analysis, Vol. 2, No. 3, pp. 243–260, 1998.
- [137] Thirion, J. P., and Gourdon, A., Computing the differential characteristics of isointensity surfaces, Computer Vision and Image Understanding, Vol. 61, No. 2, pp. 190–202, 1995, Available at: http://www.inria.fr/RRRT/RR-1881.html.
- [138] Thompson, P. and Toga, A., A surface-based technique for warping three-dimensional images of the brain, IEEE Transactions on Medical Imaging, Vol. 15, No. 4, pp. 402–417, 1996.
- [139] Thompson, P., Woods, R., Mega, M., and Toga, A., Mathematical/computational challenges in cretaing deformable and probabilistic atlases of the human brain, Human Brain Mapping, Vol. 9, pp. 81–92, 2000.
- [140] Toga, A. and Thompson, P., The role of image registration in brain mapping, Image and Vision Computing, Vol. 19, pp. 3–24, 2001.
- [141] Vaillant, M. and Davatzikos, C., Hierarchical matching of cortical features for deformable brain image registration, In: Proc. Information Processing in Medical Imaging, Kuba *et al.*, ed., No. 1613 In Lect. Not. in Comp. Sci., pp. 182–195, Springer, Jun. 1999.
- [142] Bemmel, J. H., Van and M. A. Musen, Handbook of Medical Informatics, Springer. Available at http://www.mieur.nl/mihandbook, 1997.
- [143] Elsen, P. A. Van den, Pol, E. D., and Viergever, M. A., Medical image matching: a review with classification, IEEE Engineering in Medecine and Biology, Vol. 12, pp. 26–38, 1993.

- [144] Vemuri, B., Huang, S., Sahni, S., Leonard, C., Mohr, C., Gilmore, R., and Fitzsimmons, J., An efficient motion estimator with application to medical image registration, Medical Image Analysis, Vol. 2, No. 1, pp. 79–98, 1998.
- [145] Viola, P. and Wells, W., Alignment by maximisation of mutual information, In: Proc. International Conference on Computer Vision, pp. 15–23, 1995.
- [146] Wang, Y. and Staib, L., Physical model-based non-rigid registration incorporating statistical shape information, Medical Image Analysis, Vol. 4, pp. 7–20, 2000.
- [147] West, J., Fitzpatrick, J., Wang, M., Dawant, B., Maurer, C., Kessler, R., and Maciunas, R., Retrospective intermodality registration techniques for images of the head: surface-based versus volume-based, IEEE Trans. Medical Imaging, Vol. 18, No. 2, pp. 144–150, 1999.
- [148] Witkin, A. and Welch, W., Fast animatipon and control of non-rigid structures, Computer Graphics (SIGGRAPH), Vol. 24, No. 4, pp. 243– 252, 1990.
- [149] Zhang, Y., Brady, M., and Smith, S., Segmentation of brain MR images through a hidden markov random field model and the expectationmaximization algorithm, IEEE Transactions on Medical Imaging, Vol. 20, No. 1, pp. 45–57, 2001.
- [150] Zhang, Z., Iterative point matching for registration of free-form curves and surfaces, International Journal of Computer Vision, Vol. 13, No. 2, pp. 119–152, 1994.
- [151] Zitova, B. and Flusser, J., Image registration methods: a survey, Image and Vision Computing, Vol. 21, pp. 977–1000, 2003.

# **Chapter 9**

# Elastic Registration for Biomedical Applications<sup>3</sup>

Jan Kybic<sup>1</sup> and Michael Unser<sup>2</sup>

# 9.1 Introduction

The task of *image registration* is to find homologous (corresponding) points in two images that we shall call *reference* and *test*. These images depict the same or similar objects but are not identical. See an example in Fig. 9.1. We see immediately that the task is not trivial, since some zones in one image do not have any corresponding region in the other image. This illustrates some vagueness of the registration problem which we will have to address. A multimodality biomedical registration task, Fig. 9.2, presents another difficulty: Although the images represent exactly the same object, the same slice of the same brain, the visual appearance of the tissues is radically different in both.

For us, the output of the image registration is *correspondence function*  $\mathbf{g}$  such that  $\mathbf{x}_t = \mathbf{g}(\mathbf{x}_r)$ , mapping a coordinate  $\mathbf{x}_r$  of an arbitrary point in the reference image to a coordinate  $\mathbf{x}_t$  of a corresponding point in the test image.

If we are considering image registration as an inverse problem, then *image* warping (image deformation) is the corresponding forward part (see Fig. 9.3).

<sup>&</sup>lt;sup>1</sup>Center for Applied Cybernetics, Faculty of Electrical Engineering, Czech Technical University, Technická 2, 166 27 Prague 6, Czech Republic

<sup>&</sup>lt;sup>2</sup>Biomedical Imaging Group, Swiss Federal Institute of Technology, Lausanne CH-1015 Lausanne, Switzerland

<sup>&</sup>lt;sup>3</sup>This paper is partially based on "*Fast Parametric Elastic Image Registration*" by J. Kybic and M. Unser which appeared in *IEEE Transactions on Image Processing*, November 2003. (c) 2003 IEEE

Kybic and Unser



Figure 9.1: Two versions of a penguin Tux, the Linux mascot (above). Even though they are different, corresponding points can be identified (below). (Color slide.)

While image registration takes the test and reference images and yields a correspondence (deformation) function, the image warping takes the test image and a correspondence function and outputs a warped image  $f_w(\mathbf{x}) = f(\mathbf{g}(\mathbf{x}))$  which is a deformed version of the test image. The warped test image is aligned



Figure 9.2: Corresponding magnetic resonance brain slices from EPI (left) and anatomical (right) modalities. Landmarks (white crosses) have been manually placed at corresponding locations.



Figure 9.3: Given an image and a deformation function, a deformed image is created by warping. Inversely, given two images, the corresponding deformation function is found by image registration. The registration attempts to make the warped image as similar as possible to the reference image. (Color slide.)

with the reference image and conversely, registering an original image with its warped version recovers the deformation used.

We will call a registration *elastic*, if the family of correspondence functions  $\mathbf{g}$  is sufficiently general, capable of expressing (almost arbitrary) nonlinear relations<sup>4</sup> as opposed to considering for example only linear functions  $\mathbf{g}$ .

# 9.1.1 Applications of Image Registration

Historically, some of the first applications of image registration occurred in the domain of *motion analysis* [2]. The task there is to find changes between two subsequent frames in a video sequence, assuming that these changes can be completely explained by movements of the objects in the scene or of the observer. In most cases, the inter-frame changes are relatively small and the movement smooth. The extracted motion field can be used to measure the trajectories,

<sup>&</sup>lt;sup>4</sup>Note that elasticity is used here in a wider sense than just the mechanical linear elasticity [1].

distances, speeds, and accelerations, etc. In *video compression*, this information enables us to take advantage of the temporal redundancy in the video sequence; it can be used for *object tracking* [3], *image stabilization*, or camera (observer) movement identification leading to *3D reconstruction* of the scene. *segmentation* algorithms divide the image into regions according to their speed.

Registering a pair of *stereo images* also permits the 3D reconstruction of the scene. The spatial configuration of the cameras is known here; on the other hand, we have to deal with occlusions. The general case of *n* cameras presents additional challenges in maintaining the consistency.

#### 9.1.1.1 Biomedical Applications

In the biomedical domain, there is a frequent need to compare images for analysis and diagnostic purposes. For an efficient comparison, the images need to be aligned. This can be accomplished by registration and subsequent warping.

*Intrasubject* analysis compares images of the same subject taken at different times in order to detect or quantify changes caused by the evolution of the disease or the effect of the therapy.

*Intersubject* analysis considers corresponding images from different subjects. Aligning and combining images from many subjects leads to *atlases*, which are annotated reference images. The individual subject images are then registered with the atlas for identification, segmentation, to detect abnormalities, and to quantitatively characterize the shape and size of their features.

Images of the same subject with different imaging modalities can be combined using *intermodality* registration to get a more complete picture of the subject's anatomy and physiology.

Furthermore, registration helps to compensate for geometrical distortions [4] inherent to some imaging methods, as well as for unwanted motion during the acquisition.

# 9.2 Review of Registration Techniques

Most existing registration algorithms can be cast into the following general framework:

- In the first step, some intermediate data is extracted from the two images being registered. This data lives in a *feature space*.
- The algorithm's representation of the correspondence between the two images is taken from a *search space*. This is the space in which the algorithm looks for a solution. An element from this space is returned at the end.
- To find the solution in the search space, the algorithm needs a way to measure the quality of the correspondence for different points in this space. This measure is provided by a *cost function*.
- Finally, the *search strategy* governs the movements of the algorithm in the search space in its quest for the optimum.

We classify existing registration algorithms according to their choice of the above four attributes, similarly as in [5]. We shall concentrate mostly on their biomedical applications. Figures 9.4 and 9.5 show the simplified classification according to the first two attributes in a tree form.

# 9.2.1 Feature Space

According to the feature space employed, we can identify three classes of registration algorithms: *pixel-based*, *transform-based*, and *feature-based*.



Figure 9.4: Simplified classification of registration algorithms according to the feature spaces used.



Figure 9.5: Simplified classification of registration algorithms according to the warp space used.

#### 9.2.1.1 Pixel-Based Registration

Pixel-based algorithms work directly with the (totality of) pixel values of the images being registered. *Preprocessing* is often used to suppress the adverse effects of noise and differences in acquisition [6], or to increase or uniformize pixel resolution [7]. In the continuous framework, images are often considered as functions of the continuous image coordinates, providing a consistent approach to the discretization issues. The correspondence between the discrete and continuous versions of the image is established using *interpolation*. The crudest method is the nearest-neighbor interpolation, the most often used is the linear (resp. bi- or trilinear) interpolation. Among the high-end methods, *B-spline interpolation* [8–10] provides the best trade-off between accuracy and the computational cost [11, 12].

The image model may also live in a higher-dimensional space than the original data, such as when representing 2D image as a surface in a 3D space [13], or using *level sets* [14].

#### 9.2.1.2 Transform-Based Registration

Transform-based algorithms exploit properties of the *Fourier*, *wavelet*, Hadamard, and other transforms, making use of the fact that certain deformations manifest themselves more clearly in the transform domain. These methods are used mainly in connection with linear deformation fields. Nevertheless, there are examples of methods that estimate locally linear optical flow using *Gabor filters* [15,16], *B-splines* [17] and *wavelets* [18]. The transforms are usually linear and independent of the actual image contents.

#### 9.2.1.3 Feature-Based Registration

Feature-based algorithms first reduce the dimensionality of the problem by extracting a small set of characteristic features from the images. The extraction mostly involves thresholding.

Landmark-based methods [1, 19–21] use a relatively small and sparse set of *landmarks*—important points manually or automatically identified in the images. *Extrinsic markers* are artificial features attached to the object, easily and precisely localizable. Unfortunately, they are often long to install and uncomfortable for the patient. If they are not available, we have to content ourselves with features *intrinsic* to the images. In that case, however, the automatic landmark identification is far less robust. The manual landmark identification is often tedious, time-consuming, imprecise, and unreproducible.

If the images cannot be characterized using points, it might be more appropriate to use curves such as edges [22], or volume boundaries [23]. Likewise, in the case of 3D data, surfaces can be used instead of working with the complete volumes. Popular features are also *templates*, small subimages of important regions [24, 25].

#### 9.2.2 Search Space

An important attribute of a registration algorithm is the *search space*. It is also called a *warp space* because it contains warping or correspondence functions, the candidate solutions of the registration problem. A warping function from the search space is described by a (finite) set of real parameters (from a set of permissible values) by means of a *warping model*. We classify them according to the number of parameters and the spatial extent of the area influenced by a single parameter.

#### 9.2.2.1 Local Models

At one end of the scale, we have *non-parametric*, *local methods*. The deformation function sought is basically unconstrained, or belongs to a very large and unrestrictive functional space, such as the Sobolev space  $W_2^2$  of twice differentiable functions. We seek the values of this deformation at a very fine grid, usually coinciding with pixel locations. These methods are formulated either as *variational*, defining a scalar criterion to minimize, or using partial differential equations (PDE). The continuously defined deformation function minimizes a given criterion, or solves a given PDE. The essence of these methods is thus *entirely in the criterion* (resp., PDE). The PDE come from the *optical flow* approach (gradient methods) [26], viscous fluid model [27–29], elastic deformations with physical analogs [7,30] or without them [31], or from the variational criterion [32]

The deformation function can be also modeled *indirectly*, e.g., as a potential field [33]. This reduces the dimensionality of the problem, at the expense of reducing the generality of the deformation. Displacement might be *quantized* (limited) to integer number of pixels [34].

#### 9.2.2.2 Global Models

At the other end, we have *parametric*, *global methods* that describe the correspondence function using a global model with a relatively small number of parameters [35]. The model mostly consists of expressing the warping function in a global linear [36], polynomial [37] or harmonic basis [38, 39]. For these methods, the deformation model corresponding to a specific warp space is as important as the criterion being minimized.

#### 9.2.2.3 Semi-Local Model

In between the two extremes are semi-local models, using a moderate number of parameters with local influence. A grid of *control points* is usually placed over the image and a basis function associated with each of them. Their spacing corresponds loosely to knot or landmark density. By changing the spacing, we can approach either local or global models or choose the best trade-off.

Semi-local models are instrumental for the B-spline based approach described in section 9.4 and were also used, for example, for motion estimation [40].

#### 9.2.2.4 Image Dependent Models

It is sometimes useful to adapt the warping model to the images considered. Hierarchically structured semi-local models, based on splines, wavelets, or quadtrees [41] can be refined only where needed. In feature based methods, the basis functions of the warping model can be placed where the features are and the deformation field interpolated in regions where no information is available. Typical examples are *radial basis functions* such as thin plate splines [1, 19, 42], to be discussed more in section 9.3.

## 9.2.3 Cost Function

The quality of a registration result is assessed by a cost function. It is mainly composed of a *data term*, measuring the similarity of the images after warping. Sometimes a *regularity term* is added to privilege likely (smooth) deformations.

#### 9.2.3.1 Data Term

For methods based on geometric features, we can use a (mean) distance between corresponding features in the source and target images. Note that *landmark interpolation* (section 9.3) is a limit case with infinite weights given to this distance, constraining it effectively to zero. If the pairing between the source and target features is unknown, the iterative closest point algorithm [43], can be used to determine it.

For *pixel-based* methods, the data term is a *similarity measure* on the two images. The simplest and fastest criterion is the appropriate (e.g.,  $l_1$  or  $l_2$ ) norm of the pixel-wise difference, such as in the *SSD* (sum of squared differences) criterion. However, it assumes an equivalency of intensities in both images. If the intensities only correspond up to a possibly varying linear relationship perhaps, then it is appropriate to use *correlation*, respectively, local normalized correlation [7, 44]. More general and perhaps non-functional relation between the intensities warrants the use of the *mutual information* criterion [45–48]. This is encountered, for example, in intermodality registration [49].

All three criteria lead to statistically optimal estimates of the registration under corresponding image and noise models. Their complexity, computational cost, and fragility increases in the order in which they were presented. For local criteria, such as local normalized correlation, or local mutual information, the neighborhood size must be properly chosen. *Image interpolation* is used to calculate the warped image, see also section 9.2.1.1. In *template-based* methods, the template can be compared with a specific region in the target image using any of the similarity measures suitable for pixel-based methods. For *transform-based* methods, a (semi-)norm in the transformed domain such as  $L_2$  (least-squares measure) is usually used.

#### 9.2.3.2 Regularization

In most applications, it appears to be necessary to add an additional *regularization term* to the criterion, mainly to make the problem well-posed and to stabilize the algorithm. Regularization is also used to express our a priori knowledge. For continuous, local deformation models, regularization defines the warping space in the variational sense. For instance, in the case of landmark interpolation, minimization of the norm of the Laplacian  $\|\Delta \mathbf{g}\|$  is often used in practice [1, 19], leading to a *thin-plate spline* solution [50]. Minimizing other similar measures leads to generalized splines [51] determined either directly or using PDEs. See also section 9.3.

Other regularizers are constructed by applying a non-linear (often quadratic and sometimes image dependent) function on the derivative operator. This is done mainly to preserve discontinuities [52]. Regularizers based on thresholding in wavelet domain are also used. Implicit regularization for iterative methods works by alternatively driving the intermediate solution towards the data, and applying a smoothing operator.

## 9.2.4 Search Strategy

Given a cost function, there are several basic ways of minimizing it.

## 9.2.4.1 Direct Solution

In some cases, notably if the cost function is quadratic or if higher order terms in the Taylor expansion of the criterion can be neglected, then the solution can be found in one step [38, 53]. Transform methods using directional filters are often engineered in this way [54, 55].

#### 9.2.4.2 Exhaustive Search

If the search space is finite, exhaustive search can be used. For example, when small templates are extracted from the reference image, their positions in the target image can be found by trying all possible shifts in some small neighborhoods [34, 56].

## 9.2.4.3 Dynamic Programming

For integer 1D problems, *dynamic programming* is applicable [57] with complexity proportional to the number of decisions to be taken (size of the image) and the number of possible outcomes (shifts). The stereo matching problem with cameras with known relative geometry also falls into this category [58]. These are nice examples of how the a priori knowledge about the image formation process simplifies the registration problem by imposing very strong geometrical constraints on the correspondence field.

## 9.2.4.4 Partial Differential Equations

A partial differential equation can describe an evolution in time of the deformation field converging to the desired solution [32, 52]. (Time-independent PDE's derived directly by Euler-Lagrange formulas from the variational formulation are seldom used.) The PDE parameters can vary with time, to enforce robustness at first and relax the constraints toward the end, to allow for precise registration. Various numerical methods can be used to solve the PDEs, the principal ones being finite difference relaxation method [59], finite elements method (FEM) [60, 61], hierarchical finite element bases [62], multigrid methods [63], and wavelets [64, 65].

## 9.2.4.5 Multidimensional Optimization Methods

Many non-linear registration methods lead to a non-linear multidimensional optimization problem. Various optimization methods are used, depending on the size and structure of the problem. The most popular choices include the Powell method, *gradient descent*, conjugated gradients, and variations of the Newton method, such as the *Marquardt-Levenberg* algorithm [59, 66].

Some minimization algorithms are described using different paradigms, such as 'demons' [67], but usually can be explained in terms of a coordinate descent.

#### 9.2.4.6 Multiresolution

Multiresolution [7,36,68] on the feature space (usually image size) helps to speed up the process and to increase its robustness. It is based on approaching the
solution by gradual refinements. We first solve a reduced problem using a small amount of data, then use the solution as an initial guess for the problem at a finer level. This is repeated until the finest (original) level is reached.

Multiresolution on the search space works similarly, adding degrees of freedom to the warping model at each step. We start with a simple model leading to a simple and easy to solve problem. Then gradually add a manageable amount of complexity at each step, until the desired model is reached. The model can be augmented qualitatively, such as going from translation-only to general affine transform, or quantitatively, for example, by decreasing the control node spacing in semi-local models.

Related to multiresolution are *multigrid* optimization methods, where occasional backward transitions from finer to coarser levels are used besides the coarse-to-fine refinement used in the multiresolution [63].

## 9.2.5 Other Attributes and Features

The *dimensionality* refers to the number of dimensions of the images being registered. The warping function normally works in the space of the same dimensionality, transforming one coordinate vector into another.

*Interactive algorithms* need human supervision and interaction, as opposed to fully autonomous ones. Interactive methods often perform well, taking advantage of the human expert, but are unsuitable for treating high volumes of data. Good compromise might be to use hybrid methods, requiring manual intervention or approval only in difficult cases.

In some cases there is no inherent reference and test image, both can play the same role. Then we would like the registration process to be *consistent* with respect to this choice [69, 70]. Consistency is one of the ways to enforce *invertibility* and preservation of the topology of the transformation, other possibilities involve constraining the Jacobian [71] and composition of diffeomorphic mappings [32].

## 9.2.6 Complementary Surveys

There is a wide choice of sources of information on registration algorithms. The surveys by Brown [5] and a newer one by Zitová [72] are rather general. Warfield *et al.* [73] concentrate on nonlinear registration for brain warping applications.

Bayesian interpretation of elastic matching are reviewed by Gee [74], also in the context of human neuroanatomy. The last two survey articles we mention deal specifically with medical imaging applications of image registration. An article by Van den Elsen *et al.* [75] contains very comprehensive and detailed classification of available methods. Finally, Lester and Arridge [76] emphasise the hierarchical concepts of the algorithms.

## 9.3 Landmark Registration

Landmark registration [1, 19, 20] is a two-step feature-based registration technique. In the first step, a set of landmark pairs is identified (see Fig. 9.2 for an example), either manually, or automatically [20, 77]. We get two sequences of points,  $\mathbf{x}_1, \ldots, \mathbf{x}_N$ , and  $\mathbf{z}_1, \ldots, \mathbf{z}_N$ , such that an object at coordinates  $\mathbf{x}_i$  in the reference image corresponds to the object at coordinates  $\mathbf{z}_i$  in the test image. In the second step, the correspondence function is interpolated between the landmark points [78, 79].

*Manual landmark registration* has the usual inconveniences of a manual method—poor accuracy and repeatability. On the other hand, it is robust and reliable thanks to the underlying human expert knowledge. For this reason, it is very valuable as a bootstrap method for further automatic refinement. It can also serve as a reference standard when evaluating the performance of other registration methods on real images and under realistic working conditions.

Landmark interpolation merits a study in its own right. Choosing an interpolation method or an interpolation function is difficult because the implications of this choice are not immediately apparent. On the other hand, in the variational formulation we shall present, the user is asked instead to choose a criterion of optimality, which is usually more tangible and often related to the physics (or other specificities) of the problem. The variational formulation of landmark interpolation also allows us to make an interesting link with (fractional) splines [80].

### 9.3.1 Desirable Properties

The landmark interpolation method should fulfill some basic properties:

- We agree that landmarks are points in space, as opposed to just coordinate values. Similarly, the correspondence function **g** is more than a mathematical function: it describes correspondence of real points. It is an object in space, anchored to the landmarks. Consequently, it seems reasonable to require that the interpolated function **g** be *invariant* with respect to the choice of the coordinate system. In other words, the correspondence between points in the two images should remain the same, regardless of how we measure the position of these points.
- The interpolation problem should always have a solution, if possible a unique one.
- Another property worth having is the *reproduction of identity* [81]. In addition, we might want the reproduction property for other simple transformations, such as shifts or scalings; more generally, affine transformations.
- We want the reconstructed correspondence function to be close to the (unknown) true underlying correspondence function. We want the reconstruction error to decrease rapidly with the number of landmarks—the method should have *good approximation properties* [82]. This way we can adapt the landmark density to ensure that the error is below any a priori given tolerance threshold.
- Finally, we want the interpolation procedure to accommodate easily nonexact fits, useful when the landmark positions are only known approximately. In this approximation setting, the reconstructed correspondence function will pass close to the landmarks, making a compromise between the closeness of the fit and the overall smoothness.

## 9.3.2 Thin-Plate Splines

The use of thin-plate spline technique for landmark interpolation is attributed to Bookstein [1]. Here, we present the method from the variational point of view, as a preparation for the extensions presented in section 9.3.3. Instead of imposing an empirical interpolation formula, the essence of the *variational formulation* consists of choosing a variational criterion  $J(\mathbf{g})$  and then finding among all possible functions passing through the landmarks the one that minimizes J [83, 84]. The thin-plate spline method uses the physical model of a thin steel plate [21] with small vertical displacement given by the scalar field g and calculates J as the *strain energy* of the plate:

$$J(g) = \int \left(\frac{\partial^2 g}{\partial x^2}\right)^2 + 2\left(\frac{\partial^2 g}{\partial x \partial y}\right)^2 + \left(\frac{\partial^2 g}{\partial y^2}\right)^2 dx dy = \int \left(\nabla^2 g\right)^2 dx dy \qquad (9.1)$$

where  $\nabla^2$  denotes the Laplacian and the right equality is obtained by integration by parts under some conditions on the solution space. The Laplacian energy (9.1) is a member of a more general family of scale and rotation invariant cost functions which satisfy the requirements of section 9.3.1, see also [85, 86]. It is also the simplest criterion that does not penalize affine transforms.

The criterion for the vector form  $\mathbf{g}$  is taken simply as the sum of the strain energies of the *x* and *y* components,  $J(\mathbf{g}) = J(g_x) + J(g_y)$ . As the constraints  $\mathbf{g}(\mathbf{x}_i) = \mathbf{z}_i$  can be broken into two independent sets for  $g_x$  and  $g_y$ , it follows that minimizing *J* for  $\mathbf{g}$  is equivalent to minimizing separately for  $g_x$  and  $g_y$ . Consequently, we can concentrate on the scalar case here.

#### 9.3.2.1 Interpolation Formula

The correspondence function g(x, y) minimizing (9.1) under interpolation constraints  $g(x_i, y_i) = z_i$  is given by

$$g(x, y) = \sum_{i=1}^{N} \lambda_i \varrho(\|\mathbf{x} - \mathbf{x}_i\|) + a_0 x + a_1 y + a_2$$
  
with  $\|\mathbf{x} - \mathbf{x}_i\| = \sqrt{(x - x_i)^2 + (y - y_i)^2} = r$  (9.2)

where  $\rho(r)$  is a  $\rho(r) = r^2 \log r$ . It is called radial because it only depends on the Euclidean distance r to its associated data point [87].

The generating function  $\rho(\mathbf{x}) = \rho(x, y)$  solves the associated Euler-Lagrange (or fundamental) equation

$$\nabla_{x,y}^{4} \rho(\sqrt{x^{2} + y^{2}}) = \delta(x, y)$$
(9.3)

where  $\nabla^4$  is a two times iterated Laplacian and  $\delta(x, y)$  is the Dirac distribution. The linear polynomial  $a_0x + a_1y + a_2$  in (9.2) is called a *kernel term* and it appears because it does not contribute to the criterion. The unknown parameters  $\lambda_i$  and  $a_0, a_1, a_2$  are determined from the interpolation constraints  $g(x_i, y_i) = z_i$  and from orthogonality conditions

$$\sum_{i=1}^{N} \lambda_i = 0 \qquad \sum_{i=1}^{N} \lambda_i x_i = 0 \qquad \sum_{i=1}^{N} \lambda_i y_i = 0$$
(9.4)

The method we have just briefly described is called *thin-plate spline* interpolation [1, 85, 86, 88].

#### 9.3.3 Fractional Landmark Interpolation

Although the thin-plate splines have been known to work well, in many applications we might benefit from a wider choice of interpolation functions, while keeping the general spirit and the invariance properties (affine geometrical transformations including scaling) we are interested in. The straightforward way to do it is to consider minimizing different criteria, namely fractional derivatives (in 1D) and fractional Laplacian (in multiple dimensions). In some sense, these are the *only* reasonable criteria guaranteeing the useful properties described above (see [85, 86] for a more precise statement).

#### 9.3.3.1 The Criterion and The Interpolation Formula

The Laplacian is defined in the space domain by  $\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$ . In the Fourier domain we have  $\widehat{\nabla^2 f} = \omega_x^2 \widehat{f} + \omega_y^2 \widehat{f} = \|\omega\|^2 \widehat{f}$ , provided that all quantities exist. This can be extended to *fractional orders* as  $\widehat{\nabla^{\alpha} f} = \|\omega\|^{\alpha} \widehat{f}$ , yielding a generalized version of the Laplacian based criterion (9.1):

$$J(g) = \int \|\nabla^{\alpha} g(\mathbf{x})\|^2 d\mathbf{x} \propto \int \|\omega\|^{2\alpha} |\hat{g}(\omega)|^2 d\omega$$
(9.5)

To get some intuition, note that in the univariate case we would be measuring the norm of the  $\alpha$ -th *fractional derivative* [89,90] of *g*.

There is an interesting relationship between *fractional Brownian motion* [91] and fractional derivatives, since fractional derivatives whiten the fractional Brownian motion and thus effectively yield an uncorrelated Gaussian white noise. The criterion (9.5) can be therefore interpreted as Bayesian *fractal prior* (see [88] for details and also Poggio [92] for the non-fractal case), assuming that the underlying true function is close to the fractional Brownian motion model. We then find the solution to our interpolation problem combining this knowledge with the information given by the constraints. The Euler-Lagrange equation corresponding to (9.5) is  $\nabla^{2\alpha} \rho = \delta$ . The solution for non-special  $\alpha$  (read: non-integer) is of the form

$$g = \sum_{i=1}^{N} \lambda_i \varrho(\underbrace{\|\mathbf{x} - \mathbf{x}_i\|}_{r}) \quad \text{with} \quad \varrho(r) = r^{2\alpha - 2}$$
(9.6)

The polynomial kernel term does not appear here due to technical restrictions of the Fourier domain definition of the criterion (9.5).

#### **9.3.3.2** The Influence of $\alpha$

The coefficient  $\alpha$  translates into the assumed smoothness of the deformation the higher the  $\alpha$ , the smoother the deformation. For practical purposes, we will use  $0.5 < \alpha \lesssim 5$ ; the interpolation becomes point-wise unstable for smaller  $\alpha$  and does not change much for the larger ones. For  $\alpha > 1.5$  the prior can no longer be interpreted as fractal Brownian motion, although the criterion remains usable.

The choice of the order  $\alpha$  obviously has an influence on the interpolation results. In Fig. 9.6 we present an example of this effect for the landmark (bivariate) case, (see [86] for additional examples.) We have chosen two images from a four-chamber ultrasound sequence of a heart<sup>5</sup> and declared one of them reference and the other a test. We have manually identified six pairs of conforming points in both the images and we have also put additional stable landmarks in the corners of the image. Then we warped the test image onto the reference image varying the parameter  $\alpha$ . Note that only non-integer values of  $\alpha$  were used, for which the formula (9.6) remains valid; integer values need a special treatment.

How do we choose the best  $\alpha$ ? As we have seen in the previous section, the statistically optimal  $\alpha$  can be determined directly, when the characteristics of the stochastic process generating the deformation are known. However, in practice, this is never the case . Therefore,  $\alpha$  must be found experimentally. We observe that small  $\alpha$  yields more localized and abrupt changes in the deformation field, while higher  $\alpha$  gives rise to smoother and more global changes, as predicted.

<sup>&</sup>lt;sup>5</sup>*Acknowledgements:* Images and landmark placement are the courtesy of María J. Ledesma, Universidad Politécnica de Madrid, and Laboratory of Echocardiography, Hospital General Universitario Gregorio Marañón, Madrid, Spain.



Figure 9.6: The reference (top left) and test (top right) images. The test image warped by landmark warping for  $\alpha = 0.5$  (middle left),  $\alpha = 0.9$  (middle right),  $\alpha = 1.3$  (bottom left), and  $\alpha = 2.5$  (bottom right). The landmark positions are marked with white squares and were identical in all cases.

When a sufficient number of test images and landmarks are available, a suitable  $\alpha$  for a given application can be determined by the *leave-one-out* technique: One or several landmarks are not taken into consideration when calculating the correspondence function. Their real position is then compared to their position predicted by the interpolation. Finally, the  $\alpha$  yielding the smallest average error is selected. In the present case, we found the values of  $\alpha = 1.5 \sim 2.5$  to be the most suitable.

## 9.3.4 Summary of Landmark Interpolation

We have presented the landmark registration technique with focus on the second step, the problem of landmark interpolation. This problem can be formulated very concisely in the variational setting. We choose the variational criterion to impose useful properties on the interpolation process, such as rotational, translational, and scale invariance. Most notably, when the criterion is quadratic, the solution is expressed as a linear combination of translated generating (Green) functions. The coefficients of this linear combination are determined from a linear system of equations.

The a priori non-local generating functions can be localized [88] for more efficient and more stable calculation. In some cases this localization leads to B-splines which gives an additional justification for using splines to solve this kind of problems.

# 9.4 Fast Parametric Elastic Image Registration

This section presents a practical example of a fully automatic algorithm for fast elastic multidimensional intensity-based image registration with a parametric *B-spline model* of the deformation. Its main features are high-order *B-spline models* of the deformation and of the image, pixel-based similarity criterion, double *multiresolution* strategy (for both image and the model) and sophisticated iterative multidimensional optimizer. While the algorithm presented here is based on our own work [88, 93–96], it is closely related to a number of similar, independently developed approaches, of which we can only present a very incomplete list. The use of B-spline deformation models was pioneered by Szeliski [40,41] and the different pixel criteria were studied by Studholme [48,97] and Nikou [49]. The hierarchical structure was exploited by Musse [71], Heitz [68] and Thévenaz [36], who also employed the Marquardt-Levenberg optimizer. The algorithm can be used for 2D and 3D problems, is reasonably fast, and is capable of accepting expert hints in the form of soft *landmark constraints* [1, 19–21].

#### 9.4.1 Problem Formulation

The input images are given as two *N*-dimensional discrete signals  $f_r(\mathbf{i})$  and  $f_t(\mathbf{i})$ , where  $\mathbf{i} \in I \subset \mathbb{Z}^N$ , and *I* is an *N*-dimensional discrete interval representing the set of all pixel coordinates in the image. We call  $f_r$  and  $f_t$  reference and test images, respectively. We suppose that the test image is a geometrically deformed version of the reference image, and vice versa. This is to say that the points with the same coordinate  $\mathbf{x}$  in the reference image  $f_r(\mathbf{x})$  and in the warped test image  $f_w(\mathbf{x}) = f_t^c(\mathbf{g}(\mathbf{x}))$  should correspond. Here,  $f_t^c$  is a continuous version of the test image and  $\mathbf{g}(\mathbf{x})$  is a deformation (correspondence) function to be identified.

## 9.4.2 Cost Function

The two images  $f_r$ ,  $f_w$  will not be identical because of noise and also because the assumption that there is a geometrical mapping between the two images is not necessarily correct. Therefore, we define the solution to our registration problem as the result of the minimization  $\mathbf{g} = \arg \min_{\mathbf{g} \in G} E(\mathbf{g})$ , where *G* is the space of all admissible deformation functions  $\mathbf{g}$ . We have chosen the SSD (sum of squared differences) criterion

$$E = \frac{1}{\|I\|} \sum_{\mathbf{i} \in I} e_{\mathbf{i}}^{2} = \frac{1}{\|I\|} \sum_{\mathbf{i} \in I} (f_{w}(\mathbf{i}) - f_{r}(\mathbf{i}))^{2}$$
$$= \frac{1}{\|I\|} \sum_{\mathbf{i} \in I} (f_{t}^{c}(\mathbf{g}(\mathbf{i})) - f_{r}(\mathbf{i}))^{2}$$
(9.7)

because it is fast to evaluate and yields a smooth criterion surface which lends itself well to optimization. Minimization of (9.7) yields the optimal solution  $\mathbf{g}$  in the ML (maximum likelihood) sense under the assumption that  $f_r$  is a deformed (warped) version of  $f_t$  with i.i.d. (independent and identically distributed) Gaussian noise added to each pixel. The SSD criterion proved to be robust enough, especially if preprocessing was used to equalize the image values—we mostly applied high-pass filtering and histogram normalization [98]. In principle, there is no difficulty in extending this method for more sophisticated pixel-based similarity measures, such as information-based measures [99], especially mutual information [45], or weighted  $\ell_p$  norms. Only the evaluation of the criterion and its derivatives (gradient) needs to be changed.

## 9.4.3 **B-Splines and Image Interpolation**

We have chosen to interpolate the image using uniform B-splines:

$$f_t^c(\mathbf{x}) = \sum_{\mathbf{i} \in I_b \subset \mathbb{Z}^N} b_{\mathbf{i}} \beta_n(\mathbf{x} - \mathbf{i})$$
(9.8)

where  $\beta_n(\mathbf{x})$  is a tensor product of B-splines  $\beta_n(x)$  of degree *n*, i.e.,  $\beta_n(\mathbf{x}) = \prod_{k=1}^{N} \beta_n(x_k)$ , with  $\mathbf{x} = (x_1, \dots, x_N)$ . Mirror boundary conditions were used, to ensure continuity.

Let us recall some basic facts about B-splines. Uniform symmetric B-splines [100] of degree *n* are piecewise polynomials of degree *n*. The polynomial pieces are delimited by uniformly placed knots. B-splines of degree *n* have continuous derivatives up to order n - 1 everywhere. Their integer shifts form a basis. The first (degree zero) symmetric B-spline is defined as  $\beta_0(x) = 1$  for  $x \in (-\frac{1}{2}, \frac{1}{2})$  and 0 otherwise. Higher order B-splines are defined recursively as  $\beta_{n+1} = \beta_n * \beta_0$  and their support is  $(-\frac{n+1}{2}, +\frac{n+1}{2})$ .

Using B-splines as interpolation functions has many advantages: B-splines have *good approximation properties*—for example, the error of a cubic B-spline  $(\beta_3)$  approximation decreases asymptotically as  $h^4$  (measured by any  $L_p$  or  $l_p$  norm,  $p \in \{1, 2, ..., \infty\}$ ). B-splines perform well in comparison with other bases [11, 101]. B-splines are *fast*—they have a short support (length 4 for  $\beta_3$ ), are symmetric, piecewise cubic, and separable in multiple dimensions. They are simple to compute and *scalable*—the transition from a coarse spline space with step size (knot distance) hq to a finer space with step size h is exact for integer q.

#### 9.4.4 Deformation Model Structure

So far, we have considered the deformation function  $\mathbf{g}$  to be an arbitrary admissible function  $\mathbb{R}^N \to \mathbb{R}^N$ . We will restrict it now to a family of functions described

by a finite number of parameters  $c_j$ :

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{j} \in J} \mathbf{c}_{\mathbf{j}} \varphi_{\mathbf{j}}(\mathbf{x})$$
(9.9)

where J is a set of parameter indexes and  $\varphi_j$  are the corresponding basis functions. This transforms a variational problem into a much easier finitedimensional minimization problem, for which numerous algorithms exist [59]. Moreover, the restriction of the family G of all possible functions  $\mathbf{g}$  can already guarantee some useful properties, such as the regularity (smoothness) of the solution.

#### 9.4.4.1 B-Spline Deformation Model

There are various possibilities for the choice of the basis functions  $\varphi_j$  for the deformation model (9.9). These include polynomials [37], harmonic functions [38,39], radial basis functions [1,102], and wavelets [64,65,103], all of which have been used in registration algorithms before. However, we have again chosen B-splines, basically for the same reasons that lead us to choose them to interpolate our images (see Section 9.4.3): their good approximation properties, computational efficiency [96], scalability, and additionally physical plausibility [95] (such as minimizing the "strain energy"  $||g''||_2$  by cubic B-splines [104, 105]) and low interdependency thanks to short support. Their property of being able to represent affine transformations, including rigid body motion, is welcome, too. We have also evaluated the alternative wavelet representation of the same Bspline space [106], only to find that direct B-spline representation was again more efficient [96].

The B-spline deformation model is obtained by substituting a scaled version of the B-spline (or tensor product thereof) in (9.9)

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} + \sum_{\mathbf{j} \in I_c \subset \mathbb{Z}^N} \mathbf{c}_{\mathbf{j}} \beta_{n_m} \left( \mathbf{x} / \mathbf{h} - \mathbf{j} \right)$$
(9.10)

where  $n_m$  is the degree of splines used, **h** is the knot spacing, and the division is taken elementwise. This corresponds to placing the knots on a regular grid over the image. For efficiency reasons, we require the node spacing **h** to be integer, which together with the separability of  $\beta_{n_m}(\mathbf{x})$  implies that the values of the B-spline  $\beta_{n_m}(x)$  are only needed at a very small number of points  $(n_m + 1)\mathbf{h}$  and that they can be precalculated.

## 9.4.5 Existence and Unicity

Since the useful range of the parameters  $\mathbf{c}$  is naturally compactly constrained (for example, by excluding displacements larger than the image size) and since the criterion E is non-negative and continuous with respect to  $\mathbf{c}$  (if at least linear interpolation for  $f_t^c$  is used), it is clear that E as a function of  $\mathbf{c}$  has a minimum; i.e., the proposed problem has a solution. However, depending on the images at hand, the solution does not have to be unique and there can be local minima. Fortunately, this does not pose problems in practice, thanks to a multiresolution approach (section 9.4.6.2) which smoothes out images at coarse levels and brings us sufficiently close to the solution at fine resolution levels. The algorithm will find a solution if started within the attraction basin of that solution. The virtual springs (section 9.4.7) play a role of an a priori information and a regularization term; extra regularization can be applied, if desired [88, 107] (section 9.4.9).

## 9.4.6 Optimization Strategy

#### 9.4.6.1 Optimization Algorithm

Recall from (9.7) and (9.10) that we need to minimize a criterion *E* with respect to a finite number of parameters **c**. To determine which of the many available algorithms performs best in our context, we tested four local iterative algorithms which can be cast into a following common framework: At each step *i* we take the actual estimate  $\mathbf{c}^{(i)}$  and calculate a proposed update  $\Delta \mathbf{c}^{(i)}$ . If the step is successful, i.e., the criterion decreases, then the proposed point is accepted,  $\mathbf{c}^{(i+1)} = \mathbf{c}^{(i)} + \Delta \mathbf{c}^{(i)}$ . Otherwise, a more conservative update  $\Delta \mathbf{c}^{(i)}$  is calculated, and the test is repeated.

- 1. Gradient descent with feedback step size adjustment. The update rule is:  $\Delta \mathbf{c}^{(i)} = -\mu \nabla_{\mathbf{c}} E(\mathbf{c}^{(i)})$ . After a successful step,  $\mu$  is multiplied by  $\mu_f$ , otherwise it is divided by  $\mu'_f$ .<sup>6</sup>
- 2. Gradient descent with quadratic step size estimation. We choose a step size  $\mu^*$  minimizing the following approximation of the criterion around  $\mathbf{c}^{(i)}$ :  $E(\mathbf{c}^{(i)} + \mathbf{x}) = E(\mathbf{c}^{(i)}) + \mathbf{x}^T \nabla_{\mathbf{c}} E(\mathbf{c}^{(i)}) + \alpha \|\mathbf{x}\|^2$ , where  $\alpha$  is identified from the

<sup>&</sup>lt;sup>6</sup>We used  $\mu_f = 10$  and  $\mu'_f = 15$ .

two last calculated criterion values *E*. As a fallback strategy, the previous step size is divided by  $\mu'_f$ , as above.

- 3. Conjugated gradient. This algorithm [59] chooses its descent directions to be mutually conjugate so that moving along one does not spoil the result of previous optimizations. To work well, the step size  $\mu$  has to be chosen optimally. Therefore, at each step, we need to run another internal one-dimensional minimization routine which finds the optimal  $\mu$ ; this makes it the slowest algorithm in our setting.
- 4. *Marquardt-Levenberg*. The most effective algorithm in the sense of the number of iterations was a regularized Newton method inspired by the *Marquardt-Levenberg* algorithm (ML), as in [98]. We shall examine various approximations of the Hessian matrix  $\nabla_{\mathbf{c}}^2 E$ , see Section 9.4.8.1.

The choice of the best optimizer is always application-dependent. We observe that the behavior of all optimizers is almost identical at the beginning of the optimization process (see Fig. 9.7). The main factor determining the speed



Figure 9.7: The evolution of the SSD criterion during first 18 iterations when registering the Lena image, artificially deformed with  $2 \times 4 \times 4$  cubic B-spline coefficients and a maximum displacement of about 30 pixels, without multiresolution. The optimizers used were: Marquardt-Levenberg with full Hessian (MLH), Marquardt-Levenberg with only the diagonal of the Hessian taken into account (MLdH), and gradient descent (GD). The deformation was recovered in all cases with an accuracy between 0.1 and 0.01 pixels (see also section 9.4.10).

is therefore the cost of a single iteration, which was smallest for the gradient descent (GD) algorithm. Among the GD variants we recommend the quadratic step size estimation that outperforms the classical feedback adjustment. One additional pleasant property of the GD algorithm is its tendency to leave uninfluential coefficients intact, unlike the ML algorithm. Consequently, less regularization is needed for the GD algorithm. We choose the GD optimizer for most of our image registration tasks.

When, on the other hand, we work with a small number of parameters, the criterion is smooth, and high precision is needed, the ML algorithm [36] performs the best, as its higher cost per iteration is compensated for by a smaller number of iterations due to the quadratic convergence (Fig. 9.8).

#### 9.4.6.2 Multiresolution

The robustness and efficiency of the algorithm can be significantly improved by a *multiresolution* approach: The task at hand is first solved at a coarse scale.



Figure 9.8: Comparison of gradient descent (GD), conjugated gradient (CG), and Marquardt-Levenberg (ML) optimization algorithm performances when registering SPECT images with control grid of  $6 \times 6 \times 6$  knots. The graphs give the value of the finest-level SSD criterion of all successful (i.e., criterion-decreasing) iterations as a function of the execution time. The abrupt changes are caused by transitions between resolution levels.

Then, the results are propagated to the next finer level and used as a starting guess for solving the task at that level. This procedure is iterated until the finest level is reached.

In our algorithm, multiresolution is used twice. First, we build an *image pyramid*: a set of gradually reduced versions of the original image [108, 109]. This pyramid is compatible with our image representation (9.8) and is optimal in the  $L_2$ -sense (i.e., compatible with the SSD criterion (9.7)), which ensures that the approximation made by substituting the lower resolution image is the best possible. We reduce images up to the size of  $16 \sim 32$  pixels, which works well in most cases. The coarse versions of images (half size) are generated using a reduction operator (see Section 9.4.8.4) and coarse level solutions are extrapolated to finer levels using an expansion operator (cubic spline interpolation).

Second, we use multiresolution for the warping function as well. We start with a deformation **g** described with very few parameters  $\mathbf{c}_k$ , and with a large distance **h** between knots. After the optimization of  $\mathbf{c}_k$  is complete, we halve the distance between knots. This approximately corresponds to doubling the number of knots in each direction, i.e., quadrupling (in 2D) the number of coefficients  $\mathbf{c}_k$ . Because of the two-scale spline relation, we can *exactly* represent the warping function from the old, coarse space, in the new, finer space. The sequence obeys  $\mathbf{h}^{j+1} = \mathbf{h}^j/2$ . The process starts with *g* being identity.

The global strategy combines the two multiresolutions by alternatively decreasing the scales for the image and for the model.

The consequence of using multiresolution is that the algorithm works best for images and deformations that follow the multiresolution model; i.e., when a low resolution version is a good approximation of the finer resolution version.

#### 9.4.7 Semi-Automatic Registration

We realize that although the multiresolution approach leads to a very robust registration algorithm, there are cases when it is misled by an apparent similarity of features which do not correspond physically. Therefore, we developed an extension of the algorithm which can use expert hints. The hints come in the form of a set of landmarks and are used to gear the algorithm towards the correct solution. Similar idea appeared also for non-parametric approaches [110, 111].

The landmark information is incorporated in the automatic process using the concept of *virtual springs*, tying each pair of corresponding points together. We

augment the data part of the criterion E with a term  $E_s$ , corresponding to the potential energy of the springs, and minimize the sum of the two:  $E_c = E + E_s$ . The spring term is:

$$E_s = \sum_{i=1}^{S} \alpha_i \| \mathbf{g}(\mathbf{x}_i) - \mathbf{z}_i \|^2$$
(9.11)

where *S* is the number of springs,  $\alpha_i$  are weighting factors corresponding to their stiffnesses, and  $\mathbf{x}_i$ , resp.  $\mathbf{z}_i$ , are the landmark positions in the reference, resp., test images. The spring factors  $\alpha_i$  control the influence of the particular landmark pairs.

As an example, we present the registration of an MRI slice from an atlas<sup>7</sup> with a sample MRI test image<sup>8</sup>. To identify the same structures in the test image, we register it with the unlabeled version of the atlas. Once the geometric correspondence is established, the structures and their labels from the atlas can be projected onto the test image. The unsupervised registration correctly registers some of the structures but misses others; in particular, the skull boundary (see Fig. 9.9). If we now help the algorithm by identifying several landmarks in both images (Fig. 9.10), the semi-automatic version can recover a plausible deformation, even though the landmark information alone (using e.g., thin-plate splines) would not have been enough [112].

Adding the spring term privileges likely solutions based on our a priori knowledge and makes the problem better-posed. The points need not to be imagedependent landmarks. For example, anchoring the four corners of the image prevents the solution from degenerating. In this way, the springs play in part the role of a regularization factor.

### 9.4.8 Implementation Issues

The purpose of this section is to describe some specific aspects of our implementation. These are mostly independent of the main philosophy of the algorithm but can have a major impact on its performance.

<sup>&</sup>lt;sup>7</sup>The atlas is a labeled and annotated collection of images. Courtesy of Harvard Medical School, http://www.med.harvard.edu/AANLIB/home.html

<sup>&</sup>lt;sup>8</sup>We use a proton density MR image from the Visible Human project http://www.meddean.luc.edu/lumen/meded/grossanatomy/cross\_section/index.html. Prior to registration, the histogram of the test image was matched to that of the reference.



Figure 9.9: The reference MRI proton density brain slice from the atlas with (a) and without labels (b). The sample test slice of a corresponding region (c). The superposition (in red and green) of the two images before (d) and after the registration (e). The deformation field (f). Cubic splines were used with knot spacing of h = 32. The image size was  $512 \times 512$  pixels. The difference between images is only partially corrected by the unsupervised registration. Misalignment of several structures is clearly visible. (Color slide.)



Figure 9.10: The reference (a) and test (b) images with superimposed landmarks (in red). The superimposed images after registration using the semiautomatic algorithm (c) and the deformation field found (d). Corresponding anatomical structures are well identified; the alignment is clearly superior to that in Figure 9.9. (Color slide.)

#### 9.4.8.1 Explicit Derivatives

For the optimization algorithm, we need to calculate the partial derivatives of E, as they form the gradient vector  $\nabla_{\mathbf{c}} E(\mathbf{c}^{(i)})$  and the Hessian matrix  $\nabla_{\mathbf{c}}^2 E(\mathbf{c}^{(i)})$ . Starting from Eq. (9.7), we obtain the first partial derivatives

$$\frac{\partial E}{\partial c_{\mathbf{j},m}} = \frac{1}{\|I\|} \sum_{\mathbf{i} \in I_b} \frac{\partial e_{\mathbf{i}}}{\partial f_w(\mathbf{i})} \left. \frac{\partial f_t^c(\mathbf{x})}{\partial x_m} \right|_{\mathbf{x} = \mathbf{g}(\mathbf{i})} \frac{\partial g_m(\mathbf{i})}{\partial c_{\mathbf{j},m}}$$
(9.12)

as well as the second partial derivatives

$$\frac{\partial^2 E}{\partial c_{\mathbf{j},m} \partial c_{\mathbf{k},n}} = \frac{1}{\|I\|} \sum_{\mathbf{i} \in I_b} \left( \frac{\partial^2 e_{\mathbf{i}}}{\partial f_w(\mathbf{i})^2} \frac{\partial f_t^c}{\partial x_m} \frac{\partial f_t^c}{\partial x_n} + \frac{\partial e_{\mathbf{i}}}{\partial f_w(\mathbf{i})} \frac{\partial^2 f_t^c}{\partial x_m \partial x_n} \right) \frac{\partial g_m}{\partial c_{\mathbf{j},m}} \frac{\partial g_n}{\partial c_{\mathbf{k},n}} \quad (9.13)$$

From (9.7) defining the SSD criterion, we get  $\frac{\partial e_{\mathbf{i}}}{\partial f_w(\mathbf{i})} = 2(f_w(\mathbf{i}) - f_r(\mathbf{i}))$  and  $\frac{\partial^2 e_{\mathbf{i}}}{\partial f_w(\mathbf{i})^2} = 2$ . The derivative of the deformation function (9.10) is simply  $\frac{\partial g_m}{\partial c_{\mathbf{j},m}} = \beta_{n_m} (\mathbf{x}/\mathbf{h} - \mathbf{j})$ . The deformation model is linear and all its second derivatives are therefore zero; that is the reason for the simplicity of (9.13). The partial

derivatives of  $f_t^c$  in (9.12) and (9.13) can be calculated from (9.8) as a tensor product  $\frac{\partial f_t^c}{\partial x_m}(\mathbf{x}) = \sum_{\mathbf{k} \in I} b_{\mathbf{k}} \beta'_n(x_m - k_m) \prod_{\substack{l=1 \\ l \neq m}}^N \beta_n(x_l - k_l)$ . Second-order partial derivatives of  $f_t^c$  are obtained in a similar fashion.

#### 9.4.8.2 Hessian Approximation

Because the evaluation of the Hessian matrix from (9.13) is costly, several modifications have been devised. The Marquardt-Levenberg approximation assumes that the term  $\frac{\partial e_i}{\partial f_w(\mathbf{i})}$  is negligibly small or that it sums to zero on the average. This reduces (9.13) to

$$\frac{\partial^2 E}{\partial c_{\mathbf{j},m} \partial c_{\mathbf{k},n}} = \frac{2}{\|I\|} \sum_{\mathbf{j} \in I_h} \frac{\partial f_t^c}{\partial x_m} \frac{\partial f_t^c}{\partial x_n} \frac{\partial g_m}{\partial c_{\mathbf{j},m}} \frac{\partial g_n}{\partial c_{\mathbf{j},n}}$$
(9.14)

Another simplification is to consider only diagonal terms  $\partial^2 E / \partial c_{j,m}^2$ . Obviously, this diagonal Hessian approximation only makes sense if the basis functions  $\varphi_j$  do not overlap too much. This is another argument for the B-spline model. Each such approximation makes the evaluation faster at the expense of precision which may result in slower convergence. Whether it is advantageous to use some approximation depends on many factors, including the size and the character of the data.

#### 9.4.8.3 Gradient Calculation

Similarly to the case of evaluating the deformation  $\mathbf{g}$ , the use of an integer step size  $\mathbf{h}$  leads to computational savings here too. The expanded expression for  $\frac{\partial E}{\partial c_{j,m}}$  can be transformed into a discrete separable *convolution*  $\left\{\frac{\partial E}{\partial c_{j,m}}\right\}_{\mathbf{j}} = \sum_{\mathbf{i}} w(\mathbf{i})b(\mathbf{j} \cdot \mathbf{h} - \mathbf{i}) = (w * b)_{\downarrow \mathbf{h}}$ , where we have substituted w for the first two factors in (9.12),  $b(\mathbf{q}) = \beta_{n_m}(-\mathbf{q}/\mathbf{h})$ , and  $\downarrow \mathbf{h}$  indicates downsampling as defined by the formula, with elementwise multiplication  $\mathbf{j} \cdot \mathbf{h}$ . The convolution kernel b is separable and the convolution can be calculated as a sequence of N unidimensional convolutions  $((w * b_1)_{\downarrow h_1} * \cdots b_n)_{\downarrow h_N}$ . Because of the downsampling, calculating one output value at step k consists of a scalar product with a filter  $b_k$  of length  $(n_m + 1)h_k$  and shifting this filter by  $h_k$ .

#### 9.4.8.4 Multiresolution Spline Representation

To deploy the multiresolution strategy (see section 9.4.6.2), we need to specify expansion and reduction operators. We will use the same approach for both the deformation model and the image model.

Let us consider here a 1D signal represented in a B-spline space

$$f(x) = \sum_{i} c_i \beta_n (x - i) \tag{9.15}$$

The *expansion operator*  $\mathcal{E}$  yields a twice expanded version of f which is also a spline

$$f_e = \mathcal{E}f, \quad f_e(x) = f(x/2) = \sum_i d_i \beta_n(x-i)$$
 (9.16)

with coefficients  $d_i$  given by

$$d = c_{\uparrow 2} * u^n \tag{9.17}$$

where  $c_{\uparrow 2}$  denotes upsampled version of c and  $u^n$  is a symmetrical binomial filter defined in the *z*-domain as

$$U^{n}(z) = \frac{(1+z)^{n+1}}{2^{n}} z^{-(n+1)/2}$$
(9.18)

The twice reduced signal f(2x) cannot be represented as a spline with knots at integers. We need to resort to approximation and we have chosen the  $L_2$ optimality as described in [109]. The *reduction operator*  $\mathcal{R}$  will yield a projection (denoted  $P_1$ ) in the original spline space with step size 1.

$$f_r = \mathcal{R}f, \quad f_r(x) = P_1 f(2x), \quad f_r(x) = \sum_i e_i \beta_n(x-i)$$
 (9.19)

The spline coefficients  $e_i$  are calculated as

$$e = (\mathring{h} * c)_{\downarrow 2} * b^{-(2n+1)}$$
(9.20)

with prefilter  $\mathring{h} = b^{2n+1} * u^n$ , where  $b^{2n+1}$  is the sequence of sampled values of a B-spline of degree 2n + 1,  $b^n(i) = \beta_n(i)$ . Finally,  $b^{-(2n+1)}$  is the inverse filter to  $b^{2n+1}$  and the convolution can be handled by *recursive filtering*, as described in [8,9].

Because  $\mathcal{R}$  is a projection complementary to  $\mathcal{E}$ , we have the projection identity  $\mathcal{RE}f = f$ . Extension of both operators to multiple dimensions is trivial thanks to separability.

#### 9.4.8.5 Consequences of Finite Support

All what we said so far about expansion and reduction holds for infinite signals. To adapt the method for finite signals, we considered the following requirements: the expansion must be exact in the continuous sense, the projection identity must hold, reduction followed by expansion must conserve the length of the signal, and as much information as feasible should be conserved. These requirements are useful to guarantee the best possible use of the coarse-grid results at the fine-grid level and are absolutely indispensable for multigrid minimization.

Traditionally, one represents the signal with exactly one coefficient per sample and assumes that the signal outside the region of interest follows some known pattern, such as periodicity, or mirror-on-boundary conditions. We take the mirror-on-boundary condition as an example, but the same kind of problems appear for other boundary conditions, too. First, the signal is forced to be symmetric and thus flat at boundaries. Second, the boundary conditions for both expansion and reduction are only conserved for odd number of samples, otherwise the mirror position needs to change. Third, varying the length of the signal by one does not change the length of the reduced version which makes it impossible to recover the original length by expansion.

The centered pyramids [108] conserve the mirror position for even-length signals. Unfortunately, the expansion is no longer exact. Moreover, the constraint of the size of the image being a power of two, together with the integer step size  $\mathbf{h}$ , seems to be too restrictive.

Because of these considerations, we decided to dissociate the number of B-spline coefficients from the length of the interest region. Initially, we extend the signal by  $\lceil (n-1)/2 \rceil$  samples at each extremity which allows us to represent any spline of degree *n* without constraints. We never move the boundaries of our signal when expanding, although the number of B-splines might vary. In this way, expansion is always exact while it adds extra knots at each end. Reducing expanded signal recovers exactly the original. When reducing other signals, we need to extend them to be able to use our efficient filtering technique. For this, we choose to use the mirror-on-boundary conditions.

#### 9.4.8.6 Image Size Change

The only trick when expanding and reducing the images is to adapt the deformation function accordingly. This is easily accomplished by multiplying the coefficients by 2 when expanding and 0.5 when reducing. Thanks to our choice of the expansion and reduction operators, the origin of the image does not change.

#### 9.4.8.7 Fast Spline Calculations

It is essential to take full advantage of the properties of splines. First, specialized routines are used to calculate the values of a B-spline of a specific order using a minimum number of operations. Second, as we are using tensor products of B-splines as our basis functions, many operations can be performed in a separable fashion, reducing the complexity of operations from  $O(k^N)$ , where N is the number of dimensions and k the size of the data, to O(kN). This is the case for the prefiltering step required to find the B-spline coefficients, and also for the interpolation of values of a function given by its B-spline coefficients. Third, the compact support of B-splines simplifies many of the infinite sums in the expressions given earlier, reducing them to sums over just a small number of elements.

#### 9.4.8.8 Stopping Criterion

To get a fast optimization algorithm, particular attention has to be paid to the stopping criterion. This holds for both GD and ML algorithms. Classically, the relative and absolute improvement of the criterion value is compared with a fixed threshold [59]. For our class of problems, it is often advantageous to base the stopping criterion on the changes  $\Delta \mathbf{c}$  of parameter values. We stop when the step size falls below an a priori given threshold  $\varepsilon$ . The size of a step that fails gives an indication of the accuracy of the result and is therefore easy to set. Typically, we would use the threshold of  $\varepsilon = 10^{-1} \sim 10^{-3}$  pixels for the finest level and slightly more for coarser levels, as there is usually not enough details and coherence between levels.

#### 9.4.8.9 Masking

A substantial gain in speed comes from considering only important pixels when calculating the data criterion (9.7) and its derivatives. It is possible to determine an a priori mask of significant pixels, for example,  $10 \sim 50\%$  of the total



Figure 9.11: Example of a mask selecting 10% of the pixels with the highest gradient values for the Lena image.

number of pixels, and to consider only those pixels in subsequent calculations. The contributions of individual pixels to the change of the criterion is directly proportional to the amplitude of the directional derivatives at the respective points, see (9.12). Therefore, a reasonable strategy is to construct the mask by thresholding the gradient of the image at each multiresolution level (see Fig. 9.11).

#### 9.4.9 Invertibility and Regularization

In some applications, it is useful to add an extra regularization term  $E_r$  to the difference measure E, and to look for a minimum of the combined criterion

$$E_c = E + \gamma E_r \tag{9.21}$$

The factor  $E_r$  is used to make the solution well-posed, or to privilege likely solutions based on our a priori knowledge.

First, we consider a *penalty term* designed to enforce the invertibility of the deformation, generalizing the concept from [95] and similar to [113]. Its motivation comes from the fact that if the Jacobian det( $\nabla_{\mathbf{x}} \mathbf{g}$ ) is positive everywhere,

then the deformation  $\mathbf{g}$  is locally invertible. Evaluating this constraint at pixel coordinates and converting the strict constraints into soft ones using a barrier function yields the following penalty term

$$E_p = \sum_{\mathbf{i} \in I} e^{-\alpha \det(\nabla_{\mathbf{x}} \mathbf{g}(\mathbf{i}))}$$
(9.22)

Experience shows that, for typical data, this term is never important at the solution point (to which the optimization converges). It mostly becomes useful at the beginning of the optimization process when the trial points vary a lot, especially with some optimizers. In such cases, the penalty term forces the algorithm to stay in the region of invertible deformations.

Depending on the particular task and the expected properties of the solution, various regularization terms can be used. We investigated, for example, a stabilizer penalizing non-linear deformations

$$E_t = \int \sum_{k,l,m=1}^{N} \left( \frac{\partial^2 g_k}{\partial x_l \partial x_m} \right)^2 d\mathbf{x}$$
(9.23)

and a very simple norm measuring the distance of  ${\bf g}$  from identity through the coefficients  ${\bf c}$ 

$$E_d = \sum_{\mathbf{j}} \|\mathbf{c}_{\mathbf{j}}\|^2 \tag{9.24}$$

When the corresponding weight  $\gamma$  is small, the regularization mainly smoothes the deformation function in places where little information is present in the images. As it gets bigger, the regularization gradually overrides the data term and the deformation tends towards a smooth function in the sense of the particular regularization. An alternative to regularization is the virtual spring mechanism described in section 9.4.7.

## 9.4.10 Experiments

We now illustrate the application of the presented algorithm to various problems involving medical images of several modalities. We refer the interested reader to [88,95,96], where we study in detail the accuracy, speed and robustness of the algorithm by means of a comprehensive series of experiments in a controlled environment.



Figure 9.12: The original slice of anatomical MRI brain image ((top left), original superimposed over the true deformation (top right), the recovered deformation versus the true deformation (bottom left), and the mask used to calculate the warping index (bottom right). (Color Slide)

## 9.4.10.1 Registration of Artificially Deformed MRI Brain Slices

To illustrate the behavior of the algorithm, we show its performance when recovering a known deformation of a 2D slice of an anatomical spin-echo MRI volume of the brain.<sup>9</sup> We use here artificially-deformed images because the knowledge of the ground truth permits us to better judge the performance of the algorithm.

The original image of size  $256 \times 256$  pixels is shown in Fig. 9.12, top-left. We use a cubic spline control grid with one knot for every 32 pixels. We warp the image with a deformation belonging to the warp space and consisting of displacements up to 15 pixels.<sup>10</sup> The warped image is superimposed on the original in Fig. 9.12, top-right. Then the automatic registration algorithm is run with the stopping threshold set to 0.5 pixels for all levels except the last, where we set it to 0.1 pixels. The recovered deformation was used to warp again the original image. Its warped version is shown superimposed on the image warped

<sup>&</sup>lt;sup>9</sup>First author's brain. Images courtesy of Arto Nirkko from Inselspital Hospital, Bern, Switzerland.

<sup>&</sup>lt;sup>10</sup>Approximately 14 mm.



Figure 9.13: The evolution of the optimization process. The left column displays the evolution with respect to the number of iterations, while the right column represents the same quantity respect to time. The first row shows the SSD criterion *E*, the second row the warping index  $\varpi$ . The steep (step) changes correspond to the changes in the model and image resolutions. We observe good correlation between all four graphs.

with the true deformation in Fig. 9.12, bottom-right. We note that the deformation was well recovered with no perceptible difference.

The spatial distribution of the resulting geometrical error is shown in Fig. 9.14. The maximum error is about 1.5 pixels, while the mean geometric error (warping index  $\varpi$  [36]) over the total of the brain is about 0.4 pixels. We generally observe that the error is concentrated in areas with little detail in the image. Other high-contrast regions, such as edges, are resolved much more precisely than indicated by the value of  $\varpi$ , often with subpixel accuracy. On the other hand, the agreement in the zones with low-contrast is worse and often only coincidental, since there is little or no information to guide the algorithm.

The evolution of the optimization can be studied from the graphs in Fig. 9.13. We observe the steady and correlated descent of the observable criterion being



Figure 9.14: The geometrical error after registration (green) with superposed contours of the original MRI image (red). The maximum (green) intensity corresponds to an error of 1.5 pixels. (Color slide.)

optimized (*E*) and of the warping index ( $\varpi$ ), the quantity measuring the quality of the registration. The abrupt changes in the curves are caused by the transitions between levels of the multiresolution progression; they are small thanks to the accuracy of the spline model.

Note that the final values of both E and  $\varpi$  depend strongly on the preset stopping threshold, which in turn influences the optimization time. The threshold value is a subjective compromise between the accuracy and computation time. It is perfectly possible to stop optimizing only after 7s and skip the finest resolution level altogether, if the precision of  $\varpi = 0.7$  pixels is acceptable. On the other hand, after about 4 more minutes of iteration, the error  $\varpi$  descends to less than  $10^{-4}$  pixels. However, in the authors' opinion, such super subpixel accuracy is almost never achievable on real images, because of the noise and the unknown characteristics of the acquisition process.

### 9.4.10.2 Registration of True Medical Data

Finally, we give a representative list of medical imaging registration tasks where the described algorithm was successfully used:

• Registration of ECD<sup>11</sup> and Xenon inhalation SPECT images [114] in the view of atlas creation [115].

<sup>&</sup>lt;sup>11</sup>ECD (Technetium Ethylene Cysteine Diethylester) is a radioactively marked intravenously injected agent.



Figure 9.15: The superposition of the slices of anatomical MRI images before the registration (a), after the registration (b), and the resulting deformation field (c). Quadratic splines were used with knot spacing of h = 64. (Color slide.)

- Intersubject registration of anatomical (spin-echo) MRI images, <sup>12</sup> Fig. 9.15.
- Registration of MRI images from a heart beat sequence,<sup>13</sup> see Fig. 9.16. The extracted deformation field can be used to extract trajectories of various points in the heart and to further determine the velocity and derived parameters, such as the accumulated displacement and strain, which is important for diagnostic purposes.
- Registration of standard 2D ultrasound sequences of the heart [116].
- Motion compensation for a sequence of myocardial perfusion MRI images<sup>14</sup> [117, 118], enabling the time profiles of the intensities at each tissue point to be reliable calculated (Fig. 9.17). Virtual springs were used in this case.
- Registration of two 3D computer tomography (CT) head volumes,<sup>15</sup> see Figs. 9.18 and 9.19. We worked on a reduced volume data ( $128 \times 128 \times 45$ instead of the original  $512 \times 512 \times 45$ ) with control knots placed every  $8 \times 8 \times 8$  voxels and the process took about 10 minutes to complete.<sup>16</sup>

<sup>&</sup>lt;sup>12</sup>Images courtesy of Arto Nirkko, Inselspital Hospital, Bern, Switzerland.

<sup>&</sup>lt;sup>13</sup>LECB, NIH, http://www-lecb.ncifcrf.gov/flicker/

<sup>&</sup>lt;sup>14</sup>Courtesy of J.-P. Vallée, Unité d'imagerie numérique, University Hospital, Geneva, Switzerland.

<sup>&</sup>lt;sup>15</sup>Images courtesy of Philippe Thvenaz, EPFL, Lausanne, Switzerland. The images were acquired using the same machine and the same protocol, but not preregistered.

<sup>&</sup>lt;sup>16</sup>On a 700 MHz Pentium based computer. Registering directly the undecimated volumes on the same computer takes about three hours with very minor increase in quality as relatively smooth deformations are sought. We are currently working on a optimized reimplementation of the algorithm that should reduce these times considerably.



Figure 9.16: The reference MRI image from a heart sequence with superimposed contours (a). The same contours over another image (the test image) from the same sequence before the registration (b) and after (c). The deformation field (d). Quadratic splines were used with knot spacing of h = 64, image size was  $256 \times 256$  pixels. (Color slide.)

## 9.4.11 Summary of Elastic Registration

The algorithm that has been described is a state-of-the-art example of what is available in the field of fully automatic elastic pixel-based registration. It contains many features that have been proposed in the literature and it has been streamlined for an efficient execution.

Elastic registration has numerous applications in the biomedical imaging field, all based on the basic notion of aligning two images with one other, be it intersubject, intrasubject, or intermodality. It can be used for motion and deformation detection. The deformation field itself can be used for deformation and motion compensation as well as for quantitative measurements.

The criterion, the deformation model, the regularization (penalty), and the optimization algorithm should be all adapted to the task at hand for optimum results.

By producing a specialized program taking advantage of a specific configuration, the runtime can be probably decreased by an additional factor of 10.



Figure 9.17: The first line presents original images number 6,9,11, and 14 from a sequence of originally 60 images of myocardical perfusion MRI. The second line presents the difference images between the original images and their immediate predecessors; movement artifacts can be clearly seen. On the third line you can see the difference images from the motion corrected sequence using our algorithm; the movement artifacts are significantly reduced. The same effect is also visible comparing the differences of the sequence images with the first image of the sequence on the original (fourth line) and corrected (fifth line) sequences.

This, together with the constant advances in computer technology will enable truly interactive operation of automatic and semi-automatic elastic image registration with numerous applications in medicine, biology, and any other field where deformed images need to be compared.

Kybic and Unser



Figure 9.18: The axial, sagital, and coronal views of the two CT brain volumes (one in red, second one in green) prior to registration. Meaningful comparison is difficult. (Color slide.)



Figure 9.19: The axial, sagital, and coronal views of the two CT brain volumes (one in red, second one in green) after the registration. The volumes are aligned, and the large and medium-scale differences were compensated by the registration. This permits to identify more subtle differences. (Color slide.)

## Questions

- 1. What is image registration?
- 2. What is a correspondence function?
- 3. What is image warping?
- 4. What are the usual assumptions in motion analysis?
- 5. How would you use an atlas for segmentation?
- 6. Which four attributes do the authors use to classify registration algorithms?
- 7. What is the purpose of interpolation?
- 8. What is the relation between variational methods minimizing a scalar criterion and methods described using PDEs? [For further study.]
- 9. Why should a warping model be adapted to the image?
- 10. What is the purpose of regularization in image registration?
- 11. What is the essence of multiresolution?
- 12. Explain the notion of consistency in the image registration context.
- 13. Try classify a typical landmark registration algorithm according to the framework described in the first part of this chapter.
- 14. Why does the Laplacian variational criterion not penalize affine deformation?
- 15. Why are radial basis functions called radial?
- 16. What is landmark interpolation?
- 17. Explain the notion of separability for B-spline basis functions and its impact on speed?
- 18. What is the consequence of setting the virtual spring weights  $\alpha$  too high or too low?

- 19. How does the registration result depend on the chosen optimization algorithm?
- 20. What is the relation between the reduction and expansion operators  $\mathcal{R}$  and  $\mathcal{E}$ ?
- 21. Does the positivity of the Jacobian guarantee the global invertibility of the deformation?

## Bibliography

- Bookstein, F., Morphometric Tools for Landmark Data: Geometry and Biology, Cambridge University Press, Cambridge, 1997.
- [2] Mitiche, A. and Bouthemy, P., International Journal of Computer Vision, Vol. 19, pp. 29–55, 1996.
- [3] Weese, J. et al., IEEE Transactions on Information Technology in Biomedicine, Vol. 1, pp. 284–293, December 1997.
- [4] Asari, K. V., Sanjiv, K., and Radhakrishnan, D., IEEE Transactions on Medical Imaging, Vol. 18, p. 345, April 1999.
- [5] Brown, L., ACM Computing Surveys, Vol. 24, p. 326–376, December 1992.
- [6] Kybic, J., Thévenaz, P., and Unser, M., Multiresolution spline warping for EPI registration, In: Proceedings of SPIE, Vol. 3813, pp. 571–579, SPIE, Denver, Colorado, July 1999.
- [7] Bajcsy, R. and Kovačič, S., Computer Vision, Graphics, and Image Processing, Vol. 46, pp. 1–21, 1989.
- [8] Unser, M., Aldroubi, A., and Eden, M., IEEE Transactions on Signal Processing, Vol. 41, No. 2, pp. 821–832, February 1993.
- [9] Unser, M., Aldroubi, A., and Eden, M., IEEE Transactions on Signal Processing, Vol. 41, No. 2, pp. 834–848, February 1993.
- [10] de Boor, C., A Practical Guide to Splines, Springer-Verlag, New York, 1978.
- [11] Meijering, E. H. W., Niessen, W. J., and Viergever, M. A., Quantitative Medical Image Analysis, Vol. 5, No. 2, pp. 111–126, 2001.
- [12] Thévenaz, P., Blu, T., and Unser, M., IEEE Transactions on Medical Imaging, Vol. 19, No. 7, pp. 739–758, July 2000.
- [13] Satter, M. and Goshtasby, A., Registration of deformed images, In: Image Registration Workshop, pp. 221–229, NASA Goddard Space Flight Center, Greenbelt, MD, 1997.

- [14] Faugeras, O. and Keriven, R., IEEE Transactions on Image Processing, Vol. 7, No. 3, pp. 336–344, March 1998.
- [15] Spinei, A., Pellerin, D., and Hérault, J., Signal Processing, Vol. 65, pp. 347–362, 1998.
- [16] Bernard, C., Fast optical flow computation with discrete wavelets, Technical report, Ecole Polytechnique, Paris, 1997.
- [17] Sühling, M., Arigovindan, M., Jansen, C., Hunziker, P., and Unser, M., Myocardial motion analysis and visualization from echocardiograms, In: Progress in Biomedical Optics and Imaging, Vol. 4, No. 23 Sonka, M. and Fitzpatrick, J., eds., Vol. 5032 of Proceedings of the SPIE International Symposium on Medical Imaging: Image Processing (MI'03), pp. 306–313, San Diego CA, USA, February 17-20, 2003, part I.
- [18] Wu, Y., Kanade, T., Cohn, J., and Li, C., Image registration using waveletbased motion model, In: Proceedings of Image Registration Workshop, Le Moigne, J., ed. 1997.
- [19] Rohr, K., Stiehl, H. S., Sprengel, R., Beil, W., Buzug, T. M., Weese, J., and Kuhn, M. H., Point-based elastic registration of medical image data using approximating thin-plate splines, In: Visualization in Biomedical Computing, Höhne, K. H. and Kikinis, R., eds., pp. 297–306, Springer-Verlag, Berlin 1996.
- [20] McGregor, B., Pattern Recognition, Vol. 31, No. 6, pp. 805–817, 1998.
- [21] Bookstein, F. L., IEEE Transactions Pattern Anal. Mach. Intell., Vol. 6, No. 6, pp. 567–585, June 1989.
- [22] Davatzikos, C., Prince, J., and Bryan, R., IEEE Transactions on Medical Imaging, Vol. 15, No. 1, February 1996.
- [23] Huang, C., Chang, W., Wu, L., and Wang, J., IEEE Transactions on Medical Imaging, Vol. 16, No. 5, October 1997.
- [24] Rösch, P. et al., Robust 3D deformation field estimation by template propagation, In: Proceedings of MICCAI, pp. 521–530, 2000.
- [25] Yeung, F., Levinson, F., Fu, D., and Parker, K. J., IEEE Transactions on Medical Imaging, Vol. 17, No. 6, pp. 945–956, December 1998.

- [26] Horn, B. and Schunck, B., Artificial Inteligence, Vol. 17, pp. 185–203, 1981.
- [27] Christensen, G., Deformable Shape Models for Anatomy, Ph.D. Dissertation, Washington University, Saint Louis, Mississippi, 1994.
- [28] Christensen, G., Joshi, S., and Miller, M., IEEE Trans. on Med. Imaging, Vol. 16, No. 6, pp. 864–877, December 1997.
- [29] Bro-Nielsen, M. and Gramkow, C., Fast fluid registration of medical images, In: Visualization in Biomedical Computing, Höhne, K. H. and Kikinis, R., eds., pp. 267–276, Springer-Verlag, Berlin 1996.
- [30] Iosifescu, D. V. et al., Neuroimage, Vol. 14, No. 6, pp. 14–25, 1997.
- [31] Tai, Y., Lin, K., Hoh, C., Huang, S., and Hoffman, E., IEEE Transactions on Nuclear Science, Vol. 44, No. 4, August 1997.
- [32] Hermosillo, G., Variational Methods for Multimodal Image Matching, Ph.D. Dissertation, INRIA, Available at ftp://ftp-sop.inria.fr /robotvis/html/Papers/hermosillo:02.ps.gz, 2002,
- [33] Schormann, T., Henn, S., and Zilles, K., A new approach to fast elastic alignment with applications to human brains, In: Visualization in Biomedical Computing, Höhne, K. H. and Kikinis, R., eds., Springer-Verlag, Berlin 1996.
- [34] Kjems, U., Strother, S. C., Anderson, J., Law, I., and Hansen, L. K., Enhancing the multivariate signal of <sup>15</sup>O water PET studies with a new nonlinear neuroanatomical registration algorithm, IEEE Transactions on Medical Imaging, Vol. 18, No. 4, pp. 306–319, April 1999.
- [35] Bergen, J., Anandan, P., Hanna, K., and Hingorani, R., Hierarchical model-based motion estimation, In: Second European Conference on Computer Vision (ECCV'92), pp. 237–252, Springer-Verlag, Berlin 1992.
- [36] Thévenaz, P., Ruttimann, U. E., and Unser, M., IEEE Transactions on Image Processing, Vol. 7, No. 1, pp. 1–15, January 1998.
- [37] Sicotte, N., Woods, R., and Mazziotta, J., Neuroimage, Vol. 3, No. 3, June 1996, II International Conference on Functional Mapping of the Human Brain.
- [38] Kiebel, S., Ashburner, J., Poline, J., and Friston, K., Neuroimage, No. 5, 1997.
- [39] Ashburner, J. and Friston, K. J., Nonlinear spatial normalization using basis functions, In: Brain Warping, Toga, A. W., ed., pp. 254–266, Academic Press, San Diego, 1999.
- [40] Szeliski, R. and Coughlan, J., International Journal of Computer Vision, Vol. 22, pp. 199–218, 1997.
- [41] Szeliski, R. and Shum, H.-Y., IEEE Trans. Pattern Anal. Mach. Intell., Vol. 18, No. 12, pp. 1199–1207, December 1996.
- [42] Lester, H. and Arridge, S., Summarising fluid registration by thinplate spline warps with many landmarks, In: Medical Image Understanding and Analysis, July 1997, Available at http://www.robots. ox.ac.uk/~mvl/miua97/.
- [43] Zhang, Z., International Journal of Computer Vision, Vol. 13, No. 2, pp. 119–152, December 1994.
- [44] Weese, J., Rösch, P., Netsch, T., Blaffert, T., and M., Q., Gray-value based registration of CT and MR images by maximization of local correlation, In: Proceedings of MICCAI, 1999.
- [45] Thévenaz, P. and Unser, M., IEEE Transactions on Image Processing, Vol. 9, No. 12, pp. 2083–2099, December 2000.
- [46] Kim, B., Boes, J., Frey, K., and Meyer, C., Mutual information for automated multimodal image warping, In: Visualization in Biomedical Computing, Höhne, K. H. and Kikinis, R., ed., Springer-Verlag, 1996.
- [47] Maes, F., Collignong, A., Vandermeulen, D., Marchal, G., and Suetens, P., IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, April 1997.
- [48] Studholme, C., Hill, D., and Hawkes, D., Automated Three-dimensional Medical Physics, Vol. 24, No. 1, January 1997.
- [49] Nikou, C., Heitz, F., and Armspach, J.-P., Pattern Recognition, No. 32, pp. 1351–1368, 1999.

- [50] Duchon, J., Splines minimizing rotation-invariant semi-norms in Sobolev spaces, In: Constructive Theory of Functions of Several Variables, Schempp, W. and Zeller, K., eds., pp. 85–100, Springer-Verlag, Berlin, 1977.
- [51] Suter, D. and Chen, F., IEEE Transactions on Medical Imaging, Vol. 19, No. 4, pp. 295–305, April 2000.
- [52] Hermosillo, G., Chefd'hotel, C., and Faugeras, O., ijcv, Vol. 50, No. 3, pp. 329–343, November 2002.
- [53] Davis, C. Q. and Freeman, D. M., Optical Engineering, Vol. 4, No. 37, pp. 1290–1298, April 1998.
- [54] Fleet, D. and Jepson, A., International Journal of Computer Vision, Vol. 5, No. 1, pp. 77–104, 1990.
- [55] Horn, B. and Weldon, E., The International Journal of Computer Vision, Vol. 2, No. 1, pp. 51–76, June 1988.
- [56] Tomasi, C. and Manduchi, R., IEEE Trans. Pattern Anal. Mach. Intell., Vol. 20, No. 3, March 1998.
- [57] Unser, M., Pelle, G., Brun, P., and Eden, M., Computer analysis of Mmode echocardiograms: estimation of spatial deformation with time, In: Cardiovascular Dynamics and Models, pp. 304–310, Institut National de la Santé et de la Recherche Médicale, Paris, 1988.
- [58] Gimel'farb, G., Pattern Recognition Letters, Vol. 23, No. 4, pp. 431–442, 2002.
- [59] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., Numerical Recipes in C, Cambridge University Press, 2nd edn., 1992.
- [60] Strang, G. and Fix, G., An Analysis of the Finite Element Method, Wellesley-Cambridge Press, 1988.
- [61] Martin, J., Pentland, A., Sclaroff, S., and Kikinis, R., IEEE Trans. Pattern Anal. Mach. Intell., Vol. 2, No. 2, February 1998.
- [62] Moulin, P., Krishnamurthy, R., and Woods, J., IEEE Transactions on Image Processing, Vol. 6, No. 12, pp. 1606–1620, December 1997.

- [63] Brandt, A., Math. Comput., Vol. 31, No. 138, pp. 339–390, 1977.
- [64] Yoshida, H., Removal of normal anatomic structures in radiographs using wavelet-based non-linear variational method for image matching, In: Wavelet Applications in Signal and Imaging Processing VI, Vol. 3458, pp. 174–181, SPIE, 1998.
- [65] Amit, Y., SIAM Journal on Scientific Computing, Vol. 15, No. 1, pp. 207– 224, 1994.
- [66] Dennis, Jr. J., and Schnabel, R., Numerical Methods for Unconstrained Optimisation and Nonlinear Equations, SIAM J. Math. Anal., 1997.
- [67] Thirion, J., Fast non-rigid matching of 3D medical images, Technical Report 2547, INRIA, Nice, May 1995.
- [68] Heitz, F., Perez, P., and Bouthemy, P., Comput. Vis. Graph. Image Process., Vol. 59, No. 1, pp. 125–134, 1994.
- [69] Miller, M., Joshi, S., and Christensen, G., Large deformation fluid diffeomorphisms for landmark and image matching, In: Brain Warping, Toga, A., ed., pp. 115–132, Academic Press, San Diego, 1999.
- [70] Christensen, G. and Johnson, H., IEEE Transactions on Medical Imaging, Vol. 20, No. 7, pp. 568–582, July 2001.
- [71] Musse, O., Heitz, F., and Armspach, J.-P., IEEE Transactions on Medical Imaging, Vol. 10, No. 7, pp. 1081–1093, July 2001.
- [72] Zitová, B. and Flusser, J., Image and Vision Computing, No. 21, pp. 977– 1000, 2003.
- [73] Warfield, S., Robatino, A., Dengler, J., Jolesz, F., and Kikinis, R., Nonlinear registration and template-driven segmentation, In: Brain Warping, Toga, A. W., ed., pp. 67–84, Academic Press, San Diego, 1999.
- [74] Gee, J., Pattern Recognition, No. 32, pp. 99–111, 1999.
- [75] van den Elsen, P. A., Pol, E.-J. D., and Viergever, M. A., IEEE Engineering in Medicine and Biology, pp. 26–39, March 1993.
- [76] Lester, H. and Arridge, S. R., Pattern Recognition, Vol. 32, No. 1, pp. 129– 149, January 1999.

- [77] Verard, L., Allain, P., Travere, J., Baron, J. C., and Bloyet, D., IEEE Transactions on Medical Imaging, Vol. 16, No. 5, pp. 6100–616, October 1997.
- [78] Lam, N. S.-N., The American Cartographer, Vol. 10, No. 2, pp. 129–149, 1983.
- [79] Franke, R., Mathematics of Computation, Vol. 38, No. 157, pp. 181–200, 1982.
- [80] Unser, M. and Blu, T., Wavelets and radial basis functions: A unifying perspective, In: Proceedings of the SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing VIII, Vol. 4119, pp. 487–493, San Diego, CA, 2000.
- [81] Werntges, H. W., Partitions of unity improve neural function approximators, In: Proceedings of IEEE International Conference on Neural Networks, Vol. 2, pp. 914–918, San Francisco, CA, 1993.
- [82] Mallat, S., A Wavelet Tour of Signal Processing, Academic Press, San Diego, CA, 1998.
- [83] Gabrani, M. and Tretiak, O. J., Pattern Recognition, No. 32, pp. 87–97, 1999.
- [84] Duchon, J., Revue Française d'Automatique, Informatique et Recherche Operationelle, Vol. 10, No. 12, pp. 5–12, December 1976.
- [85] Kybic, J., Blu, T., and Unser, M., IEEE Transactions on Signal Processing, Vol. 50, No. 8, pp. 1965–1976, August 2002, ftp://cmp.felk.cvut.cz/ pub/cmp/articles/kybic/Kybic-ieeeSP2002a.pdf.
- [86] Kybic, J., Blu, T., and Unser, M., IEEE Transactions on Signal Processing, Vol. 50, No. 8, pp. 1977–1985, August 2002, ftp://cmp.felk.cvut.cz/ pub/cmp/articles/kybic/Kybic-ieeeSP2002b.pdf.
- [87] Powell, M. J. D., The Theory of radial basis function approximation in 1990, In: Advances in Numerical Analysis II: Wavelets, Subdivision Algorithms and Radial Functions, Light, W. A., ed., pp. 105–210, Oxford University Press, Oxford, UK, 1992.
- [88] Kybic, J., Biomedical Image Processing by Elastic Warping, Ph.D. Dissertation, Ecole Polytechnique Fédérale de Lausanne, Lausanne,

Switzerland, 2001, http://cmp.felk.cvut.cz/kybic/thesis/jkthesis.ps.gz, No. 2439.

- [89] Liouville, J., J. Math. Pures Applications, Vol. 20, (1855), unnumbered, in French.
- [90] Unser, M. and Blu, T., SIAM Review, Vol. 42, No. 1, pp. 43-67, 2000.
- [91] Flandrin, P., IEEE Transactions Inform. Th., Vol. 38, No. 2, pp. 910–917, March 1992.
- [92] Poggio, T. and Girosi, F., Proceedings of IEEE, Vol. 78, No. 0, pp. 1481– 1497, September 1990.
- [93] Kybic, J., Thévenaz, P., and Unser, M., Compensation of unidirectional geometric distortion in EPI using spline warping, In: ICIP'99: Proceedings of the 1999 IEEE International Conference on Image Processing, Vol. II, pp. 168–172, IEEE Computer Society, 445 Hoes Lane, Piscataway, NJ, U.S.A., October 1999, ftp://cmp.felk.cvut.cz/ pub/cmp/articles/kybic/Kybic-ICIP1999.pdf.
- [94] Kybic, J. and Unser, M., Multidimensional elastic registration of images using splines, In: Proceedings of ICIP, Vol. 2, pp. 455–458, 2000, ftp://cmp.felk.cvut.cz/pub/cmp/articles/kybic/Kybic-ICIP2000.ps.gz.
- [95] Kybic, J., Thévenaz, P., Nirkko, A., and Unser, M., IEEE Transactions on Medical Imaging, Vol. 19, No. 2, pp. 80–93, February 2000,
- [96] Kybic, J. and Unser, M., IEEE Transactions on Image Processing, November 2003,
- [97] Studholme, C., Constable, R. T., and Duncan, J. S., Incorporating an image distortion model in non-rigid alignment of EPI with conventional MRI, In: International Conference on Information Processing in Medical Imaging, Lecture Notes in Computer Science, pp. 454–459, Springer-Verlag, Visegrád, Hungary, June 1999.
- [98] Kybic, J., Thévenaz, P., Nirkko, A., and Unser, M., IEEE Transactions on Medical Imaging, Vol. 19, No. 2, pp. 80–93, February 2000.
- [99] Pluim, J. P. W., Maintz, J. B. A., and Viergever, M. A., Image Registration by Maximization of Combined Mutual Information and Gradient

Information, IEEE Transactions Medical Imaging, Vol. 19, No. 8, August 2000.

- [100] Unser, M., Splines: A Perfect Fit for Signal and Image Processing, IEEE Signal Processing Magazine, Vol. 16, No. 6, pp. 22–38, November 1999.
- [101] Blu, T. and Unser, M., Quantitative Fourier analysis of approximation techniques: Part I—Interpolators and projectors, IEEE Transactions on Signal Processing, 1999.
- [102] Fornefett, M., Rohr, K., and Stiehl, H. S., Elastic registration of medical images using radial basis functions with compact support, In: Proc. Computer Vision and Pattern Recognition (CVPR'99), pp. 402–407, Fort Collins, Co, 1999.
- [103] Wu, Y.-T., Image Registration using wavelet-based motion model and its applications, Ph.D. Dissertation, University of Pittsburgh, Saint Louis, Mississippi, 1997.
- [104] Schoenberg, I., Proc. Nat. Acad. Sci., Vol. 52, pp. 947–950, 1964.
- [105] Ahlberg, J. H., Nilson, E. N., and Walsh, J. L., The theory of splines and their applications, Academic Press, New York, 1967.
- [106] Unser, M., Aldroubi, A., and Eden, M., Signal Processing, Vol. 30, No. 2, January 1993.
- [107] Kybic, J., Thévenaz, P., and Unser, M., Compensation of unidirectional geometric distortion in EPI using spline warping, In: Proceedings of the IEEE Internation Conference on Image Processing, pp. 168–172, Kobe, Japan, 1999.
- [108] Müller, F., Brigger, P., Illgner, K., and Unser, M., IEEE Transactions on Signal Processing, Vol. 46, No. 9, pp. 2555–2558, September 1998.
- [109] Unser, M., Aldroubi, A., and Eden, M., IEEE Transactions on Pattern Analysis Mach. Intelligance, Vol. 15, No. 4, April 1993.
- [110] Cachier, P., Mangin, J.-F., Pennec, X., et al., Multisubject non-rigid registration of brain MRI using intensity and geometric features, In: Proceedings of MICCAI, pp. 734–742, Utrecht, The Netherlands, 2001.

- [111] Hellier, P. and Barillot, C., Cooperation between local and global approaches to register brain images, In: Proceedings of IPMI, pp. 315–328, 2001.
- [112] Kybic, J., Biomedical Image Processing by Elastic Warping, Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2001, No. 2439, http://www-sop.inria.fr/robotvis/personnel/Jan.Kybic/ thesis/index.html.
- [113] Kara cali, B. and Davatzikos, C., Topology preservation and regularity in estimated deformation fields, In: Proceedings of IPMI, Lecture Notes in Computer Science, Springer, 2003.
- [114] Kanno, I. and Lassen, N., Journal of Computer Assisted Tomography, Vol. 1, No. 3, pp. 71–76, 1979.
- [115] Kybic, J. and Unser, M., Multidimensional elastic registration of images using splines, In: Proceedings of ICIP, Vol. 2, pp. 455–458, Vancouver, Canada, 2000.
- [116] Ledesmay-Carbayo, M. J., Kybic, J., Desco, M., Santos, A., and Unser, M., Cardiac motion analysis from ultrasound sequences using non-Rigid Registration, In: Proceedings of MICCAI, Niessen, W. J. and Viergever, M. A., eds., pp. 889–896, Springer-Verlag, Utrecht, The Netherlands, October 2001.
- [117] Vallée, J.-P. et al., Magnetic Resonance in Medicine, Vol. 38, pp. 981– 989, 1997.
- [118] Vallée, J.-P. et al., Journal of Magnetic Resonance Imaging, Vol. 9, pp. 197–203, 1999.

### **Chapter 10**

# Cross-Entropy, Reversed Cross-Entropy, and Symmetric Divergence Similarity Measures for 3D Image Registration: A Comparative Study

Yang-Ming Zhu<sup>1</sup> and Steven M. Cochoff<sup>1</sup>

### **10.1 Introduction**

Image registration is a process that quantitatively relates the information in one image to that in another image by determining a one-to-one transformation between coordinates in the two image spaces. Medical image registration is becoming increasingly useful in research and patient care (see [1, 2, 3]). Different imaging modalities often times provide unique and complementary information. Multimodality image registration makes it possible to combine structural (computed tomography or magnetic resonance images) and functional (positron emission tomography or single photon emission tomography) information to improve diagnostic accuracy and aid surgical and/or radiotherapeutic planning. Registration of the same modality images acquired at different times allows clinicians to assess lesion progression/regression or treatment effectiveness. In an interactive, image-guided surgery environment, registration of preoperative images with the physical space is an overriding requirement.

Many algorithms have been employed to register medical images and have recently been reviewed and classified (see [4, 5]). Earlier work prior to 1993 has also been reviewed (see [6, 7]). Fitzpatrick and his colleagues evaluated 16 of

<sup>&</sup>lt;sup>1</sup>PET Engineering, Nuclear Medicine Division, Philips Medical Systems, 595 Miner Road, Cleveland, Ohio 44143, USA

those algorithms visually as well as objectively [8, 9, 10]. Among those different registration algorithms, the voxel similarity approaches to image registration have attracted significant attention since these full-volume-based registration algorithms don't rely upon data reduction or segmentation, and involve little or no user interaction. They can also be fully automated and offer quantitative assessment. Maintz *et al.* [4] lists the reported paradigms and Studholme *et al.* [11], Penney *et al.* [12], and Holden *et al.* [13] compare many of them. Among various different similarity measures, mutual information is the most prominent (see [14, 15, 16], and [17]). Many papers and reports have been published on this similarity measure since its first publication and advances in this area have been recently reviewed in Pluim *et al.* [17].

A cross-entropy optimization approach to image registration was reported recently in Zhu [18]. Cross-entropy minimization as a principle was formally established by Shore and Johnson [19, 20]. They also studied the properties of cross-entropy minimization [21]. In addition to image registration, this measure has been applied to the areas of spectral analysis in Shore [22], image reconstruction in Zhuang *et al.* [23], biochemistry in Yee [24], process control in Alwan *et al.* [25], non-linear programming in Das *et al.* [26], and electron density estimation in Antolin *et al.* [27], among many others.

Cross-entropy, also known as Kullback-Leibler divergence, is an informationtheoretic measure that quantifies the difference between two probability density functions (pdf). It can be either maximized or minimized, depending on how a priori pdf is given. Cross-entropy maximization degenerates to mutual information maximization, conditional entropy minimization or joint entropy minimization under certain conditions. Cross-entropy has two close relatives known as reversed cross-entropy and symmetric divergence, which have been applied to spectral analysis (see [28, 29]) and neural networks [30]. It is reported that cross-entropy, reversed cross-entropy, and symmetric divergence spectral analyses have comparable performance. However, it is not clear how reversed cross-entropy and symmetric divergence perform as registration measures. This chapter explores their use as similarity measures for medical image registration and compares their performance.

Since 1999, the imaging vendors have been developing new imaging devices which combine two different imaging modalities into a single apparatus [31]. General Electric Medical Systems, Philips Medical Systems, and Siemens Medical Solutions/CTI all have released PET/CT devices (Discovery LS, Biograph, and Gemini, respectively, [32]. Given this hardware approach to image registration, the question arises as to the continued need for software registration techniques. It is our opinion however, that software image registration will continue to play a vital role in many instances and that the development of registration algorithms shall remain an important research area for years to come. In many cases, hardware registration is impractical or impossible and one must rely on software-based registration techniques. For example, when monitoring treatment effectiveness over time, software image registration is necessary since the single or multimodality images are acquired at different times. In addition, applications involving intersubject or atlas comparisons require software registration since the images originate from different subjects. Other applications for software registration include the correction of motion that occurs between sequential transmission and emission scans in PET and SPECT as well as the positioning of patients with respect to previously determined treatment plans. The need to offer multiple different combinations of imaging modalities (i.e., PET/MR, SPECT/MR, PET/CT, etc.) would be impractical. As most researchers agree, the hybrid devices will likely play a major role primarily in radiation oncology.

The remainder of this chapter is organized as following: section 10.2 defines all three measures and discusses how they can be used in an image registration context. Section 10.3 addresses the implementation issues. Section 10.4 details the experimental setup to test cross-entropy, reversed cross-entropy, and symmetric divergence image registration by both maximization and minimization. Section 10.5 presents the registration results, along with a discussion. Section 10.6 concludes with a brief summary.

# 10.2 Cross-Entropy, Reversed Cross-Entropy, and Symmetric Divergence

Cross-entropy, reversed cross-entropy, and symmetric divergence can be defined for pdfs of any-dimensional random variables. To make it relevant to the image registration context, in the following equations only a vector variable (u, v) is considered. In these equations, u and v are voxel gray values at corresponding points in two images, f(x, y, z) and g(x, y, z) that are known as the reference image and the floating image, respectively. This gray value pair is considered in almost all similarity measure-based image registration. Note that, if there are more than two images to be registered (see [17, 33]), these definitions can be easily extended to the n-dimensional case [18].

Assume the joint pdf of random variable (u, v) is p(u, v). Also assume a prior estimation of p(u, v) is available and denoted as q(u, v). The cross-entropy is thus defined on a compact support  $D = D_u \times D_v$ , as in Shore and Johnson [19]

$$\eta_{CE}(p,q) = \int_{D} p(u,v) \log \frac{p(u,v)}{q(u,v)} du dv,$$
(10.1)

where  $D_u$  and  $D_v$  are supports of u and v, respectively.

If the roles of q(u, v) and p(u, v) are switched, one has the reversed crossentropy,

$$\eta_{RCE}(p,q) = \int_D q(u,v) \log \frac{q(u,v)}{p(u,v)} \, du \, dv.$$
(10.2)

To make the definition symmetric with regard to p(u, v) and q(u, v), one can combine cross-entropy and reversed cross-entropy,

$$\eta_{SD}(p,q) = \int_{D} [p(u,v) - q(u,v)] \log \frac{p(u,v)}{q(u,v)} du dv, \qquad (10.3)$$

which is the definition of symmetric divergence.

In the case where cross-entropy optimization is utilized to perform image registration, if a favorable (also known as desirable or likely) priori pdf is given, an estimate of the true pdf can be found by minimizing the cross-entropy. On the other hand, if an unfavorable priori pdf is given, an estimate of the true pdf can be obtained by maximizing the cross-entropy. The same would be true when reversed cross-entropy and symmetric divergence are used as similarity measures for image registration.

A favorable pdf can be computed based on previous registration results [34]. Theoretical analysis can also provide information regarding a favorable priori. For example, the voxel values in images of the same modality and of the same patient are linearly related and it has been shown that MR image can be used to simulate a PET image [35]. Since the true pdf is expected to be close to a likely priori, the cross-entropy, reversed cross-entropy, and symmetric divergence will be minimized. It is worthy to note that the cross-entropy, reversed cross-entropy, and symmetric divergence are not the only frameworks to exploit the priori knowledge in registration. The recently proposed likelihood maximization approach can also systematically use this knowledge (see [34, 36]).

When an undesirable prior q(u, v) is given, the cross-entropy, reversed crossentropy, or symmetric divergence are maximized. When images are in registration, we would expect that the voxel values at the same coordinates in the two images are related. Therefore, unfavorable pdfs can be uniform, proportional to one of the marginal pdfs, or the product of two marginal pdfs. The case where the unfavorable pdf is the product of two marginal pdfs is worth noting. In this case, the voxel values are statistically independent since their joint pdf is the product of their marginal pdfs. Cross-entropy maximization based on these three priori estimates degenerate to entropy minimization, conditional entropy minimization, and mutual information maximization, respectively, [18].

Let's assume the marginal pdfs of u and v are  $h_f(u)$  and  $h_g(v)$ , respectively, and the unfavorable priori pdf q(u, v) is equal to the product of two marginal pdfs. Putting  $q = h_f(u)h_g(v)$  into Eqs. (10.1) (10.2) and (10.3), one arrives at

$$\eta_{CE} = \int_D p(u,v) \log \frac{p(u,v)}{h_f(u)h_g(v)} du dv, \qquad (10.4)$$

$$\eta_{RCE} = \int_{D} h_{f}(u) h_{g}(v) \log \frac{h_{f}(u) h_{g}(v)}{p(u, v)} \, du \, dv, \qquad (10.5)$$

and

$$\eta_{SD} = \int_{D} [p(u, v) - h_f(u)h_g(v)] \log \frac{p(u, v)}{h_f(u)h_g(v)} \, du \, dv.$$
(10.6)

One may notice that Eq. (10.4) is the definition of mutual information. The measures defined in Eqs. (10.4-10.6) are maximized when used in image registration.

### **10.3 Implementation**

Figure 10.1 shows the overall flow chart of the registration process. It applies to all three similarity measures as well as the minimization and maximization cases. When the similarity measures are minimized, a favorable priori is used in the similarity calculation. When they are maximized, the priori estimate is calculated from the marginal pdfs on-the-fly. As can be seen, no matter which measure is used, the registration algorithm has the same structure and they can be implemented in the same manner. Since Eqs. (10.1)-(10.6) are in a continuous form, they must be discretized to be solved numerically.



Figure 10.1: Flow chart of image registration by the cross-entropy, reversed cross-entropy, and symmetric divergence optimization. CE: cross-entropy; RCE: reversed cross-entropy; SD: symmetric divergence.

The two images involved in the registration are identified as the reference image and the floating image. The floating image will undergo rotation and translation to match the reference image. Before the automatic registration starts, an initial set of registration parameters must be set and the floating image is appropriately transformed (i.e., rotated and translated). The similarity measures between the reference image and the transformed floating image are then computed. If the similarity number is not optimal, the registration parameters are updated, otherwise, the registration process stops and the optimal registration parameters are output. The scheme to update the registration parameters is determined by the optimization algorithm employed. In the following subsections, the key steps in the registration process are expanded and discussed in detail.

### **10.3.1** Transformation

Although all three registration measures can be applied to more general transformations, we restricted the transformation to rigid-body transformations. For the rigid-body registration, the registration parameter is a six-dimensional vector,  $(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$ , where  $\theta_x, \theta_y$ , and  $\theta_z$  are rotation angles in degrees around the *x*-, *y*-, and *z*-axis, respectively, and  $t_x, t_y$ , and  $t_z$  are translation offsets along the *x*-, *y*-, and *z*-axis, respectively.

There is a  $4 \times 4$  matrix corresponding to any of the three rotations in a homogeneous coordinate system [37]. Since a successive application of the rotations amounts to matrix multiplication which is not commutative, the order of these rotations is important. It is assumed in this chapter that the rotation angles are applied around the *x*-, *y*-, and *z*-axis, in that order, and that the rotation is performed before the translation.

### 10.3.2 Marginal, Joint, and Priori Probability Density Functions

The marginal and joint pdfs of gray value pairs were estimated from normalized histograms. The maximum voxel value of image f was found first. The voxel values in image f were then divided into 64 discrete levels. Similarly, the voxel values in image g were divided into 64 discrete values. The size of the joint histogram is thus  $64 \times 64$ . Alternatively, one can change the bin size adaptively [18]. In the overlapping volume, the histograms of voxel values in images f and g, and of the voxel pairs were calculated by binning the voxel values and value pairs. The normalized histograms then give estimations of the marginal as well as the joint distributions. To compute the cross-entropy, reversed cross-entropy, and symmetric divergence, these estimated pdfs can be directly substituted into the formulas defined in Eqs. (10.1)–(10.6).

When a favorable priori estimate is employed in the registration, it can be computed in the same fashion as that for the joint pdf. The priori estimate shall be calculated with the same size of histograms as that for the joint histogram.

### **10.3.3 Interpolation**

A grid point in one volume will generally not exactly coincide with another grid point in the transformed space. Since the voxel values of the reference image are defined at specific grid points, to compute the joint pdf by binning voxel value pairs, one needs to interpolate the voxel values in the transformed space. There are different interpolation methods in three-dimensional space. Tsao [16] compared eight such methods. Interpolation artifacts are reported which introduce spurious fluctuations in the similarity measures and impact the optimization behavior. For simplicity, the trilinear interpolation was used in our study.

#### **10.3.4** Optimization

Powell's multidimensional direction set optimization is used to minimize the three similarity measures when favorable priori estimates are used using Brent's method in one-dimensional search [39]. The same technique is also used to minimize the negated similarity measures when unfavorable priori estimates are applied. The direction matrix is initialized to a unitary matrix. The vector is  $(\theta_x, \theta_y, \theta_z, t_x, t_y, t_z)$ , as explained before. A reordering of these registration parameters is possible which may improve the optimization speed as in Maes *et al.* [15]. We did not try to optimize the parameter order since the order may be image content dependent and an exhaustive trial seems impractical (there are 6! = 720 different combinations although one may try a subset of them). Furthermore, Powell's optimization may use six independent directions which do not necessarily correspond to the six desired directions as the search proceeds (see [15, 39]).

The stop criterion in the Powell's algorithm was set to ftol = 0.001.

### **10.3.5 Multiresolution**

The Powell's optimization method cannot guarantee that a global optimal value will be found since it can be easily trapped to a local optimum. To find a true global optimal value, Ritter *et al.* [40] successfully applied the simulated annealing [39] algorithm to two-dimensional image registration. It is a stochastic method and is slow, which limits its application to three-dimensional image registration. Pluim *et al.* used genetic algorithms [17] or others used multistarter methods to pinpoint a global optimum. However, the desired registration is frequently a local rather than global extremum of the similarity measure, as discussed in Fitzpatrick *et al.* [41]. This can occur when the image resolution is

low, the image size is small, or the overlap of images is limited. In practice, the multiresolution approach proves to be helpful. It can improve the optimization speed, increase the capture range and the algorithm is relatively robust [42]. In our implementation, the images were folded down to a  $16 \times 16 \times 16$  image as the most coarse image. The resolution of the successive images was doubled until the full image resolution or  $64 \times 64 \times 64$  was reached in all three dimensions. We used fine resolutions beyond  $64 \times 64 \times 64$  when the image size permitted it. But most cases did not exhibit any sizable improvements on the registration accuracy, and there was almost no effect on the success rate. To obtain the coarse images, the voxel values within a sampling volume are averaged. Although it is a little slower than a simple subsampling approach, the averaging technique results in a better registration [18].

### **10.3.6 Numerical Stability**

Referring to the definitions of cross-entropy, reversed cross-entropy, and symmetric divergence, one may find that there may be a numerical problem under some conditions. The instability is caused when the priori or joint probability is zero. Ideally, one could sample a large dataset to get a better estimate of the priori or use sophisticated sampling schemes to better estimate the joint probability. Since those are computationally demanding, we chose a simple, non-exact approach that is described below.

For cross-entropy maximization, if the joint probability is zero, the contribution to the cross-entropy is zero  $(0 \log 0 = 0)$ . If the marginal probability is zero, the joint probability will be zero and will have no contribution to the cross-entropy measure. In the cross-entropy calculation, therefore, the terms can be ignored when the joint or marginal probabilities are zero.

For the cross-entropy minimization, the contribution to the cross-entropy would be infinite when the priori probability is zero. We elected to ignore these terms since the cross-entropy is being minimized. Alternatively, one could assign a large value to the cross-entropy under this situation. When picking such a large value, one should take into account the stop condition of the optimization process. If the assigned value is too large, the optimization can prematurely terminate. We chose a positive value, that resulted in the cross-entropy minimization having a small capture range. Note that if the joint probability is zero, the terms can also be ignored. For the reversed cross-entropy maximization, if the marginal probability is zero, the contribution to the reversed cross-entropy is zero due to the fact that  $0 \log 0 = 0$ . If the joint probability is zero and the marginal probability is not zero, then the contribution to the reversed cross-entropy is infinite. Since the reversed cross-entropy is being maximized and we assume the optimized reversed cross-entropy has a finite value, this instability due to the zero joint pdf terms must be addressed. As shown later in this chapter, if the zero joint probability terms are ignored, the profile of the reversed cross-entropy will have a zig-zag shape, which adversely affects the optimization process and renders the reversed cross-entropy an inappropriate measure for registration. We elected to assign a small probability value to the zero joint probabilities, which is  $\frac{1}{binsize \times alpha}$ , where  $\alpha$  was selected empirically to be 100. We will discuss the impact of the  $\alpha$  selection in a later section.

For the reversed cross-entropy minimization, if the priori probability is zero, then the contribution to the reversed cross-entropy is zero due to the fact that  $0 \log 0 = 0$ . If the joint probability is zero and the priori probability is not zero, the contribution to the reversed cross-entropy is infinite. In the later case, since we assume the minimized reversed cross-entropy has a lower bound, their contributions are ignored.

The symmetric divergence method utilizes the cross-entropy and reversed cross-entropy components and its numerical stability is determined by those two terms.

### **10.4 Registration Experimental Setup**

#### **10.4.1** Dataset Description

MR, nuclear transmission (Tx) and nuclear emission (Em) scans were used as test volumes. They were primary brain images, with some inclusion of the neck and chest. The image data consists of slices. The *x*-axis is directed horizontally from right to left, the *y*-axis horizontally from front to back, and the *z*-axis vertically from bottom to top.

This study involved seven patients. All the pertinent image file information is tabulated in Table 10.1. The MR images were acquired with an OutLook MR scanner (Philips Medical Systems, formerly Marconi Medical Systems, Cleveland,

Patient	Modality	Image	Dimension	Voxel $(mm^3)$	Comments
A	MR	Sag	$256^{2} \times 128$	$1.0^2  imes 1.5$	m, 29 yr
	Tx	Axi	$64^2 \times 29$	$7.12^{3}$	
	Em	Axi	$64^2 \times 29$	$7.12^{3}$	
В	MR	Sag	$256^2 \times 128$	$1.0^2  imes 1.5$	m, 21 yr
	Tx	Axi	$64^2 \times 24$	$7.12^{3}$	
	Em	Axi	$64^2 \times 24$	$7.12^{3}$	
С	MR	Sag	$256^2 \times 128$	$1.0^2  imes 1.5$	f, 70 yr
	Tx	Axi	$64^2 \times 22$	$7.12^{3}$	
	Em	Axi	$64^2 \times 22$	$7.12^{3}$	
D	MR	Axi	$192\times 256\times 24$	$1.17^2  imes 6.0$	m, 45 yr
	Tx	Axi	$64^2 \times 31$	$6.23^2 \times 7.12$	
	Em	Axi	$64^2 \times 31$	$6.23^2 \times 7.12$	
Е	MR	Sag	$180^2 \times 120$	$1.41^2 \times 1.5$	m, 40 yr
	Tx	Axi	$64^2 \times 26$	$6.23^2 \times 7.12$	
	Em	Axi	$64^2 \times 26$	$6.23^2 \times 7.12$	
F	MR	Sag	$192\times 256\times 120$	$1.41^2 \times 1.5$	m, 60 yr
	Tx	Axi	$64^2 \times 53$	$3.50^{3}$	
	Em	Axi	$64^2 \times 53$	$3.50^{3}$	
G	MR	Sag	$192\times 256\times 120$	$1.41^2 \times 1.5$	m, 13 yr
	Tx	Axi	$64^2 \times 28$	$6.23^2 \times 7.12$	
	Em	Axi	$64^2 \times 28$	$6.23^2\times7.12$	

Table 10.1: Image file descriptions, where Tx stands for transmission scan,Em emission scan, Sag sagittal slices, and Axi axial slices

Ohio), and the nuclear images were acquired with a Prism 3000XP (Philips Medical Systems) scanner. In all Tx and Em scans, Technetium-99m hexamethylpropolamine-oxime (Tc-99m HMPAO) was used as the isotope. The Tx and Em images were acquired simultaneously using the STEP (simultaneous transmission emission protocol) protocol. Thus, the Tx and Em images are registered intrinsically (see [43, 44, 45, 46]).

### 10.4.2 Experimental Setup—Maximization of Similarity Measures

Cross-entropy, reversed cross-entropy, and symmetric divergence are maximized when the joint pdf is derived from an optimal registration and it differs greatly from a unlikely priori estimate. This subsection discusses the details of the experimental setup as well as the determination of the registration accuracy. In the case of retrospective multimodality registration, the correct registration parameters are typically unknown. Various evaluation methods have been used to assess registration accuracy, including phantom validation, observer assessment [11], fiducial marks [8, 9, 10], and cross validation [47]. Since the Tx and Em scans used in our study are intrinsically registered (see [43, 44, 45, 46]), we have us an ideal method to evaluate the registration accuracy. In this special case, the accuracy of a registration technique can be assessed by evaluating the registration parameters that are generated when these Tx and Em images are used as input.

The Tx and Em images were also individually registered to the MR images. Based on those registration results, the registration parameters between the Tx and Em images were indirectly calculated. If we denote the transformation matrix relating the MR and Tx images,  $M_{MR/Tx}$ , and the matrix relating the MR and Em images,  $M_{MR/Em}$ , the matrix relating the Tx and Em images can be calculated as  $M_{Tx/Em} = M_{MR/Tx}^{-1} M_{MR/Em}$ .

For this indirect registration, two sets of experiments were performed. In the first set, no attempt was made to manually preregister the images before the automatic registration started. In the second set, all images were initially brought into approximate registration using a manual method. The same manual adjustments were used in cross-entropy, reversed cross-entropy, and symmetric divergence registrations.

If the MR/Tx and MR/Em registrations deviate from the true registration in the same manner, the indirect registration evaluation may yield a misleading registration accuracy. To guard against this, all registration results were visually checked. Only if both registrations were visually acceptable, were their registration results used to indirectly calculate the Tx/Em registration. Since it is reported in Studholme *et al.* [11] that clinicians can detect registration parameter differences of  $4^{\circ}$  in x and y rotation angles,  $2^{\circ}$  in the z rotation angle, 2 mm in x and y translations, and 3 mm in z translation, the registrations were believed to be close to the truth if the results were visually acceptable.

The accuracy and robustness of direct and indirect Tx/Em registrations were analyzed in terms of the mean and standard deviation of the differences between the resultant registration parameters and the intrinsic ones. Alternatively, one can compute the difference of a group of selected points. As reported in West *et al.* [8], the mean of point differences is related to the rotation and translation registration parameters. In this chapter, we chose not to report the point differences.

To further understand the behavior of similarity measures in the vicinity of an optimal registration, the similarity measures were computed for images in the neighborhood of a manual registration. It is believed that the manual results are close to the ground truth. Thus, this calculation would shed light on the behavior of these measures. To this end, we focused on the MR/Em image pair of patient B, arbitrarily.

We also evaluated different strategies to cope with the numerical instability that occurs when maximizing the selected similarity measures. The strategies were evaluated by inspecting the similarity profiles that resulted from systematically adjusting the registration parameters from those obtained when the images underwent manual registration.

### 10.4.3 Experiment Setup—Minimization of Similarity Measures

If one has a good estimation of the joint voxel value distribution, then the crossentropy, reversed cross-entropy, and symmetric divergence measures can be minimized to find the optimal registration. Although a reasonably good prior estimation is difficult to obtain, we report on the outcome of some experiments to validate the idea. We did the experiment on the MR/Em image pair of Patient B.

The image pair was registered by four clinical experts using an interactive (manual) registration method. A joint voxel value distribution was calculated based on the averaged manual registration parameters. This distribution was then used as the priori. It is worthy to note that the priori probabilities need to be recomputed for each resolution when the multiresolution optimization is used. In this test, only two resolution levels were employed, 32 and 64.

To assess the robustness of registration by the cross-entropy, reversed crossentropy, and symmetric divergence minimization, two sets of misregistrations were randomly generated around the above-mentioned average registration results, and used as initial registrations. In the first set of 100 misregistrations (Set 1), the differences between the rotation angles and the average rotation angles were uniformly distributed over [-10, 10] degrees and the differences between the translation offsets were uniformly distributed over [-10, 10] mm. For the second set of 100 registrations (Set 2), the distributions were expanded to [-20, 20] degrees and mm.

Since the true registration parameters are unknown, we estimated them by averaging the manual registration parameters obtained from four clinical experts. A registration was judged to be successful if the registration parameters were close to the averaged manual values. The thresholds were equal to the detection thresholds [11], i.e.,  $(4^{\circ}, 4^{\circ}, 2^{\circ}, 2 \text{ mm}, 2 \text{ mm}, 3 \text{ mm})$ . Note that the thresholds are tighter than what were used in Zhu [18].

The accuracy and robustness of registrations obtained by minimizing the similarity measures were analyzed in terms of the mean and standard deviation of the differences between the resultant registration parameters and the averaged manual registration parameters.

To understand the behavior of these similarity measures in the vicinity of an optimal registration, the similarity measures were computed in the neighborhood of manual registration results. It is believed that the manual results are close to the ground truth. Thus, this calculation would shed light on the behavior of these measures.

Different strategies to handle the potential numerical instability problem in this minimization approach are compared by inspecting their similarity function profiles in the vicinity of the manual registration.

### **10.5 Results and Discussions**

## 10.5.1 Registration by Cross-Entropy, Reversed Cross-Entropy, and Symmetric Divergence Maximization

#### 10.5.1.1 Tx/Em Direct Registration

When the Em images were registered to the Tx images directly, the cross-entropy, reversed cross-entropy, and symmetric maximization successfully registered 4, 5, and 7 (out of 7) cases, respectively. The results of failed cases by the cross-entropy and reversed cross-entropy techniques were significantly different from the intrinsic registrations by visual inspection (the misalignment angles or translation offsets are too big). The mean and standard deviation of registration

Table 10.2: Average and standard deviation of Tx/Em direct registration parameters obtained from the cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) registration measures. The angles are in degrees, translation offsets in mm

Alg	$ heta_x$	$ heta_y$	$\theta_{z}$	$t_x$	$t_y$	$t_z$	Success
CE	$0.64 \pm 1.27$	$0.00\pm0.00$	$0.85 \pm 1.71$	$0.21\pm0.43$	$0.19\pm0.89$	$-0.98 \pm 1.13$	4
RCE	$0.36 \pm 1.34$	$0.35\pm0.63$	$-0.55 \pm 1.34$	$-0.14\pm1.17$	$1.18 \pm 1.40$	$1.26\pm0.69$	5
SD	$0.86 \pm 1.06$	$0.30 \pm 1.53$	$0.14\pm0.79$	$-0.07\pm0.92$	$0.79\pm0.69$	$-0.97 \pm 1.39$	7

parameters for all successful registrations obtained from the cross-entropy, reversed cross-entropy, and symmetric divergence maximization are tabulated in Table 10.2.

From the data in Table 10.2, one can conclude that registration by the symmetric divergence maximization has the higher success rate. For all cross-entropy, reversed cross-entropy, and symmetric divergence image registration, the average differences of the registration parameters from the true parameters are below the detection thresholds of a trained clinician.

The failed cases of cross-entropy and reversed cross-entropy registration are intriguing. The diameter of the field of view of the Tx and Em slice images is 46 cm. A brain with a 20 cm diameter only occupies 1/6 or 1/7 of the slice area. Thus, in the histogram estimation, many voxel pairs populate in the lowest bins, which has an adverse effect on the registration. We tried to exclude them by discarding the low end of the histogram from the similarity measure calculation and the cross-entropy successfully registered all seven cases, but with larger mean and standard deviation (data not shown here). Moreover, the same exclusion scheme reduced the success rate of MR/Tx and MR/Em registration by cross-entropy maximization. Therefore, no data exclusion techniques were used in any of the experiments reported in this chapter.

### 10.5.1.2 Tx/Em Indirect Registration

The Tx and Em images were indirectly registered through the MR images. Figure 10.2 shows typical MR/Em registration results (of patient B). The axial, sagittal, and coronal views are overlaid with a changeable transparency. All three views are correlated by a movable cross-hair token. In the top-left



Figure 10.2: A typical registration result of MR/Em (patient B).

quadrant, the surface is generated from the MR image using a marching cube algorithm. In the cut-away view, the MR and Em are overlaid. The cut plane position can also be changed by the cross-hair token. The three-dimensional view can be manipulated to select the desired view direction. This visualization tool provides an excellent means to check, visualize, and fuse the registration results.

The cross-entropy maximization registration failed to register one MR/Tx case. Thus, one Tx/Em set of indirect registration parameters were not computed. The mean and standard deviation of indirect registration parameters associating the Tx and Em images are reported in Table 10.3. Overall, the mean and standard deviation are larger than those in the direct registration (see Table 10.2), suggesting that errors in MR/Tx and MR/Em registrations may be

Table 10.3: Average and standard deviation of Tx/Em indirect registration parameters obtained from the cross-entropy, reversed cross-entropy, and symmetric divergence registration measures without manual preregistration. The angles are in degrees and translation offsets in mm

$\theta_x$	$ heta_y$	$\theta_z$	$t_x$	$t_y$	$t_z$	Success
$2.33 \pm 1.49$	$0.27\pm0.99$	$-0.67\pm0.90$	$0.16\pm0.47$	$0.05 \pm 1.05$	$3.17 \pm 1.61$	6
$3.38 \pm 1.37$	$-0.29 \pm 1.52$	$-0.77 \pm 1.97$	$-0.37 \pm 1.13$	$1.21 \pm 2.34$	$0.89 \pm 3.25$	5
$2.35\pm3.00$	$-0.03\pm0.37$	$-0.75\pm1.41$	$-0.13\pm0.45$	$1.79 \pm 2.33$	$-0.41\pm2.17$	5
	$\theta_x$ 2.33 ± 1.49 3.38 ± 1.37 2.35 ± 3.00	$\begin{array}{c} \theta_{x} & \theta_{y} \\ \\ 2.33 \pm 1.49 & 0.27 \pm 0.99 \\ 3.38 \pm 1.37 & -0.29 \pm 1.52 \\ 2.35 \pm 3.00 & -0.03 \pm 0.37 \end{array}$	$\theta_x$ $\theta_y$ $\theta_z$ 2.33 ± 1.49         0.27 ± 0.99         -0.67 ± 0.90           3.38 ± 1.37         -0.29 ± 1.52         -0.77 ± 1.97           2.35 ± 3.00         -0.03 ± 0.37         -0.75 ± 1.41	$\theta_x$ $\theta_y$ $\theta_z$ $t_x$ 2.33 $\pm$ 1.490.27 $\pm$ 0.99-0.67 $\pm$ 0.900.16 $\pm$ 0.473.38 $\pm$ 1.37-0.29 $\pm$ 1.52-0.77 $\pm$ 1.97-0.37 $\pm$ 1.132.35 $\pm$ 3.00-0.03 $\pm$ 0.37-0.75 $\pm$ 1.41-0.13 $\pm$ 0.45	$\theta_x$ $\theta_y$ $\theta_z$ $t_x$ $t_y$ 2.33 ± 1.490.27 ± 0.99-0.67 ± 0.900.16 ± 0.470.05 ± 1.053.38 ± 1.37-0.29 ± 1.52-0.77 ± 1.97-0.37 ± 1.131.21 ± 2.342.35 ± 3.00-0.03 ± 0.37-0.75 ± 1.41-0.13 ± 0.451.79 ± 2.33	$\begin{array}{c ccccccccccccccccccccccccccccccccccc$

compounded and propagated into the indirect computation of the Tx/Em registration parameters.

Both the reversed cross-entropy and symmetric divergence maximization failed to register one MR/Tx and one MR/Em case. This resulted in two Tx/Em cases that could not be registered indirectly for both the techniques. The corresponding means and standard deviations for indirectly registered Tx/Em are listed in Table 10.3. It seems that the reversed cross-entropy and symmetric divergence maximization estimated the z translation parameters more accurately. For the cross-entropy maximization, the large error in the z translation parameters that, among MR/Tx and MR/Em registrations, one overestimates that parameter and the other underestimates that parameter.

The study of cross-entropy, reversed cross-entropy, and symmetric divergence image registration includes two aspects: (1) the determination if the similarity measure is suitable for image registration and (2) how to accurately and robustly find the optimal registration associated with that measure. The lower success rates for the reversed cross-entropy and symmetric divergence techniques when registering MR/Tx and MR/Em are not sufficient to reject them as similarity measures for registration. It may simply indicate that reversed crossentropy and symmetric divergence have a very narrow capture range when used as registration similarity measures. That is, if the initial registration is far away from the optimal registration, it is hard for the iterative optimization routine to converge to an optimal solution. As a matter of fact, the angular registration parameters of MR/Tx and MR/Em can be as large as 30° and the translation parameters can be as large as 55 mm. To determine if the failed registrations were caused by the limited capture range, all image pairs were manually registered and the manual results were used as starting points for iterative optimization.

Table 10.4: Average and standard deviation of Tx/Em indirect registration parameters obtained from the cross-entropy, reversed cross-entropy, and symmetric divergence registration measures with manual pre-registration. The angles are in degrees and translation offsets in mm

Alg	$ heta_x$	$\theta_y$	$\theta_z$	$t_x$	$t_y$	$t_z$	Success
CE	$2.26 \pm 1.06$	$0.64 \pm 0.81$	$0.31 \pm 1.23$	$0.45\pm0.25$	$0.34 \pm 0.77$	$-3.69\pm2.46$	7
RCE	$2.55 \pm 1.06$	$1.29 \pm 1.16$	$1.40\pm2.26$	$0.76 \pm 2.02$	$1.91 \pm 2.34$	$0.78 \pm 2.01$	7
SD	$1.79 \pm 1.21$	$0.91 \pm 1.80$	$1.32\pm2.27$	$0.70 \pm 1.91$	$1.70\pm2.00$	$-0.27\pm2.69$	7

The results of mean and standard deviation for indirect Tx/Em registrations are tabulated in Table 10.4. Again, the large error in the z translation parameter for the cross-entropy maximization indicates that, among MR/Tx and MR/Em registrations, one overestimates that parameter and the other underestimates that parameter.

As can be seen, all three similarity measures successfully registered seven MR/Tx and seven MR/Em cases with a manual prealignment, and the means and standard deviations are comparable to those in Table 10.3. As a comparison, the Tx/Em indirect registration parameters were also computed from manual MR/Tx and MR/Em registration. The mean and standard deviation of the Tx/Em indirect registration parameters were  $(2.42 \pm 2.69, 0.22 \pm 2.20, 0.15 \pm 2.50, 2.66 \pm$  $1.94, 0.69 \pm 2.01, 2.31 \pm 3.14$ ). Since most of the manual registration results had zeros in the y- and z-axis rotation, the mean values of these parameters are small. Thus, a small mean does not necessarily mean that the manual registration is more accurate. Since the reversed cross-entropy and symmetric divergence registration has a small capture range, one might be concerned that their registration results would not deviate significantly from the manual starting registration. If this were true, the mean and standard deviation values in Table 10.4 would reflect the mean and standard deviation associated with the manual registration results. To check this possible situation, the difference of the cross-entropy, reversed cross-entropy, and symmetric divergence registration results and the manual starting solutions was checked. The angle difference was as large as  $5^{\circ}$  and the translation was as large as 6 mm. Thus, it is unlikely that the mean and standard deviation for the reversed cross-entropy and symmetric divergence registration merely reflect the mean and standard deviation in the manual registration.



Figure 10.3: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the x-rotation angle changes independently.

#### 10.5.1.3 Behavior Around True Registration

Three similarity measures were computed in the vicinity of an optimal registration. Since the ground truth is unknown, we used the manual registration result as an approximation to the ground truth. It is difficult to visualize the hypersurface in seven-dimensional space. Therefore, only one registration parameter was changed at a time when a profile was computed.

Figures 10.3–10.5 show the representative profiles of these similarity measures as a function of rotation angles about the *x*-, *y*-, and *z*-axis, respectively. The horizontal axis (registration parameters) in all profiles reported in this chapter was normalized to the manual registration parameters. All similarity values have been normalized with a maximum value of 1. As can be seen, the cross-entropy, reversed cross-entropy, and symmetric divergence measures peak around the manual registration. However, there are differences as large as  $3^{\circ}$  in the peak



Figure 10.4: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the y-rotation angle changes independently.



Figure 10.5: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the z-rotation angle changes independently.



Figure 10.6: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the x-translation changes independently.

positions for the different measures which is understandable since they are defined differently. The figures also revealed that the curve for the cross-entropy measure is slightly smoother than those associated with the reversed crossentropy and symmetric divergence measures. This rough nature in the reversed cross-entropy and symmetric divergence profiles may result in the inferior performance of iterative optimization and a small capture range for these similarity measures. This behavior may be caused by the presence of many local optima that can confuse the optimization procedure.

Similarly, Figs. 10.6–10.8 show the representative profiles of the crossentropy, reversed cross-entropy, and symmetric divergence measures as a function of x-, y-, and z-translation, in the neighborhood of a manual registration. All three profiles peak around the manual registration results. This behavior is one of the necessary conditions for a good registration measure. The difference of peak positions of three similarity measures is small (ca. 1 mm) in the x and y



Figure 10.7: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the y-translation changes independently.

directions, but sizable in the z direction (ca. 4 mm). The cross-entropy curve is the smoothest.

Although we don't know if the peak appearing in the neighborhood of the manual registration represents the global optimal, the presence of a local maxima does indicate that the three measures are suitable as registration criteria. The difference in the peak positions is expected since the similarity measures are different. Furthermore, the Em image used in our study has a low resolution (7.12 mm in all three directions) and the difference in the peak positions is not significant compared to the voxel size.

#### **10.5.1.4** Numerical Stability for Maximization

As mentioned earlier, it could cause numerical problems when there are zero joint probabilities in the reversed cross-entropy and symmetric divergence



Figure 10.8: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles of an MR/Em image pair in the vicinity of a manual registration when the z-translation changes independently.

calculation. We must cope with this numerical instability. A simple, heuristic approach is described in Section 10.3.6. In this section, the behaviors of different  $\alpha$  selections are compared. Since the symmetric divergence measure contains the reversed cross-entropy component, the experiment was only performed using the reversed cross-entropy measure. One could also ignore the contributions associated with the zero joint pdfs. As a comparison, this strategy is also evaluated.

Figures 10.9 and 10.10 show the reversed cross-entropy profiles when the *x*-rotation angle and the *x*-translation change independently. The profiles are labelled with the  $\alpha$  values used to compute a small probability value which replaces the zero pdf. The results when the zero pdfs are ignored in the profile calculation are also shown.

As these two figures reveal, the reversed cross-entropy profiles are very rough if the zero pdfs are ignored in the calculation. For the *x*-rotation angle,



Figure 10.9: The reversed cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-rotation angle changes independently. The numbers adjacent to the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdfs. The bottom profile represents the case where the contribution from zero joint pdfs was ignored in the profile computation.

there is not even a peak around the manual registration. These two undesirable features render the reversed cross-entropy measure unsuitable as a registration measure if the zero pdfs are ignored. If the zero pdfs are taken into account and as  $\alpha$  increases, the profiles become smooth and then rough again. It seems that  $\alpha = 100$  is a reasonable choice, which gives a relatively smooth profile. When  $\alpha$  is small, the extraneous pdfs replacing the zero pdfs causes the summation of the pdfs to significantly deviate from unity. Therefore, a small  $\alpha$  is not ideal. When  $\alpha$  is large, the contribution of each extraneous pdf becomes dominant, the total reversed cross-entropy increases, and the contribution from nonzero pdfs is diminished gradually. The latter behavior makes a large  $\alpha$  not ideal either. When  $\alpha = 100$ , the summation of all pdfs, including the extraneous ones, is at



Figure 10.10: The reversed cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-translation changes independently. The numbers adjacent to the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdfs. Ignored means that those pdfs are ignored when computing the profile.

most 1.01. The upper bound can be achieved only when there is 1 nonzero pdf which is 1, and all other pdfs are zero. Based on these observations, we chose  $\alpha = 100$  in our calculation.

## 10.5.2 Registration by Cross-Entropy, Reversed Cross-Entropy, and Symmetric Divergence Minimization

### 10.5.2.1 Registration Performance

The standard deviations of four independent manual registration results were, in vector form, (3.48, 0.53, 3.11, 1.68, 2.98, 1.83). The likely priori was calculated

Table 10.5: Average and standard deviation of MR/Em registration parameters obtained from cross-entropy, reversed cross-entropy, and symmetric divergence minimization. The angles are in degrees and translation offsets in mm

Alg	$\theta_x$	$ heta_y$	$\theta_z$	$t_x$	$t_y$	$t_z$	Success
Set 1							
CE	$0.03\pm0.80$	$0.18\pm0.74$	$-0.13\pm0.44$	$0.03\pm0.13$	$-0.02\pm0.42$	$-0.04\pm0.17$	46%
RCE	$-0.05\pm0.45$	$0.09 \pm 0.39$	$0.07\pm0.31$	$0.07\pm0.33$	$0.06\pm0.31$	$-0.01\pm0.10$	53%
SD	$0.00\pm0.10$	$0.06\pm0.34$	$0.05\pm0.29$	$0.05\pm0.30$	$0.02\pm0.13$	$-0.02\pm0.11$	68%
Set 2							
CE	$0.23 \pm 1.04$	$0.24\pm0.87$	$0.02\pm0.51$	$0.10\pm0.50$	$-0.07\pm0.59$	$-0.11\pm0.29$	41%
RCE	$0.02\pm0.29$	$0.10\pm0.43$	$0.04\pm0.35$	$0.06\pm0.38$	$0.00\pm0.33$	$-0.03\pm0.14$	51%
SD	$-0.06\pm0.51$	$-0.02\pm0.33$	$0.06\pm0.30$	$0.00\pm0.31$	$0.06\pm0.48$	$-0.01\pm0.10$	52%

using the average manual registration parameters, as mentioned earlier. Refer to section 10.4.3 for details on the experimental setup.

Table 10.5 lists the statistics of the registration results for two sets of experiments. As the table reveals, the registration parameters are very close to those of the manual results and the differences of these registration parameters are well below the detection threshold of a trained technician. This is expected since the priori was calculated based on the manual registration results. Also revealed by Table 10.5 is that both reversed cross-entropy and symmetric divergence outperformed cross-entropy minimization in terms of success rate, and that symmetric divergence has the highest overall success rate.

Note that the implementation of this minimization process is not yet optimized, e.g., one may use Paren density estimation with a Gaussian kernel to add some smoothness and stability to the similarity functions. Nevertheless, these results indicate that cross-entropy, reversed cross-entropy, and symmetric divergence minimization can yield a very good registration if a good prior estimation is available.

#### **10.5.2.2 Behavior Around True Registration**

The behavior of these three similarity measures was also checked in the vicinity of the manual registration, with the desirable prior calculated from that same manual registration.



Figure 10.11: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the x-rotation changes independently.

Figures 10.11–10.13 show the representative profiles of these similarity measures as a function of rotation angles about the *x*-, *y*-, and *z*-axis, respectively. Figures 10.14–10.16 show the representative profiles of these similarity measures as a function of translation along the *x*-, *y*-, and *z*-axis, respectively. As revealed by all these figures, there are valleys in the profiles around the manual registration parameters, which is one of the necessary conditions for a good registration measure. As expected, the minimum values in all cases are zero since the priori pdf and the estimated pdf at this registration are identical.

#### **10.5.2.3** Numerical Stability for Minimization

In the cross-entropy minimization calculation, if the priori pdf is zero, the crossentropy is not stable. To address this case, one can either ignore the zero priori pdf terms or substitute the zero pdfs with small positive numbers and use the



Figure 10.12: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the y-rotation changes independently.



Figure 10.13: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the z-rotation changes independently.



Figure 10.14: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the x-translation changes independently.

substituted pdfs in the calculation. The small positive pdfs can be calculated in the same fashion as discussed in sections 10.3.6 and 10.5.1.4. Figures 10.17 and 10.18 show the cross-entropy profiles as the *x*-rotation angle and *x*-translation offset change independently. The profiles are labeled in the same manner as for the reversed cross-entropy maximization case. As those two figures reveal, all of the profiles have a similar shape. The only apparent difference is the value of the cross-entropy. Since the size of the basin around the optimal value indicates the size of the capture range, it seems that either using a small  $\alpha$  or ignoring the zero pdf terms is preferred. We elected to ignore the zero pdf terms since the effect of the small  $\alpha$  on the summation of all priori pdfs is hard to quantify (the sum has an upper bound of  $1 + \frac{1}{\alpha}$ ). This strategy worked fine in our experiment.

In the reversed cross-entropy minimization, the calculation is not stable when the joint pdfs are zero. Figures 10.19 and 10.20 show the reversed cross-entropy profiles when the *x*-rotation angle and the *x*-translation change independently


Figure 10.15: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the y-translation changes independently.

and different strategies to handle the zero joint pdfs are compared. The profiles are labeled in the same manner as was used for the cross-entropy minimization case. All of the resultant profiles are similar. Again, we elected to ignore the zero pdf terms since the corresponding profiles are relatively smooth and worked well in practice.

Since the symmetric divergence measure has cross-entropy and reversed cross-entropy components, the issues with numerical stability were addressed in a similar manner.

## **10.6 Concluding Remarks**

We have applied the information-theoretic measures cross-entropy, reversed cross-entropy, and symmetric divergence to retrospective three-dimensional



Figure 10.16: The cross-entropy (CE), reversed cross-entropy (RCE), and symmetric divergence (SD) profiles (with a likely priori) of an MR/Em image pair in the vicinity of a manual registration when the z-translation changes independently.

image registration, which quantify the difference between the true joint distribution and its prior estimation of the voxel value pair in two images at corresponding points. The experimental results indicate that all three measures are suitable as registration criteria.

The cross-entropy, reversed cross-entropy, and symmetric divergence can be maximized and minimized, which is an unusual property for similarity measures. Almost all other similarity measures can be either maximized or minimized, but not both. If a likely priori pdf is given, we want our estimated pdf at registration to be as close to the priori as possible. Therefore, the similarity measures are minimized. If an unlikely priori pdf is given, we want our estimated pdf at registration to be different from the priori as much as possible. Thus, those similarity measures are maximized. There are many unlikely priori pdf. This chapter only considered the most popular one, i.e., the voxel values in two images at



Figure 10.17: The cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-rotation angle changes independently. The numbers around the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdf. The profile that is labelled "ignored" represents the case where the contribution from zero joint pdfs was ignored in the profile computation.

registration are independent of each other. This priori joint pdf is not favorable since the voxel values in two different images reflect the physical properties of the same object and thus they are most likely related in some way, at least from a statistical point of view. If more than one likely and unlikely priori pdf is available, one can define cross-entropy, reversed cross-entropy, and symmetric divergence for them and maximize or minimize them collectively, where a multiobjective optimization approach would be useful, as already illustrated in Zhu [18].

The overall flow chart for the three similarity measures as applied to image registration is presented for both maximization and minimization cases. It shall be noted that this flow chart is also consistent with those of most



Figure 10.18: The cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-translation changes independently. The numbers around the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdf. Ignored means that those pdfs are ignored when computing the profile.

similarity-based registration techniques. Therefore, incorporating those measures to existed packages would be straightforward.

The major problem with applying cross-entropy, reversed cross-entropy, and symmetric divergence optimization to image registration is the numerical stability when the estimated joint pdf or priori joint pdf is zero. If a large sample of successsful registrations are used to generate the desirable, statistically averaged pdf, it is unlikely that one would have many zero points in the priori joint pdf. Alternatively, one can assume a parametric form of the priori joint pdf. The unknown parameters can be estimated from available priori pdf data points and the pdf can be interpolated at the zero points. Unfortunately, there is no justification for one parametric form over others, particularly in the multimodality cases. To improve its accuracy and robustness, one can also use Paren



Figure 10.19: The reversed cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-rotation angle changes independently. The numbers around the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdf. Ignored means that those pdfs are ignored when computing the profile.

density estimation with a Gaussian kernel to add smoothness and stability to the similarity functions, as in Wells *et al.* [14]. Nevertheless, the simple, heuristic approach we used here proves to be effective. The main motivation for this heuristic approach is to make the similarity measure profiles smooth and to have a maximum (peak) or minimum (valley) around the unknown truth with a reasonble capture range. Both of these conditions are necessary for good registration measures. Moreover, it is important that the modified priori or estimated joint pdf do not deviate from a real pdf. That is, the sum of the modified pdfs shall still be close to 1 and the difference from the original values shall not be significant. To further improve the performance of these similarity measures, one can employ different interpolation methods. It was reported recently that different



Figure 10.20: The reversed cross-entropy profiles of an MR/Em image pair in the vicinity of a manual registration when the x-translation changes independently. The numbers around the profiles indicate the  $\alpha$  value used to compute a small probability value which is used to replace zero pdf. Ignored means that those pdfs are ignored when computing the profile.

interpolation methods can attribute to the different levels of smoothness of the similarity measures (see [16]). A smooth similarity function is desirable since the spurious peaks and valleys can trap the optimization to undesirable local optima.

## Acknowledgements

The image files were provided by Dr. M. Hartshorne. Thanks are also due to B. Jones, D. Maniawski, Z. Mercer, and C. Nortmann who manually registered the MR/SPECT image pair used in the similarity minimization study.

# Questions

- 1. What is image registration?
- 2. Why is the algorithm-based registration still required given the success of hybrid imaging devices?
- 3. What is cross-entropy? What is reversed cross-entropy? What is symmetric divergence?
- 4. Why the cross-entropy is a generalized similarity measure for image registration?
- 5. When is cross-entropy (reversed cross-entropy, symmetric divergence) maximized? When is it minimized?
- 6. How can one get a favorable priori pdf?
- 7. What are common unfavorable priori pdf's?
- 8. What are the general steps for a similarity-based registration?
- 9. Why does cross-entropy, reversed cross-entropy, and symmetric divergence based registration suffer from numerical instability? How to deal with it?
- 10. What are the requirements for a good similarity measure?

## Bibliography

- [1] Vera, P., Kaminska, A., Cieuta, C., Hollo, A., Stievenart, J.L., Gardin, I., Ville, D., Mangin, J.F., Plouin, P., Dulac, O. and Chiron, C., Use of subtraction Ictal SPECT co-registration to MRI for optimizing the localization of seizure foci in children, Journal of Nuclear Medicine, Vol. 40, pp. 786–792, 1999.
- [2] Brinkmann, B.H., O'Brien, T.J., Aharon, S., O'Connor, M.K. and Mullan, B.P., Quantitative and clinical analysis of SPECT image registration for epilepsy studies, Journal of Nuclear Medicine, Vol. 40, pp. 1098–1105, 1999.
- [3] Pfluger, T., Vollmar, C., Widmuller, A., Dresel, S., Berger, F., Suntheim, P., Leinsinger, G. and Kahn, K., Quantitative comparison of automatic and interactive methods for MRI-SPECT image registration of the brain based on three-dimensional calculation of error, Journal of Nuclear Medicine, Vol. 41, pp. 1823–1829, 2000.
- [4] Maintz, J.B.A. and Viergever, M.A., A survey of medical image registration, Medical Image Analysis, Vol. 2, No. 1, pp. 1–36, 1998.
- [5] Hajnal, J.V., Hill, D.L.G. and Hawkes, D.J., eds., Medical Image Registration, CRC Press, Boca Raton, FL: 2001.
- [6] Van den Elsen, P.A. and Viergever, M.A., Medical image matching—a review with classification, IEEE Engineering Medical Biology Vol. 12, pp. 26–39, March 1993.
- [7] Maurer, C.R., Jr. and Fitzpatrick, J.M., A review of medical image registration, In: Interactive Image-guided Neurosurgery, Maciunas, R.J. ed., American Association Neurological Surgeons, Parkridge, IL, pp. 17–44, 1993.
- [8] West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer, C.R., Jr., Kessler, R.M., Maciunas, R.J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., Van den Elsen, P.A., Napel, S., Sumanaweera, T.S., Harkness, B., Hemler, P.F., Hill, D.L.G., Hawkes, D.J., Studholme, C., Maintz, J.B.A., Viergever, M.A., Malaandain,

G., Pennec, X., Noz, M.E., Maguire, G.Q., Jr., Pollack, M., Pelizzari, C.A., Robb, R.A., Hanson, D. and Woods, R.P., Comparison and evaluation of retrospective intermodality brain image registration techniques, Journal of Computer Assisted Tomography, Vol. 21, pp. 554–566, 1997.

- [9] Fitzpatrick, J.M., Hill, D.L.G., Shyr, Y., West, J., Studholme, C. and Maurer, C.R., Jr., Visual assessment of the accuracy of retrospective registration of MR and CT images of the brain, IEEE Transactions on Medical Imaging, Vol. 17, No. 4, pp. 571–585, Aug. 1998.
- [10] West, J., Fitzpatrick, J.M., Wang, M.Y., Dawant, B.M., Maurer, C.R., Jr., Kessler, R.M. and Maciunas, R.J., Retrospective intermodality registration techniques for images of the head: Surface-based versus volumebased, IEEE Transactions on Medical Imaging, Vol. 18, No. 2, pp. 144– 150, Feb. 1999.
- [11] Studholme, C., Hill, D.L.G. and Hawkes, D.J., Automatic threedimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Medical Physics, Vol. 24, No. 1, pp. 25–35, Jan. 1997.
- [12] 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 Transactions on Medical Imaging, Vol. 17, No. 4, pp. 586–595, 1998.
- [13] Holden, M., Hill, D.L.G., Denton, E.R.E., Jarosz, L.M., Cox, T.C.S., Rohlfing, T., Goodey, J. and Hawkes, D.J., Voxel similarity measures for 3D serial MR brain image registration, IEEE Transactions on Medical Imaging, Vol. 19, No. 2, pp. 94–102, 2000.
- [14] Wells, W.M., III, Viola, P.A., Atsumi, H., Nakajima, S. and Kikinis, R., Multimodal volume registration by maximization of mutual information, Medical Image Analysis, Vol. 1, No. 1, pp. 35–51, 1996.
- [15] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P., Multimodality image registration by maximization of mutual information, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, April 1997.

- [16] Tsao, J., Interpolation artifacts in multimodality image registration based on maximization of mutual information, IEEE Transactions on Medical Imaging, Vol. 22, No. 7, pp. 854–864, July 2003.
- [17] Pluim, J.P., Maintz, J.B.A. and Viergever, M.A., Mutual informationbased registration of medical images: A survey, IEEE Transactions on Medical Imaging, Vol. 22, No. 8, pp. 986–1004, August 2003.
- [18] Zhu, Y.M., Volume image registration by cross-entropy optimization, IEEE Transactions on Medical Imaging, Vol. 21, No. 2, pp. 174–180, Feb. 2002.
- [19] Shore J.E. and Johnson, R.W., Axiomatic derivation of the principle of maximum entropy and the principle of minimum crossentropy, IEEE Transactions on Information Theory, Vol. 26, No. 1, Jan. 1980.
- [20] Johnson R.W. and Shore, J.E., Comments on and correction to Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, IEEE Transactions on Information Theory, Vol. 29, No. 6, Nov. 1983.
- [21] Shore J.E. and Johnson, R.W., Properties of cross-entropy minimization, IEEE Transactions on Information Theory, Vol. 27, No. 4, July 1981.
- [22] Shore, J.E., Minimum cross-entropy spectral analysis, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 29, No. 2, pp. 230–237, 1981.
- [23] Zhuang, T.G., Zhu, Y.M. and Zhang, X.L., Minimum cross-entropy algorithm (MCEA) for image reconstruction from incomplete projection, SPIE, Vol. 1606, pp. 697–704, 1991.
- [24] Yee, E., Reconstruction of the antibody-affinity distribution from experimental binding data by a minimum cross-entropy procedure, Journal of Theoritical Biology, Vol. 153, No. 2, pp. 205–227, Nov. 1991.
- [25] Alwan, L.C., Ebrahimi, N. and Soofi, E.S., Information theoretic framework for process control, European Journal of Operational Research, Vol. 111, No. 3, pp. 526–542, Dec. 1998.

- [26] Das, N.C., Mazumder, S.K. and De, K., Constrained non-linear programming: A minimum cross-entropy algorithm, Engineering Optimization, Vol. 31, No. 4, pp. 479–487, 1999.
- [27] Antolin, J., Cuchi, J.C., and Angulo, J.C., Minimum cross-entropy estimation of electron pair densities from scattering intensities, Physics Letters A, Vol. 26, pp. 247–252, Sept. 1999.
- [28] Zhu, Y.M. and Wang, Z.Z., Minimum reversed cross-entropy spectral analysis, Chinese Signal Processing, Vol. 8, No. 2, pp. 86–90, 1992.
- [29] Yang, Z.G. and He, Z.Y., A novel method for high resolution spectral analysis—minimum symmetric divergence spectral analysis, Acta Electronica Sinica, Vol. 17, No. 2, 1989.
- [30] Liou, C.Y. and Lin, S.L., The other variant Boltzmann machine, International Joint Conference Neural Networks, Vol. 1, pp. 449–454, 1989.
- [31] Kotz, D., Nuclear medicine in the 21st century: Integration with other specialties, Journal of Nuclear Medicine, Vol. 40, No. 7, pp. 13N–25N, July 1999.
- [32] Tarantola, G., Zito, F. and Gerundini, P., PET instrumentation and reconstruction algorithms in whole-body applications, Journal of Nuclear Medicine, Vol. 44, No. 5, pp. 756–769, May 2003.
- [33] Boes, J.L. and Meyer, C.R., Multivariate mutual information for registration, in Medical Image Computing and Computer-Assisted Intervention—MICCAI'99, Taylor C. and Colchester, A., eds., pp. 606– 612, Springer-Verlag, 1999.
- [34] Leventon, M.E. and Grimson, W.E., Multimodal volume registration using joint intensity distribution, Lecture Notes on Computer Sciences, Vol. 1496, pp. 1057–1066, Oct. 1998.
- [35] Andersson, J.L.R., Sundin, A. and Valind, S., A Method for coregistration of PET and MR brain images, Journal of Nuclear Medicine, Vol. 36, pp. 1307–1315, 1995.

- [36] Zhu, Y.M. and Cochoff, S.M., Likelihood maximization approach to image registration, IEEE Transactions on Image Processing, Vol. 11, No. 12, pp. 1417–1426, Dec. 2002.
- [37] Foley, J.D., Van Dam, A., Feiner, S.K. and Hughes, J.F., Computer Graphics: Principles and Practice, 2nd edn., Addison-Wesley, Reading, MA, 1996.
- [38] Zhu, Y.M. and Cochoff, S.M., Influence of implementation parameters on mutual information maximization image registration of MR/SPECT brain images, Journal of Nuclear Medicine, Vol. 43, No. 2, pp. 160–166, Feb. 2002.
- [39] Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., Numerical Recipes in C: the Art of Scientific Computing, Cambridge University Press, Cambridge, 1999.
- [40] Ritter, N., Owens, R., Cooper, J., Eikelboom, R.H. and Van Saarloos, P.P., Registration of stereo and temporal images of the retina, IEEE Transactions on Medical Imaging, Vol. 18, No. 5, pp. 404–418, May 1999.
- [41] Fitzpatrick, J.M., Hill, D.L.G. and Maurer, C.R., Image Registration In: Handbook of Medical Imaging: Vol. 2, Medical Image Processing and Analysis, Sonka, M. and Fitzpatrick, J.M., eds, SPIE Press, Bellingham, 2000.
- [42] Maes, F., Vandermeulen, D. and Suetens, P. Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information, Medical Image Analysis, Vol. 3, No. 4, pp. 373–386, 1999.
- [43] Pelizzari, C.A., Chen, G.T.Y., Spelbring, D.R., Weichselbaum, R.R. and Chen, C.T., Accurate three-dimensional registration of CT, PET, and/or MR images of the brain, Journal of Computer Assisted Tomography, Vol. 13, No. 1, pp. 20–26, 1989.
- [44] Costa, W.L.S., Haynor, D.R., Haralick, R.M., Lewellen, T.K. and Graham, M.M., A maximum-likelihood approach to pet emission/attenuation image registration, IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 1139–1143, 1993.

- [45] Liebig, J.R., Jones, S.M. and Wang, X., Coregistration of multimodality data in a medical imaging system, US Patent 5,672,877. 1997.
- [46] Dey, D., Slomka, P.J., Hahn, L.J. and Kloiber, R., Automatic threedimensional multimodality registration using radionuclide transmission CT attenuation maps: A phantom study, Journal of Nuclear Medicine, Vol. 40, pp. 448–455, 1999.
- [47] Kiebel, S.J., Ashburner, J., Poline, J.B. and Friston, K.J., MRI and PET coregistration—A cross validation of statistical parametric mapping and automatic image registration, Neuroimage, Vol. 5, pp. 271–279, 1997.

## Chapter 11

## Quo Vadis, Atlas-Based Segmentation?

Torsten Rohlfing,<sup>1</sup> Robert Brandt,<sup>2</sup> Randolf Menzel,<sup>3</sup> Daniel B. Russakoff,<sup>4</sup> and Calvin R. Maurer, Jr.<sup>5</sup>

## **11.1 Segmentation Concepts**

There are many ways to segment an image, that is, to assign a semantic label to each of its pixels or voxels. Different segmentation techniques use different types of image information, prior knowledge about the problem at hand, and internal constraints of the segmented geometry. Which method is the most suitable in any given case depends on the image data, the objects imaged, and the type of desired output information.

Purely intensity-based classification methods [29, 76, 81] work locally, typically one voxel at a time, by clustering the space of voxel values (i.e., image intensities). The clusters are often determined by an unsupervised learning method, for example, *k*-means clustering, or derived from example segmentations [43]. Each cluster is identified with a label, and each voxel is assigned the label of the cluster corresponding to its value. This assignment is independent of the voxel's spatial location. Clustering methods obviously require that the label for each voxel is determined by its value. Extensions of clustering methods that avoid overlapping clusters work on vector-valued data, where each voxel carries a vector of intensity values. Such data is routinely generated by multispectral

<sup>&</sup>lt;sup>1</sup>Neuroscience Program, SRI International, Menlo Park, CA, USA

<sup>&</sup>lt;sup>2</sup>Mercury Computer Systems GmbH, Berlin, Germany

<sup>&</sup>lt;sup>3</sup>Institut für Neurobiologie, Freie Universität Berlin, Berlin, Germany

<sup>&</sup>lt;sup>4</sup>Department of Neurosurgery and Computer Science Department, Stanford University, Stanford, CA, USA

<sup>&</sup>lt;sup>5</sup>Department of Neurosurgery, Stanford University, Stanford, CA, USA

magnetic resonance (MR) imaging, or by a combination of images of the same object acquired from different imaging modalities in general.

There are, however, many applications where there is no well-defined relationship between a voxel's value(s) and the label that should be assigned to it. This observation is fairly obvious when we are seeking to label *anatomical structures* rather than *tissue types*. It is clear, for example, different structures that are composed of the same tissue (e.g., different bones) cannot be distinguished from one another by looking at their intensity values in an image. What distinguishes these structures instead is their location and their spatial relationship to other structures. In such cases, spatial information (e.g., neighborhood relationships) therefore needs to be taken into consideration and included in the segmentation process.

Level set methods [37, 66, 75, 86] simultaneously segment all voxels that belong to a given anatomical structure. Starting from a seed location, a discrete set of labeled voxels is evolved according to image information (e.g., image gradient) and internal constraints (e.g., smoothness of the resulting segmented surface). Snakes or active contours [85] use an analytical description of the segmented geometry rather than a discrete set of voxels. Again, the geometry evolves according to the image information and inherent constraints.

In addition to geometrical constraints, one can take into account neighborhood relationships between several different structures [74, 84]. A complete description of such relationships is an *atlas*. In general, an atlas incorporates the locations and shapes of anatomical structures, and the spatial relationships between them. An atlas can, for example, be generated by manually segmenting a selected image. It can also be obtained by integrating information from multiple segmented images, for example, from different individuals. We shall discuss this situation in more detail in section 11.4.3.

Given an atlas, an image can be segmented by mapping its coordinate space to that of the atlas in an anatomically correct way, a process commonly referred to as registration. Labeling an image by mapping it to an atlas is consequently known as atlas-based segmentation, or registration-based segmentation. The idea is that, given an accurate coordinate mapping from the image to the atlas, the label for each image voxel can be determined by looking up the structure at the corresponding location in the atlas under that mapping. Obviously, computing the coordinate mapping between the image and the atlas is the critical step in any such method. This step will be discussed in some detail in section 11.3. A variety of atlas-based segmentation methods have been described in the literature [3, 11, 12, 15, 16, 21, 23, 24, 38, 41]. The characterizing difference between most of these methods is the registration algorithm that is used to map the image coordinates onto those of the atlas. One important property, however, is shared among all registration methods applied for segmentation: as there are typically substantial shape differences between different individuals, and therefore between an individual and an atlas, the registration must yield a non-rigid transformation capable of describing those inter-subject deformations.

In this chapter we take a closer look at an often neglected aspect of atlasbased segmentation, the selection of the atlas. We give an overview of the different strategies for atlas selection, and demonstrate the influence of the selection method on the accuracy of the final segmentation.

# 11.2 From Subject to Atlas: Image Acquisition and Processing

We illustrate the methods and principles discussed in this chapter by segmenting confocal microscopy images from 20 brains of adult, honeybee workers. Confocal laser scanning microscopy is a type of fluorescence microscopy, where a focused laser beam deflected by a set of *xy*-scanning mirrors excites the fluorescently stained specimen (i.e., the dissected brain). The emitted fluorescence is then recorded by inserting a so-called "confocal pinhole" into the microscope's optical path. This pinhole ensures that only light from the focal plane reaches the detector, thus enabling the formation of an image that can be considered an optical section through the specimen. By moving the position of the specimen along the optical axis of the microscope a three-dimensional (3D) image is generated [8, 69, 88]

The staining of the bee brains depicted in this chapter followed an adapted immunohistological protocol. Dissected and fixated brains were incubated with two primary antibodies (nc46, SYNORF1) that detect synapse proteins [28, 46]. Because cell bodies in insects reside separately from fibers and tracts, this staining ensures response from those regions in the tissue that exhibit high synaptic densities, i.e., neuropil, while somata regions remain mostly unstained. A Cy3-conjugated secondary antibody sensitive to the constant part of the primary antibody was subsequently used to render labeled regions fluorescent. After dehydration and clearing, the specimens were mounted in double-sided custom slides.

The brains were imaged with a confocal laser scanning microscope (Leica TCS 4D). The chromophor was excited with an ArKr laser, and the fluorescence was detected using a longpass filter. The intensity of the fluorescence was quantized with a resolution of 8 bits. Due to the size of the dissected and embedded brain (about  $2.5 \times 1.6$  mm laterally and about 0.8 mm axially), it cannot be imaged in a single scan. Therefore we used multiple image-stack acquisition (3D-MISA) [88]. The entire brain was scanned in  $2 \times 3$  partially overlapping single scans, each using  $512 \times 512$  pixels laterally and between 80 and 120 sections axially. The stacks were combined into a single 3D image using custom software or a script running in Amira (see next paragraph). Because of the refractive index mismatch between the media in the optical path, images exhibit a shortening of distances in axial direction that was accounted for by a linear scaling factor of 1.6 [7].

Post-acquisition image processing was done with the Amira 3D scientific visualization and data analysis package (ZIB, Berlin, Germany; Indeed – Visual Concepts GmbH, Berlin, Germany; TGS Inc., San Diego, CA). Image stacks were resampled laterally to half of the original dimensions in order to increase display speeds and allow interactive handling of the data. The final image volume contained 84–114 slices (sections) with thickness 8  $\mu$ m. Each slice had 610–749 pixels in *x* direction and 379–496 pixels in *y* direction with pixel size 3.8  $\mu$ m. In most cases no further image processing was necessary. In a few cases unsharp masking filters were applied in order to enhance contours.

Subsequently, for each brain an atlas of the neuropil areas of interest was generated by tracing them manually on each slice. We distinguished 22 major compartments, 20 of which are bilaterally symmetric on either brain hemisphere [39]. The paired structures we labeled were medulla, lobula, antennal lobe, ventral mushroom body consisting of peduncle,  $\alpha$ - and  $\beta$ -lobe, and medial and lateral lip, collar and basal ring neuropil. The unpaired structures we identified were the central body with its upper and lower division and the protocerebral lobes including the subesophageal ganglion. Examples of confocal microscopy and label images are shown in Fig. 11.1. Three-dimensional surface renderings of the segmented bee brain are shown in Fig. 11.2. The labeled structures and the abbreviations used for them in this chapter are listed in Table 11.1.



Figure 11.1: Example of bee brain confocal microscopy (*top*) and corresponding label image as defined by manual segmentation (*bottom*). Following radiological convention for axial slices, the image is seen from the cranial direction. Every gray level in the label image represents a different anatomical structure. Due to limitations of reproduction different gray levels may look alike. The correspondence between anatomical structures and abbreviations is listed in Table 11.1. Note that two structures, the left and right antennal lobes (I-AL and r-AL), are not visible in this slice, but can be seen in Fig. 11.2.

# 11.3 Fundamentals of Atlas-Based Segmentation

Mathematically speaking, an atlas A is a mapping  $A : \mathbb{R}^n \to \Lambda$  from *n*-dimensional spatial coordinates to labels from a set of classes  $\Lambda$ . It is



Figure 11.2: Three-dimensional rendering of a segmented bee brain. *From top to bottom*: View from frontal, top, and back, respectively. Note that the two symmetrical blue structures in the lower part of the brain in the frontal view (*top*) are the left and right antennal lobes (I-AL and r-AL) that were not depicted in Fig. 11.1.

Abbreviation	Structure	Abbreviation	Structure
PL-SOG	protocerebral lobes	r-medBR	right medial basal ring
CB	central body	l-latBR	left lateral basal ring
l-Med	left medulla	r-latBR	right lateral basal ring
r-Med	right medulla	l-medColl	left medial collar
l-Lob	left lobula	r-medColl	right medial collar
r-Lob	right lobula	l-latColl	left lateral collar
l-AL	left antennal lobe	r-latColl	right lateral collar
r-AL	right antennal lobe	l-medLip	left medial lip
l-vMB	left mushroom body	r-medLip	right medial lip
r-vMB	right mushroom body	l-latLip	left lateral lip
l-medBR	left medial basal ring	r-latLip	right lateral lip

 Table 11.1:
 Anatomical structures of the bee brain with abbreviations used in this chapter

conceptually very similar to an image in the same coordinate space, which is a mapping from  $\mathbb{R}^n$  to the space of gray values, a subset of  $\mathbb{R}$ . An atlas can therefore itself be considered as a special type of image, that is, a label image. In order to segment a new image R using an atlas A, we need to compute a coordinate mapping between them, that is, we need to register one image to the other. The coordinate mapping must be anatomically correct for the segmentation to be accurate.

An atlas is often generated by (manually) segmenting an actual image, say F. Therefore, we typically have access not only to a spatial map of labels, the actual atlas, but also to a corresponding realization using at least one particular imaging modality. In case multiple co-registered images from different modalities form the basis of the atlas, there may even be multiple instances of actual images. An example of an atlas and a corresponding microscopy image is shown in Fig. 11.1. This dual character is relevant insofar as, while fundamentally possible, registration of an image to the label representation of an atlas is a much harder problem than registration to the corresponding original image.

Let us consider two 3D scalar images,  $R : \mathbb{R}^3 \mapsto \mathbb{R}$  and  $F : \mathbb{R}^3 \mapsto \mathbb{R}$ . We assume that each point in one image has a corresponding equivalent in the other. For any two images, this correspondence is mathematically represented as a coordinate transformation **T** that maps the image coordinates of *R* onto those of *F*. For a given location *x* in the domain of *R*, we find the corresponding location in the domain of *F* as **T**(**x**). If *F* is associated with an atlas *A*, then we can find the correct label for any location **x** in *R* through the mapping

$$\mathbf{x} \mapsto A(\mathbf{T}(\mathbf{x})). \tag{11.1}$$

The transformation **T** is parameterized by a *p*-dimensional parameter vector  $\mathbf{p} \in \mathbb{R}^p$ . The process of finding the vector  $\mathbf{p}$  that describes the "correct" transformation is known as image registration. One of the images, *R*, remains fixed during registration, while the other, *F*, is transformed in space. The fixed image *R* is commonly referred to as the "reference image", the transformed image *F* is called the "floating image".

The terminology used in the remainder of this chapter is as follows. We refer to the already segmented image as the *atlas image* and the image to be segmented as the *raw image*. The coordinates of the raw image are mapped by registration onto those of the atlas image and thereby provide a segmentation of the former. In the context of non-rigid registration, the atlas image is to be deformed while the raw image remains fixed. The correspondence between the common terms for both images in image registration and in the present context is such that the atlas image acts as the *floating image* during registration while the raw image acts as the *reference* (or *target*) *image*.

### 11.3.1 Entropy-Based Image Similarity

It is not usually known a priori, what the correct mapping between the two images R and F is. Instead, the correctness of any given transformation is usually quantified by a so-called similarity measure. This measure is a scalar function  $S : \mathbb{R}^p \mapsto \mathbb{R}$  designed so that higher values of S correspond to better matches. That is, if for two parameter vectors,  $\mathbf{p}_1$  and  $\mathbf{p}_2$ , we have  $S(\mathbf{p}_1) > S(\mathbf{p}_2)$ , then the mapping  $\mathbf{T}_1$  parameterized by  $\mathbf{p}_1$  is assumed to be "more correct" than the mapping  $\mathbf{T}_2$  described by  $\mathbf{p}_2$ . Again, since the correct mapping is not known, Scan only be a more or less suitable approximation to the true correctness. The registration is performed by finding the parameter vector  $\mathbf{p}$  that maximizes S.

A similarity measure that has been empirically found to be particularly wellsuited for many registration applications is mutual information (MI) [36, 77, 80]. It is based on the information-theoretic entropy concept and is defined as

$$S_{\rm MI} = H_R + H_F - H_{RF},$$
 (11.2)

where  $H_R$  is the entropy of image R,  $H_F$  is the entropy of image F, and  $H_{RF}$  is the joint entropy of corresponding voxel pairs between the two images. A

modification proposed by Studholme *et al.* [72], normalized mutual information (NMI), has been found to be slightly more robust

$$S_{\rm NMI} = \frac{H_R + H_F}{H_{RF}}.$$
(11.3)

There are many different implementations of both MI and NMI, using different numerical methods to estimate the image entropies. While some use continuous methods such as Parzen windowing [77, 80], others estimate the entropies from discrete two-dimensional histograms [36, 71]. The latter techniques are more easily implemented and more common.

### 11.3.2 Rigid Registration

The first iteration of intersubject registration, such as registration of an image to an atlas, usually aims at correcting for positioning, orientation, and global size differences between the individual images. Consequently, we initially apply a 9 degree-of-freedom (DOF) affine registration algorithm that performs appropriate translation, rotation, and scaling. A technique described by Studholme *et al.* [71] has been found to produce highly accurate (rigid) transformation in an independent, blinded evaluation study [82]. The algorithm optimizes the NMI similarity measure described above using a simple but robust multiresolution optimization strategy with hierarchically resampled images. For details about our particular implementation of this algorithm the interested reader is referred to [48, 49, 50, 62].

#### 11.3.3 Non-Rigid Registration

There is typically considerable inter-individual variability in the shapes of anatomical structures in the brains of humans and animals. Figure 11.3 illustrates this for the microscopy images of bee brains that we are using to demonstrate the methods in this chapter. For MR images of human brains, Fig. 11.4 provides an analogous illustration. Therefore, in order to be effective, any registration-based segmentation method requires a registration algorithm that can compensate not only for different pose and size, but also for inter-individual shape differences between raw image and atlas (i.e., a non-rigid registration algorithm).

Many different non-rigid registration methods have been published. Some of these, such as methods based on optical flow [73] and most methods using elastic [38, 9] or fluid models [10, 32], typically require both images to be from the



Figure 11.3: Illustration of inter-subject differences between several individual bee brains. *Top:* Central axial slice from a 3D microscopy image used as the reference image for this example. *Second row:* Corresponding slice from three other bee brains after affine registration. *Third row:* Corresponding slices after non-rigid registration. *Fourth row:* Deformed coordinate grids. *Fifth row:* deformation vector fields. Note that only the 2D projection of the 3D deformed coordinate grid and vector field are shown.

same imaging modality to be able to identify corresponding features. Note that the motion model, i.e., fluid or elastic, does not require single modality images. However, most algorithms based in which fluid or elastic differential equations govern the transformation combine these with image similarity terms that are equivalent to the mean squared difference of image intensities.



Figure 11.4: Illustration of intersubject differences between several individual human brains.

Unfortunately, the particular nature of the microscopy images in our example application prohibits the use of any such method. While strictly these images are all generated by the same imaging process, they are subject to imaging artifacts that vary from acquisition to acquisition. Sources of these artifacts include individual concentration differences of the chromophor, fluctuation of laser intensity, and increasing location-dependent absorption with increasing tissue depth.

A non-rigid registration algorithm that inherently supports images originating from multiple imaging modalities was described by Rueckert *et al.* [64]. It uses the same NMI similarity measure as the affine algorithm mentioned above. The transformation model is a free-form deformation [68] **T** that is defined on a data-independent, uniformly spaced control point grid (CPG)  $\Phi$  covering the reference image. The CPG consists of discrete control points  $\phi_{i,j,k}$ , where  $-1 \leq$  $i < n_x - 1, -1 \leq j < n_y - 1$ , and  $-1 \leq k < n_z - 1$ . Points with *i*, *j*, or *k* equal to either 0 or  $n_x - 3$  ( $n_y - 3$  and  $n_z - 3$  for *j* and *k*) are located on the edge of the image data. The spacings between the control points in *x*, *y*, and *z* are denoted by  $\delta_x$ ,  $\delta_y$ , and  $\delta_z$ , respectively. For any location (*x*, *y*, *z*) in the domain of  $\Phi$ , the transformation **T** is computed from the positions of the surrounding  $4 \times 4 \times 4$ control points:

$$\mathbf{T}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_{l}(u) B_{m}(v) B_{n}(w) \phi_{i+l,j+m,k+n}$$

Here, *i*, *j*, and *k* denote the index of the control point cell containing (x, y, z), and *u*, *v*, and *w* are the relative positions of (x, y, z) inside that cell in the three spatial dimensions:

$$i = \left\lfloor \frac{x}{\delta_x} \right\rfloor - 1, \ j = \left\lfloor \frac{y}{\delta_y} \right\rfloor - 1, \ k = \left\lfloor \frac{z}{\delta_z} \right\rfloor - 1,$$

and

$$u = \frac{x}{\delta_x} - \left\lfloor \frac{x}{\delta_x} \right\rfloor, v = \frac{y}{\delta_y} - \left\lfloor \frac{y}{\delta_y} \right\rfloor, w = \frac{z}{\delta_z} - \left\lfloor \frac{z}{\delta_z} \right\rfloor.$$

The functions  $B_0$  through  $B_3$  are the approximating third-order spline polynomials [31]:

$$B_0(t) = (-t^3 + 3t^2 - 3t + 1)/6,$$
  

$$B_1(t) = (3t^3 - 6t^2 + 4)/6,$$
  

$$B_2(t) = (-3t^3 + 3t^2 + 3t + 1)/6,$$
  

$$B_3(t) = t^3/6.$$

The degrees of freedom of a B-spline based transformation **T**, and thus the elements of the parameter vector **p**, are the coordinates of the control points  $\phi_{i,j,k}$ .

The optimum parameters of the non-rigid registration transformation **T** are determined by a line search algorithm similar to the steepest descent method [44]. The target function of the optimization is the NMI similarity of the reference and the transformed floating image. We start by computing a discrete approximation of the gradient of the target function with respect to the parameters of the transformation **T**. This is achieved by a simple finite difference scheme. Despite the high-dimensional parameter space, gradient approximation can be performed very efficiently; due to the compact support of the B-spline functions, each parameter of **T** influences only a small volume in image space (i.e., the local  $4 \times 4 \times 4$  control point neighborhood). When moving any single control point, all voxels of the floating image outside this area remain in the same location. Their contribution to the similarity measure can therefore be precomputed and reused [70].

In order to capture large deformations as well as small ones, the algorithm incorporates a multiresolution deformation strategy based on multilevel B-splines [31]. After finishing optimization at one control point resolution, the spacing between the control points is reduced by a factor of 2 before registration continues. The positions of the control points in the refined grid are determined in a way that exactly preserves the current deformation [18, 58].

Using adaptive grid refinement [55, 67] and a parallel multiprocessor implementation [56], we are able to keep computation times within reasonable bounds. For example, we can complete a non-rigid registration of an image to an atlas, each about the size as described earlier in this chapter, within about 10 minutes on a modern PC (Intel Pentium 4, 3.0 GHz, hyperthreading enabled).

### 11.3.4 Regularization of the Non-Rigid Transformation

Confocal microscopy imaging is a substantially less controlled image formation process than typical medical imaging modalities. Varying concentrations of the chromophor within one structure, laser power fluctuations, tiling artifacts, and absorption of emitted light from deep structures lead to substantial imaging artifacts. As illustrated in Fig. 11.5, these artifacts can cause severe problems for the non-rigid registration, leading to grossly incorrect coordinate transformations. These can, to some extent, be prevented by regularizing the image similarity cost function with an additional constraint term that controls the geometric properties of the coordinate mapping. The total optimization function thus becomes



Figure 11.5: Illustration of the importance of constraining non-rigid registration. These microscopy images are magnified to focus on the area of the right lobula (compare Fig. 11.1 for an anatomical overview). In the reference image (*a*), the lobula appears substantially darker on the lateral side (*ellipse*). In the rigidly registered floating image (*b*) from another individual the lobula has a more homogeneous intensity. Without smoothness constraint, intensity-based non-rigid registration (*c*) computes a grossly incorrect deformation (*arrows*). A constrained non-rigid registration (*d*) does not have this problem.

a weighted sum of the data-dependent image similarity and the regularization constraint term:

$$E_{\text{total}} = (1 - w)E_{\text{NMI}} + wE_{\text{constraint}}.$$
 (11.4)

In detail, we constrain the deformation to be smooth by adding a biharmonic penalty term, which is based on the energy of a thin plate of metal that is subjected to bending deformations [4, 78]. The penalty term is composed of second-order derivatives of the deformation, integrated over the domain D of the transformation **T** as follows:

$$E_{\text{constraint}} = \int_{D} \left( \frac{\partial^{2} \mathbf{T}}{\partial x^{2}} \right)^{2} + \left( \frac{\partial^{2} \mathbf{T}}{\partial y^{2}} \right)^{2} + \left( \frac{\partial^{2} \mathbf{T}}{\partial z^{2}} \right)^{2} + 2 \left[ \left( \frac{\partial^{2} \mathbf{T}}{\partial x \partial y} \right)^{2} + \left( \frac{\partial^{2} \mathbf{T}}{\partial y \partial z} \right)^{2} + \left( \frac{\partial^{2} \mathbf{T}}{\partial z \partial x} \right)^{2} \right] d\mathbf{x}.$$
(11.5)

Since the 3D spline is the tensor product of independent 1D polynomial functions, its second-order derivative with respect to one variable, x, is easily computed as follows:

$$\frac{\partial^2}{\partial x^2} \mathbf{T}(x, y, z) = \frac{1}{\delta_x^2} \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \left( \frac{d^2}{du^2} B_l(u) \right) B_m(v) B_n(w) \phi_{i+l,j+m,k+n}.$$
 (11.6)

Computation of the derivatives of **T** is in fact very similar to computing **T** itself. Depending on the derivation variable, the spline polynomials  $B_0$  through  $B_3$  in the respective dimension are simply replaced by their respective derivatives. These derivatives are easily computed analytically. Mixed second-order derivatives with respect to two different variables are computed by substituting two spline polynomials with their respective first-order derivatives, e.g.,

$$\frac{\partial^2}{\partial x \partial y} \mathbf{T}(x, y, z)$$

$$= \frac{1}{\delta_x \delta_y} \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 \left( \frac{d}{du} B_l(u) \right) \left( \frac{d}{dv} B_m(v) \right) B_n(w) \phi_{i+l,j+m,k+n}. \quad (11.7)$$

Using the above derivative terms, the continuous integral in Eq. (11.5) is approximated as a discretely sampled sum over a set of points, for example, the  $N_D = n_x \times n_y \times n_z$  voxels in the reference image.

## **11.4 Atlas Selection Strategies**

This section will take a closer look at possible choices for atlases in atlas-based segmentation. Usually, this aspect of atlas-based segmentation receives little attention. Yet, the decision about what atlas to use has a substantial impact on the segmentation accuracy, and simple methods are not always the best as we will see below.

We describe and compare here four different atlas-based segmentation strategies with different atlas selections: segmentation with one single individual atlas, segmentation with varying single individual atlases, segmentation with an average shape atlas, and simultaneous segmentation with multiple atlases. These four strategies can be categorized according to the number of atlases used per raw image (one or multiple), the type of atlas used (individual or average), and the assignment of atlases to raw images (fixed, i.e., same atlas(es) for all raw images, or variable, i.e., different atlas image selected for each raw image).



(IND) Segmentation using a single individual atlas.



(SIM) Segmentation using the "most similar" individual atlas.



(AVG) Segmentation using an average shape atlas.



(MUL) Independent segmentation using multiple individual atlases with decision fusion.

Figure 11.6: Depiction of the atlas selection strategies discussed in this chapter. The strategies are also categorized in Table 11.2. Note that the basic atlas-based segmentation with a single atlas (IND, *gray box*) occurs in different stages in the other three strategies, in the MUL case replicated for each atlas.

A schematic graphical comparison of the four methods is given in Fig. 11.6, and a textual summary of the categorization can be found in Table 11.2.

For each strategy, the resulting segmentation accuracy was evaluated. Automatic segmentations were compared to a manual gold standard segmentation. A detailed description of the methods used for validation and accuracy evaluation,

Selection Strategy	No. of Atlases per Raw Image	Type of Atlas	Assignment of Atlas to Raw Image
IND	single	individual	fixed
SIM	single	individual	variable
AVG	single	average	fixed
MUL	multiple	individual	fixed

Table 11.2: Categorization of atlas selection strategies by number, type, and assignment of atlases. See Sections 11.4.1 through 11.4.4 for details and Fig. 11.6 for a schematic overview of the different methods

together with the results for the four atlas selection strategies, are presented in Section 11.5.

## 11.4.1 Segmentation with a Fixed, Single Individual Atlas

The most straight forward strategy for selection of an atlas is to use one individual segmented image. The selection can be random, or based on heuristic criteria such as image quality, lack of artifacts, or normality of the imaged subject. This strategy is by far the most commonly used method for creating and using an atlas [25]. It requires only one atlas, which greatly reduces the preparation effort as compared to the more complex methods described below.

Out of the 20 bee brains in our population, we picked the one that was judged to have the best image quality and least artifacts. We then used this atlas brain to segment the remaining 19 brain images. Each of the 19 raw images was registered non-rigidly to the microscopy image of the atlas, and labeled using the accordingly transformed atlas label image.

### **11.4.2** Segmentation with the Best Atlas for an Image

Suppose that instead of a single atlas, we have several atlases that originate from several different subjects. For each image that we are segmenting, there is one atlas that will produce the best segmentation accuracy among all available atlases. It is obviously desirable to use this optimum atlas, which is most likely a different atlas for each image. The problem is that we do not know what the correct segmentation for an unsegmented image is. Therefore, we can only hope to find a more or less successful heuristic for selecting the best atlas for a given image. There are at least two easily accessible characteristic numbers that describe the similarity between an image and an atlas. One is the final value of the registration criterion, or image similarity measure, after either affine or non-rigid registration. The other is the magnitude of the deformation (i.e., non-rigid transformation) that is required to map the coordinates of the image onto that of the atlas.

Based on these two concepts, we have compared four different criteria for selecting the single atlas that is most likely to produce the best segmentation of a given raw image. These criteria are:

- **NMI affine: Image similarity after affine registration.** The atlas image with the highest NMI similarity to the raw image after *affine* registration is selected and used for its segmentation. This criterion requires only an affine registration to be computed between the raw image and each of the atlases. It is therefore considerably less computationally expensive than the remaining three criteria described below.
- **NMI non-rigid: Image similarity after non-rigid registration.** The atlas with the highest NMI value after *non-rigid* registration is selected and used for segmentation.
- **DEF avg: Average deformation of the atlas over all voxels.** After non-rigid registration, the magnitude of the deformation between the raw image and each individual atlas is computed and averaged over all voxels. The atlas with the smallest average deformation is selected and used for segmentation. Whereas the above criteria are based on intensity similarity, this criterion is based on geometric (i.e., shape) similarity.
- **DEF max: Maximum deformation of the atlas over all voxels.** This criterion is identical to the previous one, except that it uses the maximum deformation over all voxels rather than the average. This criterion pays more attention to outliers. The idea is that atlases that match well overall may be substantially inaccurate in some regions.

Segmentations were generated for each of the 20 bee brains, with the remaining 19 brains as possible atlas candidates in each case. For each raw image, one of



**Criterion for Most Similar Atlas** 

Figure 11.7: Percentages of structures segmented with accuracy better than given SI thresholds using different "most similar" atlas selection criteria. Each column represents one criterion (see text for details). The stacked bars from bottom to top show the percentages of structures that were segmented with SI better than 0.95 through 0.70. For comparison, the left-most column shows the results when the atlas with the best a posteriori SI segmentation result is used for each raw image. This is the upper bound for the accuracy achievable with any criterion for selection of the best single individual atlas.

the atlas candidates was chosen using each of the criteria above. The accuracy of a segmentation was computed as the SI between the segmentation and the manual gold standard.

Figure 11.7 shows a graph of the percentages of structures segmented with varying levels of accuracy. For comparison, this graph includes results achieved when using the best atlas according to the a posteriori SI values for each raw image (left-most column, *Best SI*). In other words, this column shows the best possible result that can be achieved using only a single individual atlas, where the selection of this atlas is governed by the knowledge of the resulting segmentation accuracy (SI value). Obviously, this is not a strategy that is available in practice. However, it provides the upper bound for the segmentation accuracy that can be achieved using any criterion for selection of the best atlas for a given raw image.

Among the four criteria that do not depend on the a posteriori accuracy evaluation and thus a gold standard, the NMI image similarity after non-rigid registration performed slightly better than the other three. It was therefore selected as the criterion used for the SIM atlas selection strategy in the comparison to the other three strategies later in this chapter (section 11.5.4).

### 11.4.3 Segmentation with an Average Shape Atlas

As we stated in the previous chapter, atlas-based segmentation is an easier task if the atlas is similar to the image that is to be segmented. Smaller magnitudes of the deformation between image and atlas that the non-rigid registration algorithm has to determine typically result in a higher accuracy of the matching. If the atlas is itself derived from an individual subject, then the risk is high that this individual is an outlier in the population. In such a case segmenting other subjects using the atlas becomes a more difficult problem. A better atlas would be one that is as similar to as many individuals as possible. Such an atlas can be generated by creating an average over many individuals.

For the human brain, such an average atlas is available from the Montreal Neurological Institute (MNI) as the BrainWeb phantom [5, 13]. Note, however, that the BrainWeb phantom is an atlas of brain *tissue types*, so as we discussed in section 11.1, it is not as useful for atlas-based segmentation as an atlas of brain structures. For the human heart, an average atlas derived from cardiac MR images [45] has been used for atlas-based segmentation [35]. Similarly, an average atlas of the lung has been derived from CT images [33].

For demonstration in this chapter, we have therefore generated an average shape atlas of the structures of the bee brain using a technique outlined below [51].

#### 11.4.3.1 Iterative Shape Averaging

One way of obtaining an average shape atlas from a population of subjects is to generate an active shape model (ASM). In short, an ASM provides a statistical description of a population of subjects by means of an average shape and the principal modes of variation [14, 34]. Generating an ASM typically requires the identification of corresponding landmarks on all individuals, a tedious and errorprone process despite recent success in automating this step using non-rigid registration [20]. Other methods are based entirely on non-rigid registration, such as active deformation models (ADM) [19, 63] and a method described by Guimond *et al.* [22]. Most of these, however, require not only non-rigidly registering several individuals, but also inverting the transformations between them.

In general, non-rigid transformations are not easily inverted. Even for bijective mappings there is typically no closed-form inverse. An iterative method for generating average shape images that does not require inverse computations was first suggested by Ashburner [2]. It also does not require the explicit computation of an average transformation, the definition of which is not trivial for non-linear mappings. Instead, the average deformation is generated by the iterative process itself. This technique was later extended to segmented atlas images and applied to generate an average shape atlas of the bee brain [51].

The central idea is to first map all original individual images onto a common reference and then generate an average image. After that, the original images are mapped onto the average, and a new average image is generated. This process produces a sequence of average images that converges to an average shape image. Note that convergence and average shape in this context are not defined in a strict mathematical sense. However, the result of this iteration is sufficient for the purpose of obtaining an average atlas for atlas-based segmentation, as we will demonstrate below.

In the first step of the iteration, there is not yet an average image to register the original images to. One of the latter is therefore chosen as the reference for the initial registration. In order to avoid bias of the averaging iteration by the choice of the initial reference, the first registration is affine only, thus correcting for pose and size differences but leaving object shape unchanged. For the subsequent steps, the average image resulting from the previous registration step is chosen as the reference image of non-rigid registrations, while the individual images are used as floating images, one at a time. As a result, all floating images are independently mapped into the same reference space, thus enabling the generation of the next average image.

#### **11.4.3.2** Propagation of Transformations

It is well known that intensity-based image registration algorithms, both rigid and non-rigid, fail when started with an initial transformation estimate outside the "capture range" of the desired local optimum of the similarity measure. A wise choice of the initial transformation is therefore beneficial for the robustness of the registration method.

During the iterative averaging process, there are only minor changes in the overall shape of the average brain from one iteration to the next. Consequently, for all images n and all iterations i, the transformation  $\mathbf{T}_n^{(i+1)}$  differs from the preceding  $\mathbf{T}_n^{(i)}$  only by a small additional deformation. A similar situation, although for different reasons, is encountered when registering images from a time series to a common reference; temporally consecutive images typically differ from each other by a smaller amount than they differ from the common reference. In the context of temporal image sequence registration, a framework to incorporate deformations from previous steps into the current registration was recently proposed [58, 59].

For the iterative average image generation described here, we follow a similar approach. Our registration algorithm at each iteration takes as the initial transformation estimate the mapping found during the previous iteration (Fig. 11.8). This is the mapping used to generate the current average image. For the transition from affine to non-rigid registration, incorporation of the previous transformation is achieved by initializing the control point grid with control



**Further Iterations** 

Figure 11.8: Propagation of transformations through the iterative shape averaging algorithm. For each individual image, the transformation (affine or non-rigid) used to generate the current average image is propagated to the next iteration as the initial estimate for the next transformation [reproduced from [51]]. point positions transformed according to the individual affine transformations. For the transition from one non-rigid iteration to the next, the deformation is taken as is and used as the starting point for the optimization.

#### **11.4.3.3** Distance to the Average Shape

Let us recall the rationale behind the creation of our average shape atlas: by minimizing the deformation required to map the atlas onto a given individual, the segmentation accuracy would be improved. So does the atlas produced by the method outlined above in fact minimize this deformation? Indeed, Fig. 11.9 illustrates that the differences between a raw image and an individual atlas are on average substantially larger than the differences between a raw image and the average atlas. Most raw images are more similar in shape to the average shape atlas than to any (or at least the majority) of the remaining 19 individual atlas were the same that built the atlas in the first place, this finding is not too surprising. However, it was important to show that at least for the "training set", our shape averaging does in fact produce a reasonable approximation to the population average shape.



Figure 11.9: Comparison of deformation magnitudes between subjects vs. between a subject and the average shape atlas. The diamonds show the average deformation (over all foreground voxels) in  $\mu$ m when registering the respective raw image to the average shape atlas. The vertical lines show the range of average deformations when registering the respective raw image to the remaining 19 individual atlas images. The boxes show the 25<sup>th</sup> and 75<sup>th</sup> percentiles of the respective distributions.


Figure 11.10: Comparison of individual microscopy images and average shape atlas. Note that the average microscopy image of the average shape atlas also has a better signal-to-noise ratio and generally fewer artifacts than the original individual microscopy images (three randomly selected examples shown as rows 1 through 3 in this figure).

### 11.4.3.4 Noise and Artifacts in the Average Atlas

In Fig. 11.10, we compare some individual images that were used to build the average shape atlas with the average shape atlas itself. It is easy to see that, in addition to representing an average shape, the average atlas also comes with an average microscopy image. The latter is easily generated by averaging the gray values of the appropriately deformed original microscopy images. The average

image shows substantially reduced noise and imaging artifacts. It also shows a more homogeneous distribution of the chromophor compared to the individual brains. All these properties can potentially make registration of the atlas to a given raw image easier, and thus may aid in further improving segmentation accuracy.

## 11.4.4 Multiatlas Segmentation: A Classifier Approach

We can look at an atlas combined with a coordinate mapping from a raw image as a special type of *classifier*. The input of the classifier is a coordinate within the domain of the raw image. The classifier output, determined internally by transforming that coordinate and looking up the label in the atlas at the transformed location, is the label that the classifier assigns to the given raw image coordinate.

As we have briefly mentioned before, using a different atlas leads to a different segmentation of a given raw image. From a classifier perspective, we can therefore say that different atlases generate different classifiers for the same raw image. In the pattern recognition community, it has been well-known for some time that multiple independent classifiers can be combined, and together consistently achieve classification accuracies, which are superior to that of any of the original classifiers [27].

Successful applications of multiple classifier systems have been reported in recognizing handwritten numerals [30, 83] and in speech recognition [1, 65]. In the medical image analysis field, this principle has been applied, for example, to multi-spectral segmentation [42] and to computer-aided diagnosis of breast lesions [17, 47].

The particular beauty of applying a multiclassifier framework to atlas-based segmentation is that multiple independent classifiers arise naturally from the use of multiple atlases. In fact, multiple classifiers also arise from using the same atlas with a different non-rigid registration method. However, adding an additional atlas is typically easier to do than designing an additional image registration algorithm. One could, however, also apply the same basic registration algorithm with a different regularization constraint weight (see section 11.3.4), which would also lead to slightly different segmentations.

For the demonstration in this chapter, we performed a leave-one-out study with only one registration method, but a population of independent atlases. Each of the 20 bee brains was taken as the raw image and automatically segmented using every one of the remaining 19 brains as an atlas. This resulted in 19 segmentations per brain. These 19 segmentations were then combined into a final segmentation.

The most straightforward method for combining multiple classifications into one is the so-called "Vote Rule" decision fusion [26]. For each voxel in the raw image, the outputs of the individual atlas-based classifiers are determined. Their "votes" are then counted, and the label that received that highest number of votes is assigned to the voxel. It is worth, however, to take a closer look at the way an atlas-based classifier works: by looking up a label according to a transformed image coordinate. The label map is discrete, arranged on a 3D grid of labeled voxels. Yet the coordinates of the raw image voxels that we are trying to label hardly ever directly fall on grid points in the atlas. Therefore, looking up the correct label requires some sort of interpolation. The simplest label interpolation method is nearest neighbor (NN) interpolation, resulting in a single unique label per atlas-based classifier. These can easily be combined using vote fusion as described above.

A slightly more complex interpolation technique that can be applied to labels is partial volume interpolation (PVI) as introduced by Maes *et al.* [36]. Here, the labels of all eight neighbors of the interpolated coordinate are determined and weighted with the trilinear interpolation coefficients of their respective grid nodes. Therefore, the output of an atlas-based classifier using PVI is a vector of weights between zero and one, which are assigned to each of the possible labels. One can interpret the weights as the confidence of the classifier in the respective label being the correct answer. These weighted decisions from all classifiers can be combined by so-called "Sum Rule" fusion [26]. The weights for each label are added over all classifiers, and the label with the highest sum is taken as the combined decision.

## **11.5 Quantifying Segmentation Accuracy**

In addition to presenting selected algorithms for atlas-based segmentation, this chapter provides a quantitative comparison among different methods. For each segmentation that we perform, its accuracy is computed. The accuracies achieved for each image are then compared among different methods in order to illustrate quality differences and to identify superior algorithms.

Computing the accuracy of a segmentation requires a gold standard, or ground truth. That is, the correct segmentation needs to be known for an image in order to be able to compute the accuracy of an automatically generated segmentation of that image. While not at all guaranteed to be correct, it is commonly accepted today to use a manual segmentation by a human expert, supported by advanced semi-automatic labeling techniques such as intelligent scissors [40], as the gold standard that automatic segmentation methods are measured against.

### **11.5.1 Similarity Index**

Figure 11.11 provides a visual impression of the segmentation result for two representative slices from one segmented bee brain image. However, in order



Figure 11.11: Example of segmentation using non-rigid image registration (MUL atlas selection paradigm). The two columns show axial images at two different slice locations. *Top row:* Overlays of segmentation contours (shown in white) after non-rigid image registration. *Bottom row:* Difference images between manual and automatic segmentation. Voxels with different labels assigned by manual and automatic segmentation are shown in black.

to effectively compare different segmentation methods, we need to quantify the segmentation accuracy. One possible measure of segmentation quality is the *similarity index* (SI) [87]. For a structure *s*, the SI is computed from the set  $V_{\text{auto}}^{(s)}$  of voxels in *s* according to the automatic segmentation and the set  $V_{\text{manual}}^{(s)}$  of voxels in *s* according to the (gold standard) manual segmentation:

$$SI(s) = \frac{2 \left| V_{\text{manual}}^{(s)} \cap V_{\text{auto}}^{(s)} \right|}{\left| V_{\text{manual}}^{(s)} \right| + \left| V_{\text{auto}}^{(s)} \right|}.$$
(11.8)

For perfect mutual overlap of both segmentations, manual and automatic, the SI has a value of 1. Lesser overlap results in smaller values of SI. No overlap between the segmentations results in an SI value of 0. A major advantage of the SI measure is that it is sensitive to both over-segmentation and under-segmentation, that is, it recognizes both false positives and false negatives among the voxels of a given structure.

## 11.5.2 Bias from Structure Volume

In order to understand the SI values computed later in this chapter and to compare them with other published values, we investigated the dependence of SI values on object size. We performed a numerical simulation in which discretely sampled spheres of various radii were dilated by one or two voxels and the SI values between the original and dilated spheres were computed. The resulting SI values are plotted versus object radius in Fig. 11.12. It is also easy to derive a closed-form expression for the continuous case. The SI between two concentric spheres, one with radius *R* and the other dilated by *d*, i.e., with a radius of R + d, is

$$SI = \frac{2(R/d)^3}{2(R/d)^3 + 3(R/d)^2 + 3(R/d) + 1}.$$
 (11.9)

The SI values for the discrete and continuous cases are almost identical (Fig. 11.12). The SI value between a sphere and a concentric dilated sphere approximates the SI value for a segmentation error consisting of a uniform thickness misclassification on the perimeter of a spherical object. Inspection of Fig. 11.12 and Eq. (11.9) shows that SI depends strongly on object size and is smaller for smaller objects. A one voxel thick misclassification on the perimeter of a spherical object with a radius of 50 voxels has an SI value of 0.97, but for a radius of 10 voxels the SI value is only 0.86. Thus it is not surprising that Dawant



Figure 11.12: Dependence of SI values on size for spherical objects. The squares show SI values computed from discrete numerical simulation of dilation by one voxel. The solid line shows SI values for the continuous case (Eq. 11.9). Note that while the units on the horizontal axis are voxels for the discrete case, they are arbitrary units for the continuous case.

*et al.* [16] reported mean SI values of 0.96 for segmentation of the human brain from MR images and mean SI values of only 0.85 for segmentation of smaller brain structures such as the caudate.

In Fig. 11.13, the average volumes of the anatomical structures in the bee brain images under consideration are shown with the actual segmentation accuracies achieved for them using one of the segmentation methods discussed later (MUL). It is easy to see that the larger a structure, the more accurately it was typically segmented by the atlas-based segmentation. This confirms the theoretical treatment above and illustrates the varying bias of the SI metric when segmenting structures of different sizes.

### **11.5.3 Bias from Structure Shape**

A simple numerical measure that characterizes the shape of a geometrical object is its surface-to-volume ratio (SVR). For a discrete set of labeled voxels in a segmented structure, we can approximate the SVR  $\rho$  as the ratio of the number of surface voxels  $N_s$  to the total number of voxels  $N_t$ , that is,

$$\rho \approx \frac{N_s}{N_t} \tag{11.10}$$



Figure 11.13: Volumes of anatomical structures and corresponding segmentation accuracies. The gray bars show the volumes (in numbers of voxels) of the 22 anatomical structures, averaged over the 20 bee brains. The black vertical lines show the range of SI values achieved by the automatic segmentation (MUL paradigm) over all segmented raw images. The diamond shows the median over all segmented raw images.

A surface voxel is easily defined as one that has a neighbor with a label different from its own. When the entire surface of a structure is misclassified, this can be seen as an erosion of the structure by one voxel. The SI value computed between the original structure and the eroded structure represents the SI resulting from a segmentation that misclassifies exactly all surface voxels. From the structure's SVR  $\rho$  and its total volume *V*, this SI can be computed as

$$SI = \frac{2V(1-\rho)}{V+(1-\rho)V} = \frac{1-\rho}{1-\rho/2}.$$
(11.11)

Similarly, we can estimate the SI resulting from a misclassification of half of all surface voxels. Figure 11.14 shows the SVR values computed for all structures in all brains in our 20 bee brains, plotted versus the SI values of the automatic segmentations. The figure also shows two curves that represent the theoretical misclassification of all and half of all surface voxels, respectively.

For a typical segmentation result of a single structure, a detailed comparison of manual and automatic segmentation is shown in Fig. 11.15. The structure shown here, a right ventral mushroom body, is typical in that its volume and



Figure 11.14: Similarity index vs. surface-to-volume ratio. Each dot represents one structure in one brain (418 structures in total). The average over all individuals for one structure is marked by a  $\times$ . The solid and dashed lines show the theoretical relationship between SVR and SI for misclassification of all and half of all surface voxel, respectively.

its surface-to-volume ratio are close to the respective means over all structures (volume 141k pixels vs. mean 142k pixels; SVR 0.24 vs. mean 0.36). The segmentation accuracy for the segmentation shown was SI = 0.86, which is the median SI value over all structures and all brains.

## 11.5.4 Comparison of Atlas Selection Strategies

The results achieved using the different atlas selection strategies outlined above are visualized in Figs. 11.16–11.19. Each graph shows a plot of the distribution of the SI segmentation accuracies over 19 segmentations, separated by anatomical structure. There were 19 segmentations per strategy as one out of the 20 available bee brain images served as the fixed individual atlas for the IND strategy. Therefore, this brain was not available as the raw image for the remaining strategies, in order to avoid bias of the evaluation.

A comparison of all four strategies is shown in Fig. 11.20. It is easy to see from the latter figure that the widely used IND strategy produced the least accurate



Figure 11.15: A typical segmented structure: right ventral mushroom body (SI = 0.86). *Columns from left to right*: microscopy image, contour from manual segmentation, contour from automatic segmentation (MUL paradigm), and difference image between manual and automatic segmentation. The white pixels in the difference image show where manual and automatic segmentation disagree. *Rows from top to bottom*: axial, sagittal, and coronal slices through the right ventral mushroom body.

results of all strategies. Only slightly better results were achieved by selecting a different individual atlas for each raw image, based on the NMI after non-rigid registration criterion discussed in section 11.4.2. The AVG strategy, segmentation using an average shape atlas, outperformed both the IND and SIM strategies, but was itself clearly outperformed by the MUL strategy. Our results therefore show that the multiclassifier approach to atlas-based segmentation produced substantially more accurate segmentations than the other three strategies. This finding is, in fact, statistically significant when performing a t-test on the SI values for all structures over all segmentations, which confirms the experience of



Figure 11.16: SI by label for segmentation using a single individual atlas (IND atlas selection strategy) [reproduced from [53]].

the pattern recognition community that multiple classifier systems are generally superior to single classifiers [27, 83].

Another interesting finding is that both the AVG and the MUL strategies performed better than the theoretical upper bound of any strategy working with only a single individual atlas (series labeled "Best SI" in Fig. 11.20). We note that "Best SI" is the upper bound not only for any method with the best atlas for each raw image, but also for any possible selection of one atlas for all raw images. Therefore, it is also the upper bound of the IND strategies, which in our study can consequently never outperform the AVG or MUL strategies.



Figure 11.17: SI by label for segmentation using the most similar single individual atlas (SIM atlas selection strategy) [reproduced from [53]].



Figure 11.18: SI by label for segmentation using a single average shape atlas (AVG atlas selection strategy) [reproduced from [53]].

# 11.6 More on Segmentation with Multiple Atlases

We saw in the previous section that a multiclassifier approach to atlas-based segmentation outperforms atlas-based segmentation with a single atlas, be it an individual atlas, an average atlas, or even the best out of a database of atlases. Compared to that, the insight underlying the SIM ("most similar") atlas selection



Figure 11.19: SI by label for segmentation by combining multiple independent atlas-based segmentations (MUL atlas selection strategy) [reproduced from [53]].



Figure 11.20: Percentage of registration-based segmentations with similarity index SI better than the given threshold plotted by atlas selection strategy. The series labeled "Best SI" is the upper bound of all strategies working with a single individual atlas (see text for details).

strategy was that different atlases lead to segmentations of different accuracies. Combined, both observations lead to an even more interesting concept: combination of multiple atlas-based segmentations, weighted by estimates of their individual segmentation accuracy.

In other words, if we had estimates of how well each atlas-based classifier is performing, then we could be more confident in decisions of those classifiers that perform well, compared to the decisions of those that do not. One would hope that by concentrating on more accurate classifiers in the ensemble, the classification accuracy would be further improved.

The performance of each atlas-based classifier is obviously not known in general, due to the lack of a ground truth. However, several methods have been proposed that can estimate the performance parameters, for example, using expectation maximization (EM) methods. Two of these are outlined below, one based on a per-label binary performance model [79], and another based on a simultaneous multilabel performance model [60, 61].

For the description of both methods, we assume that an image with *N* voxels is segmented by *K* different (atlas-based) classifiers. For each voxel *x*, we denote with  $e_k(x)$  the decision by classifier *k*, which is one of the labels assigned in

the segmentation. For the sake of simplicity of the presentation, we assume classifiers that internally use NN interpolation for atlas lookup and therefore only produce one unique label as their output. If the (unknown) ground truth for voxel x is i, we say that x is in class i and write this as  $x \in C_i$ .

#### **11.6.1** A Binary Classifier Performance Model

An EM algorithm described by Warfield *et al.* [79] estimates the classifier performance for each label separately. The method is based on the common performance parameters p (sensitivity) and q (specificity), i.e., the fractions of true positives and true negatives among the classified voxels. The parameters p and q are modeled independently for each classifier k and each class  $C_i$  (label in the segmentation) as the following conditional probabilities:

$$p_i^{(k)} = P(e_k(x) = i | x \in C_i) \text{ and } q_i^{(k)} = P(e_k(x) \neq i | x \notin C_i).$$
 (11.12)

From these definitions, an EM algorithm that estimates p and q from the classifier decisions can be derived as described by Warfield *et al.* [79]. From the computed classifier performance parameters for each label, a contradiction-free final segmentation E at voxel x can be computed as

$$E(x) = \arg\max P(x \in C_i | e_1(x), \dots, e_K(x)).$$
(11.13)

Here, the probability  $P(x \in C_i | \mathbf{e})$  follows from the classifiers' decisions and their performance parameters using Bayes' rule. For details on the application of this algorithm to classifier fusion, see [60].

## 11.6.2 A Multilabel Classifier Performance Model

In a generalization of the Warfield algorithm to multilabel segmentations [60], the classifier parameters p and q are replaced by a matrix of label cross-segmentation coefficients  $\lambda_{i,j}^{(k)}$ . These describe the conditional probabilities that for a voxel x in class  $C_i$  the classifier k assigns label  $j = e_k(x)$ , that is,

$$\lambda_{i,j}^{(k)} = P(e_k(x) = j | x \in C_i).$$
(11.14)

This formulation includes the case that i = j, i.e., the classifier decision for that voxel was correct. Consequently,  $\lambda_{i,i}^{(k)}$  is the usual sensitivity of classifier k for label i. We also note that for each classifier k the matrix  $(\lambda_{i,j}^{(k)})_{i,j}$  is a row-normalized version of the "confusion matrix" [83] in Bayesian multiclassifier algorithms. This matrix, when filled with proper coefficients, expresses prior knowledge about the decisions of each classifier. Again, the coefficients can be estimated iteratively from the classifier decisions by an EM algorithm.

In the "E" step of the EM algorithm, the unknown ground truth segmentation is estimated. Given the current estimate for the classifier parameters ( $\lambda$ ) and the classifier decisions  $e_k(x)$ , the likelihood of voxel x being in class  $C_i$  is

$$W(x \in C_i) = \frac{P(x \in C_i) \prod_k \lambda_{i,e_k(x)}^{(k)}}{\sum_{i'} \left[ P(x \in C_{i'}) \prod_k \lambda_{i',e_k(x)}^{(k)} \right]}.$$
(11.15)

Note that *W* is a function of two parameters, *x* and *i*. The "M" step of our algorithm estimates the classifier parameters ( $\lambda$ ) that maximize the likelihood of the current ground truth estimate determined in the preceding "E" step. Given the previous estimates *W* of the class probabilities, the new estimates for the classifier parameters are computed as follows:

$$\hat{\lambda}_{i,j}^{(k)} = \frac{\sum_{x:e_k(x)=j} W(x \in C_i)}{\sum_x W(x \in C_i)}.$$
(11.16)

## 11.6.3 Results of Performance-Based Multiatlas Segmentation

The accuracy of the performance parameter estimation using both EM algorithms is shown in Fig. 11.21. We computed the actual performance parameters for each atlas-based classifier by comparing its output with the manual segmen-



Figure 11.21: Accuracy of classifier performance parameter estimation using EM algorithms.



**Decision Fusion Method** 

Figure 11.22: Recognition rates (foreground only) of multiple classifier systems based on binary performance model, multilabel performance model, and simple sun fusion.

tation. These a posteriori performances (conceptually equivalent to the *recognition rate* of the classifier) were then plotted versus the estimates computed by either of the EM methods. It is easy to see that there is a very good agreement between the actual and the the estimated parameters, with a slightly higher predictive accuracy following the binary performance model. The Pearson correlation coefficient between true and estimated performances was 0.94 for the binary expert model, and 0.87 for the multilabel expert model. The increased quality of the parameter estimation using the binary performance model can be explained by the substantially larger number of degrees of freedom in the multilabel model, due to the interlabel crosstalk coefficients.

As Fig. 11.22 illustrates, the accuracy of a multiclassifier segmentation can be improved considerably when the performance parameters of the individual classifiers are estimated and taken into account. Overall, the estimation method using a multilabel performance parameter model was slightly less accurate in estimating the actual parameters, but produced a slightly better segmentation accuracy than the method based on a binary performance model.

## 11.7 Conclusion

This chapter has shed light on some often overlooked aspects of atlas-based segmentation methods. We have compared four different strategies for atlas

selection and demonstrated that the accuracy of atlas-based segmentation can be improved substantially by moving beyond the use of a single, individual atlas.

Recently published works on atlas creation and atlas-based segmentation make increasing use of standard atlases that incorporate properties of a population of subjects [33, 35, 45]. Our results confirm that this is likely beneficial for improved segmentation accuracy and robustness. However, our results also suggest that the benefits of applying a multiclassifier strategy are well worth the extra effort.

On a more application-specific note regarding the accuracy of atlas-based segmentation of bee brains, we observe that the mean SI value of segmentations produced using the MUL method in this chapter is 0.86, which, given the small size of most of the structures in the bee brains considered, is comparable to the values reported by Dawant *et al.* and supports the visual assessment observation (Fig. 11.11) that the automatic segmentations described here differ from manual segmentations on average by slightly more than half of the voxels on the structure surfaces (Fig. 11.14). In fact, Zijdenbos *et al.* [87] state that "SI > 0.7 indicates excellent agreement" between two segmentations. This criterion (SI > 0.7) is satisfied by virtually all (97%) contours generated by our segmentations using the MUL method (Fig. 11.19). Furthermore, since the image quality of confocal microscopy images is inferior to clinical MR and CT images in many ways, we believe that our registration-based segmentation method represents a satisfactory intermediate solution to a segmentation problem that is appreciably harder than that of segmenting commonly used images of the human brain.

## Acknowledgments

Torsten Rohlfing was supported by the National Science Foundation under Grant No. EIA-0104114. While performing the research described in this chapter, Robert Brandt was supported by BMBF Grant 0310961. Daniel B. Russakoff was supported by the Interdisciplinary Initiatives Program, which is part of the Bio-X Program at Stanford University, under the grant "Image-Guided Radiosurgery for the Spine and Lungs." All computations were performed on an SGI Origin 3800 supercomputer in the Stanford University Bio-X core facility for Biomedical Computation. The authors would like to thank Andreas Steege and Charlotte Kaps for tracing the microscopy images.

# Questions

- 1. What are the advantages of atlas-based segmentation over other segmentation techniques?
- 2. Why is the described non-rigid registration method superior to other techniques?
- 3. What is the best value for the smoothness constraint weight of the non-rigid registration (section 11.3.4)?
- 4. What if in section 11.4.2 the atlas most similar to the raw image were selected using the following criterion I invented: ...?
- 5. When combining multiple atlas-based segmentations, what is the practical difference between NN interpolation with vote fusion and PVI with sum fusion?
- 6. If the MUL atlas selection strategy is so much better than the others, then why is it not always used?
- 7. How does atlas-based segmentation compare to manual segmentation?
- 8. Are there parallels to the multiatlas segmentation method in pattern recognition?
- 9. Could an active shape model be used as an atlas for segmentation?
- 10. Why does the binary classifier performance model predict actual performance more accurately, yet the multilabel performance model gives better combined classification results?

# Bibliography

- Altincay, H. and Demirekler, M., An information theoretic framework for weight estimation in the combination of probabilistic classifiers for speaker identification, Speech Communication, Vol. 30, No. 4, pp. 255– 272, 2000.
- [2] Ashburner, J., Computational Neuroanatomy, Ph.D. Dissertation, University College London, 2000.
- [3] Baillard, C., Hellier, P. and Barillot, C., Segmentation of brain 3D MR images using level sets and dense registration, Medical Image Analysis, Vol. 5, No. 3, pp. 185–194, 2001.
- [4] Bookstein, F. L., Principal warps: Thin-plate splines and the decomposition of deformations, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 6, pp. 567–585, 1989.
- [5] Online: http://www.bic.mni.mcgill.ca/brainweb/.
- [6] Breiman, L., Bagging predictors, Machine Learning, Vol. 24, No. 2, pp. 123–140, 1996.
- [7] Bucher, D., Scholz, M., Stetter, M., Obermayer, K. and Pflüger, H.-J., Correction methods for three-dimensional reconstructions from confocal images: I. tissue shrinking and axial scaling, Journal of Neuroscience Methods, Vol. 100, pp. 135–143, 2000.
- [8] Cheng, P. C., Lin, T. H., Wu, W. L., and Wu, J. L., eds., Multidimensional Microscopy, Springer-Verlag, New York, 1994.
- [9] Christensen, G. E. and Johnson, H. J., Consistent image registration, IEEE Transactions on Medical Imaging, Vol. 20, No. 7, pp. 568–582, 2001.
- [10] Christensen, G. E., Rabbitt, R. D. and Miller, M. I., Deformable templates using large deformation kinematics, IEEE Transactions on Image Processing, Vol. 5, No. 10, pp. 1435–1447, 1996.

- [11] Collins, D. L. and Evans, A. C., Animal: Validation and applications of nonlinear registration-based segmentation, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 11, No. 8, pp. 1271– 1294, 1997.
- [12] Collins, D. L., Holmes, C. J., Peters, T. M. and Evans, A. C., Automatic 3D model—based neuroanatomical segmentation, Human Brain Mapping, Vol. 3, No. 3, pp. 190–208, 1995.
- [13] Collins, D. L., Zijdenbos, A. P., Kollokian, V., Sled, J. G., Kabani, N. J., Holmes, C. J. and Evans, A. C., Design and construction of a realistic digital brain phantom, IEEE Transactions on Medical Imaging, Vol. 17, No. 3, pp. 463–468, 1998.
- [14] Cootes, T. F., Taylor, C. J., Cooper, D. H. and Graham, J., Active shape models—Their training and application, Computer Vision and Image Understanding, Vol. 61, No. 1, pp. 38–59, 1995.
- [15] Crum, W. R., Scahill, R. I. and Fox, N. C., Automated hippocampal segmentation by regional fluid registration of serial MRI: Validation and application in Alzheimer's Disease, NeuroImage, Vol. 13, No. 5, pp. 847– 855, 2001.
- [16] Dawant, B. M., Hartmann, S. L., Thirion, J. P., Maes, F., Vandermeulen, D. and Demaerel, P., Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and freeform transformations: Part I, methodology and validation on normal subjects, IEEE Transactions on Medical Imaging, Vol. 18, No. 10, pp. 909–916, 1999.
- [17] De Santo, M., Molinara, M., Tortorella, F. and Vento, M., Automatic classification of clustered microcalcifications by a multiple expert system, Pattern Recognition, Vol. 36, No. 7, pp. 1467–1477, 2003.
- [18] Forsey, D. R. and Bartels, R. H., Hierarchical B-spline refinement, ACM SIGGRAPH Computer Graphics, Vol. 22, No. 4, pp. 205–212, 1988.
- [19] Frangi, A. F., Rueckert, D., Schnabel, J. A. and Niessen, W. J., Automatic 3D ASM construction via atlas-based landmarking and volumetric elastic registration, In: Insana, Information Processing in Medical Imaging:

17th International Conference, IPMI 2001, Insana, M. F. and Leahy, R. M., eds., Davis, CA, USA, June 18–22, 2001, Proceedings, Vol. 2082 of Lecture Notes in Computer Science, pp. 78–91, Springer-Verlag, Berlin Heidelberg, 2001.

- [20] Frangi, A. F., Rueckert, D., Schnabel, J. A. and Niessen, W. J., Automatic construction of multiple-object three-dimensional statistical shape models: application to cardiac modeling, IEEE Transactions on Medical Imaging, Vol. 21, No. 9, pp. 1151–1166, 2002.
- [21] Gee, J. C., Reivich, M. and Bajcsy, R., Elastically deforming a threedimensional atlas to match anatomical brain images, Journal of Computer Assisted Tomography, Vol. 17, No. 2, pp. 225–236, 1993.
- [22] Guimond, A., Meunier, J. and Thirion, J.-P., Average brain models: A convergence study, Computer Vision and Image Understanding, Vol. 77, No. 2, pp. 192–210, 2000.
- [23] Hartmann, S. L., Parks, M. H., Martin, P. R. and Dawant, B. M., Automatic 3D segmentation of internal structures of the head in MR images using a combination of similarity and free-form transformations: Part II, validation on severely atrophied brains, IEEE Transactions on Medical Imaging, Vol. 18, No. 10, pp. 917–926, 1999.
- [24] Iosifescu, D. V., Shenton, M. E., Warfield, S. K., Kikinis, R., Dengler, J., Jolesz, F. A. and McCarley, R. W., An automated registration algorithm for measuring MRI subcortical brain structures, NeuroImage, Vol. 6, No. 1, pp. 13–25, 1997.
- [25] Kikinis, R., Shenton, M. E., Iosifescu, D. V., McCarley, R. W., Saiviroonporn, P., Hokama, H. H., Robatino, A., Metcalf, D., Wible, C. G., Portas, C. M., Donnino, R. M. and Jolesz, F. A., A digital brain atlas for surgical planning, model-driven segmentation, and teaching, IEEE Transactions on Visualization and Computer Graphics, Vol. 2, No. 3, pp. 232–241, 1996.
- [26] Kittler, J. and Alkoot, F. M., Sum versus vote fusion in multiple classifier systems, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 1, pp. 110–115, 2003.

- [27] Kittler, J., Hatef, M., Duin, R. P. W. and Matas, J., On combining classifiers, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, pp. 226–239, 1998.
- [28] Klagges, B. R. E., Heimbeck, G., Godenschwege, T. A., Hofbauer, A., Pflugfelder, G. O., Reifegerste, R., Reisch, D., Schaupp, M. and Buchner, E., Invertebrate synapsins: a single gene codes for several isoforms in Drosophila, Journal of Neuroscience, Vol. 16, pp. 3154–3165, 1996.
- [29] Kovacevic, N., Lobaugh, N. J., Bronskill, M. J., B., L., Feinstein, A. and Black, S. E., A robust method for extraction and automatic segmentation of brain images, NeuroImage, Vol. 17, No. 3, pp. 1087–1100, 2002.
- [30] Lam, L. and Suen, C. Y., Optimal combinations of pattern classifiers, Pattern Recognition Letters, Vol. 16, No. 9, pp. 945–954, 1995.
- [31] Lee, S., Wolberg, G. and Shin, S. Y., Scattered data interpolation with multilevel B-splines, IEEE Transactions on Visualization and Computer Graphics, Vol. 3, No. 3, pp. 228–244, 1997.
- [32] Lester, H., Arridge, S. R., Jansons, K. M., Lemieux, L., Hajnal, J. V. and Oatridge, A., Non-linear registration with the variable viscosity fluid algorithm, In: Kuba, Information Processing in Medical Imaging: 16th International Conference, IPMI'99, Visegrád, Hungary, June/July 1999. Proceedings, A., Samal, M. and Todd-Pokvopek, A., eds., Vol. 1613, pp. 238–251, Springer-Verlag, Heidelberg, 1999.
- [33] Li, B., Christensen, G. E., Hoffman, E. A., McLennan, G. and Reinhardt, J. M., Establishing a normative atlas of the human lung: Intersubject warping and registration of volumetric CT images, Academic Radiology, Vol. 10, No. 3, pp. 255–265, 2003.
- [34] Lorenz, C. and Krahnstöver, N., Generation of point-based 3D statistical shape models for anatomical objects, Computer Vision and Image Understanding, Vol. 77, pp. 175–191, 2000.
- [35] Lorenzo-Valdés, M., Sanchez-Ortiz, G. I., Mohiaddin, R. and Rueckert, D., Atlas-based segmentation and tracking of 3D cardiac MR images using non-rigid registration, In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2002: 5th International Conference,

Tokyo, Japan, September 25–28, 2002, Proceedings, Part I, Dohi, T. and Kikinis, R., eds., Vol. 2488 of Lecture Notes in Computer Science, pp. 642–650, Springer-Verlag, Heidelberg, 2002.

- [36] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P., Multimodality image registration by maximisation of mutual information, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 187–198, 1997.
- [37] Malladi, R., Sethian, J. A. and Vemuri, B. C., Shape modelling with front propagation: A level set approach, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 2, pp. 158–175, 1995.
- [38] Miller, M. I., Christensen, G. E., Amit, Y. and Grenander, U., Mathematical textbook of deformable neuroanatomies, Proceedings of the National Academy of Sciences of the U.S.A., Vol. 90, No. 24, pp. 11944– 11948, 1993.
- [39] Mobbs, P. G., Brain structure, in Kerkut, G. A. and Gilbert, L. I., eds., Comprehensive insect physiology biochemistry and pharmacology, Vol. 5: Nervous system: structure and motor function, pp. 299–370, Pergamon Press, Oxford, New York, Toronto, Sydney, Paris, Frankfurt, 1985.
- [40] Mortensen, E. N. and Barrett, W. A., Interactive segmentation with intelligent scissors, Graphical Models and Image Processing, Vol. 60, No. 5, pp. 349–384, 1998.
- [41] Musse, O., Heitz, F. and Armspach, J.-P., Fast deformable matching of 3D images over multiscale nested subspaces. application to atlas-based MRI segmentation, Pattern Recognition, Vol. 36, No. 8, pp. 1881–1899, 2003.
- [42] Paclik, P., Duin, R. P. W., van Kempen, G. M. P. and Kohlus, R., Segmentation of multispectral images using the combined classifier approach, Image and Vision Computing, Vol. 21, No. 6, pp. 473–482, 2003.
- [43] Park, H., Bland, P. H. and Meyer, C. R., Construction of an abdominal probabilistic atlas and its application in segmentation, IEEE Transactions on Medical Imaging, Vol. 22, No. 4, pp. 483–492, 2003.

- [44] Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., Numerical Recipes in C: The Art of Scientific Computing, Cambridge University Press, Cambridge, UK, 2nd ed., 1992.
- [45] Rao, A., Sanchez-Ortiz, G. I., Chandrashekara, R., Lorenzo-Valdés, M., Mohiaddin, R. and Rueckert, D., Construction of a cardiac motion atlas from MR using non-rigid registration, In: Functional Imaging and Modeling of the Heart—Second International Workshop, FIMH 2003, Lyon, France, June 5–6, 2003, Proceedings, Magnin, I. E., Montagnat, J., Clarysse, P., Nenonen, J. and Katila, T., eds., Vol. 2674 of Lecture Notes in Computer Science, pp. 141–150, Springer-Verlag, Heidelberg, 2003.
- [46] Reichmuth, C., Becker, S., Benz, M., Reisch, D., Heimbeck, G., Hofbauer, A., Klagges, B. R. E., Pflugfelder, G. O. and Buchner, E., The sap47 gene of Drosophila melanogaster codes for a novel conserved neuronal protein associated with synaptic terminals, Molecular Brain Research, Vol. 32, pp. 45–54, 1995.
- [47] Rogova, G. L. and Stomper, P. C., Information fusion approach to microcalcification characterization, Information Fusion, Vol. 3, No. 2, pp. 91–102, 2002.
- [48] Rohlfing, T., Multimodale Datenfusion für die bildgesteuerte Neurochirurgie und Strahlentherapie, Ph.D. Dissertation, Technische Universität Berlin, 2000.
- [49] Rohlfing, T., Efficient voxel lookup in non-uniformly spaced images using virtual uniform axes, In: Medical Imaging: Image Processing, Sonka, M. and Hanson, K. M., eds., Vol. 4322 of Proceedings of the SPIE, pp. 986–994, 2001.
- [50] Rohlfing, T., Incremental method for computing the intersection of discretely sampled *m*-dimensional images with *n*-dimensional boundaries, In: Medical Imaging: Image Processing, Sonka, M. and Fitzpatrick, J. M., eds., Vol. 5032 of Proceedings of the SPIE, pp. 1346–1354, 2003.
- [51] Rohlfing, T., Brandt, R., Maurer, Jr., C. R. and Menzel, R., Bee brains, B-splines and computational democracy: Generating an average shape atlas, In: IEEE Workshop on Mathematical Methods in Biomedical

Image Analysis, Staib, L., ed., pp. 187–194, IEEE Computer Society, Los Alamitos, CA, Kauai, HI, 2001.

- [52] Rohlfing, T., Brandt, R., Menzel, R. and Maurer, Jr., C. R., Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains, NeuroImage, Vol. 21, No. 4, pp. 1428–1442, 2004.
- [53] Rohlfing, T., Brandt, R., Menzel, R. and Maurer, Jr., C. R., Segmentation of three-dimensional images using non-rigid registration: Methods and validation with application to confocal microscopy images of bee brains, In: Medical Imaging: Image Processing, Sonka, M. and Fitzpatrick, J. M., eds., Vol. 5032 of Proceedings of the SPIE, pp. 363–374, 2003.
- [54] Rohlfing, T., Maurer, C. R., Bluemke, D. A. and Jacobs, M. A., An alternating-constraints algorithm for volume-preserving non-rigid registration of contrast-enhanced MR breast images, In: Biomedical Image Registration—Second International Workshop, WBIR 2003, Philadelphia, PA, USA, June 23–24, 2003, Gee, J. C., Maintz, J. B. A. and Vannier, M. W., eds., Vol. 2717 of Lecture Notes in Computer Science, pp. 291–300, Springer-Verlag, Berlin Heidelberg, 2003.
- [55] Rohlfing, T. and Maurer, Jr., C. R., Intensity-based non-rigid registration using adaptive multilevel free-form deformation with an incompressibility constraint, In: Proceedings of Fourth International Conference on Medical Image Computing and Computer-Assisted Intervention (MIC-CAI 2001), Niessen, W. and Viergever, M. A., eds., Vol. 2208 of Lecture Notes in Computer Science, pp. 111–119, Springer-Verlag, Berlin, 2001.
- [56] Rohlfing, T. and Maurer, Jr., C. R., Non-rigid image registration in sharedmemory multiprocessor environments with application to brains, breasts, and bees, IEEE Transactions on Information Technology in Biomedicine, Vol. 7, No. 1, pp. 16–25, 2003.
- [57] Rohlfing, T., Maurer, Jr., C. R., Bluemke, D. A. and Jacobs, M. A., Volumepreserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint, IEEE Transactions on Medical Imaging, Vol. 22, No. 6, pp. 730–741, 2003.

- [58] Rohlfing, T., Maurer, Jr., C. R., O'Dell, W. G. and Zhong, J., Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated MR images, In: Medical Imaging: Visualization, Display, and Image-Guided Procedures, Mun, S. K., ed., Vol. 4319 of Proceedings of the SPIE, pp. 337–348, 2001.
- [59] Rohlfing, T., Maurer, Jr., C. R., O'Dell, W. G. and Zhong, J., Modeling liver motion and deformation during the respiratory cycle using intensitybased free-form registration of gated MR images, Medical Physics, 2003 in print.
- [60] Rohlfing, T., Russakoff, D. B. and Maurer, Jr., C. R., Performancebased classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. IEEE Transactions on Medical Imaging, Vol. 23, No. 8, pp. 983–994, 2004.
- [61] Rohlfing, T., Russakoff, D. B. and Maurer, Jr., C. R., Extraction and application of expert priors to combine multiple segmentations of human brain tissue, In: Proceedings of Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Ellis, R. E. and Peters, T. M., eds., Lecture Notes in Computer Science, pp. 587–585, Springer-Verlag, Berlin Heidelberg, 2003.
- [62] Rohlfing, T., West, J. B., Beier, J., Liebig, T., Taschner, C. A. and Thomale, U.-W., Registration of functional and anatomical MRI: Accuracy assessment and application in navigated neurosurgery, Computer Aided Surgery, Vol. 5, No. 6, pp. 414–425, 2000.
- [63] Rueckert, D., Frangi, A. F. and Schnabel, J. A., Automatic construction of 3D statistical deformation models of the brain using nonrigid registration, IEEE Transactions on Medical Imaging, Vol. 22, No. 8, pp. 1014–1025, 2003.
- [64] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O. and Hawkes, D. J., Nonrigid registration using free-form deformations: Application to breast MR images, IEEE Transactions on Medical Imaging, Vol. 18, No. 8, pp. 712–721, 1999.

- [65] Saranli, A. and Demirekler, M., A statistical unified framework for rankbased multiple classifier decision combination, Pattern Recognition, Vol. 34, No. 4, pp. 865–884, 2001.
- [66] Sarti, A., de Solórzano, C. O., Locket, S. and Malladi, R., A geometric model for 3D confocal image analysis, IEEE Transactions on Biomedical Engineering, Vol. 47, No. 12, pp. 1600–1609, 2000.
- [67] Schnabel, J. A., Rueckert, D., Quist, M., Blackall, J. M., Castellano-Smith, A. D., Hartkens, T., Penney, G. P., Hall, W. A., Liu, H., Truwit, C. L., Gerritsen, F. A., Hill, D. L. G. and Hawkes, D. J., A generic framework for non-rigid registration based on non-uniform multi-level freeform deformations, In: Proceedings of IV International Conference on Medical Image Computing and Computer-Assisted Intervention (MIC-CAI 2001), Niessen, W. and Viergever, M. A., eds., Vol. 2208 of Lecture Notes in Computer Science, pp. 573–581, Springer-Verlag, Berlin, 2001.
- [68] Sederberg, T. W. and Parry, S. R., Free-form deformation and solid geometric models, Computer Graphics, Vol. 20, No. 4, pp. 151–160, 1986.
- [69] Stevens, J. K., Mills, L. R. and Trogadis, J. E., eds., Three-Dimensional Confocal Microscopy: Volume Investigation of Biological Specimens, Academic Press, London, 1994.
- [70] Studholme, C., Constable, R. T. and Duncan, J. S., Accurate alignment of functional EPI data to anatomical MRI using a physics-based distortion model, IEEE Transactions on Medical Imaging, Vol. 19, No. 11, pp. 1115– 1127, 2000.
- [71] Studholme, C., Hill, D. L. G. and Hawkes, D. J., Automated threedimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures, Medical Physics, Vol. 24, No. 1, pp. 25–35, 1997.
- [72] Studholme, C., Hill, D. L. G. and Hawkes, D. J., An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition, Vol. 32, No. 1, pp. 71–86, 1999.

- [73] Thirion, J.-P., Image matching as a diffusion process: An analogy with Maxwell's demons, Medical Image Analysis, Vol. 2, No. 3, pp. 243–260, 1998.
- [74] Tsai, A., Wells, W., Tempany, C., Grimson, E. and Willsky, A., Coupled multi-shape model and mutual information for medical image segmentation, in Taylor, C. and Noble, J. A., eds., Information Processing in Medical Imaging, Vol. 2732 of Lecture Notes in Computer Science, pp. 185–197, Springer-Verlag, Berlin Heidelberg, 2003, 18th International Conference, IPMI 2003, Ambleside, UK, July 2003.
- [75] Tsai, A., Yezzi, Jr., A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. E. and Willsky, A., A shape-based approach to the segmentation of medical imagery using level sets, IEEE Transactions on Medical Imaging, Vol. 22, No. 2, pp. 137–154, 2003.
- [76] Vannier, M. W., Pilgram, T. K., Speidel, C. M., Neumann, L. R., Rickman, D. L. and Schertz, L. D., Validation of magnetic resonance imaging (MRI) multispectral tissue classification, Computerized Medical Imaging and Graphics, Vol. 15, No. 4, pp. 217–223, 1991.
- [77] Viola, P. A., Alignment by maximization of mutual information, International Journal of Computer Vision, Vol. 24, No. 2, pp. 137–154, 1997.
- [78] Wahba, G., Spline Models for Observational Data, Vol. 59 of CBMS-NSF Regional Conference Series, SIAM, 1990.
- [79] Warfield, S. K., Zou, K. H. and Wells, W. M., Validation of image segmentation and expert quality with an expectation-maximization algorithm, in Dohi, T. and Kikinis, R., eds., Proceedings of Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Part I, Vol. 2488 of Lecture Notes in Computer Science, pp. 298–306, Springer-Verlag, Berlin Heidelberg, 2002.
- [80] Wells, W. M., Viola, P. A., Atsumi, H., Nakajima, S. and Kikinis, R., Multi-modal volume registration by maximization of mutual information, Medical Image Analysis, Vol. 1, No. 1, pp. 35–51, 1996.

- [81] Wells, III., W. M., Grimson, W. E. L., Kikinis, R. and Jolesz, F. A., Adaptive segmentation of MRI data, IEEE Transactions on Medical Imaging, Vol. 15, No. 4, pp. 429–442, 1996.
- [82] West, J. B., Fitzpatrick, J. M., Wang, M. Y., Dawant, B. M., Maurer, Jr., C. R., Kessler, R. M., Maciunas, R. J., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., van den Elsen, P. A., Napel, S., Sumanaweera, T. S., Harkness, B., Hemler, P. F., Hill, D. L. G., Hawkes, D. J., Studholme, C., Maintz, J. B. A., Viergever, M. A., Malandain, G., Pennec, X., Noz, M. E., Maguire, Jr., G. Q., Pollack, M., Pelizzari, C. A., Robb, R. A., Hanson, D. and Woods, R. P., Comparison and evaluation of retrospective intermodality brain image registration techniques, Journal of Computer Assisted Tomography, Vol. 21, No. 4, pp. 554–566, 1997.
- [83] Xu, L., Krzyzak, A. and Suen, C. Y., Methods of combining multiple classifiers and their applications to handwriting recognition, IEEE Transactions on Systems, Man and Cybernetics, Vol. 22, No. 3, pp. 418–435, 1992.
- [84] Yang, J., Staib, L. H. and Duncan, J. S., Neighbor-constrained segmentation with a 3D deformable model, in Taylor, C. and Noble, J. A., eds., Information Processing in Medical Imaging, Vol. 2732 of Lecture Notes in Computer Science, pp. 198–209, Springer-Verlag, Berlin Heidelberg, 2003, 18th International Conference, IPMI 2003, Ambleside, UK, July 2003.
- [85] Yezzi, Jr., A., Kichenassamy, S., Kumar, A., Olver, P. and Tannenbaum, A., A geometric snake model for segmentation of medical imagery, IEEE Transactions on Medical Imaging, Vol. 16, No. 2, pp. 199–209, 1997.
- [86] Zeng, X., Staib, L. H., Schultz, R. T. and Duncan, J. S., Segmentation and measurement of the cortex from 3-D MR images using coupled-surfaces propagation, IEEE Transactions on Medical Imaging, Vol. 18, No. 10, pp. 927–927, 1999.
- [87] Zijdenbos, A. P., Dawant, B. M., Margolin, R. A. and Palmer, A. C., Morphometric analysis of white matter lesions in MR images: Method and

validation, IEEE Transactions on Medical Imaging, Vol. 13, No. 4, pp. 716–724, 1994.

[88] Zuschratter, W., Steffen, T., Braun, K., Herzog, A., Michaelis, B. and Scheich, H., Acquisition of multiple image stacks with a confocal laser scanning microscope, in Proceedings of Threedimensional and Multidimensional Image Acquisition and Processing V, Vol. 3261, pp. 177–186, Proceedings of SPIE, 1998.

# Chapter 12

# Deformable Image Registration with Hyperelastic Warping

Alexander I. Veress, Nikhil Phatak, and Jeffrey A. Weiss

## **12.1 Introduction**

The extraction of quantitative information regarding growth and deformation from series of image data is of significant importance in many fields of science and medicine. Imaging techniques such as MRI, CT and ultrasound provide a means to examine the morphology and in some cases metabolism of tissues. The registration of this image data between different time points after external loading, treatment, disease or other pathologies is performed using methods known as deformable image registration.

The goal of deformable image registration is to find a transformation that best aligns the features of a "template" and "target" image (Fig. 12.1). In the ideal case, the quantity and quality of the image texture present in the template and target images, as well as the similarity in underlying anatomical structure, would yield a unique "best" transformation. In real problems, however, this is not the case. Deformable image registration is most often ill-posed in the sense of Hadamard [2–3]. No perfect transformation exists, and the solution depends on the choice of the cost function and associated solution methods.

Deformable image registration grew primarily out of the pattern recognition field where significant effort has been devoted to the representation of image ensembles (e.g., [4–13]). The approaches that are used are usually classified as either model-based or pixel-based. Model-based approaches typically require

Department of Bioengineering, University of Utah, Salt Lake City, UT.



Figure 12.1: The canonical deformable image registration problem involves the determination of the deformation map that will align a template image with a target image. In this case, the data are MR images of a heart at different times during the cardiac cycle.

some segmentation of a surface in the 3-D image dataset. This surface is then warped into alignment with features in the target image. The pixel-based approaches do not in general require a segmentation, but rather deform pixels or some sampling of the pixels.

Most methods for deformable registration incorporate a cost function so that the overall energy function to be minimized consists of one term based on the image data itself and a second term that serves to regularize the problem. The choice of this cost function can have a significant effect on the results of image registration. The dependence is most significant in regions of the template model where image texture is sparse or conflicting. In these regions, the registration solution is computed based on minimizing the deformation potential (Bayesian prior probability) portion of the particular registration cost functional [14]. A common approach is to use an analogy to a physical material by treating the original template image as an elastic sheet [12, 13, 15, 16] or a viscous fluid [17]. In general, these approaches benefit from the fact that the mapping from template to target is guaranteed to be one-to-one on the basis of the fundamentals of deformations as defined in continuum mechanics. However, the particular kinematic and constitutive assumptions can over-constrain the solution. As an example, use of the theory of linearized elasticity results in the over-penalization of large rotations, thus limiting the ability to achieve a good registration.

The objective of this chapter is to describe the theory and application of a method termed Hyperelastic Warping [16, 18–22] to problems in deformable image registration. The method is based on the principles of nonlinear solid mechanics to allow objective tracking of large deformations and rotations and the concomitant determination of stresses within the deforming body. The approach may be applied to physical deformations that arise in solid and fluid mechanics as well as to non-physical deformations such as the inter- and intrasubject registration of image data. For the physical deformation case, the goal is to quantify the kinematics and the kinetics of the deformations. In the non-physical case, only the kinematics of the deformations are sought.

# 12.2 Hyperelastic Warping

The standard notation and symbols of modern continuum mechanics are employed in the following presentation [23–25]. In particular, direct notation is used, with boldface italics for vector and tensor fields. The outer product is denoted with " $\otimes$ ", a matrix inner product is denoted with ":", and a matrix-vector product is denoted with ".". Index notation is incorporated for quantities that cannot be readily written in with direct notation. The condensed Voigt notation typically employed in finite element (FE) analysis is utilized as needed [1].

## **12.2.1** Finite Deformation Theory

A Lagrangian reference frame is assumed in the following presentation, and thus the kinematics of material points corresponding to the template image are tracked with respect to their original positions. However, it should be noted that the approach could be adapted readily to an Eulerian framework. The template and target images are assumed to have spatially varying scalar intensity fields defined with respect to the reference configuration and denoted by *T* and *S*, respectively. The deformation map is denoted  $\varphi(X) = x = X + u(X)$  where *x* are the current (deformed) coordinates corresponding to *X* and u(X) is the displacement field. *F* is the deformation gradient [26]:

$$F(X) = \frac{\partial \varphi(X)}{\partial X}.$$
 (12.1)

The local change in density is directly related to the deformation gradient through the Jacobian,  $J := \det(\mathbf{F}) = \rho_0/\rho$ , where  $\det(\mathbf{F})$  is the determinant of the deformation gradient,  $\rho_0$  is the density in the reference configuration and  $\rho$  is the density in the deformed configuration. At this point, it is assumed that T and S have a general dependence on position in the reference configuration X and the deformation map  $\varphi(X)$ .

The positive definite, symmetric right and left Cauchy-Green deformation tensors are, respectively,

$$\boldsymbol{C} = \boldsymbol{F}^T \boldsymbol{F} = \boldsymbol{U}^2 \text{ and } \boldsymbol{B} = \boldsymbol{F} \boldsymbol{F}^T = \boldsymbol{U}^2.$$
 (12.2)

The Jacobian J is defined as:

$$J := \det \boldsymbol{F} = \frac{\rho_0}{\rho}.$$
 (12.3)

## 12.2.2 Variational Framework

Most deformable image registration methods can be posed as the minimization of an energy functional E that consists of two terms. This can be defined with respect to the reference or current (deformed) configuration as:

$$E(\boldsymbol{X}, \boldsymbol{\varphi}) = \int_{\boldsymbol{\beta}_0} W(\boldsymbol{X}, \boldsymbol{\varphi}) dV - \int_{\boldsymbol{\beta}_0} U(T(\boldsymbol{X}, \boldsymbol{\varphi}), S(\boldsymbol{X}, \boldsymbol{\varphi})) dV$$
  
$$= \int_{\boldsymbol{\beta}} W(\boldsymbol{X}, \boldsymbol{\varphi}) \frac{dv}{J} - \int_{\boldsymbol{\beta}} U(T(\boldsymbol{X}, \boldsymbol{\varphi}), S(\boldsymbol{X}, \boldsymbol{\varphi})) \frac{dv}{J}$$
(12.4)

Here, *W* is an energy term that provides regularization and/or some type of constraint on the deformation map (e.g., one-to-one mapping or no negative volumes admitted), while *U* represents an energy that depends on the image data in the template and target images.  $\beta_0$  and  $\beta$  represent the volumes of integration in the reference and current configurations, respectively.

The Euler-Lagrange equations are obtained by taking the first variation of  $E(\mathbf{X}, \boldsymbol{\varphi})$  with respect to the deformation  $\boldsymbol{\varphi}$ . This can be thought of as a "virtual displacement" – a small variation in the current coordinates  $\mathbf{x}$ , denoted  $\varepsilon \eta$ . Here  $\varepsilon$  is an infinitesimal scalar. The first variation of the first energy term in Eq. (2.4) defines the forces per unit volume that arise from the regularization. The second energy term in Eq. (2.4) gives rise to an image-based force term. The first variation of Eq. (2.4) with respect to the deformation  $\boldsymbol{\varphi}(\mathbf{X})$  in direction  $\eta$  is denoted:

$$G(\boldsymbol{\varphi}, \boldsymbol{\eta}) := DE(\boldsymbol{\varphi}) \cdot \boldsymbol{\eta} = \int_{\boldsymbol{\beta}} DW(\boldsymbol{X}, \boldsymbol{\varphi}) \cdot \boldsymbol{\eta} \frac{dv}{J} + \int_{\boldsymbol{\beta}} DU(T(\boldsymbol{X}, \boldsymbol{\varphi}), S(\boldsymbol{X}, \boldsymbol{\varphi})) \cdot \boldsymbol{\eta} \frac{dv}{J} = 0.$$
(12.5)

The variations are calculated by taking the Gateaux derivative [25] of the functional *U* evaluated at  $\varphi + \varepsilon \eta$  with respect to  $\varepsilon$  and then letting  $\varepsilon \to 0$ . For general forms of *W* and *U*,

$$G(\boldsymbol{\varphi},\boldsymbol{\eta}) = \int_{\beta} \frac{\partial W}{\partial \boldsymbol{\varphi}} \cdot \boldsymbol{\eta} \frac{d\nu}{J} + \int_{\beta} \frac{\partial U}{\partial \boldsymbol{\varphi}} \cdot \boldsymbol{\eta} \frac{d\nu}{J} = 0.$$
(12.6)

## 12.2.3 Linearization

Equation (12.6) is highly nonlinear and thus an incremental-interative solution method is necessary to obtain the configuration  $\varphi$  that satisfies the equation [27]. The most common approach is based on linearization of the equations and an iterative solution using Newton's method or some variant. Assuming that the solution at a configuration  $\varphi^*$  is known, a solution is sought at some small increment  $\varphi^* + \Delta u$ . Here again,  $\Delta u$  is a variation in the configuration or a virtual displacement. The linearization of Eq. (12.6) at  $\varphi^*$  in the direction  $\Delta u$  is:

$$L\boldsymbol{\varphi}^{*}G = G(\boldsymbol{\varphi}^{*}, \boldsymbol{\eta}) + DG(\boldsymbol{\varphi}^{*}, \boldsymbol{\eta}) \cdot \Delta \boldsymbol{u} = \int_{\beta} \boldsymbol{\eta} \cdot \left(\frac{\partial W}{\partial \boldsymbol{\varphi}} + \frac{\partial U}{\partial \boldsymbol{\varphi}}\right) \frac{d\boldsymbol{v}}{J} + \int_{\beta} \boldsymbol{\eta} \cdot (\mathbf{D} + \mathbf{k}) \cdot \Delta \boldsymbol{u} \frac{d\boldsymbol{v}}{J},$$
(12.7)

where  $\mathbf{k} := \frac{\partial^2 U}{\partial \varphi \partial \varphi}$  is the *image stiffness* and  $\mathbf{D} := \frac{\partial^2 W}{\partial \varphi \partial \varphi}$  is the *regularization stiffness*. These 2nd derivative terms (Hessians) describe how small perturbations of the current configuration affect the contributions of *W* and *U* to the overall energy of the system.

# 12.2.4 Particular Forms for W and U—Hyperelastic Warping

In Hyperelastic Warping, a physical representation of the template image is deformed into alignment with the target image, which remains fixed in the reference configuration. The scalar intensity field of the template, T, is not changed directly by the deformation, and thus it is represented as T(X). Since the values of S at material points associated with the deforming template change as the template deforms with respect to the target, it is written as  $S(\varphi)$ . The formulation

uses a Gaussian sensor model to describe the image energy density functional:

$$U(\boldsymbol{X}, \boldsymbol{\varphi}) = \frac{\lambda}{2} (T(\boldsymbol{X}) - S(\boldsymbol{\varphi}))^2.$$
(12.8)

 $\lambda$  is a penalty parameter [28] that enforces the alignment of the template model with the target image data. As  $\lambda \to \infty$ ,  $(T(X) - S(\varphi))^2 \to 0$ , and the image energy converges to a finite value.

Hyperelastic Warping assumes that W is the standard strain energy density function from continuum mechanics that defines the material constitutive behavior. It depends on the right deformation tensor C. The right deformation tensor is independent of rotation and thus hyperelasticity provides an objective (invariant under rotation) constitutive framework, in contrast to linearized elasticity (see below, [29]). With these specific assumptions, Eq. (12.4) takes the form:

$$E = \int_{\boldsymbol{\beta}} W(\boldsymbol{X}, \boldsymbol{C}) \frac{d\boldsymbol{\nu}}{J} - \int_{\boldsymbol{\beta}} U(T(\boldsymbol{X}), S(\boldsymbol{\varphi})) \frac{d\boldsymbol{\nu}}{J}$$
(12.9)

The first variation of the first term in Eq. (12.9) yields the standard weak from of the momentum equations for nonlinear solid mechanics (see, e.g., [25]). The first variation of the functional U in Eq. (12.8) with respect to the deformation  $\varphi(\mathbf{X})$  in direction  $\eta$  gives rise to the image-based force term:

$$DU(\boldsymbol{\varphi}) \cdot \boldsymbol{\eta} = D\left[\frac{\lambda}{2}(T(\boldsymbol{X}) - S(\boldsymbol{\varphi}))^2\right] \cdot \boldsymbol{\eta}$$
  
=  $\lambda \left[ (T(\boldsymbol{X}) - S(\boldsymbol{\varphi} + \varepsilon \boldsymbol{\eta})) \frac{\partial}{\partial \varepsilon} (T(\boldsymbol{X}) - S(\boldsymbol{\varphi} + \varepsilon \boldsymbol{\eta})) \right]_{\varepsilon \to 0}$ . (12.10)

Noting that

$$\begin{bmatrix} \frac{\partial}{\partial \varepsilon} (T(X) - S(\varphi + \varepsilon \eta)) \end{bmatrix}_{\varepsilon \to 0} = \begin{bmatrix} -\frac{\partial S(\varphi + \varepsilon \eta)}{\partial (\varphi + \varepsilon \eta)} \cdot \frac{\partial (\varphi + \varepsilon \eta)}{\partial \varepsilon} \end{bmatrix}_{\varepsilon \to 0}$$
(12.11)
$$= -\frac{\partial S(\varphi)}{\partial \varphi} \cdot \eta,$$

Eqs. (12.10) and (12.11) can be combined to yield:

$$DU(\boldsymbol{\varphi}) \cdot \boldsymbol{\eta} = -\lambda \left[ (T(\boldsymbol{X}) - S(\boldsymbol{\varphi})) \frac{\partial S(\boldsymbol{\varphi})}{\partial \boldsymbol{\varphi}} \cdot \boldsymbol{\eta} \right].$$
(12.12)

This term drives the deformation of the template based on the pointwise difference in the image intensities and the gradient of the target intensity evaluated at material points associated with the template. A similar computation for the mechanical strain energy term *W* leads to the weak form of the momentum equations (see, e.g., [24]):

$$G(\boldsymbol{\varphi},\boldsymbol{\eta}) := DE(\boldsymbol{\varphi}) \cdot \boldsymbol{\eta} = \int_{\beta} \boldsymbol{\sigma} : \nabla \boldsymbol{\eta} \, d\boldsymbol{\nu} - \int_{\beta} \lambda \left[ (T-S) \frac{\partial S}{\partial \boldsymbol{\varphi}} \cdot \boldsymbol{\eta} \right] \frac{d\boldsymbol{\nu}}{J} = 0.$$
(12.13)

Here,  $\sigma$  is the 2nd order symmetric Cauchy stress tensor,

$$\boldsymbol{\sigma} = \frac{1}{J} \boldsymbol{F} \frac{\partial W}{\partial \boldsymbol{C}} \boldsymbol{F}^{T}.$$
(12.14)

Thus, the forces applied to the physical model of the deforming template due to the differences in the image data are opposed by internal forces that arise from the deformation of the material through the constitutive model. The particular form of *W* depends on the material and its symmetry (i.e., isotropic, transversely isotropic, etc.) [26, 30–33].

The linearization of Eq. (12.13) yields:

$$L\boldsymbol{\varphi}^{*}G(\boldsymbol{\varphi},\boldsymbol{\eta}) = \int_{\beta}\boldsymbol{\sigma}: \nabla\boldsymbol{\eta}d\boldsymbol{\nu} - \int_{\beta}\lambda \left[ (T-S)\frac{\partial S}{\partial\boldsymbol{\varphi}}\cdot\boldsymbol{\eta} \right] \frac{d\boldsymbol{\nu}}{J} + \int_{\beta}\nabla\boldsymbol{\eta}:\boldsymbol{\sigma}: \boldsymbol{\nabla}(\Delta\boldsymbol{u})d\boldsymbol{\nu} + \int_{\beta}\boldsymbol{\nabla}^{s}\boldsymbol{\eta}: \boldsymbol{\mathsf{C}}: \boldsymbol{\nabla}^{s}(\Delta\boldsymbol{u})d\boldsymbol{\nu} \quad (12.15) + \int_{\beta}\boldsymbol{\eta}\cdot\boldsymbol{\mathsf{k}}\cdot\Delta\boldsymbol{u}\frac{d\boldsymbol{\nu}}{J}$$

Here, **c** is the 4th order spatial elasticity tensor [1]:

$$\boldsymbol{c}_{ijkl} = \frac{4}{J} F_{iI} F_{jJ} F_{kK} F_{lL} \frac{\partial^2 W}{\partial C_{IJ} \partial C_{KL}}, \qquad (12.16)$$

and  $\nabla^{s}[\cdot]$  is the symmetric gradient operator:

$$\boldsymbol{\nabla}^{\boldsymbol{s}}[\boldsymbol{\cdot}] = \frac{1}{2} \left[ \frac{\partial [\boldsymbol{\cdot}]}{\partial \boldsymbol{\varphi}} + \left( \frac{\partial [\boldsymbol{\cdot}]}{\partial \boldsymbol{\varphi}} \right)^{T} \right].$$
(12.17)

In the field of computational mechanics, the first two terms in the second line of Eq. (12.15) are referred to as the *geometric* and *material stiffnesses*, respectively [1]. The 2nd order tensor representing the image stiffness for Hyperelastic Warping is:

$$\mathbf{k} = \frac{\partial^2 U}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}} = \lambda \left[ \left( \frac{\partial S}{\partial \boldsymbol{\varphi}} \right) \otimes \left( \frac{\partial S}{\partial \boldsymbol{\varphi}} \right) - (T - S) \left( \frac{\partial^2 S}{\partial \boldsymbol{\varphi} \partial \boldsymbol{\varphi}} \right) \right].$$
(12.18)

These three terms form the basis for evaluating the relative influence of the image-derived forces and the forces due to internal stresses on the converged solution to the deformable image registration problem, as illustrated in the following two sections.
#### 12.2.5 Finite Element Discretization

Hyperelastic Warping is based on an FE discretization of the template image. The FE method uses "shape functions" to describe the element shape and the arbitrary variations in configuration over the element domain [34]. In Hyperelastic Warping, an FE mesh is constructed to correspond to all or part of the template image (either a rectilinear mesh, or a mesh that conforms to a particular structure of interest in the template image). The template intensity field *T* is interpolated to the nodes of the FE mesh. The template intensity field is convected with the FE mesh and thus the nodal values do not change. As the FE mesh deforms, the values of the target intensity field *S* are queried at the current location of the nodes of the template FE mesh. To apply an FE discretization to Eq. (12.15), an isoparametric conforming FE approximation is introduced for the variations  $\eta$  and  $\Delta u$ :

$$\boldsymbol{\eta}_{e} \equiv \boldsymbol{\eta}|_{\Omega_{e}} = \sum_{j=1}^{N_{\text{nodes}}} N_{j}(\boldsymbol{\xi})\boldsymbol{\eta}_{j}, \quad \Delta \boldsymbol{u}_{e} \equiv \Delta \boldsymbol{u}|_{\Omega_{e}} = \sum_{j=1}^{N_{\text{nodes}}} N_{j}(\boldsymbol{\xi})\Delta \boldsymbol{u}_{j}, \quad (12.19)$$

where the subscript *e* specifies that the variations are restricted to a particular element with domain  $\Omega_e$ , and  $N_{\text{nodes}}$  is the number of nodes composing each element. Here,  $\boldsymbol{\xi} \in \Box$ , where  $\Box := \{(-1, 1) \times (-1, 1) \times (-1, 1)\}$  is the bi-unit cube and  $N_j$  are the isoparametric shape functions (having a value of "1" at their specific node and varying to "0" at every other node). The gradients of the variation  $\boldsymbol{\eta}$  are discretized as

$$\boldsymbol{\nabla}_{s}\boldsymbol{\eta} = \sum_{j=1}^{N_{\text{nodes}}} \boldsymbol{B}_{j}^{L}\boldsymbol{\eta}_{j}, \quad \boldsymbol{\nabla}\boldsymbol{\eta} = \sum_{j=1}^{N_{\text{nodes}}} \boldsymbol{B}_{j}^{NL}\boldsymbol{\eta}_{j}.$$
(12.20)

Where  $\boldsymbol{B}^{L}$  and  $\boldsymbol{B}^{NL}$  are the linear and nonlinear strain-displacement matrices, respectively, in Voigt notation [1]. With the use of appropriate Voigt notation, the linearized Eq. (12.15) can be written, for an assembled FE mesh, as:

$$\sum_{i=1}^{N_{\text{nodes}}} \sum_{j=1}^{N_{\text{nodes}}} (\boldsymbol{K}^{R}(\boldsymbol{\varphi}^{*}) + \boldsymbol{K}^{I}(\boldsymbol{\varphi}^{*}))_{ij} \cdot \Delta \boldsymbol{u}_{j} = \sum_{i=1}^{N_{\text{nodes}}} (\boldsymbol{F}^{\text{ext}}(\boldsymbol{\varphi}^{*}) + \boldsymbol{F}^{\text{int}}(\boldsymbol{\varphi}^{*}))_{i} \quad (12.21)$$

Equation (12.21) is a system of linear algebraic equations. The term in parentheses on the left-hand side is the (symmetric) tangent stiffness matrix.  $\Delta \boldsymbol{u}$  is the vector of unknown incremental nodal displacements – for an FE mesh of 8-noded hexahedral elements in three dimensions,  $\Delta \boldsymbol{u}$  has length [ $8 \times 3 \times N_{el}$ ], Where  $N_{el}$  is the number of elements in the mesh.  $\boldsymbol{F}^{\text{ext}}$  is the vector of external forces arising from the differences in the image intensities and gradients in Eq. (12.12), and  $F^{\text{int}}$  is the vector of internal forces resulting from the stress divergence. The material and geometric stiffnesses combine to give the mechanics regularization stiffness:

$$\boldsymbol{K}^{R} = \int_{\boldsymbol{\beta}} (\boldsymbol{B}^{NL})^{T} \boldsymbol{\sigma} \boldsymbol{B}^{NL} d\boldsymbol{\nu} + \int_{\boldsymbol{\beta}} (\boldsymbol{B}^{L})^{T} \boldsymbol{c} \boldsymbol{B}^{L} d\boldsymbol{\nu}.$$
(12.22)

The contribution of the image-based energy to the tangent stiffness is:

$$\boldsymbol{K}^{I} = -\int_{\beta} \boldsymbol{N}^{T} \boldsymbol{\mathsf{k}} \boldsymbol{N} \frac{d\nu}{\mathbf{J}}.$$
 (12.23)

Together, the terms in Eq. (12.22) and Eq. (12.23) form the entire tangent stiffness matrix. In our FE implementation, an initial estimate of the unknown incremental nodal displacements is obtained by solving Eq. (12.21) for  $\Delta u$  and this solution is improved iteratively using a quasi-Newton method [27].

#### 12.2.6 Solution Procedure and Augmented Lagrangian

In the combined energy function in Eq. (12.9), the image data may be treated as either a soft constraint, with the mechanics providing the "truth", as a hard constraint, with the mechanics providing a regularization, or as a combination. For typical problems in deformable image registration, it is desired to treat the image data as a hard constraint. Indeed, the form for U specified in Eq. (12.8) is essentially a penalty function stating that the template and target image intensity fields must be equal over the domain of interest as  $\lambda \to \infty$ . The main problem with the penalty method is that as the penalty parameter  $\lambda$  is increased, some of the diagonal terms in the stiffness matrix  $\mathbf{K}_I$  become very large with respect to others, leading to numerical ill-conditioning of the matrix. This results in inaccurate estimates for  $\mathbf{K}_I^{-1}$ , which leads to slowed convergence or divergence of the nonlinear iterations.

To circumvent this problem, the augmented Lagrangian method is used [33, 35]. With augmented Lagrangian methods, a solution to the governing equations at a particular computational timestep is first obtained with a relatively small penalty parameter  $\lambda$ . Then the total image-based body forces  $\partial U/\partial \varphi$  are incrementally increased in a second iterative loop, resulting in progressively better satisfaction of the constraint imposed by the image data. This leads to a stable

algorithm that allows the constraint to be satisfied to a user-defined tolerance. Ill conditioning of the stiffness matrix is entirely avoided.

The Euler-Lagrange equations defined in Eq. (12.13) are modified by the addition of a term that represents the additional image-based force  $\gamma$  due to the augmentation:

$$G^* = G(\boldsymbol{\varphi}, \boldsymbol{\eta}) + \int_{\beta} \boldsymbol{\gamma} \cdot \boldsymbol{\eta} \frac{d\nu}{J} = 0$$
 (12.24)

The solution procedure involves incrementally increasing  $\gamma$  at each computational timestep and then iterating using a quasi-Newton method [27] until the energy is minimized. In the context of the FE method described above, the augmented Lagrangian update procedure for timestep n + 1 takes the form:

 $\begin{aligned} \boldsymbol{\gamma}_{n+1}^{0} &= \boldsymbol{\gamma}_{n} \\ k &= 0 \\ \text{DO for each augmentation } k \text{ WHILE } \|(\boldsymbol{\gamma}_{n+1}^{k+1} - \boldsymbol{\gamma}_{n+1}^{k})/\boldsymbol{\gamma}_{n+1}^{k}\| > \text{TOL} \\ \text{Minimize } G^{*} \text{ with } \boldsymbol{\gamma}_{n+1}^{k} \text{ fixed using the BFGS method} \\ \text{Update multipliers using } \boldsymbol{\gamma}_{n+1}^{k+1} &= \boldsymbol{\gamma}_{n+1}^{k} + (\partial U/\partial \boldsymbol{\varphi})_{n+1}^{k} \\ \text{END DO} \end{aligned}$ (12.25)

This nested iteration procedure, referred to as the Uzawa algorithm [36, 37), converges quickly in general because the multipliers  $\gamma$  are fixed during the minimization of  $G^*$ . In practice, the augmentations are not performed until the penalty parameter  $\lambda$  has been incremented to the maximum value that can be obtained without solution difficulties due to ill conditioning. At this last timestep, the augmented Lagrangian method is then used to satisfy the constraint to a user defined tolerance (usually TOL = 0.05).

### 12.2.7 Sequential Spatial Filtering to Overcome Local Minima

The solution approach described above follows the local gradient to search for a minimum in the total energy (Eq. 12.4) and therefore it is susceptible to converging to local minima. This means that the registration process may get "stuck" by alignment of local image features that produce forces locking the deformation into a particular configuration. It is often possible to avoid local minima and converge to a global minimum by first registering larger image features, such as object boundaries and coarse textural detail, followed by registration of fine detail. Sequential low-pass spatial filtering is used to achieve this goal. By evolving the cut-off frequency of the spatial filter over computational time, the influence of fine textural features in the image can be initially suppressed until global registration is achieved. Fine structure can be registered subsequently by gradually removing the spatial filter.

The spatial filter is applied by convolution of the image with a kernel  $\kappa(X)$ . For the template image field *T*,

$$T^*(\mathbf{X}) = T(\mathbf{X}) * \kappa(\mathbf{X}) = \int_B T(\mathbf{X}) \kappa (\mathbf{X} - \mathbf{Z}) d\mathbf{Z}, \qquad (12.26)$$

where T(X) and  $T^*(X)$  are the original image data and the filtered data respectively in the spatial domain; X is a vector containing the material coordinates and Z is the frequency representation of X. An efficient way to accomplish this calculation is through the use of the discrete Fourier transform.

The convolution of the image data  $T(\mathbf{X})$  with the filter kernel  $\kappa(\mathbf{X})$  in Eq. (12.26) becomes multiplication of  $T(\mathbf{Z})$  with  $K(\mathbf{Z})$  in the Fourier domain.  $T(\mathbf{Z})$  is the Fourier transform of  $T(\mathbf{X})$  and  $K(\mathbf{Z})$  is the Fourier transform of  $\kappa(\mathbf{X})$ . This multiplication is applied and then the transform is inverted to obtain the convolved image in the spatial domain as shown below:

$$T^*(\mathbf{X}) = \mathfrak{I}^{-1}\{T(\mathbf{Z}) \operatorname{K}(\mathbf{Z})\}.$$
(12.27)

Because of the very fast computational algorithms available for applying Fourier transforms, this method is much faster than computing the convolution in image space. In our implementation, a 3-D Gaussian kernel is used [38]:

$$\kappa(\boldsymbol{X}) = A \exp\left(-\frac{\boldsymbol{X} \cdot \boldsymbol{X}}{2\sigma^2}\right) \tag{12.28}$$

Here,  $\sigma^2$ , the spatial variance is used to control the extent of blurring while *A* is a normalizing constant. Note that Eq. (12.28) is only valid for a 3-D vector *X*. The user specifies the evolution of the spatial filter over computational time by controlling the mask and variance. In the specific results reported below, the variance was set to a high value and evolved to remove the filtering as the computation was completed (Fig. 12.2).

The practical application of spatial filtering is complicated by the fact that the registration is nonlinear and is computed stepwise during the registration process. At each step in the computational process, the spatial distribution of



Figure 12.2: Sequential spatial filtering. (A) results of a  $10 \times 10$  pixel mask flat blur to suppress the local detail in the original image (D). (B)  $5 \times 5$  mask. (C)  $2 \times 2$  pixel mask, and (D) original image.

the template intensities changes according to the computed deformation field. Therefore, all image operations on the template during the registration process (including spatial filtering techniques) must be performed on the deformed template image, rather than the static template image before deformation. Since, in most cases, the template finite element mesh nodes are not co-located with the template image voxels, the computed deformation field must be interpolated onto the original template image in order to apply the image operations accurately.

#### 12.2.8 Regular Versus Irregular Meshes

Hyperelastic Warping accommodates an FE mesh that corresponds to all or part of the template image. A "regular mesh" is a rectilinear structured mesh that corresponds to the entire image domain. This mesh may be a subsampling of the actual image voxel boundaries. An "irregular" mesh conforms to a particular structure of interest in the template image. The template intensity field T is interpolated to the nodes of the FE mesh. As the FE mesh deforms, the values of the target intensity field S are sampled at the current location of the nodes of the template FE mesh.

Regular meshes are used primarily for nonphysical deformable image problems (Fig. 12.3). Regular meshes are simple to construct and can easily span the entire image space or a specific region of interest. However, since the mesh does not conform to any structure in the template imaged, these analyses are susceptible to element inversion prior to the completion of image registration. Typically, only a single material type is used for the entire mesh.



Figure 12.3: (A) Template and (B) deformed images of a normal mouse brain cross-section with a representation of a regular finite element mesh superimposed upon the image.

In contrast to regular meshes, irregular meshes are used primarily for physical deformation applications and conform to physical structures of interest in the domain of the image data (Fig. 12.4). Irregular meshes also support the definition of different material models and material properties for specific regions of the mesh. For example, in Fig. 12.4, the irregular mesh represents a cross-section



Figure 12.4: A – intravascular ultrasound cross-sectional image of coronary artery. B – Finite element model of Template image. C – Deformed image of artery after application of 100 mmHg internal pressure load. D – Deformed finite element model after Hyperelastic Warping analysis. The grey area of the arterial wall is represents the intima while the dark gray region represents the adventitia.

of a human coronary artery. It has two materials, each representing separate layers of the arterial wall. Each layer was assigned material properties from the literature that are appropriate for that specific layer [39]. The primary drawback of irregular meshes is that, depending upon the geometry to be modeled, they can be time consuming to construct.

#### 12.2.9 Rezoning Regular Finite Element Meshes

The large deformations that occur in the context of many deformable image registration problems can result in "element inversion" prior to complete registration. Element inversion is the generation, via deformation during the solution process, of a finite element that has a negative Jacobian. Physically, for hexahedral elements this implies an angle of greater than 180° between two adjacent edges of an element. This condition halts the solution process and thus must be remedied in order to proceed.

To overcome this problem when regular meshes are used, an FE rezoning algorithm has been implemented. The algorithm allows the tracking of large-scale deformations using a relatively coarse computational mesh. When element inversion is imminent, the FE mesh geometry is reset to its initial undeformed configuration and the deformed template image intensity T and nodal displacements u(X) are interpolated from the deformed mesh to the reset mesh. The analysis then continues until the convergence criteria are met or another rezoning is required. The rezoning process is illustrated graphically in Fig. 12.5.

The rezoning procedures require interpolation of T and u(X) from the nodes of the deformed FE mesh to the nodes of the reset mesh. For each node N



Figure 12.5: Example of rezoning a regular mesh for 2-D Warping problem. (A) Template image with a representation of the FE mesh superimposed on the image. (B) The registration process causes large deformations in the Computational mesh. (C) The mesh is reset and the analysis continues. (D) Rezoning allows for greater overall deformation during the registration process. (E) Deformed template image at the end of the analysis.

in the undeformed mesh, the element in the deformed mesh that contained the node is located using a direct search. The local coordinates of the eight nodes of the element containing node *N* are assembled into an  $8 \times 3$  matrix  $\phi(\xi_i, \eta_i, \zeta_i)$ , where  $\xi_i, \eta_i$ , and  $\zeta_i$  are the local element coordinates of the nodes composing the element; for instance, node 1 has local coordinates (-1,1,1). The local coordinates are related to the global coordinates via the interpolating polynomial coefficients arising from the shape functions as follows [40]:

$$\begin{bmatrix} \phi \end{bmatrix} = \begin{bmatrix} G \end{bmatrix} \begin{bmatrix} \alpha \end{bmatrix} \Leftrightarrow \\ \begin{bmatrix} -1 & 1 & 1 \\ & \ddots & \cdot \\ & & \ddots & 1 \\ & & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & z_1 & xy_1 & yz_1 & xz_1 & xyz_1 \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ 1 & x_8 & y_8 & z_8 & xy_8 & yz_8 & xz_8 & xyx_8 \end{bmatrix} \begin{bmatrix} \alpha_1 & \beta_1 & \gamma \\ & \ddots & \cdot \\ & & \ddots & \ddots \\ \alpha_8 & \beta_8 & \gamma_8 \end{bmatrix}$$
(12.29)

Here,  $\alpha$  is an  $8 \times 3$  matrix containing the polynomial coefficients and  $(x_i, y_i, z_i)$  are the coordinates of node *i* in the global coordinate system. The matrix  $\alpha$  is then determined for each node *N* in the reset mesh:

$$[\alpha] = [G]^{-1}[\phi]. \tag{12.30}$$

The local element coordinates  $(\xi_N, \eta_N, \zeta_N)$  of node *N* follow from  $\alpha$  and the global coordinates  $(x_N, y_N, z_N)$ :

$$[\zeta_{N} \ \eta_{N} \ \xi_{N}] = [1 \ x_{N} \ y_{N} \ z_{N} \ xy_{N} \ yz_{N} \ xz_{N} \ xyz_{N}] \begin{bmatrix} \alpha_{1} \ \beta_{1} \ \gamma \\ \alpha_{2} \ \beta_{2} \ \gamma_{2} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \alpha_{8} \ \beta_{8} \ \gamma_{8} \end{bmatrix}$$
(12.31)

The interpolated value then follows from the local coordinates, the nodal values and the trilinear shape functions. For example, the interpolated template intensity is computed using

$$T_N(\xi_{\rm N}, \eta_{\rm N}, \zeta_{\rm N}) = \sum_{i=1}^8 T_i h_i(\xi_{\rm N}, \eta_{\rm N}, \zeta_{\rm N}), \qquad (12.32)$$

where the  $T_i$  are nodal intensity values and  $h_i$  are the shape functions corresponding to each node evaluated at ( $\xi_N$ ,  $\eta_N$ ,  $\zeta_N$ ). The displacements u(X) are interpolated using the same procedure. Note that this interpolation strategy is consistent with the shape functions used in the FE solution process.

In practice, this rezoning procedure has proved to be highly efficient and effective for large three-dimensional Warping problems. It has allowed for the registration of image data sets that otherwise could not be successfully registered using Hyperelastic Warping. In the first example found below, rezoning allowed for the successful intersubject registration of mouse brain micro-MRI images. Analysis of these image data sets without rezoning led to incomplete registration of the internal structure of the brain as well as incomplete external registration of the cerebellum.

### **12.3 Applications**

The following examples illustrate the broad range of problems that have been analyzed using Hyperelastic Warping. The first example is an image registration problem in which MRI images of two normal mouse brains were registered. The second example illustrates how the results of a registration analysis of micro-CT images of the gerbil middle ear may be used to provide the boundary conditions for a second, traditional, FE analysis of the malleus bone. The remaining examples illustrate applications in cardiovascular mechanics.

### 12.3.1 Quantification of Changes in Mouse Brain Morphology

Quantification of time-dependent changes in three-dimensional morphology of brain structures and neural pathways is a fundamental challenge in anatomical studies of neurodevelopment and in tracking brain remodeling and/or progression of certain neurological diseases. The morphometric problem can be approached using *in vivo* gross-scale (submillimeter) magnetic resonance medical imaging (MRI) of the brain. Tracking anatomical changes *in vivo* has been a major motivation for the development of higher resolution CT, MRI and radiographic imaging systems. While it is currently routine in clinical MRI of humans to obtain  $1 \times 1 \times 2$  mm resolution, micro-MRI images of small animals have been obtained with isotropic resolution on the order of 40 µm resolution sometimes termed *magnetic resonance microscopy* (MRM). This type of MRI data is sufficient to resolve the neuroanatomical structures of interest but it remains difficult to extract quantitative structure-specific morphological measures

directly from this type of image data. These measures are necessary to accurately assess developmental and/or pathological changes in gross brain structures and pathways.

In order to test the efficacy of Hyperelastic Warping in the registration of normal mouse brain anatomy, normal T<sub>1</sub>-weighted micro-MRI images were obtained from two different, intact, excised mouse brains. The image datasets were  $256^3$  voxels, FOV =  $1.54 \times 1.54 \times 1.5$  cm, and had 60 µm isotropic resolution. A  $40 \times 40 \times 49$  rectilinear FE mesh was created for the 3D problems (73,008 elements). The deforming template was modeled using a neo-Hookean hyperelastic material with a shear modulus of 450 Pa and a bulk modulus of 400 Pa [22].

These 3-D results demonstrate the efficacy of Hyperelastic Warping when used on relatively large datasets. Volume-rendered images (Figs. 12.6) show that excellent external registration was achieved between the deformed template and target image a datasets. The 3-D model was rezoned three times to achieve this registration. It is interesting to note that rezoning allowed a dissection artifact in the target image dataset (Fig. 12.6C), that was not present in the template image data (Fig. 12.6A), to be extruded from the relatively smooth template to generate the same structure in the deformed template (Fig. 12.6B). Without the use of rezoning, this excellent alignment would have been impossible due to extreme mesh distortion resulting in element inversion. Examination of representative transverse and longitudinal image planes illustrated that very good internal registration was also achieved, as demonstrated by the correspondence of anatomical regions and sulci between the deformed template and target (Fig. 12.7. panels A–D).

Computational requirements for this problem were determined primarily by the size of the finite element mesh used to discretize the template and, to a lesser extent, by the size of the image datasets. The analysis required 3.38 GB of memory. Because the main computational expense in the algorithm is the inversion of a large system of linear equation resulting from the nodal degrees of freedom in the FE mesh, CPU requirements grew as the square of the size of the FE mesh. For this analysis, the mesh resulted in a linear system with 165,148 degrees of freedom. Total wall clock time for the 3-D neuroanatomical registration analysis was 14 hours, with the three mesh rezones accounting for 18% of the analysis time and the sequential spatial filtering accounting for 5% of the analysis time. The vast majority of the remaining analysis time is



Figure 12.6: Three dimensional results for inter-animal registration of two mouse neuroanatomies. (A) Surface-rendered template, (B) deformed template, and (C) target image. The arrow indicates the dissection artifact.

spent in the repeated inversion of the sparse symmetric linear system. Our code accommodates the use of several vendor-supplied parallel solvers, which can reduce the time for this phase of the solution process drastically. The analysis time can be further reduced by the reduction of the size of the computational mesh, at the potential expense of reducing the accuracy of the registration of internal structures.



Figure 12.7: A mid-brain cross-section from a normal mouse (A) and a logitudinal section (B) from the 3-D target image data and the corresponding deformed template (C and D).

### 12.3.2 Measurement of Gerbil Malleus Kinematics and Mechanics

The human auditory system is capable of transforming and distinguishing incoming acoustical signals over several orders of magnitude. The middle ear, in particular, acts as an impedance-matching transformer, allowing the mechanical vibrations of the tympanic membrane to be transformed into liquid-borne traveling waves within the cochlea. These traveling waves are in turn transformed into neural signals that the brain interprets as sound.

Finite element models have been used to study the kinematics of the middle ear bones in order to gain a better understanding of the impedance-matching function of the middle ear [41–44]. These models consist of 2-D and 3-D finite element representations of the individual bones and muscles of the middle ear as well as the tympanic membrane. The natural frequencies of the eardrum have been measured and used to excite finite element representations of the tympanic membrane to study the frequency response and kinematics of the middle ear bones [41, 42]. Ladak and Funnell [45] modeled the normal and surgically repaired cat middle ear in order to study the effects of ossicular prosthetics on the frequency response of the ossicular chain. While direct measurements of the geometry and kinematics of the tympanic membrane have been performed, measurements of the kinematics of the middle ear bones themselves have proven

to be more difficult. Toward this end, the following study was designed to examine the feasibility of using Hyperelastic Warping to determine the displacements of the ossicular chain from high-resolution CT images. These displacements would in turn provide the boundary conditions for FE models of the individual bones of the ossicular chain. This secondary analysis would be used to determine the stress distributions within bones of the middle ear. High-resolution computed tomography (CT) images  $(1024 \times 1024 \times 1024 \text{ isotropic image ma-}$ trix, 14.1 mm FOV, 10 µm isotropic resolution) were taken of the external and middle ear of an anesthetized gerbil. The images were acquired on a Skyscan 1072 80 kV micro-CT tomograph. An image data set was acquired with the tympanic membrane under no external pressure load other than atmospheric pressure. The second image set was acquired while a 3.0 kPa pressure load was placed on the external surface of the tympanic membrane. The images were cropped  $(270 \times 270 \times 172 \text{ voxels})$  to include only the tympanic membrane and the malleus bone of the middle ear. The image obtained under atmospheric loading was defined as the template image while the image under a pressure load of 3 kPa was defined as the target image. A  $41 \times 41 \times 27$  rectilinear finite element mesh was constructed that included the entire cropped image domain (11,767 elements). This deforming template mesh was modeled as a neo-Hookean hyperelastic material with a shear modulus of 450 Pa and a bulk modulus of 400 Pa. A fixed flat spatial filter  $(3 \times 3 \times 3 \text{ pixel mask})$  [38] was used in the warping analysis. The FE mesh was record twice during the analysis to determine the displacements of the malleus.

Subsequent to the deformable registration analysis, a finite element model was created to represent the malleus bone. The external boundary of the malleus was manually segmented from the *template* image data set. B-spline curves were fitted to the points generated by the segmentation and these curves were used to define the exterior surface of the malleus. A tetrahedral mesh (42,391 elements) was generated from this surface definition. The malleus was modeled as a linear elastic material using properties (elastic modulus E = 20.0 GPa, Poisson's Ratio,  $\nu = 0.3$ ) from the literature [41, 45].

The surface of the malleus model was loaded using the displacements determined from the deformable image registration analysis. The displacements for each surface node of the malleus model were defined by interpolating nodal displacements determined from the warping analysis using the rectilinear (Warping) mesh trilinear shape functions. The NIKE3-D nonlinear finite element program [46] was used to analyze the malleus model and determine the



Figure 12.8: (A) Rendered surface definition of the gerbil malleus. (B) Displacement magnitude warping results for a plane bisecting the center of the malleus. The tetrahedral mesh has been superimposed on the results to indicate the location of the malleus within the displacement field. (C) Effective stress and (D) displacement magnitude results for the surface of the malleus.

stress/strain distribution within the bone using only the surface displacements as the boundary conditions.

The results indicate that the manubrium, which is at the center of tympanic membrane, undergoes the greatest displacement and is a high stress region of the malleus (Figs. 12.8C and 12.8D). In contrast, the head of the malleus, which has attachments to the head of the incus and the superior ligament, shows the least displacement and is a low stress region. These results suggest that the malleus acts to decrease the energy being transferred to the incus. Further, this analysis demonstrates how the deformation map from a deformable image registration analysis using Hyperelastic Warping can be integrated into a traditional computational biomechanics analysis using the FE method.

### 12.3.3 Strain Measurement of the Coronary Artery using Intravascular Ultrasound

Coronary heart disease is currently the leading cause of death in the United States [47]. Plaque rupture, the structural failure of the plaque cap, is the primary

event triggering myocardial infarctions and acute coronary syndromes. The failure of the cap exposes collagen and lipid to the blood stream, which subsequently causes thrombus formation [48], often resulting in partial or complete blockage of the vessel. The exact mechanisms responsible for plaque rupture are unknown.

Finite element analyses of idealized plaque geometries have suggested that, for eccentric plaques, maximum stress levels occur at the shoulder area of the cap where the fibrous cap meets the healthy intima [49, 50]. Finite element analyses using model geometries based on atherosclerotic lesions indicate that the areas of high stress in and near the plaque correlate with the locations of plaque rupture. Fifty-eight per cent of *in vivo* plaque ruptures have been found to occur in the areas of maximum stress, while 83% of failures occurred in high stress areas [51]. FE studies have suggested that decreased cap thickness causes an increase in the peak shoulder stress when fully developed lipid layers are present. Similarly, increasing the lipid layer size increases the shoulder stress. [52–54].

Reliable predictions of stress and strain in physiologically loaded plaques *in vivo* would provide insight into plaque mechanics. Direct measurement of stress during loading of a coronary artery is currently not possible *in vivo* or *ex vivo*. However, the measurement of strain within the plaque and the wall of the coronary artery can provide an insight into the stress distribution.

Intravascular ultrasound (IVUS) yields detailed images of atherosclerotic plaques and the vessel wall. IVUS uses a catheter-mounted ultrasound transducer to acquire cross-sectional images of an artery with a spatial resolution of 80–100  $\mu$ m radially and 150–200  $\mu$ m circumferentially [55, 56]. Current IVUS catheters are as small as 0.9 mm and can interrogate most areas of the coronary tree, including coronary arteries in the range of 1.5–5.0 mm in diameter. IVUS provides a high resolution means to quantify lesion geometry [55, 56]. Our long-term goal is to use Hyperelastic Warping to determine the strain distributions within coronary plaque both *ex vivo* and *in vitro* during physiological loading as well as the loading associated with interventional techniques such as angioplasty and stent placement. The strain distributions can be correlated with the plaque histology to determine which plaque cap components are associated for use with IVUS and the details may be found in our previous publication [39].

Hyperelastic Warping was used to estimate the strain distributions in two unfixed left anterior descending (LAD) human coronary arteries. These arteries were mounted in a position approximating the artery orientation *in situ*. The left main coronary artery was cannulated, and the side branches were ligated to reduce flow until a constant physiological perfusion pressure could be maintained. IVUS images were acquired using a clinical IVUS system, comprising an HP Sonos 100 ultrasound console and a 30 MHz, 3.5 F Boston, scientific monorail intracoronary ultrasound imaging catheter using parameters typical for clinical study. The IVUS catheter was inserted into the vessel as halfway down the LAD. The arterial internal pressure monitored using a Millar 4 F pressure transducer introduced through a distal cannula placed approximately adjacent to the IVUS catheter. The vessel was then perfused with  $37^{\circ}$ C physiological saline until a 16.00 kPa (120 mmHg) internal pressure load was achieved. The IVUS images acquired under 0 kPa were designated the *template* images (Figs. 12.9A and 12.10A), while the images acquired with the artery under 16.00 kPa (120 mmHg) internal pressure load was achieved. The IVUS images acquired under 0 kPa were designated the *target* images (Figs. 12.9B and 12.10B).



Figure 12.9: (A) template image of a coronary artery with a fully formed lipid layer (arrow). (B) Corresponding target image of the artery under 16.00 kPa internal pressure load. (C) FE mesh representation of the image space. (D) circumferential stretch distribution within the arterial wall and lesion.



Figure 12.10: (A) Template image of a coronary artery that does not have a fully developed lipid core. (B) Corresponding target image of the artery under 16 kPa internal pressure load. (C) FE mesh of the image space. (D) Circumferential stretch distribution within the arterial wall and lesion.

The boundaries of the media/lesion were manually segmented in the IVUS template image of the diseased vessels. B-spline curves were fitted to the points generated by segmentation. These curves defined the boundaries of the arterial wall. A 2D plane strain FE model was constructed for each vessel that included the entire image domain Figs. 12.9 C and 12.10C). The lumen and the tissue surrounding the vessels were represented by an isotropic hypoelastic constitutive model with relatively soft elastic material properties (E = 1.0 kPa and  $\nu = 0.3$ ) to provide tethering. The outer edges of the image domain were fully constrained to eliminate rigid body motion. Transversely isotropic hyperelastic strain energy was utilized to describe nonlinear behavior of the arterial wall [57–64] and atherosclerotic lesions [50, 54, 65, 66]. This strain energy definition describes a material that consists of fibers imbedded in an isotropic ground substance. The strain energy function was defined as:

$$W = F_1(\tilde{I}_1, \tilde{I}_2) + F_2(\tilde{\lambda}) + \frac{K}{2} [\ln(J)]^2$$
(12.33)

 $F_1$  represents the behavior of the ground substance while  $F_2$  represents the behavior of the collagen fibers. The final term in the expression represents the

bulk behavior of the material. K is the bulk modulus of the material, F is the deformation gradient tensor and  $J = \det(F)$ .  $\tilde{I}_1$  and  $\tilde{I}_2$  are the first and second deviatoric invariants of the right Cauchy deformation tensor [30]. The scalar  $\tilde{\lambda}$  is the deviatoric stretch ratio along the local fiber direction, a, which was oriented circumferentially for these analyses to correspond with the collagen and smooth muscle fiber orientations in the arterial wall and plaque cap.

A neo-Hookean form was used to represent the ground substance matrix:

$$F_1(\tilde{I}_1) = \mu(\tilde{I}_1 - 3) \tag{12.34}$$

where  $\mu$  is the shear modulus of the ground substance. The stress-stretch behavior for the fiber direction was represented as exponential, with no resistance to compressive load:

$$\tilde{\lambda}W_{\lambda} = \tilde{\lambda}\frac{\partial F_2}{\partial \lambda} = 0, \quad \tilde{\lambda} < 1;$$

$$\tilde{\lambda}W_{\lambda} = \lambda\frac{\partial F_2}{\partial \lambda} = C_3 \left[\exp(C_4(\tilde{\lambda} - 1)) - 1\right], \quad \tilde{\lambda} \ge 1$$
(12.35)

where material coefficients  $C_3$  and  $C_4$  scale the fiber stress and control its rate of rise with increasing stretch, respectively. The full Cauchy stress tensor is defined as.

$$\boldsymbol{T} = 2(W_1)\boldsymbol{B} + \lambda W_\lambda \boldsymbol{a} \otimes \boldsymbol{a} + p\boldsymbol{1}$$
(12.36)

 $W_1$ ,  $W_2$  and  $W_{\lambda}$  are strain energy derivatives with respect to  $I_1$ ,  $I_2$ , and  $\lambda$  [26], and *B* is the left deformation tensor. A detailed description of the finite element implementation of this constitutive model can be found in Weiss *et al.* [19].

The material parameters for the arterial wall were determined by a nonlinear least squares fit to circumferential stress/strain values presented in the work of Cox et al. [58] for the canine coronary artery wall using the constitutive relation described above. The media region of the arterial wall was assigned material properties based on the curve fit obtained from the Cox et al. data [57]. The material constants for the media were  $\mu = 3.57$  kPa,  $C_3 = 4.99$  kPa, and  $C_4 = 5.49$ . The bulk modulus was defined as 200.00 kPa. The lesion areas were assigned identical material properties as were used for the media since the stress-strain behavior of the arterial wall falls well within the wide range of values published for the material properties of atherosclerotic lesions [67].

The warping analyses results indicate (Figs. 12.9D and 12.10D) that the presence of a fully developed lipid core increases the circumferential stretch of the plaque cap adjacent to the lipid core. These results are consistent with previous studies that suggested that the larger lipid layers increase plaque cap stress [53, 54].

#### 12.3.4 Cardiac Mechanics

Assessment of regional heart wall motion (wall motion, thickening, strain, etc.) can identify impairment of cardiac function due to hypertrophic or dilated cardiomyopathies. It can provide quantitative estimates of the impairment of ventricular wall function due to ischemic myocardial disease. The assessment of regional heart motion is used in combination measures of perfusion and metabolic uptake to diagnose and evaluate stunned/hibernating myocardium following transient ischemic events. Stunned myocardium is characterized by decreased or no contractile function but having normal perfusion and glucose utilization [68–70]. Since stunned myocardium has normal perfusion and normal viability, it can only be identified by localizing abnormal wall motion/contraction. Hibernating myocardium is characterized by persistent ventricular myocardial dysfunction with preserved viability, decreased perfusion, and normal metabolic uptake. Hibernating myocardium has been associated with a slower and incomplete restoration of contractile function as compared with stunned myocardium [71, 72]. Up to 50% of patients with ischemic heart disease and LV dysfunction have significant areas of hibernating myocardium [73, 74] and therefore would be predicted to benefit from identification and subsequent revascularization.

The assessment of the size and location of infarction, in particular, the extent of viable tissue, and the mechanical function of the tissue can be extremely valuable for predicting the utility and assessing the success of surgical interventions such as revascularization. Thus the measurement of local myocardial deformation has potential to be an important diagnostic and prognostic tool for the evaluation of a large number of patients.

The deformation of the human heart wall has been quantified via the attachment of physical markers in a select number of human subjects [75]. This approach provided valuable information but is far too invasive to be used in the clinical setting. With the development of magnetic resonance imaging (MRI) tagging techniques, noninvasive measurements of myocardial wall dynamics have been possible [76]. The most commonly clinically utilized techniques for the assessment of myocardial regional wall motion and deformation of the myocardium are echocardiography and tagged MRI. LV wall function is typically assessed using 2-D Doppler echocardiography [77–82] through the interrogation of the LV from various views to obtain an estimate of the 3-D segmental wall motion. However, these measurements are not three-dimensional in nature. Furthermore, echocardiography is limited to certain acquisition windows.

The most widely used approach for determining ventricular deformation is MR tagging [83–88]. MR tagging techniques rely on local perturbation of the magnetization of the myocardium with selective radio-frequency (RF) saturation to produce multiple, thin, tag planes during diastole. The resulting magnetization reference grid persists for up to 400 ms and is convected with the myocardium as it deforms. The tags provide fiducial points from which the strain can be calculated [85, 89]. The primary strength of MRI tagging is that noninvasive in vivo strain measurements are possible [85, 89]. It is effective for tracking fast, repeated motions in 3-D. There are, however, limitations in the use of tagged MRI for cardiac imaging. The measured displacement at a given tag point contains only unidirectional information; in order to track the full 3-D motion, these data have to be combined with information from other orthogonal tag-sets over all time frames [76]. The technique's spatial resolution is coarser than the MRI acquisition matrix. Furthermore, the use of tags increases the acquisition time for the patient compared to standard cine-MRI, although improvements in acquisition speed have reduced the time necessary for image acquisition.

Sinusas *et al.* have developed a method to determine the strain distributions of the left ventricle using untagged MRI [90]. The system is a shape-based approach for quantifying regional myocardial deformations. The shape properties of the endo-and epicardial surfaces are used to derive 3-D trajectories, which are in turn used to deform a finite element mesh of the myocardium. The approach requires a segmentation of the myocardial surfaces in each 3-D image data set to derive the surface displacements.

Our long-term goal is to use Hyperelastic Warping to determine the strain distribution in the normal left ventricle. These data will be compared with the left ventricular function due to the pathologies described above. Toward this end, the initial validation of the use of Hyperelastic Warping with cardiac cine-MRI images is described.

#### 12.3.4.1 Validation of Warping for Tracking Left Ventricular Deformation using Volumetric MRI

To validate the use of Warping for predicting LV strains from sets of volumetric cine-MRI images, a pair of 3-D cine MRI image datasets representing two states of the left ventricle during the cardiac cycle was required. Further, the deformation map between the states represented in the images had to be known to provide a gold standard for comparisons. This was achieved by first acquiring a gated 3-D cine-MRI dataset of a normal volunteer's heart during early diastole on a 1.5T Siemens scanner ( $256 \times 256$  image matrix, 378 mm FOV, 10 mm slice thickness, 10 slices). This volumetric MRI dataset was designated as the *template* image (Fig. 12.11, left). The endocardial and epicardial surfaces of the LV were hand segmented. An FE model of the left ventricular (LV) image space was created based on these segmentations (Fig. 12.12, left panel). The myocardium was represented as a transversely isotropic material with the fiber angle varying linearly from  $-90^{\circ}$  at the epicardial surface, through  $0^{\circ}$  at the Mid-wall, to  $90^{\circ}$ at the endocardial surface [91]. The material coefficients were determined by least squares fit of the transversely isotropic hyperelastic constitutive model described in Weiss et al. [30] described above in the intravascular ultrasound section, to the biaxial stress/strain values presented in the work of Humphrey et al. [31, 32].

An internal pressure load representing end-diastole was applied to the lumen and a standard "forward" nonlinear FE analysis was performed using the



Figure 12.11: Mid-ventricular slices of the template (left) and the target (right) image datasets used in the validation analyses. Left image was obtained from direct MR volumetric image acquisition, while right image was created by deforming left image using results of forward FE analysis (see text).



Figure 12.12: Left – FE mesh for forward model used to create target image. Right – A detailed view of the mesh corresponding to myocardial wall. Black arrows indicate the pressure load applied to the endocardial surface.

NIKE3-D finite element program [92] (Fig. 12.12). Using the deformation map obtained from the forward FE analysis, a deformed volumetric image dataset (target) was created by applying the deformation map to the original template MRI image (Fig. 12.12, right panel).

A Warping model was created using the same geometry and material parameters that were used in the forward model described above. The Warping analysis was performed using the template image data set and a target image dataset was created by applying the forward model's deformation map to deform the template image. This yielded a template and target with a known solution for the deformations between them. The forward FE and Warping predictions of fiber stretch (final length/initial length along the local fiber direction) were compared to determine the accuracy of the technique. The validation results indicated good agreement between the forward and the warping fiber stretch distributions (Fig. 12.13). A detailed analysis of the forward and predicted (Warping) stretch distributions for each image plane indicated good agreement (Fig. 12.14).

To determine the sensitivity of the Warping analysis to changes in material parameters,  $\mu$  and  $C_3$  were increased and decreased by 24% of the baseline values. The 24% increase and decrease corresponds to the 95% confidence interval of material parameters determined from the least-squares fit of the material model to the Humphrey et al. data [31, 32]. Since, the proper material model is often not known for biological tissue, the material model was changed from the transversely isotropic model described above to an isotropic neo-Hookean



Figure 12.13: Fiber stretch distribution for the forward (left) and warping (right) analyses. The locations for the sensitivity analysis are shown on the forward model as numbers 1–4. Locations 5–8 are at the same locations as 1–4 but at the mid-ventricle level.



Figure 12.14: Comparison of warping and forward nodal fiber stretch for each image slice. Y7 corresponds to the slice at the base of LV and Y1 is near the apex of the LV.

material model. The analysis was repeated and the results compared with the forward model results.

The forward and Warping sensitivity study results were compared at eight locations (Fig. 12.3). These results show excellent agreement (Table 3.1) for all cases indicating hyperelastic Warping is relatively insensitive to changes to material model and material parameters. These results indicate that accurate predictions can be determined even when material model and parameters are not known. This is consistent with our previous results of Warping analyses of intravascular ultrasound images [22].

caption for Fig. 3.13 above								
	1	2	3	4	5	6	7	8
Location	Upper ventricle				Mid-ventricle			
Forward	1.09	1.06	1.12	1.07	1.08	1.04	1.02	1.05
$\mu + 24\%$	1.09	1.09	1.13	1.07	1.07	1.03	1.03	1.05
$\mu - 24\%$	1.09	1.09	1.13	1.07	1.08	1.03	1.03	1.05
$C_3 + 24\%$	1.09	1.08	1.13	1.08	1.08	1.03	1.03	1.05
$C_3 - 24\%$	1.10	1.09	1.13	1.07	1.08	1.03	1.03	1.05
Neo-Hookean	1.10	1.07	1.13	1.07	1.07	1.02	1.02	1.05

Table 12.1: Effect of changes in material properties and material model on predicted fiber stretch. "Forward" indicates the forward FE solution, the "gold standard". Columns indicate locations 1–8 of the left ventricle, defined in the caption for Fig. 3.13 above

#### 12.3.4.2 Myocardial Infarction

To study changes in systolic wall function due to myocardial infarction, a warping analysis was performed on a 3-D cine-MRI image data set for an individual with a lateral wall myocardial infarction (Male, 155 lbs, 51 y/o at time of scan, diabetic w/small infarction.) The subendocardial infarction can be seen as the hyperenhancement of the lateral wall shown in the ce-MRI image (Fig. 3.15A).

Delayed contrast enhanced MRI (ce-MRI) has been shown to be able to identify regions of infraction in the myocardium as hyperenhanced [93–96]. Furthermore, studies have indicated that the transmural extent of the hyperenhancement of ce-MRI predicts recovery of function after revascularization [97, 98] and can predict improved contractility post-revascularization [94].

To acquire the ce-MRI image data sets, the patients were placed supine in a 1.5T clinical scanner (General Electric) and a phased-array receiver coil was placed on the chest for imaging. A commercially available gadolinium-based contrast agent was administered intravenously at a dose of 0.2 mmol/kg and gated images were acquired 10–15 min after injection with 10 s breath holds. The contrast-enhanced images were acquired with the use of a commercially available segmented inversion-recovery sequence from General Electric. The 3-D cine-MRI image data sets for this patient were acquired on a 1.5T GE scanner ( $256 \times 256$  image matrix, 378 mm FOV, 10 mm slice thickness, 10 slices). The volumetric MRI dataset corresponding to end-systole was designated as the *template* image (Fig. 12.15C) while the image dataset corresponding to



Figure 12.15: (A) Mid-ventricle contrast-enhanced MRI image of the left ventricle. The hyperenhancement indicates the location of the infarction (arrow in left panel). (B) Circumferential stretch distribution for systolic contraction. The arrow indicates the infarcted area of the lateral wall does not contract during systole. Mid-ventricle slices of the 3-D cine MRI image data used for the systolic function analysis. (C) Mid-systolic image (*template*). (D) End-systolic image (*target*).

end-diastole was designated the *target* image (Fig. 12.15D). A warping model and analysis was made using the methods detailed above.

The warping analysis reveals that the infarcted area undergoes little deformation during systole (circumferential stretch near 1.0). The analysis further reveals that the wall dysfunction extends over the lateral wall of the myocardium outside the area of hyperenhancement indicated in the ce-MRI images (Fig. 3.15A). These results indicate that the contractile function of the heart is significantly impaired within and adjacent to the infarcted region.

## 12.4 Discussion and Conclusions

Hyperelastic Warping is a highly flexible registration method that can be used for the registration of physical and nonphysical deformations. It makes use of either easily constructed regular meshes or irregular meshes that conform to the geometry of the structure being registered and can be used to register a particular region of interest or the entire image space. Additionally, hyperelasticity provides a physically realistic constraint for the registration of soft tissue deformation. Hyperelasticity based constitutive relations have been used to describe the behavior of a wide variety of soft tissues including the left ventricle [99–102], arterial tissue [103, 104]. skin [105] and ligaments[106–109]. Hyperelastic Warping can be tailored to the type of soft tissue being registered through the appropriate choice of hyperelastic material model and material parameters.

Deformable image registration models based other material models have been used extensively in the field of anatomical brain registration. As was described above, an energy functional is minimized in order to achieve the registration solution. This functional consists of a measure of image similarity and an internal energy term (Eq. 12.4). Measures of image similarity take the form of differences in the square of the image intensities (Eq. 12.8) [15-17, 19, 110, 111] or are based on cross-correlation methods of the intensity or intensity gradient values [112]. Since the internal energy term of the energy functional is derived from the material model through the strain energy W, the registration process takes on the characteristics of the underlying material model. For example, registration methods that use a viscous or inviscid fluid constitutive model [15,17] have been shown to provide excellent registration results. However, these models have a tendency to underpenalize shear deformations, thus producing physically unrealistic registration of solids. In other words, the deformation of the deformable template resembles that of a fluid rather than that of a solid.

Other continuum-based methods for deformable image registration use linear elasticity [12, 13, 15, 16] to regularize registration. The use of linear elasticity is attractive due to the fact that it is relatively simple to implement. However, for the large deformations involved in inter- or intra-subject registration, it has a tendency to over-penalize large deformations. This is due to the fact that linear elasticity is not rotationally invariant. For an isotropic linear elastic material, the constitutive law is:

$$\boldsymbol{T} = \lambda \operatorname{tr}(\boldsymbol{e}) + \mu \boldsymbol{e}. \tag{12.37}$$

Here,  $\lambda$  and  $\mu$  are the Lamé material coefficients, and e is the infinitesimal strain "tensor" defined in terms of the displacement gradients. The infinitesimal strain

is not a true tensor since it does not obey the transformation laws for 2<sup>nd</sup> order tensors. In detail:

$$\boldsymbol{e} = \frac{1}{2} \left( \left( \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right)^T + \frac{\partial \boldsymbol{u}}{\partial \boldsymbol{X}} \right) \cdot$$
(12.38)

But, 
$$\frac{\partial u}{\partial X} = \frac{\partial (x - X)}{\partial X} = F - 1.$$
 (12.39)

For any deformation gradient F, we can use the polar decomposition to write F as F = RU, where R is a proper orthogonal rotation and U is the positive definite symmetric right stretch tensor. With this substitution,

$$\boldsymbol{e} = \frac{1}{2} \left( (\boldsymbol{R}\boldsymbol{U} - \boldsymbol{1})^T + (\boldsymbol{R}\boldsymbol{U} - \boldsymbol{1}) \right).$$
(12.40)

As indicated in Eq. (12.40), the strain e depends directly on R, which describes the local rigid body rotation. As a result, even the smallest rotation of material axes induces stress in a linear elastic solid, making the constitutive model nonobjective.

This work has demonstrated that Hyperelastic Warping may be used to analyze a wide variety of image registration problems, using standard medical image modalities such as ultrasound, MRI, and CT. The types of analyses demonstrated range from anatomical matching typical of nonphysical image registration, to the large physical deformations present in the deformation of the left ventricle over the cardiac cycle. As demonstrated in the presented work, the method allows for the estimation of the stress distribution within the structure(s) being registered, an attribute that has not been demonstrated by other registration methods.

### Acknowledgments

Financial support from NSF Grant #BES-0134503 (JAW, AIV, NP), NIH grant #R01-EB000121 (JAW, AIV) and NIH Grant #PO1-DC01837 (AIV) is gratefully acknowledged. An allocation of computer time was provided by the Center for High Performance Computing (CHPC) at the University of Utah. The authors thank the following individuals for their contributions to this work and their continued collaboration: Grant T. Gullberg, Richard D. Rabbitt, Willem F. Decraemer, Anton E. Bowden, Bradley N. Maker, Steve A. Maas, Geoffrey D. Vince, Robert J. Gillies, Edward V. R. DiBella and Jean-Philippe Galons.

# Questions

Question 1: How are the principles of continuum mechanics used to regularize the deformable image registration problem involving the deformation of a template image into alignment with a target image? What are the primary advantages of this approach to regularization of the deformable image problem in comparison to ad hoc methods?

Question 2: What is the purpose of the regularization term W in the deformable image registration problem?

Question 3: What is meant by treating the image data as a "hard constraint" in the deformable image registration problem?

Question 4: In Hyperelastic Warping, in the limit as the penalty parameter  $\lambda \to \infty$ , the image-based energy converges to a finite value. Explain.

Question 5: Treating the image data as a hard constraint may cause the stiffness matrix to become ill-conditioned. How does the augmented Lagrangian method solve this problem?

Question 6: What is the role of the stiffness quantities in the solution procedure?

Question 7: How is sequential low-pass-filtering used in Hyperelastic Warping to keep from converging to local minima in the solution?

Question 8: When using a regular mesh for Hyperelastic Warping, why is rezoning needed?

Question 9: How is mechanical stress calculated with Hyperelastic Warping?

### References

- Bathe, K. J., Finite Element Procedures, Prentice-Hall, Englewood Cliffs, NJ, 1996.
- [2] Hadamard, J., Sur les problemes aux derivees partielles et leur signification physique, Bull. Univ. Princeton, Vol. 13, 1902.
- [3] Hadamard, J., Le probleme de Cauchy et les equations aux derivees lineaires hyperboliques, Hermann, Paris, 1932.
- [4] Grenander, U., Lectures in Pattern Theory: Pattern Analysis, Pattern Synthesis, and Regular Structures, Vols. I, II and III, Springer-Verlag, Berlin, 1976.
- [5] Grenander, U., Chow, Y., and Keenan, D., HANDS: A Pattern Theoretic Study of Biological Shapes, Springer-Verlag, New York, 1990.
- [6] Davatzikos, C., Spatial normalization of 3D brain images using deformable models, J. Comput. Assist. Tomogr., Vol. 20, No. 4, pp. 656– 665, 1996.
- [7] Davatzikos, C., Li, H. H., Herskovits, E., and Resnick, S. M., Accuracy and sensitivity of detection of activation foci in the brain via statistical parametric mapping: a study using a PET simulator, Neuroimage, Vol. 13, No. 1, pp. 176–184, 2001.
- [8] Amit, Y., Graphical shape templates for automatic anatomy detection with applications to MRI brain scans, IEEE Trans. Med. Imaging, Vol. 16, pp. 28–40, 1997.
- [9] Tagare, H. D., Deformable 2-D template matching using orthogonal curves, IEEE Transa. Med. Imaging, Vol. 16, pp. 108–117, 1997.
- [10] Miller, M. I., Christensen, G. E., Amit, Y., and Grenander, U., Mathematical Textbook of Deformable Neuroanatomies, Proc. Nat. Acad. Sci., Vol. 90, No. 24, pp. 144–148, 1993.
- [11] Hill, A., Cootes, T. F., Taylor, C. J., and Lindley, K., Medical image interpretation: A generic approach using deformable templates, Med. Inform., Vol. 19, No. 1, pp. 47–59, 1994.

- [12] Bajcsy, R., Lieberson, R., and Reivich, M., A computerized system for the elastic matching of deformed radiographic images to idealized atlas images, J. Comput. Assist. Tomogr., Vol. 7, No. 4, pp. 618–625, 1983.
- [13] Bajcsy, R. and Kovacic, S., Multiresolution elastic matching, Comput. Vis. Graph. Image Process. Vol. 46, pp. 1–21, 1989.
- [14] Gee, J. C. and Peralta, P. D., Continuum models for Bayesian image matching, in: Maximum Entropy and Bayesian Methods, Hanson K. M. and Silver, R. N. eds., Kluwer Academic Publishers, Dordrecht, pp. 109–11, 1996.
- [15] Christensen, G. E., Rabbitt, R. D., and Miller, M. I., 3D brain mapping using a deformable neuroanatomy, Phys. Med. Biol., Vol. 39, pp. 609–618, 1994.
- [16] Rabbitt, R. D., Weiss, J. A., Christensen, G. E., and Miller, M. I., Mapping of hyperelastic deformable templates, Proc. SPIE Int. Soc. Opt. Eng., Vol. 252, pp. 252–265, 1995.
- [17] Christensen, G. E., Rabbitt, R. D., and Miller, M. I., Deformable templates using large deformation kinematics, IEEE Trans. Image Process. Vol. 5, No. 10, pp. 1435–1447, 1996.
- [18] Bowden, A. E., Rabbitt, R. D., Weiss, J. A., and Maker, B. N., Anatomical registration and segmentation by warping template finite element models, Proc. SPIE Int. Soc. Opt. Eng., Vol. 3254, pp. 469–476, 1998.
- [19] Weiss, J. A., Rabbitt, R. D., and Bowden, A. E., Incorporation of medical image data in finite element models to track strain in soft tissues, Proc. SPIE Int. Soc. Opt. Eng., Biomedical Optics Symposium BiOS98, Vol. 3254, pp. 477–484, 1998.
- [20] Veress, A. I., Weiss, J. A., Gullberg, G. T., Vince, D. G., and Rabbitt, R. D., Validation for the determination of strain and displacement fields in coronary arteries imaged by intravascular ultrasound. ASME Summer Bioengineering Conference, 50:71–72, 2001.
- [21] Veress, A. I., Weiss, J. A., Klein, G. J., and Gullberg, G. T., Quantification of 3D Left Ventricular Deformation using Hyperelastic Warping: Comparisons between MRI and PET Imaging. Proceedings of

IEEE/EMBS Computers in Cardiology, Rotterdam, The Netherlands, September 22–25, 29:709–712, 2002.

- [22] Veress, A. I., Weiss, J. A., Gillies, R. J., Bowden, A. E., Galons, J. P., and Rabbitt, R. D., Deformable Image Registration of Mouse Brain MRI Data using Hyperelastic Warping. Proceedings of the 2002 ASME Advances in Bioengineering, IMECE2002-32598, 2002.
- [23] Viegas, S. F., Hillman, G. F., Elder, K., and Stoner, D., Measurement of carpal bone geometry by computer analysis of three-dimensional CT images, J. Hand Surg., Vol. 18, No. 2, pp. 341–349, 1993.
- [24] Simo, J. C. and Hughes, T. J. R., Computational Inelasticity, Springer-Verlag, New York, 1998.
- [25] Marsden, J. E. and Hughes, T. J. R., Mathematical Foundations of Elasticity, Dover, Minneola, NY, 1994.
- [26] Spencer, A. J. M., Continuum Mechanics, Longman Scientific & Technical, Essex, 1980.
- [27] Matthies, H. and Strang, G., The solution of nonlinear finite element equations, Int. J. Numeri. Methods Eng., Vol. 14, pp. 1613–1626, 1979.
- [28] Reddy, J. N., Energy and Variational Methods in Applied Mechanics, Wiley, New York, 1984.
- [29] Lai, W. M., Introduction to Continuum Mechanics, Permagon Press, New York, 1993.
- [30] Weiss, J., Maker, B., and Govindjee, S., Finite element implementation of incompressible transversely isotropic hyperelasticity, Comput. Methods Appl. Mech. Eng., Vol. 135, pp. 107–128, 1996.
- [31] Humphrey, J. D., Strumpf, R. K., and Yin, F. C., Determination of a constitutive relation for passive myocardium, Part II: Parameter estimation, J. Biomech. Eng., Vol. 112, pp. 340–346, 1990.
- [32] Humphrey, J. D., Strumpf, R. K., and Yin, F. C., Determination of a constitutive relation for passive myocardium, Part I: A new functional form, J. Biomech. Eng., Vol. 112, pp. 333–339, 1990.

- [33] Simo, J. C., Taylor, R. L., Schaewe, T. J., and Miller, M. I., Quasiincompressible finite elasticity in principal stretches: Continuum basis and numerical algorithms, Comput. Methods Appl. Mech. Eng., Vol. 85, pp. 273–310, 1991.
- [34] Hughes, T. J. R., The Finite Element Method: Linear Static and Dynamic Finite Element Analysis, Prentice Hall, Englewood Cliffs, NJ., 1987.
- [35] Fletcher, R., Practical Methods of Optimization, Wiley, New Delhi, 1989.
- [36] Powell, M. J. D., A mehod for nonlinear constraints in minimization problems, In: optimization, Fletcher, R. Ed. Academic Press, New York, 1969.
- [37] Hestenes, M. R., Multiplier and gradient methods, J. Optim. Theory Appl., Vol. 4, pp. 303–320, 1969.
- [38] Gonzalez, R. C. and Woods, R. E., Digital Image Processing. Addison-Wesley, Reading, MA, pp. 187–213, 1992.
- [39] Veress, A. I., Weiss, J. A., Gullberg, G. T., Vince, D. G., and Rabbitt, R. D., Strain measurement in coronary arteries using intravascular ultrasound and deformable images, J. Biomech. Eng., Vol. 124, No. 6, pp. 734–741, 2002.
- [40] Huebner, K. H., The Finite Element Method for Engineers, Wiley, New York, 1995.
- [41] Wada, H. and Metoki, T., Analysis of dynamic behavior of human middle ear using a finite-element method, J. Acoustical Soc. Am. Vol. 92, No. 6, pp. 3157–3168, 1992.
- [42] Wada, H., Ando, M., Takeuchi, M., Sugawara, H., Koike, T., Kobayashi, T., Hozawa, K., Gemma, T., and Nara, M., Vibration measurement of the tympanic membrane of guinea pig temporal bones using time-averaged speckle pattern interferometry, J. Acoustical Soc. Am., Vol. 111, No. 5pt. 1, pp. 2189–2199, 2002.
- [43] Koike, T., Wada, H., and Kobayashi, T., Modeling of the human middle ear using the finite-element method, J. Acoustical Soc. Am. Vol. 111, No. 3, pp. 1306–1317, 2002.

- [44] Huber, A. M., Schwab, C., Linder, T., Stoeckli, S. J., Ferrazzini, M., Dillier, N., and Fisch, U., Evaluation of eardrum laser doppler interferometry as a diagnostic tool, Laryngoscope, Vol. 111, No. 3, pp. 501–507, 2001.
- [45] Ladak, H. M. and Funnel, W. R. J., Finite-element modeling of the normal and surgically repaired cat middle ear, J. Acoustical Soc. Am. Vol. 100, No. 2, pp. 933–944, 1996.
- [46] Maker, B., Ferencz, R., and Hallquist, J., NIKE3D: A nonlinear implicit three-dimensional finite element code for solid and structural mechanics, UC-Lawrence Livermore National Laboratory Report, Vol. UCRL-MA-105268 rev 1, 1995.
- [47] National Health and Nutrition Examination Survey (NHANES), CDC/NCHS and the American Heart Association, 1996.
- [48] MacIsaac, A. I., Thomas, J. D., and Topol, E. J., Toward the quiescent coronary plaque, J. Am. Coll. Cardiol., Vol. 22, No. 4, pp. 1228–1241, 1993.
- [49] Richardson, P. D., Davies, M. J., and Born, G. V., Influence of plaque configuration and stress distribution on fissuring of coronary atherosclerotic plaques, Lancet, Vol. 2, No. 8669, pp. 941–944, 1989.
- [50] Loree, H. M., Kamm, R. D., Stringfellow, R. G., and Lee, R. T., Effects of fibrous cap thickness on peak circumferential stress in model atherosclerotic vessels, Circ. Res., Vol. 71, No. 4, pp. 850–858, 1992.
- [51] Cheng, G. C., Loree, H. M., Kamm, R. D., Fishbein, M. C., and Lee, R. T., Distribution of circumferential stress in ruptured and stable atherosclerotic lesions. A structural analysis with histopathological correlation, Circulation, Vol. 87, No. 4, pp. 1179–1187, 1993.
- [52] Veress, A. I., Cornhill, J. F., Herderick. E. E., and Thomas, J. D., Age-related development of atherosclerotic plaque stress: A population-based finite-element analysis, Coron. Artery Dis., Vol. 9, No. 1, pp. 13–19, 1998.
- [53] Veress, A. I., Vince, D. G., Anderson, P. M., Cornhill, J. F., Herderick, E. E., Klingensmith, J. D., Kuban, B. D., Greenberg, N. L., and Thomas,

J. D., Vascular mechanics of the coronary artery, Z. Kardiol., Vol. 89 (Suppl. 2), pp. 92–100, 2000.

- [54] Loree, H. M., Tobias, B. J., Gibson, L. J., Kamm, R. D., Small, D. M., and Lee, R. T., Mechanical properties of model atherosclerotic lesion lipid pools, Arteriosclerosis Thrombosis, Vol. 14, No. 2, pp. 230–234, 1994.
- [55] Nissen, S. E., Gurley, J. C., Grines, C. L., Booth, D. C., McClure, R., Berk. M., Fischer, C., and DeMaria, A. N., Intravascular ultrasound assessment of lumen size and wall morphology in normal subjects and patients with coronary artery disease, Circulation, Vol. 84, No. 3, pp. 1087–1099, 1991.
- [56] Bermejo, J., Botas, J., Garcia, E., Elizaga, J., Osende, J., Soriano, J., Abeytua, M., and Delcan, J. L., Mechanisms of residual lumen stenosis after high-pressure stent implantation: a quantitative coronary angiography and intravascular ultrasound study, Circulation, Vol. 98, No. 2, pp. 112–118, 1998.
- [57] Cox, R. H., Anisotropic properties of the canine carotid artery in vitro, J. Biomech., Vol. 8, No. 5, pp. 293–300, 1975.
- [58] Cox, R. H., Passive mechanics and connective tissue composition of canine arteries, Am. J. Physiol., Vol. 234, No. 5, pp. H533–541, 1978.
- [59] Dobrin, P. B, Biaxial anisotropy of dog carotid artery: estimation of circumferential elastic modulus, J. Biomech., Vol. 19, No. 5, pp. 351–358, 1986.
- [60] Dobrin, P. B. and Doyle, J. M., Vascular smooth muscle and the anisotropy of dog carotid artery, Circ. Res., Vol. 27, No. 1, pp. 105–119, 1970.
- [61] Dobrin, P. B. and Mrkvicka, R., Estimating the elastic modulus of non-atherosclerotic elastic arteries, J. Hypertension, Vol. 10, No. 6, (Suppl.), pp. S7–10, 1992.
- [62] Frobert, O., Gregersen, H., and Bagger, J. P., Mechanics of porcine coronary arteries ex vivo employing impedance planimetry: A new intravascular technique, Ann. Biomed. Eng. Vol. 24, No. 1, pp. 148–155, 1996.

- [63] Patel, D. J. and Janicki, J. S., Static elastic properties of the left coronary circumflex artery and the common carotid artery in dogs, Circ. Res., Vol. 27, No. 2. pp. 149–158, 1970.
- [64] Patel, D. J., Janicki, J. S., and Carew. T. E., Static anisotropic elastic properties of the aorta in living dogs, Circ. Res., Vol. 25, No. 6., pp. 765–779, 1969.
- [65] Salunke, N. V. and Topoleski, L. D., Biomechanics of atherosclerotic plaque, Crit. Rev. Biomed. Eng., Vol. 25, No. 3, pp. 243–285, 1997.
- [66] Weizsacker, H. W., Lambert, H., and Pascale, K., Analysis of the passive mechanical properties of rat carotid arteries, J. Biomech., Vol. 16, No. 9, pp. 703–715, 1983.
- [67] Humphrey, J. D., Cardiovascular Solid Mechanics Cells, Tissues, and Organs, Springer-Verlag, New York, 2002.
- [68] Mazhari, R. and McCulloch, A. D., Integrative models for understanding the structural basis of regional mechanical dysfunction in ischemic myocardium. Ann. Biomed. Eng., Vol. 28, No. 8, pp. 979–900, 2000.
- [69] Kraitchman, D. L., Hillenbrand, H. B., Oznur, I., Lima, J. A., McVeigh, E. R., Zerhouni, E. A., and Bluemke, D. A., Noninvasive assessment of myocardial stunning from short-term coronary occlusion using tagged magnetic resonance imaging, J. Cardiovasc. Magn. Reson., Vol. 2, No. 2, pp. 123–136, 2000.
- [70] Dilsizian, V. and Bonow, R. O., Current diagnostic techniques of assessing myocardial viability in patients with hibernating and stunned myocardium, Circulation, Vol. 87, No. 1, pp. 1–20, 1993.
- [71] Haas, F., Jennen, L., Heinzmann, U., Augustin, N., Wottke, M., Schwaiger, M., and Lange, R., Ischemically compromised myocardium displays different time-courses of functional recovery: correlation with morphological alterations?, Eur. J. Cardiothorac Surg., Vol. 20, No. 2, pp. 290–298, 2001.
- [72] Bax, J. J., Visser, F. C., Poldermans, D., Elhendy, A., Cornel, J. H., Boersma, E., van Lingen, A., Fioretti, P. M., and Visser, C. A., Time course of functional recovery of stunned and hibernating

segments after surgical revascularization, Circulation, Vol. 104, No. 12 (Suppl. 1), pp. 1314–1318, 2001.

- [73] Bonow, R. O., Identification of viable myocardium, Circulation, Vol. 94, No. 11, pp. 2674–2480, 1996.
- [74] al-Mohammad, A., Mahy, I. R, Norton, M. Y., Hillis, G., Patel, J. C., Mikecz, P., and Walton, S., Prevalence of hibernating myocardium in patients with severely impaired ischaemic left ventricles, Heart, Vol. 80, No. 6, pp. 559–564, 1998.
- [75] Ingels, N. B., Daughters, G. T., Stinson, E. B., and Alderman, E. L., Measurement of midwall myocardial dynamics in intact man by radiography of surgically implanted markers, Circulation, Vol. 52, No. 5, pp. 859–867, 1975.
- [76] Ozturk, C. and McVeigh, E. R., Four-dimensional B-spline based motion analysis of tagged MR images: introduction and in vivo validation, Phys. Med. Biol., Vol. 45, No. 6, pp. 1683–1702, 2000.
- [77] Weidemann, F., Dommke, C., Bijnens, B., Claus, P., D'Hooge, J., Mertens, P., Verbeken, E., Maes, A., Van de Werf, F., De Scheerder, I., and Sutherland, G. R, Defining the transmurality of a chronic myocardial infarction by ultrasonic strain-rate imaging: implications for identifying intramural viability: an experimental study, Circulation, Vol. 107, No. 6, pp. 883–888, 2003.
- [78] Weidemann, F., Eyskens, B., Jamal, F., Mertens, L., Kowalski, M., D'Hooge, J., Bijnens, B., Gewillig, M., Rademakers, F., Hatle, L., and Sutherland, G. R., Quantification of regional left and right ventricular radial and longitudinal function in healthy children using ultrasoundbased strain rate and strain imaging, J. Am. Soc. Echocardiogr., Vol. 15, No. 1, pp. 20–28, 2002.
- [79] Weidemann, F., Jamal, F., Kowalski, M., Kukulski, T., D'Hooge, J., Bijnens, B., Hatle, L., De Scheerder, I., and Sutherland, G. R., Can strain rate and strain quantify changes in regional systolic function during dobutamine infusion, B-blockade, and atrial pacing—Implications for quantitative stress echocardiography, J. Am. Soc. Echocardiogr., Vol. 15, No. 5, pp. 416–424, 2002.
- [80] Weidemann, F., Kowalski, M., D'Hooge, J., Bijnens, B., and Sutherland, G. R., Doppler myocardial imaging. A new tool to assess regional inhomogeneity in cardiac function, Basic Res. Cardiol., Vol. 96, No. 6, pp. 595–605, 2001.
- [81] D'Hooge, J., Konofagou, E., Jamal, F., Heimdal, A., Barrios, L., Bijnens, B., Thoen, J., Van de Werf, F., Sutherland, G., and Suetens, P., Two-dimensional ultrasonic strain rate measurement of the human heart in vivo, IEEE Trans. Ultrason. Ferroelectr. Freq. Control., Vol. 49, No. 2, pp. 281–286, 2002.
- [82] D'Hooge, J., Heimdal, A., Jamal, F., Kukulski, T., Bijnens, B., Rademakers, F., Hatle, L., Suetens, P., and Sutherland, G. R., Regional strain and strain rate measurements by cardiac ultrasound: principles, implementation and limitations, Eur. J. Echocardiogr., Vol. 1, No. 3, pp. 154–170, 2000.
- [83] Bolster, B. D., Jr., McVeigh, E. R., and Zerhouni, E. A., Myocardial tagging in polar coordinates with use of striped tags, Radiology, Vol. 177, No. 3, pp. 769–772, 1990.
- [84] Zerhouni, E. A., Parish, D. M., Rogers, W. J., Yang, A., and Shapiro, E. P., Human heart: Tagging with MR imaging—A method for noninvasive assessment of myocardial motion, Radiology, Vol. 169, No. 1, pp. 59–63, 1988.
- [85] Buchalter, M. B., Weiss, J. L., Rogers, W. J., Zerhouni, E. A., Weisfeldt, M. L., Beyar, R., and Shapiro, E. P., Noninvasive quantification of left ventricular rotational deformation in normal humans using magnetic resonance imaging myocardial tagging, Circulation, Vol. 81, No. 4, pp. 1236–1244, 1990.
- [86] McVeigh, E. R. and Atalar, E., Cardiac tagging with breath-hold cine MRI, Magn. Reson. Med., Vol. 28, No. 2, pp. 318–327, 1992.
- [87] McVeigh, E. R. and Zerhouni, E. A., Noninvasive measurement of transmural gradients in myocardial strain with MR imaging, Radiology, Vol. 180, No. 3, pp. 677–683, 1991.

- [88] Axel, L., Goncalves, R. C., and Bloomgarden, D., Regional heart wall motion: Two-dimensional analysis and functional imaging with MR imaging, Radiology, Vol. 183, No. 3, pp. 745–750, 1992.
- [89] Ungacta, F. F., Davila-Roman, V. G., Moulton, M. J., Cupps, B. P., Moustakidis, P., Fishman, D. S., Actis, R., Szabo, B. A., Li, D., Kouchoukos, N. T., and Pasque, M. K., MRI-radiofrequency tissue tagging in patients with aortic insufficiency before and after operation, Ann. Thorac Sur., Vol. 65, No. 4, pp. 943–950, 1998.
- [90] Sinusas, A. J., Papademetris, X., Constable, R. T., Dione, D. P., Slade, M. D., Shi, P., and Duncan, J. S., Quantification of 3-D regional myocardial deformation: Shape-based analysis of magnetic resonance images, Am. J. Physiol. Heart Circ. Physiol., Vol. 281, No. 2, pp. H698–714, 2001.
- [91] Berne, R. M. and Levy, M. N., Physiology, Mosby Year Book, St. Louis, MO, 1998.
- [92] Maker, B. N., Ferencz, R., and Hallquist, J., NIKE3D: A nonlinear, implicit, three-dimensional finite element code for solid and structural mechanics. Lawrence Livermore National Laboratory Report #UCRL-MA-105268, Rev. 1, 1995.
- [93] Ramani, K., Judd, R. M., Holly, T. A., Parrish, T. B., Rigolin, V. H., Parker, M. A., Callahan, C., Fitzgerald, S. W., Bonow, R. O., and Klocke, F. J., Contrast magnetic resonance imaging in the assessment of myocardial viability in patients with stable coronary artery disease and left ventricular dysfunction, Circulation, Vol. 98, No. 24, pp. 2687–2694, 1998.
- [94] Kim, R. J., Wu, E., Rafael, A., Chen, E. L., Parker, M. A., Simonetti, O., Klocke, F. J., Bonow, R. O., and Judd, R. M., The use of contrastenhanced magnetic resonance imaging to identify reversible myocardial dysfunction, N. Engl. J. Med., Vol. 343, No. 20, pp. 1445–1453, 2000.
- [95] Szolar, D. H., Saeed, M., Wendland, M. F., Sakuma, H., Roberts, T. P., Stiskal, M. A., Derugin, N., and Higgins, C. B., MR imaging characterization of postischemic myocardial dysfunction ("stunned myocardium"): Relationship between functional and perfusion abnormalities, J. Magn. Reson. Imaging, Vol. 6, No. 4, pp. 615–624, 1996.

- [96] Wu, E., Judd, R. M., Vargas, J. D., Klocke, F. J., Bonow, R. O., and Kim, R. J., Visualisation of presence, location, and transmural extent of healed Q-wave and non-Q-wave myocardial infarction, Lancet, Vol. 357, No. 9249, pp. 21–28, 2001.
- [97] Gerber, B. L., Garot, J., Bluemke, D. A., Wu, K. C., and Lima, J. A., Accuracy of contrast-enhanced magnetic resonance imaging in predicting improvement of regional myocardial function in patients after acute myocardial infarction, Circulation, Vol. 106, No. 9, pp. 1083–1089, 2002.
- [98] Choi, K. M., Kim, R. J., Gubernikoff, G., Vargas, J. D., Parker, M., and Judd, R. M., Transmural extent of acute myocardial infarction predicts long-term improvement in contractile function, Circulation, Vol. 104, No. 10, pp. 1101–1107, 2001.
- [99] Costa, K. D., Hunter, P. J., Wayne, J. S., Waldman, L. K., Guccione, J. M., and McCulloch, A. D., A three-dimensional finite element method for large elastic deformations of ventricular myocardium, Part II: Prolate spheroidal coordinates, J. Biomech. Eng. Vol. 118, No. 4, pp. 464–472, 1996.
- [100] Costa, K. D., Hunter, P. J., Rogers, J. M., Guccione, J. M., Waldman, L. K., and McCulloch, A. D., A three-dimensional finite element method for large elastic deformations of ventricular myocardium, Part I: Cylindrical and spherical polar coordinates, J. Biomech. Eng., Vol. 118, No. 4, pp. 452–463, 1996.
- [101] Guccione, J. M., McCulloch, A. D., and Waldman, L. K., Passive material properties of intact ventricular myocardium determined from a cylindrical model, J. Biomech. Eng., Vol. 113, No. 1, pp. 42–55, 1991.
- [102] Needleman, A., Rabinowitz, S. A., Bogen. D. K., and McMahon. T. A., A finite element model of the infarcted left ventricle, J. Biomech., Vol. 16, No. 1, pp. 45–58, 1983.
- [103] Vito, R. P. and Hickey, J., The mechanical properties of soft tissues, Part II: The elastic response of arterial segments, J. Biomech., Vol. 13, No. 11, pp. 951–957, 1980.

- [104] Vito, R. P., The mechanical properties of soft tissues, Part I: A mechanical system for bi-axial testing, J. Biomech., Vol. 13, No. 11, pp. 947–950, 1980.
- [105] Allaire, P. E., Thacker, J. G., Edlich, R. F., Rodeheaver, G. J., and Edgerton, M. T., Finite deformation theory for in vivo human skin, J. Bioeng., Vol. 1, No. 3, pp. 239–249, 1977.
- [106] Gardiner, J. C., Weiss, J. A., and Rosenberg, T. D., Strain in the human medial collateral ligament during valgus loading of the knee, Clin. Orthopaedics, No. 391, pp. 266–274, 2001.
- [107] Gardiner, J. C. and Weiss, J. A., Effects of flexion angle and valgus rotation on stresses in the human medial collateral ligament, Proc. ASME Bioeng. Conf., Vol. BED-35, pp. 27–28, 1997.
- [108] Quapp, K. M. and Weiss, J. A., Material characterization of human medial collateral ligament, J. Biomech. Eng., Vol. 120, No. 6, pp. 757–763, 1998.
- [109] Weiss, J. A., Gardiner, J. C., and Quapp, K. M., Material models for the study of soft tissue mechanics, Proceeding of the International Conference on Pelvic and Lower Extremity Injuries, December 4–6, 1995, pp. 249–261.
- [110] Johnson, H. J. and Christensen, G. E., Consistent landmark and intensity-based image registration, IEEE Trans. Med. Imaging, Vol. 21, No. 5, pp. 450–461, 2002.
- [111] Christensen, G. E. and Johnson, H. J., Consistent image registration, IEEE Trans. Med. Imaging, Vol. 20, No.7, pp. 568–582, 2001.
- [112] Maintz, J. B. and Viergever, M. A., A survey medical image registration, Med. Image Anal., Vol. 2, No. 1, pp. 1–36, 1998.

## **Chapter 13**

## **Future of Image Registration**

Jasjit Suri,<sup>1</sup> Baowei Fei,<sup>2</sup> David Wilson,<sup>2</sup> Swamy Laxminarayan,<sup>3</sup> Chi-Hsiang Lo,<sup>4</sup> Yujun Guo,<sup>5</sup> Cheng-Chang Lu,<sup>5</sup> and Chi-Hua Tung<sup>6</sup>

#### **Future Application of Image Registration** 13.1

Image registration will have more and more applications in the future. Below are a few predictions on where image registration will lead to in the next few years, and where to expect significant progress.

#### 13.1.1 **Small Animal Imaging**

Small animal imaging is a fast-growing field that has numerous applications in the studies of functional genomics, the biology of disease, and therapeutics. Since commonly, functional imaging modalities such as single photon emission tomography (SPECT), positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) have little anatomic information, images acquired from computed tomography (CT) or MRI are used to provide structural identification and localization of organs/regions of interest and may also provide additional diagnostic information. Automatic image registration and fusion visualization methods will be very useful for this new and important application.

<sup>&</sup>lt;sup>1</sup>Senior IEEE Member, CWRU, Cleveland, USA.

<sup>&</sup>lt;sup>2</sup>Biomedical Engineering Department and Department of Radiology, CWRU, Cleveland, USA. <sup>3</sup>Idaho State University, Pocatello, ID.

<sup>&</sup>lt;sup>4</sup>National Ilan University, Taiwan.

<sup>&</sup>lt;sup>5</sup>Kent State University, USA.

<sup>&</sup>lt;sup>6</sup>Philips Medical Systems, USA.

## 13.1.2 Perfusion Studies

Perfusion imaging is likely to have many clinical applications. For example, cardiac MR imaging is progressing fast and the applications in tumor metabolism and angiogenesis is driving advances in MR imaging for oncology. Image registration will be essential, enabling technologies for perfusion imaging where patients may not be able to maintain the same position during long dynamic studies.

## 13.1.3 Registration for Image-Guided Interventions

Interventional MR, CT, X-ray fluoroscopy, and ultrasound, as well as optical images from endoscopes, microscopes, and arrays of free-standing cameras are used for image guided procedures. However, many image-guided surgery systems are currently restricted to applications in which patient anatomy can be treated as a rigid body. These technologies have great potential in soft tissues away from the bone. Registration methods could be used to update the spatial information in accurate and detailed representations of the patient, generated from preoperative images. This information could be incorporated into interventional procedures that often use incomplete and much lower quality information from intraoperative images.

# 13.1.4 Registration of Electronic Data Set with Anatomical Images

Multimedia electronic data sets can be incorporated with radiological images as well as other context-based information: for example, registration of EEG with MR images. Integrating this information and relating it to atlas data could be achieved transparently with the potential for improved diagnosis and decision support.

## 13.1.5 Deformation Fields Generated by Nonrigid Registration

Nonrigid image registration methods have great potential beyond simply lining up images. The deformation field produced by nonrigid registration algorithms can quantify normal development and contribute to an understanding of disease processes and aging. Nonrigid registration algorithms will be reliable enough for clinical applications and provide valuable tools for diagnosis and for monitoring disease progression.

## 13.1.6 Combination of Registration and Segmentation

Good segmentation can be achieved by lining up images to an atlas using the image registration algorithm. Labeled structures in the atlas can then be used to split up the images into anatomical and pathological components for visualization or quantification. For example, registration and segmentation of plaque images may allow the detecting of much of more subtle changes.

## 13.1.7 Registration to Data Acquisition

Registration is beginning to be used to improve data acquisition. For example, online registration can be used to dynamically adjust slice position in MR scans to compensate prospectively for motion correction. Spectroscopic or perfusion acquisitions can be defined to interrogate specific tissues of interest, delineated in a previously acquired high resolution image, rather than a fixed region relative to the scanner. Specific tissue regions could be followed as the patient moves or is repositioned. These applications are likely to grow as algorithms become faster and scanner computing power increases.

# 13.2 A Multiresolution Approach to Medical Image Registration Using Binning Technique

Medical image registration has been applied to the diagnosis of cardiac studies and a variety neurological disorders including brain tumors. Image registration, or alignment, is the process of aligning two images so that corresponding features of the images can be easily related. Registration using different modalities is widely used in many medical applications. In practice, the complementary information acquired from the registration of multiple imaging modalities can be used for medical diagnosis and treatment planning. In recent years many registration algorithms for medical imaging have been designed. Typically an algorithm falls into one of three categories: algorithms that use a landmarkbased method, algorithms that use a surface-based method, or algorithms that work directly on the image intensities (voxel-based). Maintz *et al.* [1] gave a survey of registration methods.

For automated registration a quality measure of the registration is necessary in order to find an optimal registration. Maximization of mutual information (MI) of voxel intensities, the registration method independently proposed by Wells *et al.* [2] and Maes *et al.* [3], is one of the most popular registration methods for three-dimensional multimodality medical image registration. This method measures the statistical dependence between the image intensities of corresponding voxels in two images; this statistical dependence is maximal when the images are totally aligned.

Intensity-based methods regard all voxels in the images as independent, thus, anatomical information is not taken into consideration. Maurer *et al.* [4] and Gall *et al.* [5] exploited landmark-based methods. Audette *et al.* [6] gave an overview on surface registration techniques. Landmark-based methods and surface-based methods utilize features extracted from the images. The required preprocessing is usually time-consuming and the accuracy of the registration is dependent on the correctness of the landmark or surface extraction.

We have developed a two-stage method, which is both feature-based and intensity-based. Three binning methods were utilized and the performance of each is compared in this chapter. In the first stage, we segment the images using region-growing. Then we perform one of the three binning methods on the full foreground before the down-sampled images are registered. In the second stage, the results from the first stage are taken as the starting point for the registration of the full original images. Experiments show that this new two-stage method gives improved accuracy without loss of speed, compared to multiresolution registration without bin preprocessing. Of the three binning techniques used, the nonlinear binning method gave the best performance. Normalized mutual information (NMI) is used as the similarity measure, and downhill simplex method is taken as the optimization method due to its quickness in practice.

## 13.2.1 Image Registration Using Binning Technique

Registration based on maximization of mutual information uses an iterative approach in which an initial estimate of the transformation is gradually refined.

One of the difficulties with this approach is that it can converge to a local optimum instead of a global optimum. Using multiresolution in conjunction with maximization of mutual information is very helpful when tackling this problem. The work of Maes *et al.* [8], Studholme *et al.* [7] and Pluim *et al.* [9, 10], has proved this. The idea of a multiresolution hierarchical approach is to register the coarse (low resolution) image first and then to use the result as the starting point for finer (high resolution) image registration, and so on.

In order to use NMI, an estimation of the intensity distribution is required. There are a couple of methods used to estimate the intensity distribution of an image. Colligon *et al.* [11] used joint entropy as the registration criterion. Viola [12] obtained the estimation by Parzen, windowing an intensity distribution. Camp *et al.* [13] proposed a binning method for registration using normalized mutual information. The image intensities are assigned to a histogram bin by a binning technique. The most commonly used binning method is equidistant binning. With equidistant binning, once the bin number is given, the intensities range assigned to each bin is also determined, after the overall image intensity range is distributed evenly among all the bins. The weakness of the equidistant binning method is that it totally ignores the anatomical information of the image. From typical histograms of CT and MR images, as shown in Fig. 13.1 and Fig. 13.2, we can spot the same property: a giant peak around the intensities of the background region. In our approach, we use region-growing to separate the



Figure 13.1: A typical histogram for a CT image.



Figure 13.2: A typical histogram for an MR image.

background region first. Then all the background region voxels are put into one bin, while the foreground region voxels are binned using a binning technique. Two-level multiresolution registration method is applied next. For binning we have experimented three techniques.

#### 13.2.1.1 Region-growing

Region-growing is an approach to image segmentation that has received considerable attention in the computer vision segment of the artificial intelligence community [14]. The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighboring voxels that have properties similar to the seed [15].

In our approach, a seed point is selected automatically near the left-upper corner of the given CT and MR images. This is based on the observation that there is always a large background area and that the object is always centered.

This seed point is used to begin region-growing for the background. There are two criteria for a voxel to be annexed to the background: (1) the absolute intensity difference between the voxel and the seed must be less than a threshold; and (2) the voxel must be adjacent to at least one voxel in the background. The threshold for the region is determined by the histogram, see Figs. 13.1 and 13.2. In the figures the spike at the left delimits the background and so the valley to

its right is picked as the threshold. Adjacency includes the eight directions: N, S, E, W, NE, NW, SE, and SW.

The region-growing procedure for the background/foreground segmentation is fully automatic and takes about 2% of the total registration time.

#### 13.2.1.2 Three Binning Techniques

The first binning technique is a binarization approach. All the background voxels are put into one bin and the remaining voxels (foreground) are put into the other bin. Lo *et al.* [16] developed this approach. The binarized, 2-bin images, see Fig. 13.3, are down-sampled by a factor of 2 along the x, y, and z axis directions. The down-sampled images are then used as the input to the first level of the multiresolution registration. The result of the first level provides the initial estimate for the second level, where the registration of the original images is performed.

The second binning technique, linear binning, is based on equidistant binning. All background voxels are put into one bin, then equidistant binning is applied to the remaining voxels, which assigns them to the rest of the bins. The binned images are then down-sampled and two-level registration is performed.

The third technique is a nonlinear binning approach [25]. The background of the image is first segmented into one bin and the remaining voxels are assigned to the remaining (k - 1) bins by *k*-means clustering [23].



Figure 13.3: A CT image (upper) and an MR image (lower) are shown with the original image on the left and the binarized image on the right.

### 13.2.1.3 K-means Clustering

In a digital image, regions with the same structure may have corresponding spikes in their intensity distributions. Intervals of the intensity distribution are more likely to have a higher variance if the same structure region is not in the same interval. For example, in standard CT or MR images, which in general contain different structure regions, such as background, cortical bone, white matter, gray matter, and cerebrospinal fluid (CSF). In the joint-histogram we would like to see the same structure is assigned to the same bin, i.e., for each bin to have a high probability of the same structure so that less dispersion in the joint-histogram is achieved. The method proposed here is to make all the voxels of background map to one bin by background segmentation using region growing, and have the remaining voxels map to the rest of bins (with a variable bin size for each bin) by k-means clustering, i.e., minimizing the variance of intensities within each bin [22].

Following is the *k*-means clustering algorithm used:

- 1. Initially partition the image voxels into *k* bins.
  - (1a) Put all the background voxels into bin 0.
  - (1b) Calculate the step size for the other k-1 bins using  $\frac{MaxIntensity-MinIntensity}{k-1}$ . Each bin will be assigned all voxels whose intensity falls within the range of its boundary.
  - (1c) Calculate the centroid of each bin.
- 2. For each voxel in the image, compute the distances to the centroids of its current, previous, and next bin, if it exists; if it is not currently in the bin with the closest centroid, switch it to that bin, and update the centroids of both bins.
- 3. Repeat step 2 until convergence is achieved; that is, continue until a passthrough all the voxels in the image causes no new assignments, or until a maximum number of iterations is reached. The maximum number of iterations was set to 500.

### 13.2.1.4 Normalized Mutual Information

Mutual information (MI) can be thought of as a measure of how well one image explains the other, and when maximized indicates optimal alignment.

Mutual information I(A, B) was proposed for intermodality medical image registration by several researchers [2, 3, 11]. The formula to compute mutual information is:

$$I(A, B) = \sum_{a} \sum_{b} P_{AB}(a, b) \log \frac{P_{AB}(a, b)}{P_{A}(a) \cdot P_{B}(b)}$$

where  $P_A(a)$ ,  $P_B(b)$  denote the marginal probability distributions of the two random variables, *A* and *B*, and  $P_{AB}(a, b)$  is their joint probability.

The formulation of normalized mutual information (NMI) as described by Studholme *et al.* [17] is used:

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)}$$

where

$$H(A) = -\sum_{a} P_A(a) \log P_A(a)$$

and

$$H(A, B) = -\sum_{a} \sum_{b} P_{AB}(a, b) \log P_{AB}(a, b)$$

H(A) and H(B) are the entropies of A and B, and H(A, B) is their joint entropy. NMI, devised to overcome the sensitivity of MI to change in image overlap, has been shown to be more robust than standard mutual information [17]. Image registration using NMI is performed in the following manner:

- 1. Take one of the two input images as floating image, the other as reference image.
- 2. Choose starting parameters for the transformation.
- 3. Apply the transformation to the floating image. Evaluate the NMI between reference image and transformed floating image.
- 4. Stop if convergence is achieved. If not, choose a new set of parameters, repeat steps 3 and 4.

When the registration parameters (three translations and three rotations) are applied to the floating image, the algorithm iteratively transforms the floating image with respect to the reference image while making the NMI measure calculated from the voxel intensities optimal. While all samples are taken at grid points of the floating image, their transformed position will, in general, not coincide with a grid point of the reference image and interpolation of the reference image is needed to obtain the image intensity value at this point.

The NMI registration criterion states that the images are geometrically aligned by the transformation  $T_{\lambda}$  for which NMI(A, B) is maximal. (The registration parameters are denoted by  $\lambda$  and  $T_{\lambda}$  is the transformation based on the six registration parameters given above.) To decide if the NMI is optimal, algorithms to measure optimality are applied. If the NMI is not optimal, a new set of parameters will be chosen and evaluated. The Downhill simplex is used to determine optimality. This method, as shown by NeIder and Mead [19, 20], only requires evaluation of the cost function, the derivative computation need not be redone.

#### 13.2.1.5 Two-Stage Multiresolution Registration

Multiresolution approach [7–10] is widely used in medical image registration due to the following two features:

- 1. Methods for detecting optimality cannot guarantee that a global optimal value will be found. Parameter spaces for image registration are usually not simple. There are often multiple optima within the parameter space, and registration can fail if the optimization algorithm converges to the wrong optimum.
- 2. Time to evaluate the registration criterion is proportional to the number of voxels: Medical images consist of a large number of voxels. During registration the main portion of computational time is consumed by the resampling voxels of the floating image with respect to the reference image according to actual geometrical transformation.

The idea of a multiresolution hierarchical approach is to register a coarse (low resolution) image first and then to use the result as the starting point for finer (high resolution) image registration, and so on [24]. In practice, the multiresolution approach proves to be helpful. It can improve the optimization speed, improve the capture range, and the algorithm is relatively robust.

In the first level of our two-stage multiresolution registration method, one of the three binning techniques, described above, is applied to the segmented images before they are down-sampled and the registration is performed on the coarser level. Nearest neighbor interpolation method is used in this level. In the second level, the original images are registered and a trilinear partial volume interpolation method [18] is employed. This process is done for each of the three binning techniques.

### 13.2.2 Experiments and Discussion

#### 13.2.2.1 Data Set

This study involved the data sets of seven patients, each consisting of CT and a subset of six Magnetic-Resonance (MR) volumes (spin-echo T1, PD, T2 and the rectified version of each of these three). The MR-T1-rectified image was not available for patient six. Thus, 41 image pairs were available to be registered. These images were provided by *The Retrospective Registration Evaluation Project* database maintained by J. Michael Fitzpatrick from Vanderbilt University, Nashville, TN, USA [26]. For each patient data set, all CT images were registered to the MR image using the MR image as the reference image.

All image pairs were registered using the maximization of NMI. Registration transformation was limited to six-parameter rigid-body type (three translations and three rotations) transformations. Registrations were conducted on a PC, having a 2.4 Ghz Intel Pentium 4 processor, and a 512MB DDR SDRAM memory.

The experiments were performed using a two-stage multiresolution approach described above. When the binarization approach was used, the foreground/foreground bin was given an additional weight when the joint histogram was computed in the first level. The additional weight was heuristically determined using the ratio of foreground to background voxels. The number of histogram bins used to compute the normalized mutual information criterion was set to 256 bins for both CT and MR.

The Downhill-simplex optimization method was used throughout the experiments. Optimization for each pair of images started from the same initial position with all three translation offsets set to zero millimeters and all three rotation angles set to zero degrees. Convergence was declared when the fraction difference between the lowest and the highest function value evaluated at the vertices of the simplex was smaller than  $10^{-5}$ .

#### 13.2.2.2 Results

All 41 CT-MR image pairs were used in the experiments. All CT images were registered to the MR images using the MR as the reference image. A typical superposition, before and after registration, of CT-MR images is shown in Fig. 13.4.

Observed from the registration results using the binarization approach, the translation parameter ranged from -26.42 to 8.37 mm and the angle parameter ranged from  $-4.3^{\circ}$  to  $1.59^{\circ}$ . The running time for each CT-MR pair took 10.4 to 27.8 minutes. The average time for all 41 CT-MR pairs registration is 18.4 minutes.

From our results, using the nonlinear binning method, the running time for each CT-MR pair took 7.6 to 18.3 minutes. The average time for all 41 CT-MR pairs registration was 12.2 minutes. When we compare the registration time using the nonlinear binning with the registration time using the binarization approach, we observe the average speedup when using nonlinear method is 51%. Figure 13.5 shows the timing results.

The running time for the approach using the linear binning technique is between the above two methods.

For each experiment the registration solution was obtained using fiducial markers as provided by *The Retrospective Registration Evaluation Project*. This was used as the gold standard to evaluate registration accuracy.



Figure 13.4: A typical superposition of CT-MR images. The left images are before registration and the right ones are after.



Figure 13.5: Time required to perform registration for each of the 41 CT-MR pairs: (a) the binarization approach and (b) the nonlinear binning method.

#### **13.2.2.3** Accuracy

The accuracy of all 41 experiments with respect to the fiducial-mark-base gold standard can be found on the web (see http://www.vuse.vanderbilt.edu/ image/registration/).

In addition, we compare the results of our approaches to those of several other approaches published in the literature. The comparison is based on a methodology proposed by West and Fitzpatrick [21], who let selected researchers access a standard set of image volumes to be registered. They also act as a repository for the ideal registration transformations (gold standard) acquired by a prospective method using fiducial markers. These markers are removed before the volumes are disclosed to the investigators, who then face a retrospective blind registration task. After registration, they email back a set of transformation parameters that are compared to the gold standard.

Tables 13.1 and 13.2 show the median and maximum accuracy reached by the investigators taking part in that project [21]. All errors are in units of mm. The method using binarization approach is labelled LO1, while the method using nonlinear binning technique is labelled LO2. We observed that both techniques give impressive results for CT-MR registration.

#### **13.2.2.4** Optimization Steps

In the downhill simplex optimization method, the number of optimization steps is measured by the number of times the cost function is called. The mean number and the standard deviation of optimization steps for the three binning techniques, for each patient data set, is compared for CT-MR registration in Table 13.3 and Table 13.4.

The binarization approach needs the least number of optimization steps in the first level (Table 13.3). The methods using linear binning and nonlinear binning need 30% to 102% more steps than the binarization approach. Of the three binning techniques, the binarization approach has the most stable performance for all the patients. The reason is that binarization can give an extreme blurring of the images and so eliminates local optima. The performances of linear and nonlinear binning are pretty much the same, while nonlinear binning is better in four out of seven patients.

Table 13.1: Median error for CT-to-MR registration in 'mm' between the prospective gold-standard and several retrospective registration techniques. The label "rect." indicates that the MR image was corrected for geometrical distortion before registration. The result of the technique using the binarization approach is labelled LO1, while the one using the nonlinear binning method is labelled LO2.

	BA	СО	$\mathbf{EL}$	HA	HE	HI	MAI	MAL	NO	PE	RO1	LO1	LO2
CT-T1	1.6	1.5	1.6	3.4	1.4	1.2	5.1	4.3	3.3	2.7	4.2	1.26	1.24
CT-PD	1.9	1.5	2.0	3.1	2.4	1.9	4.1	4.0	7.8	1.9	4.5	1.67	1.90
CT-T2	2.5	1.5	1.6	4.2	4.7	1.5	3.9	5.0	3.9	2.5	4.5	1.64	1.47
CT-T1-rect.	1.4	0.7	0.9	3.3	1.0	0.7	4.9	5.4	3.4	2.2	5.9	0.65	0.96
CT-PD-rect.	1.7	0.8	1.1	3.0	1.7	0.7	3.0	4.0	4.6	2.1	5.9	0.85	0.90
CT-T2-rect.	2.1	0.8	1.6	3.5	1.6	0.8	4.3	5.3	4.2	2.9	5.5	0.81	0.89

Table 13.2: Maximum error for CT-to-MR registration. See notes in Table 13.1.

	BA	СО	EL	HA	HE	HI	MAl	MAL	NO	PE	RO1	LO1	LO2
CT-T1	6.4	6.7	6.0	51.8	11.0	2.8	12.8	61.4	10.4	7.3	26.0	3.15	2.76
CT-PD	6.9	3.6	6.6	49.6	10.4	4.1	19.0	59.0	13.9	4.3	25.9	3.64	3.91
CT- T2	9.1	3.4	4.1	50.6	13.6	4.2	6.3	59.5	9.7	7.2	26.7	3.64	4.64
CT-T1-rect.	5.8	3.8	2.6	48.2	2.1	2.3	14.2	60.9	9.6	5.9	27.8	1.98	1.95
CT-PD-rect.	5.9	2.5	5.3	45.9	3.7	2.3	9.9	62.7	11.5	4.6	27.5	2.13	1.81
CT-T2-rect.	7.4	4.3	5.2	49.1	14.3	3.0	6.5	63.2	10.2	9.0	27.1	2.65	2.05

Table 13.3: Number of optimization steps used by each of the three binning techniques for first level.

	Binari	zation	Linear	binning	Nonlinear binning		
Data Set	Mean	δ	Mean	δ	Mean	δ	
Patient 1	98	21.85	165	51.66	139	17.72	
Patient 2	140	22.98	244	65.85	208	56.72	
Patient 3	128	17.51	244	66.02	251	23.25	
Patient 4	163	38.94	246	68.41	211	50.42	
Patient 5	151	42.66	195	24.28	236	52.2	
Patient 6	153	32.69	199	53.76	268	56.55	
Patient 7	112	27.12	228	67.98	226	71.77	

	Bina	rization	Linear	binning	Nonlinear binning		
Data Set	Mean	δ	Mean	δ	Mean	δ	
Patient 1	127	24.31	123	18.81	113	17.16	
Patient 2	171	37.50	129	25.3	120	27.76	
Patient 3	218	43.81	212	36.75	188	42.06	
Patient 4	185	86.14	160	50.48	149	56.73	
Patient 5	197	46.25	187	70.06	157	50.12	
Patient 6	274	110.03	198	55.19	149	41.49	
Patient 7	260	74.11	163	34.40	151	39.41	

Table 13.4: Number of optimization steps used by each of the three binning techniques for second level.

When it comes to the second level (Table 13.4), the binarization approach turns out to be the worst method for all seven patients. Not only the number of optimization steps, but also the standard deviation for each patient data sets is inferior. The nonlinear binning based method performs the best, on average, while linear binning is in the middle.

The binarization approach can accelerate the registration in the coarser level, where the image volumes are greatly decreased and so the time to do the entire registration is reduced, as seen in the results section. When applied to the original images, binarization requires more steps to be taken to make up for the oversimplified bin structure in the previous level. This results in a longer registration time than the other two techniques.

The nonlinear binning method is slightly better than the linear method in the first level, but it dominates in the second level. Because the nonlinear binning method takes anatomical information into consideration, and gives more details than the linear binning, its superiority over the other two is no surprise. Running times given in Section 13.2.2.2 also proves this. The nonlinear binning technique can accelerate the convergence of the registration thereby reducing total registration time, compared to the binarization approach or linear binning, without loss of accuracy.

#### 13.2.3 Summary

In practice, the multiresolution approach has proven to be helpful for multimodality image registration. In our approach, we implemented a two-stage multiresolution approach for CT-MR registration using normalized mutual information. Before the registration starts, the background is segmented in both images using region-growing. Then we performed one of three binning techniques on the foreground details, while the background is put into one bin. Three binning techniques were investigated in this manner.

Our results have shown that all three approaches can reach a subvoxel accuracy with no loss of speed. The approach using the nonlinear binning technique shows improvement in accuracy and speed, compared to the other two binning techniques, since it can achieve less dispersion in the joint-histogram computation.

## References

- Maintz, J., and Viergever, M., A survey of medical image registration, Med. Image Anal., Vol. 2, No. 1, pp. 1–36, 1998.
- [2] Wells, W., Viola, P., Atsumi, H., Nakajima, S., and Kikinis, R., Multimodal volume registration by maximization of mutual information, Med. Image Anal., Vol. 1, No.1, pp. 35–51, 1996.
- [3] Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., and Suetens, P., Multimodality image registration by maximization of mutual information, IEEE Trans. Med. Imag., Vol. 16, No. 2, pp. 187–198, 1997.
- [4] Maurer, C., Fitzpatrick, J., Wang, M., Galloway, R., Maciunas, R., and Allen, G., Registration of head volume images using implantable fiducial markers, Medical Imaging 1997: Image Processing, Proc. SPIE, Vol. 3034, pp. 561–579, 1997.
- [5] Gall, K., and Verhey, L., Computer-assisted positioning of radiotherapy patients using implanted radiopaque fiducials, Med. Phys., Vol. 20, No. 4, pp. 1153–1159, 1993.
- [6] Audette, M., Ferrie, F., and Peters, T., An algorithmic overview of surface registration techniques for medical imaging, Med. Image Anal., Vol. 4, No. 3, pp. 201–217, 2000.
- [7] Studholme, C., Hill, D., and Hawkes, D., Automated 3D registration of MR and PET brain images by multi-resolution optimization of voxel similarity measures, Med. Phys., Vol. 24, No. 1, pp. 25–36, 1997.
- [8] Maes, F., Vandermeulen, D., and Suetens, P., Comparative evaluation of multiresolution optimization strategies for multimodality image registration by maximization of mutual information, Med. Image Anal., Vol. 3, No. 4, pp. 373–386, 1999.
- [9] Pluim, J., Maintz, J., and Viergever, M., A multiscale approach to mutual information matching, Medical Imaging 1998: Image Processing, SPIE Vol. 3338, pp. 1334–1344, 1998.

- [10] Pluim, J., Maintz, J., and Viergever, M., Mutual information matching in multiresolution contexts, Image Vis. Comput., Vol. 19, No. 1–2, pp. 45–52, 2001.
- [11] Collignon, A., Maes, F., Delaere, D., Vandermeulen, D., Suetens, P., and Marchal, G., Automated multimodality image registration using information theory, In: Information Processing in Medical Imaging (IPMI'95), Bizais, Y., Barillot, C., and Di Paola, R. eds. Kluwer, Dordrecht, pp. 263– 274, 1995.
- [12] Viola, P., Alignment by maximization of mutual information, PhD Thesis, Massachusetts Institute of Technology, 1995.
- [13] Camp, J., and Robb, R., A novel binning method for improved accuracy and speed of volume image coregistration using normalized mutual information, In: Medical Imaging 1999: Image Processing Proc. SPIE, Vol. 3661, pp. 24–31, 1999.
- [14] Castleman, K., Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1996.
- [15] Gonzalez, R., and Woods, R., Digital Image Processing, 2nd edn., Prentice-Hall, Englewood Cliffs, NJ, 2002.
- [16] Lo, C., Guo, Y., and Lu, C., A binarization approach for CT-MR registration using normalized mutual information, In Proceedings of the 5th IASTED International Conference on Signal and Image Processing (SIP'03), pp. 496–501, 2003.
- [17] Studholme, C., Hill, D., and Hawkes, D., An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition, Vol. 32, No. 1, pp. 71–86, 1999.
- [18] Lehmann, T., Gonner, C., and Spitzer, K., Survey: Interpolation methods in medial image processing, IEEE Trans. Med. Imaging, Vol. 18, No. 11, pp. 1049–1075, 1999.
- [19] Press, W., Flannery, B., Teukolsky, S., and Vetterling, W., Numerical Recipes in C: The Art of Scientific Computing, 2nd edn., Cambridge University Press, Cambridge, U.K., 1999.

- [20] NeIder, J., and Mead, R., Comput. J., Vol. 7, pp. 308–313, 1965.
- [21] West, J., Fitzpatrick, J., Wang, M., Dawant, B., Maurer, C., Kessler, R., Maciunas, R., Barillot, C., Lemoine, D., Collignon, A., Maes, F., Suetens, P., Vandermeulen, D., vanden Elsen, P., Napel, S., Sumanaweera, T., Harkness, B., Hemler, P., Hill, D., Hawkes, D., Studholme, C., Maintz, J., Viergever, M., Malandin, G., Pennec, X., Noz, M., Maguire G., Pollack, M., Pelizzari, C., Robb, R., Hanson, D. and Woods, R., Comparison and evaluation of retrospective intermodality brain image registration techniques, J. Comput. Assist. Tomogr., Vol. 21, No. 4, pp. 554–566, 1997.
- [22] Knops, Z., Maintz, J., Viergever, M., and Pluim, P., Normalized mutual information based registration using k-means clustering based histogram binning, in Medical Imaging 2003: Image Processing, Proc. SPIE, Vol. 5032, pp. 1072–1080, 2003.
- [23] Anderberg, M., Cluster Analysis for Applications, Academic Press, New York, 1973.
- [24] Lester, H., and Arridge, S., A survey of hierarchical non-linear medical image registration, Pattern Recognition, Vol. 32, No. 1, pp. 129–149, 1999.
- [25] Lo, C., Medical image registration by maximization of mutual information, Ph.D. Dissertation, Kent Stete University., 2003.
- [26] http://www.vuse.vanderbilt.edu/image/registration

## **The Editors**



**Dr. Jasjit S. Suri** received his BS in Computer Engineering with distinction from Maulana Azad College of Technology, Bhopal, India, his MS in Computer Sciences from University of Illinois, Chicago, and Ph.D. in Electrical Engineering from University of Washington, Seattle. He has been working in the field of Computer Engineering/Imaging Sciences for 20 years. He has published more than 125 technical papers in body imaging. He is a lifetime member of research engineering societies Tau-Beta-Pi, Eta-Kappa-Nu, and Sigma-Xi and a member of NY Academy of Sciences, Engineering in Medicine and Biology Society (EMBS), SPIE, ACM, and is also a Senior Member of IEEE. He is in the editorial board/reviewer of several international journals such as *Real Time Imaging, Pattern Analysis and Applications, Engineering in Medicine and Biology, Radiology, Journal of Computer Assisted Tomography, IEEE Transactions of Information Technology in Biomedicine and IASTED Board.* 

He has chaired image processing tracks at several international conferences and has given more than 40 international presentations/seminars. Dr. Suri has written four books in the area of body imaging (such as Cardiology, Neurology, Pathology, Mammography, Angiography, Atherosclerosis Imaging) covering Medical Image Segmentation, image and volume registration, and physics of medical imaging modalities like MRI, CT, X-ray, PET, and Ultrasound. He also holds several United States Patents. Dr. Suri has been listed in Who's Who seven times, is a recipient of President's Gold medal in 1980, and has received more than 50 scholarly and extracurricular awards during his career. He is also a Fellow of American Institute of Medical and Biological Engineering (AIMBE) and ABI. Dr. Suri's major interest are Computer Vision, Graphics and Image Processing (CVGIP), Object Oriented Programming, Image Guided Surgery, and Teleimaging. Dr. Suri had worked for Philips Medical Systems and Siemens Medical Research Divisions. He is also a Visiting Professor with Department of Computer Science, University of Exeter, Exeter, England. Currently, Dr. Suri is with JWT, Inc., as Director of Biomedical Engineering Division (in Opthalmology Imaging) in conjunction with Biomedical Imaging Laboratories, Case Western Reserve University, Cleveland.



**Dr. David Wilson** is a Professor of Biomedical Engineering and Radiology, Case Western Reserve University. He has research interests in image analysis, quantitative image quality, and molecular imaging, and he has a significant track record of federal research funding in these areas. He has over 60 refereed journal publications and has served as a reviewer for several leading journals. Professor Wilson has six patents and two pending patents in medical imaging. He has been active in the development of international conferences; he was Track Chair at the 2002 EMBS/BMES conference, and he was Technical Program Co-Chair for the 2004 IEEE International Symposium on Biomedical Imaging.

Professor Wilson teaches courses in biomedical imaging and biomedical image processing and analysis. He has advised many graduate and undergraduate students, all of whom are quite exceptional, and has been primary research advisor for over 16 graduate students since starting his academic career. Prior to joining CWRU, he worked in X-ray imaging at Siemens Medical Systems at sites in New Jersey and Germany. He obtained his Ph.D. from Rice University. Professor Wilson has actively developed biomedical imaging at CWRU. He has led a faculty recruitment effort, and he has served as PI or has been an active leader on multiple research and equipment developmental awards to CWRU, including an NIH planning grant award for an In Vivo Cellular and Molecular Imaging Center and an Ohio Wright Center of Innovation award. He can be reached at dlw@po.cwru.edu.



**Dr. Swamy Laxminarayan** currently serves as the Chief of Biomedical Information Engineering at the Idaho State University. Previous to this, he held several senior positions both in industry and academia. These have included serving as the Chief Information Officer at the National Louis University, Director of the Pharmaceutical and Health Care Information Services at NextGen Internet (the premier Internet organization that spun off from the NSF sponsored John von Neuman National Supercomputer Center in Princeton), Program Director of Biomedical Engineering and Research Computing and Program Director of Computational Biology at the University of Medicine and Dentistry in New Jersey, Vice-Chair of Advanced Medical Imaging Center, Director of Clinical Computing at the Montefiore Hospital and Medical Center and the Albert Einstein College of Medicine in New York, Director of the VocalTec High Tech Corporate University in New Jersey, and the Director of the Bay Networks Authorized Center in Princeton. He has also served as an Adjunct Professor of Biomedical Engineering at the New Jersey Institute of Technology, a Clinical Associate Professor of Health Informatics, Visiting Professor at the University of Brno in Czech Republic, and an Honorary Professor of Health Sciences at Tsinghua University in China.

As an educator, researcher, and technologist, Prof. Laxminarayan has been involved in biomedical engineering and information technology applications in medicine and health care for over 25 years and has published over 250 scientific and technical articles in international journals, books, and conferences. His expertise are in the areas of biomedical information technology, high performance computing, digital signals and image processing, bioinformatics and physiological systems analysis. He is the coauthor of the book on State-of-the-Art PDE and Level Sets Algorithmic Approaches to Static and Motion Imagery Segmentation published by Kluwer Publications and the book on Angiography Imaging: State-of-the-Art Acquisition, Image Processing and Applications Using Magnetic Resonance, Computer Tomography, Ultrasound and X-ray, Emerging Mobile E-Health Systems, published by the CRC Pres and two volumes of the Handbook of Biomedical Imaging to be published by the Kluwer publications. He has also authored as the editor/coeditor of 20 international conferences and has served as a keynote speaker in international conferences in 43 countries.

He is the Founding Editor-in-Chief and Editor Emeritus of the IEEE Transactions on Information Technology in Biomedicine. He served as an elected member of the administrative and executive committees in the IEEE Engineering in Medicine and Biology Society and as the Society's Vice President for 2 years. His other IEEE roles include his appointments as Program Chair and General Conference Chair of about 20 EMBS and other IEEE Conferences, an elected member of the IEEE Publications and Products Board, member of the IEEE Strategic Planning and Transnational Committees, Member of the IEEE Distinguished Lecture Series, Delegate to the IEEE USA Committee on Communications and Information Policy (CCIP), U.S. Delegate to the European Society for Engineering in Medicine, U.S. Delegate to the General Assembly of the IFMBE, IEEE Delegate to the Public Policy Commission and the Council of Societies of the AIMBE, Fellow of the AIMBE, Senior Member of IEEE, Life Member of Romanian Society of Clinical Engineering and Computing, Life Member Biomedical Engineering Society of India, and US Delegate to IFAC and IMEKO Councils in TC13. He was recently elected to the Administrative Board of the International Federation for Medical and Biological Engineering, a worldwide organization comprising 48 national members, overseeing global biomedical engineering activities. He was also elected to serve as the Publications Co-Chairman of the Federation.

His contributions to the discipline have earned him numerous national and international awards. He is a Fellow of the American Institute of Medical and Biological Engineering, a recipient of the IEEE 3rd Millennium Medal, a recipient of the Purkynje award from the Czech Academy of Medical Societies, a recipient of the Career Achievement Award, numerous outstanding accomplishment awards, and twice recipient of the IEEE EMBS distinguished service award. He can be reached at s.n.laxminarayan@ieee.org.

## Index

AC-PC referential, 274, 276 Active contours, 48-55, 426. See also Deformable models; Geodesic active contours/deformable models: Geometric active contours/deformable models gradient-based level set, 51-52 registering for multimodal segmentation, 69 - 70Active deformation models (ADM), 455 Active shape model (ASM), 454 Advection term, 200 Affine registration atlas-based segmentation and, 452, 455, 456-457 global, 313 nine degree-of-freedom algorithm, 443 Affine transformations, 6 CL-TPS image registration and, 235 global, 302 retinal image registration and, 158-160 AffineTransformOp, 159 Alzheimer's disease, 72 Amodal contours, 56 Analysis of variance (ANOVA), 204, 207t, 211Anatomical landmarks, 5 Anatomical structures, labeling of, 436 Aneurysms. See Brain aneurysms ANIMAL, 280 Aortic abdominal aneurysms, 190 Approximate nearest neighbors (ANN) algorithm, 200 Artifacts, 458-459 Atherosclerotic aneurysms, 186 Atlas-based registration, 7-8, 342 elastic image registration and, 365, 376

electronic, 273-274 level set segmentation and, 63-65 Atlas-based segmentation, 435–474 with an average shape atlas, 449, 450f, 454-459, 466-467, 468 bias from structure shape, 463-465 bias from structure volume, 462-463 binary performance model for, 469, 470, 472 comparison of selection strategies, 465 - 467concepts in, 435-437 fundamentals of, 439-449 image acquisition and processing, 437-438 with the most similar atlas, 449, 450f, 451-454, 466, 468-469 multilabel performance model, 469, 470-471, 472with multiple atlases, 449, 450f, 459-460, 466-467, 468-472, 473 quantifying accuracy of, 460-467 with a single, individual atlas, 449, 450f, 451, 465-467, 468 Atlas image, 442, 443, 449 Atlas of Talairach, 274 Automatic registration, 6 Balloons method, 277 Bayesian-based regularizers, 57-58 Bayesian fractal prior, 354 Biharmonic equation, 278 Bilinear interpolation, 160-161, 162, 174, 176, 179Binarization approach to binning, 541, 545, 546, 548 - 550Binary classifier performance model, 469, 470, 472

Binning technique, 551 accuracy of, 548 binarization approach, 541, 545, 546, 548-550 equidistant, 539 k-means clustering and, 542 linear, 541, 548, 550 nonlinear, 541, 546, 547f, 548, 550 Biomedical applications, 342. See also Elastic image registration Bland-Altman analysis, 87, 204, 208, 209f, 211 Block-matching minimization scheme, 284 Bone imaging, 55 Brain aneurysms, 185-213 atherosclerotic, 186 clinical protocol for, 202-203 computerized protocol for, 203 fusiform, 186, 187f image features, 193-194 implicit deformable models for, 190-191 mycotic, 186 non-saccular, 186, 187f planning interventions for, 188-190 prevalence of, 187 saccular (berry), 186, 187f segmentation of, 193-202 traumatic, 186 Brain imaging, 230 artificially deformed, 374-376 atlas-based segmentation of, 437-438, 439f, 440f, 441f, 443, 444f, 445f, 451, 452-453, 454, 459, 460, 461, 463, 464-465, 473 cross entropy, symmetric divergence, and, 402elastic image registration and, 374-376 elastic models and, 279 electronic atlases and, 273-274 geometric deformable models and, 79-81 in honeybees (see Honeybee brain images) hyperelastic warping and, 502-504, 505f I-SICLE algorithm and, 234 level set segmentation of, 62, 63, 66, 70-73, 78-79 in mice, 502-504, 505f non-rigid registration and, 273 Romeo and, 293-294, 313-315 3D non-rigid body registration and, 122 Brain tumors, 262-264, 265, 266f, 537 BrainWeb phantom, 72, 301-306, 316, 454 Breast imaging, 54, 122, 459 Brent line minimizations, 34, 400 Brightness constancy constraint, 286-287 Brownian motion, fractal, 354, 355 B-spline(s), 344 atlas-based segmentation and, 446, 447

elastic image registration and, 359, 368, 369, 370, 371 fast calculations for, 371 hyperelastic warping and, 506, 510 inverse consistent registration and, 227 multiresolution representation of, 369 non-rigid registration and, 276-277, 284 semi-local models and, 346 B-spline deformation model, 357, 360 B-spline interpolation, 344 Canny edge detectors, 78, 265, 267 Cardiac imaging, 537 atlas-based segmentation of, 454 hyperelastic warping and, 512-518 level set segmentation of, 54, 55, 62, 68, 73-77 neuractive pyramid and, 282 Cauchy-Green deformation, 490 Cauchy stress tensor, 493, 511 CCMC algorithm. See Connectivity consistent marching cubes algorithm Cerebral cortex imaging, 81-84 Cerebrospinal fluid (CSF), 542 intensity correction and, 289-290 level set segmentation and, 71, 78, 79, 81-84 Chromophor, 445, 447, 459 Circle of Willis, 186, 187f, 192, 202 CL-TPS image registration. See Inverse consistent landmark-based thin-plate spline image registration Clustering-based regularizers, 57 Coarse-to-fine method, 256, 350 Coiling, 212. See also Guglielmi detachable coils Color images, 157, 161, 174 Composite registration, 34 Computed tomography (CT), 2-3, 21 binning technique and, 539, 540, 542, 545, 546, 548, 549t, 551 elastic image registration and, 377 hyperelastic warping and, 502, 505-507 interventional, 536 level set segmentation and, 55, 63, 69-70, 74, 76,88 of lung cancer, 40f, 41 of pelvis and prostate, 109 small animal imaging and, 535 symmetric similarity cost functions and, 224 - 225Computed tomography angiography (CTA), 3 of brain aneurysms, 190, 191, 192-212 Computer-aided volumetric segmentation, 251 - 269shape editing and, 262, 263-264 shape voxel parametrizing and, 256-260 surface approximation and, 255-256

Conditional entropy, 156, 157, 397 Confocal laser scanning microscopy, 437–438, 447 Confusion matrix, 470-471 Conjugated gradients, 349, 362, 363f Connectivity consistent marching cubes (CCMC) algorithm, 80, 81 Continuum mechanical models, 226-227, 278, 280 Contours active (see Active contours) amodal, 56 inverse consistent registration and, 220 modal, 56 to volume from, 277-278 Control point(s) (CPs) atlas-based segmentation and, 446-447 computer-aided volumetric segmentation and, 256, 260-264 least squares computation of, 260-263 semi-local models and, 346 shape editing and, 263-264 Control point grid (CPG), 446-447, 456-457 Control point (CP) optimization, 122, 123-124, 136-137, 141 Control point (CP) selection, 122, 127-128, 129f, 137-138, 139, 140 Coronary artery imaging, 507-512 Correlation, 347. See also Cross-correlation Correlation coefficient (CC) in 3D non-rigid body registration, 132, 133f in 3D rigid body registration, 105-106, 107-108, 111, 120-121, 136, 139 Correspondence function, 339 COSMOS surface registration, 22, 24-26 Cost functions binning technique and, 544 CL-TPS image registration and, 235-236 computer-aided volumetric segmentation and, 253 data term in, 347-348 deformable image registration and, 488 elastic image registration and, 358-359 intensity variance, 225 inverse consistent registration and, 223-225 non-rigid registration and, 296, 297 purpose of, 347 regularization term in, 347, 348 Romeo and, 286, 289, 290-291, 292, 295 search strategy and, 343, 348-350 symmetric similarity, 223-225 UL-TPS image registration and, 234 Coupled-surfaces propagation, 81 Coupling-surfaces regularizers, 57, 58-59

Courant-Friedrichs-Levy (CFL) condition, 200

Crest lines, 276 Cross-correlation, 6, 154, 278, 280-281 Cross-entropy, 394, 395-427. See also Reversed-cross entropy implementation of, 397-402 maximization of, 394, 396, 397, 403-405, 406-417 minimization of, 394, 396, 401, 405-406, 417-422 numerical stability and, 401-402, 414-417, 419-422 optimization of, 394, 400, 401 registration experimental set-up for, 402-406 CT. See Computed tomography CTA. See Computed tomography angiography Curvature constraints, 201 Curve(s) Euclidean, 54 in non-rigid registration, 276-277 Curved (elastic) transformations, 6 Curve matching methods, 277 Data acquisition improvements in, 537 in 3D rigid body registration, 111-112 Data sets in binning technique, 545 for cross entropy and symmetric divergence, 402-403 electronic, 536 partial, 9, 10 for Romeo experimentation, 315-318 Data term, in cost functions, 347-348 Deformable image registration, 5, 487–521. See also Hyperelastic warping Deformable models. See also Active contours; Geodesic active contours/deformable models; Geometric active contours/deformable models; Implicit deformable models; Parametric deformable models active, 455 B-spline, 357, 360 elastic image registration and, 359-360 Deformation(s) atlas-based segmentation and, 452, 456-457 Cauchy-Green, 490 free-form, 277, 446 left ventricular, 512, 513, 514-516 thin-plate spline, 277-278, 316 Deformation fields, 536-537 Deformation gradient, 489 Deformation maps, 489, 490, 515 Demon's algorithm, 225, 278, 281, 349 DFD. See Displaced frame difference

Diabetic retinopathy, 152 Diffeomorphic transformations, 221, 226 Digital shapes, 254–255, 256–260 Digital subtraction angiography (DSA), 3, 190 Dirac measures, 66, 353 Direct registration, 406-407 Dirichlet problem, 234 Displaced frame difference (DFD), 282, 287, 295Distance-based registration techniques, 10-21 Doppler echocardiography, 3, 513 Drusen deposits, 152 DSA. See Digital subtraction angiography Dynamic programming, 349 Echo-planar imaging (EPI), 315-316 Eikonal equation, 200-201 Elastic image registration, 357-379. See also Romeo algorithm consequences of finite support, 370 cost functions in, 358-359 defined, 341 deformable model structure, 359-360 existence and unicity of, 361 experiments in, 373-377 implementation issues, 365-372 masking and, 371-372 optimization strategy for, 361-364 problem formulation, 358 semi-automatic, 364-365 size change and, 370-371 stopping criterion in, 371 Elastic models atlas-based segmentation and, 443-444 non-rigid registration and, 278-279 Electronic atlases, 273-274 Element inversion, 500 EM algorithm. See Expectation-maximization algorithm Entropy, 31, 52. See also Cross-entropy; Joint entropy; Reversed cross-entropy atlas-based segmentation and, 442-443 conditional, 156, 157, 397 information-theoretic, 442 minimization of, 397 retinal image registration and, 156, 171 Epilepsy, 3, 73 Equidistant binning technique, 539 Euclidean curve length, 54 Euclidean distance functions, 13, 16, 195, 353 Euclidean norms, 238, 243 Eulerian framework, 489 Euler-Lagrange equations, 67, 349 brain aneurysm treatment and, 197-198 deformable image registration and, 490, 496

geodesic active contours and, 54 landmark interpolation and, 353, 355 Exhaustive search method, 163, 176, 348-349 Expansion operator, 369 Expectation-maximization (EM) algorithm, 299f, 301 atlas-based segmentation and, 469, 470, 471-472 intensity correction and, 289-290, 297, 308 level set segmentation and, 78 Explicit derivatives, 367–368 Extrinsic markers, 345 Extrinsic registration, 4-5 Fast marching method, 77, 191, 200-201 Fast spline calculations, 371 Feature-based inverse consistent registration, 236 Feature-based registration, 10, 21-34, 343, 347. See also Mutual information maximization; Surface-based registration techniques binning technique in, 538 description of, 345 Feature correspondence, 21 Feature extraction, 21 Feature space, 343-345 Filter-based edge detection, 47 Filters and filtering Gabor, 344 Gaussian, 284, 288, 291, 292 Gaussian convolution, 51 Gaussian isotropic, 288 Kalman, 14, 276 recursive, 369 sequential spatial, 496–498 Finite deformation theory, 489-490 Finite difference relaxation method, 349 Finite element (FE) analysis, 349, 502, 506, 507, 508, 510, 511, 514-515 Finite element (FE) discretization, 494-495 Finite element (FE) meshes, 494-495, 503, 506 regular vs. irregular, 498-500 rezoning regular, 500-502 FLAIR magnetic resonance imaging (MRI), 36, 37f Floating image atlas-based segmentation and, 442, 447, 448, 455 binning technique and, 543-544 cross entropy, symmetric divergence, and, 395, 398 retinal image registration and, 157, 160, 161, 162, 165, 174

Floating volume, 31-32 multiple sclerosis studies and, 38-39 Romeo and, 316 3D non-rigid body registration and, 122, 124-125, 136-137 3D rigid body registration and, 105 Fluid models, 279-280, 443-444 Fluorescein images, 152 Forward transformations CL-TPS image registration and, 235-236 fluid models and, 280 inverse consistent registration and, 219-220, 222, 227-228, 240, 242, 243-244 Fourier series parameterization, 227-228, 280 Fourier transforms, 153, 154, 171, 344, 497 Fractional landmark interpolation, 354-357 Fractional orders, 354 Free-form deformations, 277, 446 Free-form surfaces, 21, 22 Freehand scanning, 75 Functional magnetic resonance imaging (fMRI), 3, 316, 318, 535 Fusiform aneurysms, 186, 187f Gabor filters, 344 Gateaux derivative, 491 Gaussian filters, 284, 288, 291, 292 convolution, 51 isotropic, 288 Gaussian kernels, 193, 195, 202, 282, 418, 426, 497Gaussian noise, 358 white, 354 Gaussian PDF models, 193, 199 Gaussian sensor models, 492 Gauss-Seidel iterative solver, 295, 297 GCP. See Grid closest point Genetic algorithms (GAs), 15-21, 34 computer-aided volumetric segmentation and, 253-254 in lung cancer studies, 41 major drawbacks of, 19-21 retinal image registration and, 162 Geodesic active contours/deformable models, 55, 66, 68, 77 brain aneurysms and, 190-191, 197-200, 208 properties of, 53-55 region-based information in, 198-200 RT3D ultrasound and, 84-86 speed terms of, 57 Geodesic active regions (GAR) brain aneurysms and, 191, 192, 193, 197-202 curvature constraints and, 201 numerical schemes for, 200

Geometric active contours/deformable models, 48, 51, 55, 65 speed terms of, 57 topology preserving, for brain reconstruction, 79-81 Geometric indexing, 22 Geometric landmarks, 5 Geometric methods, in non-rigid registration, 275-278, 284-285, 320 Geometric stiffness, 493, 495 Geometric transformations, 278, 283 Gerbil malleus, 502, 505-507 Glaucoma, 152, 166 Global affine registration, 313 Global affine transformations, 302 Global models, 346 Global optimum, 539 Global transformations, 6, 275 Good approximation properties, 352, 359 Gradient-based level set active contours, 51 - 52Gradient calculation, 368 Gradient descent algorithms, 349, 361-362, 363 Gray matter, 61, 71, 78, 79, 81-84, 225, 317, 542coupling-surfaces regularizers and, 58, 59geometric deformable models and, 81 intensity correction and, 289-290 multiple sclerosis and, 36 overlap with white matter, 313-315 Green images, 152, 157 Grid closest point (GCP), 15-21 Guglielmi detachable coils (GDC), 188, 189f, 193Hadamard transforms, 344 Hamilton-Jacobi equations, 51, 60, 200 Hard constraints, 495 Hausdorff measure, 61 Heart imaging. See Cardiac imaging Heaviside function, 66, 69 Hessian matrix, 193-194, 195, 367, 368 Hibernating myocardium, 512 Hierarchical finite element bases, 349 Histogram(s) joint (see Joint histograms) marginal, 33, 162 Histogram clustering, 153 Histogram dispersion, 153 Honeybee brain images, 437-438, 439f, 440f, 441f, 444f, 451, 452-453, 460, 461, 464-465, 473Hough transform, 155

Hyperelastic warping, 487-521 applications of, 502-518 finite deformation theory and, 489-490 finite element discretization and, 494-495 linearization in, 491 overcoming local minima in, 496-498 sequential spatial filtering in, 496-498 solution procedure, 495-496 variational framework for, 490-491 Hypertrophic (dilated) cardiomyopathy, 512IBSR. See Internet Brain Segmentation Repository ICP algorithm. See Iterative Closest Point algorithm Image cropping, 108, 116-117 Image dependent models, 346-347 Image interpolation, 347 Image pyramids, 364 Image registration, 1-42, 339-357, 535-551. See also specific types applications of, 341-342 classification of, 3-8 clinical validations of, 70-88 2D/2D, 4 2D/3D, 4, 6 3D/3D, 4 defined, 8 dimensionality of, 4 distance-based techniques in, 10-21 extrinsic, 4-5 future application of, 535-537 general theory of, 8-10 interaction levels in, 6-7 intrinsic, 4, 5 level set segmentation and, 63-70 modalities involved in, 7 practical examples of, 34-42 review of techniques, 342-351 segmentation combined with, 537 subject types, 7-8 Image segmentation. See Segmentation Image stabilization, 342 Image stiffness, 491 Image warping. See Warping Implicit deformable models for brain aneurysms, 190-191 without gradients, 86-88 Important points, 28-29 Indirect registration, 407-410 Information-theoretic entropy, 442 Initialization, 200-201

Inner ear imaging, 234 Insight Segmentation and Registration Toolkit (ITK), 77-78 Intensity-based inverse consistent registration feature-based registration and, 236 landmark-based registration and, 240-245 Intensity-based methods, 538. See also Photometric methods Intensity-based small deformation inverse consistent linear elastic (I-SICLE) image registration, 233-236 Intensity correction geometric transformations and, 278, 283 Romeo and, 289-290, 297-301, 307-308 Intensity variance cost functions, 225 Interactive algorithms, 350 Interactive registration, 6-7 Intermodality registration, 342 Internet Brain Segmentation Repository (IBSR), 71-72, 83 Interpolation, 344 bilinear, 160-161, 162, 174, 176, 179 B-spline, 344 cross entropy, symmetric divergence, and, 399-400 of elastic images, 359 image, 347 landmark (see Landmark interpolation) linear, 344 nearest neighbor (see Nearest neighbor interpolation) partial volume, 460 retinal image registration and, 160-161, 162, 174, 176, 179 thin-plate spline, 352-354 Intersubject registration, 7 defined, 342 non-rigid, 273-321 Intrasubject registration, 7, 342 Intravascular ultrasound, 507-512 Intrinsic markers, 345 Intrinsic registration, 4, 5 Invasive markers, 5 Inverse consistency constraint (ICC), 222, 225, 236, 244 Inverse consistency error, 219, 220, 230-231, 239, 243-244 Inverse consistent image registration, 219 - 245algorithms, 233-236 intensity and feature-based, 236 problem statement, 222-224 regularization constraint, 226-227

Inverse consistent landmark-based thin-plate spline (CL-TPS) image registration, 234-236, 237-245 Inverse consistent transformations, 222 computation of, 225-226 invertibility property of, 230-231 parameterization of, 227-228 transitivity property of, 231-233 Invertibility, 350 elastic image registration and, 372-373 of inverse consistent transformations, 230-231 Invertibility constraint, 221 Ischemic myocardial disease, 512 I-SICLE. See Intensity-based small deformation inverse consistent linear elastic image registration Iterative Closest Point (ICP) algorithm, 10-15, 26, 276-277, 283. See also Genetic algorithms Iterative shape averaging, 454–455 ITK. See Insight Segmentation and Registration Toolkit Jacobians, 350 elastic image registration and, 372 hyperelastic warping and, 489-490, 500 inverse consistent registration and, 225-226, 227, 229, 240, 242-243, 245 Joint entropy, 31-32, 153, 156, 157, 442 Joint histograms, 32-33, 545, 551 cross entropy, symmetric divergence, and, 399 k-means clustering and, 542 retinal image registration and, 153, 162, 167, 169 - 171Joint probability density function (pdf), 156, 171, 399, 401, 402, 415, 424-426 estimation of, 161-162 Kalman filters, 14, 276 Kernel term, 353, 355 K-means clustering, 542 k-Nearest Neighbor (kNN) classifiers, 192, 194, 195, 199, 200, 211

Lagrangian framework augmented, 495–496 for deformable image registration, 489 for level sets, 49 Lamé coefficients, 279, 519 Landmark(s) anatomical, 5 geometrical, 5

in pelvis and prostate imaging, 104, 109-110, 112, 116-117, 118, 120, 121 Landmark-based registration, 345, 351-357, 538 intensity registration and, 240-245 inverse consistent registration and, 220, 221, 236, 237-240 manual, 351 Landmark constraints, 358 Landmark error, 81 Landmark interpolation, 347, 348, 351-357 desirable properties of, 351-352 formula for, 353-354 fractional, 354-357 Laparascopy, 3 Laplacian energy, 353 Laplacian level set methods, 78 Laryngoscopy, 3 Least squares techniques, 11, 14 Leave-one-out technique, 356 Leclerc contour detector function, 202 Left ventricular (LV) deformation, 512, 513, 514 - 516Level set equation, 200, 201 Level set segmentation, 47-89, 344, 436 of brain aneurysms, 198 clinical validations of, 70-88 framework for, 48-51 image registration and, 63-70 important clinical problems, 70-77 limitations of, 88-89 open source software tools for, 77-78 shape priors in, 66-69 Level set speed functions gradient-based active contours and, 51-52 with regularizers, 56, 57-59 tuning for, 55-57 Linear binning technique, 541, 548, 550 Linear elasticity, 226-227, 492 Linear interpolation, 344 Live-wire method, 253 Local models, 345-346 Local optimum, 539 Local transformation, 6 Luminance conservation hypothesis, 289 Lung cancer, 39-42 Lung imaging, 137, 234, 454 Macular degeneration, 152 Magnetic resonance angiography (MRA), 3, 190 Magnetic resonance imaging (MRI), 2, 3, 21, 32-33, 61, 436

atlas-based segmentation and, 443, 454, 463 binning technique and, 539, 540, 542, 545, 546, 548, 549t, 551
Magnetic resonance imaging (MRI) (cont.) coupling-surfaces regularizers and, 59 cross entropy, symmetric divergence, and, 396, 402-403, 404-405, 407-410, 411f, 412f, 413f, 415f, 416f, 417f, 418f, 419f, 420f, 421f, 423f, 424f, 425f, 427f displaced frame difference and, 287 elastic image registration and, 365, 366f, 374-376, 377, 378f FLAIR, 36, 37f functional, 3, 316, 318, 535 geodesic active contours and, 54, 55 hyperelastic warping and, 502-504, 512-513, 514-516 interventional, 103, 140 inverse consistent registration and, 230, 240, 241, 245 level set segmentation and, 62, 63-64, 66, 67, 68, 69-72, 73-74, 76, 78-84 for multiple sclerosis, 36-38 non-rigid registration and, 275, 278, 283 online registration and, 537 of pelvis and prostate, 103-142 perfusion studies and, 536 Romeo and, 289, 291, 297, 299f, 308, 313-315, 316, 318 of small animals, 535 symmetric similarity cost functions and, 224-225 3D, 81-84 volumetric, 514-516, 517-518 Magnetic resonance microscopy (MRM), 502 - 503Magnetic resonance spectroscopy, 103 Malleus imaging, 502, 505-507 Manual registration, 418 cross entropy, symmetric divergence, and, 405, 406, 409-410, 411-414, 418-419, 420f, 421f, 422f, 423f, 424f landmark, 351 Manual segmentation, 462, 464, 471-472 MAP classifiers. See Maximum a posteriori classifiers Marching cubes algorithm, 203 Marginal histograms, 33, 162 Marginal probability density function (pdf), 156, 397, 399, 402 estimation of, 161-162 Marquardt-Levenberg algorithm, 349, 357, 362-363, 368 Masking, 371-372 Material stiffness, 493, 495 Maximization cross-entropy, 394, 396, 397, 403-405, 406 - 417

mutual information (see Mutual information maximization) numerical stability for, 414-417 reversed cross-entropy, 397, 402, 403-405, 406 - 417symmetric divergence, 397, 403-405, 406-417 Maximum a posteriori (MAP) classification, 192, 195-197, 201 Maximum Intensity Projection (MIP), 189f, 190, 191, 203, 212 Mean square error (MSE), 302-305, 313 Mechanical scanning, 75-76 Minimization block-matching scheme, 284 Brent line, 34, 400 cross-entropy, 394, 396, 401, 405-406, 417-422 entropy, 397 multigrid scheme, 292-294, 298f numerical stability for, 419-422 reversed cross-entropy, 396, 405-406, 417-422 symmetric divergence, 396, 405-406, 417-422 MIP. See Maximum Intensity Projection Mixture model, 289-290 MNI phantom. See BrainWeb phantom Modal contours, 56 Modality to model registration, 7 Model-based registration, 487-488 Model to modality registration, 7 Monomodal registration, 3, 7, 39 Motion analysis, 341-342 Motion transformations, 11, 14 Mouse brain, 502-504, 505f MRA. See Magnetic resonance angiography MRI. See Magnetic resonance imaging MUGA. See Multigated angiography Multidimensional optimization methods, 349 Multigated angiography (MUGA), 73, 76 Multigrid methods, 349 Multigrid minimization scheme, 292-294, 298f Multilabel classifier performance model, 469, 470-471, 472 Multimodal registration, 7 composite approach for, 34 non-rigid, 284, 296-297 Multimodal segmentation, 69-70 Multiple sclerosis (MS), 3, 35-39, 72 Multiresolution strategies, 349-350 binning technique in, 537-550 B-spline representation and, 369 for elastic image registration, 357, 363-364 for inverse consistent image registration, 228-229 for optical flow computation, 291-292

for retinal image registration, 162, 164-165, 178 - 179for 3D image registration, 400-401 two-stage, 544-545 Mumford-Shah segmentation, 56, 61-62, 66, 67, 86 Mutual information (MI), 347, 394 atlas-based segmentation and, 442-443 inverse consistent registration and, 225 in lung cancer studies, 41 normalized (see Normalized mutual information) Romeo and, 288 3D non-rigid body registration and, 132, 133f, 1363D rigid body registration and, 105-106, 107-108, 110, 114, 118, 120-121, 136, 139 Mutual information (MI) maximization, 6, 21, 397 algorithm, 31-34 binning technique and, 538-539 computation of metric, 32-33 non-rigid registration by, 284 optimization methods for, 34 retinal image registration by, 151-179 Romeo and, 286, 315 Mycotic aneurysms, 186 Myelin, 36 Myocardial infarction, 508, 517-518 Navier-Stokes equation, 279-280 Nearest neighbor interpolation, 179, 344 atlas-based segmentation and, 460, 470 binning technique and, 545 retinal image registration and, 160-161, 174 Neo-Hookean hyperelastic material, 503, 506, 511, 515-516 Neuractive pyramid, 282 Newton's method, 349, 362, 491 Noise atlas-based segmentation and, 458-459 computer-aided volumetric segmentation and, 262 elastic image registration and, 358 Gaussian, 358 Gaussian white, 354 level set segmentation and, 63, 71, 73, 80, 84, 86, 87, 88 retinal image registration and, 152, 154, 167 Romeo and, 286, 301, 302, 304t, 316 3D rigid body registration and, 105 white, 154, 354 Non-invasive markers, 5 Nonlinear binning technique, 541, 546, 547f,

548, 550

Non-parametric, local methods, 345-346 Non-parametric probability density function (pdf), 194-195 Non-parametric tissue probability estimation, 194 - 195Non-rigid registration, 273-321 atlas-based segmentation and, 443-447, 452, 454-455, 456-457, 466 deformation fields generated by, 536-537 geometric methods in, 275-278, 284-285, 320 multimodal, 284, 296-297 overview of methods, 274-285 photometric methods in, 275, 278-284, 285, 320 Romeo algorithm and (see Romeo algorithm) 3D (see Three-dimensional non-rigid body registration) Non-rigid transformations atlas-based segmentation and, 446, 447-449, 452.455 regularization of, 447-449 surface signature surface registration and, 29 Non-saccular aneurysms, 186, 187f Normalized mutual information (NMI) affine, 452atlas-based segmentation and, 443, 446, 447, 452, 454, 466 binning technique and, 538, 539, 542-544, 545non-rigid, 452, 454, 466 Nuclear emission (Em), 402-403, 404-405, 406-410, 411f, 412f, 413f, 415f, 416f, 417f, 418f, 419f, 420f, 421f, 423f, 424f, 425f direct registration by, 406-407 indirect registration by, 407-410 Nuclear transmission (Tx), 402, 404-405, 406-410 direct registration by, 406-407 indirect registration by, 407-410 Numerical stability, 401-402 for maximization, 414-417 for minimization, 419-422 Nyquist's criterion, 291 Object tracking, 342 Online image registration, 537 Optical flow, 346 atlas-based segmentation and, 443 displaced frame difference and, 282, 287 multiresolution incremental computation of, 291 - 292robust multigrid elastic registration based on (See Romeo algorithm) Optical flow constraint (OFC), 287 Optical flow hypothesis, 286-287

Optimization atlas-based segmentation and, 447-448 binning technique and, 548-550 control point (see Control point optimization) cross-entropy, 394, 400, 401 elastic image registration and, 361-364 multidimensional, 349 mutual information maximization, 34 non-rigid registration and, 275 Powell's method, 34, 162, 289, 297, 349, 400 procedure for, 7 retinal image registration and, 162-165, 174, 176, 178 reversed cross-entropy, 400, 401, 402 Simplex downhill (see Simplex downhill optimization) symmetric divergence, 400, 401 Optimization algorithm, 361-363 Pairwise consistent transformations, 222 Parametric correction, 290 Parametric deformable models, 80, 81, 87 main limitations of, 47-48 for Romeo, 294-295 Parametric global methods, 346 Parametric surfaces, 13 Partial data sets, 9, 10 Partial differential equations (PDEs), 63, 65, 197, 346, 348, 349 Partial volume interpolation (PVI), 460 Parzen density estimation method, 194, 418, 425-426 Pasha algorithm, 281 Pattern recognition field, 487 PDEs. See Partial differential equations pdf. See Probability density function Peaking phenomenon, 211 Pelvis and prostate imaging, 103-142 3D non-rigid body registration in, 104, 121-141 3D rigid body registration in, 104, 105-121 Penalty method, 495 Penalty term atlas-based segmentation and, 448 elastic image registration and, 372-373 Perfusion studies, 536 PET. See Positron-emission tomography Phased arrays, 76 Photometric methods, 275, 278-284, 285, 320 Pixel-based clustering, 47 Pixel-based registration, 343, 344, 347, 487, 488 Planar scintigraphy, 3

Plaque rupture, 507-512 Point(s) control (see Control point(s)) grid closest, 15-21 important, 28-29 in non-rigid registration, 276 seed, 540 Point matching, 176-177 Point signature surface registration, 22, 24 Positron emission tomography (PET), 2, 3, 103 cross entropy, symmetric divergence, and, 395, 396 level set segmentation and, 63, 70, 72-73, 74, 76.88small animal imaging and, 535 Powell's optimization method, 34, 162, 289, 297, 349, 400 Preprocessing, 344 Principal axes methods, 5 Priori probability density function (pdf), 394, 396, 397, 399, 401, 402, 405, 417-418, 419-422, 423-426 Probability density function (pdf) brain aneurysms and, 193, 194-195, 196f, 199, 211cross-entropy, symmetric divergence, and, 394, 395-397, 399 estimation of, 195, 196f joint (see Joint probability density function) marginal (see Marginal probability density function) non-parametric, 194–195 priori (see Priori probability density function) retinal image registration and, 156, 157 zero, 402, 415-417, 419-422 Projective transformations, 6 Prostate centroid displacement, 111, 120, 121, 132, 138-139, 141 Prostate imaging. See Pelvis and prostate imaging Quasi-Newton methods, 495, 496 Radial basis functions, 347 Rational Gaussian (RaG) surface, 255-256, 260-263, 267-269 Raw image, 442, 443, 449, 452-453, 457, 459, 460, 465, 466, 467 Real-time three-dimensional (RT3D) ultrasound geodesic deformable models and, 84-86 implicit deformable models and, 86-88

Real-time three-dimensional (RT3D) volumetric imaging, 76

## Index

Recursive filtering, 369 Reduction operator, 369 Reference image, 339, 340, 546 atlas-based segmentation and, 442, 447, 448, 455 binning technique and, 543-544 cross entropy, symmetric divergence, and, 395, 398 elastic image registration and, 358 retinal image registration and, 157, 161, 162, 165 Reference volume, 31, 32 multiple sclerosis studies and, 39 Romeo and, 316 3D non-rigid body registration and, 122, 124, 136-137  $3\mathrm{D}$  rigid body registration and, 105Region-based level set active contours, 61-63 Region growing, 47, 539-541, 551 Registration. See Image registration Regularization cost functions and, 347, 348 elastic image registration and, 372-373 non-rigid registration and, 275 non-rigid transformations and, 447-449 Regularization constraint, 226-227 Regularization stiffness, 491, 495 Regularizers, 56, 57-59 Bayesian-based, 57-58 clustering-based, 57 coupling-surfaces, 57, 58-59 shape-based, 57, 58 RegViz (software), 109-110, 120, 122, 125, 126 Reinitialization, 200-201 Reproduction of identity, 352 Residual transformations, 222 Retinal image registration, 151-179 description of files, 165-166 review of, 154-156 software implementation and architecture, 165success rate, speed, and accuracy of, 174-177 Retrospective Registration Evaluation Project, 545, 546 Reversed cross-entropy, 394, 395-427 implementation of, 397-402 maximization of, 397, 402, 403-405, 406-417 minimization of, 396, 405-406, 417-422 numerical stability and, 401-402, 414-417, 419-422 registration experimental set-up for, 402-406 Reverse transformations CL-TPS image registration and, 236

fluid models and, 280 inverse consistent registration and, 219-220, 222, 227-228, 240, 242, 243-244 Riemannian space, 53-54, 197 Rigid registration, 6, 8 atlas-based segmentation and, 443 level set segmentation and, 67, 69 limitations of, 273 Romeo and, 288-289 3D (see Three-dimensional rigid body registration) **Rigid segmentation**, 5 Rigid transformations, 6, 151 cross entropy, symmetric divergence, and, 399 level set segmentation and, 66 shape-based regularizers and, 58 spin image surface registration and, 26 surface signature surface registration and, 31 Robust adaptive segmentation, 78-79 Robust multigrid elastic registration based on optical flow. See Romeo algorithm Robustness of level set segmentation, 88-89 of 3D non-rigid body registration, 139 of 3D rigid body registration, 135, 139 Romeo algorithm, 274, 285-319 experiments on multimodal datasets, 315-318 experiments on real data, 313-315, 318 experiments on simulated data, 301-307, 316-318 general formulation for, 286-288 intensity correction and, 289-290, 297-301, 307-308 multigrid minimization scheme for, 292-294, 298f parametric models for, 294-295 rigid registration step, 288-289 robust estimators for, 290-291, 306-307, 320 Rotational angiography (3DRA), 191, 192f Rotational misalignment, 9 Rotation angles, 21 cross entropy, symmetric divergence, and, 399, 405, 411, 412f, 415-416, 418f, 419, 420f, 421, 424f level set segmentation and, 69-70 retinal image registration and, 154, 157, 158, 159, 160, 163, 171, 173f, 174, 176, 177, 178Rotation matrix, 8, 18-19 Rotation transformations, 11 Roulette wheel selection, 19

RT3D ultrasound. See Real-time three-dimensional ultrasound Saccular (berry) aneurysms, 186, 187f Scaling feature-based registration and, 21 retinal image registration and, 154, 157, 158, 159, 160, 163, 165, 171, 173f, 176, 177, 178 Search space, 343, 345-347 Search strategy, 343, 348-350 Seed points, 540 Segmentation aneurysm, 193-202 atlas-based (see Atlas-based segmentation) level set (see Level set segmentation) manual, 462, 464, 471-472 multimodal, 69-70 Mumford-Shah, 56, 61-62, 66, 67, 86 registration combined with, 537 rigid, 5 shape detection, 77 volumetric (see Computer-aided volumetric segmentation) Segmentation algorithms, 342 SEM algorighm. See Stochastic expectation-maximization algorithm Semi-automatic registration, 7, 364-365 Semi-local models, 346 Sequential spatial filtering, 496–498 Shape-based regularizers, 57, 58 Shape detection segmentation, 77 Shape editing, 262, 263-264 Shape index, 25 Shape models, 64, 65, 454 Shape priors, 66-69 Shape spectrum, 26 Shape voxel parametrizing, 256-260 Signal-to-noise ratio, 63 Similarity index (SI), 453, 461-463, 464, 465, 466-467, 468f, 469f Similarity measures. See also Cross-entropy; Entropy; Mutual information; Reversed cross-entropy; Symmetric divergence in non-rigid registration, 275 in pixel-based registration, 347 in Romeo, 286 in 3D image registration, 393-427 in 3D rigid body registration, 105-106, 120 - 121Similarity transformation matrix, 8-9 Simplex downhill optimization, 108, 136 binning technique and, 545, 548 retinal image registration and, 156, 162, 163-164, 165, 178 Simulated annealing, 162, 174, 176

Simulated Brain Database (SBD), 72. See also BrainWeb phantom Single photon emission computed tomography (SPECT), 2, 3, 103, 395 elastic image registration and, 376 level set segmentation and, 63, 70, 74, 76 small animal imaging and, 535 Small animal imaging, 535-536 Smoking, 40 Snake(s), 253, 254, 436. See also Active contours; Deformable models Snake Automatic Partitioning (SNAP), 77 Sobolev space, 345 Soft constraints, 495 Spatial elasticity tensor, 493 Spatial registration, 4 SPECT. See Single photon emission computed tomography Spin image surface registration, 22, 26–27, 28f Spin maps, 27 Splash surface registration, 22-24 SPM spatial normalization approach, 282 Squared-error similarity functions, 225 SSD. See Sum of Squared Differences Stereo images, 342 Stereo retinal images, 156, 166, 167-168f, 169, 176, 177, 178 Stochastic expectation-maximization (SEM) algorithm, 78, 79, 289-290, 297, 300f, 301 Stopping criterion, 371 Strain energy, 353, 360, 492, 493 Strain measurement, of coronary artery, 507-512 Structural indexing, 22-24 Stunned myocardium, 512 Subarachnoid hemorrhage (SAH), 187-188 Sum of Squared Differences (SSD), 281, 282, 347, 358-359, 362f, 364, 367 Sum Rule fusion, 460 Surface(s), 12-13 free-form, 21, 22 inverse consistent registration and, 220 non-rigid registration and, 277 parametric, 13 Rational Gaussian (see Rational Gaussian surface) Surface-based registration techniques, 21-31, 34.538 COSMOS, 22, 24-26 point signature, 22, 24 spin image, 22, 26-27, 28f splash, 22-24 surface signature, 22, 28-31, 34 Surface signature surface registration, 22, 28-31, 34

Symmetric divergence, 394, 395-427 implementation of, 397-402 maximization of, 397, 403-405, 406-417 minimization of, 396, 405-406, 417-422 numerical stability and, 401-402, 414-417, 419-422 registration experimental set-up for, 402-406 Symmetric gradient operator, 493 Symmetric similarity cost functions, 223-225 Talairach proportional squaring system, 276 Talairach stereotaxic space, 276 Tangent stiffness, 495 Target image, 348, 487, 489, 490, 492, 503, 509 Template image, 345 data term for, 348 deformable image registration and, 487, 488, 489, 490, 491, 497, 498, 499f, 503, 506, 509, 514, 515, 517-518 Template matching, 29 Temporal retinal images, 156, 166, 167-168f, 170, 171, 176, 177, 178 Test image, 339, 340, 358 Thin-plate spline (TPS) deformations, 277-278, 316 Thin-plate spline (TPS) interpolation, 352-354 Thin-plate spline (TPS)-RPM algorithm, 278 Thin-plate spline (TPS) solutions, 348 Thin-plate spline (TPS) transformations, 124, Thin-plate spline (TPS) warping, 122, 125, 128, 129, 130f, 131, 132, 133f, 134f, 135, 138, 139, 140 Three-dimensional digital subtraction angiography (DSA), 190 Three-dimensional echocardiography (3DUS), 75 - 76Three-dimensional image registration, 103–142, 393-427. See also Cross-entropy; Reversed cross-entropy; Symmetric divergence; Three-dimensional non-rigid body registration; Three-dimensional rigid body registration Three-dimensional level set segmentation, 78-88 in brain reconstruction, 79-81 of the cortex, 81-84 robust adaptive, 78-79 Three-dimensional non-rigid body registration, 104, 121-141 algorithm, 122-125 algorithmic implementation, 135-137 algorithmic robustness and efficiency, 139 applications of, 137-138, 140 evaluation of, 125-126, 138-139

examples and results, 126-128 quantitative evaluation of, 132-135 rationale for, 121-122 3D rigid body registration comparison, 129 - 132Three-dimensional rigid body registration, 104, 105-123, 125, 135-136, 139, 140, 141 accuracy of, 118-119, 121-122 algorithm, 120-121 algorithm with special features, 106-108 assessment of pelvic, 112-115 assessment of procedure, 119-120 assessment of prostate, 115-116 computer implementation of, 121 consistency of, 110 evaluation of, 109-111 examples and results, 111-116 implementation issues, 117-118 3D non-rigid body registration comparison, 129-132 Thresholding outliers, 13 Threshold level set methods, 78 Time series in registration, 4 Tissue types, labeling of, 436 TPS. See Thin-plate spline Transformation(s) affine (see Affine transformations) atlas-based segmentation and, 455-457 cross entropy, symmetric divergence, and, 399 curved (elastic), 6 deformable image registration and, 487 diffeomorphic, 221, 226 finding parameters of, 7 forward (see Forward transformations) geometric, 278, 283 global, 6, 275 ICP algorithm and, 10–11 inverse consistent (see Inverse consistent transformations) local, 6 motion, 11, 14 nature of, 6 non-rigid, 452 pairwise consistent, 222 projective, 6 propagation of, 455-457 residual, 222 retinal image registration and, 157, 158 - 160reverse (see Reverse transformations) rigid (see Rigid transformations) rotation, 11 thin-plate spline, 124, 137 translation, 11

Transformation matrix, 31 ICP algorithm and, 11 similarity, 8-9 Transformation parameterization, 227-228 Transform-based registration, 343, 344, 348 Transitivity, 231-233 Translation, 8, 9, 21 cross entropy, symmetric divergence, and, 399, 405, 415, 418f, 419, 421, 422f, 423f, 425f, 427f genetic algorithms and, 18 level set segmentation and, 69-70 retinal image registration and, 154, 157, 158, 159, 160, 163, 165, 171, 172f, 176, 177, 178 Translation transformations, 11 Traumatic aneurysms, 186 Triangular meshes, 13 True registration, 411-414, 418-419 Tumors, 2, 3, 72, 73. See also Brain tumors Two-dimensional echocardiography (2DE), 74 - 75U (energy term), 490, 491-493, 495 UL-TPS image registration. See Unidirectional thin-plate spline image registration Ultrasound, 3 interventional, 536 intravascular, 507-512 level set segmentation and, 54, 62, 63, 67, 68, 76RT3D (see Real-time three-dimensional ultrasound) Unidirectional thin-plate spline (UL-TPS) image registration, 234-235, 237-245 Uzawa algorithm, 496 Variational formulation, 346, 352 Video compression, 342 Virtual springs, 364-365

Virtual springs, 364–365 Viscous fluid models, 226, 346, 488 Visible Human Project, 77 Voigt notation, 489, 494 Volume bias from, in segmentation, 462-463 from contour to, 277-278 floating (see Floating volume) inverse consistent registration and, 220 reference (see Reference volume) 3D non-rigid body registration and, 127 voxel-based, 21 Volume of interest (VOI), 123-125, 136-137, 138, 139 Volumetric image segmentation. See Computer-aided volumetric segmentation Volumetric magnetic resonance imaging (MRI), 514 - 516Vote Rule decision fusion, 460 Voxel-based registration, 5-6, 103-104, 122, 538 Voxel-based volumes, 21 Voxel displacement, 110-111, 117 Voxel similarity approaches, 394 W (energy term), 490, 491-493, 519 Warfield algorithm, 470 Warping, 339-340, 345, 346-347, 364, 374-375, 376. See also Hyperelastic warping; Thin-plate spline warping Warp space. See Search space Wavelets, 227, 344, 346, 349 White matter, 71, 78, 79, 81-84, 225, 317, 542 coupling-surfaces regularizers and, 59 geometric deformable models and, 80-81 intensity correction and, 289-290 multiple sclerosis and, 36 overlap with gray matter, 313-315 White noise, 154, 354 Window levelling, 190

X-ray(s), 2, 41 X-ray fluoroscopy, 536