

Ayman S. El-Baz  
Rajendra Acharya U  
Andrew F. Laine  
Jasjit S. Suri *Editors*

Volume II

# Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies



# Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies



Ayman El-Baz • Rajendra Acharya U •  
Andrew F. Laine • Jasjit S. Suri  
Editors

# Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies

Volume II



Springer

*Editors*

Ayman El-Baz  
University of Louisville  
Department of Bioengineering  
Louisville  
USA  
aselba01@louisville.edu

Rajendra Acharya U  
Ngee Ann Polytechnic  
School of Engineering  
Clementi Road 535  
599489 Singapore  
Blk 7 Level 2  
Singapore  
aru@np.edu.sg

Andrew F. Laine  
Columbia University  
Department of Biomedical Engineering  
New York  
USA

Jasjit S. Suri  
Biomedical Technologies, Inc.  
Denver, CO, USA  
Global Biomedical Technologies, Inc.  
California, USA  
(Aff.) Idaho State University  
Pocatello, ID, USA  
jsuri@comcast.net

ISBN 978-1-4419-8203-2      e-ISBN 978-1-4419-8204-9

DOI 10.1007/978-1-4419-8204-9

Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011924210

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden. The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

*Ayman El-Baz would like to dedicate this book to  
his wife, daughter, son, mother, and father.*

*Rajendra Acharya U would like to dedicate this book to  
his students, collaborators, and colleagues.*

*Andrew F. Laine would like to dedicate this book to  
his late beloved father Jason Laine.*

*Jasjit S. Suri would like to dedicate this book to  
his students and collaborators all over the world.*



# Preface

Segmentation of medical images in 2-D, 3-D, and 4-D has taken new dimension. This is mainly because for real-time image guidance in diagnosis and therapeutics. The application of image/volume registration is prominently seen in reconstruction, multimodality fusion, and now propagating in atlas-based analysis. The focus of this book is to share the application of combined segmentation and registration in medical imaging.

Some of the cutting edge topics in segmentation covered in the book are graph cut, energy minimization methods, cine loop processing in cardiac applications, parametric and geometric deformable models in combination with principal component analysis, breast mass classification, classification of thyroid lesions into benign and malignant using vasculature properties, classification of autistic versus nonautistic brain data sets, and geometric modeling for prosthesis design. Some of the advanced topics in image/volume registration are optimizers for image registration, such as Amoeba, conjugate gradient, and gradient-based methods and techniques for alignment of images of a textured object with a given prototype, multimodality approaches to image registration. The book also presents unified variation frameworks where prior segmentation information can be seamlessly integrated into nonrigid registration procedures for brain structure registration. The detail layout of the book is as follows.

Chapter 1 gives a brief overview of the most popular medical image segmentation techniques and discusses their capabilities and basic advantages and limitations. Moreover, the state-of-the-art techniques of the last decade are outlined.

Chapter 2 presented a graph cut-based segmentation approach to segment cerebral white matter (CWM) from magnetic resonance images (MRIs). This approach is based on using a probabilistic variation of the traditional graph cut algorithm with an improved parameter selection mechanism for the energy function, to be optimized in a graph cut problem.

Chapter 3 introduced a new framework to analyze noncontrast agent cine cardiac magnetic resonance (CMR). The framework includes two automated algorithms (1) to segment the inner and outer borders (or walls) from the surrounding tissues in cine CMR images based on using a new stochastic speed function to control the evolution of geometric deformable model, and (2) to estimate the local wall thickness, and thickening function indexes based on solving Laplace's equation

in the segmented wall to find point-to-point correspondences between the inner and outer border of the segmented wall.

Chapter 4 presented a novel shape-based segmentation approach using learned prior and current appearance model. The target shapes are approximated directly with linear combinations of distance vectors describing positions of the mutually aligned training shapes with respect to their common centroid. Such a vector space is now closed with respect to the linear operations and it is of much smaller dimensionality than the 2-D distance maps. Prior knowledge of visual appearance of the object is represented by Gibbs energies of its gray levels. To accurately separate the object from its background, each current empirical marginal probability distribution of gray values within a deformable boundary is also approximated with an adaptive linear combination of discrete Gaussians. Both the shape/appearance priors and the current probabilistic appearance description control the boundary evolution.

An improved level set-based mass segmentation is presented in Chap. 5 to classify the segmented mass to malignant or benign based on the estimated texture and morphological features. The new mass segmentation algorithm is based on a new energy functional, which combines energy terms from several traditional algorithms. In Chap. 6, an image-based framework is presented to diagnose single thyroid nodule. This image-based approach is based on three main steps: (1) image preprocessing; (2) morphological aperture for contrast agent bubbles detection; (3) thresholding and 3-D lesion reconstruction. Chapter 7 covers a novel framework to analyze the cortex of the patient with autism using spherical harmonics analysis. Chapter 8 covers state-of-the-art approaches related to surface reconstruction and geometric methods that have been used for digital prosthesis design. Chapter 9 covers the necessary background information required to understand medical image registration, the basic tools required to implement registration algorithms, and to demonstrate a complete application for various types of registration between different modalities using freely available and maintained software.

Chapter 10 presented a novel approach to align an image of a textured object with a given prototype using a new similarity measure. Visual appearance of the images, after equalizing their signals, is modeled with a Markov–Gibbs random field (MGRF) with pairwise interaction. Similarity to the prototype is measured by a Gibbs energy of signal co-occurrences in a characteristic subset of pixel pairs derived automatically from the prototype. Chapter 11 covers three registration methods that address some of the technical choices that are part of specific engineering solutions. These, while driven by concrete clinical needs, may generalize well to similar problems in other domains requiring image registration. Chapter 12 presented registration-based framework to estimate ocular surface temperature. Chapter 13 presented simultaneous segmentation and registration approach to track brain morphological changes in MRIs.

# Contents

<b>1</b>	<b>Medical Image Segmentation: A Brief Survey .....</b>	1
	Ahmed Elnakib, Georgy Gimel'farb, Jasjit S. Suri, and Ayman El-Baz	
<b>2</b>	<b>Cerebral White Matter Segmentation using Probabilistic Graph Cut Algorithm .....</b>	41
	Mainak Sen, Ashish K. Rudra, Ananda S. Chowdhury, Ahmed Elnakib, and Ayman El-Baz	
<b>3</b>	<b>A New Image-Based Framework for Analyzing Cine Images .....</b>	69
	Fahmi Khalifa, Garth M. Beache, Georgy Gimel'farb, Guruprasad A. Giridharan, and Ayman El-Baz	
<b>4</b>	<b>Medical Images Segmentation Using Learned Priors .....</b>	99
	Ayman El-Baz and Georgy Gimel'farb	
<b>5</b>	<b>Classification of Breast Mass in Mammography with an Improved Level Set Segmentation by Combining Morphological Features and Texture Features .....</b>	119
	Jinshan Tang and Xiaoming Liu	
<b>6</b>	<b>Segmentation and Skeletonization of 3D Contrast Enhanced Ultrasound Images for the Characterization of Single Thyroid Nodule .....</b>	137
	Filippo Molinari, Alice Mantovani, Maurilio Deandrea, Paolo Limone, Rajendra Acharya U, Roberto Garberoglio, and Jasjit S. Suri	
<b>7</b>	<b>Shape-Based Detection of Cortex Variability for More Accurate Discrimination Between Autistic and Normal Brains .....</b>	161
	Matthew Nitzken, Manuel F. Casanova, Fahmi Khalifa, Guela Sokhadze, and Ayman El-Baz	

<b>8 Surface Reconstruction and Geometric Modeling for Digital Prosthesis Design .....</b>	187
Luiz C.M. de Aquino, Gilson A. Giraldi, Paulo S.S. Rodrigues, Antônio Lopes A. Junior, Jaime S. Cardoso, and Jasjit S. Suri	
<b>9 Medical Image Registration .....</b>	227
Usaf E. Aladl and Terry Peters	
<b>10 Robust Image Registration Based on Learning Prior Appearance Model .....</b>	247
Ayman El-Baz and Georgy Gimel'farb	
<b>11 Image Registration in Medical Imaging: Applications, Methods, and Clinical Evaluation .....</b>	263
Gerardo Hermosillo, Christophe Chefd'hotel, K.-H. Herrmann, Guillaume Bousquet, Luca Bogoni, Kallol Chaudhuri, Dorothee R. Fischer, Christian Geppert, Rolf Janka, Arun Krishnan, Berthold Kiefer, Ines Krumbein, Werner Kaiser, Michael Middleton, Wanmei Ou, Jürgen R. Reichenbach, Marcos Salganicoff, Melanie Schmitt, Evelyn Wenkel, Susanne Wurdinger, and Li Zhang	
<b>12 The Applications of Feature-Based Image Metamorphosis and Eyelashes Removal in the Investigations of Ocular Thermographic Sequences .....</b>	315
Jen Hong Tan, E.Y.K. Ng, Rajendra Acharya U, and Jasjit S. Suri	
<b>13 Segmentation-Assisted Registration for Brain MR Images .....</b>	335
Jundong Liu	
<b>About the Editors .....</b>	355
<b>Index .....</b>	357

# Chapter 1

## Medical Image Segmentation: A Brief Survey

Ahmed Elnakib, Georgy Gimel’farb, Jasjit S. Suri, and Ayman El-Baz

**Abstract** Accurate segmentation of 2-D, 3-D, and 4-D medical images to isolate anatomical objects of interest for analysis is essential in almost any computer-aided diagnosis system or other medical imaging applications. Various aspects of segmentation features and algorithms have been extensively explored for many years in a host of publications. However, the problem remains challenging, with no general and unique solution, due to a large and constantly growing number of different objects of interest, large variations of their properties in images, different medical imaging modalities, and associated changes of signal homogeneity, variability, and noise for each object. This chapter overviews most popular medical image segmentation techniques and discusses their capabilities, and basic advantages and limitations. The state-of-the-art techniques of the last decade are also outlined.

**Keywords** Medical image segmentation · Survey · Segmentation features · Appearance · Shape · Edges · Segmentation techniques · Statistical models · Atlas · Deformable models

### 1.1 Introduction

Image segmentation is one of the most interesting and challenging problems in computer vision generally and medical imaging applications specifically. Segmentation partitions an image area or volume into nonoverlapping, connected regions, being homogeneous with respect to some signal characteristics [1]. Let  $\mathbf{R}$ ,  $\mathbf{Q}$ , and  $\mathbf{K}$  be a finite 2-D arithmetic lattice (grid) of the size  $XY$  supporting pixels (or a 3-D one of the size  $XYZ$  supporting voxels), a finite set of  $Q$  intensities, and a finite set of  $K$  object, or region labels, respectively:

---

A. El-Baz (✉)

BioImaging Laboratory, Department of Bioengineering, University of Louisville, 423 Lutz Building, Louisville, KY 40292, USA  
e-mail: [aselba01@louisville.edu](mailto:aselba01@louisville.edu)

$$\mathbf{R} = (\mathbf{r} = (x, y) : x = 0, 1, \dots, X - 1, \quad y = 0, 1, \dots, Y - 1)$$

or

$$(\mathbf{r} = (x, y, z) : x = 0, 1, \dots, X - 1, \quad y = 0, 1, \dots, Y - 1, \quad z = 0, 1, \dots, Z - 1)$$

$$\mathbf{Q} = (0, 1, \dots, Q - 1)$$

$$\mathbf{K} = (0, 1, \dots, K)$$

Where  $x$ ,  $y$ , and  $z$  denote Cartesian coordinates of the lattice sites. Let  $g : \mathbf{R} \rightarrow \mathbf{Q}$  and  $m : \mathbf{R} \rightarrow \mathbf{K}$  denote a digital image and a region, or object map, respectively.

If  $H(\cdot)$  is a homogeneity predicate for signals from a connected subset of image pixels or voxels, the segmentation divides an image  $g$  into  $K$  connected subimages,  $g_k : \mathbf{R}_k \rightarrow \mathbf{Q}; k \in \mathbf{K}; \mathbf{R}_k \subset \mathbf{R}$  that cover the whole lattice,  $\bigcup_{k \in \mathbf{K}} \mathbf{R}_k = \mathbf{R}$ , without overlaps,  $\mathbf{R}_k \cap \mathbf{R}_{\kappa} = \emptyset$  for all pairs  $(k, \kappa) : k \neq \kappa$ , and keep the homogeneity individually,  $H(g_k) = \text{true}$  for all  $k \in \mathbf{K}$ , but not in combination with their immediate neighbors,  $H(g_k \cup g_{\kappa}) = \text{false}$  for all the pairs of adjacent regions  $\mathbf{R}_k$  and  $\mathbf{R}_{\kappa}$  [2], as exemplified in Fig. 1.1.

Medical image segmentation is of considerable importance in providing noninvasive information about human body structures that helps radiologists to visualize and study the anatomy of the structures [3], simulate biological processes [4], localize pathologies [5], track the progress of diseases [6, 7], and evaluate the need for radiotherapy or surgeries [8, 9]. For these reasons, segmentation is an essential part of any computer-aided diagnosis (CAD) system, and functionality of the system depends heavily on segmentation accuracy. However, accurate segmentation of medical images meets with many challenges. First of all, many anatomical structures are inhomogeneous with respect to spatial repetitiveness of individual pixel/voxel intensities or their grouped co-occurrences. Two typical examples (lung tissues on a CT image and kidney on an MRI) are shown in Fig. 1.2: both the lungs and kidneys are difficult to separate because of similar lung vessels and chest tissues visual appearances as well as overlapping kidney and background appearances, respectively.

Low contrast medical images present additional challenges: e.g., object boundaries on typical kidney and cardiac MRI in Fig. 1.2 are difficult to determine even manually. The segmentation becomes even more complex in the presence of

**Fig. 1.1** Sub-images  $g_1$ ,  $g_2$ , and  $g_3$  under the homogeneity predicate requiring the same grey level for all the pixels. The union of any two adjacent regions,  $\mathbf{R}_1 \cup \mathbf{R}_3$  or  $\mathbf{R}_2 \cup \mathbf{R}_3$  is inhomogeneous

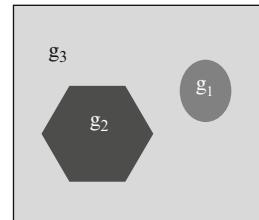


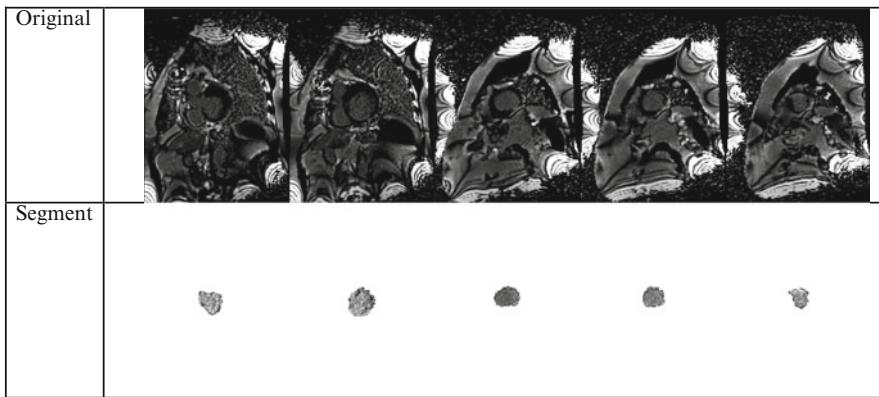
Image	Inhomogeneity		Low contrast	
Original				
Segment				
Image	Imaging noise		Occlusions	
Original				
Segment				

**Fig. 1.2** Segmentation challenges: image inhomogeneity, low contrast, noise, and occlusions

occlusions or signal deviations and noise caused by different imaging modalities and image capturing processes. Some examples of prostate and corpus callosum MRI and digital photos of natural objects are given in Fig. 1.2.

3-D and 4-D (spatio-temporal) images present further challenges, such as e.g., inter and intrasubject variability of shapes and textures of an object-of-interest in images collected over a period of time for the same subject or a group of different subjects. Consequently, a particular segmentation technique may work well for one, but not for another subject, or only on certain images of one subject. Figure 1.3 exemplifies the shape variability across a 3-D cardiac MRI data set for a single subject.

To overcome these challenges, many segmentation techniques have been developed over many years and reported in the literature. Generally, no segmentation technique works for all the applications, and various approaches of different accuracy, speed, and degree of complexity have been explored for each practical



**Fig. 1.3** Large intrasubject shape changes: Inner cavities in a 3-D cardiac MRI of the same subject

CAD problem. Their categorization is based on goals and specific aspects of data processing involved. In particular, contour-, region-, and pixel/voxel-based techniques aim at getting boundaries of goal objects, forming connected regions occupied by these objects, or performing just pixel-/voxel-wise classification, respectively. Most of the known region- or pixel-based techniques frequently involve a few contour-based steps using e.g., deformable models (boundaries), i.e., closed 2-D curves or 3-D surfaces evolving to the goal object contour or surface under internal forces reflecting the current model shape and/or external forces depending on the signal homogeneity inside and outside the model.

With respect to the user's intervention, the segmentation can be user-interactive, user-initialized, or fully automated. Many segmentation problems are stated in terms of specific cost or energy functions depending on the homogeneity of goal regions, and their solutions are found by minimizing the cost or energy over the whole set of all the possible region maps or its feasible subset. Typically, these functions are multimodal in the search space, and exact or approximate global optimization methods such as iterative graph-cut, loopy belief propagation, dynamic programming, or genetic algorithms are used to perform the best segmentation.

In contrast to unsupervised segmentation, the supervised techniques invoke additional training data (images with manually obtained region maps) for adapting to particular types of objects and images. Atlas-based segmentation simplifies the adaptation due to the use of one or more available standard medical anatomical maps, or atlases of sections across the human body, where each map point is labeled with the corresponding object class. The simplest atlas-based technique aligns an image to a relevant anatomical map to associate the object labels with the image.

This chapter overviews in brief most popular medical image segmentation techniques and discusses their fundamental pros and cons. Section 1.2 outlines different types of image features guiding the segmentation. Section 1.3 presents the known techniques and outlines the state-of-the-art ones proposed for the last decade. Concluding remarks are presented in Sect. 1.4.

## 1.2 Image Features

Accurate segmentation of medical structures relies on quantitative visual appearance and shape descriptors (image features) that help to discriminate between the structures and their background on a current image to be segmented, as well as on available prior information about the appearance and shape features for a set of manually segmented training images. Basic appearance, shape, and combined shape/appearance features to guide the segmentation process are sketched below.

### 1.2.1 Appearance Features

Typically, visual appearances of goal objects are associated with individual pixel- or voxel-wise intensities (gray values) and spatial interactions between the intensities in terms of pair-wise or higher-order intensity co-occurrences in an image.

*Intensity:* Individual intensities can guide the segmentation if their ranges or, more generally, probability distributions for an object-of-interest and its background differ to a large extent. Then the whole object or at least, most of its pixels/voxels can be easily separated from the background by comparing the intensities to one or more thresholds derived from the learned ranges or distributions of the intensities. However, generally, more powerful discriminative features have to be used.

*Spatial interaction:* The appearance of certain textures can be associated with spatial patterns of local pixel/voxel intensity variations or empirical probability distributions of intensity co-occurrences, called co-occurrence matrices [2] in the pairwise cases. Probabilistic spatial interaction models consider images as samples of a certain random field of interdependent intensities specified by their joint probability distribution.

- *Spatial domain filtering* provides the most direct way to capture local spatial interactions. Early attempts focused on intensity edges because fine textures have a higher spatial density of the edges than the coarser textures [10]. Malik and Perona [11] used nonlinear filters computing differences of two offset Gaussians (DOOG) to model preattentive human visual texture perception. The nonlinearity introduced to discriminate between the textures with the same local first- and second-order statistics of intensities worked well for both synthetic and natural images. Similar work by Unser and Eden [12] also used spatial filters in conjunction with a nonlinear operator. Other spatial filters exploit morphological (see [13] for more details) and edge operators that also aim to extract image edges. The latter, due to their extensive use, will be detailed in Sect. 1.2.3.
- *Frequency domain filtering* stems from the assumption that human visual system decomposes an image for textural analysis into oriented frequency components [14]. A multiband filtering approach by Coggins and Jain [15] used a set of

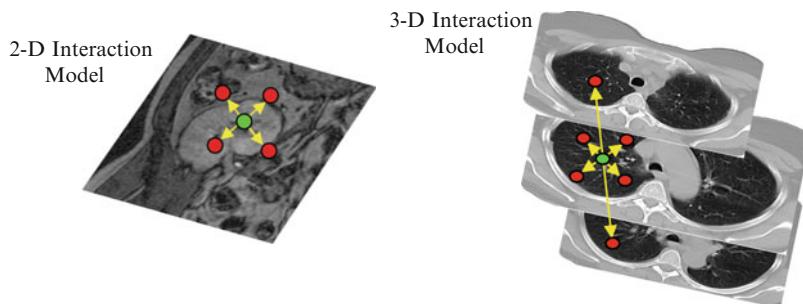
frequency and orientation selective filters to segment and classify a variety of natural and synthetic textures. Smith [16] used a set of band-pass filters followed by zero-crossing detection to feed a successful tree classifier of textures. In addition, the wavelet frequency components decomposing an image at different scales and orientations are also used to guide the segmentation [17]. These components are able to localize spatial frequencies and detect the high frequency content of the image, such as lines and edges [18].

- *Fractal models* are useful in modeling statistical roughness and self-similarity at different scales of many natural surfaces. Fractal geometry of the nature was first introduced in the nineteen seventies by Mandelbrot [19] and explored further by Barnsley [20]. A short introduction to the fractal models was given by Haindl [21]. These models are scale independent, display self-similarity, and are able to model natural textures such as clouds, leaves, or rivers as well as some medical structures (e.g., breast background structures in digital mammograms [22]). An image can be segmented in accord with one or more fractal parameters associated with the goal regions. Chaudhuri et al. [23] presented a few fractal based texture segmentation techniques. Advanced multifractal features improve capturing the texture content to guide segmentation (see e.g., [24]). However, more flexible models are needed to model medical images.
- *Probabilistic Gaussian models* assume continuous image signals in an infinite range (although digital images have a finite set  $\mathbf{Q}$  of intensities). The joint distribution  $p(g)$  is unimodal and depends on the two parameters, the mean image of the size  $XY$  or  $XYZ$  and the symmetric covariance matrix of the size  $(XY)^2$  or  $(XYZ)^2$ , respectively, to be learned from the training images. Due to an extremely large parameter space, only models with special simple covariance matrices, e.g., independent Gaussian fields with diagonal covariance matrices, were used in practice. Samples of a Gaussian model can be obtained reasonably quickly via the fast Fourier transform (FFT) [25]. However, the model imposes significant and practically unjustified restrictions on the images, such as only unimodal Gaussian distributions of intensities for any subset of pixels/voxels as well as only specific unimodal pairwise interactions between the pixels/voxels in spite of the large parameter space involved [26]. In practice, empirical distributions of image intensities and pairs of intensities are mostly multimodal and thus do not comply with this model.
- *Simultaneous autoregressive models* of images generalize the like 1-D time series models [27]: by assumption, errors  $\varepsilon(\mathbf{r})$  of linear prediction of each individual intensity  $g(\mathbf{r})$  from the neighboring intensities within a fixed window of neighbors  $\mathbf{w}$  are considered as a sample of an independent random field with the same 1-D probability distribution,  $\varphi(\cdot)$ , of each pixel- or voxel-wise component. The image probability is set equal to the joint probability of the prediction errors:

$$p(g) = \prod_{\mathbf{r} \in \mathbf{R}} \varphi \left( g(\mathbf{r}) - \sum_{\delta \in \mathbf{w}} \alpha_\delta g(\mathbf{r} + \delta) \right)$$

where,  $\varphi(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\varepsilon^2}{2\sigma^2}\right)$  for the Gaussian model of errors with zero mean image and uniform variance  $\sigma^2$ . Depending on an assumed mean image, one can use the simultaneous moving average (SMA), simultaneous autoregressive (SAR), or simultaneous autoregressive with moving average (SARMA) model [21]. For every SAR model, there exists a unique conditional Markov random field (MRF) model, but the converse is not true, except for the Gaussian case [28]. The major advantage of these models is in generally lesser numbers of parameters (equal to the size  $|\mathbf{w}|$  of the window) comparing to the equivalent MRF models. The main drawback is that the autoregressive random processes are too specific and thus are rarely encountered in practical image analysis.

- *MRF models* are most common in image modeling due to their capabilities to explicitly capture local spatial interactions between the pixels or voxels. Each individual intensity depends on a certain subset of the neighboring intensities (see e.g., Fig. 1.4), which is invariant to translations or more general geometric transformations of the lattice coordinates. A vast class of the MRF models can be equivalently described by exponential Gibbs probability distributions factored over cliques, or complete subgraphs of the interaction graph on the lattice, each factor being a function of signals supported by the clique [28, 29]. Nodes of the interaction graph are associated with the pixels or voxels,  $\mathbf{r} \in \mathbf{R}$ , and edges of the graph connect the interacting nodes. The factors, or – more frequently – their logarithms, are called clique potentials in the Gibbs distribution. A Markov–Gibbs random field (MGRF) is specified by its order (the largest cardinality of its cliques) and a set of characteristic cliques and their potentials. Due to their extensive use, the MGRF models are under continuous development to improve the segmentation accuracy, including more accurate analytical or numerical parameter estimates (see e.g., [30–33]), efficient techniques for minimizing the cost or energy functions (e.g., [34–36]), and spatial interactions of better descriptive abilities (e.g., high-order MGRF with cliques of order three or more [37–42]).

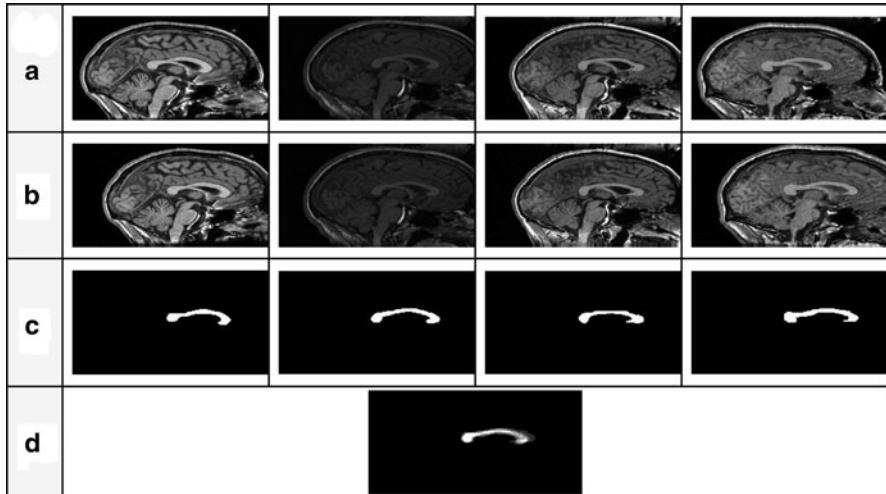


**Fig. 1.4** Spatial 2-D and 3-D interaction of each pixel/voxel intensity with the neighboring ones

### 1.2.2 Shape Features

Appearance features (i.e., intensity and spatial interaction) by themselves may fail to capture goal objects in an image in the presence of sensor noise, under poor image resolution, and/or diffused boundaries or occluded shapes. Medical structures often have well-constrained forms within a family of shapes [43]. Thus, their segmentation can be enhanced by incorporating probabilistic prior shape models specifying an average shape and variation for a certain object. The shape prior is estimated from a training database of coaligned images of the object. Figure 1.5 shows basic steps of reconstructing a shape prior for corpus callosum in four brain MRI samples.

Shape features have been extensively used in medical image segmentation (e.g., [8, 44–63]). To enhance the quality of segmentation, Cootes et al. [64] developed a statistical point distribution model, called active shape model (ASM), to represent the object shape. A number of landmarks are labeled in multiple training images of the object, and the mean positions of the landmarks and their modes of variation were used to establish the ASM. Davatzikos et al. [56] formed a hierarchy of the ASMs using the wavelet transform to capture the shape variability. Leventon et al. [65] used a shape prior with the variance obtained through the principal component analysis (PCA). Chen et al. [66] averaged an aligned training set to form a shape prior, while Litvin and Karl [67] formed a shape boundary prior from the boundary features. Gong et al. [57] modeled a 3-D prostate shape with a super ellipsoid affected by a simple parametric deformation, whereas Tutar et al. [58] represented the prostate shape using spherical harmonics (SH) of the eighth degree. Tsai et al. [68] used

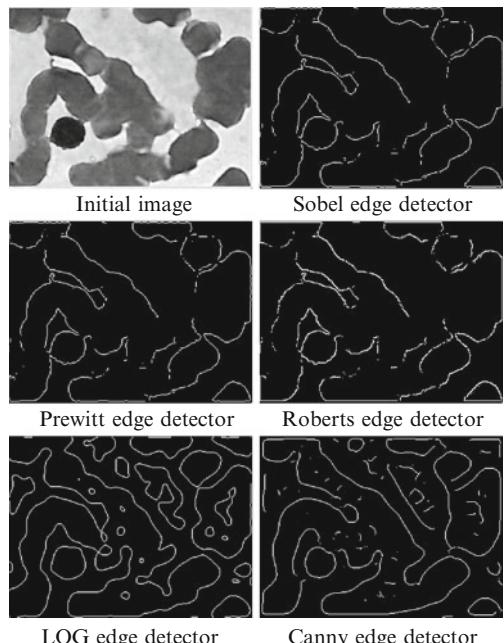


**Fig. 1.5** Estimating the shape prior: initial brain MRI (a), affine alignment by maximizing mutual information (b), manual segmentation of *corpus callosum* (c), and the estimated shape prior (d) averaging the training set (i.e., manually segmented corpora callosa in (c))

expectation-maximization (EM) estimators of the shape contours. A deterministic model was used in [45] to represent the goal shape with a weighted linear combination of signed 2-D distance maps. The weights were estimated by minimizing a mutual information-based cost function. Yang and Duncan [69] described the shape with a multidimensional Gaussian probability model of these weights. More details about the shape features can be found in [70, 71].

### 1.2.3 Combined Shape/Appearance Edge and Scale-Space Features

*Edges* belong to the most popular features for guiding image segmentation. Image analysis allows us to specify and extract different types of edges, such as intensity edges, texture edges, parametric edges, and watersheds. The intensity edges relate to large discontinuities or fast continuous changes in the intensities across an image. Such changes frequently are detected by thresholding first- and second-order spatial derivatives of the intensities (i.e., the intensity gradient and Laplacian, respectively). Popular edge detectors, such as the Sobel, Prewitt, and Roberts operators, provide finite difference approximations of the gradient [2]. Figure 1.6 shows results of applying several edge detectors to a cell image. All the detectors exploiting the first- or second-order derivatives are too sensitive to image noise [72].



**Fig. 1.6** Intensity edges on a cell image

To reduce the noise, image smoothing is typically applied before the edge detection. For example, the Marr-Hildreth Laplacian-of-Gaussian (LOG) operator [73] combines a smoothing Gaussian filter with a Laplacian zero-detecting operator to simultaneously reduce the noise and find blobs, or image regions that are consistently brighter or darker than their surroundings. Canny edge detector [74] finds the blobs by searching for the local gradient maxima using the derivative of Gaussian (DOG) filter. Figure 1.6 allows us to compare the simple derivative-based edge detectors to the LOG and Canny ones.

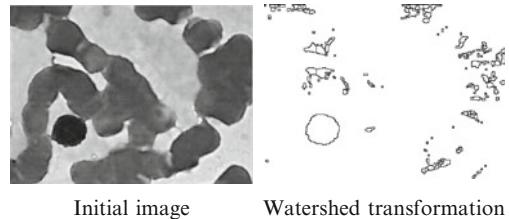
In addition to blobs, more complex appearance and shape patterns can be found by texture edge detection. Texture descriptors, such as, e.g., Gabor-like spectral coefficients, are often used to represent different scale and directional properties of a texture and thus can guide different segmentation techniques [44, 75].

The intensity or texture edges allow for detecting parametric edge models such as 2-D circles, straight lines, ellipses, 3-D spheres, ellipsoids, or more complex shapes. For example, an edge detector followed by thresholding is applied to an image. Then, the Hough transform [76] of the binary map of the detected edges detects the required parametric curves even if they are rotated or scaled. For complex shapes, such a parametric representation can be obtained by learning a prior model from a training set although such a framework might turn to be too slow and difficult to implement due to a high-dimensional parameter space. In addition, the efficiency of the Hough transform depends on both the image quality and the quality of edge detection. To enhance the latter, only appropriate connected edges have to be selected from the edge map produced by the detector. This can be done, for instance, by assigning weights to each detected candidate edge and using graph search techniques [77, 78] or dynamic programming [79] to find optimal connected edges minimizing a certain cost function.

Watershed transformation based on object morphology is a simple, fast, and intuitive alternative method for dividing a complete image of even poor contrast into separate homogeneous regions. Originally developed by Digabel and Lantuejoul [80], this edge detection technique needs no additional mechanisms for joining contours. The transformation treats pixel-wise gradient magnitudes over a 2-D image as a 3-D topographic surface with catchment basins as regions to be extracted and basins boundaries as goal watershed lines. Two possible types of the watersheds simulate either flooding, or drainage. The flooding-type transformation, also called an immersion, associates every local minimum of the topological surface with a hole, and immerses the surface slowly in water. Dams ensuring that no water from any two neighboring basins might merge represent the watershed lines. The drainage-type transformation analyzes the downstream steepest descent paths from each point of the surface and relates to the same basin any two points draining to the same local minimum (i.e., the same hole).

Watersheds have been used in a number of medical applications such as e.g., follicle segmentation from ovarian ultrasound images [81], delineation of knee cartilage and full brain white matter and gray matter segmentation [6], multispectral chromosome classification [82], lung lobe segmentation [83], lesion segmentation in uterine cervix images [84], etc. The watershed transformation frequently leads

**Fig. 1.7** Watershed transformation of a cell image



to over-segmentation into many small regions exemplified in Fig. 1.7 and caused, among other factors, by image noise amplification of the gradient operator. Region-growing mechanisms following the watershed transformation remedy the over-segmentation. Pre-smoothing and anisotropic filters also decrease it, but affect the capabilities of detecting thin structures.

*Scale-space features* extracted at multiple scales of image smoothing, to guide medical image segmentation (see e.g., [85–88]), mimic the ability of human vision to extract shapes, appearances, and/or edges of fine local image details together with large features crossing the whole image. Conventionally, spatial or frequency domain filters are applied repeatedly to scaled images of different levels of smoothing [13]. The same holds for multiple-scale shapes and/or edges of objects.

## 1.3 Segmentation Techniques

Large complexity and variability of appearances and shapes of anatomical structures make medical image segmentation one of the most challenging and essential tasks in any CAD system. Due to diversity of objects-of-interest, image modalities, and CAD problems, no universal feature set and general segmentation technique exist. Some popular rule-based, statistical, atlas-based, and deformable models-based techniques, and their key strengths and weaknesses are outlined below.

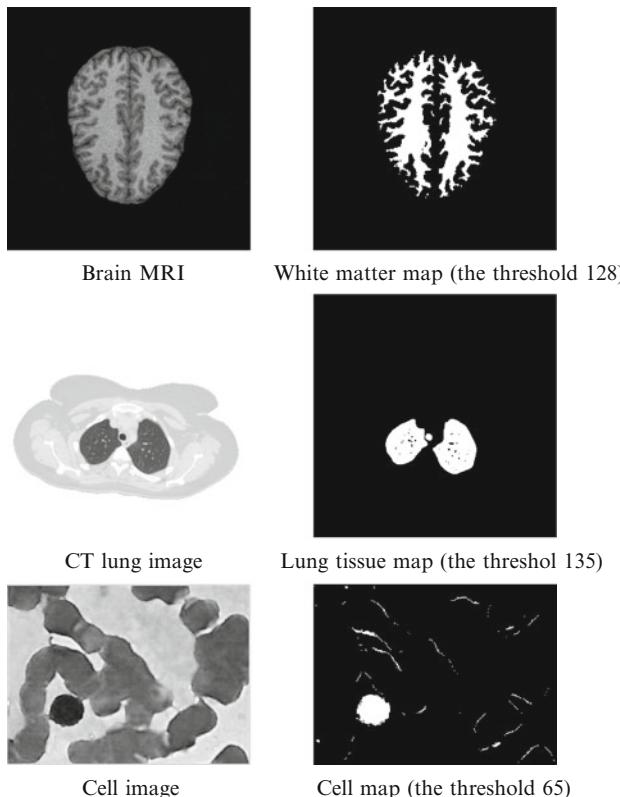
### 1.3.1 Rule-Based Segmentation

In this case, image features over an individual region comply with a set of heuristic rules. Simple and straightforward feature thresholding is widely used, due to its computational simplicity and speed, for fast initial segmentation or at intermediate stages of various segmentation scenarios, but usually it cannot stand alone for the final segmentation. The simplest thresholding divides an image into two regions related to a goal object and its background, respectively: e.g., an object label is assigned to each pixel or voxel if its intensity exceeds a certain threshold; otherwise, it is classified as the background. The threshold can be fixed through all the image (global thresholding) or vary according to the pixel/voxel location

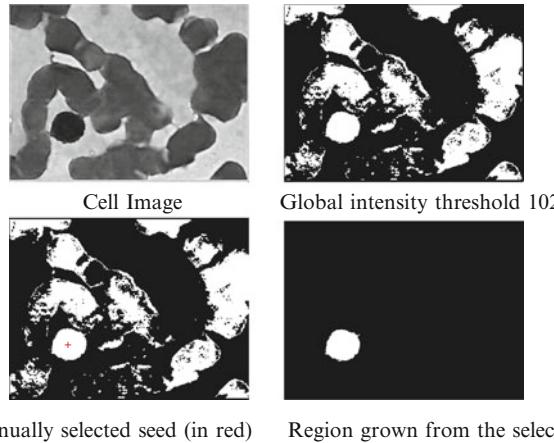
(adaptive, or local thresholding). Frequently, it is selected by statistical analysis of image intensities over the whole image or a certain vicinity of the pixel or voxel under consideration, e.g., peaks and valleys of the gray-level histogram, or by optimizing a certain image-dependent criterion, e.g., minimizing the cross-entropy between an input gray-level image and the output binary image [89], or by statistical inference for probabilistic image models discussed later in Sect. 1.3.2. Figure 1.8 illustrates the global-threshold-based segmentation.

If the intensity distributions for the object and background totally or partially intersect, simple comparisons to a global threshold either fail or produce too inaccurate results. Moreover, the thresholding does not guarantee connectedness of the found objects, which is a basic requirement in many medical imaging and CAD applications. More details about these techniques can be found in [90–92].

*Region-growing* (also called region merging) techniques guarantee a connected region for each segmented object. After initial seeds are selected, their neighbors are added up, and the group continues to grow by including adjacent pixels/voxels that comply with a predefined criterion specifying the required properties of the regions. Segmentation results depend on the latter criterion and rules for selecting



**Fig. 1.8** Segmentation of anatomical structures with a global intensity threshold



**Fig. 1.9** Segmentation by thresholding followed by region growing

the seeds and specifying the neighbors. Figure 1.9 illustrates a simple region-growing procedure. First, a binary image is produced by global thresholding of the initial gray-scale image. Then a seed is selected manually within the region of interest, and begins to grow by testing and adding the immediate eight neighbors having the same properties (the intensity below the threshold). The process is repeated for each added pixel until no more connected neighboring pixels have the same intensity. For textured images, more complex region-growing procedures are to be involved, e.g., using the first- and second-order statistics (e.g., mean and standard deviation) for the current region and the candidate intensity to decide whether the neighbor should be added to the region [92]. The process is repeated and the statistics are recomputed for each added pixel until no more pixels are accepted. Obviously, the region growing techniques are very sensitive to initialization and often need user assistance to select proper seed points.

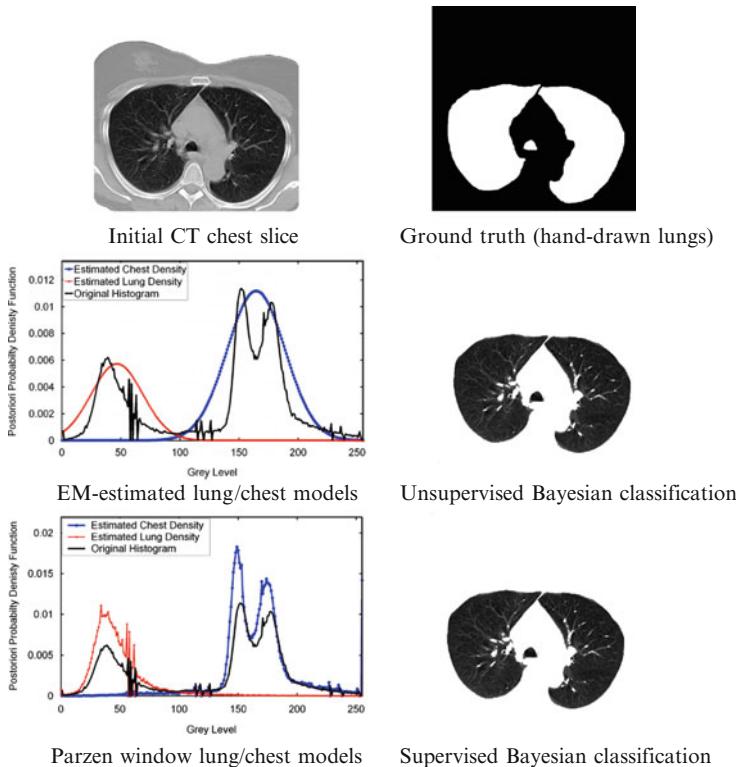
*Region split-and-merge* techniques partition an image initially into a number of regions and then iteratively merge and/or split the current regions in accord with a certain region homogeneity rule. Just as the region growing, the split-and-merge approach is also sensitive to the initialization. However, its known applications to medical image segmentation include large brain lesions [93], cavity deletion [94], retinal blood vessels [95], and pulmonary nodules [96]. Additional information about the split-and-merge and region growing segmentation can be found in [72, 92, 97].

### 1.3.2 Segmentation by Optimal Statistical Inference

Statistical image segmentation involves parametric or nonparametric probability models of appearance and shape of goal objects and optimal, e.g., Bayesian or maximum likelihood inference [29, 98–100]. Popular nonparametric probability

density models are built using the k-nearest neighbor and Parzen-window estimators [101]. Popular parametric models exploit tractable analytical representations (such as Gaussian ones in Sect. 1.2.1) that allow for analytical or computationally feasible numerical parameter learning. In particular, the maximum likelihood estimates (MLE) of parameters of a Gaussian model are analytical, namely, the mean and the covariance matrix for a given set of training samples, while parameters of a Gaussian mixture model (the means, covariance matrices, and prior probabilities of the Gaussian components) are learned in part numerically and in part analytically with expectation-maximization (EM) techniques [99, 100, 102].

Figure 1.10 shows results of segmenting lung tissues on a CT chest image using 1-D probability distributions of intensities as region models. The mixed marginal empirical distribution of pixel-wise intensities over the whole image is approximated and separated into the chest and lung probability densities using either a parametric bi-modal Gaussian mixture model and the unsupervised EM estimation of its parameters or a nonparametric density model, and the supervised Parzen-window estimation of both the classes. Then the obtained chest and lung models are used for Bayesian classification. Given the estimated prior probabilities  $p(k); k \in \{1, 2\}$  of the classes (i.e., relative frequencies of the chest,  $k = 1$ , and



**Fig. 1.10** Segmentation of lung tissues using different probability density models

lung,  $k = 2$ , tissues) and the class models  $p(q|k); q \in \mathbf{Q}$  (i.e., relative frequencies of intensities  $q$  over each class  $k$ ), the classifier relates each pixel  $\mathbf{r} \in \mathbf{R}$  to the most probable class:  $k^* = \arg \max_k p(k|g(\mathbf{r})) = \arg \max_k (p(k)p(g(\mathbf{r})|k))$ . More information about the parametric and nonparametric density estimation methods and their implementations can be found in [101, 103–105].

The above marginal conditional intensity models are very simple and allow for fast unsupervised or even faster supervised learning and fast classification. The main limitation is that the tails of the object distributions usually overlap, so it is difficult to find thresholds (if at all) that separate these objects accurately unless the object intensities considerably differ from the background. However, this is generally not the case in medical images so that most of the applications need more adequate probability models.

Two-level Markov-Gibbs random field (MGRF) image/map models assume that each observed image to be segmented is a sample of two interrelated random processes, the conditional low- and the unconditional high-level one representing the original images  $g$  and their desired region maps  $m$ , respectively. The two-level MGRF is specified by the joint probability density;  $p(g, m) = p(g|m)P(m)$  is the conditional probability density of an image  $g$  given the region map  $m$ , and  $P(m)$  is the unconditional probability distribution of the maps  $m$ . Both the parameter learning and Bayesian inference for this MGRF are considerably simplified if the low level is modeled with a conditionally independent random field of pixel-wise intensities as in the example in Fig. 1.10:  $p(g|m) = \prod_{\mathbf{r} \in \mathbf{R}} p(g(\mathbf{r})|m(\mathbf{r}))$ , or

$$p(g|m) = \prod_{k \in \mathbf{K}} \prod_{q \in \mathbf{Q}} p(q|k)^{|\mathbf{R}_{q,k}|} \equiv \prod_{k \in \mathbf{K}} \prod_{q \in \mathbf{Q}} p(q|k)^{|\mathbf{R}|f_m(k)f_{g,m}(q|k)} \quad (1.1)$$

where,  $\mathbf{R}_{q,k} \subset \mathbf{R}$  denotes a subset of pixels/voxels from the region  $k$  in the map  $m$  having the intensity  $q$  in the image,  $g : \mathbf{R}_k = \bigcup_{q \in \mathbf{Q}} \mathbf{R}_{q,k}$  and  $\mathbf{R} = \bigcup_{k \in \mathbf{K}} \mathbf{R}_k$ . The cardinality  $|\mathbf{R}_{q,k}| = |\mathbf{R}|f_m(k)f_{g,m}(q|k)$  of the subset is apparently related to the empirical frequency  $f_m(k) = |\mathbf{R}_k|/|\mathbf{R}|$  of the region labels  $k$  in the map  $m$  and the empirical frequency  $f_{g,m}(k) = |\mathbf{R}_{q,k}|/|\mathbf{R}_k|$  of the intensity value  $q$  in the image  $g$  over the region  $k$ . Then the individual class densities  $p(q|k)$  can be learned using the nonparametric Parzen-window estimator or parametric EM estimator with a Gaussian mixture model.

Assuming that each region label  $m(\mathbf{r}) \in \{1, 2\}$  depends uniformly on its nearest neighbors, e.g.,  $(x \pm 1, y)$  and  $(x, y \pm 1)$  for  $\mathbf{r} = (x, y)$ , the high-level object/background model  $K = 2$  can be specified by an unconditional Gibbs probability distribution as shown.

$$P(m) = \frac{1}{Z_\gamma} \exp - \sum_{\mathbf{r} \in \mathbf{R}} \left( \sum_{\delta \in N} V(m(\mathbf{r}), m(\mathbf{r} + \delta)) \right) \quad (1.2)$$

where,  $V(k, \kappa); k, \kappa \in \{1, 2\}$ , denotes a Gibbs potential function that quantifies pairwise interactions between the map pixels or voxels ( $V(k, k) = -\gamma$  for equal and  $V(k, \kappa) = \gamma$  for unequal  $(k \neq \kappa)$  labels;  $\gamma \geq 0$ ),  $N$  is a set of the coordinate

offsets to define the pixel/voxel neighbors in the lattice  $\mathbf{R}$  (two offsets  $N = \{(1,0),(0,1)\}$  for the 2-D 4-neighborhood; three offsets  $N = \{(1,0,0),(0,1,0), (0,0,1)\}$  for the 3-D 6-neighborhood, and, respectively, 4 and 13 offsets for the 2-D 8- and 3-D 26-neighborhoods), and  $Z_\gamma$  is the normalizing factor called the partition function. Such a model integrates the information provided by the intensity density estimation with the anticipated spatial interactions between the object pixels or voxels. In spite of simplicity, this two-level MGRF improves in many cases the segmentation accuracy.

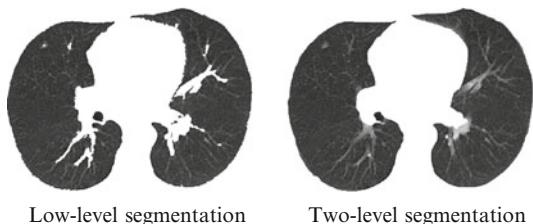
Given the above model (i.e., the value  $\gamma$  and class densities  $p(q|k)$ ) and the Bayesian *maximum a posteriori* (MAP) decision rule with the anti-diagonal unit loss matrix [100], the goal region map  $m^*$  for a given image  $g$  maximizes the posterior probability  $p(m|k)$ , or what is the same, the joint probability  $p(g,m)$ . Equivalently, it maximizes the log-likelihood,  $\mathcal{L}(m|g) = \log p(g,m) = \log p(g|m) + \log P(m)$ , or minimizes the energy  $E(m|g) = E(g|m) + E(m)$  over the search space  $\mathcal{M}$  of all the possible region maps. Here,  $E(g|m) = -\log p(g|m)$  denotes the conditional image energy and  $E(m) = -\log P(m) + \log Z_\gamma$  is the unconditional region map energy.

$$E(m) = \sum_{\mathbf{r} \in \mathbf{R}} \sum_{\delta \in N} V(m(\mathbf{r}), m(\mathbf{r} + \delta)) \equiv |\mathbf{R}| \rho_N \gamma (1 - 2F_{\text{eq}}(m))$$

where,  $\rho_N = |C_N|/|\mathbf{R}|$  is the relative cardinality of the family of pairwise cliques,  $C_N$ , associated with the set  $N$  (e.g.,  $\rho_N \approx 2$  for the 2-D 4-neighborhood and  $\approx 13$  for the 3-D 26-neighborhood) and  $F_{\text{eq}}(m)$  denotes the relative frequency of equal pairs of labels in all the cliques over  $C_N$ .

Figure 1.11 compares qualitatively the accuracy of segmentation with the two-level MGRF model and the low-level intensity model only. In this case, invoking the pairwise interactions between the adjacent pixels of the lungs map helps to fill in gaps caused by the misclassified intensities.

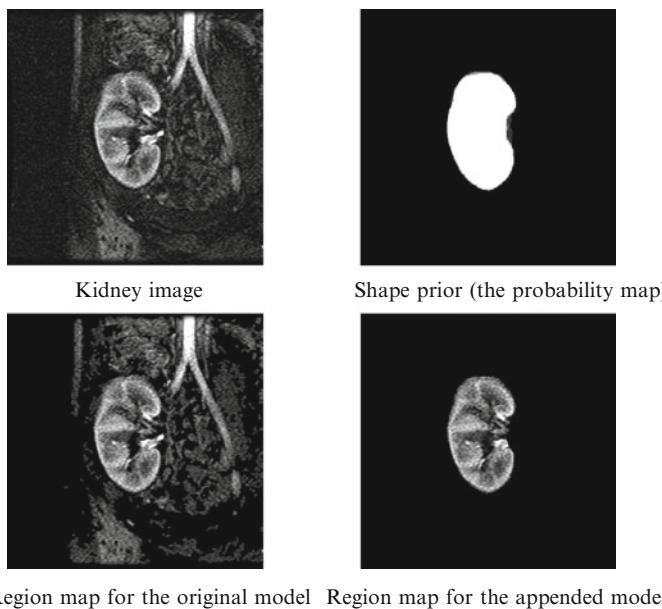
However, large image noise and/or occluded objects could notably decrease the accuracy of image segmentation with this simple appearance-based MGRF model. One possibility to improve it is to add shape constraints for goal objects of specific forms such as most of medical structures. A shape prior introduced in [106] is defined as a probability map  $s : \mathbf{R} \rightarrow [0.0, 1.0]$  built from a coaligned training set of object/background maps. Such a prior is embedded to the conditional low-level image model so that the shape prior components  $s$  ( $\mathbf{r}$  constitute weights of the



**Fig. 1.11** Bayesian segmentation of lungs using the low-level Parzen-window intensity models vs. the two-level image/map MGRF model of (1.1) and (1.2)

conditional densities  $p(g(\mathbf{r})|m(\mathbf{r}))$ ). Unlike the original simple two-level MGRF, an image to be segmented has now to be prealigned to the same training set (i.e., the shape prior acts then as a rejection filter, and the segmentation accuracy depends heavily on the accurate image alignment). Results of segmenting a kidney image using the original and appended two-level MGRF models are shown in Fig. 1.12.

*More advanced MGRF models* attempt to accurately describe the visual appearance of more diverse textured objects to be segmented using at the lower level also an MGRF with pairwise or higher-order intra and interobject interactions (Sect. 1.2.1 refers to a few such models). However, the more complex the two-level MGRF, the larger the computational complexity of statistical inference. In this case, both the joint image-map and posterior map-image models are not necessarily the MGRFs because the probability density,  $P(g|m) = \frac{1}{Z(m)} \exp(-E(g|m))$ , for the conditional image-map MGRF combines the computable energy  $E(g|m)$  of intensity interactions over the image  $g$ , given the region map  $m$ , with the computationally unfeasible partition function  $Z(m)$ , which also depends on the map  $m$ . Thus, unlike the simplified model, in this case, both the maximum likelihood and Bayesian MAP decision rules result in the minimization of the energy augmented with the noncomputable term  $\log Z(m)$ . To circumvent the problem and retain only the pure energy minimization, a single-level conditional random field (CRF) model [107] directly forms a posterior MGRF  $P(m|g) = \frac{1}{Z_g} \exp(-E(m|g))$ , so that the goal map minimizes the computable energy function  $E(m|g)$ .



**Fig. 1.12** Segmentation of a kidney image using the original two-level MGRF model and the same model appended with the shape prior

Search for a local energy minimum close to a heuristically chosen initial position in the search space  $M$  is relatively simple (e.g., pixel/voxel-wise relaxation or gradient-based techniques). But unless the initial position is in the close vicinity of the goal global minimum, such a solution is mostly too far from the latter and thus is too inaccurate. Due to both  $x$ - and  $y$ -directed pixel interactions in 2-D and  $x$ -,  $y$ -, and  $z$ -directed voxel interactions in 3-D region maps, the search for the goal map that converges to the global energy minimum is largely NP-hard. Only in very special cases, e.g., tree-like (loopless) interaction structures, some computationally feasible exact optimization techniques of the polynomial time and space complexity exist, such as max-sum and sum-product algorithms like dynamic programming and belief propagation [100]. Stochastic relaxation-based global optimization (called simulated annealing) [29] has theoretically the exponential convergence time so that the latter can hardly be reached in practice. The fast-converging simplified variant such as iterated conditional modes (ICM) [28] or even simpler relaxation labeling perform only the local optimization.

The iterative graph-cut optimization [108–111] is widely used in image segmentation today. Due to its provable convergence to a solution within a fixed factor from the global minimum, the factor value depending on the pixel/voxel interactions involved, the graph-cut techniques, have considerably improved solutions of various energy minimization problems in image processing, computer vision, and pattern recognition over older (mostly heuristic) techniques [36–42, 112–114]. The graph-cut solution was first applied by Greig et al. [115] to find the global energy minimum for a simple MGRF model of noisy binary images in order to perform the Bayesian de-noising of a given image. Later on, this approach was extended to convex multilabel problems [116] and to approximate global energy minimization for more general multilabel MGRF models [34]. The graph-cut optimization employs the minimum-cut/maximum-flow algorithms for directed networks of the polynomial (e.g., from quadruple to cubic by the lattice cardinality  $|R|$ ) time complexity, e.g., the Ford–Fulkerson path augmenting and Goldberg–Tarjan push-relabel techniques. A comprehensive comparison of graph-cut algorithms for discrete energy minimization problems in image analysis can be found in [35].

### 1.3.3 *Atlas Based Segmentation*

The use of anatomical atlases as reference images to guide segmentation of new images is very popular in different medical applications, e.g., for segmenting brain and its internal structures or segmenting pathological lungs, lung lobes, heart and aorta, and internal abdominal organs [48, 117–133]. The atlas typically depicts prototypical locations and shapes of anatomical structures together with their spatial relations [134]. All the known atlas-based methods can be classified into single and multi atlas-based segmentation.

*Single atlas-based segmentation* uses an atlas constructed from one or more labeled segmented images. Once the atlas is created, it is registered to the target

image, and the goal region map is obtained by so called label propagation that transfers the labels from the atlas back to the image using the same geometric mapping as the registration. Obviously, the segmentation accuracy depends on the registration (if the latter fails, so does the segmentation). The registration always involves time consuming and complex local deformations. Also, the segmentation is affected by the ability of the atlas to represent the whole population of images under consideration.

A single image to construct the atlas can be selected randomly, or by visual inspection based on practical criteria, or made artificially [131]. If the atlas is constructed from several images, one image can be selected as a reference and all other images are registered to it. To increase the signal-to-noise ratio, all the registered images are averaged, and the segmented average image is used as the atlas [124]. Alternatively, the atlas can be built by transforming the reference to the average image and segmenting the transformed reference [119]. Probabilistic atlases built by averaging the transformed images and analyzing the corresponding labels [120, 123] provide different weights of each pixel. However, an average atlas does not handle elastic deformations of internal structures during the registration process. To overcome this problem, Leemput [129] proposed a mesh-based atlas representation instead of the average atlas. Also, an iterative atlas generation uses the output of each iteration as the input of the next iteration [134].

*Multi atlas-based segmentation* registers many independently built atlases to a target image and then combines their segmentation labels. The underlying idea is that fusion of multiple independent classifiers might produce better classification [135]. There exist different ways for segmenting a particular target image, e.g., to select all the atlases or only their subset as well as to choose one or another strategy of combining the selected atlases to produce the goal region map. The pre or postregistration selection of atlases can be based on certain matching criteria such as the mutual information or the degree of deformation of the object of interest (obviously, the atlases of the highest local mutual information or the least object deformation are preferable).

Popular strategies of combining the selected atlases to segment the target image include decision fusion (also called majority voting, majority rule, or label voting). In this strategy, the label of each pixel or voxel is selected as the label that most of the segmentations agree on [132]. Another strategy, called simultaneous truth and performance level estimation (STAPLE), evaluates the performance of each classifier iteratively, weighs the corresponding segmentation accordingly, and uses the EM approach to find the best final segmentation [121]. Isgum et al. [131] combined the propagated labels by spatially variant decision fusion weights derived from the local assessment of the registration accuracy, and Rohlfing and Maurer [48] proposed a shape-based averaging strategy based on the Euclidean distance map to perform the combining.

In general, the accuracy of the atlas-based segmentation depends on the accuracy of the atlas-to-target registration (i.e., the segmentation problem is reduced to the registration one). To provide accurate segmentation results, local elastic registration is always mandatory, although it is time consuming and source exhausting.

The single atlas-based segmentation could suffer from the possible insufficient representation of the whole image population. The multi atlas approach can overcome this drawback, but it requires a large number of atlas-to-target registration steps in order to produce the final segmentation.

Recent methods integrate the created atlases into larger frameworks that account for other image features in order to improve the segmentation accuracy. Marroquin et al. [118] proposed an EM-based classification that relies on an atlas and image intensities to segment brain and cerebellum volumes from 3-D MRI. Park et al. [120] incorporated a probabilistic atlas into Bayesian framework to segment four abdomen organs from thoracic CT images. Sluimer et al. [123] added a voxel classification step after an atlas-based segmentation step in order to refine the segmentation of pathological lungs from CT images. Zhang et al. [125] segmented the oblique labor fissure from CT chest image using a fuzzy reasoning framework that incorporates intensity, anatomy smoothness constraints, and an atlas.

### **1.3.4 Segmentation with Deformable Models**

Since the seminal paper by Kass et al. [136], deformable models became a dominant technique and gave rise to one of the most dynamic and successful areas in image segmentation, edge detection, shape modeling, and tracking of medical image structures. These techniques have quickly gained widespread popularity due to superiority over other segmentation techniques. In particular, these models end up with a continuous boundary of an object of interest in spite of possible large shape variations, image noise and inhomogeneities, and discontinuous object boundaries due to occlusions. In addition, they incorporate both the shapes and appearances of the objects of interest as their most significant discriminative features, and can be aided by interactions with the users. A deformable or active model is a curve in a 2-D digital image or a surface in a 3-D image that evolves to outline a goal object. Its evolution is guided by internal and external forces, which are defined so that the deformable model will eventually conform to the object boundary. The internal forces, coming from the curve/surface itself, are designed to keep the evolved curve smooth and unified. The external forces, depending on the image, propagate the evolution towards the object boundary. By representation and implementation, the deformable models are divided into two broad classes: parametric (e.g., [49–64, 136–194]) and geometric (e.g., [45–47, 195–227]).

*Parametric deformable models* are contours or surfaces specified in an explicit parametric analytical form [228]. For example, a conventional parametric deformable 2-D model by Kass et al. [136], called also an active contour, or snake, is a single-parameter curve  $\mathbf{r}_m = (x_m, y_m)$  with  $M$  control points,  $[\mathbf{r}_m : m = 1, \dots, M]$ , the integer index  $m$  being generalized into a real number in the interval  $\mathbf{M} = [1, M]$  to indicate continuous positions  $\mathbf{r}_m$  from  $\mathbf{r}_1$  to  $\mathbf{r}_M$  around the curve. The curve moves through the continuous image plane to minimize the total energy

$$\varepsilon_{\text{tot}} = \varepsilon_{\text{int}} + \varepsilon_{\text{ext}} = \int_{m \in M} \zeta_{\text{int}}(\mathbf{r}_m) + \zeta_{\text{ext}}(\mathbf{r}_m) dm \quad (1.3)$$

where,  $\zeta_{\text{int}}(\mathbf{r}_m)$  and  $\zeta_{\text{ext}}(\mathbf{r}_m)$  denote the internal and external forces, respectively, that control the point-wise movements of the point  $\mathbf{r}_m = (x_m, y_m)$  along the curve. The internal energy term keeps the curve as a single continuous unit and the external energy term attracts the model to the goal object boundary. The total internal force is typically defined as  $\zeta_{\text{int}}(\mathbf{r}_m) = \alpha_{\text{int}} |\mathbf{r}'_m|^2 + \beta_{\text{int}} |\mathbf{r}''_m|^2$ , where the weights  $\alpha_{\text{int}}$  and  $\beta_{\text{int}}$  control the curve's tension and rigidity, respectively, and  $\mathbf{r}'_m$  and  $\mathbf{r}''_m$  denote the first and second derivatives of the curve in the point  $\mathbf{r}_m$  by the positional parameter  $m$ . The external forces  $\zeta_{\text{ext}}(\mathbf{r}_m) = -|\nabla g(\mathbf{r}_m)|^2$  or  $-|\nabla[G * g(\mathbf{r}_m)]|^2$  where  $\nabla$ ,  $G$ , and  $*$  denote a gradient operator, a 2-D Gaussian smoothing kernel, and the kernel-image convolution, respectively, have been used in [136] to lead an active contour toward stepwise intensity edges in a grayscale image  $g$ .

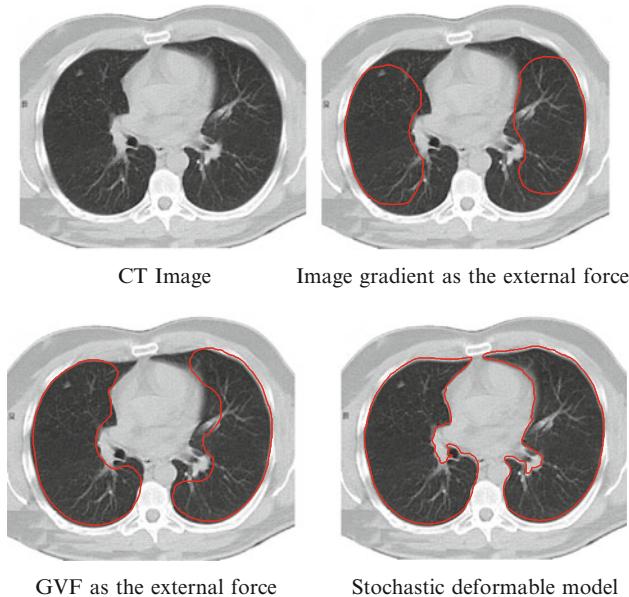
The model performance depends obviously on adequate choices of their parametric forms and forces. Known parametric models employ a wide range of formal representations, initializing mechanisms, user interaction modes, driving forces to guide evolution, and optimality criteria for final contours found by evolution. In particular, deformable curves and surfaces are represented by superquadrics, splines, spectral and point distribution models, etc. Their initialization is either manual or automatic, e.g., by using Hough transform, statistical region or contour models, and region-growing techniques. The evolution guidance could be interactive or fully automatic, and the external driving forces can account for different image and/or object properties such as e.g., intensity edges, gradient vector flows, appearance and shape features, and so on.

Superquadrics [229] are second-order algebraic surfaces that can be utilized to model only symmetric shapes such as ellipsoids, spheres, cylinders, and parallelepipeds [58]. For example, Gong et al. [57] model a prostate with a superellipsoid. The main advantage is that a deformable model is represented with only a few parameters. Also, superquadrics can be combined with free-form splines to add the capability to taper, bend, and twist the model through global deformation [141]. Spline models with many local degrees of freedom through their shape control variables were used to represent different medical structures such as brain structures [54], left ventricle surfaces [147], and vessel boundary [154]. Point distribution models take account of the object's shape variability. Cootes et al. [64] developed a statistical point distribution model, called an active shape model (ASM), to represent specific object shapes. A number of landmarks are labeled in multiple training images of an object, and the mean positions of the landmarks and their modes of variation establish the ASM. The model gained popularity and was applied to different medical structures [49, 52, 55, 193]. Its two main limitations are that one-to-one point-wise correspondence between the different boundaries of the object in the training images should be established first and a large number of landmarks is required for representing complex shapes. Many extensions to the ASM model have been proposed, such as e.g., an active appearance model (AAM)

[156] and an oriented active shape model [62]. One more way for representing the deformable models is to use Fourier descriptors or spherical harmonics. Fourier descriptors model shapes by compact decomposition via sinusoidal basis functions. The 2-D Fourier descriptors were originally proposed by Persoon et al. [230]. To express fine details, more sinusoidal basis functions are needed. Staib and Duncan [142, 143] used them for 2-D segmentation and limited to star-shaped surfaces 3-D segmentation. Szekely et al. [152] used Fourier descriptors to represent spherical object topology with spherical harmonics. Spherical harmonic basis can be used instead of sinusoidal one to represent closed surfaces of spherical topology. Like with the Fourier descriptors, fine details of an object can be captured with a large enough number of harmonics.

Generally, all the parametric deformable models are initialization-sensitive, and different initializations may produce different results. In the majority of models, the initial contour or surface should be placed initially close to the goal object boundary. Manual delineation of an initial contour is one solution. In addition, the users may initialize only two or more points as seeds for the initial contour (e.g., in [60, 183]). Fully automatic initialization has also been used, even though it could involve extensive complex computations. Atkins et al. [159] initiated a deformable model for fully automated brain segmentation in MRI by an intensity histogram analysis followed by nonlinear anisotropic filtering. Madabhushi and Metaxas [172] used a probabilistic intensity and texture based classifier followed by region growing to obtain an initial segmentation of breast lesions in ultrasound images. Ecabert et al. [8] used a 3-D Hough transform to automatically initialize a deformable model for heart segmentation in CT images. Prior information from training images about the goal object was also employed to automatically determine initial positions of parametric contour or surface models, e.g., the mean object shape obtained from the training images was used for model initialization in [44, 51, 52, 63, 187].

Depending on the application, certain internal and external forces can be added to or excluded from the model. Davatzikos and Prince [148] used only elastic internal forces to find the human cortex. A fast and precise boundary approximation by Xu and Prince [157] introduced a gradient vector flow (GVF) as a new external force. Due to a larger capture range, the GVF allows a snake to move into the boundary concavities. Another fast and precise snake by Mokhtarian and Mohanna [165] removed the curvature part from the internal energy and used instead the curvature scale-space filtering to smooth the boundary. A class of probabilistic guiding forces in [59] offers the fast and robust model evolution due to the use of joint probabilities of signals and region labels in individual pixels as external guiding forces. The probabilities are derived from a two-level MGPF image/map model detailed in Sect. 1.3.2. Conditional probability distributions  $p(q|k)$  of signals in each region  $k$  are recovered from a mixed empirical signal distribution over the whole image using accurate mixture approximation with a linear combination of sign-alternate Gaussian densities (LCG) or discrete Gaussians (LCDG) [105]. The resulting class of stochastic deformable models outperforms some conventional ones because of automated initialization, insensitivity to the initialization, and



**Fig. 1.13** Lung segmentation in a CT image with the gradient-based, GVF-based, and stochastic deformable models

ability to follow complex boundary shapes with concavities. Figure 1.13 compares the segmentation of a 2-D low-dose CT image with the latter model to results obtained with the conventional snakes that use the image gradient or the gradient vector flow (GVF) as their external forces.

Live-wire technique for 2-D objects (e.g., [144]) and its extension, known as wirelane, for 3-D objects (e.g., [163, 164]) guide a deformable model by simple interactive operations to facilitate the user's manual segmentation.

To guide the evolution towards the optimal, in terms of its total energy, contour, the original iterative energy minimization in [136] and further in [139] was based on a finite-difference solution of closed-form Eulerian differential equations. However, the desired global minimum could be unstable, its iterative search is usually trapped in multiple local minima, and it is difficult to constrain both the forces and the whole energy. Amini et al. [140] pointed to these shortcomings and improved the minimization in [136] by representing a snake as a linked chain of control points and minimizing the total energy of the chain with dynamic programming. This approach allowed for rigid constraints on the energy function, so that the minimization became more stable. However, the process was too time consuming and needed a very careful adjustment of its control parameters. Staib et al. [150] segmented medical structures by the conjugate gradient search for an optimal 3-D active surface. Still, the gradient search is as computationally expensive as the dynamic programming. The greedy algorithm by Williams and Shah [145] simplify computations because of its linear time complexity in the numbers of control points

and their neighbors taken into account for energy minimization. This algorithm is much more stable and simultaneously more than an order of magnitude faster than the previous techniques. However, it could be stuck in the local minimum. An alternative finite element based optimization divides the image lattice into a large number of connected sublattices. The object shape is constructed via a mesh of lattice nodes separating these sublattices. This approach has limited approximation continuity along the boundaries, and it is difficult to generate the mesh from the consistent sublattices [231]. To closely approach a realistic boundary with concavities, Grzeszczuk et al. [155] control the snake evolution by simulated annealing that in theory will eventually reach the global energy minimum and in principle escapes local traps. Actually, as underlined in [35], the simulated annealing solutions typically stop very far from the global minimum. The iterative graph-cut algorithms by Boykov and Kolmogorov [232] produce geodesic contours or 3-D surfaces such that their total energies are probably within a fixed factor from the global minimum under rather flexible boundary conditions.

Flexible *geometric level-set based deformable models* introduced by Osher and Sethian [195] are popular and powerful parameterization-independent tools for segmenting medical images. An object-background boundary in the  $(x, y)$ -plane is represented at each time instant  $t$  by zero level  $\phi_t(x, y) = 0$  of an implicit level-set function  $\phi_t(x, y)$ . The latter is a map of signed minimum Euclidean distances from every point  $(x, y)$  to the boundary, the distance being negative for an interior and positive for an exterior point [195]. The distance map is evolving iteratively [214]:

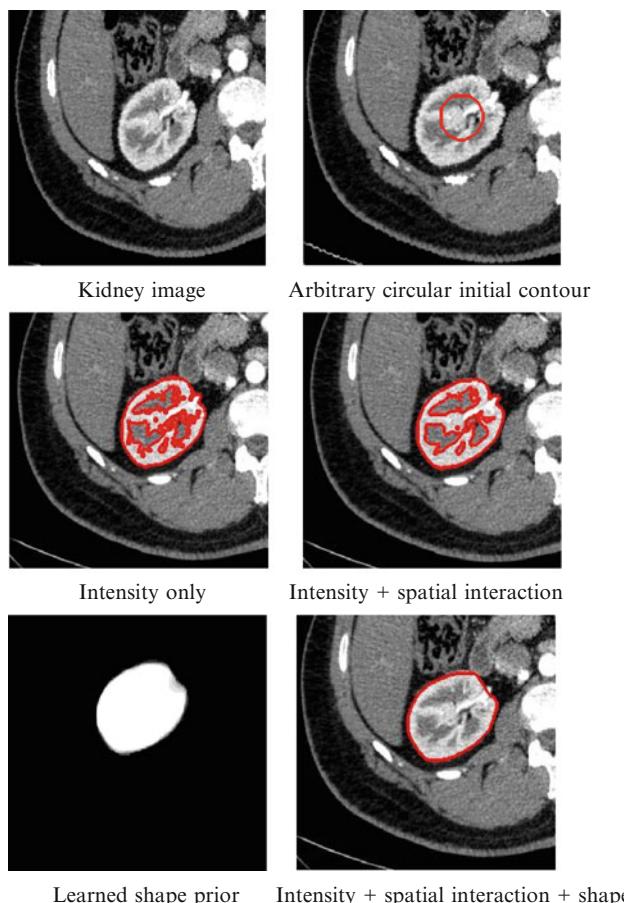
$$\phi_{k+1}(x, y) = \phi_k(x, y) - \tau V_k(x, y) |\nabla \phi_k(x, y)| \quad (1.4)$$

where,  $k$ , indicates iterations, or discrete instants  $t = k\tau$  with a step  $\tau > 0$ ,  $V_k(x, y)$  is an evolution speed at iteration  $k$ , and  $\nabla \phi_k(x, y) = \left[ \frac{\partial \phi_k}{\partial x}, \frac{\partial \phi_k}{\partial y} \right]$  is the gradient of  $\phi_k(x, y)$ .

Conventional speed functions  $V_k(x, y)$  accounting for image intensities (e.g., [203]), object edges (e.g., [197]), an intensity gradient vector flow (e.g., [157]), etc., are unsuccessful on very noisy images with low object-background intensity gradients. Additional prior shape knowledge results in more efficient segmentation (see e.g., [45, 204, 220]). Leventon et al. [65] and Shen and Davatzikos [198] augmented speed functions with terms attracting the level set front to more likely shapes specified by analyzing principal components of a training set of objects. Chen et al. [204] used a prior “average shape,” whereas Tsai et al. [45] approximated an object shape with a linear combination of distance maps for a set of mutually aligned training images of goal objects. However, due to a very large dimensionality  $|\mathbf{R}|$  of the distance maps, only a few top rank principal components can be included into the linear combination to keep the model computationally feasible. To overcome over-segmentation, Cheng et al. [220] incorporated a prior shape knowledge into the improved Chan-Vese’s level set model [200] using a variational approach. However, the speed functions based solely on the shape priors result in notable segmentation errors under discontinuous object boundaries and large image noise and inhomogeneities (see e.g., [196]). Also, the approximation of

a distance map  $\phi_k(x, y)$  with the closest linear combination of the maps from a set of coregistered training images (see e.g., [45, 204, 220]) has a serious inherent drawback; because the space of distance maps is not closed with respect to linear operations, a linear combination of the training maps may result in a physically impossible map so that its zero level may produce an invalid or even physically impossible boundary. A stochastic level set framework by Khalifa et al. [227] circumvents this problem (some segmentation results for this approach are shown in Fig. 1.14). More about the geometric deformable models can be found in [233].

Deformable models often rely on proper threshold-, region-, and/or prior-based initialization and are embedded into larger data processing frameworks to ensure accurate segmentation. Pang et al. [177] combined a bi-elliptical deformable template with an active contour model to segment tongue in Chinese medicine.



**Fig. 1.14** Kidney segmentation with geometric deformable models using level set functions driven by image intensities, their spatial interactions, and a shape prior learned from a coaligned set of training kidney images

Yang et al. [46] built a Bayesian MAP segmentation framework using the joint prior information about the object shape and appearance (in terms of intensities), the shape-intensity model having been specified by level-set functions. Tsai et al. [210] used an expectation-maximization (EM) based classification of a shape database with simultaneous estimation of the best classes' shapes through the level-set contours. Vukadinovic et al. [225] combined a level-set framework, a feature classifier, and a 2-D elliptic deformable shape model to segment the outer vessel wall of the common carotid artery in CT angiography.

Hybrid segmentation-by-registration frameworks use image registration to aid the segmentation. Close relations between the segmentation and registration are clearly outlined by Yezzi et al. [207]. Zagrodsky et al. [184] used a mutual-information based registration to initialize a wire-mesh template to segment the left-ventricle from 3-D echocardiography data. Baillard et al. [201] used an automatic registration to initialize an evolving 3-D surface to segment the brain MRI. Vemuri et al. [206] used a level-set deformable model to segment the hippocampus shape in 3-D brain MRI by registration using an atlas. Simultaneous registration-segmentation frameworks have been also proposed: e.g., Yezzi et al. [207] used the deformable level-set models to simultaneously segment and register objects in multiple MRI and CT images.

## 1.4 Concluding Remarks

Accurate segmentation of objects of interest is one of the basic requirements of any medical imaging and CAD system. At present, a variety of different shape/appearance features and decision techniques based on these features are developed, tested, and used for solving application-specific segmentation problems. Due to the rapid development and extension of medical image modalities, new segmentation problems are emerging and new methodologies are being introduced and explored for their solution. The most successful approaches combine multiple image/object features and data processing techniques. However, though experiments bring more accurate results, the segmentation often becomes too complex and time consuming. The likely future research will focus on accurate, robust to image noise, and fast segmentation techniques that can be used in practical CAD systems.

## References

1. Pham DL, Xu C, Prince JL (2000) A survey of current methods in medical image segmentation. *Annu Rev Biomed Eng* 2:315–338
2. Gonzalez RC, Woods RE (2008) Digital image processing, 3rd edn. Pearson Prentice Hall, Upper Saddle River, NJ
3. Farag A, Ahmed MN, El-Baz A, Hassan H (2005) Advanced segmentation techniques. In: Wilson DL, Laxminarayan S, Suri JS (eds) *Handbook of biomedical image analysis*,

- volume I: segmentation models. Kluwer Academic/Plenum Publishers, New York, pp 479–534, Chapter 9
4. Prastawa M, Bullitt E, Gerig G (2009) Simulation of brain tumors in MR images for evaluation of segmentation efficacy. *Med Image Anal* 13(2):297–311
  5. El-Baz A, Farag A, Gimel'farb G, Falk R, El-Ghar M, Eldiasty T (2006) A framework for the automatic segmentation of lung nodules from low dose chest CT scans. In: Proceedings of the international conference on pattern recognition (ICPR'06), Hong Kong, 20–24 August 2006, pp 611–614
  6. Grau V, Mewes AUJ, Alcaniz M, Kikinis R, Warfield SK (2004) Improved watershed transform for medical image segmentation using prior information. *IEEE Trans Med Imaging* 23(4):447–458
  7. Greenspan H, Ruf A, Goldberger J (2006) Constrained Gaussian mixture model framework for automatic segmentation of MR brain images. *IEEE Trans Med Imaging* 25(9):1233–1245
  8. Ecabert O, Peters J, Schramm H, Lorenz C, von Berg J, Walker MJ, Vembar M, Olszewski ME, Subramanyan K, Lavi G, Weese J (2008) Automatic model-based segmentation of the heart in CT images. *IEEE Trans Med Imaging* 27(9):1189–1201
  9. El-Baz A, Farag A, Yuksel S, Abou El-Ghar M, Eldiasty T, Ghoneim M (2007) Application of deformable models for the detection of acute renal rejection. In: Suri JS, Farag A (eds) Handbook of parametric and geometric deformable models: biomedical and clinical applications, II. Springer, New York, pp 293–333, Chapter 10
  10. Laws KI (1980) Textured image segmentation. Ph.D. Thesis, University of Southern California, Los Angeles, CA
  11. Malik J, Perona P (1990) Preattentive texture discrimination with early vision mechanisms. *J Opt Soc Am* 7(5):923–932
  12. Unser M, Eden M (1990) Nonlinear operators for improving texture segmentation based on features extracted by spatial filtering. *IEEE Trans Syst Man Cybern B Cybern* 20(4):804–815
  13. Schowengerdt RA (2007) Remote sensing: models and methods for image processing, 3rd edn. Elsevier, Burlington, MA
  14. Campbell FW, Robson J (1968) Application of Fourier analysis to the visibility of gratings. *J Physiol* 197(3):551–566
  15. Coggins JM, Jain AK (1985) A spatial filtering approach to texture analysis. *Pattern Recognit Lett* 3(3):195–203
  16. Smith GM (1998) Image texture analysis using zero crossing information. Ph.D. Thesis, University of Queensland, Australia
  17. Choi H, Baraniuk RG (2001) Multiscale image segmentation using wavelet-domain hidden Markov models. *IEEE Trans Image Process* 10(9):1309–1321
  18. Jin Y, Angelini E, Laine A (2005) Wavelets in medical image processing: denoising, segmentation, and registration. In: Suri JS, Wilson DL, Laxminarayan S (eds) Handbook of biomedical image analysis, vol I. Kluwer Academic/Plenum Publishers, New York, pp 305–358, Segmentation Models, Chapter 6
  19. Mandelbrot BB (1983) The fractal geometry of nature. W. H. Freeman and Company Limited, New York
  20. Barnsley M (1993) Fractals everywhere. Academic, Boston, MA
  21. Haindl M (1991) Texture synthesis. *CWI Q* 4:305–331
  22. Li H, Liu KJR, Lo S-CB (1997) Fractal modeling and segmentation for the enhancement of microcalcifications in digital mammograms. *IEEE Trans Med Imaging* 16(6):785–798
  23. Chaudhuri B, Sarker N, Kundu P (1993) Improved fractal geometry based texture segmentation technique. *IEEE Proc Comput Digit Tech* 140(5):223–241
  24. Xia Y, Feng D, Zhao R (2006) Morphology-based multifractal estimation for texture segmentation. *IEEE Trans Image Process* 15(3):614–623
  25. Dubes RC, Jain AK (1989) Random field models in image analysis. *J Appl Stat* 16(2):131–164

26. Chellappa R (1989) Two dimensional discrete Gaussian Markov random field models for image processing. *J Inst Electron Telecommun Eng* 35(2):114–120
27. Kashyap R (1984) Characterization and estimation of two dimensional ARMA models. *IEEE Trans Inform Theory* 30(5):736–745
28. Besag J (1974) Spatial interaction and statistical analysis of lattice systems. *J R Stat Soc Series B Methodol* 36(2):192–236
29. Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images. *IEEE Trans Pattern Anal Mach Intell* 6(6):721–741
30. Gimel'farb G (1996) Texture modeling by multiple pairwise pixel interactions. *IEEE Trans Pattern Anal Mach Intell* 18(11):1110–1114
31. Zhu SC, Wu Y, Mumford D (1998) Filters, random fields, and maximum entropy (FRAME): towards a unified theory for texture modeling. *Int J Comput Vis* 27(2):107–126
32. El-Baz A, Farag AA (2003) Image segmentation using GMRF models: parameters estimation and applications. In: Proceedings of IEEE international conference on image processing (ICIP'03), Barcelona, Spain, 14–18 September 2003, vol. II, pp 177–180
33. Roth S, Black MJ (2009) Fields of experts. *Int J Comput Vis* 82(2):205–229
34. Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 23(11):1222–1239
35. Boykov Y, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans Pattern Anal Mach Intell* 26(9):1124–1137
36. Veksler O (2007) Graph cut based optimization for MRFs with truncated convex priors. In: IEEE conference on computer vision and pattern recognition (CVPR'07), Minneapolis, MN, 17–22 June 2007
37. Kohli P, Kumar MP, Torr PHS (2007) P3 & beyond: solving energies with higher order cliques. In: IEEE conference on computer vision and pattern recognition (CVPR'07), Minneapolis, MN, 17–22 June 2007
38. Kohli P, Ladicky L, Torr P (2008) Robust higher order potentials for enforcing label consistency. In: IEEE conference on computer vision and pattern recognition (CVPR'08), Anchorage, AK, 23–28 June 2008
39. Ramalingam S, Kohli P, Alahari K, Torr P (2008) Exact inference in multi-label CRFs with higher order cliques. In: IEEE conference on computer vision and pattern recognition (CVPR'08), Anchorage, AK, 23–28 June 2008
40. Ali AM, Farag AA, Gimel'farb GL (2008) Optimizing binary MRFs with higher order cliques. In: Proceedings of the 10th European conference on computer vision (ECCV'08), Marseille, France, 12–18 October 2008, part III, pp 98–111
41. Ishikawa H (2009) Higher-order clique reduction in binary graph cut. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR'09), Miami, FL, 20–25 June 2009, pp 2993–3000
42. Rother C, Kohli P, Feng W, Jia J (2009) Minimizing sparse higher order energy functions of discrete variables. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR'09), Miami, FL, 20–25 June 2009, pp 1382–1389
43. Rousson M, Paragios N, Shape priors for level set representations. In: Proceedings of the 7th European conference on computer vision (ECCV'02), Denmark, Copenhagen, 27 May–2 June 2002, part II, pp 78–92
44. Shen D, Zhan Y, Davatzikos C (2003) Segmentation of prostate boundaries from ultrasound images using statistical shape model. *IEEE Trans Med Imaging* 22(4):539–551
45. Tsai A, Yezzi A Jr, Wells W, Tempany C, Tucker D, Fan A, Grimson WE, Willsky A (2003) A shape based approach to the segmentation of medical imagery using level sets. *IEEE Trans Med Imaging* 22(2):137–154
46. Yang J, Staib LH, Duncan JS (2004) Neighbor-constrained segmentation with level set based 3-D deformable models. *IEEE Trans Med Imaging* 23(8):940–948

47. Tsai A, Wells W, Tempany C, Grimson E, Willsky A (2004) Mutual information in coupled multi-shape model for medical image segmentation. *Med Image Anal* 8(4):429–445
48. Rohlfing T, Maurer CR Jr (2007) Shape-based averaging. *IEEE Trans Image Process* 16(1): 153–161
49. Duta N, Sonka M (1998) Segmentation and interpretation of MR brain images: an improved active shape model. *IEEE Trans Med Imaging* 17(6):1049–1062
50. Shen D, Herskovits EH, Davatzikos C (2001) An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures. *IEEE Trans Med Imaging* 20(4): 257–270
51. Tao X, Prince JL, Davatzikos C (2002) Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain. *IEEE Trans Med Imaging* 21(5):513–524
52. van Ginneken B, Frangi AF, Staal JJ, ter-Haar Romeny BM, Viergever MA (2002) Active shape model segmentation with optimal features. *IEEE Trans Med Imaging* 21(8):924–933
53. Joshi S, Pizer S, Fletcher PT, Yushkevich P, Thall A, Marron JS (2002) Multiscale deformable model segmentation and statistical shape analysis using medial descriptions. *IEEE Trans Med Imaging* 21(5):538–550
54. Pitiot A, Toga AW, Thompson PM (2002) Adaptive elastic segmentation of brain MRI via Shape-model-guided evolutionary programming. *IEEE Trans Med Imaging* 21(8):910–923
55. Behets G, Maes F, Vandermeulen D, Suetens P (2002) Deformable evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models. *Med Image Anal* 6(1):47–62
56. Davatzikos C, Tao X, Shen D (2003) Hierarchical active shape models, using the wavelet transform. *IEEE Trans Med Imaging* 22(3):414–423
57. Gong L, Pathak SD, Haynor DR, Cho PS, Kim Y (2004) Parametric shape modeling using deformable super ellipses for prostate segmentation. *IEEE Trans Med Imaging* 23(3): 340–349
58. Tutar IB, Pathak SD, Gong L, Cho PS, Wallner K, Kim Y (2006) Semiautomatic 3-D prostate segmentation from TRUS images using spherical harmonics. *IEEE Trans Med Imaging* 25(12):1645–1654
59. El-Baz A, Gimel’farb G (2008) Image segmentation with a parametric deformable model using shape and appearance priors. In: IEEE conference on computer vision and pattern recognition (CVPR’08), Anchorage, AK, 23–28 June 2008
60. Shi Y, Qi F, Xue Z, Chen L, Ito K, Matsuo H, Shen D (2008) Segmenting lung fields in serial chest radiographs using both population-based and patient-specific shape statistics. *IEEE Trans Med Imaging* 27(4):481–494
61. Nguyen H, Ji Q (2008) Shape-driven three-dimensional watersnake segmentation of biological membranes in electron tomography. *IEEE Trans Med Imaging* 27(5):616–628
62. Liu J, Udupa JK (2009) Oriented active shape models. *IEEE Trans Med Imaging* 28(4): 571–584
63. Yan P, Xu S, Turkbey B, Kruecker J (2010) Discrete deformable model guided by partial active shape model for TRUS image segmentation. *IEEE Trans Biomed Eng* 57(5): 1158–1166
64. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. *Comput Vis Image Understanding (GVIU)* 61(1):38–59
65. Leventon ME, Grimson WEL, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR’2000), Hilton Head Island, SC, 13–15 June 2000, pp 316–323
66. Chen Y, Thiruvenkadam S, Tagare HD, Huang F, Wilson D, Geiser EA (2001) On the incorporation of shape priors into geometric active contours. In: Proceedings of IEEE workshop on variational and level set methods, Vancouver, BC, Canada, 2001, pp 145–152
67. Litvin A, Karl WC (2003) Levelset based segmentation using data driven shape prior on feature histograms. In: IEEE workshop on statistical signal processing, pp 166–169

68. Tsai A, Wells W, Warfield SK, Willsky A (2004) Level set methods in an EM framework for shape classification and estimation. In: Proceedings of the international conference on medical image computing and computer-assisted intervention (MICCAI'04), Saint-Malo, France, 26–30 September 2004, pp 1–9
69. Yang J, Duncan JS (2004) 3D image segmentation of deformable objects with joint shape-intensity prior models using level sets. *Med Image Anal* 8(3):285–294
70. Krim H, Yezzi A Jr (2006) Statistics and analysis of shapes (modeling and simulation in science engineering and technology). Springer Science and Business Media LLC, Boston
71. Davies R, Twining C, Taylor C (2008) Statistical models of shape optimisation and evaluation. Springer-Verlag London Limited, London, UK
72. Rogowska J (2000) Overview and fundamentals of medical image segmentation. In: Bankman I (ed) *Handbook of medical image processing and analysis*. Elsevier, Amsterdam, The Netherlands, pp 69–85
73. Marr D, Hildreth E (1980) Theory of edge detection. *Proc R Soc London Series B Biol Sci* 207(1167):187–217
74. Canny J (1986) A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell* 8(6):679–698
75. Grigorescu SE, Petkov N, Kruizinga P (2002) Comparison of texture features based on Gabor filters. *IEEE Trans Image Process* 11(10):1160–1167
76. Duda RO, Hart BE (1972) Use of the Hough transformation to detect lines and curves in pictures. *Commun ACM* 15(1):11–15
77. Lester JM, Williams HA, Weintraub BA, Brenner JF (1978) Two graph searching techniques for boundary finding in white blood cell images. *Comput Biol Med* 8(4):293–308
78. Thedens DR, Skorton DJ, Feagle SR (1995) Methods of graph searching for border detection in image sequences with applications to cardiac magnetic resonance imaging. *IEEE Trans Med Imaging* 14(1):42–55
79. Martelli A (1972) Edge detection using heuristic search methods. *Comput Graph Image Process* 1(2):169–182
80. Digabel H, Lantuejoul C (1978) Iterative algorithms. In: Chermant J-L (ed) 2nd European symposium on quantitative analysis of microstructures in material science, biology and medicine, Riederer Verlag, Stuttgart, Germany, pp 85–99
81. Krivanek A, Sonka M (1998) Ovarian ultrasound image analysis: follicle segmentation. *IEEE Trans Med Imaging* 17(6):935–944
82. Karvelis PS, Tzallas AT, Fotiadis DI, Georgiou I (2008) A multichannel watershed-based segmentation method for multispectral chromosome classification. *IEEE Trans Med Imaging* 27(5):697–708
83. Ukil S, Reinhardt JM (2009) Anatomy-guided Lung lobe segmentation in X-Ray CT images. *IEEE Trans Med Imaging* 28(2):202–214
84. Alush A, Greenspan H, Goldberger J (2010) Automated and interactive lesion detection and segmentation in uterine cervix images. *IEEE Trans Med Imaging* 29(2):488–501
85. Vincken KL, Koster ASE, Viergever MA (1997) Probabilistic multiscale image segmentation. *IEEE Trans Pattern Anal Mach Intell* 19(2):109–120
86. Snel JG, Venema HW, Grimbergen CA (1998) Detection of the carpal bone contours from 3-D MR images of the wrist using a planar radial scale-space snake. *IEEE Trans Med Imaging* 17(6):1063–1072
87. Gauch JM (1999) Image segmentation and analysis via multiscale gradient watershed hierarchies. *IEEE Trans Image Process* 8(1):69–79
88. Petrovic A, Escoda OD, Vanderghenst P (2004) Multiresolution segmentation of natural images: from linear to nonlinear scale-space representations. *IEEE Trans Image Process* 13(8):1104–1114
89. Li CH, Tam PKS (1998) An iterative algorithm for minimum cross-entropy thresholding. *Pattern Recognit Lett* 19(8):771–776

90. Sezgin M, Sankur B (2004) Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging* 13(1):146–168
91. Wirjad O (2007) Survey of 3D image segmentation methods. Technical report, Fraunhofer ITWM, no. 123
92. Haralick RM, Shapiro LG (1985) Image segmentation techniques. *Comput Vis Graph Image Process* 29(1):100–132
93. Hojjatoleslami SA, Kruggel F (2001) Segmentation of large brain lesions. *IEEE Trans Med Imaging* 20:666–669
94. Wan S-Y, Higgins WE (2003) Symmetric region growing. *IEEE Trans Image Process* 12(9): 1007–1015
95. Mendonca AM, Campilho A (2006) Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Trans Med Imaging* 25: 1200–1213
96. Dehmeshki J, Member HA, Valdovieso M, Ye X (2008) Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach. *IEEE Trans Image Process* 27(4):467–480
97. Zucker SW (1976) Region growing childhood and adolescence. *Comput Graph Image Process* 5(3):382–399
98. Bezdek JC, Hall LO, Clarke LP (1993) Review of MR image segmentation techniques using pattern recognition. *Med Phys* 20(4):1033–1048
99. Duda RO, Hart PE, Stork DG (2001) Pattern classification, 2nd edn. Wiley, New York
100. Bishop CM (2006) Pattern recognition and machine learning. Springer Science and Business Media LLC, Singapore
101. Scott DW (1992) Multivariate density estimation: theory, practice, and visualization. Wiley, New York
102. Hastie T, Tibshirani R, Friedman J (2001) The elements of statistical learning: data mining, inference, and prediction. Springer, Canada
103. McLachlan G, Peel D (2000) Finite mixture models. Wiley, New York
104. Frey BJ, Jojic N (2005) A comparison of algorithms for inference and learning in probabilistic graphical models. *IEEE Trans Pattern Anal Mach Intell* 27(9):1392–1416
105. Farag AA, El-Baz AS, Gimel'farb G (2006) Precise segmentation of multimodal images. *IEEE Trans Image Process* 15(4):952–968
106. El-Baz A, Gimel'farb G (2009) Robust medical images segmentation using learned shape and appearance models. In: Proceedings of the international conference on medical image computing and computer-assisted intervention (MICCAI'09), London, UK, 20–24 September 2009, pp 281–288
107. Lafferty J, McCallum A, Pereira F (2001) Conditional random fields: probabilistic models for segmenting and labeling sequence data. In: Proceedings of the 18th international conference on machine learning (ICML01), Williamstown, MA, 28 June–1 July 2001, pp 282–289
108. Kim J, Fisher JW, Tsai A, Wible C, Willsky A, Wells WM (2000) Incorporating spatial priors into an information theoretic approach for fMRI data analysis. In: Proceedings of the international conference on medical image computing and computer-assisted intervention (MICCAI'2000), Pennsylvania, 11–14 October 2000, pp 62–71.
109. Boykov Y, Jolly M-P (2000) Interactive organ segmentation using graph cuts. In: Proceedings of the international conference on medical image computing and computer-assisted intervention (MICCAI'2000), Pennsylvania, 11–14 October 2000, pp 276–286
110. Kim J, Zabih R (2003) Automatic segmentation of contrast-enhanced image sequences. In: Proceedings of the international conference on computer vision (ICCV'03), Nice, France, 13–16 October 2003, pp 502–509
111. Boykov YY, Jolly M-P (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: Proceedings of the 8th IEEE international conference on computer vision (ICCV'01), Vancouver, BC, Canada, 7–14 July 2001, pp 105–112

112. Freedman D, Drineas P (2005) Energy minimization via graph cuts: settling what is possible. In: Proceedings of IEEE conference on computer vision and pattern recognition (CVPR'05), San Diego, CA, 20–26 June 2005, vol II, pp 939–946
113. Kolmogorov V, Zabih R (2002) Multicamera scene reconstruction via graph cuts. In: Proceedings of the 7th European conference on computer vision (ECCV'02), Denmark, Copenhagen, 27 May–2 June 2002, part III, pp 82–96
114. Kolmogorov V, Zabih R (2004) What energy functions can be minimized via graph cuts? *IEEE Trans Pattern Anal Mach Intell* 26(2):147–159
115. Greig DM, Porteous BT, Seheult AH (1989) Exact maximum a posteriori estimation for binary images. *J R Stat Soc Series B Methodol* 51(2):271–279
116. Roy S, Cox IJ (1998) A maximum-flow formulation of the n-camera stereo correspondence problem. In: Proceedings of the 6th international conference on computer vision (ICCV'98), Bombay, India, 4–7 January 1998, pp 492–499
117. Hartmann SL, Parks MH, Martin PR, Dawant BM (1999) Automatic 3-D segmentation of internal structures of the head in MR images using a combination of similarity and freeform transformations: part II, validation on severely atrophied brains. *IEEE Trans Med Imaging* 18(10):917–926
118. Marroquin JL, Vemuri BC, Botello S, Calderon F, Fernandez-Bouzas A (2002) An accurate and efficient Bayesian method for automatic segmentation of brain MRI. *IEEE Trans Med Imaging* 21(8):934–945
119. Li B, Christensen GE, Hoffman EA, McLennan G, Reinhardt JM (2003) Establishing a normative atlas of the human lung: intersubject warping and registration of volumetric CT images. *Acad Radiol* 10(3):255–265
120. Park H, Bland PH, Meyer CR (2003) Construction of an abdominal probabilistic atlas and its application in segmentation. *IEEE Trans Med Imaging* 22(4):483–492
121. Rohlfing T, Russakoff DB, Maurer CR (2004) Performance-based classifier combination in atlas-based image segmentation using expectation-maximization parameter estimation. *IEEE Trans Med Imaging* 23(8):983–994
122. Cuadra MB, Pollo C, Bardera A, Cuisenaire O, Villemure J-G, Thiran J-P (2004) Atlas-based segmentation of pathological MR brain images using a model of lesion growth. *IEEE Trans Med Imaging* 23(10):1301–1314
123. Sluimer I, Prokop M, van Ginneken B (2005) Toward automated segmentation of the pathological lung in CT. *IEEE Trans Med Imaging* 24(8):1025–1038
124. Stancanello J, Romanelli P, Modugno N, Cerveri P, Ferrigno G, Uggeri F, Cantore G (2006) Atlas-based identification of targets for functional radiosurgery. *Med Phys* 33(6):1603–1611
125. Zhang L, Hoffman EA, Reinhardt JM (2006) Atlas-driven lung lobe segmentation in volumetric X-Ray CT images. *IEEE Trans Med Imaging* 25(1):1–16
126. O'Donnell LJ, Westin C-F (2007) Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Trans Med Imaging* 26(11):1562–1575
127. Han X, Fischl B (2007) Atlas renormalization for improved brain MR image segmentation across scanner platforms. *IEEE Trans Med Imaging* 26(4):479–486
128. Bello M, Ju T, Carson J, Warren J, Chiu W, Kakadiaris IA (2007) Learning-based segmentation framework for tissue images containing gene expression data. *IEEE Trans Med Imaging* 26(5):728–744
129. Leemput KV (2009) Encoding probabilistic brain atlases using Bayesian inference. *IEEE Trans Med Imaging* 28(6):822–837
130. Sabuncu MR, Balci SK, Shenton ME, Golland P (2009) Image-driven population analysis through mixture modeling. *IEEE Trans Med Imaging* 28(9):1473–1487
131. Isgum I, Staring M, Rutten A, Prokop M, Viergever MA, van Ginneken B (2009) Multiatlas-based segmentation with local decision fusion-application to cardiac and aortic segmentation in CT scans. *IEEE Trans Med Imaging* 28(7):1000–1010
132. Artaechevarria X, Munoz-Barrutia A, Ortiz-de-Solorzano C (2009) Combination strategies in multi-atlas image segmentation: application to brain MR data. *IEEE Trans Med Imaging* 28(8):1266–1277

133. van Rikxoort EM, Prokop M, de Hoop B, Viergever MA, Pluim J, van Ginneken B (2010) Automatic segmentation of pulmonary lobes robust against incomplete fissures. *IEEE Trans Med Imaging* 29(6):1286–1296
134. Rohlfing T, Brandt R, Menzel R, Russakoff DB, Maurer CR Jr (2005) Quo vadis, atlas-based segmentation? In: Suri JS, Wilson DL, Laxminarayan S (eds) *Handbook of Biomedical Image Analysis*, vol III: Registration Models. Kluwer Academic/Plenum Publishers, New York, pp 435–486, chapter 11
135. Kittler J, Hatef M, Duin RPW, Matas J (1998) On combining classifiers. *IEEE Trans Pattern Anal Mach Intell* 20(3):226–239
136. Kass M, Witkin A, Terzopoulos D (1987) Snakes: active contour models. *Int J Comput Vis* 1(4):321–331
137. Terzopoulos D (1987) On matching deformable models to images. Technical report no. 60, Schlumberger Palo Alto research, Palo Alto, CA, 1986. Reprinted in Topical meeting on machine vision, technical digest series, vol 12. pp 160–163
138. Terzopoulos D, Fleischer K (1988) Deformable models. *Vis Comput* 4:306–331
139. Terzopoulos D, Witkin A, Kass M (1988) Constraints on deformable models: recovering 3D shape and nonrigid motion. *Artif Intell* 36:91–123
140. Amini AA, Weymouth TE, Jain RC (1990) Using dynamic programming for solving variational problems in vision. *IEEE Trans Pattern Anal Mach Intell* 12(9):855–867
141. Terzopoulos D, Metaxas D (1991) Dynamic 3-D models with local and global deformations: deformable superquadrics. *IEEE Trans Pattern Anal Mach Intell* 13(7):703–714
142. Staib LH, Duncan JS (1992) Boundary finding with parametrically deformable models. *IEEE Trans Pattern Anal Mach Intell* 14(11):1061–1075
143. Staib LH, Duncan JS (1992) Deformable Fourier models for surface finding in 3-D images. *Vis Biomed Comput* 14:90–104
144. Mortensen E, Morse B, Barrett W, Udupa J (1992) Adaptive boundary detection using livewire two dimensional dynamic programming. In: Proceedings of computers in cardiology, Durham, NC, 11–14 October 1992, pp 635–638
145. Williams DJ, Shah M (1992) A fast algorithm for active contour and curvature estimation. *Comput Vis Image Process Understand* 55(1):14–26
146. McInerney T, Terzopoulos D (1995) Topologically adaptable snakes. In: Proceedings of 5th international conference on computer vision (ICCV'95), Cambridge, MA, 20–23 June 1995, pp 840–845
147. McInerney T, Terzopoulos D (1995) A dynamic finite surface model for segmentation and tracking in multidimensional medical images with application to cardiac 4D image analysis. *Comput Med Imaging Graph* 19(1):69–83
148. Davatzikos CA, Prince JL (1995) An active contour model for mapping the cortex. *IEEE Trans Med Imaging* 14(1):65–80
149. Lobregt S, Viergever MA (1995) A discrete dynamic contour model. *IEEE Trans Med Imaging* 14(1):12–24
150. Staib LH, Duncan JS (1996) Model-based deformable surface finding for medical images. *IEEE Trans Med Imaging* 15(5):720–731
151. McInerney T, Terzopoulos D (1996) Deformable models in medical image analysis: a survey. *Med Image Anal* 1(2):91–108
152. Szekely G, Kelemen A, Brechbuhler C, Gerig G (1996) Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Med Image Anal* 1(1):19–34
153. Yezzi A Jr, Kichenassamy S, Kumar A, Olver P, Tenenbaum A (1997) A geometric snake model for segmentation of medical imagery. *IEEE Trans Med Imaging* 16(2):199–209
154. Klein A, Lee F, Amini AA (1997) Quantitative coronary angiography with deformable spline models. *IEEE Trans Med Imaging* 16(5):468–482
155. Grzeszczuk RP, Levin DN (1997) Brownian strings: segmenting images with stochastically deformable contours. *IEEE Trans Pattern Anal Mach Intell* 19(10):1100–1114

156. Cootes TF, Edwards GJ, Taylor CJ (1998) Active appearance models. In: Proceedings of the 5th European conference on computer vision (ECCV'98), Freiburg, Germany, 2–6 June 1998, vol 2, pp 484–498
157. Xu C, Prince JL (1998) Snakes, shapes, and gradient vector flow. *IEEE Trans Image Process* 7(3):359–369
158. Mikic I, Krucinski S, Thomas JD (1998) Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates. *IEEE Trans Med Imaging* 17(2):274–284
159. Atkins MS, Mackiewich BT (1998) Fully automatic segmentation of the brain in MRI. *IEEE Trans Med Imaging* 17(1):98–107
160. Lachaud J-O, Montanvert A (1999) Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction. *Med Image Anal* 3(2):187–207
161. McInerney T, Terzopoulos D (1999) Topology adaptive deformable surfaces for medical image volume segmentation. *IEEE Trans Med Imaging* 18(10):840–850
162. McInerney T, Terzopoulos D (2000) T-snakes: topology adaptive snakes. *Med Image Anal* 4(2):73–91
163. Falcao AX, Udupa JK (2000) A 3D generalization of user-steered live-wire segmentation. *Med Image Anal* 4(4):389–402
164. Falcao AX, Udupa JK, Miyazawa FK (2000) An ultra-fast user-steered image segmentation paradigm: live wire on the fly. *IEEE Trans Med Imaging* 19(1):55–62
165. Mokhtarian F, Mohanna F (2001) Fast active contour convergence through curvature scale space filtering. In: Proceedings of image and vision computing New Zealand, Dunedin, New Zealand, 2001, pp 157–162
166. Fenster SD, Kender JR (2001) Sectored snakes: evaluating learned-energy segmentations. *IEEE Trans Pattern Recognit Mach Intell* 23(9):1028–1034
167. Ghanei A, Soltanian-Zadeh H (2002) A discrete curvature-based deformable surface model with application to segmentation of volumetric images. *IEEE Trans Inf Technol Biomed* 9(4):285–295
168. Snel JG, Venema HW, Grimborgen CA (2002) Deformable triangular surfaces using fast 1-D radial Lagrangian dynamics-segmentation of 3-D MR and CT images of the wrist. *IEEE Trans Med Imaging* 21(8):888–903
169. Goldenberg R, Kimmel R, Rivlin E, Rudzsky M (2002) Cortex segmentation: a fast variational geometric approach. *IEEE Trans Med Imaging* 21(12):1544–1551
170. McInerney T, Hamarneh G, Shenton M, Terzopoulos D (2002) Deformable organisms for automatic medical image analysis. *Med Image Anal* 6(3):251–266
171. Serment M, Forest C, Pennec X, Delingette H, Ayache N (2003) Deformable biomechanical models: application to 4D cardiac image analysis. *Med Image Anal* 7(4):457–488
172. Madabhushi A, Metaxas DN (2003) Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. *IEEE Trans Med Imaging* 22(2):155–169
173. Brusseau E, de Korte CL, Mastik F, Schaar J, van der Steen AFW (2004) Fully automatic luminal contour segmentation in intracoronary ultrasound imaging-A statistical approach. *IEEE Trans Med Imaging* 23(5):554–566
174. Yao J, Miller M, Franaszek M, Summers RM (2004) Colonic polyp segmentation in CT colonography-based on fuzzy clustering and deformable models. *IEEE Trans Med Imaging* 23(11):1344–1352
175. Lowell J, Hunter A, Steel D, Basu A, Ryder R, Fletcher E, Kennedy L (2004) Optic nerve head segmentation. *IEEE Trans Med Imaging* 23(2):256–264
176. Ghebreab S, Smeulders AWM (2004) Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. *IEEE Trans Med Imaging* 23(10):1821–1829
177. Pang B, Zhang D, Wang K (2005) The bi-elliptical deformable contour and its application to automated tongue segmentation in Chinese medicine. *IEEE Trans Med Imaging* 24(8):946–956

178. Kaus MR, von Berg J, Weese J, Niessen W, Pekar V (2004) Automated segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 8(3):245–254
179. de Bruijne M, van Ginneken B, Viergever MA, Niessen WJ (2004) Interactive segmentation of abdominal aortic aneurysms in CTA images. *Med Image Anal* 8:127–138
180. Montagnat J, Delingette H (2005) 4D deformable models with temporal constraints: application to 4D cardiac image segmentation. *Med Image Anal* 9:87–100
181. Chen T, Metaxas D (2005) A hybrid framework for 3D medical image segmentation. *Med Image Anal* 9(6):547–565
182. Chrastek R, Wolf M, Donath K, Niemann H, Paulus D, Hothorn T, Lausen B, Lammer R, Mardin CY, Michelson G (2005) Automated segmentation of the optic nerve head for diagnosis of glaucoma. *Med Image Anal* 9(4):297–314
183. Olabarriaga SD, Rouet J-M, Fradkin M, Breeuwer M, Niessen WJ (2005) Segmentation of thrombus in abdominal aortic aneurysms from CTA with nonparametric statistical grey level appearance modeling. *IEEE Trans Med Imaging* 24(4):477–485
184. Zagrodsky V, Walimbe V, Castro-Pareja CR, Qin JX, Song J-M, Shekhar R (2005) Registration-assisted segmentation of real-time 3-D echocardiographic data using deformable models. *IEEE Trans Med Imaging* 24(9):1089–1099
185. Freedman D, Radke RJ, Zhang T, Jeong Y, Lovelock DM, Chen GTY (2005) Model-based segmentation of medical imagery by matching distributions. *IEEE Trans Med Imaging* 24(3):281–292
186. Dufour A, Shinin V, Tajbakhsh S, Olivo-Marin J-C, Guillen-Aghion N, Zimmer C (2005) Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. *IEEE Trans Med Imaging* 14(9):1396–1410
187. Zhan Y, Shen D (2006) Deformable segmentation of 3-D ultrasound prostate images using statistical texture matching method. *IEEE Trans Med Imaging* 25(3):256–272
188. Xu M, Thompson PM, Toga AW (2006) Adaptive reproducing kernel particle method for extraction of the cortical surface. *IEEE Trans Med Imaging* 25(6):755–767
189. van Assen HC, Danilouchkine MG, Frangi AF, Ordas S, Westenberg JJM, Reiber JHC, Lelieveldt BPF (2006) SPASM: a 3D-ASM for segmentation of sparse and arbitrarily oriented cardiac MRI data. *Med Image Anal* 10:286–303
190. Li B, Acton ST (2007) Active contour external force using vector field convolution for image segmentation. *IEEE Trans Image Process* 16(8):2096–2106
191. Huang X, Metaxas DN (2008) Metamorphs: deformable shape and appearance models. *IEEE Trans Pattern Anal Mach Intell* 30(8):1444–1459
192. Bernard O, Friboulet D, Thevenaz P, Unser M (2009) Variational B-spline level-set: a linear filtering approach for fast deformable model evolution. *IEEE Trans Image Process* 18(6): 1179–1191
193. Fripp J, Crozier S, Warfield SK, Ourselin S (2010) Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. *IEEE Trans Med Imaging* 29(1):55–64
194. Manzke R, Meyer C, Ecabert O, Peters J, Noordhoek NJ, Thiagalingam A, Reddy VY, Chan RC, Weese J (2010) Automatic segmentation of rotational X-Ray images for anatomic intra-procedural surface generation in atrial fibrillation ablation procedures. *IEEE Trans Med Imaging* 29(2):260–272
195. Osher S, Sethian JA (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation. *J Comput Phys* 79:12–49
196. N Paragios, R Deriche (1999) Unifying boundary and region-based information for geodesic active tracking. In: Proceedings of the IEEE international conference in computer vision and pattern recognition (CVPR'99), Collins, CO, 23–25 June 1999, vol 2, pp 300–305
197. Ma W-Y, Manjunath BS (2000) Edge flow: a technique for boundary detection and image segmentation. *IEEE Trans Image Process* 9(8):1375–1388
198. Shen D, Davatzikos C (2000) An adaptive-focus deformable model using statistical and geometric information. *IEEE Trans Pattern Anal Mach Intell* 22(8):906–913

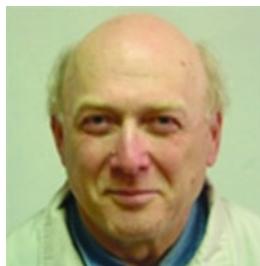
199. Tsai A, Yezzi A Jr, Willsky A (2001) Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Trans Image Process* 10(8):1169–1186
200. Chan TF, Vese LA (2001) Active contours without edges. *IEEE Trans Image Process* 10(2):266–277
201. Baillard C, Hellier P, Barillot C (2001) Segmentation of brain 3D MR images using level sets and dense registration. *Med Image Anal* 5(3):185–194
202. Lorigo LM, Faugeras OD, Grimson WEL, Keriven R, Kikinis R, Nabavi A, Westin C-F (2001) CURVES: curve evolution for vessel segmentation. *Med Image Anal* 5:195–206
203. Carson C, Belongie S, Greenspan H, Malik J (2002) Blobworld: image segmentation using expectation-maximization and its application to image querying. *IEEE Trans Pattern Anal Mach Intell* 24(8):1026–1038
204. Chen Y, Tagare HD, Thiruvenkadam S, Huang F, Wilson D, Gopinath KS, Briggs RW, Geiser EA (2002) Using prior shapes in geometric active contours in a variational framework. *Int J Comput Vis* 50(3):315–328
205. Sebastian TB, Tek H, Crisco JJ, Kimia BB (2003) Segmentation of carpal bones from CT images using skeletally coupled deformable models. *Med Image Anal* 7(1):21–45
206. Vemuri BC, Ye J, Chen Y, Leonard CM (2003) Image registration via level-set motion: applications to atlas-based segmentation. *Med Image Anal* 7(1):1–20
207. Yezzi A, Zollei L, Kapur T (2003) A variational framework for integrating segmentation and registration through active contours. *Med Image Anal* 7(2):171–185
208. Cates JE, Lefohn AE, Whitaker RT (2004) GIST: an interactive, GPU-based level set segmentation tool for 3D medical images. *Med Image Anal* 8(3):217–231
209. Kim J, Fisher JW III, Yezzi A, Cetin M, Willsky AS (2005) A nonparametric statistical method for image segmentation using information theory and curve evolution. *IEEE Trans Med Imaging* 14(10):1486–1502
210. Tsai A, Wells WM, Warfield SK, Willsky AS (2005) An EM algorithm for shape classification based on level sets. *Med Image Anal* 9(5):491–502
211. Yan P, Kassim AA (2006) Segmentation of volumetric MRA images by using capillary active contour. *Med Image Anal* 10:317–329
212. Manniesing R, Velthuis BK, van Leeuwen MS, van der Schaaf IC, van Laar PJ, Niessen WJ (2006) Level set based cerebral vasculature segmentation and diameter quantification in CT angiography. *Med Image Anal* 10(2):200–214
213. Holtzman-Gazit M, Kimmel R, Peled N, Goldsher D (2006) Segmentation of thin structures in volumetric medical images. *IEEE Trans Image Process* 15(2):354–363
214. Osher S, Fedkiw R (2006) Level set methods and dynamic implicit surfaces, applied mathematical sciences, vol 153. Springer, New York
215. Chang H-H, Valentino DJ, Duckwiler GR, Toga AW (2007) Segmentation of brain MR images using a charged fluid model. *IEEE Trans Biomed Eng* 54(10):1798–1813
216. Abd El Munim HE, Farag AA (2007) Curve/surface representation and evolution using vector level sets with application to the shape-based segmentation problem. *IEEE Trans Pattern Anal Mach Intell* 29(6):945–958
217. Gelas A, Bernard O, Friboulet D, Prost R (2007) Compactly supported radial basis functions based collocation method for level-set evolution in image segmentation. *IEEE Trans Image Process* 16(7):1873–1887
218. Sun W, Cetin M, Chan R, Willsky AS (2008) Learning the dynamics and time-recursive boundary detection of deformable objects. *IEEE Trans Image Process* 17(11):2186–2200
219. Law YN, Lee HK, Yip AM (2008) A multiresolution stochastic level set method for Mumford-Shah image segmentation. *IEEE Trans Image Process* 17(12):2289–2300
220. Cheng K, Gu L, Wu J, Li W, Xu J (2008) A novel level set based shape prior method for liver segmentation from MRI images. In: Proceedings of the 4th international workshop on medical imaging and augmented reality, Tokyo, Japan, 2008, pp 150–159

221. Gooya A, Liao H, Matsumiya K, Masamune K, Masutani Y, Dohi T (2008) A variational method for geometric regularization of vascular segmentation in medical images. *IEEE Trans Image Process* 17(8):1295–1312
222. Krinidis S, Chatzis V (2009) Fuzzy energy-based active contours. *IEEE Trans Image Process* 18(12):2747–2755
223. Huang A, Abugharbieh R, Tam R (2009) A hybrid geometric-statistical deformable model for automated 3-D segmentation in brain MRI. *IEEE Trans Biomed Eng* 56(7):1838–1848
224. Wu X, Luboz V, Krissian K, Cotin S, Dawson S (2011) Segmentation and reconstruction of vascular structures for 3D real time simulation. *Med Image Anal* 15(1):22–34
225. Vukadinovic D, van Walsum T, Manniesing R, Rozie S, Hameeteman R, de Weert TT, van der Lugt A, Niessen WJ (2010) Segmentation of the outer vessel wall of the common carotid artery in CTA. *IEEE Trans Med Imaging* 29(1):65–76
226. Duan C, Liang Z, Bao S, Zhu H, Wang S, Zhang G, Chen JJ, Lu H (2010) A coupled level set framework for bladder wall segmentation with application to MR cystography. *IEEE Trans Med Imaging* 29(3):903–915
227. Khalifa F, El-Baz A, Gimelfarb G, Ousephand R, Abu El-Ghar M (2010) Shape-appearance guided level-set deformable model for image segmentation. In: IEEE international conference on pattern recognition (ICPR'2010), Istanbul, Turkey, August 2010 [in print]
228. Xu C, Pham DL, Prince JL (2000) Image segmentation using deformable models. In: Fitzpatrick JM, Sonka M (eds) *Handbook on medical imaging*. SPIE Press, Bellingham, WA, pp 129–174, Chapter 3
229. Jaklic A, Leonardis A, Solina F (2000) Segmentation and recovery of superquadrics. Kluwer Academic Publisher, Dordrecht, Netherland
230. Persoon E, Fu KS (1977) Shape discrimination using Fourier descriptors. *IEEE Trans Syst Man Cybern* 7(3):170–179
231. Zienkiewicz OC, Taylor RL, Zhu JZ (2005) *The finite element method: its basis and fundamentals*, 6th edn. Elsevier, Oxford, UK
232. Boykov Y, Kolmogorov V (2003) Computing geodesics and minimal surfaces via graph cuts. In: Proceedings of 9th IEEE international conference on computer vision (ICCV'03), Nice, France, 13–16 October 2003, vol 1, p 2633
233. Suri JS, Liu K, Singh S, Laxinarian SN, Zeng X, Reden L (2002) Shape recovery algorithms using level sets in 2-D/3-D medical imagery: a state-of-the-art review. *IEEE Trans Inf Technol Biomed* 6(1):8–28

## Biography



**Ahmed Elnakib** received his B.Sc. and M.Sc. degrees in Electronics and Communication Engineering from Mansoura University, Egypt, in 2003 and 2007, respectively. Mr. Elnakib joined the BioImaging Laboratory as research assistance in September 2008. His current research includes developing new computer assisted diagnosis systems, image modeling, image segmentation, 2-D and 3-D registration, and shape analysis of detected lung nodules.



**Georgy Gimel'farb** graduated from Kiev Polytechnic Institute, Kiev, Ukraine, and received the Ph.D. degree in engineering cybernetics from the Institute of Cybernetics, Academy of Sciences of the Ukraine, and the D.Sc. (Eng) degree in control in engineering from the Higher Certifying Commission of the USSR, Moscow, Russia. He joined the University of Auckland, Auckland, New Zealand, in July 1997. His research is focused on image analysis, computer vision, and statistical pattern recognition.



**Jasjit S. Suri** is an innovator, scientist, a visionary, an industrialist and an internationally known world leader in Biomedical Engineering. Dr. Suri has spent over

20 years in the field of biomedical engineering/devices and its management. He received his Doctorate from University of Washington, Seattle; and Business Management Sciences from Weatherhead, Case Western Reserve University, Cleveland, Ohio. Dr. Suri was crowned with the President's Gold medal in 1980 and the Fellow of American Institute of Medical and Biological Engineering for his outstanding contributions.



**Ayman El-Baz** received B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering department, University of Louisville, in August 2006. His current research is focused on developing new computer assisted diagnosis systems for different diseases and brain disorders.

# **Chapter 2**

## **Cerebral White Matter Segmentation using Probabilistic Graph Cut Algorithm**

**Mainak Sen, Ashish K. Rudra, Ananda S. Chowdhury, Ahmed Elnakib, and Ayman El-Baz**

**Abstract** Accurate and efficient computer-assisted brain image segmentation methods are of great interest to both scientific and clinical researchers of the human central neural system. Cerebral white matter segmentation from Magnetic Resonance Imaging (MRI) data of brain remains a challenging problem due to a combination of several factors: noise and imaging artifacts, partial volume effects, intrinsic tissue variation due to neurodevelopment and neuropathologies, and the highly convoluted geometry of the cortex. We propose here a probabilistic variation of the traditional graph cut algorithm (IEEE international conference on computer vision, pp 105–112) with an improved parameter selection mechanism for the energy function, to be optimized in a graph cut problem. In addition, we use a simple yet effective shape prior in form of a series of ellipses to increase the computational efficiency of the proposed algorithm and improve the quality of the segmentation by modeling the contours of the human skull in various 2D slices of the sequence. Qualitative as well as quantitative segmentation results on T1-weighted MRI input, for both 2D and 3D cases are included. These results indicate that the proposed probabilistic graph cut algorithm outperforms some of the state-of-the art segmentation algorithms like the traditional graph cut (IEEE international conference on computer vision, pp 105–112) and the expectation maximization segmentation (IEEE Trans Med Imaging 20(8):677–688, 2001).

### **2.1 Introduction**

Cerebral white matter segmentation has always been a challenging task for a computer assisted diagnosis system. The problem is of interest for researchers in computer vision and biomedical imaging as well as for neurologists and neurosurgeons. The segmentation problem becomes very difficult due to a combination of several factors

---

A.S. Chowdhury (✉)

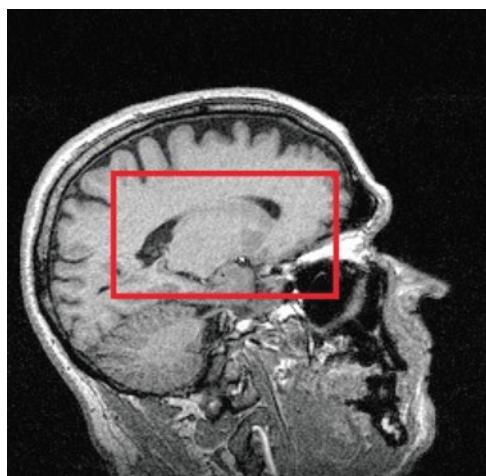
Department of Electronics and Telecommunication Engineering, Jadavpur University, Calcutta 700032, India

e-mail: ananda.chowdhury@gmail.com

like (a) noise and imaging artifacts, (b) partial volume effects, (c) intrinsic tissue variation due to neurodevelopment and neuropathologies, and (d) the highly convoluted geometry of the cortex. So, this problem poses great challenge to the computer vision community. In addition to the technical challenges, the problem has a lot of clinical significance, e.g., it helps us in the understanding of brain disorders like autism [1]. Furthermore, study of abnormalities and lesions in the white matter is extremely helpful for the study of multiple sclerosis, Alzheimer's disease [2], and cognitive defects in elderly subjects [3].

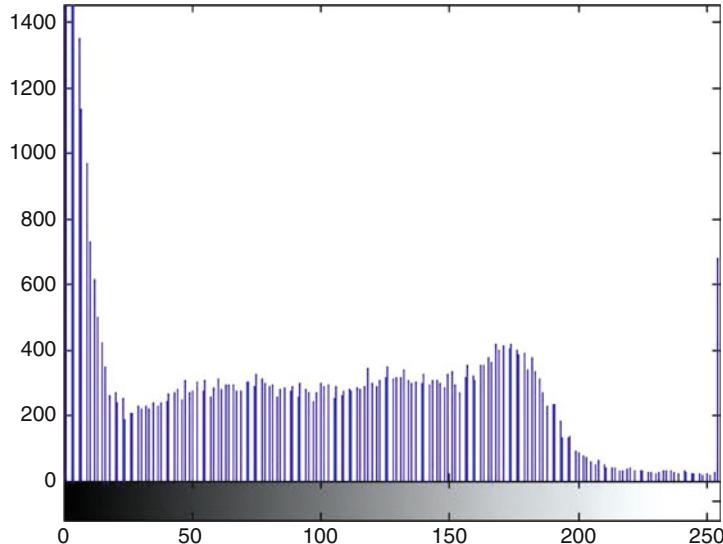
Graph cuts are applied over the last few years to accurately solve complex segmentation problems [4–8]. The main advantage of graph cuts is that they yield a globally optimal segmentation which can potentially overcome the problems of oversegmentation and undersegmentation. However, the success of the graph cut algorithm depends on the variation in the capacities of different edges in the graph, which in turn depends on existing contrast in the image. T1-weighted input MRI images for our problem exhibit significant low contrast and are hence are not particularly amenable to graph cut as prescribed by Boykov and Jolly [4].

A 2D slice from a T1-weighted MRI input is shown in Fig. 2.1. Note that there is pronounced low contrast between our target object, i.e., the cerebral white matter and the surrounding brain structures. One such region is indicated by a red rectangle. For better illustration, this region is zoomed in Fig. 2.2. To understand the distribution of the intensity levels (among the pixels), we present a histogram of the above region in Fig. 2.3. The histogram does not show significant variations among the intensity levels. Therefore, it becomes evident that the cerebral white matter and the neighboring structures have quite similar intensity profiles and do not exhibit high contrast. Thus it becomes necessary to modify the graph cut algorithm in [4] to have accurate segmentation results. In this work, we propose a probabilistic version of



**Fig. 2.1** A T1-weighted MRI input data with the rectangle (in red) indicates a typical low-contrast region containing a part of cerebral white matter

**Fig. 2.2** A zoomed view of the selected area in Fig. 2.1



**Fig. 2.3** Histogram of the selected area in Fig. 2.1

the graph cut algorithm, described in [4] by modifying the edge capacity functions according to the probabilities of voxels to belong to different segmentation classes.

The Expectation Maximization Segmentation (EMS) algorithm proposed by Leemput et al. [9] is one important statistical method, which has been extensively used as a tissue classifier for brain MRI. In order to show the strength of our method, we have compared our 3D segmentation results with [4, 9]. An earlier version of this work, where we focused on 2D segmentation and restricted our comparison with [4], appears in [10].

The rest of the chapter is organized in the following manner: in Sect. 2.2, we present the literature review and highlight our contributions. In Sect. 2.3, we describe various components of our proposed technique, namely, the probabilistic graph cut algorithm, the variation of the weighting parameter in the energy function of the graph cut and the imposition of the shape prior. In Sect. 2.4, we discuss the experimental results for both 2D and 3D segmentation. Finally, in Sect. 2.5, we summarize and conclude the work.

## 2.2 Literature Review

Popular computer vision techniques for extracting a brain structure from the MRI data can be grouped into three broad categories, namely, (a) *statistical models*-based segmentation, (b) *atlas*-based segmentation, and (c) *deformable models*-based segmentation. The statistical approaches [9, 11–13] suffer from the fact that the tails of the distributions of the brain tissues (e.g., white matter, gray matter), which are modeled as a mixture of well-known distributions, are always overlapped. Thus, it becomes quite difficult to find thresholds that separate them accurately. Success of atlas-based methods [11, 12, 14] depends heavily on the high accuracy of registration between the atlas brain and the unsegmented brain, which may be hard to obtain in many cases. The main drawbacks of the deformable models [15–17] include slow speed, strong sensitivity to initialization, and incomplete segmentation in case of complicated object boundary with concavities. Graph cuts [4–8] have been applied to medical image segmentation problems including structure detections in the human brain [18].

To overcome the limitations of the various prevalent methods, we propose a probabilistic graph cut algorithm with an improved parameter selection strategy in conjunction with a simple shape prior to segment the cerebral white matter. The probabilistic graph cut algorithm effectively handles low contrast between the cerebral white matter and the surrounding anatomical structures in the MRI data. Note that the choice of the weighting parameter for the data and the smoothness terms in the energy function of the graph cut algorithm affects the segmentation performance [19]. An improved parameter selection procedure for each individual voxel is, based on the probabilities of its surrounding voxels to belong to the object class and the background class, is suggested. We further employ a simple yet effective shape prior in form of a series of ellipses to model the human skull in individual 2D slices constituting the input sequence. In sharp contrast to the methods described in [20–22], where the shape constraints are added in form of additional terms in the energy function, we use the shape constraint to generate a subgraph by pruning a significant part of the 3D graph (constructed from the input). The proposed algorithm is applied on the newly formed subgraph. So, three major contributions are (a) proposition of a probabilistic graph cut algorithm, (b) introduction of an improved parameter selection strategy for the energy function of the probabilistic graph cut algorithm and (c) the manner in which a geometric shape prior is imposed to improve computational efficiency as well as accuracy of segmentation.

## 2.3 Methods

Since our probabilistic graph cut method is developed on the framework of the graph cut algorithm proposed by Boykov and Jolly [4], some of its salient features are included. We will henceforth refer this algorithm as the traditional graph cut. In traditional graph cut, an user marks some pixels as being part of the object of

interest, and some pixels as lying outside the object, *i.e.*, within the background. The segmentation is governed by the following criteria:

1. Each pixel inside the object is given a value according to whether its intensity matches the object's appearance model; low values represent better matches.
2. Each pixel in the background is given a value according to whether its intensity matches the background's appearance model; low values represent better matches.
3. A pair of adjacent pixels, where one is inside the object and the other is outside, is given a value according to whether the two pixels have similar intensities; low values correspond to contrasting intensities (*i.e.*, to an edge).

Given these criteria, the goal is to design an algorithm that can find an optimal segmentation. The edge weights in this graph cut can be formulated such that a binary partitioning derived from the solution of the min-cut/max-flow algorithm minimizes certain energy function. As mentioned earlier, success of the graph cut algorithm depends on correct identification of inexpensive edges. Thus, coining of a proper energy function to capture the variations in the capacities of the edges becomes the most essential step. In the present problem, the existing intensity distribution is not particularly amenable to a correct minimum cut, which thus leads to an improper segmentation. Therefore, some probabilistic adaptation of the standard graph cut is proposed.

We first discuss the construction of the graph from the input sequence. The input MRI sequence is modeled as a weighted undirected 3D graph  $G = G(V, E)$ . Let  $X$  denote the set of all voxels. Each voxel  $x \in X$  is a node/vertex in  $G$ . In addition, two special terminal nodes, called the “source”  $s$  and the “sink”  $t$ , are considered [23]. There are two types of edges/links in the present network, namely, the  $t$ -links ( $T$ ) and the  $n$ -links ( $N$ ). The  $t$ -links connect all the individual voxels  $x$  to the source node  $s$  and the sink node  $t$ . A neighborhood set  $\text{Ne}(x)$  is assumed for each node  $x$ . Here, we have used the 26-neighborhood, *i.e.*,  $|\text{Ne}(x)| = 26$ . The  $n$ -links are constructed between each voxel  $x$  and its neighboring voxels  $y$ , where,  $y \in \text{Ne}(x)$ . Therefore, we have  $V = X \cup s \cup t$  and  $E = T \cup N$ . Note that for 2D segmentation from a single slice, a weighted undirected 2D graph is constructed. An eight-neighborhood is used in such case.

### 2.3.1 Probabilistic Graph Cut

Here, we describe our probabilistic graph cut algorithm. Let  $A$  define a segmentation, *i.e.*, classification of all voxels into either “object” or “background”. Then, following [4], we can write the energy function to be minimized via graph cut over the binary variable  $\delta(A)$  (where  $\delta(A) = 1$  when  $A = \text{object}$  and  $\delta(A) = 0$  when  $A = \text{background}$ ) as:

$$E(A) = B(A) + \lambda R(A) \quad (2.1)$$

where  $B(A)$  and  $R(A)$  respectively denote the smoothness term/boundary properties of segmentation A and the data term/region properties of segmentation A. Mathematical expressions for these two terms are given below:

$$B(A) = \sum_{\substack{x \in X \\ y \in \text{Ne}(x)}} B_{\{x,y\}} \quad (2.2)$$

$$R(A) = \sum_{x \in X} R_x(A_x) \quad (2.3)$$

Let  $O$ ,  $B$ , and  $I_x$  be the histogram of the object seeds, histogram of the background seeds, and intensity of any voxel  $x$  respectively. Further, let  $\Pr(I_x|O)$  and  $\Pr(I_x|B)$  denote the probability of a certain voxel  $x$  to belong to the class “object” (*i.e.*, white matter) and to the class “background” (*i.e.*, everything else in the image) respectively. We next define  $B_{\{x,y\}}$ ,  $R_x(\text{object})$ , and  $R_x(\text{background})$  in the following equations:

$$B_{\{x,y\}} = K_{\{x,y\}} \exp\left(\frac{-(I_x - I_y)^2}{2\sigma^2}\right) * \frac{1}{d(x,y)} \quad (2.4)$$

$$R_x(\text{object}) = -K_x^o \ln(\Pr(I_x|O)) \quad (2.5)$$

$$R_x(\text{background}) = -K_x^b \ln(\Pr(I_x|B)) \quad (2.6)$$

The term  $d(x, y)$  denotes distance between two voxels  $x$  and  $y$ . The three terms  $K_{\{x,y\}}$ ,  $K_x^o$ ,  $K_x^b$  appearing respectively in (2.4)–(2.6) are evaluated within the probabilistic graph cut algorithm.

### Algorithm: Probabilistic Graph Cut

*Step 1: Evaluation of  $K_{\{x,y\}}$*  – Compare the probabilities of any two voxels  $x \in X$  and  $y \in \text{Ne}(x)$  to belong to the same segmentation class using the terms  $\Pr(I_x|O)$ ,  $\Pr(I_x|B)$ ,  $\Pr(I_y|O)$ , and  $\Pr(I_y|B)$ . If both  $x$  and  $y$  have higher probabilities to belong to the same class, assign  $K_1$  to  $K_{\{x,y\}}$ . Otherwise, assign  $K_2$  to  $K_{\{x,y\}}$ .

*Step 2: Evaluation of  $K_x^o$  and  $K_x^b$*  – Compare the probability of a single voxel  $p$  to belong to the segmentation class “object” or “background” using the terms  $\Pr(I_x|O)$  and  $\Pr(I_x|B)$ . If  $x$  has a higher probability to belong to the “object” class, assign  $K_1$  to  $K_x^o$  and  $K_2$  to  $K_x^b$ . Otherwise, assign  $K_2$  to  $K_x^o$  and  $K_1$  to  $K_x^b$ .

*Step 3: Apply probabilistic graph cut using (2.2)–(2.6).*

In a standard graph cut algorithm, we do not consider any variation of the above three terms. Here, we define all of them within a probabilistic framework. Note that the nature of the energy function, as defined in (2.1) still remains convex, in spite of the probabilistic modifications. Thus, we can clearly apply the max-flow min-cut algorithm to find a global minimum of this energy function, which corresponds to the optimal segmentation [24]. In the next subsection, we illustrate the comparative

performances of the proposed probabilistic graph cut algorithm and the traditional graph cut algorithm on a synthetic image. The value of  $\lambda$  in (2.1) is modified according to the technique described in the Sect. 2.3.3. There exists a lower limit of the ratio ( $K_1/K_2$ ), which depends on the contrast of the image. For all values of this ratio above the lower limit, the accuracy of the segmented output remains almost constant.

### 2.3.2 *Probabilistic Graph Cut vs. Traditional Graph Cut*

In this section, we show how the proposed probabilistic graph cut algorithm achieves a better segmentation with respect to the traditional graph cut with the help of a small synthetic image containing only four pixels.

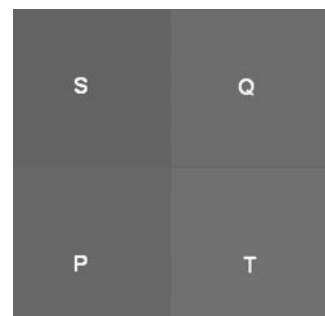
Figure 2.4 shows a low-contrast image with just four pixels, (S, Q, P, and T) having intensities  $I_p$ ,  $I_q$ ,  $I_s$ , and  $I_t$  respectively. Let us assume, without any loss of generality, that the right half of the image (*i.e.*, the pixels Q and T) is our region of interest (ROI). In other words, our objective is to segment these 2 pixels from the set of 4 pixels. The contrast difference between the ROI pixel set (Q, T) and foreground pixel set (S, P) is extremely low.

#### 2.3.2.1 *Segmentation using Traditional Graph Cut*

We form the graph (G) from the synthetic image using the capacity functions of the traditional graph cut [4]. The four pixels (S, P, Q, and T) are the nodes of the graph constructed from the image. To find the minimum cut [23], we consider node S and node T as the source and the sink respectively without any loss of generality. As the intensity difference between the pixel P and Q is very low, the  $N$ -link capacity formed between them will be extremely high. Let us introduce the following symbols:

$C_1$  = capacity of the T-link formed between source(S) and pixel P

$C_3$  = capacity of the T-link formed between source(S) and pixel Q



**Fig. 2.4** A synthetic low-contrast image with four pixels S, Q, P, and T

$C1 - X$  = capacity of the T-link formed between sink (T) and pixel P (where  $X > 0$ )

$C3 + Y$  = capacity of the T-link formed between sink (T) and pixel Q (where  $Y > 0$ )

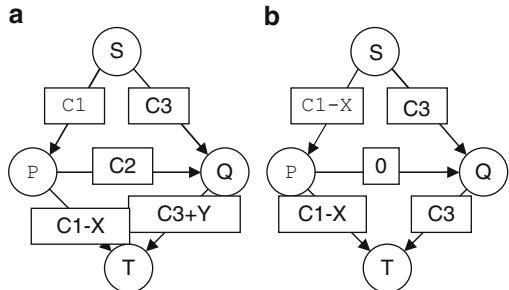
$C2$  = capacity of the N-link formed between pixel P and pixel Q

In Fig. 2.5a,  $C2$  (the capacity of the edge between P and Q) is very high and the value of  $X$  and  $Y$  is quite low due to the low contrast of the image. Now, when we calculate the maximum flow, the flow through the augmenting path 1 ( $S \rightarrow P \rightarrow T$ ) and path 2 ( $S \rightarrow Q \rightarrow T$ ) will be as shown in Fig. 2.5b. Now, in the case of augmenting path 3 ( $S \rightarrow P \rightarrow Q \rightarrow T$ ), the excess capacities in the edges (top to bottom respectively) are  $X$ ,  $C2$ , and  $Y$ . As,  $C2$  is greater than  $X$  and  $Y$ , we have only three cases.

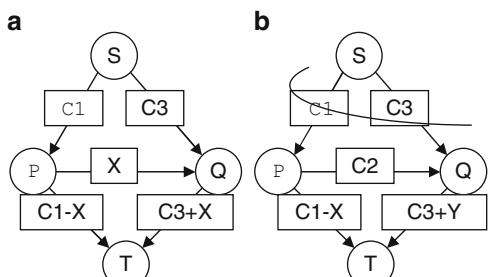
**Case I.**  $X < Y$ : In this case, we can send only  $X$  units of flow along the augmenting path 3. The updated residual graph is shown in Fig. 2.6a.

Here, the maximum amount flow that can be augmented from source(S) to sink (T) is  $(C1 + C3)$ . Figure 2.6b shows the minimum cut, which is equal to the maximum flow in the graph. We can see that it is a case of oversegmentation.

**Case II.**  $X > Y$ : Under this condition, the maximum amount of flow we can send down this augmenting path is  $Y$ . Here, total flow is  $(C1 + C3 - X + Y)$ . The modified flow is shown in Fig. 2.7a and the solid curved line in Fig. 2.7b shows

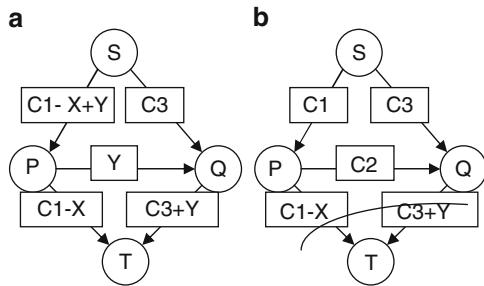


**Fig. 2.5** (a) Graph constructed from the low-contrast image, (b) Flow augmented along the paths in the graph



**Fig. 2.6** (a) Updated flows along the edges after augmenting flow along the path 3 for  $X < Y$ ; (b) The solid curved line shows the minimum cut in the graph

**Fig. 2.7** (a) Updated flows along the edges after augmenting flow along the path 3 for  $X > Y$ ; (b) The solid curved line shows the minimum cut in the graph



the minimum cut. In this case, we get an undersegmented output by applying the traditional graph cut algorithm.

**Case III.  $X = Y$ :** In this case we will get any of the two segmented outputs, obtained in the previous two cases.

So, this flow augmentation based analysis using the maxflow-mincut algorithm [23] shows that for a very low-contrast image, the traditional graph cut segmentation can not produce accurate results.

### 2.3.2.2 Segmentation Using Probabilistic Graph Cut

Now, we again form the graph ( $G$ ) from the Fig. 2.4 incorporating our probabilistic modification in the edge capacities of the traditional graph cut algorithm. For simplicity, we take the probabilistic scaling functions as mentioned earlier, *i.e.*,  $K_1 = K$  and  $K_2 = 1$ . Alternatively, we can say,  $K$  is the ratio of the two scaling functions  $K_1$  and  $K_2$ .

Figure 2.3a shows the residual graph constructed from the low-contrast image. To calculate the maximum flow in  $G$ , we consider only two augmenting paths: path 1 which is  $S \rightarrow P \rightarrow T$  and path 2 which is  $S \rightarrow Q \rightarrow T$ . These two paths are shown in Fig. 2.8b. Now, in the case of path 3 which is  $S \rightarrow P \rightarrow Q \rightarrow T$ , the excess capacities of the three edges are  $(K - 1)*C1 + X$ ,  $C2$  and  $(K - 1)*C3 + K*Y$  from source to sink respectively. Now, if we chose the value of  $K$  such that it satisfies the two following conditions:

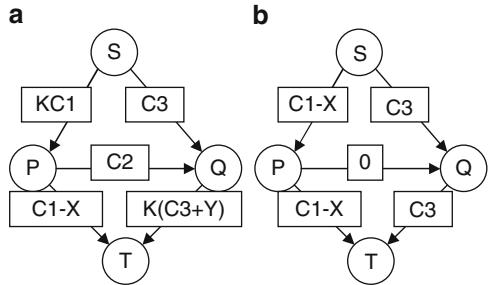
$$(K - 1)C1 + X \geq C2 \quad (2.7a)$$

$$(K - 1)C3 + KY \geq C2 \quad (2.7b)$$

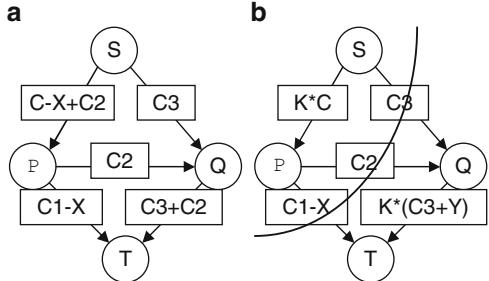
then, the maximum flow through the augmenting path 3 will be  $C2$ . The updated flows in the graph are shown in Fig. 2.9a.

Here, the total flow is  $C1 - X + C2 + C3$  and the mincut is shown in the Fig. 2.9b. In this case, we see that the segmented output is the pixel set  $(Q, T)$

**Fig. 2.8** (a) Graph (G); (b) Flows along the edges after augmenting flow along path 1 and path 2



**Fig. 2.9** (a) Augmented flow along the path 3 and the updated flows in the graph; (b) the solid curved line showing the minimum cut in the graph



which is our desired region of interest (ROI). So, we can conclude that probabilistic graph cut produces more accurate segmentation than the traditional graph cut when the input image is of low contrast.

In the previous section, we have established the criteria (see (2.7a) and (2.7b)) for  $K$ , which need to be satisfied to get the desired segmentation results. We now derive a lower limit of  $K$ , say  $K_L$ , by rewriting the eqs.(2.7a) and (2.7b). Thus, we get:

$$K \geq \frac{(C_2 + (C_1 - X))}{C_1} \quad (2.7c)$$

$$K \geq \frac{C_2 + C_3}{C_3 + Y} \quad (2.7d)$$

From (2.7c) and (2.7d), we can write  $K_L = \max((C_2 + (C_1 - X))/C_1, (C_2 + C_3)/(C_3 + Y))$ . So, the value of  $K_L$  depends upon the edge capacities. Note that the capacities  $C_1, (C_1 - X)$  and  $C_3, (C_3 + Y)$  depend on  $\Pr(I_p/O)$ ,  $\Pr(I_p/B)$ ,  $\Pr(I_q/O)$ , and  $\Pr(I_q/B)$ , and  $C_2$  depends on  $(I_p - I_q)$  or the contrast of the image according to (2.4)–(2.6). In our algorithm, we use the probability functions to formulate the edge capacities weighted by  $K_1$  or  $K_2$ . Then, for a given image, the  $K_L$  essentially depends on the contrast of the image.

### 2.3.3 Parameter Selection Strategy

The parameter  $\lambda$  determines the relative importance of the smoothness term and the data term in a graph cut-based segmentation framework (see (2.1)). Proper choice of this parameter largely influences the quality of the segmented output [4, 19]. In particular, large values of  $\lambda$  result in a nonsmooth boundary thereby leading to oversegmentation. On the other hand, small values of  $\lambda$  make the boundary too smooth causing undersegmentation. Thus selecting a correct  $\lambda$ , which would cause neither oversegmentation nor undersegmentation, remains a fundamental problem. Significant user interaction may be necessary at many cases to obtain optimal segmentation. In order to circumvent the above problem, Peng and Veksler [19] developed an automatic parameter selections strategy with a combination of different image features using AdaBoost. However, the above method involves many steps like feature extraction, feature normalization, and training the AdaBoost classifier and hence potentially requires significantly more time for completion compared to that of the standard graph cut. In this work, we have undertaken a purely probabilistic strategy for modulating the relative importance of the data term and the smoothness term, based on the location of a particular voxel in the image. The proposed method, unlike [19] is quite fast and has very similar execution time as that of the standard graph cut.

Instead of assigning a single value of  $\lambda$  to all the voxels, we estimate a separate value of  $\lambda_x$  for each individual voxel  $x$ . The value of  $\lambda_x$  depends on the location of the voxel  $x$ . If a voxel lies well within the foreground or the background region, the smoothness term gets more importance. In such cases, a voxel is typically surrounded by voxels whose probabilities to belong to any of the segmentation class are considerably higher than the other class. Consequently, we make the contribution of the smoothness term high by setting a lower value of  $\lambda_x$ . In contrast, if a voxel is close to the boundary, then we should give more importance to the data term. This is accomplished by setting a higher value of  $\lambda_x$ . Let  $A_y$  be the set of neighbors of any voxel  $x$  for which  $\Pr(I_y|O) > \Pr(I_y|B)$ . Similarly, let  $B_y$  be the set of neighbors of any voxel  $x$  for which  $\Pr(I_y|B) > \Pr(I_y|O)$ . Then, we set the value of  $\lambda_x$  using the following equation:

$$\lambda_x = \lambda_{\max} \exp(-((S_A - S_B)/D)^2) \quad (2.9)$$

Expressions for  $S_A$ ,  $S_B$ , and  $D$  are given below:

$$S_A = \sum_{i=1}^{|A_y|} \Pr(I_y|O) / d(x, y_i) \quad (2.10)$$

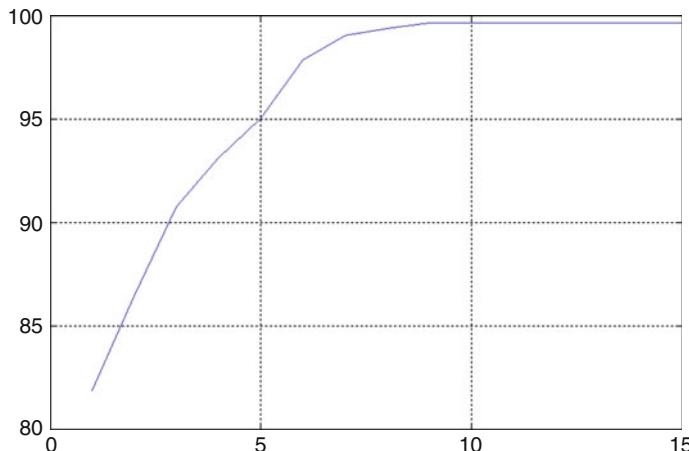
$$S_B = \sum_{i=1}^{|B_y|} \Pr(I_y|B) / d(x, y_i) \quad (2.11)$$

$$D = \sum_{i=1}^{|\text{Ne}(x)|} (1/d(x, y_i)) \quad (2.12)$$

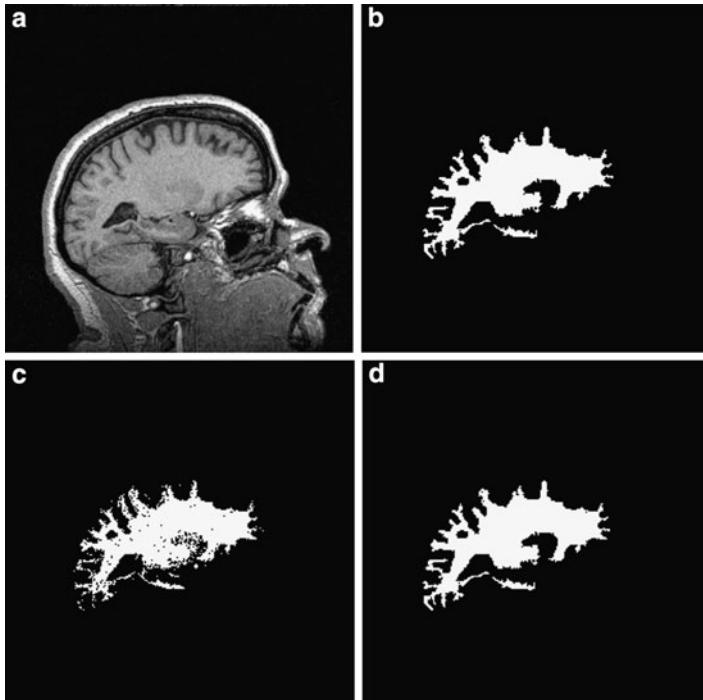
In (2.10),  $S_A$  is used to denote the total distance-weighted probability of the neighboring voxels in the set  $A_y$  to belong to the class “object”. Likewise, in (2.11),  $S_B$  is used to denote the total distance-weighted probability of the neighboring voxels in the set  $B_y$  to belong to the class “background”. The term  $D$  in (2.12) is used as the normalizing factor. The value of  $\lambda_{\max}$  is experimentally determined along the lines of [19]. When a voxel is well within the object region, we have  $S_A \gg S_B$ . Similarly, when a voxel is well within the background region, we have  $S_B \gg S_A$ . In both these cases, following (2.11), we get:  $\lambda_x \ll \lambda_{\max}$ . Consequently, the smoothness term gets more priority. When a voxel is on or near the edge of the foreground and the background, we have  $(S_A - S_B)^2 \rightarrow 0$ . Then, following (2.9), we get:  $\lambda_x \rightarrow \lambda_{\max}$ . Consequently, the data term gets more importance. In this manner, our probabilistic strategy provides appropriate weights to the data term and the smoothness term, based on the position of a voxel in the image.

### 2.3.4 Imposition of Shape Prior

A simple geometric shape prior [25] is applied to extract a subgraph  $G_1 = G_1(V_1, E_1)$  by pruning the original graph  $G = G(V, E)$ . As an alternative to more complex methods for shape modeling like Gradient Vector Flow (GVF) snakes [26], the current approach does not need any further user inputs (other than seeds) and has negligible computational overhead. The contour of the skull in the input axial data is



**Fig. 2.10** Variation of DC for the probabilistic graph cut algorithm with  $K = (K1/K2)$



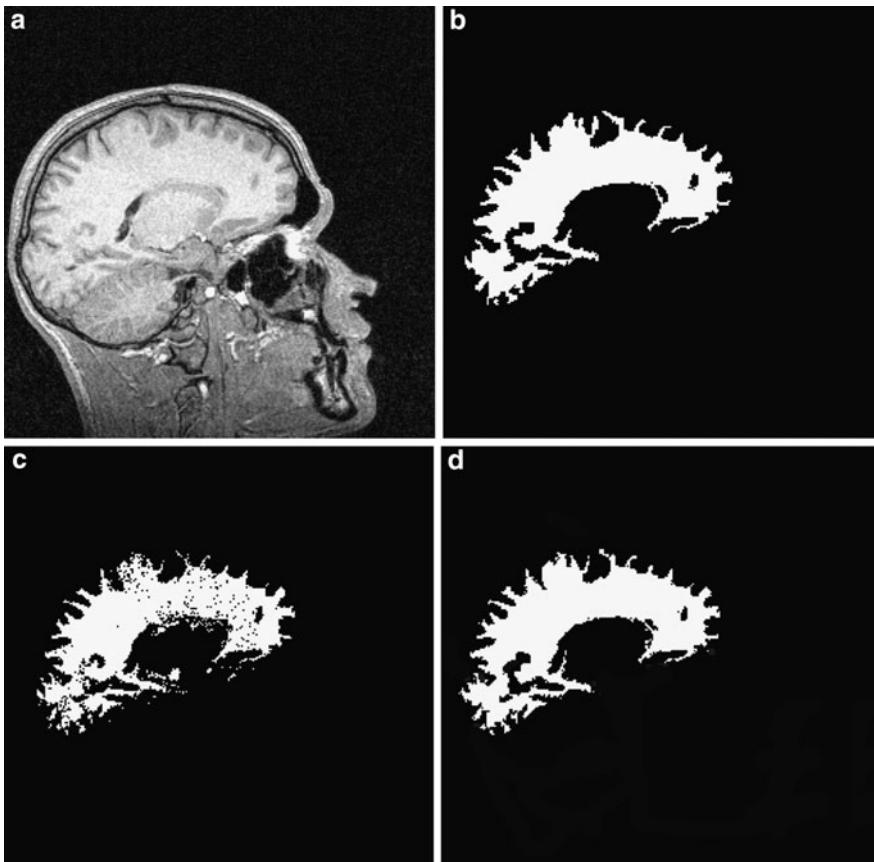
**Fig. 2.11** 2D Segmentation results from one dataset: (a) input, (b) ground-truth, (c) traditional graph cut (DC = 84.56%), (d) probabilistic graph cut (DC = 97.23%)

modeled as an ellipse with center  $(\alpha, \beta)$ , semi-major axis  $a$ , and semi-minor axis  $b$ . The equation for the ellipse is:

$$\frac{(x - \alpha)^2}{a^2} + \frac{(y - \beta)^2}{b^2} = 1 \quad (2.13)$$

The contours of the skull appearing in various 2D slices in the input can be appropriately modeled by a series of ellipses. The center of such a typical ellipse in any slice is obtained from the seeds inputted by the user in that particular slice. The other two parameters of the ellipse, namely,  $a$  and  $b$  are extracted from the image using simple geometry. In this way, we can get the ellipse in only few slices (where the user has provided seeds). For the rest of the slices, bilinear interpolation is employed to obtain the corresponding ellipses. Let  $\{p_x(k), p_y(k)\}$  and  $\{p_x(m), p_y(m)\}$  respectively denote the coordinates of a pixel  $p$  lying on the ellipse in the slice  $k$  and that of in the slice  $m$  respectively. Then, the coordinates of the pixel  $p$  lying on the ellipse in the slice  $l$ , where  $k < l < m$ , are given by:

$$p_x(l) = \frac{(l - k)}{(m - k)} p_x(k) + \frac{(m - l)}{(m - k)} p_x(m) \quad (2.14a)$$



**Fig. 2.12** 2D Segmentation results from a second dataset: **(a)** input, **(b)** ground-truth, **(c)** traditional graph cut ( $DC = 87.29\%$ ), **(d)** probabilistic graph cut ( $DC = 96.31\%$ )

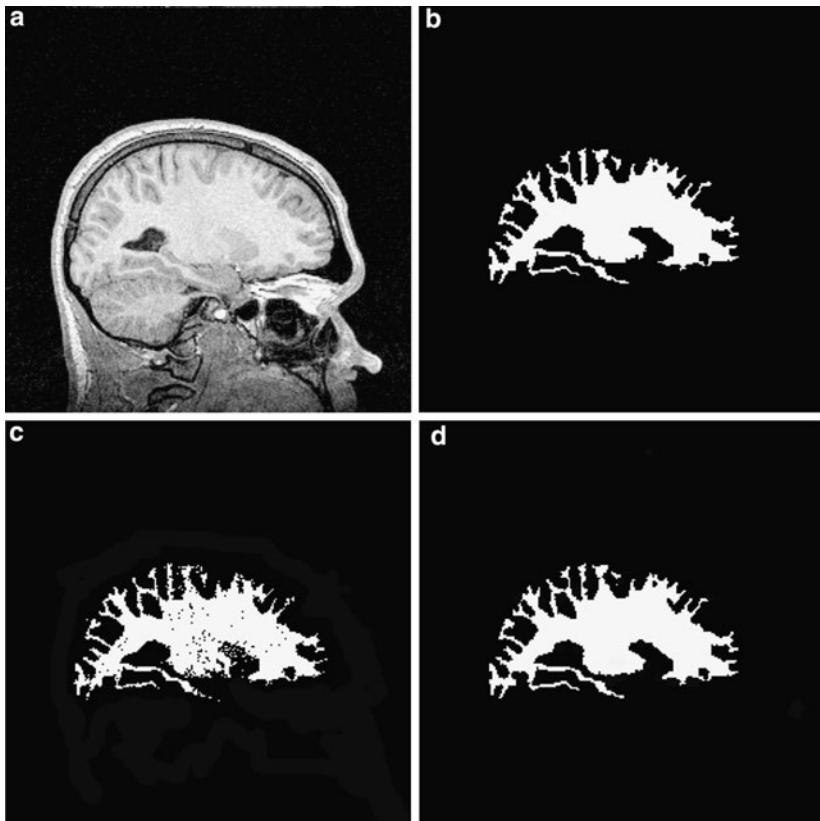
$$p_y(l) = \frac{(l - k)}{(m - k)} p_y(k) + \frac{(m - l)}{(m - k)} p_y(m) \quad (2.14b)$$

The potential target voxels lie inside the contour of the skull. Once these target voxels, *i.e.*, the vertices of the newly formed subgraph ( $P_1 \subset P$ ) are obtained, the  $t$ -links ( $T_1$ ) and the  $n$ -links ( $N_1$ ) are constructed. It is obvious that  $T_1 \subset T$  and  $N_1 \subset N$ . So, we can write  $V_1 = P_1 \cup s \cup t$  and  $E_1 = T_1 \cup N_1$ . The subgraph construction has two distinct advantages. Firstly, the time-complexity of the Ford-Fulkerson algorithm, used to find a minimum cut, is reduced considerably. Note that the augmenting paths are detected using the Edmonds-Karp algorithm [27]. The time-complexity of the Edmonds-Karp method, on the original graph  $G = G(V, E)$  is  $\mathbf{O}(VE^2)$  and that on the subgraph  $G_1 = G_1(V_1, E_1)$  is  $\mathbf{O}(V_1E_1^2)$ . Let us assume that:  $V_1 = \gamma V$  and  $E_1 = \delta E$ , where  $\gamma$  and  $\delta$  are two fractions. Let us put  $\mu = \gamma\delta^2$ . Obviously, the value of  $\mu \ll 1$ . Now, the time-complexity of the Edmonds-Karp algorithm on the subgraph

becomes  $\mathbf{O}(\mu VE^2)$ ; which is  $\ll \mathbf{O}(VE^2)$ . The second advantage is that the proposed construction has significantly improved the segmentation accuracy as will be shown in the next section.

## 2.4 Experimental Results

The results reported in this section are based on experimentation with different sets of T1-weighted MRI sequences. We have run our algorithm on both 2D and 3D brain images. For 2D experiments, typically, a single slice (of size  $256 \times 256$  pixels) is chosen from an entire 3D dataset in a completely unbiased manner while for 3D, we experiment with an entire dataset. The probabilistic graph cut-based segmentation results are compared with the ground-truth as well as with the results obtained from other standard segmentation algorithms like the traditional graph cut and the EMS. The ground-truth has been obtained from manual segmentation done



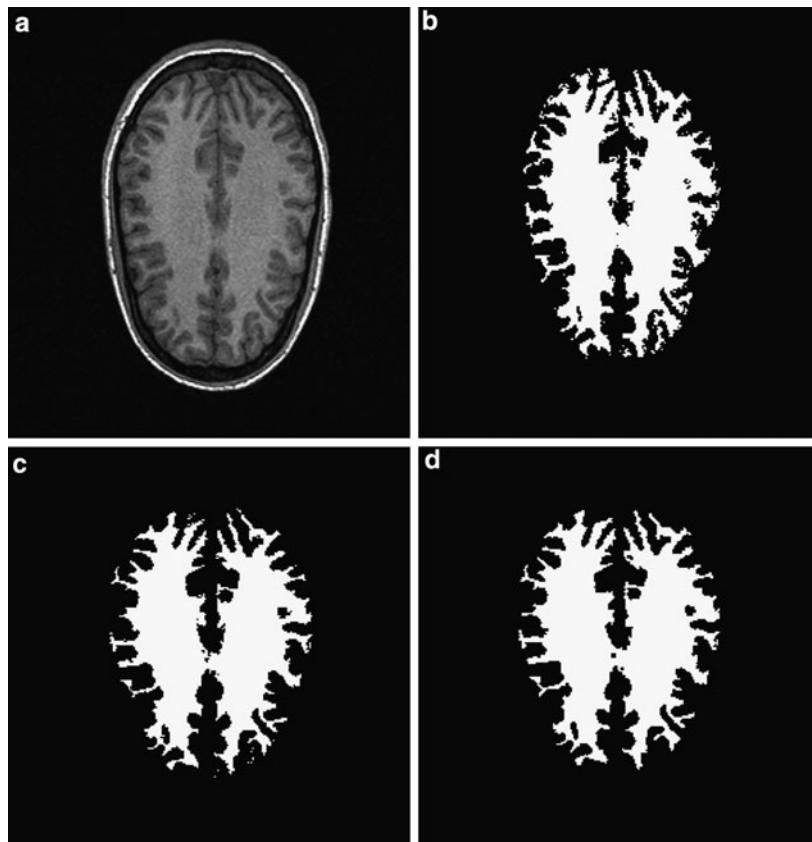
**Fig. 2.13** 2D Segmentation results from a third dataset: (a) input, (b) ground-truth, (c) traditional graph cut ( $DC = 90.85\%$ ), (d) probabilistic graph cut ( $DC = 97.17\%$ )

by the domain experts. In addition to qualitative comparisons, we also undertake quantitative validations using Dice coefficient. Dice coefficient DC ( $A, B$ ) between two sets  $A$  and  $B$  is defined below:

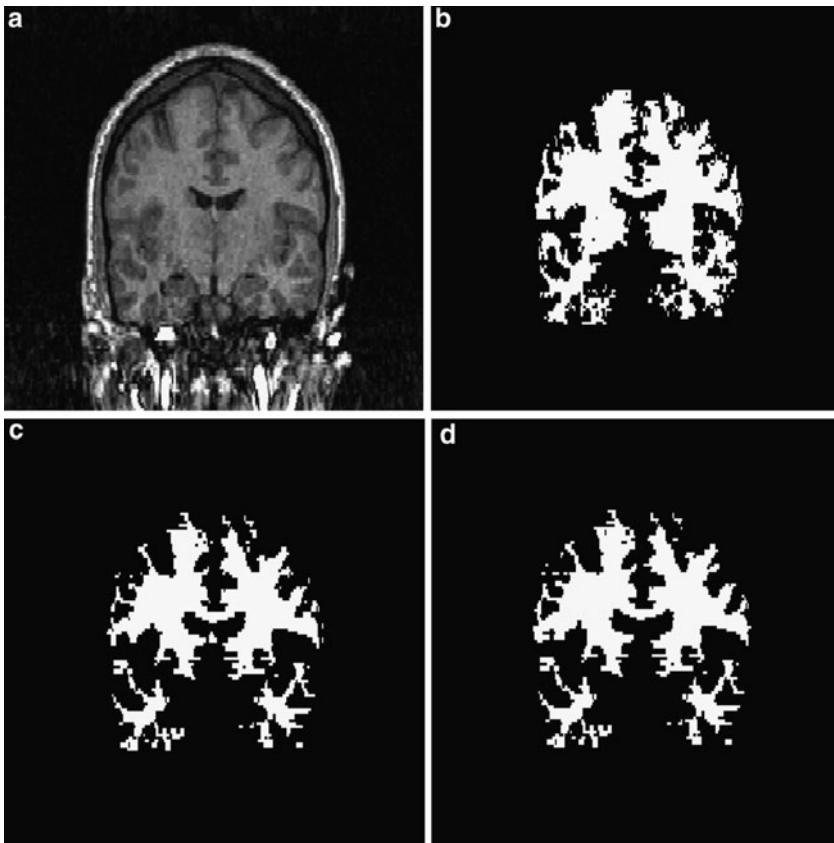
$$\text{DC}(A, B) = 2(|A \cap B|)/(|A| \cup |B|) \quad (2.15)$$

Values of dice coefficients for different segmentation algorithms with respect to the ground-truth are used to judge the relative performances of the different segmentation algorithms.

As a part of the experimental results, we first discuss the choice and estimation of different parameters for both 2D and 3D experiments. The seeds for the probabilistic graph cut algorithm are manually entered in only a few slices for any given dataset. The centers of the ellipses are obtained from the manually entered seeds in these slices using image geometry. For the rest of the slices, the ellipses are generated through bilinear interpolation. Both bilinear and biquadratic interpolations are



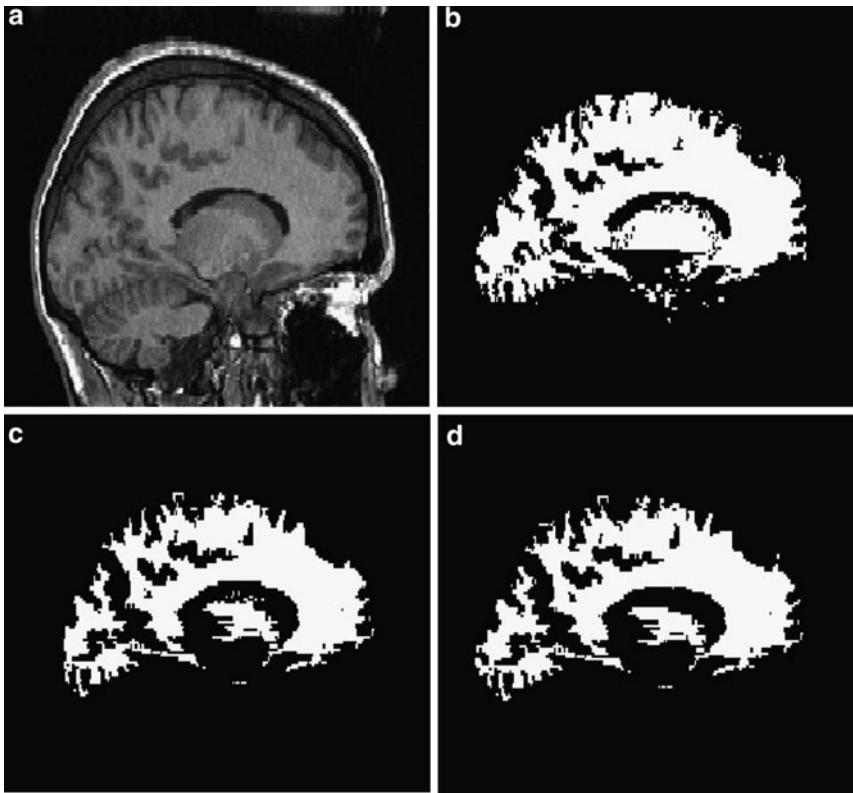
**Fig. 2.14** Axial view: (a) input, (b) graph cut output, (c) probabilistic graph cut output, (d) ground truth



**Fig. 2.15** Coronal view: (a) input, (b) graph cut output, (c) probabilistic graph cut output, (d) ground truth

employed. No noticeable differences are observed in the results from the above two interpolation techniques. So, we stick to the less complex bilinear interpolation. From the above discussion, it is also evident that very little user input is required in the entire process. The parameter  $\lambda_{\max}$  is kept fixed at 70. The value of the parameter  $\sigma$ , which can be treated as “camera noise”, is estimated as 3. It is shown in the Sect. 3.2, that for a low-contrast image, there is a lower limit of the parameter  $K$  (the ratio of  $K_1$  and  $K_2$ ) above which the accuracy of segmentation is independent of the value of  $K$ . If a suitable value of  $K$  (above the lower limit) is chosen, then the desired segmentation results can be achieved for the images of same or higher contrast. A plot of Dice coefficient of the probabilistic graph cut algorithm with variation of  $K$  for a typical dataset with 5% contrast is shown in Fig. 2.10.

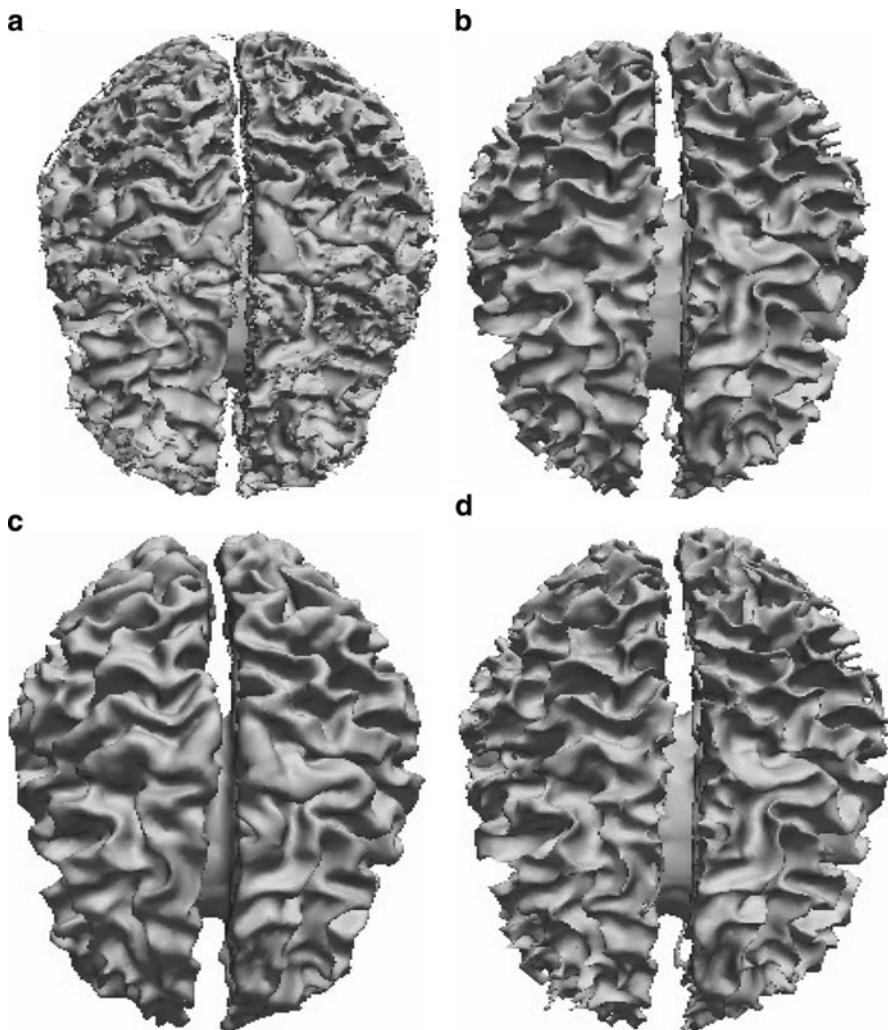
In Figure 2.10, we show that after  $K$  has reached a value of 8.7, the value of the Dice coefficient remains almost constant. We fix the value of  $K$  at 10 throughout because the other experimental datasets have contrasts greater than or equal to 5%.



**Fig. 2.16** Sagittal view: (a) input, (b) graph cut output, (c) probabilistic graph cut output, (d) ground truth

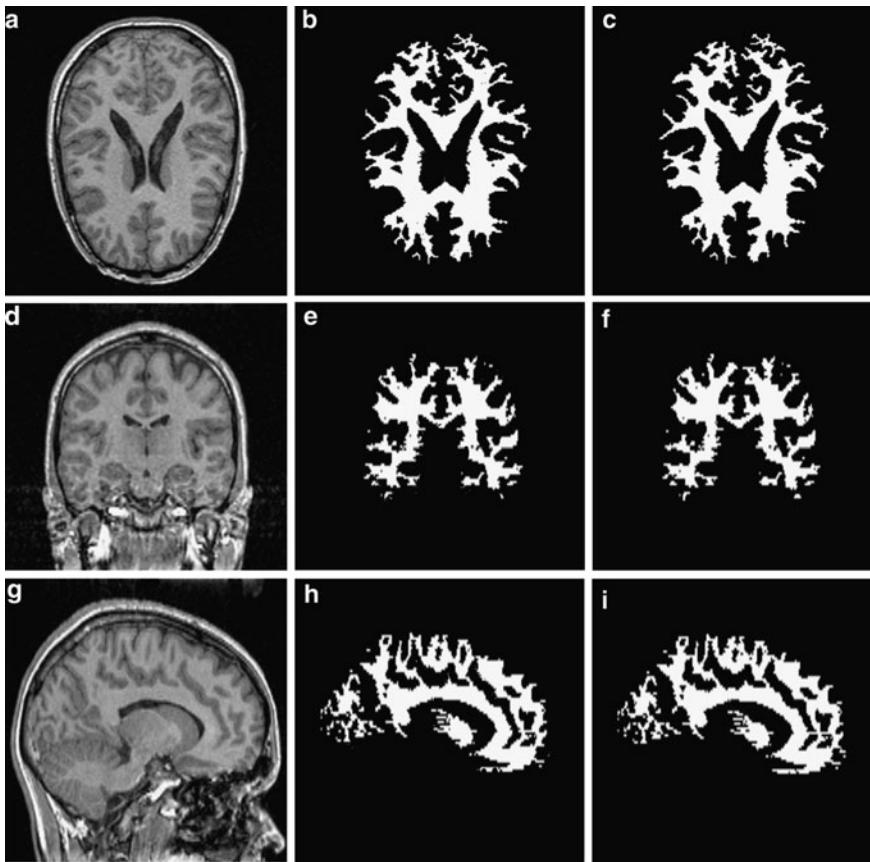
We next present the results of segmentation on three different 2D slices randomly selected from three different datasets. In Figs. 2.11–2.13, we show the outputs of the probabilistic graph cut-based segmentation, the traditional graph cut-based segmentation, and the ground-truth for these 2D slices. The figures clearly demonstrate that the probabilistic graph cut-based segmentation results are much closer in appearance to the respective ground-truths. For example, in Fig. 2.11, DC from the probabilistic graph cut (97.23%) is much higher than DC from the traditional graph cut (84.56%), which indicates superior performance of the probabilistic graph cut algorithm. Similarly, in Fig. 2.12, DC from probabilistic graph cut (96.31%) is once again much higher than DC from traditional graph cut (87.29%) which points to the consistent better performance of the probabilistic graph cut algorithm. Same nature of results is also corroborated by Fig. 2.13.

Now, we show the results of 3D segmentation using probabilistic graph cut algorithm for two MRI sequences. Each 3D sequence contains 124 slices (axial scans) and as mentioned earlier each slice has dimensions of 256 pixels  $\times$  256 pixels. Three different views, namely, the axial, the coronal, and the sagittal of the



**Fig. 2.17** 3D Segmented surfaces: (a) from graph cut ( $DC = 88.18\%$ ), (b) from probabilistic graph cut ( $DC = 96.99\%$ ), (c) from EMS ( $DC = 93.25\%$ ), and (d) ground truth

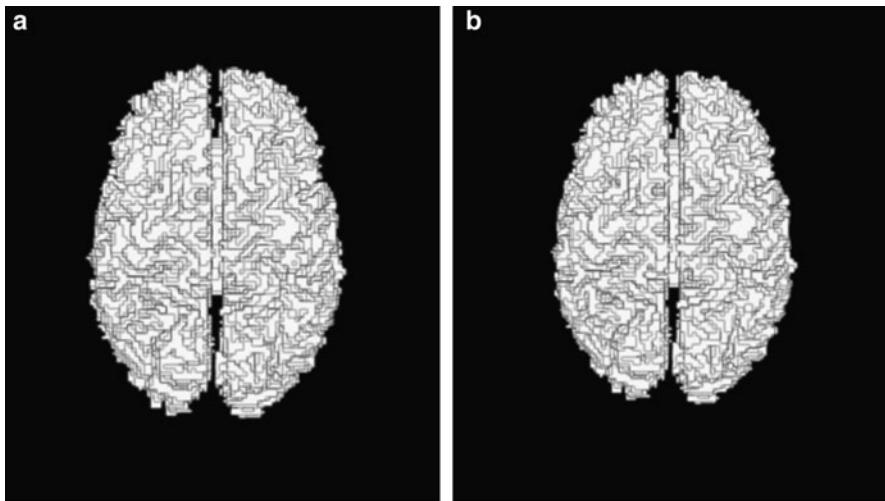
same 3D dataset are presented. Figure 2.14 shows the axial view of the input, graph cut-based segmented output, probabilistic graph cut-based segmented output, and ground-truth respectively. Figures 2.15 and 2.16 respectively illustrate the coronal and the sagittal view of the same. In Fig. 2.17, we show the 3D segmented surfaces from the probabilistic graph cut, traditional graph cut, EMS, and ground-truth. The segmented surface from the probabilistic graph cut algorithm bears closest resemblance with the surface obtained from ground-truth. As far as quantitative evaluations are concerned, DC obtained from the probabilistic graph cut (96.99%) is higher than both the traditional graph cut (88.18%) and the EMS (93.25%). In



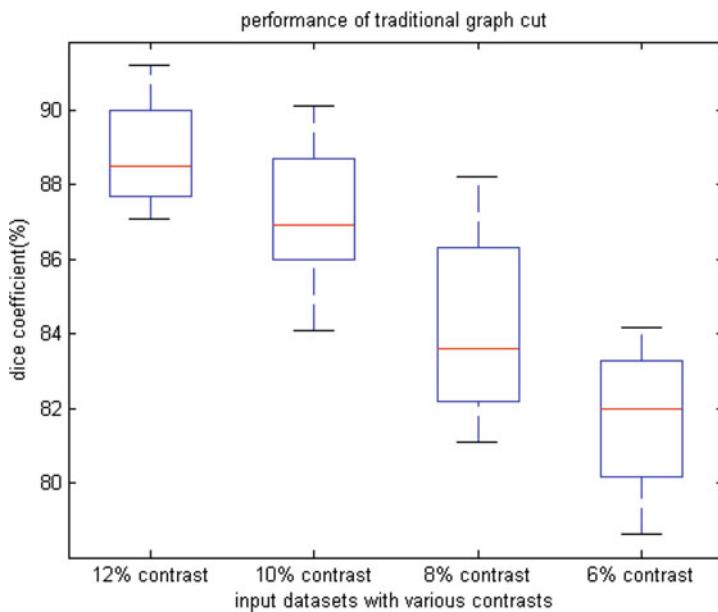
**Fig. 2.18** Axial view: (a) input, (b) probabilistic graph cut output, (c) ground truth. Coronal view: (d) input, (e) probabilistic graph cut output, (f) ground truth. Sagittal view: (g) input, (h) probabilistic graph cut output, (i) ground truth

Fig. 2.18, we present the axial, coronal and sagittal views of the input, probabilistic graph cut-based segmented output, and ground-truth of a second 3D dataset. Similarly, in Fig. 2.19, the 3D segmented surfaces from the probabilistic graph cut and the ground-truth are shown.

We end this section with a discussion on the robustness of our proposed method. The proposed method is fairly robust as its performance is found to be quite insensitive to the user inputs. For example, even if the centers of the ellipses are off by few pixels, only a negligible change ( $< 0.5\%$ ) is noticed in the value of the Dice Coefficient. The only basic requirement was that object (white matter in the present case) seeds had to be placed on object and the background seeds had to be placed on the background. In Figs. 2.20 and 2.21, we show two box plots [28] of DC of the traditional graph cut algorithm and that of the probabilistic graph cut algorithm with variations in contrast in the input MRI sequences respectively.

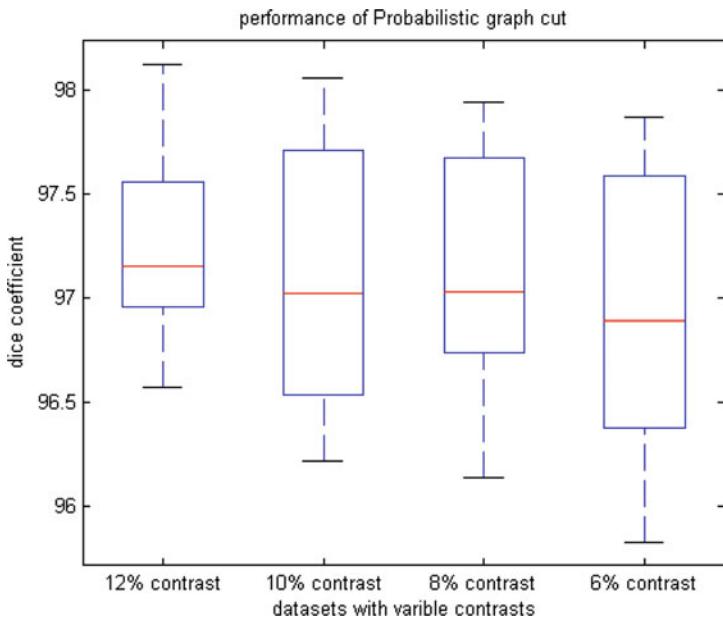


**Fig. 2.19** 3D Segmented surfaces of another MRI dataset: (a) from probabilistic graph cut (DC = 97.12%), (b) ground truth



**Fig. 2.20** Box plot of performance of the traditional graph cut algorithm with variation in contrast in the input MRI sequences

From Fig. 2.20, it can be easily observed that the mean DC (indicated by the red line) decreases considerably with decrease in contrast in the input datasets. On the other hand, Fig. 2.21 clearly demonstrates that the mean DC reduces only by a little amount



**Fig. 2.21** Box plot of performance of the probabilistic graph cut algorithm with variation in contrast in the input datasets

with the same decrease in contrast in the input datasets. Therefore, it can be inferred from these two figures that the probabilistic graph cut outperforms the traditional graph cut, more so for images with low contrasts.

## 2.5 Conclusion

In this paper, we have addressed the problem of segmentation of cerebral white matter from T1-weighted MRI data. The problem is of great interest to both the computer vision as well as the medical community. The segmentation problem becomes very challenging due to the complicated shape of the white matter and extremely poor contrast between the object and the surrounding structures in the input data. From the medical point of view, accurate segmentation of cerebral white matter will enable us to better understand different neurological disorders. We propose a probabilistic graph cut algorithm, where a probabilistic modification is applied to the traditional graph cut algorithm. We also adopt a probabilistic strategy for estimating the value of the weighting parameter for the smoothness and the data terms for each individual voxel in the input. In addition, we impose a simple yet effective shape prior in form of a series of ellipses to model the human skull in the axial 2D slices constituting an input sequence which increases the speed of execution as well as the accuracy of the segmentation. The method proposed in this paper

is extremely fast and has outperformed both the traditional graph cut and the EMS algorithms. More sophisticated techniques like GVF snakes can be explored in conjunction with the proposed probabilistic graph cut algorithm for modeling the human skulls to further improve the segmentation accuracy.

## References

1. Levy SE, Mandell DS, Schultz RT (2009) Autism. *Lancet* 374(9701):1627–1638
2. Warfield SK, Westin CF, Guttman CRG, Albert M, Jolesz FA, Kikinis R (1999) Fractional segmentation of white matter. In: Taylor C, Colchester A (eds) MICCAI 1999. LNCS, vol 1679. Springer-Verlag Berlin Heidelberg, pp 62–72
3. Yang F, Jiang T, Zhu W, Kruggel F (2004) White matter lesion segmentation from volumetric MR images. In: Yang G-Z, Jiang T (eds) MIAR 2004. LNCS, vol 3150. Springer-Verlag Berlin Heidelberg, pp 113–120
4. Boykov Y, Jolly MP (2001) Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: IEEE international conference on computer vision, pp 105–112
5. Boykov Y, Veksler O, Zabih R (2001) Fast approximate energy minimization via graph cuts. *IEEE Trans Pattern Anal Mach Intell* 23(11):1222–1239
6. Ishikawa H (2003) Exact optimization for Markov random fields with convex priors. *IEEE Trans Pattern Anal Mach Intell* 25(10):1333–1336
7. Kim J, Zabih R (2003) A segmentation algorithm for contrast-enhanced images. In: IEEE international conference on computer vision, pp 502–509
8. Boykov Y, Funka-Lea G (2006) Graph cuts and efficient N-D image segmentation. *Int J Comput Vis* 70(2):109–131
9. Van Leemput K, Maes F, Vandermeulen D, Colchester A, Suetens P (2001) Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans Med Imaging* 20(8):677–688
10. Chowdhury AS, Rudra AK, Sen M, Elnakib A, El-Baz A (2010) Cerebral white matter segmentation from MRI using probabilistic graph cuts and geometric shape priors. In: IEEE international conference on image processing, (to appear)
11. Atkins M, Mackiewich B (1998) Fully automatic segmentation of the brain in MRI. *IEEE Trans Med Imaging* 17(1):98–107
12. Farag A, El-Baz A, Gimel’farb G (2006) Precise segmentation of multi-modal images. *IEEE Trans Image Process* 15(4):952–968
13. Thacker NA, Jackson A (2001) Mathematical segmentation of grey matter, white matter and cerebral spinal fluid from MR image pairs. *Br J Radiol* 74:234–242
14. Collins D, Peters T, Dai W, Evans A (1994) Model based segmentation of individual brain structures from MRI data. *SPIE Vis Biomed Comput* 1808:10–19
15. Kass M, Witkin A, Terzopoulos D (1987) Snakes: active contour models. *Int J Comput Vis* 1:321–331
16. Caselles V, Kimmel R, Sapiro G (1997) Geodesic active contours. *Int J Comput Vis* 22:61–79
17. Zeng X, Staib L, Duncan J (1998) Volumetric layer segmentation using coupled surface propagation. In: IEEE international conference on computer vision pattern recognition, pp 708–715
18. Wolz R, Aljabar P, Rueckert D, Heckemann AR, Hammers A (2009) Segmentation of subcortical structures in brain MRI using graph-cuts and subject-specific a-priori information. In: IEEE international symposium on biomedical imaging, pp 470–473
19. Peng B, Veksler O (2008) Parameter selection for graph cut based image segmentation. In: British machine vision conference, pp 1–10

20. Freedman D, Zhang T (2005) Interactive graph cut based segmentation with shape priors. In: IEEE international conference on computer vision, pp 755–762
21. Song Z, Tustison N, Avants B, Gee J (2006) Adaptive graph cuts with tissue priors for brain MRI segmentation. In: IEEE international symposium on biomedical imaging, pp 762–765
22. Zhang J, Wang Y, Shi X (2009) An improved graph cut segmentation method for cervical lymph nodes on sonograms and its relationship with node's shape assessment. Comput Med Imaging Graph 33:602–607
23. Ford LR, Fulkerson DR (1962) Flows in networks. Princeton University Press, Princeton, NJ
24. Kolomogorov V, Zabih R (2004) What energy functions can be minimized via graph cuts. IEEE Trans Pattern Anal Mach Intell 26(2):147–159
25. Slabaugh G, Unal G (2005) Graph cuts segmentation using an elliptical shape prior. In: IEEE international conference on image processing, pp 1222–1225
26. Xu C, Prince JL (1998) Snakes, shapes and gradient vector flow. IEEE Trans Image Process 7(3):359–369
27. Cormen TH, Leiserson CE, Rivest RL, Stein C (2001) Introduction to algorithms. MIT Press, Cambridge, MA
28. Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading, MA

## Biography



**Mainak Sen** passed his 12th standard from Ramakrishna Mission Vivekananda Centenary College, West Bengal in 2007. He is currently an undergraduate student of Bachelor in Engineering (B.E.) in Department of Electronics and Telecommunication Engineering., Jadavpur University, India. He worked as research assistant in Department of Computing Science, University of Alberta, Canada. His research interest includes computer vision, pattern recognition, and biomedical image processing.



**Ashish Kumar Rudra** got his 12th standard degree from Krishnanagar Collegiate School, West Bengal in 2007. Now, he is a 4th year student of Bachelor in Engineering (B.E.) in Department of Electronics and Telecommunication Engineering, Jadavpur University, India. He worked as research assistant in CIMS Lab, Department of Computing Science, University of Alberta, Canada. His research interest includes computer vision, pattern recognition and biomedical image processing. He is also interested in abstract algebra, literatures recitation and painting.



**Ananda S Chowdhury** received a B.Sc. (Hons.) in Physics from Presidency College, Calcutta, India in 1996, a B.Tech. in Electronics Engineering from the Institute of Radiophysics and Electronics, Calcutta, India in 1999, and an M.E. in Computer Engineering from Jadavpur University, Calcutta, India in 2001. He earned his Ph.D. in Computer Science at the University of Georgia, Athens, Georgia in July 2007. From July 2007 to December 2008, he was as a postdoctoral fellow in Radiology and Imaging Sciences at the National Institutes of Health, Bethesda, Maryland. At present, he is working as a Reader in the Electronics and Telecommunication Engineering department of Jadavpur University, Calcutta, India.



**Ahmed Elnakib** received his B.Sc. and M.Sc. degrees in Electronics and Communication Engineering from Mansoura University, Egypt, in 2003 and 2007, respectively. Mr. Elnakib joined the BioImaging Laboratory as research assistance in September 2008. His current research includes developing new computer assisted diagnosis systems, image modeling, image segmentation, 2-D and 3-D registration, and shape analysis of detected lung nodules.



**Ayman El-Baz** received the B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering Department, University of Louisville, in August 2006. His current research is focused on developing new computer-assisted diagnosis systems for different diseases and brain disorders

# Chapter 3

## A New Image-Based Framework for Analyzing Cine Images

Fahmi Khalifa, Garth M. Beach, Georgy Gimel’farb,  
Guruprasad A. Giridharan, and Ayman El-Baz

**Abstract** Noncontrast agent cine Cardiac Magnetic Resonance (CMR) acquisitions using gated images taken through the cardiac cycle to characterize function are a well-established part of a comprehensive CMR exam. Although these methods have long been explored, in practice, efficient parameter determination and effective displays remain an area that could benefit from continued improvement to encourage their routine use. In this chapter, we propose a novel and automated framework to (1) segment the inner and outer myocardial borders (contours) from the surrounding tissues in cine CMR images based on using a new stochastic speed function to control the evolution of a geometric (level set) deformable model, and (2) to estimate local wall thickness and thickening function index based on solving Laplace equation in the segmented wall to find point-to-point correspondences between the inner and outer border of the segmented wall. Effects of segmentation errors are reduced and a continuity analysis of the wall thickening values is performed by iteratively computing the Maximum A Posteriori (MAP) estimate of a pairwise energy function of a generalized Gauss-Markov random field (GGMRF) probabilistic model. We also use cavity areas derived from the inner region to index global function parameters (e.g., ejection fraction). Experiments with *in vivo* CMR data collected from eight independent data sets confirm higher robustness and accuracy of the proposed framework over the existing ones.

**Keywords** Wall thickening analysis · Regional myocardial function · Level set · Markov-Gibbs random field · Shape prior · Laplace equation

---

A. El-Baz (✉)

BioImaging Laboratory, Department of Bioengineering, University of Louisville, 423 Lutz Hall Building, Louisville, KY 40292, USA  
e-mail: [aselba01@louisville.edu](mailto:aselba01@louisville.edu)

### 3.1 Introduction

Wall mass, cavity volumes, and global performance indexes (such as end systolic and end diastolic volumes), and classical function indexes (i.e., wall thickness and thickening) are all important indicators of the status of integrative cardiac physiology in health and disease [1]. These indexes help clinicians to detect heart diseases, characterize major adult heart conditions that lead to heart attack or failure, and guide optimized therapies for individual patients. Note that more continuum mechanics-based indexes using strain or strain rate have been previously explored; however, these approaches have evolved their own specialized analysis approaches and will not be addressed in this treatise.

*Related Work in Cardiac Image Segmentation:* To determine ventriculometrics requires accurate segmentation of the inner and outer myocardial borders (contours) in cardiac images. Specifically, the segmentation of the Left Ventricle (LV) in cine Cardiac Magnetic Resonance (CMR) images has been the motivation of many research works, and many methods for 2-D, 3-D, or 4-D analysis have been presented. However, the segmentation is a challenge due to difficulties of delineating the inner border from the cavity, protrusions of papillary muscle structures into the cavity, partial influence of adjacent structures such as the diaphragm and abdominal contents, and artifacts from blood circulation within the ventricular cavity [2]. Thus, there remains room for improvement of existing image segmentation methods (e.g., [3–5]) for efficient, robust, and accurate assessment.

Most successful known approaches have addressed the challenge by modeling object appearance and shape. In particular, Mitchell et al. [6] presented a multi-stage hybrid Active Appearance Model (AAM) for the segmentation of 2-D MR image sections of the heart. They later extended their work to a full 3-D AAM in [7]. A multi-view AAM for consistent segmentation of multiple cardiac views in long- and short-axis cardiac MR images was proposed by Lelieveldt et al. [8]. Qian et al. [9] combined the hybrid AAM with an Active Shape Model (ASM) to segment the tagged MR images. Lin et al. [10] proposed an automatic framework to segment the left and right ventricular inner (endocardium) and outer (epicardium) borders in CMR images using a simple contour detection method combined with temporal Fourier analysis of image cines. Lötvönen et al. [11] presented a method based on a statistical point distribution model and a mean grayscale model.

However, segmentation frameworks that rely on parametric shape models have some drawbacks. First, they require landmark points that are carefully chosen and manually placed on all examples by experts to assure good correspondences. Second, manual landmark placement makes the training process time-consuming, which can be particularly tedious in 3-D cases. Furthermore, parametric shape models become unsuitable for discontinuous objects due to a small number of distinct cardiac landmarks. Additionally, appearance modeling may fail in the presence of large grey value variability across subjects and time.

To partially overcome the parametric shape model limitations, Chen et al. [12] incorporated the shape statistics and an edge image in a variational framework for the segmentation of the outer border of the LV wall in MR images. Taleb-Ahmed et al. [13] located the inner and outer contours in a signal processing context by using nonlinear filters that are based on the invariance properties of the image. de Bruijne and Nielsen [14] extracted the LV by an adopted shape particle filtering by use of a maximum likelihood shape inference that is based on pixel classification. Zhu et al. [15] proposed a dynamical shape prior for the segmentation of the LV to simultaneously account for intersubject variability as well as cardiac dynamics. A semiautomated framework by Li et al. [16], tested on CT, perfusion MR, and cine CMR images, exploits two different energy functionals for segmenting the inner and the outer borders and forms a weighted estimate of the myocardial thickness. The functionals are derived from a dual “object-background” intensity model and a shape prior.

Deformable model-based segmentation has been successfully employed in many CMR segmentation studies. For example, a variational framework by Paragios [17] combined the boundary- and region-based LV segmentation in the CMR images by using a Geodesic Active Region (GAR) model. To obtain a more accurate automatic level set-based segmentation of the rat CMR images, Pluempiwiriyawej et al. [18] added an elliptic shape prior of the heart to probabilistic region- and edge-based information. Recently, Bernard et al. [19] coupled a level set-based approach with a probability model of intensity distributions in echographic heart images. However, both active contours and level set tend to fail in the case of excessive noise, poor image resolution, diffuse boundaries, or occluded shapes, if they do not incorporate a priori models (e.g., appearance, shape).

Xu and Prince [20] explored a Gradient Vector Flow (GVF) field and used it as an additional external force term to constrain snake constructs in medical images. Liang et al. [2] proposed an automated radial GVF approach to segment the LV contours from short-axis MR images using the Hough transform. A Generalized Fuzzy Gradient Vector Flow (GFGVF) was proposed by Zhou et al. [21] to guide a parametric deformable model and by Qin et al. [22] to guide a level set-based deformable model. Recently, a semiautomated version of [20] was proposed by Chen et al. [23]. They used a nonparametric model of the background intensity distribution and a self-adaptive parametric shape prior model.

In summary, several classes of segmentation approaches have been proposed in the last few years to address the challenges of cardiac image segmentation. However, in most of these methods, the segmentation reliability is not very high due to the following reasons:

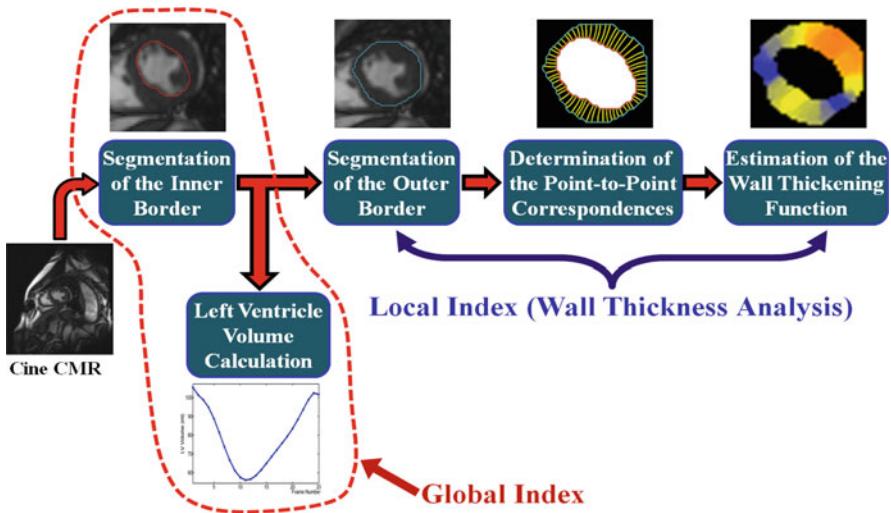
1. The complexity of the LV, especially that of the outer contour, due to broken or discontinuous object boundaries, large image noise, and inhomogeneities.
2. Large variations of goal shape, especially that of the inner region of the LV, within a specific time series of the same image section in the same subject, as well as goal shape variations across subjects.

3. Shape-guided approaches are impractical for clinical applications because they typically require a large training set and a tedious training process, and often fail to capture the variability across subjects.
4. Most of the proposed approaches required user input and interaction.

*Related Work in Wall Thickness Measurements:* Measuring wall thickness throughout the cardiac cycle in CMR images provides key local information for deriving integrated physiological parameters of an individual in health and disease [1]. To reliably measure the LV wall thickness and assess the wall thickening, one has to accurately segment inner and outer wall boundaries in the CMR images. After segmenting the borders, the wall thickening is typically assessed by visual inspection, being prone to considerable intra and interobserver variability. Automated or semiautomated wall thickness measurement techniques (see e.g., [24–32]) overcome this drawback. The measurements often exploit the centerline algorithm [33] to determine planar distances between the inner and outer boundaries for each image. For 3-D wall thickness determination, a 3-D surface is constructed at mid-myocardium through all the midpoints of the centerline of the 2-D images.

However, these methods are typically sensitive to the contours' imperfections, subject to errors in noisy images, and still require a search method to collocate the corresponding borders points. Additionally, the point-to-point correspondences are miscalculated in the case of broken or irregular boundaries which are common in CMR images. Recently, Prasad et al. [34] proposed to measure the myocardial thickening in CMR images more reliably by solving a partial differential Laplace equation. However, their approach was based on manual tracing of the segmented contours by a cardiologist, i.e., was prone to the interobserver variability, and did not address the one-pixel distance scenario where both the inner and outer contours share common pixels.

To overcome the aforementioned limitations of the LV wall borders segmentation challenges as well as wall thickness measurements, we propose an automatic framework to analyze cine CMR images (Fig. 3.1). The inner and outer contours of the LV wall are segmented from surrounding anatomic structures in a cine CMR image with a geometric deformable model guided by a new stochastic speed function that accounts for a shape prior and features of image intensity and spatial interaction. We used a probabilistic map to describe the prior shape of the outer contour of LV wall. Due to the dynamical nature of the heart, we developed an adaptive approach to automatically update the prior calculated probabilistic map to accurately handle the inner dynamical variability of the inner contour of the LV. The wall thickness is accurately measured on the basis of point-to-point correspondences between the inner and the outer wall borders found by solving the Laplace equation. Then the measurements are refined using their Gauss-Markov random field model and Maximum A Posteriori (MAP) estimates. Quantitative results with *in vivo* CMR real-time (RT) series (an image time sequence through the cardiac cycle, taken of the same image section) for eight independent data sets confirm higher robustness and accuracy of the proposed framework over the existing ones.



**Fig. 3.1** The proposed framework for analyzing cine CMR images

## 3.2 Methods and Data Acquisition

### 3.2.1 MRI Data Collection

To acquire images for ventriculometrics, we used images collected as time series data, through a cardiac cycle. Modern acquisitions use the True FISP technique [35] which provides high contrast resolution, with adequate temporal resolution to depict key physiological events, without the administration of external contrast agents. In our studies, images were acquired using a Siemens 1.5T system, with multichannel phased array reception coils. Typical spatial resolution was  $1.4 \times 3.1 \times 8.0 \text{ mm}^3$ .

### 3.2.2 Myocardial Borders Segmentation

The segmentation of anatomical structures in medical images is a fundamental step for many high-level processing tasks such as image analysis, Computer-Assisted Diagnosis (CAD), geometric modeling of anatomical structures, and the construction of biomechanical models used for surgical simulation [36]. However, accurate segmentation of anatomical structures still remains a challenging problem due to many reasons: (1) poor image quality, (2) large variations of goal shapes, (3) image noise and inhomogeneities, and (4) discontinuous boundaries due to occlusions, similar visual appearance of adjacent parts of different structures, etc. [37].

Geometric (level set) deformable models are very popular in today's medical image analysis due to the flexibility of an evolving contour on the  $xy$ -plane and the lack of need for parameterization [38]. The object-background boundary at each moment  $t$  is represented by a zero level  $\phi_t(x, y) = 0$  of an implicit level set function, namely, a distance map  $\phi_t(x, y)$  containing the signed minimum Euclidean distances from every point  $(x, y)$  to the boundary (negative for interior and positive for exterior points) [39]. The distance map is evolved iteratively [40], the evolution being guided by a speed function  $V_n(x, y)$ :

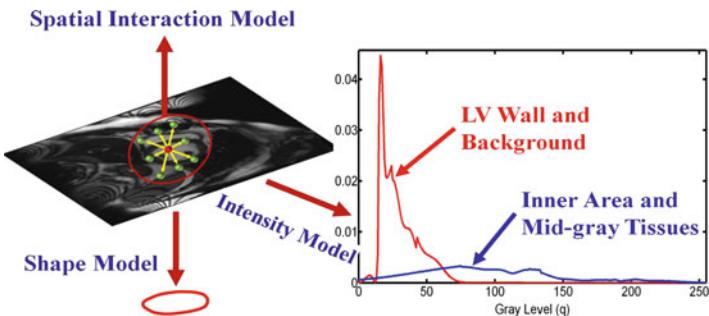
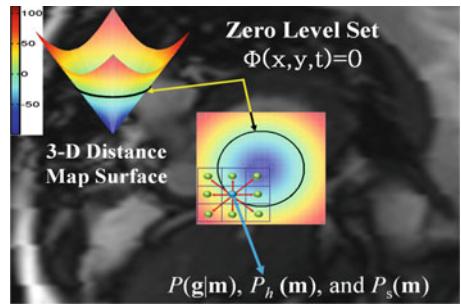
$$\phi_{n+1}(x, y) = \phi_n(x, y) - \tau V_n(x, y) |\nabla \phi_n(x, y)| \quad (3.1)$$

where  $n$  indicates the time instants  $t = n\tau$  (taken with a step  $\tau > 0$ ) and  $\nabla \phi_n(x, y) = \left[ \frac{\partial \phi_n}{\partial x}, \frac{\partial \phi_n}{\partial y} \right]$  is the gradient of  $\phi_n(x, y)$ .

Conventional speed functions, which account for image intensities (e.g., [41, 42]), object edges (e.g., [43, 44]), and intensity gradient vector flow (e.g., [4, 45, 46]), etc., are unsuccessful on very noisy images with low object-background intensity gradients. More effective segmentation is obtained after adding shape priors (e.g., [47–53]). Chen et al. [47] calculated an average of a registered training set to be a shape prior, and defined an energy functional that basically minimizes a Euclidean distance between a given point and its shape prior. Tsai et al. [48] approximated an object shape with a linear combination of distance maps for a set of mutually aligned images of goal objects. However, due to the very large dimensionality of the distance maps' space, only a few top-rank principal components can be included in the linear combination to make the model computationally feasible. To overcome oversegmentation, Cheng et al. [50] incorporated a prior shape knowledge into the improved Chan-Vese level set model [5] using a variational approach. Leventon et al. [51] and Shen and Davatzikos [52] augmented speed functions with terms attracting the level set front to more likely shapes specified by analyzing principal components of a training set of objects. Litvin and Karl [53], formed a shape boundary prior from the features of the boundary, and this boundary was used within a level set framework.

However, the speed function based only on the shape priors still results in large errors under discontinuities in object boundaries, large image noise, and other image inhomogeneities [54]. Also, the approximation of a distance map  $\phi_n(x, y)$  with the closest linear combination of the maps for a set of coregistered training images has a serious inherent drawback: the distance maps' space is not closed with respect to linear operations, so the zero level of a linear combination of the training maps may produce an invalid or even physically impossible boundary. The proposed stochastic speed function overcomes the above limitations by accounting for the shape prior and features of image intensity and spatial interaction. These features being integrated with a two-level joint Markov-Gibbs Random Field (MGRF) model of the LV and background (see Fig. 3.2).

**Fig. 3.2** Integrating image features for the level set evolution



**Fig. 3.3** Aligning a joint MGRF model to a shape prior

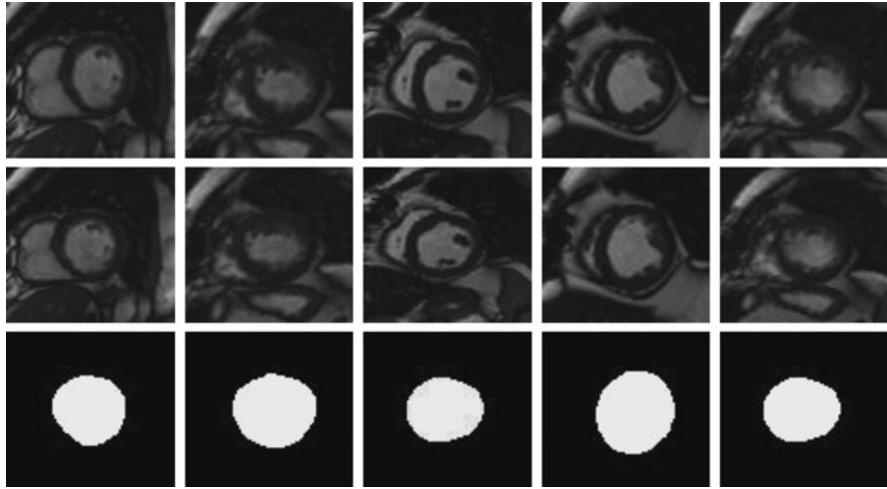
### 3.2.2.1 Joint MGRF Model of an LV and Background

Let  $\mathbf{g} : \mathbf{R} \rightarrow \mathbf{Q}$  and  $\mathbf{m} : \mathbf{R} \rightarrow \mathbf{L}$  be a grayscale image to be segmented and its binary “object – background” map, respectively, supported by a 2-D arithmetic  $(x, y)$ -lattice  $\mathbf{R}$ . Here,  $\mathbf{Q} = \{0, 1, \dots, Q - 1\}$  is a set of  $Q$  integer gray levels and  $\mathbf{L} = \{0, 1\}$  is a set of object (“1”)/background (“0”) labels. The image  $\mathbf{g}$  and its map  $\mathbf{m}$  are described with a two-level joint MGRF model combining a 2nd-order MGRF map model and a conditionally independent random field of image intensities given the map as shown in Fig. 3.3.

### 3.2.2.2 Shape Model

In the area of medical imaging, organs have well constrained forms within a family of shapes. Thus, additional constraints based on the shape of the objects are greatly needed to overcome the problems associated with discontinuous objects and image inhomogeneities and therefore enhance the segmentation accuracy. Below, we will describe in details the steps of building the shape priors for both inner and outer contours.

*Outer Contour Shape Prior:* To segment the outer contour, we used a probabilistic shape prior (built at the learning stage from different subjects) that accounts



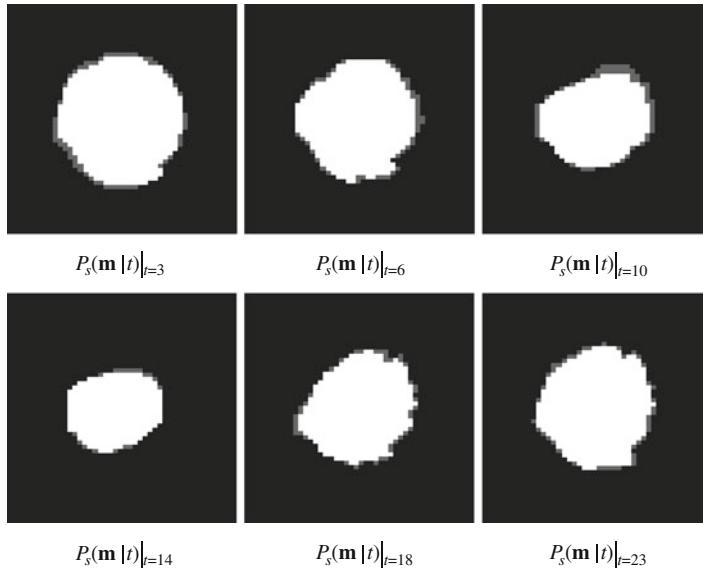
**Fig. 3.4** Forming outer shape prior: top row – training samples; middle row – their affine MI-based alignment, and bottom row – manually segmented outer regions

**Fig. 3.5** Gray-coded outer region shape prior



for all possible variabilities. In this work, the outer shape prior is integrated into the two-level joint MGRF model as follows: the two-level joint MGRF model is described with the probability distribution  $P(\mathbf{g}, \mathbf{m}) = P(\mathbf{g}|\mathbf{m})P(\mathbf{m})$ . Here, the map model  $P(\mathbf{m}) = P_h(\mathbf{m})P_s(\mathbf{m})$  has two independent parts: a second-order MGRF model  $P_h(\mathbf{m})$  of a spatially homogeneous evolving map  $\mathbf{m}$  for the image  $\mathbf{g}$  and a shape prior being a spatially variant independent random field of region labels  $P_s(\mathbf{m}) = \prod_{(x,y) \in \mathbb{R}} p_{s,x,y}(m_{x,y})$ , where the factors are the empirical pixel-wise object,  $p_{s,x,y}(1)$ , and background,  $p_{s,x,y}(0) = 1 - p_{s,x,y}(1)$ , probabilities, respectively, for the manually segmented training images coaligned by a rigid 2-D transformation.

Figures 3.4 and 3.5 illustrate the process of building the outer contour shape prior at the midventricular level. The CMR images from the training set are coaligned by a rigid 2-D transformation maximizing their mutual information (MI) [55]. The outer areas are manually segmented by a medical imaging expert, and their pixel-wise object prior probabilities  $p_{s,x,y}(1)$  are estimated by counting the frequencies of occurrence of the pixel  $(x,y)$  in the outer area of the LV (Fig. 3.5). Parameters of the 2-D rigid transformation are estimated using the genetic algorithm.



**Fig. 3.6** Samples of the inner region adaptive shape prior at different time instances

*Inner Contour Shape Prior:* Due to the dynamical nature of the heart, we used a subject-specific dynamical shape prior that is adaptively updated to accurately handle the dynamical variability of the inner contour of the LV wall. The inner contour shape prior at time  $t$  is described with a spatially variant independent random field of region labels  $P_s(\mathbf{m}|t) = \prod_{(x,y) \in \mathbb{R}} P_{s,x,y}(m_{x,y}|t)$ , and adaptively constructed (see Fig. 3.6) using the segmented inner regions starting from  $t = 0$  up to  $t - 1$ .

### 3.2.2.3 Intensity Model

Just as in [42, 56], specific visual appearance of the LV in each data set to be segmented is taken into account by modeling a marginal gray level distribution within each current evolving boundary with a mixture of the object and background components. Such a model adapts the segmentation to changing visual appearance, i.e., nonlinear intensity variations caused by patient weight and data acquisition system (scanner type and scanning parameters). Close approximation with a linear combination of discrete Gaussians (LCDG) separates each factor of the joint empirical gray level distribution,  $P(\mathbf{g}) = \prod_{(x,y) \in \mathbb{R}} P_{\text{mix}}(g_{x,y})$  mix, into two (object and background) components, ( $p(q|\lambda); q \in \mathbf{Q}$ , and  $\lambda \in \mathbf{L}$ ). The LCDG modeling restores transitions between these components more accurately than conventional mixtures of only positive Gaussians, thus yielding a better initial region map formed by pixel-wise classification of the image gray values.

### 3.2.2.4 Spatial Interaction Model

For a smoother level set evolution and more accurate segmentation, spatially homogeneous interactions between the region labels are modeled with the popular Potts model (the Markov-Gibbs random field with the nearest 8-pixel neighborhood) having bi-valued Gibbs potentials (see Fig. 3.3), depending only on whether the nearest pairs of labels are equal or not. Let,  $f_{\text{eq}}(\mathbf{m})$  denote the relative frequency of equal labels in the neighboring pixel pairs  $\{(x, y), (x + \xi, y + \eta) : (x, y), (x + \xi, y + \eta) \in \mathbf{R}^2; (\xi, \eta) \in \{\pm 1, 0\}, (0, \pm 1), (\pm 1, \pm 1)\}$ . The initial region map (obtained by the pixel-wise classification) results in the approximate analytical estimates of the potentials (see e.g., [42]):

$$V_{\text{eq}} = -V_{\text{ne}} = 2 \left( f_{\text{eq}}(\mathbf{m}) - \frac{1}{2} \right) \quad (3.2)$$

that allow for computing the pixel-wise probabilities  $p_{h:x,y}(m_{x,y})$  of labels,  $m_{x,y} = \lambda; \lambda \in \mathbf{L}$ , given their 8-neighbors, at each step of the contour evolution.

### 3.2.2.5 Stochastic Speed Function

Let  $\kappa$  be the mean contour curvature and let  $\vartheta(x, y)$  specify the evolution magnitude and direction at the contour point  $(x, y)$ . The stochastic speed function in (3.1) is defined as:

$$V(x, y) = \kappa \vartheta(x, y) \quad (3.3)$$

where,

$$\vartheta(x, y) = \begin{cases} -P_{1:x,y} & \text{if } P_{1:x,y} > P_{0:x,y} \\ P_{0:x,y} & \text{Otherwise} \end{cases} \quad (3.4)$$

where,

$$\begin{aligned} P_{1:x,y} &= \frac{W_{1:x,y}}{W_{1:x,y} + W_{0:x,y}}; \quad P_{0:x,y} = 1 - P_{1:x,y} \\ W_{1:x,y} &= p(g_{x,y}|1)p_{h:x,y}(1)p_{s:x,y}(1) \\ W_{0:x,y} &= p(g_{x,y}|0)(1 - p_{h:x,y}(1))(1 - p_{s:x,y}(1)) \end{aligned}$$

$p_{h:x,y}(1)$  is the probability of transition to the object label at the pixel  $(x, y)$  of the region map  $\mathbf{m}$  in the Potts model  $P(\mathbf{m})$  at the current evolution step, and  $p_{s:x,y}(1)$  is the pixel-wise object probabilities. Note that for the inner contour segmentation  $p_{s:x,y}(1|t)$  will be used instead of  $p_{s:x,y}(1)$ .

For inner contour segmentation, the initial distance map  $\phi_o(x, y)$  is computed with respect to a circular contour initialized inside the heart cavity on the initial region map and then incremented or decremented at every pixel in line with (3.4). If a pixel  $(x, y)$  is classified as the object, the level set function is decremented and the contour expands; otherwise, it is incremented and the boundary contracts. The resulting segmentation is used to compute the initial distance map for the outer contour segmentation.

During the contour evolution, the level set function  $\phi(x, y)$  has to remain the signed distance function as suggested in [57]. Due to the small area of the cavity and the LV wall, we constrained the level set reinitialization based on the  $|\nabla\phi|$  to ensure a smooth and accurate segmentation. All the experimental results below were obtained with the level set function recomputed when  $||\nabla\phi| - 1| \leq 0.1$ . The steps of the proposed segmentation approach are summarized in Algorithm 1.

#### **Algorithm 1: Segmentation Algorithm**

1. Construct the shape prior.
2. Estimate the marginal intensity model  $P(\mathbf{g}|\mathbf{m})$  by identifying the bimodal LCDG.
3. Form of an initial region map  $\mathbf{m}$  by using the estimated marginal intensity model.
4. Identify the MGRF model from the initial map.
5. Estimate the stochastic speed function by using (3.4).
6. Evolve the level set using the above estimated speed.

#### **3.2.2.6 Performance Evaluation of the Proposed Segmentation Approach**

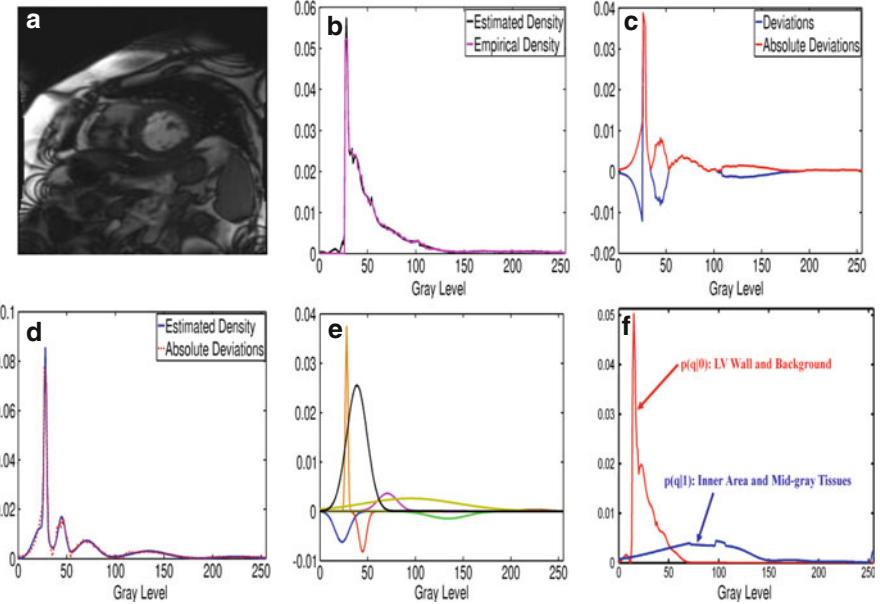
To evaluate the performance of the proposed segmentation approach, we measured the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) segmentations. Let  $|\mathbf{C}|$ ,  $|\mathbf{G}|$ , and  $|\mathbf{g}|$  denote areas (by the number of pixels) of the segmented region,  $\mathbf{C}$ , its ground truth (manually traced by an imaging expert),  $\mathbf{G}$ , and the whole image,  $\mathbf{g}$ , respectively. Then  $TP = |\mathbf{C} \cap \mathbf{G}|$  is the overlapping between  $\mathbf{C}$  and  $\mathbf{G}$ ;  $TN = |\mathbf{g} - \mathbf{C} \cup \mathbf{G}|$ ;  $FP = |\mathbf{C} - \mathbf{C} \cap \mathbf{G}|$ ; and  $FN = |\mathbf{G} - \mathbf{C} \cap \mathbf{G}|$ . Then the segmentation error is defined as:

$$e = \frac{1}{|\mathbf{g}|} (FP + FN) \quad (3.5)$$

Based on this estimated error, we calculated the error statistics (i.e., minimum, maximum, mean, and standard deviation) for all data sets (2,225 images) employed in this study.

#### **3.2.2.7 Segmentation Results**

The proposed segmentation approach has been tested on eight independent data sets of 2-D cine CMR images collected as time series (consisting each of 25 frames of

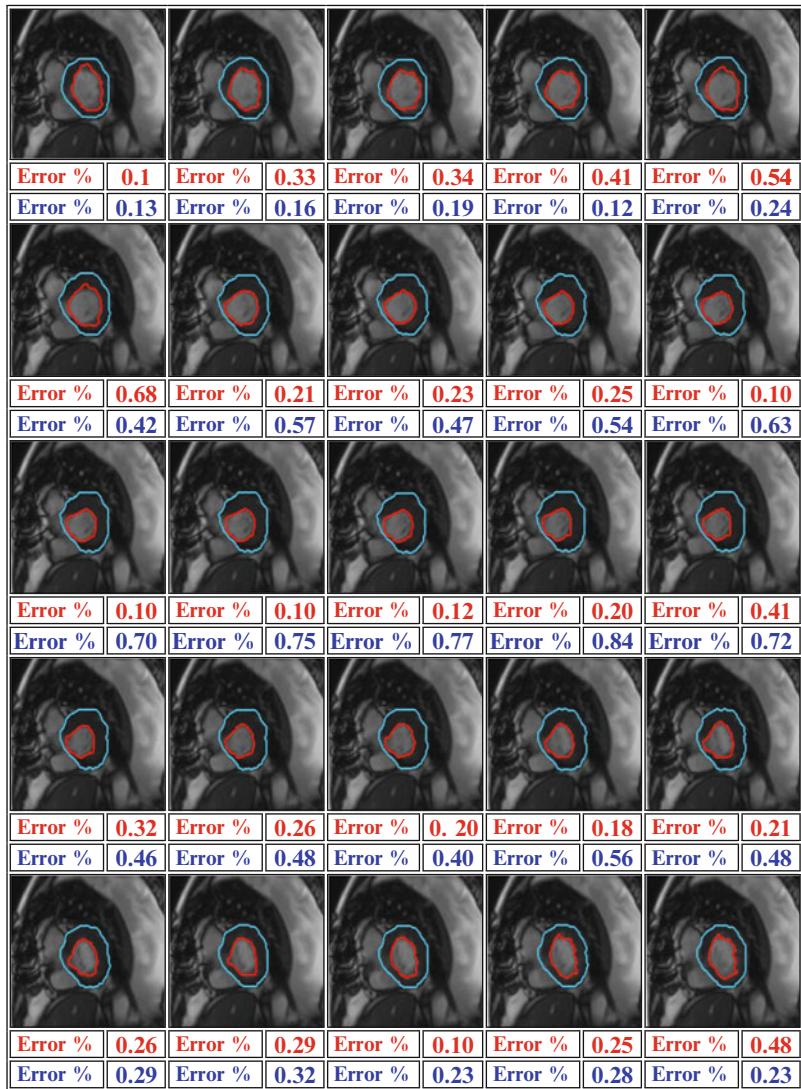


**Fig. 3.7** (a) Typical cine CMR image, (b) its gray level density estimation using the LCDG model [56], (c) deviations and absolute deviations between the estimated density and the empirical density, (d) estimated density of the absolute deviation, (e) LCDG components, and (f) final estimated marginal density for each class

the same image section). Figure 3.7 shows the LCDG models [56] of the marginal intensity distributions of a time series image that represent two dominant modes: for the inner area and mid-gray tissues, and for the LV wall and darker image parts. Segmentation results for one series of the test data sets employed in this study and their associated errors (evaluated by manual tracing of an imaging expert) are shown in Fig. 3.8. Note that the segmentation errors are calculated as described in Sect. 3.2.2.6.

More results of segmenting the LV wall borders on the test cine CMR images for different subjects together with the errors comparing them to the manual segmentation by an expert are shown in Fig. 3.9. Table 3.1 presents the segmentation results for all time series of one of the test data sets.

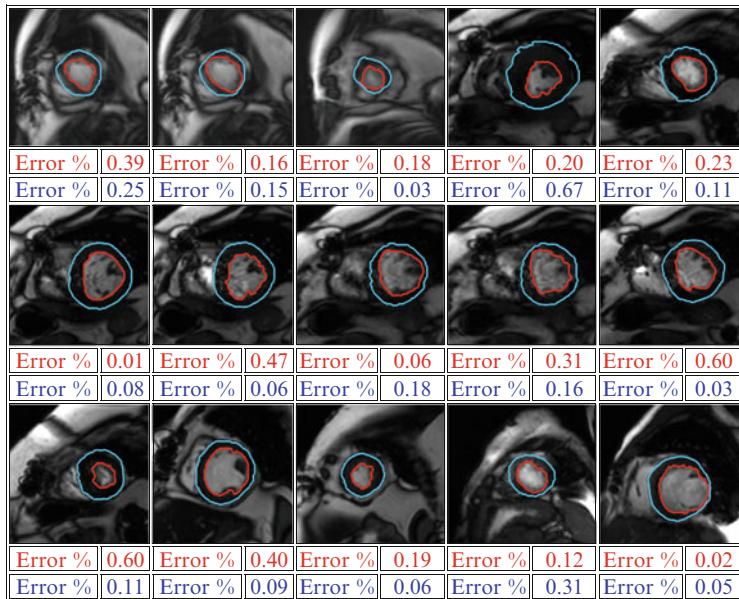
For comparison, all the time series images have been segmented using the shape-based approach proposed in [48]. Results for selected image sections for different subjects are shown in Fig. 3.10, and Table 3.2 compares the segmentation over all the test data sets with the ground truth obtained by manual tracing by an imaging expert. (Segmentation errors are calculated as described in Sect. 3.2.2.6). Differences between the mean errors for our approach and the shape-based approach in [48] are statistically significant by the unpaired *t*-test.



**Fig. 3.8** Segmentation errors for the inner (red) and outer (cyan) borders for one complete time series with respect to the manual tracing

### 3.3 Applications

Heart ventricular metrics used in clinical and research applications pursue the goal of determining quantitative global or local indexes (regional) that assess the entire field-of-view of interest and characterize the integrated physiology of a patient.



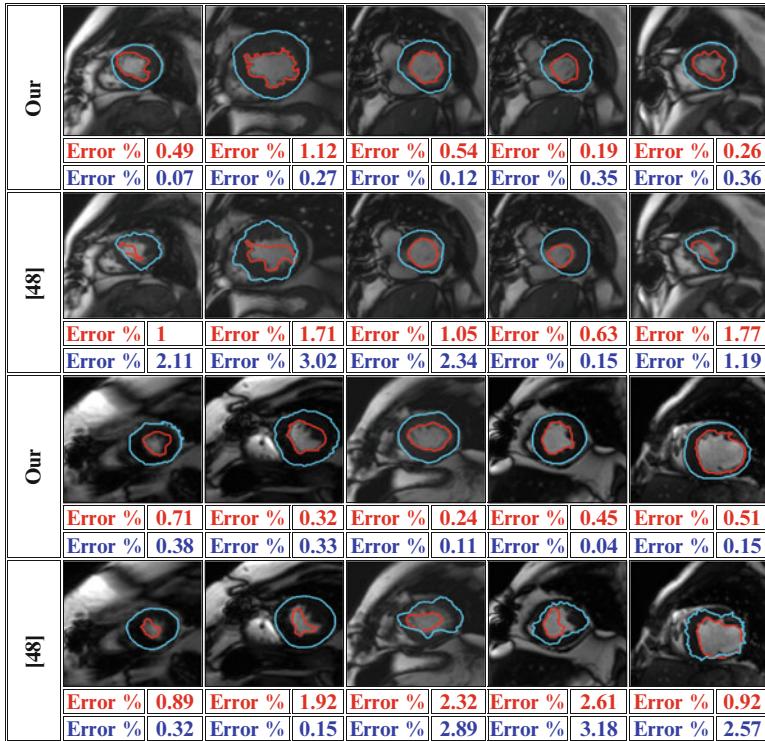
**Fig. 3.9** Additional segmentation results of cine CMR images

**Table 3.1** Segmentation errors for seven time series of one of the test data sets. (Each series consists of 25 image frames)

	Inner Contour				Outer Contour			
	Minimum	Maximum	Mean	Standard deviation	Minimum	Maximum	Mean	Standard Deviation
Series #1	0.0	0.3313	0.2788	0.0716	Series #1	0.0850	0.8878	0.4836 0.2459
Series #2	0.0	0.3841	0.3364	0.0799	Series #2	0.0302	1.1781	0.5897 0.2885
Series #3	0.0	0.3777	0.2263	0.1738	Series #3	0.3169	1.4152	0.6828 0.2712
Series #4	0.0	0.3662	0.3237	0.0763	Series #4	0.2747	0.9418	0.5879 0.2062
Series #5	0.0	0.3223	0.2561	0.0632	Series #5	0.1836	0.8327	0.5079 0.1703
Series #6	0.0	0.2646	0.2122	0.0523	Series #6	0.0852	0.9511	0.4807 0.2363
Series #7	0.0	0.1668	0.1229	0.0464	Series #7	0.0973	0.6717	0.4182 0.1351

### 3.3.1 Global Index

Characteristic of global indexes, we use total inner region (cavity) volume (time varying LV volume) data to estimate the ejection fraction. This ejection fraction is defined as the difference between the LV end-diastolic volume (the greatest inner area volume) and the end systolic volume (the smallest inner area volume) divided by the end-diastolic volume. The ejection fraction is a key clinical indicator of LV function – normal values of ejection fraction are between 0.55 and 0.75, while patients with significant ventricular dysfunction have ejection fractions from 0.15 to 0.4. To estimate the LV volume, we approximate the LV as a prolate spheroid



**Fig. 3.10** Accuracy of our segmentation in comparison with [48]

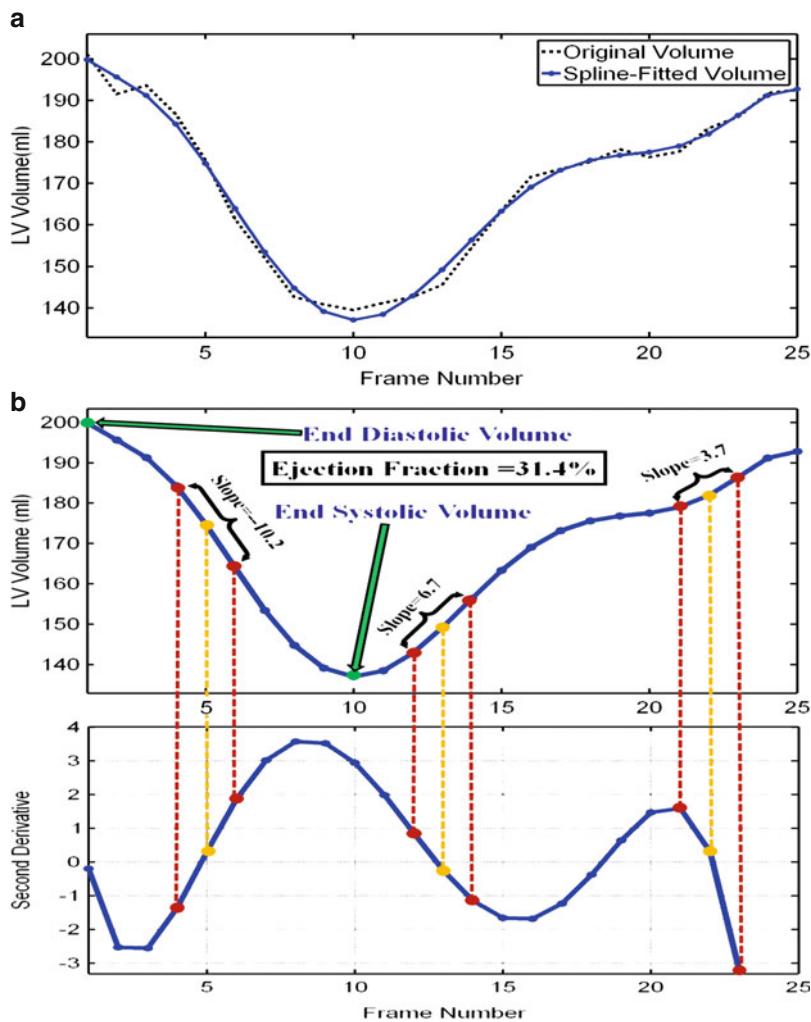
**Table 3.2** Accuracy of our segmentation approach over all test data sets (2225 image frames)

	Inner Contour		Outer Contour			
	Algorithm		Algorithm			
	Our	[48]	Our	[48]		
Minimum error %	0	0	Minimum error %	0.01	0	
Maximum error %	1.2503	3.0724	Maximum error %	1.4152	4.6532	
Average error %	0.5356	1.2615	Average error %	0.2919	1.8695	
Standard deviation %	0.3385	0.6687	Standard deviation %	0.2399	1.2033	
Two-tailed <i>P</i> -value	0.0001		Two-tailed <i>P</i> -value	0.0001		

(similar shape to the American football). The Simpson's rule is then used to estimate total LV volume by summing the contributions of enclosed areas from the individual image sections.

For a given image section, we delineate the inner contour to derive the cavity area (LV volume). From the time series (over the cardiac cycle for a given image section), we plot out a curve that depicts the physiology of cavity area changes over the cardiac cycle. There are four major physiological “events” of interest, namely time point corresponding to the greatest volume (end diastolic volume), time point

corresponding to smallest volume (end systolic volume), time point corresponding to most rapid emptying, and time point corresponding to most rapid filling. As the cardiac cycle is periodic, namely, it repeats over the period of a heartbeat, we can approximate this physiological or function curve using several methods; however, we chose a spline fit of order five (see Fig. 3.11a). This resulted in an adequate representation of key physiological details corresponding to the well-documented ventricular function curve, based on known cardiac physiology. From such a fitted curve, events such as smallest volume, and the greatest filling and emptying rates can be automatically be extracted using first and second derivatives of this curve

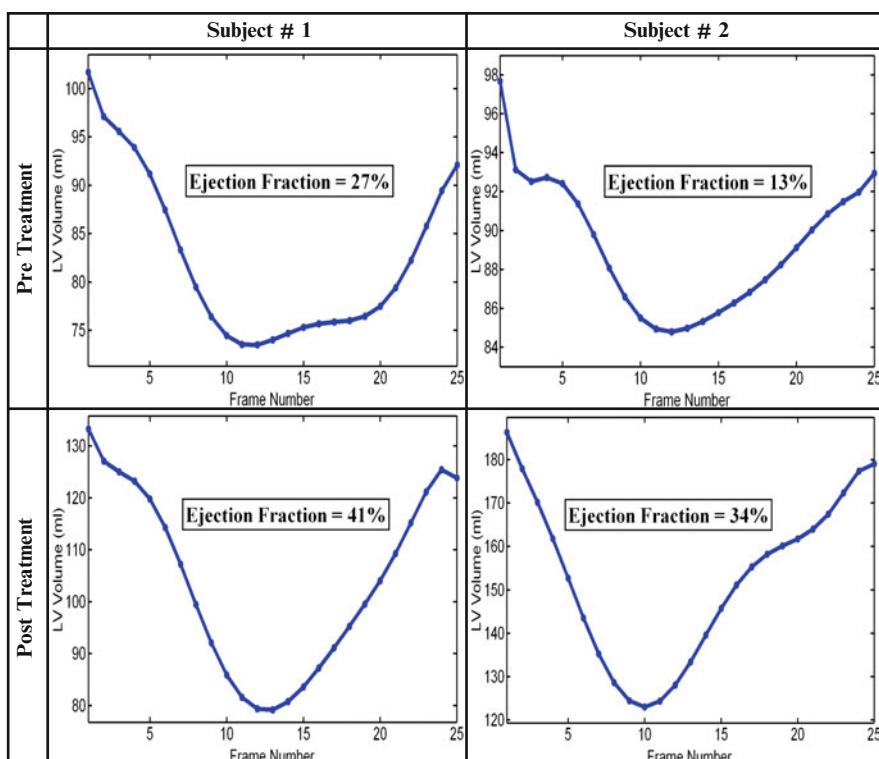


**Fig. 3.11** Ventricular function curve over one cardiac cycle: (a) curve fitting using fifth-order spline, and (b) estimation of the LV global indexes

over a cardiac cycle. An illustrative example of the depiction and estimation of the LV physiology is shown in Fig. 3.11b.

Specifically, in healthy individuals, most rapid filling rate is faster than most rapid emptying rate. For data taken sufficiently late into the filling portion of the cardiac cycle, termed diastole, there are two characteristic filling rates, the earlier being greater than the latter, in healthy individuals. Figure 3.12 depicts pre and posttreatment LV volume curves over one complete cardiac cycle for two subjects. Note that the generic “U” shape of the pretreatment (upper row in Fig. 3.12) curve is an indicative of decreased emptying and filling rates, with a preponderance of the effect evident on the filling portion of the curve. With treatment, the improved rates result in a “V” shape of the curves as shown in the bottom row in Fig. 3.12.

In our ongoing studies in heart attack patients, we are comparing clinical indicators of functional disability, namely how the patient functions in the daily activity, as indexed by measures such as New York Heart Association functional class (NYHA class I to class IV), to global functional indexes derived from our function curve, including the ejection fraction, the ratio of the peak filling rate to the peak emptying rate, and the ratio of the peak early filling rate to the peak late filling rate.



**Fig. 3.12** Pre- and posttreatment ventricular function curves for two subjects over one cardiac cycle

### 3.3.2 Local Index

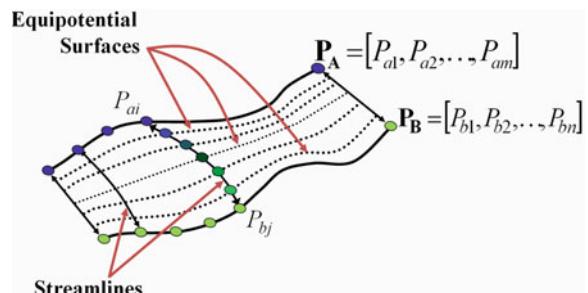
Representative of a local index, we examined the wall thickening, i.e., changes in wall thickness during the cardiac cycle. Wall thickening is a popular and well-documented local index showing, in particular, that patients with normal LV function have characteristically higher wall thickening than the patients with dilated cardiomyopathies. The wall thickness is measured with the Euclidean distance between two corresponding points on the inner and outer borders. In this chapter, the correspondences, or matches between the borders' points are found by solving, in the segmented wall, the Laplace equation.

Laplace's equation is a second-order linear Partial Differential Equation (PDE) for a scalar field  $\Psi$ . It arises in a variety of applications including fluid mechanics, electromagnetism, potential theory, solid mechanics, heat conduction, geometry, probability, etc. Specifically, Laplace-based approaches have been previously used in medical imaging applications such as colon surface flattening and centerline extraction [58, 59], thickness measurements [34, 60, 61], and nonrigid registration for kidney local motion correction [62]. Mathematically, the planer Laplace equation takes the form:

$$\nabla^2 \Psi = \frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} = 0 \quad (3.6)$$

A real-valued function,  $\Psi(x, y)$ , that satisfies (3.6) is known as harmonic function. The solution  $\Psi$  of (3.6) can be obtained by knowing the values of  $\Psi$  on all boundaries (Dirichlet boundary conditions). Solving Laplace's equation with Dirichlet boundary conditions between the enclosed boundaries (e.g.,  $\mathbf{P}_A$  and  $\mathbf{P}_B$  in Fig. 3.13) will induce point-to-point correspondences between both boundaries through the field lines (streamlines).

Generally, the solution of a planer Laplace equation between two boundaries results in intermediate equipotential surfaces (dashed lines in Fig. 3.13) and streamlines that connect both boundaries. These streamlines are defined as being everywhere orthogonal to all equipotential surfaces (e.g., the line connecting the points  $P_{ai}$  and  $P_{bj}$  in Fig. 3.13) and are used to find the point-to-point correspondences



**Fig. 3.13** 2-D correspondences by a potential field

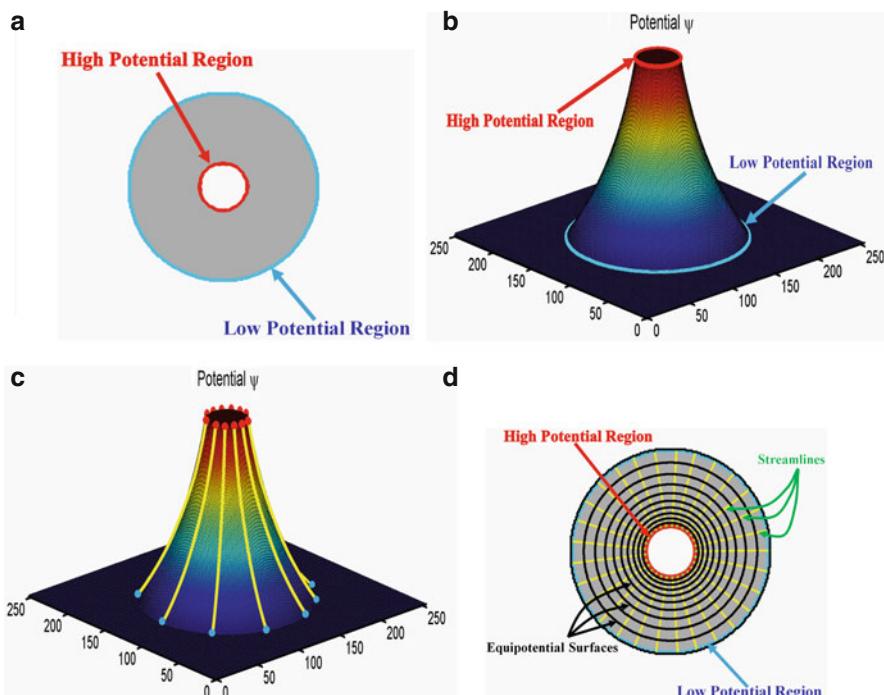
between both boundaries. The steps of the Laplace-potential-based wall thickness estimation are summarized in Algorithm 2.

### Algorithm 2: Wall Thickness Estimation

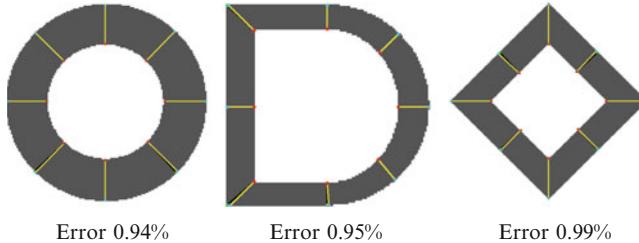
1. Initial condition: the potential  $\Psi$  is set maximum at the inner contour and minimum (equal zero) at the outer contour.
2. Solve (3.6) numerically in the segmented wall using the above initial condition.
3. Estimate the electric field vectors  $E_x = \partial\Psi/\partial x$  and  $E_y = \partial\Psi/\partial y$ .
4. Generate the streamlines between the inner and outer borders to find the point-to-point correspondence, then calculate the Euclidian distance between each pair of points.

#### 3.3.2.1 Validation of the PDE-Based Approach for the Estimation of the Point-to-Point Correspondence

To demonstrate the ability of the solution of the Laplace's equation to collocate the corresponding pixel pairs between two boundaries, we use a synthetic phantom (Fig. 3.14) of two concentric circles representing inner (red) and outer (cyan) borders.



**Fig. 3.14** Laplace-potential-based matching: inner (red) and the outer (cyan) borders of a synthetic phantom (a); 3-D potential surface  $\Psi(x, y)$  between the borders (b); induced streamlines (yellow) between the corresponding points (c), and their 2-D projection (d)



**Fig. 3.15** Validation of the Laplace-potential-based wall thickness estimates: ground truth (black lines) vs. estimated (yellow lines) point-to-point correspondences

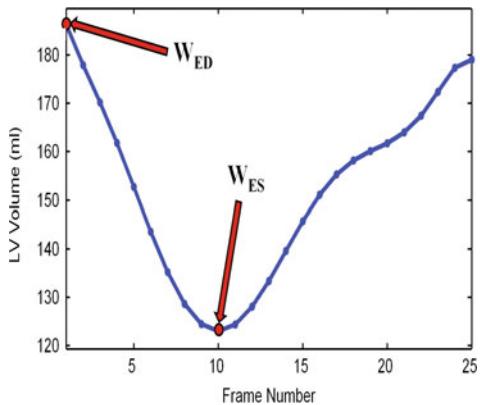
myocardial borders. The potential  $\Psi$  is found in the interior area enclosed by the borders by solving (3.6).  $\Psi$  is set to maximum – high potential – at the inner border and minimum – zero potential – at the outer border. After the potential field  $\Psi$  between the borders is found, its gradient vectors induce the streamlines linking the corresponding border points. To evaluate the performance, this approach was tested on three synthetic 2-D phantoms of different shape (Fig. 3.15) that mimic enclosed areas between two boundaries as is found with the inner and outer myocardial borders in cine CMR images. For each inner contour point, we need to estimate the error corresponding to the distance between its outer “ground-truth” match and the point found by solving the Laplace equation. This was normalized with respect to the estimated local thickness, i.e., the distance between this inner contour point and the corresponding outer point. The relative errors, i.e., differences between the found and true distances, is close to 1.0% in the average over all the inner contour points in each phantom.

### 3.3.2.2 Physiological Parameter (Wall Thickening, $\delta$ ) Estimation

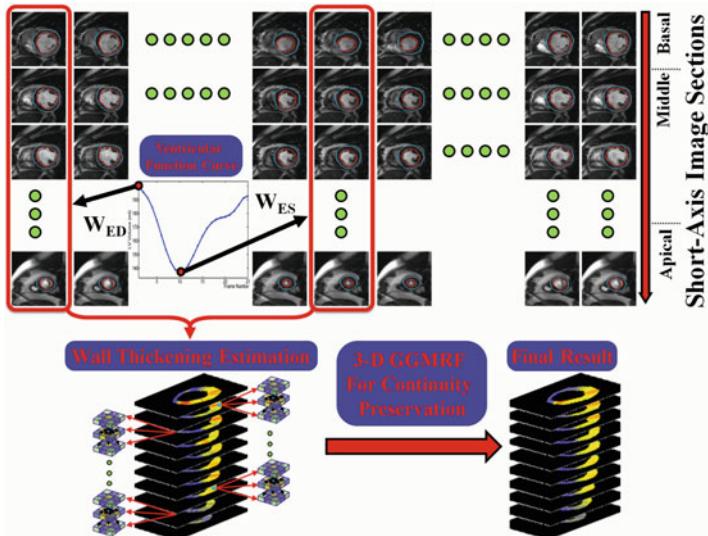
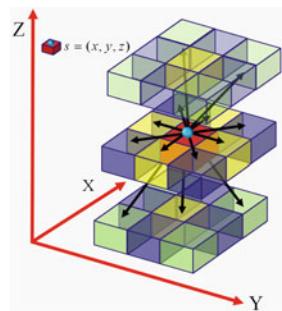
For each image section  $k$ , we calculated the wall thickness changes,  $\delta_k = W_{\text{ES}} - W_{\text{ED}}$  between the wall thickness values  $W_{\text{ES}}$  and  $W_{\text{ED}}$  (Fig. 3.16) corresponding to the end systolic volume (the smallest inner area volume) and end diastolic volume (the greatest inner area volume), respectively. For continuity preservation and inconsistencies removal (i.e., smoothed  $\delta$ ), the estimated wall thickening values are modeled with a generalized 3-D Gauss-Markov random field (GGMRF) proposed in [63] with the 26-neighborhood voxel shown in Fig. 3.17. For each patient, the estimated wall thickening values in all image sections were used to build a 3-D model of the LV (see Fig. 3.18). The continuity of the wall thickening values in the constructed 3-D volume is preserved by estimating its Maximum A Posteriori (MAP) estimate as shown in [63]:

$$\hat{\delta}_s = \arg \min_{\tilde{\delta}_s} \left\{ |\delta_s - \tilde{\delta}_s|^\alpha + \sigma^\alpha \lambda^\beta \sum_{r \in v_s} \eta_{s,r} |\tilde{\delta}_s - \delta_r|^\beta \right\} \quad (3.7)$$

**Fig. 3.16** Ventricular function curve showing time points within the cardiac cycle used to estimate wall thickening  $\delta$



**Fig. 3.17** Characteristic pairwise voxel interaction in a 3-D GGMRF image model



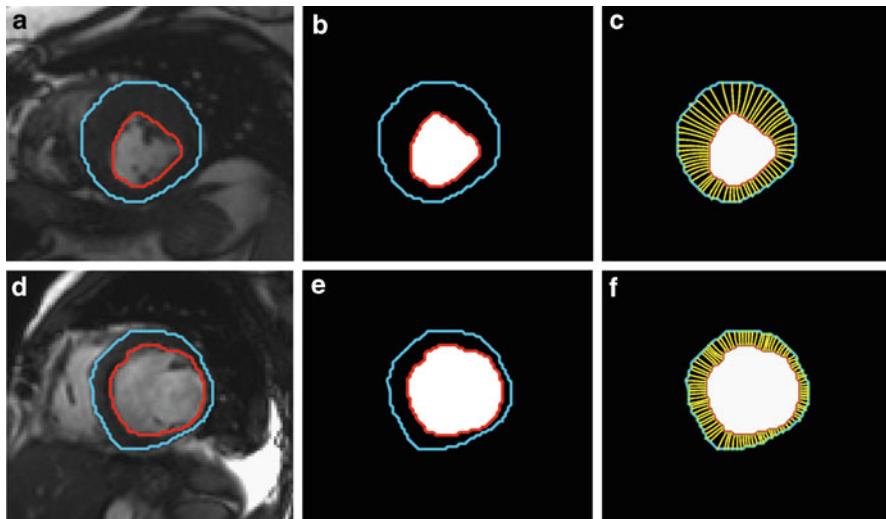
**Fig. 3.18** Wall thickening estimation and continuity analysis using the 3-D GGMRF model [63]

where,  $\delta_s$  and  $\tilde{\delta}_s$  are the changes in the wall thickness values at the observed 3-D location  $s = (x, y, z)$  and its neighbors, respectively, from the 26-neighborhood voxel set  $v_s$  (Fig. 3.18);  $\eta_{s,r}$  is the GGMRF potential, and  $\sigma$  and  $\lambda$  are scaling factors. The parameter  $\beta \in [1.01, 2]$  controls the smoothing (e.g., smooth,  $\beta = 2$  vs. relatively abrupt edges,  $\beta = 1.01$ ). The parameter  $\beta \in [1, 2]$  determines the Gaussian,  $\alpha = 2$ , or Laplace,  $\alpha = 1$ , prior distribution of the MAP estimator. In our experiments below, the results are obtained using the following values:  $\sigma = 1$ ,  $\lambda = 5$ ,  $\beta = 1.01$ ,  $\alpha = 2$ , and  $\eta_{s,r} = \sqrt{2}$ .

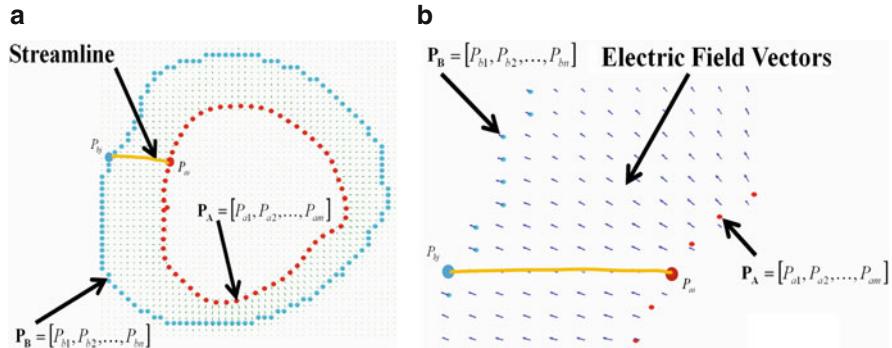
### 3.3.2.3 Wall Thickness Results

Figure 3.19 illustrates the process of determining the wall thickness from the point-to-point border correspondences on the segmented cine CMR images. Figure 3.20 is an illustration of the use of the induced streamlines that establish the point-to-point correspondence between inner and outer border of the segmented wall.

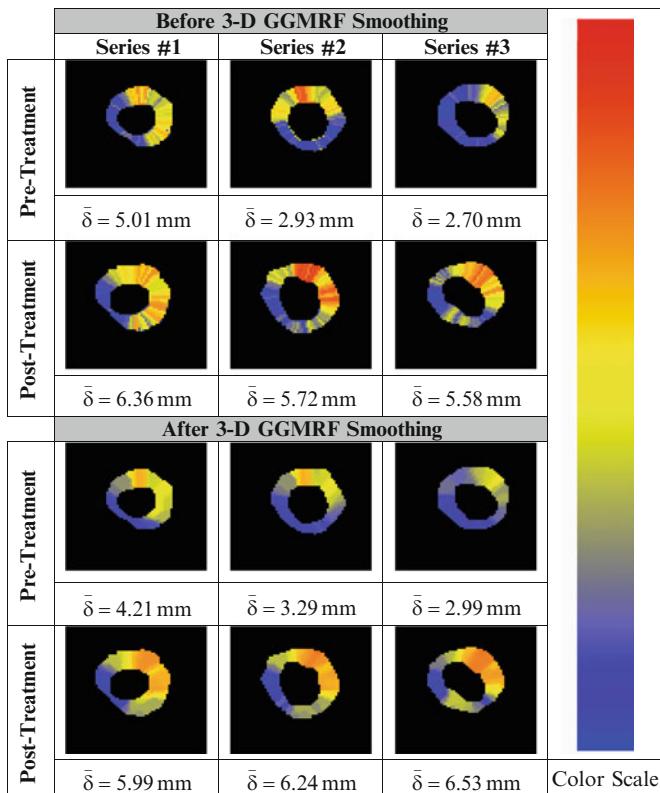
The wall thickening function is displayed as pixel-wise parametric color-coded maps that characterize the physiological data. To visualize the color map, the wall thickness changes,  $\delta$ , are normalized by relating them to the maximum wall thickness change measured in all the image sections for both the pre and posttreatments for each patient. Figure 3.21 presents parametric maps of the changes  $\delta$  before and after the 3-D GGMRF smoothing over multiple image sections covering the LV of the heart. Note that  $\bar{\delta}$  shown in Fig. 3.21 is the average wall thickening value calculated over the LV wall. The improvement of the wall thickening



**Fig. 3.19** Typical cine CMR images (**a, d**) with inner (in red) and outer (in cyan) borders, and (**c, f**) the streamlines obtained by the solution of Laplace's equation



**Fig. 3.20** (a) Electric field vectors between the inner ( $\mathbf{P}_A$ ; red) and outer ( $\mathbf{P}_B$ ; cyan) LV wall borders, and (b) an enlarged section around the indicated streamline indicating the vector field induced by the solution of Laplace's equation



**Fig. 3.21** Color-coded maps for the changes of wall thickness for a pre and posttreatment test data set before and after the 3-D GGMRF smoothing with  $\sigma = 1$ ,  $\lambda = 5$ ,  $\beta = 1.01$ ,  $\alpha = 2$ , and  $\eta_{s,r} = \sqrt{2}$ . The red and blue ends of the color scale relate to the maximum and minimum thickness, respectively

function  $\delta$  obtained by the proposed framework (see Fig. 3.21) can be correlated with other clinical indexes of the LV such as ejection fraction (see Fig. 3.12). There was a tentative agreement between the improvement of both wall thickening results and the ejection fraction after treatment.

### 3.4 Conclusions

Ventriculometry based on finding the inner and outer myocardial boundaries results in well-established and documented parameters such as the change in the wall thickness (wall thickening), end diastolic and end systolic volumes, and ejection fraction. The proposed efficient and robust framework determines these parameters automatically by accurate segmentation of myocardial inner and outer borders in cine CMR Images and accurate searching for corresponding points on the myocardial borders. In our initial experiments, this approach has been applied to the cine CMR images for patients undergoing restorative therapy after heart damage due to heart attacks. The tentative agreement between our measurements and other clinical indexes of the LV (e.g., ejection fraction) is very promising and indicates good potentials for future clinical applications of the proposed approach.

## References

1. Beache G, Wedeen V, Dinsmore R (1993) Magnetic resonance imaging evaluation of left ventricular dimensions and function and pericardial and myocardial disease. Coron Artery Dis 4(4):328–333
2. Liang J, Ding G, Wu Y (2008) Segmentation of the left ventricle from cardiac MR images based on radial GVF snake. In: Proceedings of the international conference on biomedical engineering and informatics, Sanya, China, 27–30 May 2008, vol 2, pp 238–242
3. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331
4. Xu C, Prince JL (1998) Snakes, shapes, and gradient vector flow. IEEE Trans Image Process 7(3):359–369
5. Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10(2):266–277
6. Mitchell S, Lelieveldt B, van der Geest R, Bosch H, Reiver J, Sonka M (2001) Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. IEEE Trans Med Imaging 20(5):415–423
7. Mitchell S, Bosch H, Lelieveldt B, van der Geest R, Reiver J, Sonka M (2002) 3-D active appearance models: segmentation of cardiac MR and ultrasound images. IEEE Trans Med Imaging 21(9):1167–1178
8. Lelieveldt BF, Üzümçü M, van der Geest RJ, Reiber JC, Sonka M (2003) Multi-view active appearance models for consistent segmentation of multiple standard views: application to long- and short-axis cardiac MR images. In: Proceedings of the 17th international congress and exhibition on computer assisted radiology and surgery (CARS 0'3), London, UK, vol 1256, pp 141–1146

9. Qian Z, Metaxas DN, Axel L (2005) A learning framework for the automatic and accurate segmentation of cardiac tagged MRI images. In: Liu Y, Jiang T, Zhang C (eds) Computer vision for biomedical image applications (CVIBA'01) Workshop, Beijing, China, 21 October 2005, LNCS 3765, pp 93–102, Springer-Verlag Berlin Heidelberg
10. Lin X, Cowan B, Young A. Automatic detection of left ventricle in 4D MR images: experience from a large study. In: Proceedings of the MICCAI'06, LNCS 4190, pp 728–735
11. Lötjönen J, Kivistö S, Koikkalainen J, Smutek D, Lauerma K (2004) Statistical shape model of atria, ventricles and epicardium from short- and long-axis MR images. *Med Image Anal* 8:371–386
12. Chen Q, Zhou ZM, Tang M, Heng PA, Xia D-S (2006) Shape statistics variational approach for the outer contour segmentation of left ventricle MR images. *IEEE Trans Inf Technol Biomed* 10(3):588–597
13. Taleb-Ahmed A, Leclerc X, Michel TS (2001) Semi-automatic segmentation of vessels by mathematical morphology: application in MRI. In: Proceedings of the international conference on image processing (ICIP'01), Palermo, Italy, 7–10 October, vol 3, pp 1063–1066
14. de Bruijne M, Nielsen M (2004) Shape particle filtering for image segmentation. Lecture Notes in computer science, medical image computing and computer-assisted intervention (MICCAI'04), Rennes, France, 26–30 September 2004, pp 168–175
15. Zhu Y, Papademetris X, Sinusas AJ, Duncan JS (2009) A dynamical shape prior for LV segmentation from RT3D echocardiography. Lecture notes in computer science, medical image computing and computer-assisted intervention (MICCAI'09), London, UK, 20–24 September, vol 1, pp 206–213
16. Li C, Jia X, Sun Y (2009) Improved semi-automated segmentation of cardiac CT and MR images. In: IEEE international symposium on biomedical imaging: from nano to micro (ISBI'09), Boston, MA, 28 June–1 July 2009, pp 25–28
17. Paragios N (2002) A variational approach for the segmentation of the left ventricle in MR cardiac images. *Int J Comput Vis* 50(3):345–362
18. Pluemptiwiriyawej C, Moura JMF, Wu Y-JL, Ho C (2005) STACS: new active contour scheme for cardiac MR image segmentation. *IEEE Trans Med Imaging* 24(5):593–603
19. Bernard O, Friboulet D, Thevenaz P, Unser M (2009) Variational B-spline level set: a linear filtering approach for fast deformable model evolution. *IEEE Trans Image Process* 18(6): 1179–1191
20. Xu C, Prince JL (2000) Gradient vector flow deformable models. In: Isaac Bankman (ed) Handbook of medical imaging, Academic: Orlando, FL, pp 159–169
21. Zhou S, Liang B, Chen W (2004) A new method for robust contour tracking in cardiac image sequences. In: IEEE international symposium on biomedical imaging: from nano to micro (ISBI'04), Arlington, VA, 15–18 April 2004, pp 15–18
22. Qin A, Feng Q, Chen W (2007) Cardiac MR image segmentation with modified active contour model. In: IEEE/ICME international conference on complex medical engineering, (CME'07), Beijing, China, 23–27 May 2007, pp 619–622
23. Chen T, Babb J, Kellman P, Axel L, Kim D (2008) Semiautomated segmentation of myocardial contours for fast strain analysis in cine displacement-encoded MRI. *IEEE Trans Med Imaging* 27(8):1084–1094
24. van der Geest RJ, de Roos A, van der Wall EE, Reiber JHC (1997) Quantitative analysis of cardiovascular MR images. *Int J Card Imaging* 13(3):247–258
25. Holman ER et al (1997) Detection and quantification of dysfunctional myocardium by magnetic resonance imaging: a new three-dimensional method for quantitative wall-thickening analysis. *Circulation* 95:924–931
26. Buller VGM, van der Geest RJ, Kool MD, Reiber JHC (1995) Accurate three-dimensional wall thickness measurement from multi-slice short-axis MR imaging. In: Proceedings of computers in cardiology, Vienna, Austria, 10–13 September, pp 245–248
27. Beohar N, Flaherty JD, Davidson CJ, Vidovich MI, Brodsky A, Lee DC (2007) Quantitative assessment of regional left ventricular function with cardiac MRI: three-dimensional center-surface method. *Catheter Cardiovasc Interv* 69:721–728

28. Saeed M, Wendland MF, Yu KK, Lauferma K, Li HT, Derugin N, Cavagna FM, Higgins CB (1994) Identification of myocardial perfusion with echo planar magnetic resonance imaging; discrimination between occlusive and reperfused infarctions. *Circulation* 90:1492–1501
29. Saeed M, Wendland MF, Szolar D, Sakuma H, Geschwind JF, Globits S, Derugin N, Higgins CB (1996) Quantification of extent of area at risk with fast contrast-enhanced magnetic resonance imaging in experimental coronary artery stenosis. *Am Heart J* 132(5):921–932
30. Maier SE, Fischer SE, McKinnon GC, Hess OM, Krayenbuehl HP, Boesiger P (1992) Evaluation of left ventricular segmental wall motion in hypertrophic cardiomyopathy with myocardial tagging. *Circulation* 86:1919–1928
31. Lima JA, Jeremy R, Guier W, Bouton S, Zerhouni EA, McVeigh E, Buchalter MB, Weisfeldt ML, Shapiro EP, Weiss JL (1993) Accurate systolic wall thickening by NMR imaging with tissue tagging: correlation with sonomicrometers in normal and ischemic myocardium. *J Am Coll Cardiol* 21:1741–1751
32. Dong SJ, MacGregor JH, Crawley AP, McVeigh E, Belenkie I, Smith ER, Tyberg JV, Beyar R (1994) Left ventricular wall thickness and regional systolic function in patient with hypertrophic cardiomyopathy three dimensional tagged magnetic resonance imaging study. *Circulation* 90:1200–1209
33. Sheehan FH, Bolson EL, Dodge HT, Mathey DG, Schofer J, Woo HW (1986) Advantages and applications of the centerline method for characterizing regional ventricular function. *Circulation* 74(2):293–305
34. Prasad M, Ramesh A, Kavanagh P, Gerlach J, Germano G, Berman DS, Slomka PJ (2009) Myocardial wall thickening from gated Magnetic Resonance images using Laplace's equation. In: Proceeding s of the SPIE, medical imaging 2009: Computer-aided diagnosis, Lake Buena Vista, FL, February 2009, vol 7260, 726021, pp 1–8
35. Miller S, Simonetti O, Carr J, Kramer U, Finn JP (2002) MR Imaging of the heart with steady-state precession: influence of spatial and temporal resolutions on left ventricular functional parameters. *Radiology* 223:263–269
36. Montagnat J, Delingette H (2005) 4D deformable models with temporal constraints: application to 4D cardiac image segmentation. *Med Image Anal* 9(1):87–100
37. El-Baz A, Gimel'farb G (2009) Robust image segmentation using learned priors. In: Proceedings of the international conference on computer vision (ICCV'09), Kyoto, Japan, 27 September–4 October, 2009, vol 6, pp 61–79
38. Han X, Xu C, Prince JL (2009) A topology preserving level set method for geometric deformable models. *IEEE Trans Pattern Anal Mach Intell* 25(6):61–79
39. Osher S, Sethian J (1988) Fronts propagating with curvature dependent speed: algorithms based on the Hamilton-Jacobi formulation. *J Comput Phys* 79:12–49
40. Osher S, Fedkiw R (2006) Level set methods and dynamic implicit surfaces (applied mathematical sciences), 153. Springer, New York
41. Carson C, Belongie S, Greenspan H, Blobworld JM (2002) Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans Pattern Anal Mach Intell* 24(8):1026–1038
42. Farag A, El-Baz A, Gimel'farb G (2006) Precise segmentation of multimodal images. *IEEE Trans Image Process* 15(4):952–968
43. Ma WY, Manjunath BS (2000) Edgeflow technique for boundary detection and image segmentation. *IEEE Trans Image Process* 9(8):1375–1388
44. Martin D, Fowlkes C, Malik J (2004) Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Trans Pattern Anal Mach Intell* 26(5):530–549
45. Tang J (2009) A multi-direction GVF snake for the segmentation of skin cancer images. *Pattern Recognit* 42(6):1172–1179
46. Ng HP, Foong KWC, Ong SH, Goh PS, Nowinski WL (2007) Medical image segmentation using feature-based GVF snake. In: Proceedings of the IEEE international conference of engineering in medicine and biology society (EMBS), Lyon, France, 22–26 August 2007, pp 800–803

47. Chen Y, Thiruvenkadam S, Tagare HD, Huang F, Wilson D, Geiser EA (2002) On the incorporation of shape priors into geometric active contours. In: IEEE workshop on variational and level set methods (VLSM'01), Vancouver, BC, Canada, 13 July 2002, pp 145–152
48. Tsai A, Yezzi A, Wells W, Tempany C, Tucker D, Fan A, Grimson E, Willsky A (2003) A shape-based approach to curve evolution for segmentation of medical imagery. *IEEE Trans Med Imaging* 22(2):137–154
49. Tsai A, Wells W, Tempany C, Grimson E, Willsky A (2004) Information in coupled multi-shape model for medical image segmentation. *Med Image Anal* 8(4):429–445
50. Cheng K, Gu L, Wu J, Li W, Xu J (2008) A novel level set based shape prior method for Liver segmentation from MRI Images. In: Proceedings of the international conference on information technology and applied biomedicine (ITAB'08), Shenzhen, 30–31 May 2008, pp 144–147
51. Leventon M, Grimson WL, Faugeras O (2000) Statistical shape influence in geodesic active contours. *IEEE Int Conf Comput Vis* 1:316–323
52. Shen D, Davatzikos C (2000) An adaptive-focus deformable model using statistical and geometric information. *IEEE Trans Pattern Anal Mach Intell* 22(8):906–913
53. Litvin A, Karl WC (2003) Level set-based segmentation using data driven shape prior on feature histograms. In: Proceedings of IEEE workshop on statistical signal process (SSP'03), St. Louis, MO, 28 September–1 October 2003, pp 166–169
54. Paragios N, Deriche R (1999) Unifying boundary and regionbased information for geodesic active tracking. In: Proceedings of IEEE international conference on computer vision and pattern Recognition, (CVPR'99), Fort Collins, CO, June 23–25, vol 2, pp 300–305
55. Viola PA, Wells WM (1997) Alignment by maximization of mutual information. *Int J Comput Vis* 24(2):137–154
56. El-Baz A, Gimel'farb G (2007) EM Based approximation of empirical distributions with linear combinations of discrete Gaussians. In: Proceedings of IEEE international conference on image processing (ICIP'07), San Antonio, TX, 16–19 September 2007, vol 4, pp 373–376
57. Sussman M, Smereka P, Osher S (1994) A level set approach for computing solutions to incompressible two-phase flow. *J Comput Phys* 114:146159
58. Haker S, Angenent S, Tannenbaum A, Kikinis R (2000) Nondistorting flattening maps and the 3D visualization of colon CT images. *IEEE Trans Med Imaging* 19(7):665–670
59. Nain D, Haker S, Eric W, Grimson L, Cosman E Jr, Wells WW, Ji H, Kikinis R, Westin CF (2002) Intra-patient prone to supine colon registration for synchronized virtual colonoscopy. Lecture Notes in computer science, medical image computing and computer-assisted intervention (MICCAI'02)
60. Jones S, Buchbinder B, Aharon I (2000) Three-dimensional mapping of cortical thickness using Laplace's equation. *Hum Brain Mapp* 11:12–32
61. Yezzi A, Prince JL (2003) A PDE approach for measuring tissue thickness. In: Proceedings of IEEE computer society conference on computer vision and pattern recognition (CVPR'01), Kauai, Hawaii, 8–14 December 2001, vol 1, pp 87–92
62. Khalifa F, El-Baz A, Gimel'farb G, Abu El-Ghar M (2010) Non-Invasive image-based approach for early detection of acute renal rejection. In: Jiang T et al. (eds) Medical image computing and computer-assisted intervention (MICCAI'10), Part I, LNCS 6361, Springer-Verlag, Berlin/Heidelberg, pp 10–18
63. Bouman C, Sauer K (1993) A Generalized Gaussian image model for edge preserving MAP estimation. *IEEE Trans Image Process* 2(3):296–310

## Biography



**Fahmi Khalifa** received his B.Sc. and M.S. degrees in Electrical Engineering from Mansoura University, Mansoura, Egypt, in 2003 and 2007, respectively. In May 2009 he joined the BioImaging Laboratory at University of Louisville, Louisville, KY, USA, as a research assistant. His current research includes developing new *Computer Aided Diagnosis* systems, image modeling, shape analysis, and 2-D and 3-D image segmentation, and registration with main focus on *Automatic Diagnosis of Lung Cancer using Contrast Enhancement Computed Tomography Images*.



**Garth M. Beaché** is an Associate Professor of radiology, at the University of Louisville School of Medicine, Louisville, KY. He completed a fellowship in magnetic resonance research at Massachusetts General Hospital, Harvard Medical School, and has served on the faculty of Johns Hopkins School of Medicine. His research interests include advanced radiological imaging methods, including registration, and functional visualization methods.



**Georgy Gimel'farb** graduated from Kiev Polytechnic Institute, Kiev, Ukraine, and received the Ph.D. degree in engineering cybernetics from the Institute of Cybernetics, Academy of Sciences of the Ukraine, and the D.Sc. (Eng) degree in control in engineering from the Higher Certifying Commission of the USSR, Moscow, Russia. He joined the University of Auckland, Auckland, New Zealand, in July 1997. His research is focused on image analysis, computer vision, and statistical pattern recognition.



**Guruprasad A. Giridharan** received his B.Tech in Chemical Engineering from the University of Madras, India in 1998, and M.S. and Ph.D. degrees in Chemical Engineering from the University of Utah, Salt Lake City, UT in 2002. He is currently an Assistant Professor of Bioengineering at the University of Louisville, Louisville, KY. His current research includes circulatory system modeling, development of physiologic control strategies for blood pumps, and biomedical device

development and testing. He has published over 80 peer reviewed manuscripts, conference proceedings, and abstracts.



**Ayman El-Baz** received the B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering department, University of Louisville, in August 2006. His current research is focused on developing new computer assisted diagnosis systems for different diseases and brain disorders.

# Chapter 4

## Medical Images Segmentation Using Learned Priors

Ayman El-Baz and Georgy Gimel'farb

**Abstract** Objects of specific shapes in an image are typically segmented with a deformable model being zero level of a geometric-level set function specifying sign-alternate shortest distances to the object boundary from each pixel. The goal shapes are approximated by a linear combination of such 2D distance maps built for mutually aligned images of given training objects. Unfortunately, the approximate shapes may deviate much from the training ones because the space of the distance maps is not closed with respect to linear operations and the map for zero level of a particular linear combination need not coincide with the latter. To avoid this drawback, we propose a parametric deformable model with the energy tending to approach the learned shapes. Instead of the level sets formalism, the target shapes are approximated directly with linear combinations of distance vectors describing positions of the mutually aligned training shapes with respect to their common centroid. Such a vector space is now closed with respect to the linear operations and it is of much smaller dimensionality than the 2D distance maps. Thus our shape model is easily simplified with the PCA, and shape-dependent energy terms guiding the boundary evolution obtain very simple analytic form. Prior knowledge of visual appearance of the object is represented by Gibbs energies of its gray levels. To accurately separate the object from its background, each current empirical marginal probability distribution of gray values within a deformable boundary is also approximated with an adaptive linear combination of discrete Gaussians. Both the shape/appearance priors and the current probabilistic appearance description control the boundary evolution, the appearance-dependent energy terms having also simple forms due to analytical estimates. Experiments with natural images confirm robustness, accuracy, and high speed of the proposed approach.

---

A. El-Baz (✉)

BioImaging Laboratory, Department of Bioengineering, University of Louisville, 423 Lutz Building, Louisville, KY 40292, USA  
e-mail: [aselba01@louisville.edu](mailto:aselba01@louisville.edu)

## 4.1 Introduction

Parametric and geometric (based on level set techniques) deformable models are widely used for image segmentation. However, in many applications, especially in medical image analysis, accurate segmentation with these models is a challenging problem due to noisy or low-contrast 2D/3D images with fuzzy boundaries between goal objects, e.g., anatomical structures; both objects of interest and multiple background objects even in small areas or volumes in a typical image; objects of almost the same shape but different visual appearance (grayscale or color pattern), and frequently discontinuous boundaries because of occlusions or almost similar visual appearance of the adjacent parts of objects of different shapes [1, 2]. This is why prior knowledge about the goal shape and/or visual appearance may help in its solution [2].

*Relationship to the prior works:* Traditional deformable models (active contours, or “snakes”) proposed, initially by Kass et al. [3], account for no prior shape constraints and search for strong signal discontinuities, or grayscale/color edges in an image. Further modifications, e.g., [4], focus on numerically stable model evolution to intricate object boundaries by using level set techniques [5]. But the evolution guided by signal edges and general continuity/curvature limitations encounters difficulties when the goal object is not clearly separated from background. More accurate and robust results, but at the expense of a considerably reduced speed, are obtained by evolving a deformable surface under more detailed limitations of the interior grayscale or color pattern [6]. Nonetheless, without the prior knowledge, these algorithms cannot meet the above challenges.

Today’s image analysis actively explores 2D/3D deformable models based on learning the prior shape and/or appearance constraints from a given training set of manually segmented images. Most of the known approaches get by with only the goal shapes. Leventon et al. [7] and Shen and Davatzikos [8] augment a level-set based energy function guiding the model evolution with special terms attracting the deformable model to more likely shapes specified with the principal component analysis (PCA) of the training set. Chen et al. [9] use a prior “average shape” in the geometric active contour, while Cootes et al. [10] develop a parametric point distribution model to describe the evolving curve. The latter model describe variations from the average shape with a linear combination of eigenvectors. The points are matched to strong image gradients by determining shape and pose parameters of the point distribution. Another variant of this approach is proposed by Pentland and Sclaroff [11].

Staib and Duncan [12] introduced a parametric point model based on an elliptic Fourier decomposition of landmark points, the model parameters ensuring the best match between the evolving curve and the points of high signal gradient. Chakraborty et al. [13] extend this approach to a hybrid model incorporating information about both the gradient and the region appearance homogeneity.

Pizer et al. [14] and Styner et al. [15] segment 3D medical images by coarse-to-fine deformation of a shape-based medial representation (“m-rep”). A geometric

deformable contour proposed by Tsai et al. [1] evolves as zero level of a level set function being the 2D map of the signed shortest distances between each pixel and the contour. The goal shapes are approximated by a linear combination of the training distance maps for a set of mutually aligned training images. To simplify the model, only a few top-rank principal components are included to the linear combination. Unfortunately, the approximate shapes may differ from the training ones because the distance maps space is not closed with respect to linear operations, so that the map for zero level of a particular linear combination need not coincide with the latter. A deformable model of Huang et al. [16] integrates region, shape, and interior signal features assuming an approximate region shape is a priori known and aligned with the image to initialize the model.

Typically, the total energy function in all these models is based on a simple predefined appearance model assuming significantly different means or variances of gray values in the object and background (see e.g., [1, 17]). More promising is to learn the appearance model, too, although the segmentation in this case is typically based on a pixel- or voxel-wise classification [18]. An alternative approach of Joshi [19] performs nonparametric warping of a goal surface to a deformable atlas. The atlas contours are then transferred to the goal volume. But due to no shape prior, the resulting boundaries need not approach the actual ones. Speed of the iterative warping of a whole image volume is also quite low.

To overcome these problems, some active contour models, e.g., the models of Paragios and Deriche [20] and Cootes and Taylor [21], assume the prior probabilistic appearance characteristics have to be learned in order to jointly model both the shape and appearance of a goal object using the PCA. These shape/appearance models have been successfully applied to segment complex 3D medical images, e.g., brain, heart, and articular cartilage MRI [22–24]. But because of establishing pixel- or voxel-wise correspondences between the model and the image, such a segmentation becomes extremely slow.

*Our approach* follows the latter ideas of involving both the prior shape and appearance knowledge. But instead of using the level set formalism that run into problems when combining the training 2D distance maps, we control the evolution of a simple parametric deformable model with energy functions that directly use both the learned shape and appearance data in order to approach the goal shapes.

Each goal shape is represented with a piecewise-linear boundary having a predefined number of control points. These latter are positioned along a number of equiangular rays given off the common center being the centroid of the control points. The boundary is described with a vector of square distances from the center to the circularly ordered control points. All the goal shapes are represented with a linear combination of the training distance vectors, the training boundaries being mutually aligned to have the common centroid and similar poses. Then the linear combination has the same centroid as its components, i.e., the training boundaries.

Given the same center and the same equiangular system of rays, the space of the distance vectors is closed with respect to linear operations. Also, its dimensionality is much smaller than of the space of the conventional 2D distance maps. Thus the

above shape representation is easily simplified using the PCA and excluding the noisy bottom-rank eigenvectors with small eigenvalues. An evolving boundary is analytically approximated with the closest shape prior. The evolution is controlled by simple energy terms attempting to decrease the approximation errors.

Grayscale patterns manifested as visual appearance of the training objects are roughly described by Gibbs energies for a spatially homogeneous central-symmetric Markov-Gibbs random field (MGRF) model. The model involves only first-order and central-symmetric second-order statistics of gray levels. Because Gibbs potentials are analytically estimated from the training set of images, the energy terms depending on the appearance prior are also simple. Simultaneously, the object and background patterns in an image to be segmented are specified with the marginal gray value probability distributions learned by approximating the mixed empirical gray value distribution for the image with a linear combination of discrete Gaussians (LCDG). The LCDG-models of the current object and background take part in the energy terms controlling the model evolution, too.

The paper is organized as follows. Section 4.2 describes the proposed shape priors and discusses briefly how the robust scale invariant feature transform (SIFT) by Lowe [25, 26] is used to affinely align the images to either build the priors or segment. Sections 4.3 and 4.4 introduce the MGRF-based appearance prior and the LCDG-models of a current image to be segmented, respectively. Section 4.5 presents with full details the evolution of the parametric model guided by the learned shape/appearance priors and the current image appearance. Several experimental results and conclusions are given in Sect. 4.6.

#### *Basic Notation:*

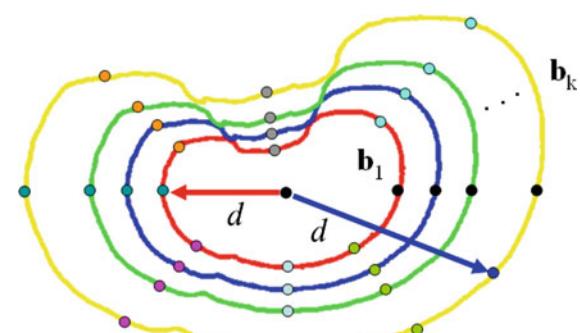
- $(x, y)$  – Cartesian coordinates of points in the image plane.
- $\mathbf{R} = [(x, y) : x = 0, \dots, X - 1; y = 0, \dots, Y - 1]$  – a finite arithmetic lattice supporting digital images and their region maps.
- $g = [g_{x,y} : (x, y) \in \mathbf{R}; g_{x,y} \in \mathbf{Q}]$  – a grayscale digital image taking gray values from a finite set  $\mathbf{Q} = \{0, \dots, Q - 1\}$ .
- $\mathbf{m} = [m_{x,y} : (x, y) \in \mathbf{R}; m_{x,y} \in \mathbf{L}]$  – a region map taking region labels from a finite set  $\mathbf{L} = \{\text{ob}, \text{bg}\}$ ; each label  $m_{x,y}$  indicates whether the pixel  $(x, y)$  in the corresponding image  $\mathbf{g}$  belongs to the goal object ( $m_{x,y} = \text{ob}$ ) or to the background ( $m_{x,y} = \text{bg}$ ).
- $\mathbf{b} = [\mathbf{p}_k : k = 1, \dots, K]$  – a deformable piecewise-linear boundary with the  $K$  control points  $\mathbf{p}_k = (x_k, y_k)$  forming a circularly connected chain of line segments  $(\mathbf{p}_1, \mathbf{p}_2), \dots, (\mathbf{p}_{K-1}, \mathbf{p}_K), (\mathbf{p}_K, \mathbf{p}_1)$ .
- $\mathbf{d} = [d_k^2 : k = 1, \dots, K]$  – a vector description of the boundary  $\mathbf{b}$  in terms of the square distances  $d_k^2 = (x_k - x_0)^2 + (y_k - y_0)^2$  from the control points to the model centroid  $p_0 = (x_0 = \frac{1}{K} \sum_{k=1}^K x_k, y_0 = \frac{1}{K} \sum_{k=1}^K y_k)$ , i.e., to the point at the minimum mean square distance from all the control points.
- $\mathbf{S} = \{(g_t, m_t, b_t, d_t) : t = 1, \dots, T\}$  – a training set of grayscale images of the goal objects with manually prepared region maps and boundary models.
- $|\mathbf{A}|$  – the cardinality of a finite set  $\mathbf{A}$ .

## 4.2 Shape Prior to Control a Parametric Model

Attractive and repulsive forces that cause a deformable model to evolve from an initial form to the goal object boundary have to account for the learned prior knowledge of the shape and visual appearance (grayscale or color pattern) of the goal object. To build a shape prior, all the training objects in  $S$  are mutually aligned to have the same centroid and unified orientations and scales of all the objects boundaries (see e.g., Fig. 4.1). For the definiteness, each training boundary  $\mathbf{b}_t \in S$  is described with the  $K$  control points positioned on the same polar system of  $K^\circ$  equiangular rays emitted from the common centroid  $\mathbf{p}_0$ . The rays are enumerated clockwise and have the angular pitch  $2\pi/K^\circ$  and zero angle for the first position  $\mathbf{p}_{t,1}$  of each boundary. We assume the alignment centers and uniquely orients each boundary with respect to that polar system. Generally, there may be rays with no or more than one intersection of a particular boundary, so that the number of the control points  $K$  may differ from the number of the rays  $K^\circ$ . For simplicity,  $K = K^\circ$  below, although the model holds for the general case, too.

Because the training boundaries  $\mathbf{b}_t = \mathbf{b} \in S; t = 1, \dots, T$ , share the same centroid  $\mathbf{p}_0$ , any linear combination  $\mathbf{d} = \sum_{t=1}^T w_t \mathbf{d}_t$  of the training distance vectors defines a unique new boundary  $\mathbf{b}$  with the same centroid. Typically, shapes of the training objects are very similar, and their linear combinations should be simplified by the PCA to escape singularities when adjusting to a given boundary.

Let  $\mathbf{D} = [\mathbf{d}_1 \mathbf{d}_2 \cdots \mathbf{d}_T]$  and  $\mathbf{U} = \mathbf{D}\mathbf{D}^T$  denote the  $K \times T$  matrix with the training distance vectors as columns and the symmetric  $K \times K$  Gram matrix of sums of squares and pair products  $\sum_{t=1}^T d_{t,k} d_{t,k}; k, k' = 1, \dots, K$  of their components, respectively. The PCA of the matrix  $\mathbf{U}$  produces  $K$  eigenvectors  $[\mathbf{e}_i : i = 1, \dots, K]$  sorted by their eigenvalues  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K \geq 0$ . Due to identical or very similar training shapes, most of the bottom-rank eigenvalues are zero or very small, so that the corresponding “noise” eigenvectors are discarded. Only a few top-rank eigenvectors actually represent the training shapes, the top distance eigenvector  $\mathbf{e}_1$  corresponding to an “average” shape and several others determining its basic variability. For simplicity, we select the top-rank subset of the eigenvectors



**Fig. 4.1** Mutually aligned training boundaries

$(\mathbf{e}_i : i = 1, \dots, K')$ ;  $K' < K$  by thresholding:  $\sum_{i=1}^{K'} \lambda_i \approx \theta \sum_{i=1}^K \lambda_i$  with an empirical threshold  $\theta = 0.8, \dots, 0.9$ .

Let an arbitrary boundary  $\mathbf{b}_c$  be aligned with the training set to get the vector description  $\mathbf{d}_c$  with the square distances of its control points from the centroid along the same  $K$  equiangular rays. Then the prior shape closely approximating this boundary is specified by the following linear combination of the training distance vectors:

$$\mathbf{d}^* = \sum_{i=1}^{K'} w_i^* \mathbf{e}_i \equiv \sum_{i=1}^{K'} (\mathbf{e}_i^T \mathbf{d}_c) \mathbf{e}_i \quad (4.1)$$

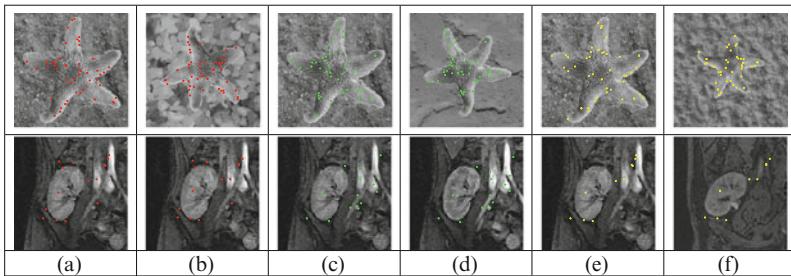
The signed differences  $\Delta_k = d_k^* - d_{c,k}$  determine the directions and forces to control the evolution of the boundary  $\mathbf{b}_c$  towards the closest shape prior  $\mathbf{b}^*$  specified by the distance vector  $\mathbf{d}^*$  in (4.1).

Just as the more conventional geometric (level-set based) deformable models with the shape priors, e.g., [1], our approach depends essentially on a proper mutual alignment of the objects of similar shapes at both the training and segmentation stages. In the latter case, to initialize the deformable model, an image  $\mathbf{g}$  to be segmented is to be aligned with one of the training images, say,  $\mathbf{g}_1$ , arbitrarily chosen in  $S$  as a reference image.

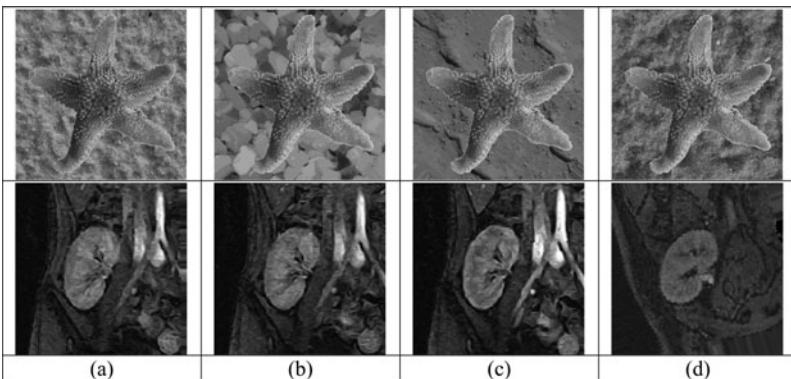
We establish first a number of point-wise correspondences between the two images assuming affine relative geometric distortions and limited contrast/offset signal distortions. Then an affine transform aligning  $\mathbf{g}$  most closely to  $\mathbf{g}_1$  is determined by minimizing the mean squared positional error between the corresponding points with the gradient descent method. To reliably determine image correspondences, we use the SIFT proposed by Lowe [25, 26]. Comparisons to several other approaches in [27] confirm its higher robustness with the affine geometric distortions present. SIFT determines the correspondences in the three main steps: (1) detection of interest points, (2) descriptor building, and (3) descriptor matching.

The interest points are located at the extrema in a difference-of-Gaussian (DOG) image pyramid. The points are localized with the subpixel accuracy to have more stable local features [28, 29]. The key elements of the SIFT descriptor are local gradient-orientation histograms in the neighbors of an interest point at the same pyramidal level. The orientations are relative to a canonical orientation of the interest point to make the descriptor invariant to the global object orientation. The translation invariance follows directly from the locality of the features, and the DOG pyramid decreases the scale changes of the descriptors.

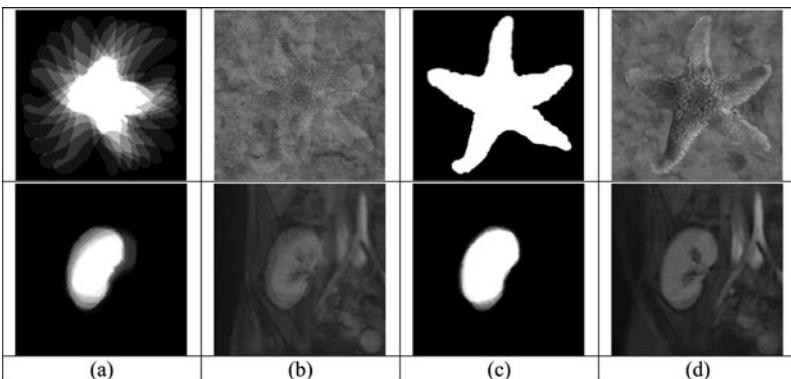
After the closest SIFT descriptors of the image  $\mathbf{g}$  and the reference  $\mathbf{g}_1$  are determined, the global 2D affine transform giving the best alignment of these images is estimated using the locations, scales, and canonical orientations of each corresponding pixel pair. Figure 4.2 shows the corresponding points obtained by SIFT for two types of natural images: digital images of starfishes and dynamic magnetic resonance images (DMRI) of human kidneys. The corresponding affinely



**Fig. 4.2** Corresponding points found by SIFT in each image pair **(a, b)**, **(c, d)**, and **(e, f)** (shown in *red*, *green*, and *yellow*, respectively)



**Fig. 4.3** Alignment of the images shown in Fig. 4.2a, b, d, and f



**Fig. 4.4** Comparison of the shape overlaps in the training data sets: **(a, c)** region maps and **(b, d)** grayscale images before **(a, b)** and after **(c, d)** the alignment

aligned images in Fig. 4.3 have roughly the same center, similar orientation, and approximately equal sizes. Figure 4.4 evaluates quality of the SIFT-based affine alignment with the region maps and grayscale images being the pixel-wise averages

of all the training maps  $\mathbf{m}_t$  and images  $\mathbf{g}_t; t = 1, \dots, T$ , respectively, before and after the mutual alignment of the training set  $\mathbf{S}$ . The similar shapes are overlapped more significantly after the alignment.

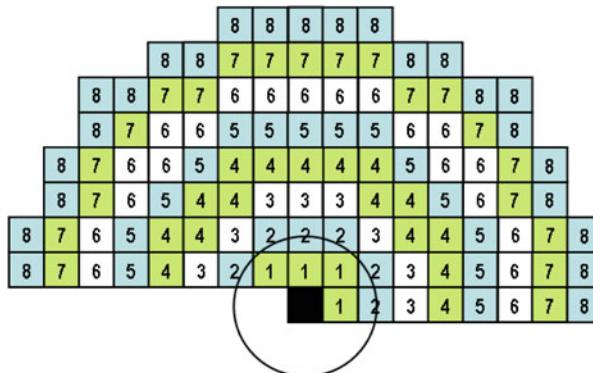
### 4.3 MGRF Image Model as an Appearance Prior

Although mutually aligned training images make it possible, in principle, to construct the appearance priors, typical complex grayscale or color patterns of the goal objects force to use only rough image descriptors, e.g., based on first- and second-order grayscale statistics. The latter implicitly “homogenize” each image, i.e., substitute certain spatially homogeneous patterns with the same statistics for the actual object. Nonetheless, even such limited prior knowledge is useful in many practical cases to segment the initial complex images.

We describe the aligned training objects as samples of a generic piecewise homogeneous Markov–Gibbs random field (MGRF) with pixel-wise and central-symmetric pair-wise pixel interactions. Let  $\mathbf{N} = \{1, 2, 3, \dots\}$ ,  $\mathbf{n}_v$  and  $\mathbf{V}$  denote an index set for successive semi-open intervals of characteristic interpixel distances  $(d_{v,\min}, d_{v,\max}]$  specifying central-symmetric pixel neighborhoods, the latter neighborhood as a set of  $(x, y)$ -coordinate offsets between each pixel  $(x, y)$  and its neighbor  $(x', y')$  such that  $d_{v,\min} < \sqrt{(x - x')^2 + (y - y')^2} \leq d_{v,\max}$ , and a set of Gibbs potential functions of gray values in individual pixels and gray value co-occurrences in the neighboring pixel pairs, respectively:

$$\mathbf{V} = [\mathbf{V}_{\text{pix}}; \{\mathbf{V}_v : v \in \mathbf{N}\}] \equiv [\{\mathbf{V}_{\text{pix}}(q) : q \in \mathbf{Q}\}; \{V_v(q, q') : v \in \mathbf{N}; (q, q') \in \mathbf{Q}^2\}]$$

The neighborhood sets  $n^\circ$  corresponding to the uniform distance ranges  $(v - 0.5, v + 0.5); v = 1, \dots, 8$ .



**Fig. 4.5** Central-symmetric neighborhood sets for the eight distance ranges  $(v - 0.5, v + 0.5); v = 1, \dots, 8$

Let  $\mathbf{R}_t = \{(x, y) : (x, y) \in \mathbf{R} \wedge m_{t,x,y} = \text{ob}\}$  and  $\mathbf{C}_{v,t}$  denote the part of the lattice  $\mathbf{R}$  supporting the training object in the image-map pair  $(\mathbf{g}_t, \mathbf{m}_t) \in \mathbf{S}$  and the family of all pixel pairs in  $\mathbf{R}_t^2$  such that each pair has the coordinate offset  $(\zeta, \eta) \in \mathbf{n}_v$ , respectively. Let  $\mathbf{F}_{\text{pix},t}$  and  $\mathbf{F}_{v,t}$  denote joint empirical probability distributions of gray values and gray value co-occurrences in the training object from the image  $\mathbf{g}_t$ , respectively:

$$\mathbf{F}_{\text{pix},t} = \left[ f_{\text{pix},t}(q) = \frac{|\mathbf{R}_{t,q}|}{|\mathbf{R}_t|}; \sum_{q \in \mathbf{Q}} f_{\text{pix},t}(q) = 1 \right]$$

$$\mathbf{F}_{v,t} = \left[ f_{v,t}(q, q') = \frac{|\mathbf{C}_{v,t;q,q'}|}{|\mathbf{C}_{v,t}|}; \sum_{(q, q') \in \mathbf{Q}^2} f_{v,t}(q, q') = 1 \right]$$

where,  $\mathbf{R}_{t,q} = \{(x, y) : (x, y) \in \mathbf{R}_t \wedge g_{x,y} = q\}$  is a subset of pixels supporting the object's gray value  $q$  in the training image  $\mathbf{g}_t$  and  $\mathbf{C}_{v,t;q,q'}$  is a subfamily of all the pixel pairs  $\mathbf{c}_{\zeta,\eta}(x, y) = ((x, y), (x + \zeta, y + \eta)) \in \mathbf{R}_t^2$  supporting the object's gray value co-occurrence  $(q, q')$  in the same image, respectively.

The MGRF model of the  $t$ -th object is specified by the joint Gibbs probability distribution on the sublattice  $\mathbf{R}_t$ :

$$P_t = \frac{1}{Z_t} \exp \left( \sum_{(x,y) \in \mathbf{R}_t} \left( V_{\text{pix}}(g_{x,y}) + \sum_{v \in \mathbf{N}} \sum_{\mathbf{c}_{\zeta,\eta}(x,y) \in \mathbf{C}_{v,t}} V_v(g_{x,y}, g_{x+\zeta,y+\eta}) \right) \right) \quad (4.2)$$

$$= \frac{1}{Z_t} \exp \left( |\mathbf{R}_t| \left( \mathbf{V}_{\text{pix}}^T \mathbf{F}_{\text{pix},t} + \sum_{v \in \mathbf{N}} \rho_{v,t} \mathbf{V}_{v,t}^T \mathbf{F}_{v,t} \right) \right)$$

where,  $\rho_{v,t} = |\mathbf{C}_{v,t}|/|\mathbf{R}_t|$  is an average cardinality of  $\mathbf{n}_v$  over the sublattice  $\mathbf{R}_t$ .

Because the areas and shapes of the aligned training objects are similar, then  $\mathbf{R}_t \approx R_{\text{ob}}$  and  $|\mathbf{C}_{v,t}| \approx C_{v,\text{ob}}$  for  $t = 1, \dots, T$ , where  $R_{\text{ob}}$  and  $C_{v,\text{ob}}$  are the average cardinalities over the training set  $\mathbf{S}$ . Assuming the independent samples, the joint probability distribution of gray values for all the training objects has the similar form

$$P_S = \frac{1}{Z} \exp \left( T R_{\text{ob}} \left( \mathbf{V}_{\text{pix}}^T \mathbf{F}_{\text{pix}} + \sum_{v \in \mathbf{N}} \rho_v \mathbf{V}_v^T \mathbf{F}_v \right) \right)$$

Here,  $\rho_v = C_{v,\text{ob}}/R_{\text{ob}}$ , and the empirical probability distributions  $\mathbf{F}_{\text{pix},\text{ob}}$  and  $\mathbf{F}_{v,\text{ob}}$  describe now the gray values and the gray value co-occurrences in all the objects from the training set. Zero empirical probabilities caused by a relatively small volume of the training data available to identify the above MGRF model are eliminated if fractions defining the empirical probabilities in terms of cardinalities of the related sublattices or subfamilies are modified as follows:  $(\langle \text{nominator} \rangle + \varepsilon)/(\langle \text{denominator} \rangle + S\varepsilon)$ . The

Bayesian quadratic loss estimate suggests the offset  $\varepsilon = 1$  and  $S = Q$  for the first-order or  $S = Q^2$  for the second-order interactions. More conservative approach in [30] suggests  $\varepsilon = 1/S$ . Using the same analytical approach as in [31], the Gibbs potentials are approximated with the scaled centered empirical probability distributions:

$$\begin{aligned} V_{\text{pix,ob}}(q) &= \lambda \left( f_{\text{pix,ob}}(q) - \frac{1}{Q} \right); \quad (q) \in Q; \\ V_{v,\text{ob}}(q, q') &= \lambda \left( f_{v,\text{ob}}(q, q') - \frac{1}{Q^2} \right); \quad (q, q') \in Q^2; \quad v \in N \end{aligned} \quad (4.3)$$

where, the common scaling factor  $\lambda$  is also computed analytically. It can be omitted ( $\lambda = 1$ ) when only relative potential values are involved, e.g., for relative energies  $E_{v,\text{rel}}$  of the central-symmetric pairwise pixel interactions in the training objects. These latter energies are equal to the variances of the co-occurrence distributions.

$$E_{v,\text{rel}} = \sum_{q,q' \in Q^2} f_{v,\text{ob}}(q, q') \left( f_{v,\text{ob}}(q, q') - \frac{1}{Q^2} \right)$$

and rank the neighborhoods  $\mathbf{n}_v$  for choosing most characteristic top-rank prior appearance descriptors  $N' \subset N$  specifying the central-symmetric neighborhoods of (4.3). Under these priors, a grayscale pattern within a deformable boundary  $\mathbf{b}$  in an image  $\mathbf{g}$  is described by its Gibbs energy

$$E(g, b) = V_{\text{pix,ob}}^T F_{\text{pix,ob}}(g, b) + \sum_{v \in N'} V_{v,\text{ob}}^T F_{v,\text{ob}}(g, b)$$

Here,  $N'$  is a selected top-rank index subset of the neighborhoods and the empirical probability distributions are collected within the boundary  $\mathbf{b}$  in  $\mathbf{g}$ .

#### 4.4 LCDG Probability Model of a Current Appearance

To account for a current image appearance in addition to the learned shape and appearance priors, the gray-level distribution within the evolving boundary  $\mathbf{b}$  in an image  $\mathbf{g}$  is described with a dynamic mixture of the object and background distributions. To identify the latter, it is closely approximated with a bimodal linear combination of discrete Gaussians (LCDG) and then split into the object and background LCDG-submodels. The approximation is performed with our EM-based techniques previously introduced in [32] and adapted in [33] to the DGs.

The DG  $\Psi_\theta = (\varphi(q|\theta) : q \in Q)$  integrates the continuous Gaussian density over the successive intervals related to the discrete gray values in  $Q$ , that is,

$\varphi(0|\theta) = \phi_\theta(0.5)$ ,  $\varphi(q|\theta) = \phi_\theta(q + 0.5) - \phi_\theta(q - 0.5)$  for  $q = 1, \dots, Q - 2$ , and  $\varphi(Q - 1|\theta) = 1 - \phi_\theta(Q - 1.5)$ . Where,  $\phi_\theta(q)$  is the cumulative Gaussian probability function with a shorthand notation  $\theta = (\mu, \sigma^2)$  for its mean,  $\mu$ , and variance,  $\sigma^2$ .

Each current appearance is represented with the bimodal LCDG having two dominant positive components that approximate the marginal empirical probability distribution of gray values within the evolving boundary and a number of subordinate positive and negative DGs describing deviations of the empirical distribution from the dominant mixture. Let  $C_p$  and  $C_n$  such that  $C_p \geq 2$  denote the total numbers of the positive and negative components. Then the LCDG model is as follows:

$$p_{w,\Theta}(q) = \sum_{r=1}^{C_p} w_{p,r} \varphi(q|\theta_{p,r}) - \sum_{l=1}^{C_n} w_{n,l} \varphi(q|\theta_{n,l}) \quad (4.4)$$

where the non-negative weights  $w = [w_p, w_n]$  meet the following obvious restriction:

$$\sum_{r=1}^{C_p} w_{p,r} - \sum_{l=1}^{C_n} w_{n,l} = 1$$

To identify the model of (4.4) the numbers  $C_p$  and  $C_n$  of its positive and negative components and the parameters  $w, \Theta$  (weights, means, and variances) of the positive and negative DGs are estimated first with a sequential EM-based initializing algorithm producing a close initial LCDG-approximation of the empirical distribution. Then under the fixed numbers  $C_p$  and  $C_n$ , all other parameters are refined with an modified EM algorithm that accounts for the alternating components. The final LCDG-model is partitioned into the two LCDG-submodels  $p_{\text{pix},l} = [p_{\text{pix},l}(q) : q \in Q]$ , one per class  $l \in L$ , by associating the subordinate DGs with the dominant terms such that the misclassification rate is minimal. The complete process is detailed in [32, 33].

## 4.5 Model Evolution

The evolution  $b_\tau \rightarrow b_{\tau+1}$  of the deformable boundary  $b$  in discrete time,  $\tau = 0, 1, \dots$ , is specified by the following difference equations system:

$$p_{k,\tau+1} = p_{k,\tau} + F(p_{k,\tau}) u_{k,\tau}; \quad k = 1, \dots, K \quad (4.5)$$

where,  $F(p)$  is a velocity function for the control point  $p$  and  $u_{k,\tau}$  is the unit vector along the  $k$ -th ray from the centroid to the control point  $p$  of the current boundary. Our velocity function depends on the shape prior, the current appearance evaluated

with the LCDG model, and the appearance prior based on the MGRF model as follows:

$$F(p = (x, y)) = \begin{cases} e^{-\beta\Delta} p_{\text{pix,ob}}(g_{x,y}) \pi_p(g_{x,y}|S) & \text{if } \Delta \geq 0 \\ e^{-\beta\Delta} p_{\text{pix,ob}}(g_{x,y}) \pi_p(g_{x,y}|S) & \text{if } \Delta \geq 0 \end{cases} \quad (4.6)$$

Here,  $\Delta$  is the signed distance between the current control point  $p \in b_\tau$  and the like one in the closest shape prior along the ray from the current boundary centroid. The constant factor determines the evolution speed ( $0 < \beta < 1$  for a smooth propagation). The marginal probabilities  $p_{\text{pix,ob}}(q)$  and  $p_{\text{pix,bg}}(q)$  of the gray value  $q$  are estimated with the LCDG-submodels for the object and its background, respectively, specified in Sect. 4.4. The prior conditional probability  $\pi_p(q|S)$  of the gray value  $q$  in the pixel  $p$ , given the current gray values in its particular central-symmetric neighborhood, is estimated in line with the MGRF model of the object appearance in Sect. 4.3:

$$\pi_p(g_{x,y}|S) = \frac{\exp(E_p(g_{x,y}|S))}{\sum_{q \in Q} \exp(E_p(q|S))}$$

where  $E_p(q|S)$  is the pixel-wise Gibbs energy for the gray value  $q$  in the pixel  $p$ , given the fixed gray values in its neighborhood:

$$E_p(q|S) = V_{\text{pix,ob}}(q) + \sum_{v \in N'} \sum_{(\zeta, \eta) \in n_v} (V_{v,\text{ob}}(g_{x-\zeta, y-\eta}, q) + V_{v,\text{ob}}(q, g_{x+\zeta, y+\eta}))$$

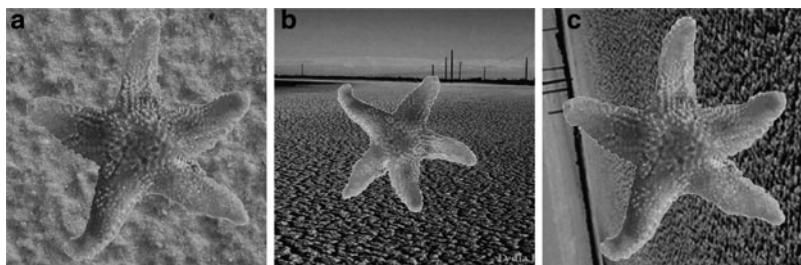
In total, the proposed segmentation algorithm is summarized as follows:

1. Initializing stage ( $\tau = 0$ ):
  - (a) Find the mutually matching pixel pairs in a given image  $g$  and one of the aligned training images  $g_{\text{tr}} \in S$  using the SIFT technique.
  - (b) Find the affine alignment of  $g$  to  $g_{\text{tr}}$  with the gradient descent technique.
  - (c) Initialize the deformable boundary with the training one  $b_{\text{tr}}$  for  $g_{\text{tr}}$ .
  - (d) Initialize the two dominant DGs representing the object and the background with the mean gray values for the arbitrarily selected pixels within and outside the deformable boundary, respectively.
  - (e) Use the modified EM algorithm in [32, 33] to estimate the LCDG-submodels  $p_{\text{pix,ob}}$  and  $p_{\text{pix,bg}}$  representing the current object and its background.
2. Evolution stage ( $\tau \leftarrow \tau + 1$ ):
  - (a) Evolve the deformable boundary in accord with (4.6).
  - (b) If the overall absolute deformation  $\sum_{k=1}^K |d_{k,\tau+1} - d_{k,\tau}| \leq \alpha$  (a small predefined threshold), terminate the process; otherwise return to Step 2a.
3. Segmentation stage: the final boundary is transferred to the initial (nonaligned) image  $g$  by the inverse affine transform.

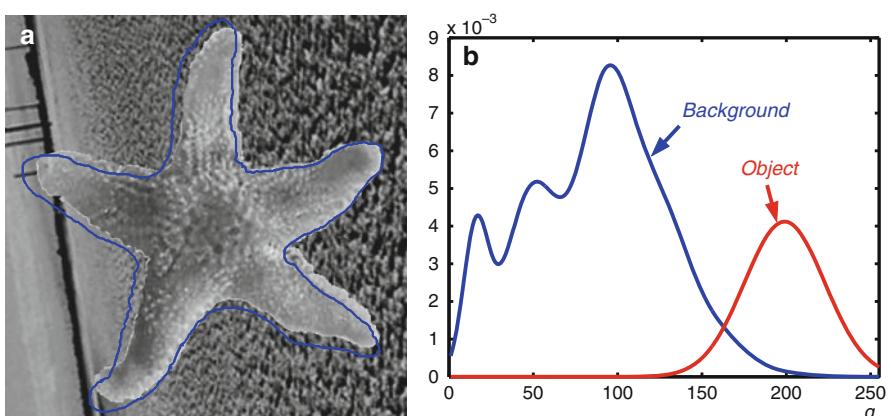
## 4.6 Experimental Results

Performance of the proposed parametric deformable model is evaluated below on two different types of images: the digitized photos of a starfish with the known ground truth (actual object boundaries) and the Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) of human kidneys. The latter images have been acquired with a Signa Horizon GE 1.5T scanner using a Gadolinium-DTPA contrast agent. After injecting the agent, the abdomen is scanned rapidly and repeatedly, the fast image acquisition having resulted in the high image noise. Moreover, the kidney contrast is continuously changing in the images in accord with perfusion of the contrast agent into the kidney and becomes very low at some perfusion stages. This necessitates the use of the shape priors for segmentation.

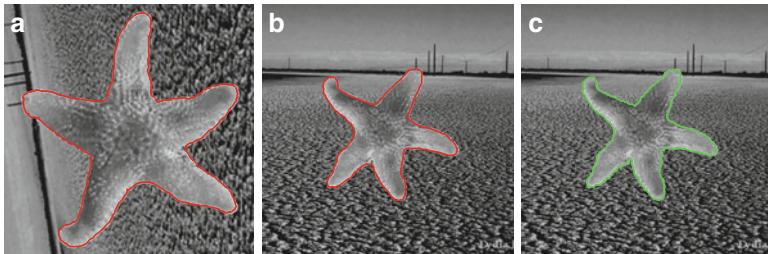
Basic stages of segmenting a starfish image are shown in Figs. 4.6–4.9. In particular, Fig. 4.6 demonstrates one of the aligned training images of a starfish, an image to be segmented, and the SIFT-based alignment of the latter to the former. Figure 4.7 illustrates the initialization of our deformable model with a training



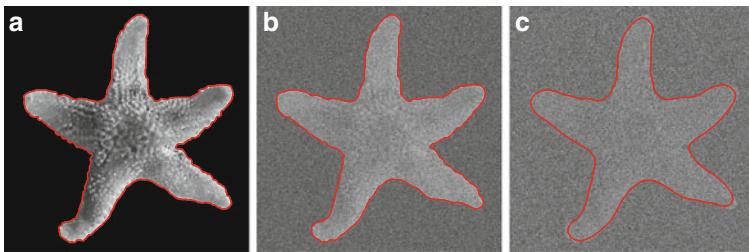
**Fig. 4.6** One of the mutually aligned training starfish images (a), an image to be segmented (b), and its alignment to the training set (c)



**Fig. 4.7** Initialization of the deformable model (a) shown in blue, and the LCDG-estimates of the marginal gray value distributions  $p_{pix,ob}$  and  $p_{pix,bg}$  (b)



**Fig. 4.8** Final segmentation of the image aligned to the training set (a) and the same result (b) after its inverse affine transform to the initial image in Fig. 4.6b in comparison to the ground truth (c); the total error 0.008% (our boundaries and the ground truth are in red and green, respectively)

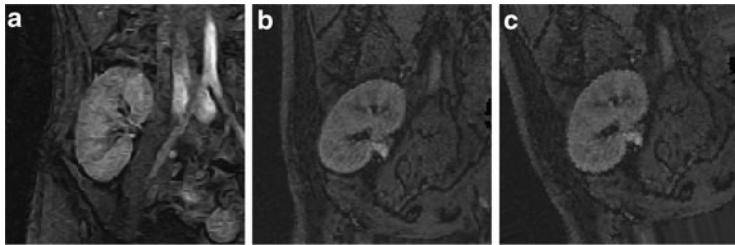


**Fig. 4.9** Segmentation errors vs. the SNR (our boundaries are in red)

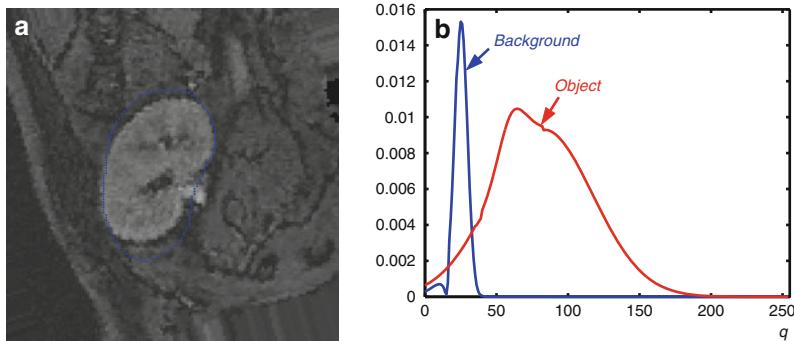
boundary and shows the LCDG-estimates of the marginal probability distributions over the object (starfish) and background. The resulting segmentation in Fig. 4.8 has the error of 0.008% with respect to the ground truth. Figure 4.9 demonstrates the robustness of our segmentation in the case when the same starfish is placed to the uniform background and the whole image is distorted by the uniform independent random noise. When the signal-to-noise ratio (SNR) changes from 20 to  $-1.5$  dB, the segmentation error increases from 0.00013% to only 2.9%.

The same segmentation stages for the typical DCE-MRI kidney images are shown in Figs. 4.10–4.12. Here, the total error of our segmentation is 0.63% with respect to the ground truth provided by a radiologist. Figure 4.13 shows more experimental results for the DCE-MRI with the different kidney contrast. The geometric deformable models with the level set shape priors in [1, 17] produce the like accurate results on the contrast images such as the eleven top images in Fig. 4.13, but degrade with the decreasing contrast, e.g., on the nine bottom images there.

These and other experiments with various natural images confirm the promising potential of embedding the shape and appearance priors to the parametric deformable models. Our approach assumes the boundaries of the training and goal objects are similar under the affine transform and the SIFT technique reliably detects the matching points to automatically align the images of the goal objects in spite of



**Fig. 4.10** One of the mutually aligned training kidney images (a), an image to be segmented (b), and its alignment to the training set (c)

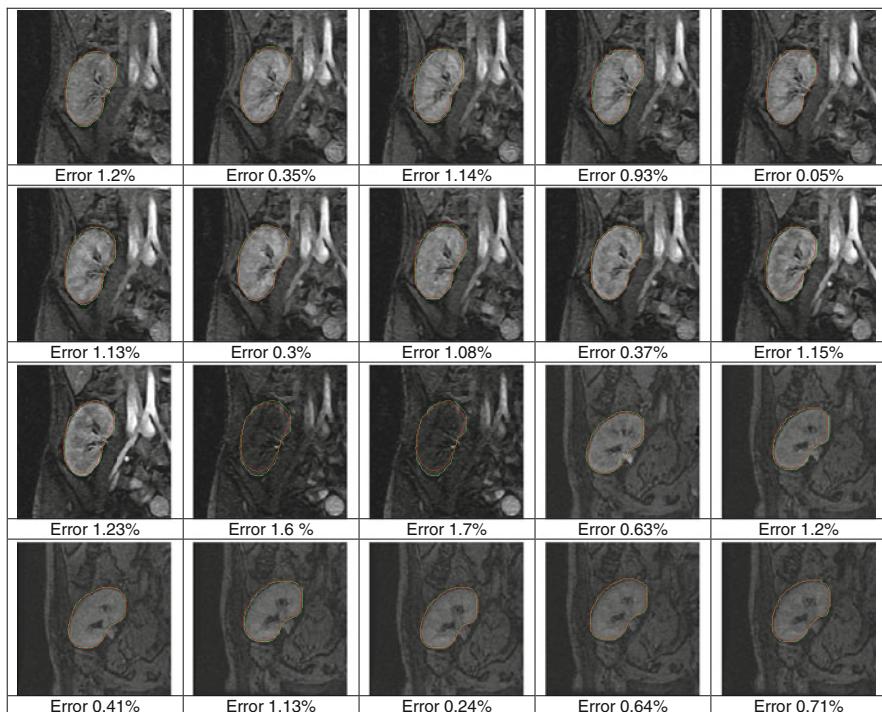


**Fig. 4.11** Initialization of the deformable model (a) shown in *blue*, and the LCDG-estimates of the marginal gray value distributions  $p_{\text{pix},\text{ob}}$  and  $p_{\text{pix},\text{bg}}$  (b)



**Fig. 4.12** Final segmentation of the image aligned to the training set (a) and the same result (b) after its inverse affine transform to the initial image in Fig. 4.10b in comparison to the ground truth (c); the total error 0.63% (our boundaries and the ground truth are in *red* and *green*, respectively)

their different backgrounds. While the assumptions restrict the application area comparing to the more conventional parametric deformable models, these latter totally fail on the above and similar images. Comparing to more accurate geometric models with linear combinations of the training distance maps as the shape priors,



**Fig. 4.13** Segmentation of 20 other kidney DCE-MRI: our final boundaries and the radiologist ground truth are shown in red and green, respectively

our approach escapes theoretical inconsistencies of these latter, is computationally much simpler and faster, and has the similar performance on the contrast images but outperforms on the low-contrast ones.

## References

1. Tsai A, Yezzi A, Wells W, Tempany C, Tucker D, Fan A, Grimson W, Willsky A (2003) A shape based approach to the segmentation of medical imagery using level sets. *IEEE Trans Med Imaging* 22:137–154
2. Rousson M, Paragios N (2002) Shape priors for level set representations. In: Proceedings of the 7th European conference computer vision, Copenhagen, Denmark, June 2002, part 2 (Lecture notes in computer science), vol 2351. Springer, Berlin, pp 78–92
3. Kass M, Witkin A, Terzopoulos D (1987) Snakes: active contour models. *Int J Comput Vis* 1:321–331
4. Caselles V, Kimmel R, Sapiro G (1997) On geodesic active contours. *Int J Comput Vis* 22:61–79
5. Osher S, Sethian J (1988) Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulation. *J Comput Phys* 79:12–49
6. Chan T, Vese L (2001) Active contours without edges. *IEEE Trans Image Process* 10:266–277

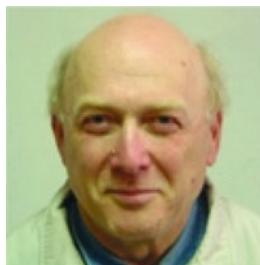
7. Leventon M, Grimson E, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proceeding of the IEEE conference computer vision pattern recognition, Hilton Head, South Carolina, 13–15 June 2000, IEEE Computer Society Press, Los Alamitos, pp 316–323
8. Shen D, Davatzikos C (2000) An adaptive-focus deformable model using statistical and geometric information. *IEEE Trans Pattern Anal Mach Intell* 22:906–913
9. Chen Y, Thiruenkadam S, Tagare H et al (2001) On the incorporation of shape priors into geometric active contours. In: Proceeding IEEE workshop on variational and level set methods, Vancouver, Canada, 13 July 2001, IEEE Computer Society Press, Los Alamitos, pp 145–152
10. Cootes T, Taylor C, Cooper D, Graham J (1995) Active shape models – their training and application. *Comput Vis Image Underst* 61:38–59
11. Pentland A, Sclaroff S (1991) Closed-form solutions for physically based shape modeling and recognition. *IEEE Trans Pattern Anal Mach Intell* 13:715–729
12. Staib L, Duncan J (1992) Boundary finding with parametrically deformable contour models. *IEEE Trans Pattern Anal Mach Intell* 14:1061–1075
13. Chakraborty A, Staib L, Duncan J (1994) An integrated approach to boundary finding in medical images. In: Proceeding of the IEEE workshop biomedical image analysis, Seattle, Washington, 24–25 June 1994, IEEE Computer Society Press, Los Alamitos, pp 13–22
14. Pizer S, Gerig G, Joshi S, Aylward S (2003) Multiscale medial shape-based analysis of image objects. *Proc IEEE* 91:1670–1679
15. Styner M, Gerig G, Pizer S, Joshi S (2002) Automatic and robust computation of 3D medial models incorporating object variability. *Int J Comput Vis* 55:107–122
16. Huang X, Metaxas D, Chen T (2004) MetaMorphs: deformable shape and texture models. In: Proceeding of the IEEE conference computer vision pattern recognition, Washington, DC, 27 June–2 July 2004, vol 1. IEEE Computer Society Press, Los Alamitos, 2004, pp 496–503
17. Tsai A, Wells W, Tempany C, Grimson E, Willsky A (2003) Coupled multi-shape model and mutual information for medical image segmentation. In: Proceeding of the 18th international conference information processing in medical imaging, Ambleside, UK, 20–25 July 2003 (Lecture notes in computer science), vol 2732. Springer, Berlin, pp 185–197
18. Van Leemput K, Maes F, Vandermeulen D, Suetens P (2003) A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans Med Imaging* 22(1):105–119
19. Joshi S (1997) Large deformation diffeomorphisms and Gaussian random fields for statistical characterization of brain submanifolds. PhD Thesis, Washington University, St. Louis
20. Paragios N, Deriche R (2000) Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans Pattern Anal Mach Intell* 22:266–280
21. Cootes T, Taylor C (2001) Statistical models of appearance for medical image analysis and computer vision. In: Sonka M, Hanson KM (eds) Proceeding of the SPIE (Medical imaging 2001: image processing), vol 4322, pp 236–248
22. Hill A, Thornham A, Taylor C (1993) Model-based interpretation of 3-D medical images. In: Proceeding of the 4th British machine vision conference, Guildford, September 1993, BMVA, pp 339–348
23. Kelemen A, Szekely G, Gerig G (1999) Elastic model-based segmentation of 3-D neuroradiological data sets. *IEEE Trans Med Imaging* 18:828–839
24. Wang Y, Staib LH (1998) Integrated approaches to nonrigid registration in medical images. Proceeding of the 4th IEEE workshop applications of computer vision, Princeton, NJ, 19–21 October 1998, IEEE Computer Society Press, Los Alamitos, pp 102–108
25. Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceeding of the international conference computer vision, Kerkyra, Corfu, Greece, 20–25 September 1999, vol 2. IEEE Computer Society Press, Los Alamitos, pp 1150–1157
26. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60:91–110
27. Mikolajczyk K, Schmid C (2003) A performance evaluation of local descriptors. Proceeding of the IEEE conference computer vision pattern recognition, Madison, Wisconsin, 18–20 June 2003, vol. 2. IEEE Computer Society Press, Los Alamitos, pp 257–263

28. Brown M, Lowe D (2002) Invariant features from interest point groups. Proceeding of the 14th British machine vision conference, Cardiff, 2–5 September 2002, BMVA, pp 656–665
29. Johansson B, Soderberg R (2004) A repeatability test for two orientation based interest point detectors, Department of Electrical Engineering, Linköping University, LiTH-ISY-R-2606, SE-581 83, April 2004
30. Titterington DM, Murray GD, Murray LS et al (1981) Comparison of discrimination techniques applied to a complex data set of head injured patients. *J R Stat Soc Series A* 144:145–175
31. Gimel'farb GL (1999) Image textures and Gibbs random fields. Kluwer Academic, Dordrecht
32. Gimel'farb G, Farag AA, El-Baz A (2004) Expectation-maximization for a linear combination of Gaussians. Proceeding of the IAPR International Conference Pattern Recognition, 22–27 August 2004, Cambridge, UK, vol 3. IEEE Computer Society Press, Los Alamitos, pp 422–425
33. El-Baz A, Farag AA, Gimel'farb G (2005) Unsupervised segmentation of multi-modal images by a precise approximation of individual modes with linear combinations of discrete Gaussians In: Proceeding of the IEEE workshop on learning in computer vision and pattern recognition, June 2005, San Diego, CA, USA, p 54

## Biography



**Ayman S. El-Baz** received the B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering Department, University of Louisville, in August 2006. His current research is focused on developing new computer assisted diagnosis systems for different diseases and brain disorders.



**Georgy Gimel'farb** graduated from Kiev Polytechnic Institute, Kiev, Ukraine, and received the Ph.D. degree in engineering cybernetics from the Institute of Cybernetics, Academy of Sciences of the Ukraine, and the D.Sc. (Eng) degree in control in engineering from the Higher Certifying Commission of the USSR, Moscow, Russia. He joined the University of Auckland, Auckland, New Zealand, in July 1997. His research is focused on image analysis, computer vision, and statistical pattern recognition.

# Chapter 5

## Classification of Breast Mass in Mammography with an Improved Level Set Segmentation by Combining Morphological Features and Texture Features

Jinshan Tang and Xiaoming Liu

**Abstract** In this chapter, an improved level set-based mass segmentation is proposed. The new mass segmentation algorithm is based on a new energy functional, which combines energy terms from several previous algorithms. Experimental results show that the proposed level set segmentation algorithm can get better segmentation results than the original one. Based on the contours obtained by the improved level set algorithm, mass regions are classified into malignant or benign using linear discriminant analysis (LDA) and support vector machine (SVM) by combining morphological features and texture features. Experiments show that the accuracy of mass classification by combining morphological features and textures features is higher than the accuracy of mass classification using only morphological features or texture features.

### 5.1 Introduction

In the world, breast cancer is the second most common type of non-skin cancer and the fifth most common cause of cancer death [1]. In order to reduce the death rate of breast cancer, early detection of breast cancer is very important. Mammography is one of the effective tools to detect breast cancer in the early stage [2–4]. In the application of mammography to the diagnosis of breast cancer, computer aided detection (CAD) technologies play a very important role and many CAD systems have been developed. By adopting image enhancement technology, pattern recognition, computer aided breast detection systems can assist the radiologists [5] to find the abnormal regions in the breast easily and effectively.

---

J. Tang (✉)

Department of Advanced Technologies, Alcorn State University, 1000 ASU Dr.Lorman, MS 39096, USA

e-mail: jtang@alcorn.edu

X. Liu (✉)

College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, 430081, China

In computer aided breast cancer detection systems, mass classification technology is one of the key technologies used in CAD systems. Some papers have been published in the past on mass classification. Based on the existing publications, the performance of mass classification depends on the performance of mass segmentation [6]. Segmentation is the basis of classification. There are a lot of work on mass segmentation and classification.

For mass segmentation, the methods could be performed manually [7], semiautomatically [8], or automatically [9]. Besides the work in [7–9], many other results on mass segmentation have been published. In [10], a two-step method for mass segmentation was developed. In the proposed method, regions of interest were first extracted from the images by adaptive thresholding and then the initial segmentation was further refined with a modified Markov random field (MRF) model to achieve the final segmentation. In [11], a mass segmentation method on dense breast mammograms was presented. The initial segmentation was obtained using thresholding and the final segmentation was obtained by a watershed segmentation method. In [12], two methods for mass segmentation were proposed. One was based on radial gradient index (RGI) feature and the other was based on probabilistic models.

After the mass is segmented, classification is applied to classify the segmented masses into benign or malignant. Several papers have investigated mass classification. In [6], Sahiner et al. investigated mass classification and compared the mass classification accuracies using the contours obtained by manual segmentation and computer based segmentation respectively. Linear discriminant analysis with step-wise feature selection was used for classification. Their research indicated that classification accuracies obtained by computer based mass segmentation and manual segmentation were similar. In [13], Rangayyan et al. studied mass classification using morphological features. In their research, mass boundary was segmented into concave and convex parts and polygonized. Features quantifying the extent of the spiculated nature of the boundary and the degree of narrowness of the spicules were extracted for mass classification. In [14], Huo et al. extracted mass with region growing method from a manually extracted square shaped region of interest. Features quantifying the marginal spiculation of a mass with radial edge-gradient analysis were extracted and simple thresholding was used for classification. They found that the characterization of spiculation resulting from radial edge-gradient analysis was comparable to that of experienced radiologists. In [15], Rojas Dominguez and Nandi investigated mass classification based on automated segmentation of masses. Dynamic programming based boundary tracking and constrained region growing segmentation methods were used for segmentation. Six features from mass margins were extracted, and three classifiers (Bayesian classifier, linear discriminant analysis, and support vector machine) were compared for classification. They showed that features extracted from automated contours can benefit the diagnosis of breast masses. In [16], Yuan et al. proposed a two-stage method for automated mass segmentation. A radial gradient index based segmentation method was used to obtain initial contour and a geometric active contour model was used to refine the segmentation. The evolution of the contour was decided on a dynamic stopping criterion based on lesion boundaries.

In this chapter, we investigate both mass segmentation and classification. For mass segmentation, we investigate the method based on level set. For feature extraction, we investigate morphological and texture features. For classifiers, we study LDA and SVM based classifiers.

## 5.2 An Improved Level Set Algorithm for Mass Segmentation

Our mass segmentation algorithm is based on level set. Level set method is a variation of active contour model [17] and it provides a flexible mechanism which permits topological change of target objects, such as object splitting and merging [18, 19].

Level set method is based on region based snake. In region based snake, an image is assumed to be composed of two homogenous regions and the contour which separates the two regions can be formulated with the following energy functional [18, 19]

$$\begin{aligned} E^{\text{CV}}(C, c_1, c_2) = & \lambda_1 \int_{\text{inside}(C)} |I_0(x, y) - c_1|^2 dx dy \\ & + \lambda_2 \int_{\text{outside}(C)} |I_0(x, y) - c_2|^2 dx dy + \mu |C| \end{aligned} \quad (5.1)$$

where,  $\lambda_1, \lambda_2 \geq 0$  and  $\mu \geq 0$  are the weights,  $C$  is the evolving contour and  $|C|$  is the length of the contour,  $\text{inside}(C)$  and  $\text{outside}(C)$  represent the regions inside and outside the contour  $C$  respectively.  $c_1$  and  $c_2$  are two constants that approximate the image intensity in  $\text{inside}(C)$  and  $\text{outside}(C)$ . The values of optimal  $c_1$  and  $c_2$  can be shown to be the averages of intensities in the regions  $\text{inside}(C)$  and  $\text{outside}(C)$  [19].

Equation (5.1) can be represented by level set [19]. In level set theory, a two-dimensional contour  $C \subset \Omega$  is represented by the zero level set of a Lipschitz function [19]

$$\phi : \Omega \rightarrow \mathbb{R}, \quad \text{i.e.,} \quad C = \{(x, y) \in \Omega, \phi(x, y) = 0\} \quad (5.2)$$

$\phi : \Omega \rightarrow \mathbb{R}$  is called a level set function. In level set, the evolution of the curve is implemented by updating the level set function  $\phi(x, y)$  at fixed coordinates [19]. The value of the level set function  $\phi$  is negative/positive when it is inside/outside of the contour  $C$ .

The original level set has some shortcomings when it is applied to mass segmentation [17]. For example, the original level set assumes that each region is homogeneous; however, this is not the case in mass segmentation because a mass region is generally inhomogeneous. In order to deal with this issue, we adopt the level set developed by Li et al. [20]. The level set developed in [20] can overcome the shortcoming by introducing a region-scalable fitting (RSF) energy term in the energy functional. Besides, a level set regularization term is also introduced to make the function approximate the signed distance function. The energy functional is defined as [20]

$$\begin{aligned}
E(\phi, f_1, f_2) = & \lambda_1 \int \left( \int K_\sigma(x-y) |I(y) - f_1(x)|^2 H(\phi(y)) dy \right) dx \\
& + \lambda_2 \int \left( \int K_\sigma(x-y) |I(y) - f_2(x)|^2 (1 - H(\phi(y))) dy \right) dx \\
& + \mu \int |\nabla H(\phi(x))| dx + v \int \frac{1}{2} (|\nabla \phi(x)|^2 - 1) dx
\end{aligned} \tag{5.3}$$

where  $\phi$  is a signed distance function and has the intrinsic property that  $\|\nabla \phi\| = 1$ ,  $H$  is Heaviside function, and  $K_\sigma$  is a Gaussian function with  $\sigma > 0$ , which is defined as

$$K_\sigma(x) = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-|x|^2/2\sigma^2} \tag{5.4}$$

In (5.3), the first two terms are region-scalable fitting energy terms and the last term is the level set regularization term. In (5.3),  $f_1(x)$  and  $f_2(x)$  are two functions which approximate original grayscale values in region  $\phi(x) > 0$  and  $\phi(x) < 0$  respectively. It was shown that  $f_1(x)$  and  $f_2(x)$  are weighted averages of the grayscale values in a neighborhood of the pixel  $x$  [20].

Although the level set in [20] can solve the inhomogeneous issue, it is not sensitive to edge information which is very important for classification of masses, thus we integrate a new energy term from [21] into the energy functional (5.3). The new energy term from [21] is as follows

$$E_g = \int_{\Omega} g |\nabla H(\phi)| dx dy = \int_{\Omega} g \delta(\phi) |\nabla \phi| dx dy \tag{5.5}$$

where,

$$g = \frac{1}{1 + |\nabla G_\sigma * I|^2} \tag{5.6}$$

In the above equation,  $G_\sigma$  is the Gaussian kernel with standard deviation  $\sigma$ . The energy is low when the curve is located at the object boundaries [20]. A mass segmentation algorithm with energy term (5.5) has been proposed for mass segmentation in our previous work [17] and has obtained better performance.

To further improve the performance of the proposed level set, global data fitting energy as in traditional level set method [19] is integrated into the optimization problem. Thus the final energy functional we used in the proposed level set is

$$\begin{aligned}
E(\phi, f_1, f_2) = & \lambda_1 \int \left( \int K_\sigma(x-y) |I(y) - f_1(x)|^2 H(\phi(y)) dy \right) dx \\
& + \lambda_2 \int \left( \int K_\sigma(x-y) |I(y) - f_2(x)|^2 (1 - H(\phi(y))) dy \right) dx \\
& + \gamma_1 \int |I(x) - c_1|^2 (1 - H(\phi(x))) dx \\
& + \gamma_2 \int |I(x) - c_2|^2 H(\phi(x)) dx + \mu \int |\nabla H(\phi(x))| dx \\
& + v \int g \delta(\phi) |\nabla \phi| dx + w \int \frac{1}{2} (|\nabla \phi(x)|^2 - 1) dx
\end{aligned} \tag{5.7}$$

where,  $c_1$  and  $c_2$  are the averaged grayscale values inside the region and outside the region. For numerical implementation, the Heaviside function can be approximated with [19]

$$H_\varepsilon(\phi) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan\left(\frac{\phi}{\varepsilon}\right) \right] \tag{5.8}$$

and the corresponding Dirac measure is  $\delta_\varepsilon(\phi) = \frac{dH_\varepsilon(\phi)}{d\phi} = \frac{\varepsilon}{\pi(\varepsilon^2 + \phi^2)}$ , as given in [19].

### 5.3 Mass Classification: Feature Extraction and Classification

After mass region is segmented out, we will extract the features from the boundary of the mass and the margin region between the mass and background for mass classification. From the past work [7, 22], we know that a typical benign mass has a round, smooth, and well-circumscribed boundary, while a malignant tumor usually has a spiculated, rough, and blurry boundary [7, 22]. Thus, we can use boundary analysis to classify the masses into benign or malignant. In this chapter, we will investigate both morphological features and texture features [17]. Although these features have been used in different publications, we will combine them together in this chapter to get better performance.

#### 5.3.1 Morphological Features

The morphological features used in this chapter are compactness, moments, Fourier descriptor, normalized radial length (NRL) based features, and relative gradient orientation. These aforementioned features were adopted in the past work for mass classification [17, 22].

### 5.3.1.1 Compactness

Compactness was used by Kilday et al. [8] as a feature for mass classification. It is a measure of contour complexity versus the enclosed area [8] and is defined as

$$F_1 = C = 1 - \frac{4\pi a}{p^2} \quad (5.9)$$

where,  $p$  and  $a$  are the object perimeter and area respectively. As is shown in [8], malignant masses with concavities or spicules will have higher compactness values than smooth and round benign masses.

### 5.3.1.2 Moments

Moments have been used in pattern recognition widely and used successfully in mass classification [23]. Let the contour with  $N$  pixels be  $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ , and the centroid of the contour be  $(\bar{x}, \bar{y})$ ,  $z(i)(i = 1, 2, \dots, N)$  be the Euclidean distance from the  $i$ th pixel on the contour to the centroid, then the  $p$ th moment is defined as [23]

$$m_p = \frac{1}{N} \sum_{i=1}^N [z(i)]^p \quad (5.10)$$

and the  $p$ th centered moment is given by Liang et al. [23]

$$M_p = \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^p \quad (5.11)$$

We adopt the moment features used in [23] for mass classification. The moment features are defined as [23]

$$F_2 = \frac{(M_2)^{1/2}}{m_1} = \frac{\left[ \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^2 \right]^{1/2}}{\frac{1}{N} \sum_{i=1}^N z(i)} \quad (5.12)$$

$$F_3 = \frac{(M_3)^{1/3}}{m_1} = \frac{\left[ \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^3 \right]^{1/3}}{\frac{1}{N} \sum_{i=1}^N z(i)} \quad (5.13)$$

$$F_4 = \frac{(M_4)^{1/4}}{m_1} = \frac{\left[ \frac{1}{N} \sum_{i=1}^N [z(i) - m_1]^4 \right]^{1/4}}{\frac{1}{N} \sum_{i=1}^N z(i)} \quad (5.14)$$

### 5.3.1.3 Fourier Descriptor (FD)

We will use the features based on Fourier descriptor developed by Liang et al. [23]. The Fourier descriptor measure defined in [23] is robust to noise because it emphasize low-frequency components and thus it is useful in discriminating between shapes with smooth and irregular contours. In order to obtain the Fourier descriptor measurement defined in [23], a normalized Fourier descriptor is defined first. Let  $\{(x_0, y_0), (x_1, y_1), \dots, (x_{N-1}, y_{N-1})\}$  be the point set on the contour, then the normalized Fourier descriptor (NFD) is defined as [23]

$$\text{NFD}(k) = \begin{cases} 0 & k = 0 \\ |a_k|/|a_1| & k = 1, 2, \dots, \frac{N}{2} \\ |a_{k+N}|/|a_1| & k = -1, -2, \dots, -\frac{N}{2} + 1 \end{cases} \quad (5.15)$$

$$a_n = \frac{1}{N} \sum_{i=0}^{N-1} z_i e^{-j2\pi ni} \quad (5.16)$$

where  $|a|$  is the norm of  $a$  and  $z_i = x_i + y_i * j$  ( $0 \leq i \leq N - 1$ ) is a complex number. It is obvious that NFD is now independent of translation, scaling, rotation, and the start point  $z_0$  on the object contour [23]. Based on the normalized Fourier descriptor, the Fourier descriptor measurement is defined as [23]

$$F_5 = \text{FD} = \frac{\sum_{k=-N/2+1, k \neq 0}^{N/2} \text{NFD}(k)/|k|}{\sum_{k=-N/2+1, k \neq 0}^{N/2} \text{NFD}(k)} \quad (5.17)$$

### 5.3.1.4 Normalized Radial Length (NRL) Based Features

NRL is defined as the Euclidean distance from the mass center to each of the boundary coordinates and normalized by dividing by the maximum radial length  $L_{\max}$  [8]. For a mass contour, the NRL vector is given as  $R = \{r_i, 0 \leq i \leq N - 1\}$ , where  $N$  is the number of edge pixels on the contour and  $r_i \leq 1$ . The features we extract for mass classification are NRL mean, NRL standard deviation, NRL entropy, and NRL area ratio, which are defined as follows [8]:

---

NRL mean:

$$F_6 = \mu_{\text{NRL}} = \frac{1}{N} \sum_{i=0}^{N-1} r_i$$

NRL standard deviation:

$$F_7 = \sigma_{\text{NRL}} = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (r_i - \mu_{\text{NRL}})^2}$$

NRL entropy:

$$F_8 = E = \sum_{k=1}^{N_h} p_k \log(p_k)$$

NRL area ratio:

$$F_9 = \text{AR}_{\text{NRL}} = \left\{ \frac{1}{N \times \mu_{\text{NRL}}} \sum_{k=0}^{N-1} (r_k - \mu_{\text{NRL}}) : r_k > \mu_{\text{NRL}} \right\}$$


---

where,  $p_k$  is the probability that the radial length lies between  $(k-1) \times L_{\max}$  and  $k \times L_{\max}$ , and  $N_h$  is the number of histogram bins.

### 5.3.1.5 Relative gradient orientation

We use the features from gradient information proposed in [24]. The features are called relative gradient orientation (RGO) measurement, which is used to measure spiculation. RGO is defined as the acute angle  $\theta$  between radial direction of a point on the contour and the gradient direction of the point [24]. For a mass contour, the RGO vector is given by  $\Theta^\circ = \{\theta_i^\circ, 0 \leq i \leq N-1\}$ . Here  $N$  is the number of edge pixels on the contour. To normalize the value to  $[0, 1]$ , the angle  $\theta_i$  is divided by  $\pi/2$ , that is  $\Theta = \{\theta_i, \theta_i = \frac{2}{\pi}\theta_i^\circ\}$ . Several features are extracted from RGO vector:

---

RGO mean:

$$F_{10} = \mu_{\text{RGO}} = \frac{1}{2} \sum_{i=0}^{N-1} \theta_i$$

RGO standard deviation:

$$F_{11} = \sigma_{\text{RGO}} = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (\theta_i - \mu_{\text{RGO}})^2}$$

RGO entropy:

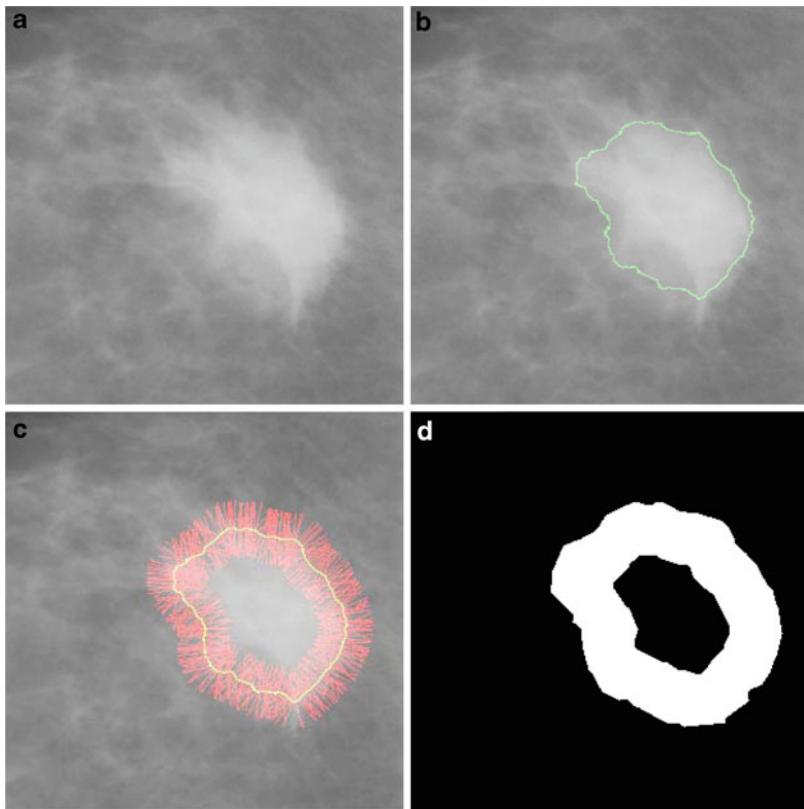
$$F_{12} = E = \sum_{k=1}^{N_h} p_k \log(p_k)$$


---

where,  $p_k$  is the probability that the radial length lies between  $(k-1) \times \theta_{\max}$  and  $k \times \theta_{\max}$ ,  $N_h$  is the number of histogram bins.

### 5.3.2 Texture Features

Besides the shape information of a mass contour, the texture information of the region surrounding the mass boundary also contains important information to discriminate benign and malignant masses [7, 20]. Thus we also use texture information for mass classification. The texture features we extract are based on Gray Level Co-occurrence Matrix (GLCM) [25, 26]. GLCM has been widely used in several applications, including analysis of mammographic masses [7, 17]. We use the same region used in [7] for texture analysis. The region used in [7] for the texture



**Fig. 5.1** An example of feature extraction based on level set segmentation. (a) a  $301 \times 301$  pixel section of a mammography containing a malignant mass; (b) segmentation result on the ROI; (c) extraction of normal pixels on the boundary of the mass, which is the same as [7]; (d) band of pixels across the boundary of the mass extracted for texture features

analysis is a band around the closed contour of the mass, as shown in Fig. 5.1. The width of extracted ribbon is limited to be 8 mm across the boundary (4 mm or 20 pixels on either side of the boundary, where available, at a resolution of  $200 \mu\text{m}$  per pixel, as used in [7]). The texture features used in the book chapter include energy, entropy, correlation, inverse difference moment, contrast, cluster shade, and variance. Although these features have been used in different papers [7, 17], we will still give the mathematical equation for easy reading. In order to do that, we need to describe the definition of GLCM. As is defined in [25, 26] and many other publications, a GLCM matrix  $M$  is defined as a  $G \times G$  matrix, whose rows and columns are indexed by the image gray levels  $i = 1, \dots, G$ , where  $G = 2^n$  for a  $n$ -bit image. An element  $M_d(i, j)$  reflects the distribution of the probability of occurrence of a pair of gray levels  $(i, j)$  separated by a given distance  $(d)$ . Based on the definition of GLCM, the texture features aforementioned are computed as follows [7, 17]:

---

Energy:	$\text{Energy} = f_1 = \sum_{i=1}^G \sum_{j=1}^G M(i,j)^2$
Entropy:	$\text{Entropy} = f_2 = - \sum_{i=1}^G \sum_{j=1}^G M(i,j) \log(M(i,j))$
Correlation:	$\text{Correlation} = f_3 = \sum_{i=1}^G \sum_{j=1}^G \frac{(i - \mu)(j - \mu) \times M(i,j)}{\sigma^2}$
Inverse difference moment:	$\text{Inverse difference moment} = f_4 = \sum_{i=1}^G \sum_{j=1}^G \frac{M(i,j)}{1 + (i - j)^2}$
Contrast:	$\text{Contrast} = f_5 = \sum_{i=1}^G \sum_{j=1}^G (i - j)^2 \times M(i,j)$
Cluster shade:	$\text{Cluster shade} = f_6 = \sum_{i=1}^G \sum_{j=1}^G ((i - \mu) + (j - \mu))^3 \times M(i,j)$
Variance:	$\text{Variance} = f_7 = \sum_{i=1}^G \sum_{j=1}^G (i - \mu)^2 \times M(i,j)$

---

where  $\mu$  and  $\sigma$  are the mean and standard deviation values of GLCM in row or column.

In the computation of the features, we scale the gray level to 16 and four GLCMs are constructed by scanning each mass ribbon in  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  directions with pixel distances. The seven features described above are calculated for each direction; thus we obtain 28 features. The distance used is  $d = 1$ , as in [7]. Note for  $45^\circ$  and  $135^\circ$ , the actual distance used is  $d = \sqrt{2}$ .

### 5.3.3 Classification Methods

Two classification methods are investigated in the book chapter. One is the classification method based on linear discriminant analysis (LDA) and the other is support vector machine. The basic idea of LDA is to find a linear transform matrix which projects a  $D$ -dimensional feature vector  $\mathbf{x}$  in the original space to a vector  $\mathbf{y}$  which belongs to a lower dimensional space (feature space). In this book chapter, LDA is used to reduce the space dimension. After dimension reduction, KNN is usually used for final classification.

The other classification method is support vector machine. For a two-class problem, SVM finds a hypersurface which separates samples from the two classes, while the hypersurface is selected such that its distance to the nearest training data on either side of the surface is maximized [26]. In SVM, the choice of kernel is very important because it determines the performance of the SVM. Several kernels have been proposed for SVM till now. The kernels include linear kernel, polynomial kernel, and Gaussian radial basis functions (RBFs). In this book chapter, we use the RBF kernel. The RBF kernel is defined as [27]

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|}{2\sigma^2}\right) \quad (5.18)$$

where,  $\sigma > 0$  is a constant that defines the kernel width.

## 5.4 Experiment Results

### 5.4.1 Image Dataset

The images for the experiments were from the DDSM database [28]. A set of 309 mammograms were used for the validation of the methods. The mammograms were resampled at a pixel size of  $200 \mu\text{m} \times 200 \mu\text{m}$  and 256 gray levels. One hundred and forty-two of the dataset were benign and 167 of them were malignant. The sizes of the benign masses varied in the range 16–362 mm with a mean size of 108 mm.

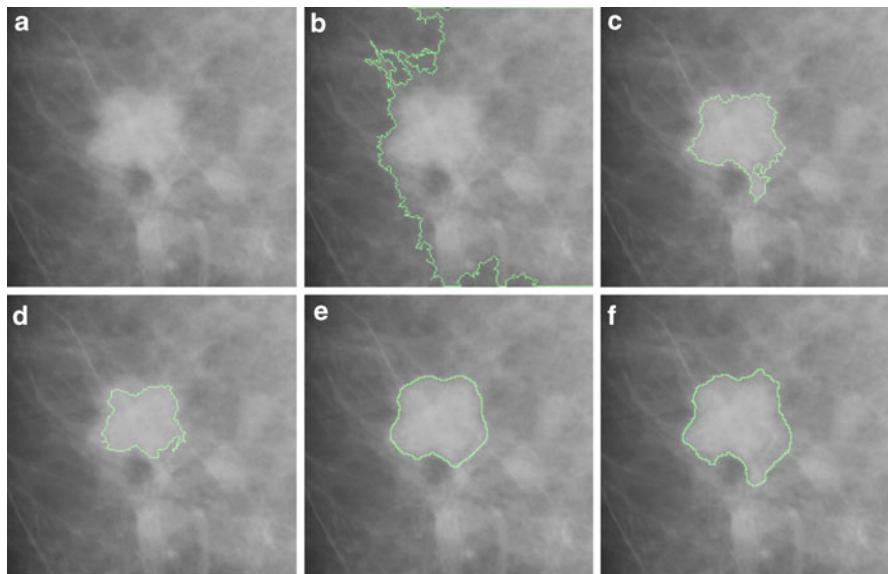
### 5.4.2 Experimental Results for Segmentation

We first investigated the performance of the proposed segmentation algorithm. In the segmentation experiments, the parameters in the experiments were set as  $\lambda_1 = \lambda_2 = 1.0$ ,  $\gamma_1 = \gamma_2 = 2.0$ ,  $\mu = 0.01 \times 255^2$ ,  $v = 1.0$ ,  $\omega = 1.0$ . The experimental results are shown in Figs. 5.2 and 5.3. Figure 5.2a shows the original ROI with a benign mass, and Fig. 5.2b shows the segmentation result obtained by grayscale thresholding. Figure 5.2c–f show the segmentation results obtained by region growing, MRF method, traditional level set, and the proposed method respectively. From the segmentation results, we can find that gray threshold failed in this case, and the traditional level set with only global grayscale difference and the proposed algorithm can get better results. Figure 5.3 shows the results of several segmentation methods on a malignant mass with fuzzy boundary. Among all the methods, the proposed method can achieve the best performance with accurate boundary outline.

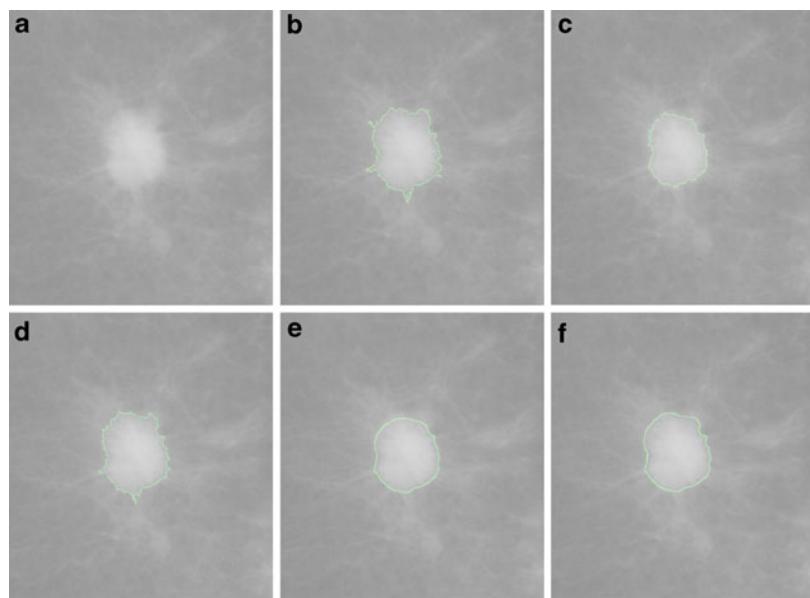
### 5.4.3 Experimental Results for Classification

Three experiments were performed in the research. In each experiment, both the LDA based classification and SVM based classification method were used for benign/malignant mass classification. The first experiment was designed to evaluate the effectiveness of morphological features for mass classification. The second experiment was designed to investigate the GLCM based texture features for mass classification. The third experiment was designed to evaluate the combination of morphological and texture features for mass classification.

In SVM classification method, the selection of the parameters is very important and it will affect the performance of the classification. Two parameters are important to the performance of classification. One is the misclassification penalty parameter  $C$  [26] and the other is the parameter  $\sigma$  in RBF. In order to find the best parameter values, we used five cross-validation to select the optimal values. In



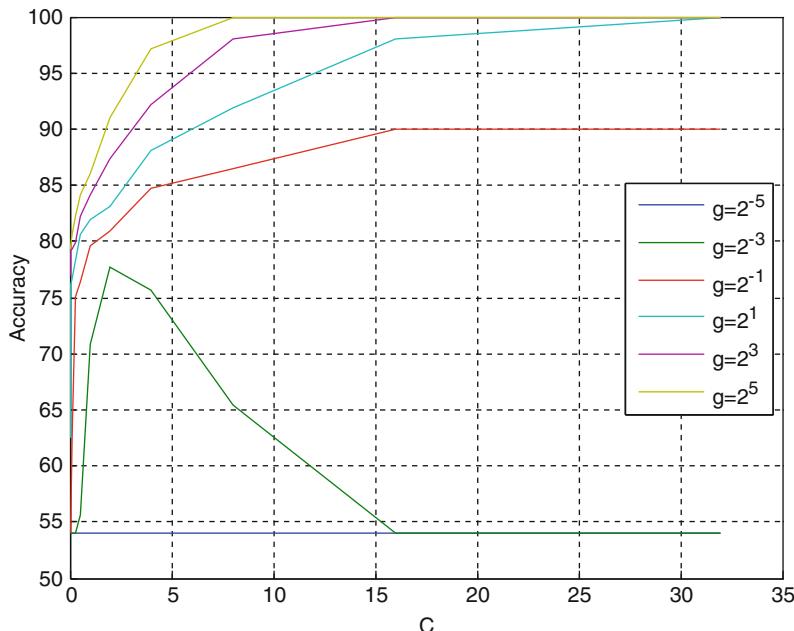
**Fig. 5.2** Segmentation results of different methods on a mammography containing a benign mass. (a) Original ROI, (b) result with grayscale threshold, (c) result with region growing, (d) result with Markov random field, (e) result with traditional level set, (f) result with our method



**Fig. 5.3** Segmentation results of different methods on a mammography containing a malignant mass. (a) Original ROI, (b) result with grayscale threshold, (c) result with region growing, (d) result with Markov random field, (e) result with traditional level set, (f) result with the proposed method

five cross-validation, the training samples were split in five sets with equal size. Four of them were used for training and the left one set was used for testing. For each setting  $C$  and  $\sigma$ , five classification experiments were performed, and the values of  $C$  and  $\sigma$  with highest average accuracy were selected for mass classification. In the experiments, the range of the parameter  $C$  and  $\sigma$  were  $[2^{-5}, 2^5]$ , and the tool box LIBSVM [27] was used for classification. Figure 5.4 shows the accuracies of SVM with morphological features and level set-based segmentation under different parameter values. It can be shown that the values need to be tuned for good performance. It should be noted that all the features used for SVM were normalized into  $[0, 1]$ .

In the first experiment, 12 morphological features were first extracted from the mass boundaries, and then LDA and SVM were used for mass classification. To evaluate the performance of a classifier, we used the receiver operating characteristic (ROC) analysis. The ROC analysis used in the experiments were the LABROC program [29] and the area under the ROC curve ( $A_z$ ) was used to measure the classification accuracy. Table 5.1 shows the classification accuracies of LDA and SVM based on level set segmentation contour. From Table 5.1, we can find that the classification results obtained by LDA and SVM with morphological features based on level set segmentation resulted in accuracies of 73% and 76% respectively. The areas under the ROC curve ( $A_z$ ) are 0.8 and 0.85 for LDA and SVM respectively. The experiments show that SVM was better than LDA when morphological features were used.



**Fig. 5.4** The classification performance obtained by SVM with morphological features under different parameter settings

**Table 5.1** Comparison of classification results using different classifiers with morphological features

	Benign (%)	Malignant (%)	Total (%)	$A_z$
LDA	65	80	73	0.8
SVM	68	83	76	0.85

**Table 5.2** Classification results with texture features from level set-based segmentation

	Benign (%)	Malignant (%)	Total (%)	$A_z$
LDA	0.59	0.71	0.65	0.66
SVM	0.46	0.80	0.65	0.70

**Table 5.3** Classification results with both morphological and texture features from level set-based segmentation

	Benign (%)	Malignant (%)	Total (%)	$A_z$
LDA	72	83	78	0.84
SVM	75	79	0.77	0.87

In the second experiment, we evaluated mass classification using texture features. Twenty-eight features were extracted from each mass ribbon. Table 5.2 shows the benign/malignant classification accuracies obtained by LDA and SVM with texture features respectively. The  $A_z$  values for LDA and SVM are 0.66 and 0.70 respectively. Experimental results show that SVM was better than LDA when texture features were used.

In the third experiment, we investigated mass classification using the combination of morphological features and texture features. The classification results from the combined features are shown in Table 5.3. The total accuracy for LDA is 78% and 77% for SVM. The  $A_z$  for LDA and SVM are 0.84 and 0.87 respectively. It indicates that the morphological and texture information are both useful for mass classification. The performance used LDA and SVM classifier was similar. But Table 5.3 shows that mass classification from combined morphological and texture features yielded better performance than mass classification using only morphological or texture features for both LDA and SVM classifiers.

## 5.5 Conclusion

In this chapter, we investigated level set-based segmentation for mass classification. After segmentation, morphological and texture features were extracted from the boundary and a ribbon region around the boundary. LDA and SVM were investigated for classification. The experiments were performed using a database of 309 clinical mammograms. Experimental results demonstrated that mass classification from combined morphological and texture features yielded better performance than mass classification using only morphological or texture features.

## References

1. A. C. Society (2007) *Breast cancer facts & figures 2007–2008*. In: American Cancer Society (ed), Atlanta
2. Tang J, Rangayyan R, Xu J, Naqa I, Yang Y (2009) Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. *IEEE Trans Inf Technol Biomed* 13(2): 236–251
3. Elter M, Horsch A (2009) CADx of mammographic masses and clustered microcalcifications: A review. *Med Phys* 36(6):2052–2068
4. Liu X, Tang J, Zhang X (2009) A multiscale image enhancement method for calcification detection in screening mammograms. In: Proceedings of IEEE international conference on image processing, pp 677–680, Cairo, Egypt, Nov. 7–10, 2009
5. Chan H, Sahiner B, Helvie M, Petrick N, Roubidoux M, Wilson T, Adler D, Paramagul C, Newman J, Sanjay-Gopal S (1999) Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: an ROC study. *Radiology* 212(3): 817–827
6. Sahiner B, Petrick N, Chan H, Hadjiiski L, Paramagul C, Helvie M, Gurcan M (2001) Computer-aided characterization of mammographic masses: accuracy of mass segmentation and its effects on characterization. *IEEE Trans Med Imaging* 20(12):1275–1284
7. Mudigonda N, Rangayyan R, Desautels J (2000) Gradient and texture analysis for the classification of mammographic masses. *IEEE Trans Med Imaging* 19(10):1032–1043
8. Kilday J, Palmieri F, Fox M (1993) Classifying mammographic lesions using computerized image analysis. *IEEE Trans Med Imaging* 12(4):664–669
9. Shi J, Sahiner B, Chan H, Ge J, Hadjiiski L, Helvie M, Nees A, Wu Y, Wei J, Zhou C (2008) Characterization of mammographic masses based on level set segmentation with new image features and patient information. *Med Phys* 35:280
10. Li H, Kallergi M, Clarke L, Jain V, Clark R (1995) Markov random field for tumor detection in digital mammography. *IEEE Trans Med Imaging* 14(3):565–576
11. Santos V, Schiabel H, Goes C, Benatti R (2002) A segmentation technique to detect masses in dense breast digitized mammograms. *J Digit Imaging* 15:210–213
12. Kupinski M, Giger M (1998) Automated seeded lesion segmentation on digital mammograms. *IEEE Trans Med Imaging* 17(4):510–517
13. Rangayyan R, Mudigonda N, Desautels J (2000) Boundary modelling and shape analysis methods for classification of mammographic masses. *Med Biol Eng Comput* 38(5):487–496
14. Huo Z, Giger M, Vyborny C, Bick U, Lu P, Wolverton D, Schmidt R (1995) Analysis of spiculation in the computerized classification of mammographic masses. *Med Phys* 22:1569
15. Rojas Dominguez A, Nandi A (2009) Toward breast cancer diagnosis based on automated segmentation of masses in mammograms. *Pattern Recognit* 42(6):1138–1148
16. Yuan Y, Giger M, Li H, Suzuki K, Sennett C (2007) A dual-stage method for lesion segmentation on digital mammograms. *Med Phys* 34:4180
17. Liu X, Liu J, Zhou D, Tang J (2010) A benign and malignant mass classification algorithm based on an improved level set segmentation and texture feature analysis. In: Fourth International Conference on Bioinformation and Biomedical Engineering
18. Yezzi AJ, Tsai A, Willsky A (1999) A statistical approach to snakes for bimodal and trimodal imagery. In: Proceedings of international conference on computer vision, pp 898–903
19. Chan T, Vese L (2001) Active contours without edges. *IEEE Trans Image Process* 10(2): 266–277
20. Li C, Kao C, Gore J, Ding Z (2008) Minimization of region-scalable fitting energy for image segmentation. *IEEE Trans Image Process* 17(10):1940–1949
21. Lankton S, Tannenbaum A (2008) Localizing region-based active contours. *IEEE Trans Image Process* 17(11):2029–2039
22. Guliatto D, Rangayyan R, Carvalho J, Santiago S (2008) Polygonal modeling of contours of breast tumors with the preservation of spicules. *IEEE Trans Biomed Eng* 55(1):14–20

23. Liang S, Rangayyan RM, Leo DJE (1994) Application of shape analysis to mammographic calcifications. *IEEE Trans Med Imaging* 13(2):263–274
24. Sahiner B, Chan H, Petrick N, Helvie M, Hadjiiski L (2001) Improvement of mammographic mass characterization using spiculation measures and morphological features. *Med Phys* 28: 1455
25. Haralick R, Dinstein I, Shanmugam K (1973) Textural features for image classification. *IEEE Trans Syst Man Cybern* 3(6):610–621
26. Gonzalez RC, Woods RE (2002) Digital image processing. Prentice Hall, Upper Saddle River, NJ
27. Chang C, Lin C (2001) LIBSVM: a library for support vector machines (ed)
28. Heath M, Bowyer K, Kopans D, Moore R, Kegelmeyer P (2000) The digital database for screening mammography. In: Proceedings of the fifth International Work on Digital Mammography, pp 212–218
29. Metz C, Herman B, Shen J (1998) Maximum likelihood estimation of receiver operating characteristic(ROC) curves from continuously distributed data. *Stat Med* 17(9):1033–1053

## Biography



**Jinshan Tang** was born in Chenzhou, Hunan, P.R. China. He received the Ph.D. degree from Beijing University of Posts and Telecommunications, Beijing, China, in 1998. He joined ATR Media Integration and Communication Research Laboratories, Japan after he graduated. In 2000, he joined Harvard Medical School as a Postdoctoral Researcher. In 2001 he joined the University of Virginia, and was there for about 3 years. He also worked as a senior engineer in Intel and a visiting fellow in NIH. Since 2006, he has been an Assistant Professor at Alcorn State University, Alcorn State, MS.



**Xiaoming Liu** received his Ph.D. degree from Zhejiang University, China, in 2007. He visited the Image Processing and Bio-imaging Research Laboratory, Alcorn State University, USA, as a Postdoctorate Research Associate from 2008 to 2009. Currently he is a lecturer in the College of Computer Science and Technology, Wuhan University of Science and Technology, Wuhan, China.

# **Chapter 6**

## **Segmentation and Skeletonization of 3D Contrast Enhanced Ultrasound Images for the Characterization of Single Thyroid Nodule**

**Filippo Molinari, Alice Mantovani, Maurilio Deandrea, Paolo Limone, Rajendra Acharya U, Roberto Garberoglio, and Jasjit S. Suri**

**Abstract** The thyroid nodules are common findings in clinical practice. Recent studies demonstrated that thyroid nodules can be found in about 66% of the adult population. Nevertheless, only 5–7% of the thyroid nodules is malignant. In the presence of a suspicious nodule, thyroidectomy is performed. As a consequence, the patient will lack the thyroid hormones, which will be integrated by using drugs.

Currently, the differential diagnosis of thyroid nodules is still problematic. The most used technique is the ultrasound examination, but it has been proved that several interpretative limitations of the ultrasound nodule appearance still remain. Color and Power Doppler slightly improved the accuracy of ultrasound characterization, but performance is still inadequate for clinical practice. Other possible diagnostic techniques are MRI and nuclear medicine. MRI is expensive and still not very diffuse for thyroid analysis, while nuclear medicine essentially differentiates cold from hot nodules.

In this chapter, we will describe a novel strategy for the characterization of thyroid nodules based on 3-D ultrasound images acquired after the injection of intravascular contrast agent. The objective of this processing technique is the characterization of the intranodular vascularization, under the hypotheses that malignant nodules had a greater vascularization than benign.

The processing technique consists of three steps: (1) preprocessing and segmentation of the contrast agent distribution in the nodule volume; (2) thinning (by means of a combination of Distance Transform and Ma and Sonka skeleton) of the segmented 3-D volumes; and (3) computation of intranodular vascularization descriptors.

Our results showed that malignancy is associated to a higher vascularization. Malignant nodules had a higher number of vascular trees, which had several branches and a marked tortuosity. Conversely, benign lesions were overall poorly vascularized.

---

F. Molinari (✉)

Biolab, Department of Electronics, Politecnico di Torino, Torino 10129, Italy  
e-mail: filippo.molinari@polito.it, molinari.filippo@gmail.com

This segmentation and skeletonization technique is a first step toward a 3-D contrast agent-based ultrasound technique for the differential diagnosis of thyroid nodules.

**Keywords** Contrast-enhanced ultrasound · Thyroid nodule · Skeleton · Vascular assessment · Tumor

## 6.1 Introduction to the Thyroid Pathology

Thyroid pathologies can be divided in two main classes:

1. Pathologies compromising the thyroid functionality.
2. Pathologies altering the thyroid morphology.

Such pathologies can be further subdivided into pathologies causing hypo-functionality of the thyroid gland, and those causing hyper-functionality.

The nodular pathologies are commonly associated with a hypo-functionality of the thyroid. In a recent review, De Martino et al. [10] showed that asymptomatic thyroid nodules were found in two-third of adult subjects, instead of the 50% detected by clinicians [15]. Nevertheless, only 7% of thyroid nodules are malignant [15].

The most diffused nodular pathologies are the following:

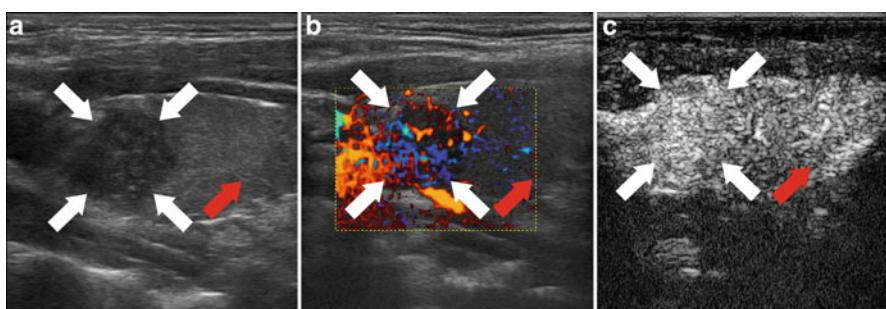
- Nontoxic goiter nodule. It is a nodule that does not cause functional alterations of the thyroid gland. Usually, it is not associated to inflammation and cancer [18,19]. This pathology is coupled to euthyroidism and patients undergo examination when the thyroid mass starts increasing its dimensions. The vascularization of a goiter nodule is similar to that of the normal thyroid parenchyma; larger vessels are present at the periphery of the nodule and only few and small vessels are observed within the nodule. This pathology can cause, in a chronic stage, a multinodular aspect of the thyroid.
- Follicular adenoma. It is the most common type of benign tumor of the thyroid. Usually, it is characterized by the presence of a single nodule in euthyroidism. The average dimension of adenoma nodules is 1–3 cm of diameter [27]. The nodule might protrude from the thyroid tissue, causing inflammation and, in some cases, hemorrhage. There are six variants of this pathology [11]: (1) the embryonic adenoma; (2) the fetal adenoma; (3) the adenoma simplex; (4) the colloid adenoma; (5) the Hürthle cells adenoma; and (6) the atypical adenoma. Cytology is usually the common adopted exam for differential diagnosis of this kind of disease.
- Papillary carcinoma. It is the most common type of thyroid tumor. It has an incidence peak between 20 and 50 years, with a female:male ration of 3:1. This tumor usually originates a single nodule that is palpable. If not properly treated, this tumor rapidly invades the lymph nodes creating metastases. If diagnosed by scintigraphy, the tumor appears as a cold nodule. The degeneration of portions of the nodule in microcalcifications is observed in about 50% of the subjects.

- Follicular carcinoma. It is a follicular malignant pathology with high incidence in the subjects with low iodine intake by diet. It is a soft nodule, with a well-defined capsula, similar to the follicular adenoma (which is a benign lesion). It can be distinguished by the adenoma by relying on the cytology examination, since the follicular carcinoma generates nuclear mitosis. The risk of invasion of the thyroid parenchyma is high and this tumor tends to expand via the vascular tree. If the papillary carcinoma usually expands via the lymphatic circulation, the follicular carcinoma expands via the blood vessels.

There are some other types of nodular thyroid pathology, but they are rare or specific of some classes of subjects. For all the malignant pathologies, early diagnosis is the key for a positive prognosis. For example, when the follicular carcinoma is not infiltrated, subjects have 95% of probability of complete remission, but if it invades the vessels, the percentage drops to 50%.

## 6.2 Thyroid Nodules Diagnosis

High-resolution ultrasonography (HRUS) is the most frequently used tool for identification, assessment and follow-up of thyroid lesions (Fig. 6.1a). This is because it can reveal formations as small as 1 mm, noninvasive, and low cost [25]. Even though it has been demonstrated that malignancy is related to common features in HRUS B-Mode images (i.e., microcalcifications, marked hypoechoicity, irregular margins, and the absence of a hypoechoic halo around the nodule [5]), interpretative pitfalls remains. Recently, D’Souza et al. measured



**Fig. 6.1** (a) HRUS image of a papillary carcinoma. The white arrow indicates the nodule; the red arrow indicates the healthy thyroid parenchyma. (b) Color Doppler image of the papillary carcinoma showing marked vascularization of the lesion (white arrows) compared to healthy parenchyma (red arrow). (c) CEUS image of the papillary carcinoma showing marked enhancement of the lesion (white arrows) compared to healthy parenchyma (red arrow). All three images are same projection and slice

the HRUS sensitivity and specificity in 80.8 and 81.6%, respectively [9], thus demonstrating limited performance in detecting malignancy.

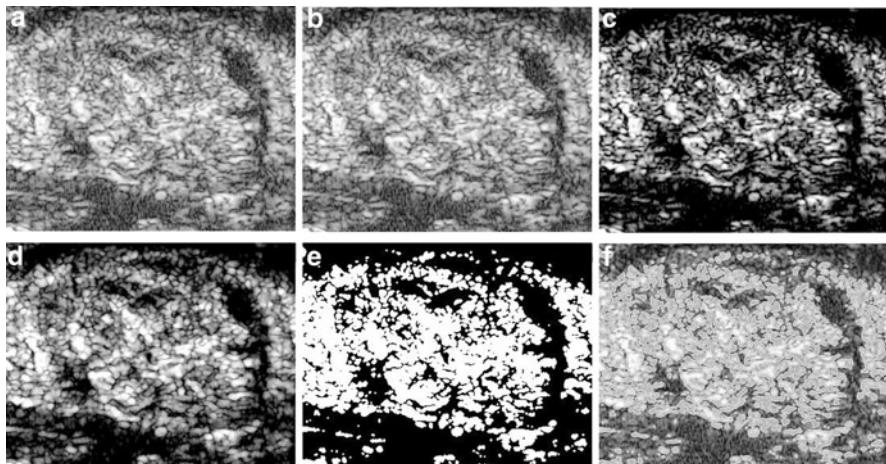
Color Doppler ultrasound imaging (Fig. 6.1b) has been used to differentiate thyroid nodules. Results showed that benign nodules tend to have no or minimal internal flow with presence or absence of a peripheral ring, whereas malignant nodules tend to have a peripheral ring with an extensive internal flow [17]. These studies evidenced the need for a quantitative evaluation of the internal nodule flow, to avoid subjective interpretations and partial visions.

Contrast-enhanced ultrasound imaging (CEUS) has been introduced to improve the differential diagnosis of thyroid nodules. The contrast agent enhances the parenchyma representation, given the abundant vasculature of the thyroid gland. Therefore, by using CEUS, it should be possible to gain a better representation of the thyroid vascular pattern, overcoming the limitations given by the low spatial resolution of the Color Doppler imaging (Fig. 6.1c). CEUS studies of the thyroid nodules, conducted by analyzing the increase in the echogenicity of the nodule due to perfusion, did not prove more performing than traditional HRUS [4]. On the contrary, CEUS perfusion analysis of neoplastic nodules proved effective in differentiating benign from malignant lesions, even if performances were poor when compared to fine needle aspiration examinations [1]. CEUS enhancement patterns were found different in benign and malignant lesions, and ring enhancement was indicative of benign lesions, whereas heterogeneous enhancement of malignancy [30].

In this chapter, we will cover a methodological description of a procedure for the 3D analysis of thyroid nodules and for an objective representation and numerical characterization of the 3D nodular vascularization. Earlier studies we conducted [7,23] showed that Color and Power Doppler imaging were not suitable to 3D micro-vessel detection, due to poor spatial resolution and to color blooming that may occur in highly perfused nodules (blooming can be seen in Fig. 6.1b). We used CEUS and developed a processing technique to represent the 3D internal flow that we used to characterize 14 solitary thyroid nodules. This methodology represents a step forward with respect to an earlier study we started in 2008 [7]. In our former procedure, the characteristics of the 3D CEUS images were qualitatively scored by an expert sonographer, whereas in this contribution we will show our numerical approach to describe the internal nodular vasculature.

### 6.3 CEUS Image Processing Strategy

The image processing system (ThyScreenCEUS) we developed was structured in three main steps: (a) image preprocessing; (b) morphological aperture for contrast agent bubbles detection; (c) thresholding and 3D lesion reconstruction. Figure 6.2 depicts the steps of the algorithm on a malignant thyroid nodule (papillary carcinoma).



**Fig. 6.2** Processing of a single frame of the 3D volume of a thyroid nodule. (a) Raw image; (b) image after the application of a linear ramp to compensate the depth artifact; (c) intensity normalized image; (d) morphological opening with a disk-shaped structuring element of 5 pixel radius; (e) segmentation mask obtained by thresholding; (f) contour of the segmentation overlapped to the original image

### 6.3.1 Preprocessing

All the images of the 3D volume were first converted into double precision format. When we insonated a homogeneous phantom with the same probe and the same settings that were used in the study, we found a linear dependence of the average gray level from depth (i.e., from the vertical dimension of the image). A linear ramp with decreasing intensity from 1 to 0 was applied to the original image (Fig. 6.2a) in order to attenuate the intensity artifact that was present in the distal (i.e., deeper) portion of the images (the result after intensity correction in Fig. 6.2b).

To make the algorithm as independent as possible on the acquisition scanner/probe and on the optical integrated density of the image (due to different concentrations of the contrast agent in the 3D slices), each image was normalized. Non-perfuse zones were mapped to black; highly perfuse regions were mapped to white. This step required the manual selection of two  $3 \times 3$  ROIs on an image of the 3D volume.

The images were then low-pass filtered by using a  $7 \times 7$  Gaussian kernel with standard deviation of 0.1 pixels. Finally, speckle noise was attenuated by using the first order local statistic filter, defined by the following equation:

$$J_{x,y} = \bar{I} + k_{x,y}(I_{x,y} - \bar{I})$$

where  $I_{x,y}$  is the intensity of the noisy pixel,  $\bar{I}$  is the mean intensity of a  $N \times M$  pixel neighborhood, and  $k_{x,y}$  is a local statistic measure. The noise-free pixel is indicated by

$J_{x,y}$ . We followed the proposal by Loizou et al. [20], who defined  $k_{x,y} = \frac{\sigma_I^2}{I^2\sigma_I^2 + \sigma_n^2}$ , where  $\sigma_I^2$  represents the variance of the pixels in the neighborhood, and  $\sigma_n^2$  the variance of the noise in the cropped image. We found that a suitable neighborhood size was  $7 \times 7$ . The normalized, filtered and despeckled image is reported in Fig. 6.2c.

### 6.3.2 Micro-Bubble Detection Using Morphological Aperture

To detect the micro-bubbles in the acquired volume, we processed every 2D slice separately. The images had an axial resolution of  $62.5 \mu\text{m}/\text{pixel}$  and the contrast agent was constituted by micro-bubbles of less than  $10 \mu\text{m}$  diameter [14]. Hence, we used a structuring element consisting of a disk with radius of 5 pixels, which gave maximum sensibility on groups of micro-bubbles. The images were then morphologically opened (Fig. 6.2d).

### 6.3.3 Threshold Process and 3D Reconstruction

Threshold process was used to create a mask for each volumetric slice. At this stage, we considered the image histogram as composed of two Gaussians, one relative to the foreground (i.e., the presence of micro-bubbles) and the other relative to the background (i.e., the absence of micro-bubbles). Thus, this is a two-class problem. Since a threshold can separate the statistical distribution of the two classes, Otsu's method was adapted. The threshold was made local since it was computed separately for each frame of the 3D volume. Figure 6.2e, f depicts the obtained segmentation mask and the final segmentation, respectively.

The original 3D volume was then multiplied by the segmentation masks and was reconstructed in order to spatially visualize the distribution of contrast agent in the thyroid lesion (we used OsiriX software – <http://www.osirix-viewer.com/> by the Osirix Foundation, Genève, Switzerland).

## 6.4 Distance Transform, Skeleton Techniques, and Skeleton Characteristic Parameters

In this section, we describe our original procedure for the skeletonization of the segmented 3D CEUS images. We coupled a distance transform based technique, to the Ma and Sonka skeletonization architecture, in order to preserve the connectivity of the images. In this specific case, in fact, since we are using skeletons to reconstruct the 3D vascular tree of a lesion, the preservation of connectivity is crucial for a proper characterization of the lesion itself.

### 6.4.1 *Skeletonization Strategy*

The skeletonization procedure aims at reducing the amount of information in an image into a simpler image that still preserve morphology. In other words, the skeleton is a technique for extracting the innermost and essential information of an image [21].

We applied skeletonization technique to our images to reduce the representation of the contrast agent in the nodule. Specifically, since the spatial resolution of our probe was more than ten times lower than the micro-bubbles diameter, we could only track micro-bubbles that were grouped together. Hence, the contrast agent diffusion in the thyroid was still over-represented and little useful from a clinical point of view. We adopted a skeletonization procedure that ensured an effective reduction of the vessels appearance while preserving connectivity.

We adopted the skeletonization technique proposed by Ma and Sonka [21]. This algorithm was a 3D technique for thinning and reconstruction of binary images that was applied to medical images. Basically, this algorithm defined four classes (called A–D in the original algorithm) that were used to derive templates. Each class was translated, so that a total number of 38 templates could be derived (6 templates from class A, 12 from B, 8 from C, and 12 from D). A template was a cube of  $3 \times 3 \times$  pixels. In each voxel of the template, there was a coded number that specified the presence of the background (mapped by the 0), of the foreground (mapped by 1), or a “don’t care” voxel (mapped by –1). So, the 38 templates test each point of the image (or volume) and the point was deleted if it matched one of the templates. Full details about the Ma and Sonka algorithm [21] can also be found in the study by Palàgyi and Kuba [24].

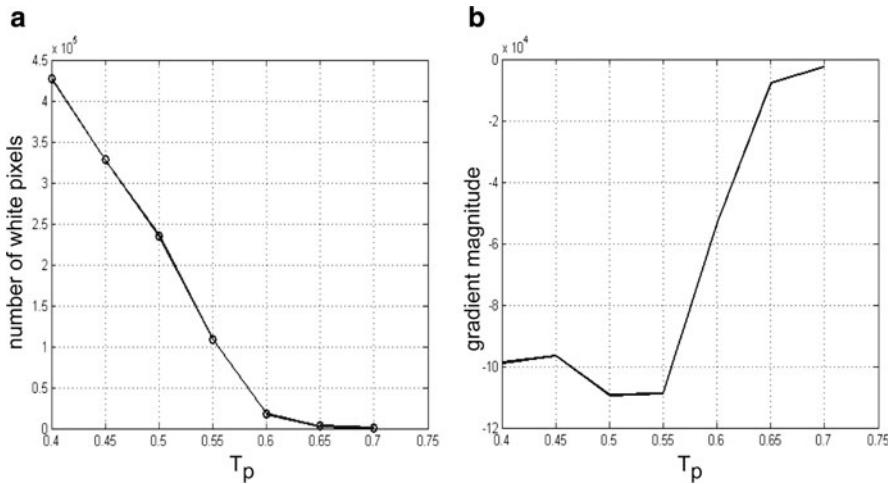
Previous studies showed that this skeleton technique does not preserve connectivity [29]. In our study, however, connectivity was the key to correctly reconstruct the vascularization of a lesion. Therefore, we modified the Ma and Sonka algorithm by inserting a preliminary skeletonization based on the distance transform (DT).

The DT of a point belonging to an object is defined as its lower distance from the borders of the object. Therefore, points located in the middle of a vessel have high values of DT, whereas points located near the vessel borders have low DT values. Let DTI be an image containing the DT of each point. A point is assigned to the skeleton if its distance from the borders differs from the average distance of its neighborhood less than a predefined threshold  $T_p$ . We can express the definition of such skeleton by using the following formula:

$$DSk = DTI_F < (DTI - T_p) \quad (6.1)$$

where DS<sub>k</sub> is the skeleton image and DTI<sub>F</sub> is the DTI image after a low-pass filtering with a 3D  $10 \times 10 \times 10$  Gaussian kernel (with sigma value equal to 1).

To select the optimal value of  $T_p$ , we randomly selected four nodules (two carcinomas and two goiter nodules) from the image database. The images were skeletonized using  $T_p$  values equal to 0.40, 0.45, 0.50, 0.55, 0.60, 0.65, and 0.70. After having obtained the skeleton, we counted the total number of white points



**Fig. 6.3** (a) The graph shows the average number of points belonging to the skeleton of four randomly selected nodules (two malignant and two benign) with respect to the selectivity threshold  $T_p$ . The curve shows an inflection for middle values of  $T_p$ . (b) The gradient of the image shows a marked increase after 0.55. The  $T_p$  values considered as optimal was chosen equal to 0.55

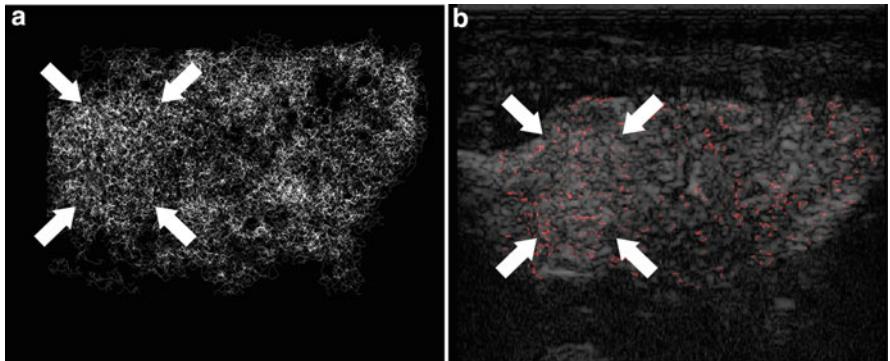
(i.e., the total number of pixels constituting the skeleton). We averaged the computed number of white points on the four nodules (Fig. 6.3a). The curve showed an inflection for middle values of  $T_p$ . We also computed the fifth-order polynomial curve fitting the white points  $T_p$  curve (Fig. 6.3b). The gradient showed a marked increase for a  $T_p$  value of 0.55. Therefore, we selected 0.55 as optimal value for our skeletonization procedure.

Qualitative results repeated on all the remaining images confirmed that skeletons with higher  $T_p$  values were poor of information (i.e., very thinned) and computationally heavy. Conversely, skeletons obtained with lower  $T_p$  values resulted in too noisy representations, even if computation time was reduced.

The summary of skeletonization procedure consisted of the following steps:

1. DT was applied to the 3D CEUS image. This step roughly removed the smallest objects and cleaned the points located far from the middle of the vessels.
2. Application of the Ma and Sonka [21] skeleton. This technique considered five consecutive frames to preserve connectedness (i.e., for each frame, the two preceding and the two successive frame, thus a total of five). The points that matched one or more of the 38 templates were removed and the skeletonized frame was saved.
3. The step 2 was repeated for each frame until there were no more points to remove.

Basically, the DT thinning procedure, which is faster even if less accurate, allowed for a quick removal of small objects and points far from the vessel



**Fig. 6.4** (a) Maximum intensity projection (MIP) of the 3D skeleton. (b) CEUS image of a papillary carcinoma (background) with corresponding skeleton overlaid in red (for a single slice). The *white arrows* indicate the nodule position. The lesion shows increased vascularity with respect to healthy thyroid tissue

centerline, whereas the Ma and Sonka skeleton enabled an accurate reduction of the points. The use of DT prior to the application of the Ma and Sonka skeleton also reduced the computational cost of the algorithm. Figure 6.4 reports a sample result of the skeleton technique, showing a nodular vascularization greater than that of normal tissue.

#### 6.4.2 Skeleton Characteristics

The skeleton characteristics were analyzed and computed in order to provide a quantitative measure of the thyroid vascularization. In fact, considering the skeleton as the minimal representation of the thyroid lesion vascularization, we depicted that the skeleton characteristics modeled the lesion morphology.

The skeleton is a binary 3D image containing “1” in correspondence of the contrast agent (i.e., therefore of a vessel) and “0” otherwise. Hence, the skeleton can be thought of as a set of paths, some of which are connected. Many techniques were proposed to analyze the skeleton characteristics [8,16]. We used the Breadth First Search (BFS). The BFS technique consists in searching all the skeleton nodes starting from a given node. The procedure iterates on all the nodes and stops when all the nodes have been visited. The BFS algorithm maps the skeleton as a series of vascular trees ( $T$ ):

$$\text{DSk} = \{T_1, T_2, \dots, T_N\} \quad (6.2)$$

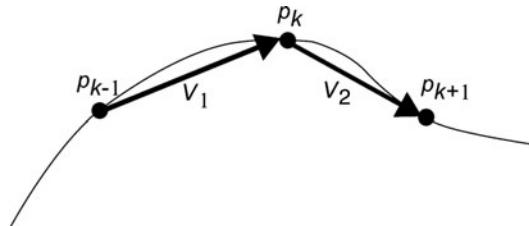
where  $N$  is the number of the vascular trees. Each vascular tree  $T_i$  can be thought as the sequence of  $m$  nodes  $p_j : T_i = \{p_1, p_2, \dots, p_m\}$ .

By using the BFS we computed the following skeleton characteristics [16]:

1. *Number of vascular trees* (NT): defined as the number of trees in which the BFS decomposes the skeleton image. This parameter is important since we expect a higher number of vascular trees in malignant lesions compared to benign lesions.
2. *Density of the vascular structure* (VD): defined as the number of white points belonging to the skeleton ( $M$ ) over the size of the ROI. Considering a ROI of dimensions  $h$ ,  $l$ , and  $w$ , we can thus define  $VD = M/(h \cdot l \cdot w)$ . This parameter is expected to be higher in malignant lesions than in benign.
3. *Number of branching nodes* (NB): defined as the number of branches of the structure. Again, given the characteristics of the malignant neovascularization, we expect that NB is higher for tumors than for nonmalignant nodules.
4. *Average vessel radius* (MR): defined as the average radius of the vessels in the ROI. Once a set of connected vessels has been defined, the average vessel radius is calculated for each vessel by summing the radius defined at each vessel skeleton point and by dividing by the number of points. The vessels created by a tumor should have lower MR than normal vessels.
5. *2D vascular tortuosity (distance metrics – DM)*: this is a measure of the tortuosity of a vascular tree  $T_i$ . DM can be defined as the Euclidean distance between the first and the last point of a tree divided by its real length. Let's consider a tree consisting of  $m$  nodes:  $T_i = \{p_1, p_2, \dots, p_m\}$ . The DM can be defined as:  $DM = \frac{|p_m - p_1|}{\sum_{i=1}^{m-1} |p_i - p_{i+1}|}$ , where  $|\cdot|$  denotes the Euclidean distance measure.
6. *Inflection count metric* (ICM): defined as DM divided by the number of flexes in the path. This parameter is an estimate of the average number of times the vascular tree changes its shape from convex to concave, and vice versa.
7. *3D vascular tortuosity (sum of angles metric – SOAM)*: this parameter is used to evaluate the vessel torsion in 3D. With reference to Fig. 6.5, let's consider a node  $p_k$ , and the two adjacent nodes  $p_{k-1}$  and  $p_{k+1}$ . Let's  $V_1$  and  $V_2$  be the vectors connecting the three nodes as in Fig. 6.5. The planar angle  $IP_k$  between the vectors  $V_1$  and  $V_2$  can be expressed as

$$IP_k = \cos^{-1} \left( \left( \frac{V_1}{|V_1|} \right) \cdot \left( \frac{V_2}{|V_2|} \right) \right).$$

**Fig. 6.5** Schematic representation of the SOAM computation. The vectors  $V_1$  and  $V_2$  define the planar angle  $IP_k$ , relative to the node  $p_k$ . By extending such computation to the 3D geometry and then to all the  $m$  nodes of the path, it is possible to compute the SOAM



The angle defining the torsion of the curve in a node  $p_k$  can be expressed as

$$\text{TP}_k = \cos^{-1} \left( \left( \frac{\mathbf{V}_1 \times \mathbf{V}_2}{|\mathbf{V}_1 \times \mathbf{V}_2|} \right) \cdot \left( \frac{\mathbf{V}_2 \times \mathbf{V}_3}{|\mathbf{V}_2 \times \mathbf{V}_3|} \right) \right),$$

where  $\mathbf{V}_3$  is the vector between the nodes  $p_{k+1}$  and  $p_{k+2}$ . The total angle in a node  $p_k$  is then defined as  $\text{CP}_k = \sqrt{\text{IP}_k^2 + \text{TP}_k^2}$ . Hence, the SOAM can be computed by averaging the total angle on all the nodes of the path

$$\text{SOAM} = \frac{\sum_{k=1}^{m-3} \text{CP}_k}{\sum_{k=1}^{m-1} |p_k - p_{k+1}|}$$

where  $m$  is the number of nodes of the considered path.

The skeletonization procedure was applied to the entire thyroid volume. We manually selected a cubic ROI of  $90 \times 90 \times 20$  pixels of size centered on the lesion and we computed the seven skeleton characteristics only in the traced ROI. The dimension of the ROI was chosen by considering the average size of a thyroid lesion, which resulted in about  $150 \times 150$  pixels and in a number of slices ranging between 20 and 35. Our ROI was smaller than the average nodule size, so that it comprised a significant portion of the lesion without incorporating parenchyma pixels. This ROI size was compatible with all the lesions in our database. The next section presents the results of the application of these skeleton characteristics.

## 6.5 Nodules Selection and Their Characterization

We enrolled in this study 30 patients with a previous diagnosis of thyroid nodule. We excluded patients having multinodule goiter. We preselected subjects showing HRUS features of malignancy: intranodular microcalcifications, hypoechoic appearance, and irregular margins [5]. Patients were thus nonconsecutive. All the patients signed an informed consent prior of participating to the experiment. The experimental protocol was approved by the ethical committee of the Endocrinology Section of the “Umberto I” Hospital of Torino (Italy).

All the subjects underwent a clinical examination, hormonal profile, and ultrasound (HRUS and Color Doppler) examination of the lesion. Then, a bolus of 2.5 ml of ultrasound contrast agent (Sonovue, Bracco Imaging, Milano, Italy) was administered intravenously and a gray scale 3D volume containing the lesion was acquired. Due to bulkiness and weight of external mechanical scanning systems and the variability associated with the nodules dimension and its position, we preferred to perform a freehand scanning. A trained operator with more than 30 years of experience in neck ultrasonography (R.G.) performed all the scans 50 s after the

injection of the contrast agent bolus. The freehand scanning was performed by sliding the probe along the patient's neck, in order to cover the nodule extension. The high frame rate of the device compared to the slow movement of the probe ensured that there was no gap between adjacent frames. The average frame rate of the device during acquisitions was 16 frames/s.

Images were acquired by a MyLab70 ultrasound scanner (Biosound-Esaote, Genova, Italy) equipped by a LA-522 linear probe working in the range 4–10 MHz. All the images were acquired at 5 MHz. We calibrated gain by acquiring an empty volume (probe in air) and in order to set the background average intensity to be equal to 5 in a linear scale 0–255. Gain was constant for all patients. The TGC curve was set to flat in the middle of the scale and kept constant. The volumes were transferred in DICOM format to an external workstation Apple PowerPc (Apple, Cupertino, CA), dual 2.5 GHz, 8 GB of RAM, and equipped with processing and reconstruction software.

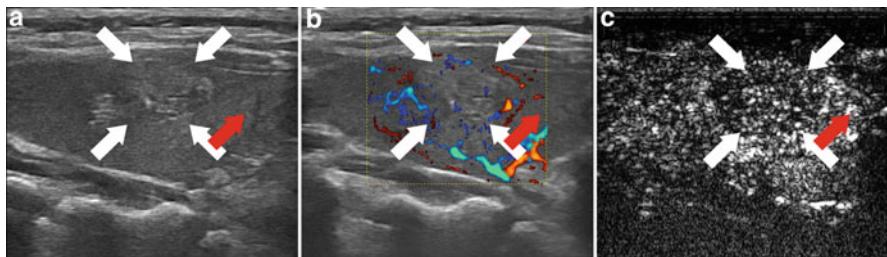
All the subjects underwent ultrasound-guided fine needle aspiration biopsy (FNAB) of the thyroid lesion. Five nodules resulted benign (goiter nodules) and were classified as THY2 according to cytological criteria [22]. In 25 nodules, cytology suggested a follicular neoplasm (THY3). These patients underwent thyroidectomy. We discarded from the study five patients who had a nodule of diameter greater than 6 cm, which was twice the average size of the other nodules, and which caused problems in the manual scanning. Three patients were discarded due to motion artifacts during the CEUS examination (swallowing and cough). Two patients were discarded since FNAB demonstrated a concomitant thyroiditis. Of the 15 THY3 patients, 5 patients resulted in benign (follicular neoplasm) and 10 malignant (seven papillary, one follicular and two Hürthle cells carcinoma). Overall, we had ten benign and ten malignant lesions.

The results of the FNAB (or of the histopathological analysis for surgically removed tumors) were kept blinded to the researchers until all the image acquisitions and processing were completed.

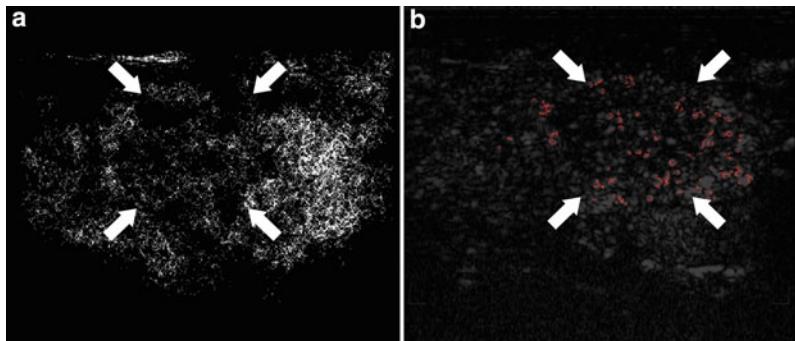
Ten subjects were males (age:  $53.5 \pm 13.3$  years; range: 22–71 years) and ten females (age:  $50.1 \pm 10.8$  years; range: 25–68 years). The average size of the nodules was  $31.7 \pm 17.9$  mm, range: 10–52 mm (malignant:  $34.4 \pm 18.8$  mm; benign:  $29.3 \pm 17.3$  mm).

The average perfusion time-to-peak of the lesions was  $40.8 \pm 12.3$  s. The intensity enhancement caused by the contrast agent was stable for about 20–30 s. The average scanning time of a 3D volume was 3.5 s, thus considerably fast if compared to the duration stability of the intensity enhancement. This ensured an optimal representation of the intra-nodular intensity enhancement.

Figure 6.6 reports the HRUS (Fig. 6.6a), Color Doppler (Fig. 6.6b) and CEUS (Fig. 6.6c) of a benign nodule (goiter nodule). Figure 6.7 reports the 3D reconstruction of the skeleton (Fig. 6.7a) and the overlay of a single slice of the skeleton on the corresponding CEUS image (Fig. 6.7b). The benign lesion shows an enhancement lower than the surrounding healthy tissue. The skeleton of the malignant nodule (Fig. 6.8a, b) is clearly richer than that of the benign (Fig. 6.8c, d). The malignant lesion has a greater NT value (83.1 against 18.1 of the benign), a more dense



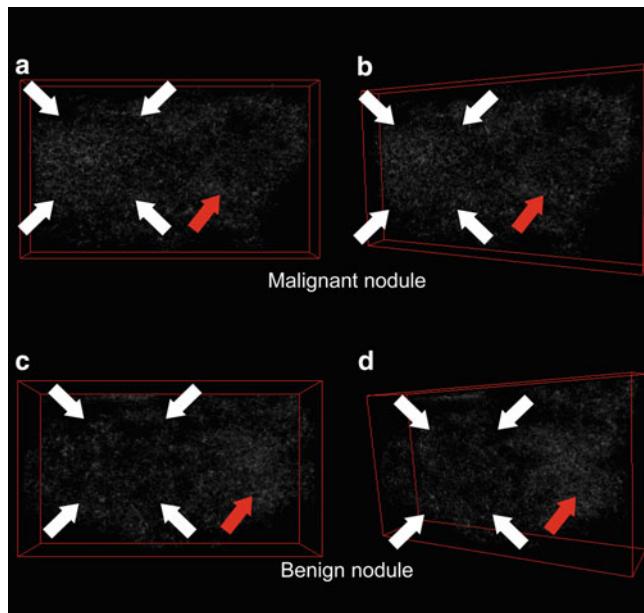
**Fig. 6.6** (a) HRUS of a benign lesion (goiter nodule). The white arrows indicate the position of the lesion; the red arrow indicates healthy thyroid parenchyma. (b) Color Doppler image of the goiter nodule showing a weak vascularization (white arrows) compared to healthy tissue (red arrow). (c) CEUS image of the goiter nodule. The lesion shows a low enhancement (white arrows) if compared to healthy parenchyma (red arrow)



**Fig. 6.7** (a) Maximum intensity projection of the 3D skeleton. (b) CEUS image of a goiter nodule with corresponding skeleton overlaid in red. The white arrows indicate the nodule position. The lesion shows lower vascularity than healthy thyroid tissue

vascular structure (VD for the tumor is equal to 0.04 against 0.01 of the benign lesion), a higher NB (1,453 against 552), and increased values of DM (51 against 18), ICM (19.9 against 8.7), and SOAM (26 against 11). The average vessel radius MR resulted almost equal (1.8 for both).

Table 6.1 reports the average skeleton values measured on the sample population, showing that the malignant nodules had higher values of NT, VD, NB, and SOAM. The MR, DM, and ICM parameters resulted higher in carcinomas, but not at a significant level. We performed an outlier removal procedure (based on the Mahalanobis distance and the  $T^2$  Hotelling test) and we found no outliers in the malignant and one outlier in the benign group. We removed the outlier and reran the Mann–Whitney U-test. This experiment showed that DM and ICM parameters became statistically different between the malignant and benign populations. The new values of DM and ICM after outlier removal were  $17.9 \pm 10.9$  and  $8.6 \pm 6.5$ , respectively.



**Fig. 6.8** (a, b) 3D skeleton representation of a malignant nodule (papillary carcinoma). The *white arrow* indicates the nodule position, whereas the *red arrow* indicates a portion of healthy thyroid parenchyma of the same lobe. Note that the nodule has a brighter appearance with respect to normal tissue, thus indicating increased vascularity. (c, d) 3D skeleton of a benign nodule (goiter). Note that the nodule has a lower enhancement with respect to normal tissue, thus showing a weak vascularization

The algorithms for preprocessing, skeletonization, and computation of the skeleton characteristics were implemented in MATLAB environment (TheMatWorks, Natick, MAUSA) and required on average  $100 \pm 12$  min of processing. We are in a process of porting the scientific system toward faster commercial version.

## 6.6 Possible Role of Computer-Aided Diagnosis and CEUS Imaging in Thyroid Nodules Characterization and Future Perspectives

We presented a strategy for the processing of 3D contrast-enhanced ultrasound images of single thyroid lesions. Morphological operations were coupled to a thinning technique to obtain a minimal representation of the lesion vasculature. We preferred to use manual instead of mechanically driven scanning. This is mainly due to: (a) almost negligible bulkiness and (b) relatively higher cost of the motor controlled systems and maintenance.

**Table 6.1** Average values of the seven skeleton characteristics measured on the sample population of ten patients with malignant nodules (first row) and ten patients with benign nodules (second row)

	NT	VD	NB	MR	DM	ICM	SOAM
Malignant	76.0 ± 38.8 <sup>a</sup>	0.03 ± 0.01 <sup>a</sup>	1,385.0 ± 796.7 <sup>a</sup>	1.9 ± 1.1	47.0 ± 24.7	20.5 ± 11.7	25.9 ± 13.4 <sup>a</sup>
Benign	33.2 ± 27.4	0.01 ± 0.01	659.0 ± 982.8	1.8 ± 1.1	20.8 ± 18.4	9.5 ± 6.2	10.8 ± 7.3

Skeleton parameters were *NR* number of vascular trees, *VD* density of vascular structure, *NB* number of branches, *MR* average vessel radius, *DM* 2D tortuosity, *ICM* inflection count metric, *SOAM* 3D vascular tortuosity

<sup>a</sup>The corresponding skeleton characteristic value for malignant lesions was statistically different when compared to benign (Mann–Whitney *U*-test, *p* < 0.05)

From a technical point of view, our procedure proved robust in processing the 3D CEUS images and in extracting the corresponding skeleton. We used a combination of distance transform and Ma and Sonka method [21] to optimize the thinning algorithm of our volumes. We also optimized the threshold parameter  $T_p$  in order to ensure the optimal representation of the nodule inner vasculature without allowing disconnected patterns. The value of  $T_p$  was critical. If the  $T_p$  value was low, a point was assigned to the skeleton if its DT value is slightly higher than that of its neighborhood. Conversely, if the value of  $T_p$  was high, a point is assigned to the skeleton if and only if the DT value is substantially higher than that of its neighborhood. Hence, higher  $T_p$  values produce a very selective skeleton, resulting in a strong thinning of the image. The probability of obtaining disconnected points increases when the  $T_p$  value increases. We kept  $T_p$  value same for all the subjects.

This procedure was semiautomatic. Two points were manually selected: darkest and brightest spots in the image, representing the presence (perfused zone) and absence (non-perfused zone) of the contrast agent. These two extreme points were used for normalization or scaling of the intensities in the rectangular ROI manually drawn by the trained operator (Endocrinologist and sonographer – R.G., with 30 years of experience). A possible bias from this step could lead into erroneous ROI determination, which as a result could affect the normalization stage. An incorrect normalization could therefore affect the region-based segmentation. This rectangular ROI stage of the procedure is still user-dependent since the relatively low signal-to-noise ratio of the images posed some difficulties in automating this step.

The nodule in the ultrasound image was manually delineated too. Manual tracing of the boundaries was made by the same expert sonographer who performed the acquisitions. Also, the images of the patients undergoing thyroidectomy were further evaluated by surgeons. This was a double check and helped in minimizing the tracing the errors. ROIs were manually traced too, but it was simple to place them in the middle of the lesion once the expert had outlined the nodule. The authors are now working towards the development of a completely automated processing procedure.

The skeleton parameters measured on malignant lesions were clearly different from those of benign ones. After the outlier rejection removal, the only parameter that did not prove statistically different was MR (i.e., the average radius of the vessels). Therefore, the average radius of the vessels did not differ between the benign and the malignant lesions. This result is in accordance to studies based on the Color Doppler technique, which did not show evidence on differences in the vessels size in malignancy [6].

Conversely, malignant lesions showed a higher number of the vascular trees (NT) and the number of branches (NB). Also, the density of the vascular structure (VD) was higher than in benign lesions. These findings suggest that malignant lesions are highly perfused by a dense vascular bed. Such vessels also originate several branches, thus increasing the overall density of the vasculature in the lesion. This result confirms previous findings. Fukunari et al. showed that follicular carcinomas can be diagnosed by using Color Doppler given their high degree of internal vascularization [13]. Bakhshaei et al. found that dense internal vasculature is highly correlated to malignancy [3]. Turner et al., in their review, pointed out

that many studies confirmed the increased vascularization of thyroid tumors with respect to normal tissue and benign lesions [28].

The use of CEUS in the diagnosis of thyroid nodules is a promising technique, even though results are still discordant. Spiezia et al. conducted a perfusion study based on Color Doppler imaging after contrast agent injection [26]. They showed that carcinomas had a lower time-to-peak of the perfusion curve, but that the peak intensity and the washout were not statistically different from benign nodules. Conversely, in another study, the time-intensity CEUS curves did not prove useful in differentiating between benign and malignant nodules [12]. Argalia et al. demonstrated that the CEUS time-intensity curves can provide an indirect description of the intranodular vascularization [2]. They observed that 89% of malignant nodules (characterized by an “anarchical” vascularization) had polyphasic wash-out curves, whereas benign lesions showed a monophasic curve. These studies demonstrate the diagnostic possibilities of CEUS, but also manifest the need for numerical evaluation of the vascular pattern of the lesions.

Overall, our results are in agreement with the above-referenced studies. We obtained a clear indication of “anarchical” vascularization of malignant nodules, which we demonstrated and quantified by higher values of the skeleton parameters.

Whilst the dimension of our sample was relatively low, this study gave encouraging results. We are currently working to enlarge our sample population in order to (a) enhance the statistical significance of our results; (b) derive a classification scheme based on the skeleton parameters.

**Acknowledgement** The study was granted by the Fondazione Scientifica Mauriziana ONLUS of Torino.

## References

- Appetecchia M, Bacaro D, Brigida R, Milardi D, Bianchi A, Solivetti F (2006) Second generation ultrasonographic contrast agents in the diagnosis of neoplastic thyroid nodules. *J Exp Clin Cancer Res* 25:325–330
- Argalia G, De Bernardis S, Mariani D, Abbattista T, Taccaliti A, Ricciardelli L, Faragona S, Gusella PM, Giuseppetti GM (2002) Ultrasonographic contrast agent: evaluation of time-intensity curves in the characterisation of solitary thyroid nodules. *Radiol Med* 103:407–413
- Bakhshaei M, Davoudi Y, Mehrabi M, Layegh P, Mirsadaee S, Rad MP, Leyegh P (2008) Vascular pattern and spectral parameters of power Doppler ultrasound as predictors of malignancy risk in thyroid nodules. *Laryngoscope* 118:2182–2186
- Bartolotta TV, Midiri M, Galia M, Runza G, Attard M, Savoia G, Lagalla R, Cardinale AE (2006) Qualitative and quantitative evaluation of solitary thyroid nodules with contrast-enhanced ultrasound: initial results. *Eur Radiol* 16:2234–2241
- Bastin S, Bolland MJ, Croxson MS (2009) Role of ultrasound in the assessment of nodular thyroid disease. *J Med Imaging Radiat Oncol* 53:177–187
- Bozbora A, Erbil Y, Ozarmagan S, Barbaros U, Sari S, Degirmenci B (2002) Color Doppler sonography in cold thyroid nodules for malignancy prediction. *Acta Chir Belg* 102:259–262

7. Carraro R, Molinari F, Deandrea M, Garberoglio R, Suri JS (2008) Characterization of thyroid nodules by 3-D contrast-enhanced ultrasound imaging. *Conf Proc IEEE Eng Med Biol Soc* 2008:2229–2232
8. Chen Z, Molloj S (2002) Vascular tree object segmentation by deskeletonization of valley courses. *Comput Med Imaging Graph* 26:419–428
9. D’Souza MM, Marwaha RK, Sharma R, Jaimini A, Thomas S, Singh D, Jain M, Bhalla PJ, Tripathi M, Tiwari A, Mishra A, Mondal A, Tripathi RP (2010) Prospective evaluation of solitary thyroid nodule on 18F-FDG PET/CT and high-resolution ultrasonography. *Ann Nucl Med* 24:345–355
10. De Martino E, Pirola I, Gandossi E, Delbarba A, Cappelli C (2008) Thyroid nodular disease: an emerging problem. *Minerva Endocrinol* 33:15–25
11. Dubreau L (2008) The cell of origin of ovarian epithelial tumours. *Lancet Oncol* 9:1191–1197
12. Friedrich-Rust M, Sperber A, Holzer K, Diener J, Grunwald F, Badenhoop K, Weber S, Kriener S, Herrmann E, Bechstein WO, Zeuzem S, Bojunga J (2010) Real-time elastography and contrast-enhanced ultrasound for the assessment of thyroid nodules. *Exp Clin Endocrinol Diabetes* 118:602–609
13. Fukunari N, Nagahama M, Sugino K, Mimura T, Ito K (2004) Clinical evaluation of color Doppler imaging for the differential diagnosis of thyroid follicular lesions. *World J Surg* 28:1261–1265
14. Gorce JM, Ardit M, Schneider M (2000) Influence of bubble size distribution on the echogenicity of ultrasound contrast agents: a study of SonoVue. *Invest Radiol* 35:661–671
15. Hoang JK, Lee WK, Lee M, Johnson D, Farrell S (2007) US features of thyroid malignancy: pearls and pitfalls. *Radiographics* 27:847–860, discussion 61–65
16. Huang SF, Chang RF, Moon WK, Lee YH, Chen DR, Suri JS (2008) Analysis of tumor vascularity using three-dimensional power Doppler ultrasound images. *IEEE Trans Med Imaging* 27:320–330
17. Ivanac G, Brkljacic B, Ivanac K, Huzjan R, Skreb F, Cikara I (2007) Vascularisation of benign and malignant thyroid nodules: CD US evaluation. *Ultraschall Med* 28:502–506
18. Lewinski A (2002) The problem of goitre with particular consideration of goitre resulting from iodine deficiency (I): classification, diagnostics and treatment. *Neuro Endocrinol Lett* 23:351–355
19. Lewinski A (2002) The problem of goitre with particular consideration of goitre resulting from iodine deficiency (II): management of non-toxic nodular goitre and of thyroid nodules. *Neuro Endocrinol Lett* 23:356–364
20. Loizou CP, Pattichis CS, Christodoulou CI, Istepanian RS, Pantziaris M, Nicolaides A (2005) Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery. *IEEE Trans Ultrason Ferroelectr Freq Control* 52:1653–1669
21. Ma MC, Sonka M (1996) A fully parallel 3D thinning algorithm and its applications. *Comput Vis Image Underst* 64:420–433
22. Mihai R, Parker AJ, Roskell D, Sadler GP (2009) One in four patients with follicular thyroid cytology (THY3) has a thyroid carcinoma. *Thyroid* 19:33–37
23. Molinari F, Carraro R, Deandrea M, Garberoglio R (2008) Skeletonization of 3-D contrast enhanced ultrasound images for the accurate characterization of thyroid nodules. In: Group NB (ed) First Italian national congress of bioengineering, Pisa, Pàtron Editore, 3–5 July, pp 593–594
24. Palàgyi K, Kuba A (1998) A 3D 6-subiteration thinning algorithm for extracting medial lines. *Pattern Recognit Lett* 19:613–627
25. Polyzos SA, Kita M, Avramidis A (2007) Thyroid nodules – stepwise diagnosis and management. *Hormones (Athens)* 6:101–119
26. Spiezio S, Farina R, Cerbone G, Assanti AP, Iovino V, Siciliani M, Lombardi G, Colao A (2001) Analysis of color Doppler signal intensity variation after levovist injection: a new approach to the diagnosis of thyroid nodules. *J Ultrasound Med* 20:223–231, quiz 33

27. Thompson LD, Wenig BM, Adair CF, Shmookler BM, Heffess CS (1997) Primary smooth muscle tumors of the thyroid gland. *Cancer* 79:579–587
28. Turner HE, Harris AL, Melmed S, Wass JA (2003) Angiogenesis in endocrine tumors. *Endocr Rev* 24:600–632
29. Wang T, Basu A (2007) A note on “A fully parallel 3D thinning algorithm and its applications”. *Pattern Recognit Lett* 28:501–506
30. Zhang B, Jiang YX, Liu JB, Yang M, Dai Q, Zhu QL, Gao P (2010) Utility of contrast-enhanced ultrasound for evaluation of thyroid nodules. *Thyroid* 20:51–57

## Biography



**Filippo Molinari, PhD** is Assistant Professor in Biomedical Engineering at the Department of Electronics of the Politecnico di Torino (Torino, Italy). His main research interests include biomedical signal processing and medical imaging. He is expert in automated techniques for aiding the diagnostic process in ultrasound imaging. Vascular sonography, atherosclerosis, and tumor detection are Dr. Molinari's main research topics. He published more than 50 peer reviewed contributions and 3 collaborative books. Dr. Molinari is member of the IEEE EMBS, of the European Society for Molecular Imaging (ESMI), and of the American Institute of Ultrasounds in Medicine (AIUM). He is Regional Associate Editor for Europe of the Journal of Medical Imaging and Health Informatics and he is on the Editorial Board of other five international journals.



**Alice Mantovani, MS** was born in 1985 in Savigliano (CN). He graduated at the Politecnico di Torino in Biomedical Engineering in 2010, where she conducted her master thesis on ultrasound imaging (CEUS) for thyroid lesions detection. From May 2009 to December 2010 she had an internship at the “Ospedale Mauriziano” (Torino). She is expert in medical imaging, segmentation, and classification techniques. She is currently working as a software specialist at DiaSorin Spa (System Development & Integration area).



**Maurilio Deandrea, MD** received the Italian Laurea Degree in 1987, and is working as Endocrinology Consultant at the Unit of Endocrinology of “Umberto I” Mauriziano Hospital in Turin since 1995. Dr. Deandrea has carried out his basic and clinical research mainly in the field of thyroid diseases, with particular expertise in the field of diagnostic and interventional ultrasound, diagnosis and treatment of thyroid cancer, as documented by many publications and book chapters on these topics.



**Paolo Piero Limone, MD** received the Italian Laurea degree in Medicine in 1977; from 1980 to 2004 he was Research Assistant Professor at the University of Turin. On July 2004 he was appointed Head of the Endocrinology Unit at the “Umberto I” Mauriziano Hospital of Turin. Dr. Limone has carried out research in the field of neuroendocrinology, metabolism and thyroid diseases; his activity is documented in over 70 publications and many chapters in books. He is lecturer at the Endocrinology Specialization School of the University of Turin.



**Rajendra Acharya U, Ph.D., D.Engg.** is a visiting faculty in Ngee Ann Polytechnic, Singapore and adjunct faculty in Manipal Institute of Technology, Manipal, India. He received his Ph.D. from National Institute of Technology Karnataka, Surathkal, India and D.Engg. from Chiba University, Japan. He has published more than 120 papers in refereed international journals (94), international conference proceedings (30), books (9) over the last years. Currently, he is on the editorial board of the following journals: Journal of Medical Imaging and Health Informatics, Open Access Medical Informatics Journal, Open Access Artificial Intelligence Journal, Journal of Biomedical Nanotechnology, and Journal of Mechanics in Medicine and Biology and Biomedical Imaging and Intervention Journal. He served as Guest Editor for many journals.



**Roberto Garberoglio, MD**, is endocrinologist, sonographer, and specialist in nuclear medicine. He has been director of the Echographic and Doppler Diagnostic Service of the “Umberto I” Hospital of Torino and he is now consultant of the “Fondazione Scientifica Mauriziana ONLUS”. He is founder of the “Advanced School of Ultrasound Thyroid Pathologies”. He is currently AdCom member of the Società Italiana di Ultrasuoni in Medicina (SIUMB), affiliated to the World Federation of Ultrasound in Medicine and Biology (WFUMB). He is internationally renowned thyroid expert, as documented by his vast publication history in this field.



**Jasjit S. Suri, Ph.D.** is an innovator, scientist, a visionary, an industrialist, and an internationally known world leader in Biomedical Engineering/Imaging Sciences and its Management and has spent over 20 years in this field. He received his Doctorate from the University of Washington, Seattle and MBA from Ivy League Weatherhead School of Mgt., Case Western Reserve University, Cleveland, Ohio. Dr. Suri has published over 300 technical articles. Dr. Suri was crowned with the President's Gold Medal in 1980 and elected as a Fellow of the American Institute for Medical and Biological Engineering and won over 50 awards during his career.

## **Chapter 7**

# **Shape-Based Detection of Cortex Variability for More Accurate Discrimination Between Autistic and Normal Brains**

**Matthew Nitzken, Manuel F. Casanova, Fahmi Khalifa, Guela Sokhadze,  
and Ayman El-Baz**

**Abstract** Autism is a complex developmental disability that typically appears during the first 3 years of life, and is the result of a neurological disorder that affects the normal functioning of the brain, impacting development in the areas of social interaction and communication skills. Early detection allows for treatments to be attempted, thus minimizing the impact of the autism on the individual. Given currently available diagnostic instruments, autism and other pervasive developmental disorders are difficult to detect in very young children. While shape based statistical analysis methods for autism are still in their early stages, current results show positive outlooks on the ability to detect differences between autistic and non-autistic patients. A framework is proposed that is capable of taking two-dimensional images from a standard medical scanner, and be able to construct a three-dimensional representation of the object and examine it through combination of its weighted linear spherical harmonics. The desired outcome is that a distinction can be made between the analysis of autistic and non-autistic brain data. The reconstruction analysis process involves linearly combining spherical harmonics of the corresponding mesh. It was expected that due to the complexity of the brain of an autistic subject it would require more iterations of reconstruction to reach convergence of the same error level as compared to the brain of a non-autistic subject. This was confirmed by the data. Using this method of analyzing the data a significant difference can be demonstrated between groups of examined subjects. The research clearly demonstrates that the non-autistic subjects' data converges both faster and with a lower rate of error level than the data taken from a person with autism.

---

A. El-Baz (✉)

Bioimaging Laboratory, Department of Bioengineering, University of Louisville, 304 Lutz Building, Louisville, KY 40292, USA  
e-mail: [aselba01@louisville.edu](mailto:aselba01@louisville.edu)

## 7.1 Introduction

Autism is a complex developmental disability that typically appears during the first 3 years of life, and is the result of a neurological disorder that affects the normal functioning of the brain, impacting development in the areas of social interaction and communication skills. Difficulties can be identified in both children and adults with autism. The symptoms are identifiable in verbal and nonverbal communication, social interactions, and leisure or play activities. The classic form of autism involves a triad of impairments, these are typically in social interaction, in communication and the use of language, and in limited imagination as reflected in restricted, repetitive, and stereotyped patterns of behavior and activities [1].

The reported incidence of autism spectrum disorders has increased markedly over the past decade [2]. It is believed that autism affects the information processing found in the brain through the alteration of nerve collections and their synapses [3]. From Congress to popular media, speculation is increasing that more children have autism than ever before. The three classifications of autism include autism spectrum disorders (ASD), Asperger syndrome (AS) and pervasive developmental disorder (PDD) [4].

A study done in 2008 by Rapin and Tuchman [5], shows that autism is now recognized in 1 out of 150 children making it a prevalent disorder [5, 6]. Additional studies by DiGuiseppe show a high prevalence among screened American children with as high as 6.4% of screened children showing at least a mild form of an autistic spectrum disorder [7, 8]. According to the Centers for Disease Control and Prevention (CDC) in 2009, about 1 in 110 American children will fall somewhere in the autistic spectrum. Although the cause of autism is still largely not clear, researchers have suggested that genetic, developmental, and environmental factors may be the cause or the predisposing effects towards developing autism [9].

There are other mathematical relationships between incidence and prevalence. An important nuance about prevalence is that its accuracy is only as good as the degree to which each individual who actually has the condition is counted (the numerator or top number of the fraction), and the completeness with which the “general” or other population has been counted (the denominator or bottom number of the fraction.) Accuracy in these two figures can be hard to achieve. In fact, there are no scientifically based epidemiological prevalence estimates for ASD in the USA at this time. Federal agencies have, however, called upon researchers to submit proposals that will develop better prevalence rates [4].

Until research in the USA results in more accurate figures, the National Institutes of Health (NIH) have suggested the following prevalence rates for ASD based upon research in other Westernized, developing nations:

- 10/10,000 people with “classic” autism
- 20/10,000 people with ASD, including PDD
- 50/10,000 people with ASD, including PDD and Asperger syndrome

These estimates are inclusive; that is, the third estimate includes people in the first two groups. This means that in a given large population, on average 0.5%, one-half percent of the population could be diagnosed with an ASD [4].

Early detection allows for treatments to be attempted, thus minimizing the impact of the autism on the individual. Given currently available diagnostic instruments, autism and other PDD are difficult to detect in very young children. This may be due to several factors: presentation of symptoms varies from case to case; social and language deficits and delays may not be identified until the child is given the opportunity for peer interaction in preschool, low incidence leads to a low index of suspicion, and motor milestones are usually unaffected. Furthermore, there is no standard and easily administered screening instrument for young children. For all of these reasons, pediatric evaluations rarely identify autism before the age of three. However, evidence indicates that there is a large gap between the age of the child at the parents' first concern, the age of the first evaluation, and the age of a definitive diagnosis [8]. Parents are typically first concerned between the ages of 15 and 22 months (earlier for children who have co-morbid mental retardation), but the child is often not seen by a specialist until 20–27 months [14]. In addition, there is often further delay between the first visit to a specialist and a definitive diagnosis. However, evidence shows that this delay in diagnosis causes additional distress to parents, as well as wasting valuable intervention time, indicating that professionals in the field of autism need instruments to aid in the detection of autism in very young children [10].

Some forms of autism merely result in the individual exhibiting low social interaction, but more severe forms can result in severe mental retardation. These individuals may be prone to self-injuring and aggressive behavior. There is no current cure for any forms of autism. However, educational, behavioral, or skill-oriented therapies were designed to remedy specific symptoms in each individual. Such therapies can result in a notable improvement for the individual, especially when begun at a young age [10].

### 7.1.1 *Neuropathology of Autism*

In identification of autism, the analysis of the neuropathology is important. The role of single-stranded microdeletions and epigenetic influences on brain development has dramatically altered our understanding of the etiology of the autisms. Recent research has focused on the role of synapse structure and function as central to the development of autism and suggests possible targets of interventions. Brain under connectivity has been a focus in recent imaging studies, and has become a central theme in conceptualizing autism. Despite increased awareness of autism, there is no “epidemic” and no one cause for autism. Data from the sibling studies are identifying early markers of autism and defining the broader autism phenotype [5].

The three sections of the brain analyzed are the gray matter, the white matter, and the corpus callosum. Examination of the individual sections shows significant

changes to the neuropathology of autistic individuals, suggesting a higher complexity in the autistic brain than the normal brain [11].

The gray matter is the brain cortex that contains the nerve cells responsible for routing sensory or motor stimuli to interneurons of the central nervous system. In autistic individuals Abel et al. [12] identified a decreased gray matter volume relative to a control group in the right paracingulate sulcus, the left inferior frontal gyrus, and an increased gray matter volume in amygdala and periamygdaloid cortex, middle temporal gyrus, inferior temporal gyrus, and in regions of the cerebellum. Additionally Boddaert et al. [13] found significant decreases of gray matter concentration in the superior temporal sulcus when comparing autistic child patients to normal child patients. The autistic children also demonstrated a decrease in white matter concentration located in the right temporal pole and in the cerebellum. Herbert et al. [14] applied a voxel-based-morphometry (VBM) approach to male patients between the ages of 7 and 11 years and showed that those with autism had a significantly larger volume of cerebral white matter (CWM) while cerebral cortex and hippocampus–amygdala had smaller volumes. The corpus callosum is largest single fiber bundle in the brain and is responsible for connecting the two hemispheres of the brain. It has been proposed that there are significant differences between the CC of autistic and normal patients [15, 16].

The concept that the cerebellum might play a role in the coordination of attention in a fashion analogous to the role it plays in motor control and that in autism, cerebellum mal-development is a consistent feature that renders the child unable to adjust his or her mental focus of attention to follow the rapidly changing verbal, gestural, postural, tactile, and facial cues that signal changes in a stream of social information [17]. Such cues signal the normal child to move his or her “spotlight of attention” from one source of information (e.g., auditory) to another (e.g., visual). This process involves disengaging attention from one source and then moving and reengaging it on another (i.e., inhibition of one source and enhancement of another). To selectively adjust the focus of attention, the nervous system must quickly and accurately alter the pattern of neural responsiveness to sensory signals – from an enhanced neural response to certain stimuli (e.g., vocalizations) to an enhanced response to other stimuli (e.g., gestures), and from inhibited neural response to some stimuli to inhibited response to others [18, 19].

### 7.1.2 Related Work

One method of early detection intervention, utilizes medical providers to screen children using the Modified Checklist for Autism in Toddlers (M-CHAT) as they were referred for early intervention services. The M-CHAT is designed to screen for early identification of autism spectrum disorder (ASD) in toddlers over the age of 12 months. Ideally, it is given at the 18–24 month well baby check. Parents complete the items on the checklist independently or by interview. Meeting the criteria suggests the risk of ASD and indicates a positive diagnosis for autism. The

purpose is to survey parents to determine how their child responds to varied stimuli from toddler locomotion to a child's reaction to other people. M-CHAT users also incorporate the M-CHAT Follow-up Interview into the screening process, given that recent findings demonstrate that the interview greatly reduces the false positive rate, which avoids unnecessary referrals [14]. Therefore, these children were considered to be at risk for a developmental disorder, but none had received any specific diagnoses and none had received more than several weeks of minimal intervention services.

In new experiments at Yale University, the researchers studied a group of 2-year-olds with autism, as well as typically developing children with developmental disabilities other than autism. The Yale program of research focuses on mechanisms of socialization and their disruption in the autism spectrum disorders. This work includes a close collaboration with Warren Jones in the development of novel techniques to quantify social processes using eye-tracking technologies with a view to visualize and measure the ontogeny of social engagement. New data analysis strategies have been used with children, adolescents, and adults with autism spectrum disorders revealing abnormalities of visual scanning behaviors when viewing naturalistic social approaches and situations. In this study autistic children showed a preference for audiovisual synchronicity in the use of "pat-a-cake" videos, while the other children were more interested in the figure's movements regardless of audiovisual synchronicity. That pattern could be a clue about brain development and early signs of autism [20–23].

Dr. Klin of Yale University explains that within a few days after birth, normal developing children prefer watching biological motion, the movement of living beings, such as their parents, and that preference is an important survival skill and a building block for relationships [21]. But Klin's group found that autistic children were more interested in "non-social contingencies", which are synchronicities that do not have any social meaning, like two balls colliding and making a sound, or a stone falling when someone drops it [20].

Researchers hope that a simple brain scan performed in infants and toddlers can presage the development of autism, leading to early detection and early intervention. The test involved using functional magnetic resonance imaging (MRI) to measure brain responses to spoken words in sleeping children. For this study, Dr. Eyler and her colleagues monitored the brain activity of 30 children with an autism spectrum disorder (aged 14 months to 46 months) and 14 "typical" children of roughly the same age [24].

Children slept in the MRI machine while researchers read them bedtime stories. This allowed the investigators to see which parts of the brain were being activated in typical children vs. children with autism. "In the typically developing children, both sides of the brain involved in language processing were activated. In the youngest children, the activation was about equal in both the right and left hemisphere, while in the older children, activity became more pronounced on the left side, which is similar to adult patterns and to be expected", Dr. Lisa T. Eyler explained. But in the autistic children, there was slightly more right hemisphere response than left hemisphere, and there was no change in activity across the age range [24].

David G. Amaral stated that “this leads to the conclusion that, in many children with autism, there are alterations either in structure growth or connectivity of the brain, but we really don’t understand the implications of that for core features of autism, one of which is the problem with communication. This provides more evidence for abnormal connectivity in the brain” [24].

Further analysis of neurological MRI scans has been pursued in automated computer analysis of specific components of the brain. Approaches by El-Baz et al. [11] examine the shape model comparison between the corpus callosum in individuals with and without autism. This analysis focuses on comparison of the three-dimensional voxel positioning. In such an automated technique, specific areas of MRI scan images are extracted. These images are then placed in a stack to recreate a volume of the image. The difference between regions of this volume can be statistically measured. While statistical analysis methods are still in their early stages, current results show positive outlooks on the ability to detect differences between autistic and normal patients based on voxel-based analysis. The positive findings from automated analysis research provide the basis for the research done in this chapter.

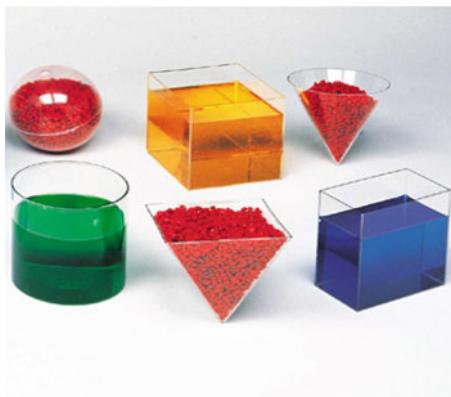
There are many different methods that can be used to examine objects for difference. In the area of autism detection and analysis, Gyral Window Index (GWI) and volumetric descriptions of brain structures (volume analysis) are the most common ones.

GWI analysis involves counting the gyrations found along the brain in various areas and comparing these counts to make an assumption. This technique is heavily prone to error in that brains that drastically differ in shape and complexity may show little to no difference and brains of similar shape and complexity could show a large difference.

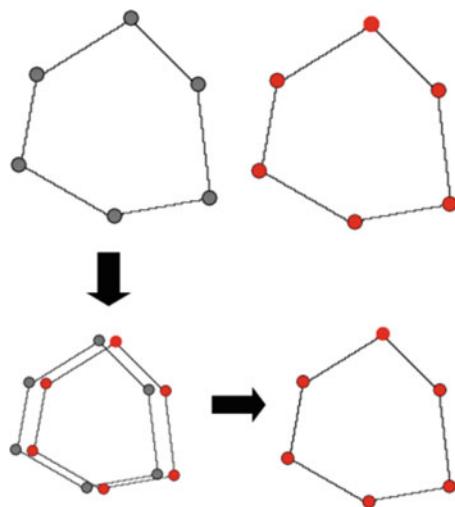
Volumetric descriptions of brain structures examine the volume of an object. This is also commonly known as volume analysis. In this the volumes of objects are compared. While this does relay important information, it is not optimal for examining a complex organ. Imagine that you have two bowls. One of these is shaped like a large flat saucer and the other is shaped like a very thin vase. Assume that both of these will hold the same amount of liquid when filled. This example illustrates the flaws of volume based analysis on an object. While the volumes are identical for the two objects, the shape and complexity is drastically different. An example of this type of phenomenon can be seen in Fig. 7.1 illustrated by a clear plastic geometry set.

Shape analysis is an alternative method to the previous two. In these the shape is represented by a complex mesh of nodes. These nodes fall within respective areas on the surface and can be compared to one another to determine differences. Instead of examining the volume or visual changes along the cortex of the brain the locations of points on the brain are examined and compared based on their locations in space. This examination performs a comparison between specific points as illustrated in Fig. 7.2. It is important to maintain strict rules on construction and examinations of objects when using this technique to ensure accuracy. This chapter will focus on a framework for the analysis of the cortex using shape analysis through the construction and examination of spherical harmonics.

**Fig. 7.1** Containers that all hold an identical volume with dramatically different shapes



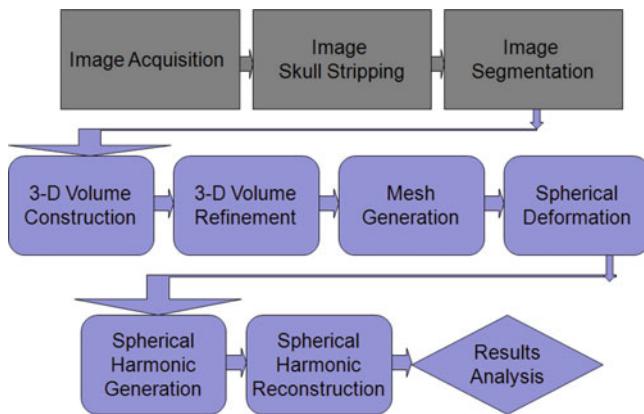
**Fig. 7.2** Illustrating the comparison between points representing an object in shape analysis instead of examining the object as one volumetric entity



## 7.2 Methods

The proposed analysis system is subdivided into the parts found in Fig. 7.3. This section will primarily cover the second and third rows in the flowchart which pertain to three-dimensional mesh construction and to the analysis of a mesh using spherical harmonics. The initial data consists of segmented MRI images of data.

First, the raw recorded image data is passed through a system to remove the skull from the image. This is done using the software called MRIcro developed by Chris Rorden from the Neuropsychology Lab and Georgia Institute of Technology in Atlanta, GA [25–28]. This software examines the raw image and removes and pixel data pertaining to the skull leaving only brain matter remaining as shown in



**Fig. 7.3** Flowchart of the proposed analysis system

**Fig. 7.4** (a) Typical MRI image, (b) image after removal of skull from around the brain using MRIcro

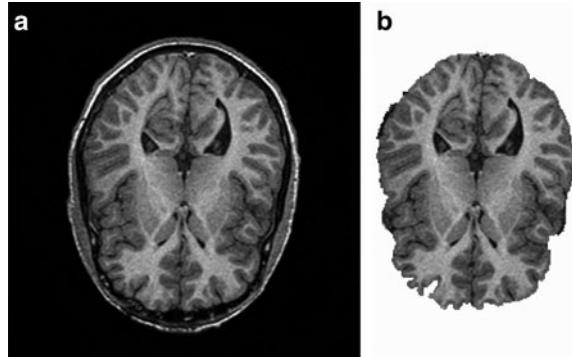
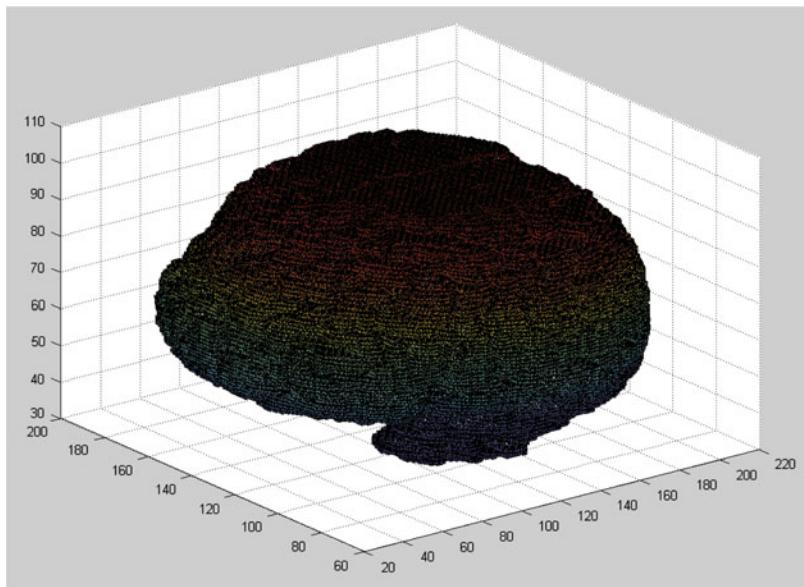
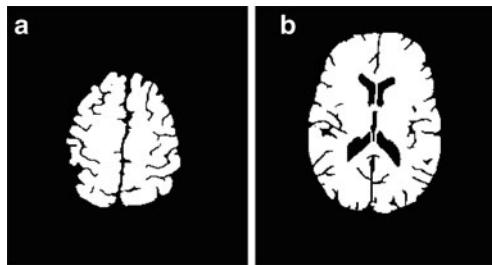


Fig. 7.4. After this the image is segmented to a binary representation. The image is inverted so that the brain received by the framework has a black background and the brain data is the white area.

For each image, a black pixel represented the background or portions of the image where no data existed, and the white pixels represented areas where data existed (Fig. 7.5). Often images formats do not store data in binary form and data is stored using a variation of a grayscale format. The individual pixel values range from 0 to the maximum grayscale value (in many cases this value was 255 and represented a white pixel.) It can be assumed then that any pixel with a value greater than zero would constitute a pixel containing data. Setting all pixels to the desired limits can be accomplished by dividing the matrix by the maximum value and then taking using a ceiling command on the data.

After each image is converted to a binary representation, the images are assembled in a three-dimensional matrix stack. Each matrix is represented by an X, Y, and Z dimension. The X dimensions represent the rows in an image, the Y dimension represents the columns in an image, and the Z dimension represented the layers in

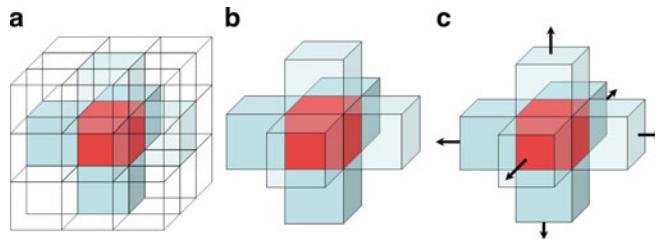
**Fig. 7.5** Examples of segmented brain data used as input data



**Fig. 7.6** Three-dimensional binary representation of a brain

the volume, with each layer encompassing a separate distinct image. Each coordinate ( $x, y, z$ ) represented a single voxel in three-dimensional space and the combination of the voxels is a rough three-dimensional representation of the object to be analyzed. Figure 7.6 illustrates a generic representation of a three-dimensional volume as viewed using Matlab.

In this framework the images are loaded into the memory without preprocessing to remove holes in the images. This is done because when the images are compiled into a shape in three-dimensional space, holes can be created by the placement of layers on top one another. It is a simpler process to simply execute an algorithm to remove these holes using one processing step than to perform both pre and post processing on the volume. While initially the decision was made to remove all holes in the two-dimensional images prior to being loaded into the program, this proved ineffective after the images were placed in a three-dimensional volume. To remove these holes from the image, a hole detection algorithm can be used.



**Fig. 7.7** (a) The original cube of pixels, (b) showing the starting neighbor pixels in the six cardinal directions, (c) the outward directions of movement for the detection iterations

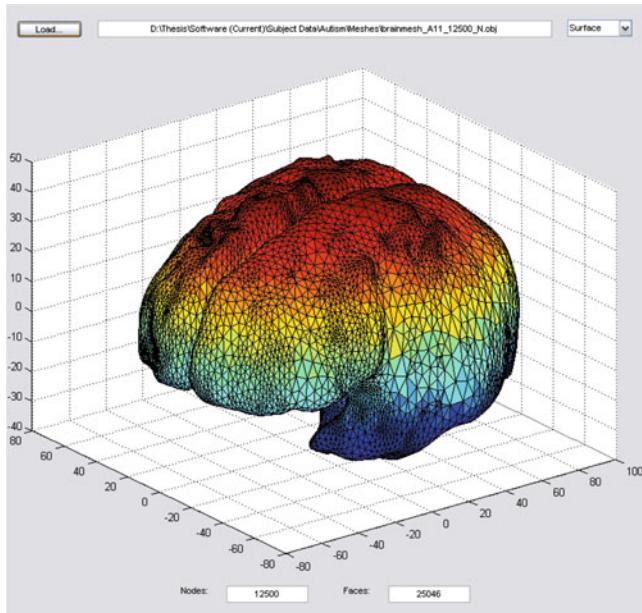
The algorithm begins by loading the images into the three-dimensional volume matrix as previously described. After creating the three-dimensional volume matrix, an iterative pass is made across the  $X$ ,  $Y$ , and  $Z$  axis of the volume using the Matlab “fillholes” command. To accomplish this, a two-dimensional image slice is removed on each plane of the volume. This image slice was then passed through the “fillholes” algorithm, and the modified image slice was reinserted into the volume. In this way, the holes were removed from the two-dimensional representations in the  $X$ ,  $Y$ , and  $Z$  directions.

Following this procedure, additional small holes not removed using the previous method are scanned for and removed. To remove these additional holes, each individual pixel is iterated in the three-dimensional matrix. For each pixel, it is calculated if there was a pixel belonging to the image in the  $\pm X$ ,  $\pm Y$ , and  $\pm Z$  directions between the origin pixel and the edge of the volume. In Fig. 7.7a, the origin pixel is defined as the red pixel, the blue neighbor pixels are the  $\pm X$ ,  $\pm Y$ , and  $\pm Z$  directional pixels, and the clear pixels represent unchecked points. Figure 7.7b shows the same image with the unchecked pixels removed for easier visibility. From the image points, outgoing vectors were tested in each of the six directions. Figure 7.7c illustrates the path of testing that each neighbor pixel would expand from towards the outer edge. If a pixel value of 1 is found in a given direction, a value of true is marked for the boolean corresponding with that direction.

If a pixel is found radiating out in all six directions, it is determined that the pixel was a hole in the image. A pixel that does not contain pixels surrounding it on all six sides is ignored. This procedure is repeated until holes are removed from the image. This is often accomplished in a single iteration. To ensure that the volume is clean, an additional iteration is repeated that verifies that no holes remain in the volume.

### 7.2.1 Mesh Generation

Once the three-dimensional volume has been properly constructed, the mesh can be generated. In this framework, the mesh generation is performed using a modified version of the iso2mesh Matlab based mesh generation system, written by Qianqian



**Fig. 7.8** Mesh render for the cortex of an autistic person

Fang and David Boas [29, 30]. A sample mesh render is illustrated in Fig. 7.8. This system is built on the CGAL Delaunay Triangulation mesh engine. This is a nonrigid mesh generation engine, and points are not constrained to contain a specific number of neighboring nodes. The modifications to the original iso2mesh code are to enable restriction of node counts and to generate a raw nonprocessed mesh. The restriction of node counts is important to ensure that objects being compared have the same properties. The raw mesh is recovered from the iso2mesh package before any of the post processing and surface analysis is performed on it. For this framework, these alterations are not necessary and can lead to complexity problems in the mesh.

The data received from the mesh generation is a collection of nodes and triangulations. The nodes matrix is of dimensions  $3 \times N$ , and the triangulations matrix is of dimensions  $4 \times T$ , where  $N$  is the maximum allowable amount of nodes and  $T$  varies based on the positions used by the mesh engine and the triangulations created between points. The combination of three triangulation points connected with one another represents a face found on the corresponding mesh and each row in the triangulations matrix contains data representing a unique face. The fourth column in the triangulation matrix is a boolean value representing that the face connectivity at this point exists. Following mesh creation, all values in this fourth column will be equal to true and this column can thus be removed to create the traditional  $3 \times T$  representation found in a mesh triangulation matrix.

Once the initial mesh is created, it is necessary to reposition it in three-dimensional space and to resize the mesh to appropriate proportions. The centroid

of the mesh is calculated in the  $X$ ,  $Y$ , and  $Z$  directions. Using the coordinates of the centroid, the mesh is repositioned so that it is centered on the origin in three-dimensional Cartesian space ( $x = 0$ ,  $y = 0$ ,  $z = 0$ ). Because of data acquisition limitation the original mesh is not scaled properly. It is necessary then to appropriately resize the mesh using the original image slice acquisition scaling. The mesh coordinates are then repositioned according to the  $X$ ,  $Y$ , and  $Z$  magnification parameters of the scanner used to acquire the images. The  $X$  and  $Y$  planes are multiplied by a magnification factor of 0.93 in this example, taken from the MRI scanner acquisition parameters. The  $Z$  plane is scaled by a factor of 1.5 in this example, the distance between slices taken during MRI acquisition. The result is a properly scaled mesh that is centered around the origin in three-dimensional space.

The raw mesh generation returns a node and face cluster where vertex points contain a large amount of noise. To remove this noise, a vertex-based anisotropic smoothing filter is applied to the data. The filter performs a low-pass filter Laplacian smoothing algorithm across the exterior of the mesh. The low-pass filter Laplacian smoothing algorithm used in this framework is based on code written by Zhang and Hamza [31]. The smoothing algorithm is applied iteratively three times during the procedure, and is configured to smooth at a minimal value with each pass. The final mesh is accurately scaled and correctly positioned in three-dimensional space and contains smooth and accurate contours representing the various locations found on the cortex.

The file format chosen to save the mesh data in this framework is the Wavefront OBJ format, developed by Wavefront Technologies. The standardized Wavefront OBJ format is used because it is a common mesh file representations used in a majority of commercial and open-source application mesh analysis applications and will allow for potential future integration of the framework with third-party software. Custom algorithms have been written to write and save meshes in this format in a way that is universally compatible. The algorithms used follow all standard guidelines for the common Wavefront OBJ format.

### **7.2.2 *Spherical Deformation***

Following the generation of a stable, hole-free initial mesh, it was necessary to generate a corresponding unit sphere. The accuracy of the sphere creation is relative to the accuracy of the statistical analysis. There were several techniques attempted including Cartesian and spherical registration methods.

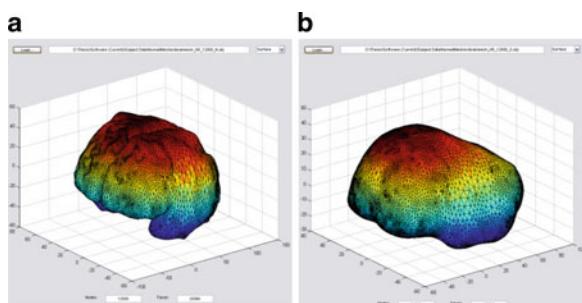
The initial attempt to create a unit sphere involved simply inflating the original mesh into a unit sphere. This technique was done by reducing all points in the mesh to a maximum distance of 1.0 from the origin. Once the mesh had been scaled, all points with a value less than 1.0 were up scaled to a value of 1.0. This inflation technique resulted in numerous problems, with the most problematic being vertex overlap and poor distribution of points. The natural shape of the brain sulci and the valley located between hemispheres in the brain created this overlap by inverting

points and face connectivity, as they were forced to the outer edges of the inflated sphere. This occurred due to the fact that multiple points existed along the same or incorrect  $\theta$  and  $\varphi$  angles in spherical space.

In the creation of a unit sphere, it became imperative that all points remain in their correct orientation with their neighbor points during the deformation process. This means that during deformation, the triangulation connections could not become crossed. A unit sphere that contained crossed triangulations produced an erroneous spherical representation. This is because during calculation of the harmonics it assumes that the detected neighboring nodes are truly the neighbor nodes for a given vertex. If the points become misplaced or shifted to an erroneous location away from its true position than a neighboring node may appear to be falsely connected to additional nodes. The spherical harmonics are based on angular values, and points with incorrectly crossed angles caused the system to produce poor resultant data making a spherical reconstruction not possible.

The accuracy of the spherical representation was also based on the distribution of the vertex coordinates throughout the sphere and the accuracy to it being a true unit sphere. A sphere with clusters of vertices and other areas of sparse vertex placement produced significantly more error and had a greater difficulty converging. The ideal spherical representation will have all vertices spaced equidistant from one another across the surface of the sphere. While it is possible to perform an analysis using an improperly spaced unit sphere, the results were less than desirable. The original method leaves large clusters of vertices around areas of significant sulcus curvature in the brain.

The process, used in the creation of a unit sphere is a multiple phase deformation technique, created for the purpose of deforming the brain meshes. Before running the spherical deformation, the brain was heavily smoothed using a Laplacian based smoothing algorithm. The mesh was loaded into the freely available software, MeshLab v1.2.3b written by Paolo Cignoni [32, 33]. It was determined through trial and error that an average smoothing of 400 iterations per 12,500 nodes deformed the mesh so that no existing points in the mesh could be found residing on the same  $\theta$  and azimuth angles, when examined in spherical coordinate space (Fig. 7.9). MeshLab was used to improve processing speed during the smoothing phase. The comparable Matlab algorithm took approximately 30 seconds to run one



**Fig. 7.9** (a) An original brain render and (b) the render after 400 iterations of Laplacian smoothing

complete Laplacian smoothing pass. The same algorithm run in MeshLab was capable of performing 2000 Laplacian smoothing passes in slightly under one minute.

### **Algorithm 1: Spherical Deformation Algorithm**

1. *Deformation of original mesh into smoothed mesh model.*
2. *Expansion of smoothed mesh model distances to distance of unity sphere.*
3. *Application of attraction–repulsion algorithm for node positioning refinement on expanded mesh.*
4. *Expansion of refined mesh model distances to distance of unity sphere.*

The Spherical Deformation Algorithm was used to transform the shape of a constructed mesh to that of a unity sphere. In the Laplacian deformation phase, the brain is smoothed until all curves, peaks, and valleys of the cortex are flattened across the brain. This ensures that no more than one point will be found for a given set (azimuth and  $\theta$  angles) of spherical coordinates. This mesh is then resaved as a smooth mesh object which will be loaded for the second through fourth phases.

The unit sphere distance deformation involves inflating the smoothed mesh to the distance of a unit sphere. To increase the speed of the spherical inflation it is necessary to manipulate the mesh in the spherical coordinate domain. The Cartesian mesh is first converted to spherical coordinates using the *cart2sph* command in Matlab. This returns a  $\theta$  angle, azimuth ( $\varphi$ ) angle, and  $R$  distance for each point in the mesh. The  $R$  distance is removed and replaced with a matrix of the same size with all values equal to 1.0. The coordinates are then reconverted to Cartesian using the *sph2cart* command in Matlab. This provides a fast and effective conversion that is extremely efficient, because it requires no iterative processing and does not involve examining each node individually or performing any distance calculations. These inflated points are then passed on to the third phase.

In the third phase, the attraction–repulsion algorithm is used to refine the inflated sphere. The attraction–repulsion algorithm is based on a standard spring algorithm. The algorithm was inspired by a concept used in Graphic Art Design for inflating objects. The purpose of this algorithm is to reposition the nodes in the mesh. After the second phase of the spherical deformation, there are large clusters of nodes spaced near one another in areas of the mesh that contain a high density of nodes. Looking at the mesh, it is possible to see large clusters near the lower lobes of the brain and around complex areas. As previously mentioned, it is necessary to evenly space the nodes across the sphere. To accomplish this, the attraction–repulsion algorithm was created.

The attraction–repulsion algorithm is composed of two steps, the attraction step and the repulsion step. Each node is altered by being processed in the attraction step and then the repulsion step. In the attraction step, the node is altered based on its neighbor nodes. The distance between a node and each of its neighbor nodes is calculated. The node is then pulled based on numerical weighting, so that it becomes centered between its neighbors. The attraction for iteration  $i$  is defined as:

$$A_{i+1} = A_i + (\overline{Q_N P_J})(D^2)(0.01) + (\overline{P_J Q_N}) \left( \frac{0.00001}{D} \right), \quad (7.1)$$

$1 \leq j \leq \text{number of nodes}, \quad 1 \leq N \leq \text{number of neighbors}$

where  $A_{i+1}$  represents the new node coordinate and  $A_i$  is the original node coordinate.  $P$  represents the original coordinate of the unmodified node that is and  $Q$  is the coordinate of neighbor node  $N$ . The distance ( $D$ ) from  $P$  to  $Q$  is a three-dimensional Euclidean distance.

After the node has been centered between its neighbors it is processed in the repulsion step. Here the node is slightly readjusted by every node in the mesh. Each node minimally repels one another so that the nodes do not cross or touch. The repulsion is defined as:

$$R_{i+1} = R_i + (\overline{Q_k P_J})(T) \left( \left( \frac{1.5}{(\overline{Q_k P_J})^2} \right) \left( \frac{1}{2N} \right) \right), \quad (7.2)$$

$1 \leq j \leq \text{number of nodes}, \quad 1 \leq k \leq \text{number of nodes}$

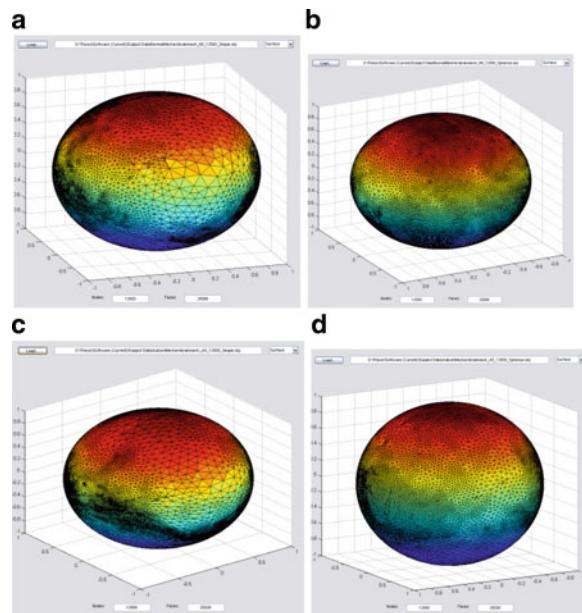
where  $R_{i+1}$  represents the new node coordinate and  $R_i$  is the original node coordinate.  $N$  represents the total number of nodes in the mesh.  $T$  is a value between 0 and 1 and stands for the time step of the algorithm. A larger time step enables the algorithm to converge faster, but increases the chance of error as nodes are capable of moving larger distances.

This algorithm also causes an inflation effect to occur during repeat iterations as each node is gently repelled from interior angles by nodes opposite it on the unit sphere. Because there are no nodes outside the unit sphere to repel the nodes back toward the center this inflation occurs. This step is repeated several times until a satisfactory node distribution is reached (Figs. 7.10 and 7.11). Once the attraction–repulsion algorithm has completed, the nodes are more evenly spaced on the sphere. Due to the previously described interior repulsion, the sphere is also much larger than a unit sphere of radius 1 at this point. To alleviate this phenomenon, the unit sphere distance deformation algorithm is run a second time. While this algorithm does not alter the angular placement of the nodes, it will reduce the  $R$  values back to 1 for all nodes. This is the same algorithm that is run during phase two. Following this deformation, the newly created unit sphere mesh is written back into a Wavefront OBJ file.

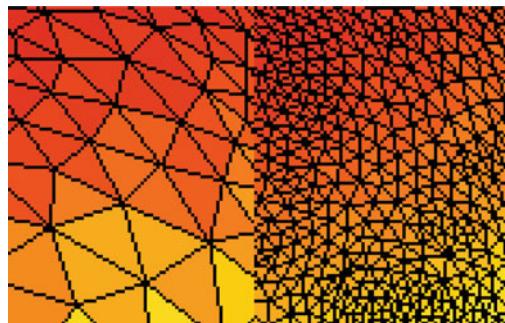
### 7.2.3 Spherical Harmonic Analysis

The weighted spherical harmonic representation (weighted-SPHARM) is a surface modeling framework that can be used in encoding cortical shape information. The technique was developed by Moo K. Chung, Kim M. Dalton and Richard J. Davidson.

**Fig. 7.10** (a, c) Simple spherical inflations of subject brains. (b, d) The respective corresponding unit spheres after three iterations of the attraction–repulsion algorithm have been applied



**Fig. 7.11** Detailed view of the same region before and after application of the attraction–repulsion algorithm



The weighted-SPHARM representation is a spectral method [34], where a linear combination explicit basis functions is used to represent noisy cortical surface data. The basis expansion corresponds to the solution of an isotropic heat equation on a unit sphere. The result of the weighted-SPHARM is explicitly given as a weighted linear combination of spherical harmonics. This provides a more natural setting for statistical modeling. The representation can be further used in surface registration that reduces the improper alignment of brain sulcus folding patterns between subjects and across hemispheres within a subject.

The system generates harmonics from the original distribution of coordinates. This ensures that the same neighbor nodes are connected in the same order between the original and the reconstructed mesh. The harmonics are specifically generated

from the unique unit sphere corresponding to the original mesh. The original Cartesian coordinate system of the corresponding unit sphere mesh parameterizes a coordinate  $v_i = v(x, y, z)$  to spherical coordinates with a polar angle between 0 and  $\pi$  and an azimuth angle between 0 and  $2\pi$ . Each point is then represented as a spherical coordinate that can be expressed as  $v_i = v(\theta, \varphi)$ . The distance from the origin, or  $R$  value, in spherical space is always equal to 1.0 for any given point on a unit sphere [35]. The paper by Chung et al. defines the values of  $\theta$  and  $\varphi$  for purposes of calculation as:

$$\theta_{\text{SPHARM}} = \frac{\pi}{2} - \varphi, \quad \varphi_{\text{SPHARM}} = \pi + \theta. \quad (7.3)$$

The spherical harmonic  $Y_{lm}$  of degree  $l$  and order  $m$  [36, 37] is defined as:

$$Y_{lm} = \begin{cases} c_{lm} P_l^{|m|}(\cos \theta) \sin(|m|\varphi), & -l \leq m \leq -1 \\ \frac{c_{lm}}{\sqrt{2}} P_l^{|m|}(\cos \theta), & m = 0 \\ c_{lm} P_l^{|m|}(\cos \theta) \cos(|m|\varphi), & 1 \leq m \leq l \end{cases}, \quad (7.4)$$

$$c_{lm} = \sqrt{\frac{2l+1}{2\pi}} \frac{(l-|m|)!}{(l+|m|)!}, \quad (7.5)$$

where  $P_l^{|m|}$  is the associated Legendre polynomial of order  $m$  [35, 36]. These form a form a polynomial sequence of orthogonal polynomials [35].

For the purpose of this chapter only nonnegative harmonics were used. For each degree  $Y_{lm}$  represents the Fourier coefficients capable of reconstructing the spherical harmonic as specified in code written by Moo K. Chung. This code saves the coefficients of the spherical harmonic in a new file for each degree. These can then be reloaded to expedite future calculations [34].

A final reconstruction is created by iteratively using the desired number of harmonics to reconstruct the original brain in a linear fashion. As each harmonic is loaded into memory, it is multiplied by a factor sigma which is equivalent to the smoothing of the harmonic. A larger sigma value indicates a higher degree of smoothing, while a smaller sigma value preserves more of the data from the current harmonic degree. These values are then linearly added to the previous coordinate value, and the new coordinate is formed. During reconstruction of the original mesh, the surface coordinates can be modeled independently according to the equation:

$$v_i(\theta, \varphi) = h_i(\theta, \varphi) + \varepsilon_i(\theta, \varphi), \quad (7.6)$$

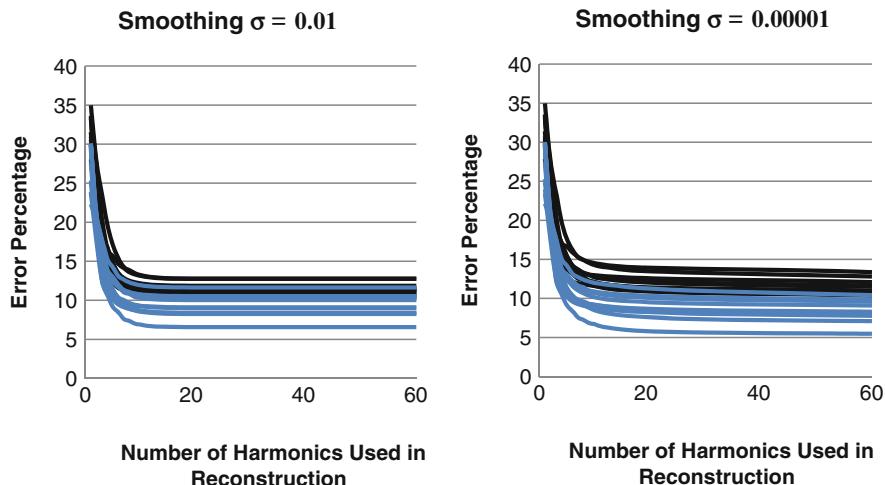
where  $v_i(\theta, \varphi)$  represents the new coordinate,  $h_i(\theta, \varphi)$  represents the original coordinate, and  $\varepsilon_i(\theta, \varphi)$  is the linear modifier constructed from the combination of the smoothing  $\sigma$  value and the residual Fourier values for the specified coordinate calculated for the current spherical harmonic degree. This procedure is repeated for

the desired number of harmonics, and a final resulting mesh is produced and returned to the user [37].

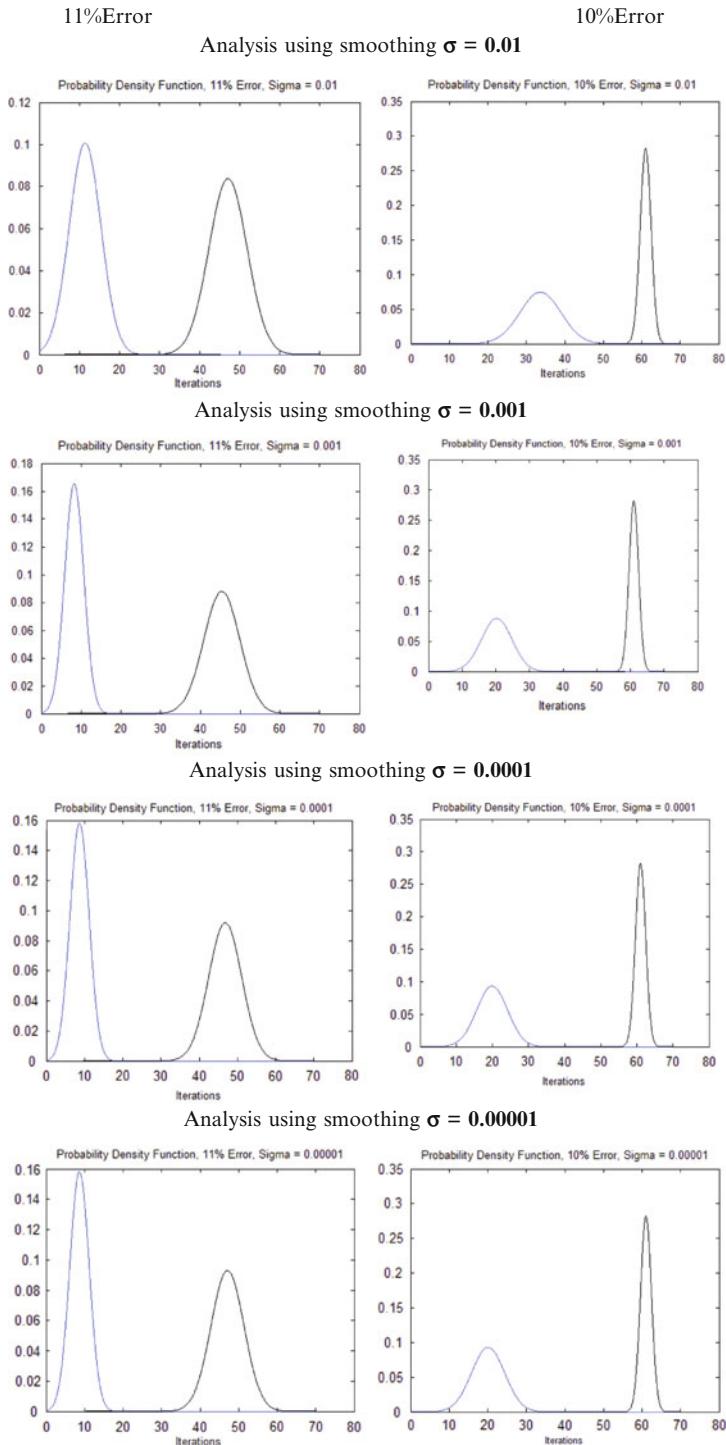
The error between the reconstructed brain mesh and the original brain mesh is found by calculating the three-dimensional Euclidean distance between corresponding points in the original mesh and the reconstructed mesh. Due to the strict ordering of the data storage in the harmonics, the reconstructed mesh and the original mesh are already registered to one another, and thus do not require additional registration, simplifying the calculations required. Additionally, for visualization purposes, the same connected set of faces is used for both the original and reconstructed mesh [37].

All subject data was processed identically, and results were analyzed for  $\sigma$  smoothing values of 0.01, 0.001, 0.0001, and 0.00001. There were 10 subjects in the autistic group and 12 subjects in the non-autistic group. The error for each reconstruction was calculated for every set of harmonics in each of the patients. These errors were then analyzed to find the iteration at which a reconstructed brain demonstrated accuracy below a certain threshold. The thresholds of 11 and 10% were selected for accuracy. Due to the slow convergence of many of the autistic brains, an iteration value of 61 has been used to represent that the brain did not converge below the threshold within the initial 60 iterations used.

As the smoothing due to Fourier reconstruction across the mesh is reduced, the amount of data used from each harmonic is increased. The harmonics begin to converge much faster in these reconstructions. The error curves for each data set show the rate of convergence and the maximum convergence visually for a specific brain. The convergence curves at the two extremes for smoothing can be seen in Fig. 7.12.



**Fig. 7.12** Convergence curves for the two extreme values of sigma



**Fig. 7.13** Probability Density Function plots at decreasing levels of  $\sigma$

It is important to note the overlap between the blue and black lines in the first image in the figure and the sharp distinction between the blue and black lines in the second image. The blue lines represent convergence of subjects not demonstrating autism while the black lines represent subjects who positively demonstrated autistic characteristics.

Probability density functions (pdf) were generated for each error threshold as well using the mean and standard deviation applied to a Gaussian fit pdf. This gives the likelihood of classification for a given number of iterations to convergence. The black lines (right peaks in all images) represent the data for the subjects demonstrating autism while the blue lines (left peaks in all images) represent subjects not demonstrating autism. This shows that the probability of a subject requiring a low number of reconstruction iterations to converge below a given threshold is likely to not be autistic.

### 7.3 Conclusion and Future Work

Considering the data from previous methods including Modified Checklist for Autism in Toddlers, eye-tracking technologies, and the prevalence of the disorder, it is essential to find alternate and scientific methods that include brain analysis as a diagnostic and evaluative method. The research of doctors at Yale has determined that brain scans can predict the development of autism, leading to early detection and early intervention.

Using this method of analyzing data can demonstrate accurate differences in normal and autistic brains. The research that has been generated in this chapter can clearly demonstrate that the normal brain data converged both faster and with a lower rate of error level than the Autistic brain data. This result proves that the autistic brain is a more complex structure, and would be more difficult to reconstruct using this shape-based detection of cortex variability process.

The flexibility of the created package is of additional importance to the expansion of the project. The algorithms and theories introduced in this chapter can readily be applied to any object freely. The object can be converted from two-dimensional scans into a three-dimensional mesh, deformed and analyzed. This would allow for future analysis of many other organs including cancerous growths and individual components of the brain. Examining the difference between reconstructions can enable change in an object to be tracked over time and compared as well. This would allow for an analysis of the rate of the progression of autism in the individual. This can provide valuable information that can be used to improve treatments by providing physicians with a detailed mathematical representation of the current state of their patient.

It is also plausible to use a technique similar to this one to understand what areas of the brain begin to alter at different times in the subject and to track their impact on the overall autism. Gaining a detailed understanding of the progression of autism in patients can help lead to meaningful solutions for autistic patients.

After numerous attempts to create a package for shape-based detection of cortex variability, the primary difficulties arose in generating an accurate mesh from a variety of data and deforming the mesh into an accurate unit sphere while preserving the integrity and positioning of nodes within the mesh. Through the combination of techniques from a variety of fields including engineering, computer science and graphical art design these challenges were overcome. Using this technique it will be possible to analyze a large variety of MRI scans to compare the complexities of normal and Autistic brains.

Future plans are to improve the efficiency of the algorithms to allow accurate deformation and analysis of larger and more detailed mesh structures. Improvements in the algorithm will allow for faster and more accurate analysis of the subject. As previously mentioned, it is also planned to make the package more flexible so that it can readily be applied to a variety of structures and used as a meaningful evaluation technique for multiple disorders.

## References

1. Minshew NJ, Payton JB (1988) New perspectives in autism, part I: the clinical spectrum of autism. *Curr Probl Pediatr* 18(10):561–610
2. Schaefer GB, Mendelsohn NJ (2008) Clinical genetics evaluation in identifying the etiology of autism spectrum disorders. *Genet Med* 10(4):301–305
3. Losh M, Adolphs R, Poe MD, Couture S, Penn D, Baranek GT, Piven J (2009) Neuropsychological profile of autism and the broad autism phenotype. *Arch Gen Psychiatry* 66(5):518–526
4. Jacobson JW (2000). Is autism on the rise? *Science in Autism*, vol 2, no 1
5. Rapin I, Tuchman RF (2008) What is new in autism? *Curr Opin Neurol* 21(2):143–149
6. Department of Developmental Services (1999) Changes in the population of persons with autism and pervasive developmental disorders in California's developmental services system 1987 through 1998. Author, Sacramento, CA
7. Department of Education (1999) Digest of education statistics, 1998. Author, Washington, DC (GPO No. 065-000-01174-3, NCES No. NCES 1999036)
8. DiGiuseppe C, Hepburn S, Davis JM, Fidler DJ, Hartway S, Lee NR, Miller L, Ruttenber M, Robinson C (2010) Screening for autism spectrum disorders in children with Down syndrome: population prevalence and screening test characteristics. *J Dev Behav Pediatr* 31(3):181–191
9. Stevens M, Fein D, Dunn M, Allen D, Waterhouse LH, Feinstein C, Rapin I (2000) Subgroups of children with autism by cluster analysis: a longitudinal examination. *J Am Acad Child Adolesc Psychiatry* 39:346–352
10. Robins DL, Fein D, Barton M, Green J (2001) The modified checklist for autism in toddlers: an initial study investigating the early detection of autism and pervasive developmental disorders. *J Autism Dev Disord* 31(2):131–144
11. El-Baz A, Elnakib A, Casanova M, Gimel'farb G, Switala A, Jordan D, Rainey S (2010) Accurate automated detection of autism related corpus callosum abnormalities. *J Med Syst*. doi:10.1007/s10916-010-9510-3
12. Abell F, Krams M, Ashburner J, Passingham R, Friston K, Frackowiak R, Happe F, Frith C, Frith U (1999) The neuro-anatomy of autism: a voxel-based whole brain analysis of structural scans. *Neuroreport* 10(8):1647–1651
13. Boddaert N, Chabane N, Gervais H, Good CD, Bourgeois M, Plumet MH, Barthélémy C, Mouren MC, Artiges E, Samson Y, Brunelle F, Frackowiak RSJ, Zilbovicius M (2004)

- Superior temporal sulcus anatomical abnormalities in childhood autism: a voxel-based morphometry MRI study. *Neuroimage* 23:364–369
- 14. Herbert MR, Ziegler DA, Deutsch CK, O'Brien LM, Lange N, Bakardjiev A, Hodgson J, Adrien KT, Steele S, Makris N, Kennedy D, Harris GJ, Caviness VS Jr (2003) Dissociations of cerebral cortex, subcortical and cerebral white matter volumes in autistic boys. *Brain* 126:1182–1192
  - 15. Egaas B, Courchesne E, Saitoh O (1995) Reduced size of corpus 576 callosum in autism. *Arch Neurol* 52(8):794–801
  - 16. Schumann CM, Barnes CC, Lord C, Courchesne E (2009) Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. *Biol Psychiatry* 66(10):942–949
  - 17. Courchesne E, Press GA, Yeung-Courchesne R (1993) Parietal lobe abnormalities detected with MR in patients with infantile autism. *AJR Am J Roentgenol* 160(2):387–93
  - 18. Schumann CM, Bloss CS, Barnes CC, Wideman GM, Carper RA, Akshoomoff N, Pierce K, Hagler D, Schork N, Lord C, Courchesne E (2010) Longitudinal magnetic resonance imaging study of cortical development through early childhood in autism. *J Neurosci* 30 (12):4419–4427
  - 19. Kennedy DP, Courchesne E (2008) Functional abnormalities of the default network during self- and other-reflection in autism. *Soc Cogn Affect Neurosci* 3(2):177–190
  - 20. Hitti M, Chang L (2009) Autism: new clue to earlier detection. *WebMD Health News*. March 29, 2009
  - 21. Klin A, Lin DJ, Gorrindo P, Ramsay G, Jones W (2009) Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature* 459(7244):257–261
  - 22. Jones W, Klin A (2009) Heterogeneity and homogeneity across the autism spectrum: the role of development. *J Am Acad Child Adolesc Psychiatry* 48(5):471–473
  - 23. Jones W, Carr K, Klin A (2008) Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Arch Gen Psychiatry* 65(8):946–954
  - 24. Gardner A (2010) New hope for early detection of autism. *US News: Health Day*. May 19, 2010
  - 25. Damasio H, Damasio AR (1989) *Lesion analysis in neuropsychology*. New York, Oxford
  - 26. Frey R, Woods DL, Knight RT, Scabini D (1987) Defining functional cortical areas with “averaged” CT scans. *Soc Neurosci Abstr* 13:1266–1267
  - 27. Rorden C, Brett M (2000) Stereotaxic display of brain lesions. *Behav Neurol* 12(4):191–200
  - 28. Brett M, Leff AP, Rorden C, Ashburner J (2001) Spatial normalization of brain images with focal lesions using cost function masking. *Neuroimage* 14(2):486–500
  - 29. Fang Q, Boas DA (2009) Tetrahedral mesh generation from volumetric binary and gray-scale images. In: *Proceedings of IEEE international symposium on biomedical imaging 2009*, pp 1142–1145
  - 30. Fang Q, Carp SA, Selb J, Moore R, Kopans DB, Miller EL, Brooks DH, Boas DA (2008) A multi-modality image reconstruction platform for diffuse optical tomography, *Biomedical optics 2008, BIOMED poster session II (BMD)*, March 2008.
  - 31. Zhang Y, Hamza A (2006) Vertex-based anisotropic smoothing of 3D mesh data, In: *IEEE CCECE*, pp 202–205
  - 32. Cignoni P, Gobbetti E, Pintus R, Scopigno R (2008) Color enhancement for rapid prototyping. The 9th international symposium on vast international symposium on virtual reality, archaeology and cultural heritage, *Eurographics*, pp 9–16
  - 33. Pitzalis D, Cignoni P, Menu M, Aitken G (2008) 3D enhanced model from multiple data sources for the analysis of the Cylinder seal of Ibni-Sharrum, The 9th international symposium on VAST international symposium on virtual reality, archaeology and cultural heritage, *Eurographics*, pp 79–84
  - 34. Fornberg B (1998) *A practical guide to pseudospectral methods*. Cambridge University Press, Cambridge, UK

35. Chung MK, Shen L, Dalton KM, Evans AC, Davidson RJ (2007) Weighted Fourier representation and its application to quantifying the amount of gray matter. *IEEE Trans Med Imaging* 26:566–581
36. Courant R, Hilbert D (1953) Methods of mathematical physics, II. Interscience, New York, English edition
37. Chung MK, Dalton KM, Shen LL, Evans AC, Davidson RJ (2007) Weighted Fourier series representation and its application to quantifying the amount of gray matter. *IEEE Trans Med Imaging* 26:566–581

## Biography



**Matthew Nitzken** is currently a student of the University of Louisville working on his Ph.D. studies. He holds a B.Eng. and M.Eng. degree in bioengineering from the University of Louisville, Louisville, KY. He has worked on shape analysis and medical imaging studies in his senior and graduate years of the program and plans to pursue this path in his future studies.



**Manuel F. Casanova** received his M.D. from the University of Puerto Rico in 1979. He received specialty training in neurology, neuropathology, and psychiatry. In 2003 he came to the University of Louisville where he is the Vice Chair for Research, Department of Psychiatry. His main focus of research is neurodevelopmental conditions, primarily autism.



**Fahmi Khalifa** received his B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Mansoura, Egypt, in 2003 and 2007, respectively. In May 2009 he joined the BioImaging Laboratory at University of Louisville, Louisville, KY, USA, as a research assistance. His current research is focused on simultaneous image segmentation and registration with main focus on Automatic Diagnosis of Lung Cancer using Contrast Enhancement Computed Tomography Images.



**Guela Sokhadze** is a Psychology Major, and an undergraduate student of the University of Louisville. As well as working part time at BioImaging Laboratory, he assists at the Cognitive Neuroscience Lab in the Health Sciences Center. He is currently involved in multiple Neuroimaging studies of ADHD and Autism.



**Ayman El-Baz** received the B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering Department, University of Louisville, in August 2006. His current research is focused on developing new computer assisted diagnosis systems for different diseases and brain disorders.

# Chapter 8

## Surface Reconstruction and Geometric Modeling for Digital Prosthesis Design

**Luiz C.M. de Aquino, Gilson A. Giraldi, Paulo S.S. Rodrigues,  
Antônio Lopes A. Junior, Jaime S. Cardoso, and Jasjit S. Suri**

**Abstract** The restoration and recovery of a defective skull can be performed through operative techniques to implant a customized prosthesis. Recently, image processing, surface reconstruction, and geometric methods have been used for digital prosthesis design. In this chapter, we review state-of-the-art approaches in this field and discuss related issues. The field of prosthesis modeling may include methods for segmentation and surface reconstruction, geometric modeling, multi-scale methods, and user interaction approaches. So, we present the background in the area by reviewing methods in isosurface extraction from 3D images, deformable models, wavelets, and subdivision surfaces. Then, we discuss some proposals in this area: a balloon model for slice-by-slice bone reconstruction, the T-Surfaces plus isosurface generation models as a general framework for surface reconstruction and user interaction, wavelets-based multiscale methods, and filling holes methods. We describe also experimental results and a computational tool that we are developing for image processing and visualization which can be used for digital prosthesis design. We offer a discussion by presenting some perspectives and issues related to the models described on previous sections. Finally, we present the conclusions of our work.

### 8.1 Introduction

Mending an imperfectskull is necessary not just for cosmetic reasons but also since sizable imperfections might endanger a fundamental section of the brain. Cranioplasty (the procedure of mending imperfections in the cranium with cranial engrafts) is feasible in these cases [38]. However, cranioplasty continues to be a challenge to craniofacial surgeons and neurosurgeons which motivates the development of computational tools for surgical planning [10]. It is even more difficult

---

L.C.M. de Aquino (✉)  
National Laboratory for Scientific Computing, Petrópolis, Brazil  
e-mail: lcaquino@lncc.br

when there is great damage or if it occurs in an area where the structure is fragile or if it involves critical tissues such as the eyes.

So far, a common procedure has not been developed for the conceptualization and fabrication of premanufactured cranial engrafts. Many techniques exist but the treatment is differentiated according to the methodology specifics. Furthermore, each technique fulfills personal customization differently [29, 46].

The use of medical image processing, computer graphics and rapid prototyping (RP) technology permits personal customization to decrease patients' injuries during surgeries [23]. The creation of the prosthesis starts with a computed tomography (CT) of the cranium. Next, image processing methods are implemented for segmentation. In addition, a surface reconstruction technique is implemented in order to acquire a three-dimensional model of the prosthesis, which will become the insert for the following stage when the prosthesis is fabricated using RP technology. Finally, the cranial imperfection is replaced by the prosthesis [26, 29, 48].

There are two categories of approaches in this field, in relation to the image processing and geometric modeling. In the first category, surface rebuild is applied in order to have a representation of the imperfection, which is used later to create the digital model of the missing piece. In [48], we find such method which divides each voxel into five tetrahedra and apply the marching cubes technique to get a polygonal representation of the defective skull. Then, the method takes user-defined guiding points and apply computational geometry algorithms to generate the digital version of the prosthesis. In this process, the geometric model of the prosthesis is obtained taking into account only the polygonal mesh.

The second category includes techniques which rebuild the cranium digitally. Moreover, Boolean operation and the Marching Cubes technique are employed to acquire the prosthesis' digital model. The technique of reflection, built on the assumption of cranium symmetry, is part of this category. From this point, the axis of symmetry is calculated in order to reflect the segmented image in respect to this axis. However, this technique may only be used if the injury does not intersect that axis. If it does, then we must rebuild the missing piece according to the geometry of the cranium surrounding the injury and the soft tissues of the brain. The separation of bone is achieved without difficulty through thresholding techniques. But, separation of soft tissues is not easily accomplished.

For example in [33] there is a case study of a young woman who sustained significant skull damage from a car accident. In her case, computational techniques for quickly creating a customized prosthesis aided the doctors, who used Mimics software [3] to search for a threshold to separate the bone using the CT scan information.

The doctors needed to remove the meninges to make sure that the inside of the engraft did not obstruct the meninges and also to acquire another reference for the creation of the prosthesis. Next, another software, the 3-Matic [1] was used to draw guiding curves on the usable references (part of the meninges and bone particles) to indicate the necessary curvature of the implant. These curved lines compiled to a convex path over the notch of the meninges to preclude the prosthesis from applying force to the meninges.

Because MRI is better at differentiating soft tissue, it is generally preferred over CT for brain imaging which makes the research on CT brain segmentation relatively scarce. However, for prosthesis design, MRI is contraindicated because the main target is the bone tissue. In [26], a study is presented that considers the effectiveness of existing algorithms for segmenting brain tissue in CT images. Three methods (Bayesian classification, Fuzzy c-Means, and Expectation Maximization) were used to segment brain and cerebrospinal fluid. The results shown that these methods outperformed commonly used threshold-based segmentation and points toward the necessity for new imaging protocols for optimizing CT imaging for soft tissue scanning.

Aside from segmentation and image processing, problems with geometric models may occur. Surface reconstruction, filling holes techniques, and multiscale methods must be considered in this avenue.

Among the isosurface methods [44], two types are considered in this chapter: continuation and marching methods. Continuation methods propagate the surface from a set of seed cells. The key idea is to use the spatial coherence of the isosurface during its extraction [8, 44]. In Marching Cubes, each surface-finding phase visits all cells of the volume, normally by varying coordinate values in a triple “for” loop [32]. As each cell that intersects the isosurface is found, the necessary polygon(s) to represent the portion of the isosurface within the cell is generated.

The reconstruction of the lost part of the bone can be cast in the problem of constructing a *n-sided* surface patch that smoothly connects the surfaces that surrounds the polygonal hole. According to [29], n-sided patches generation algorithms fall basically into two classes. In the first class, the polygonal domain is subdivided in the parametric plane. So, triangular or rectangular elements are put together or recursive subdivision methods are applied. In the second class, one uniform equation is used as a combination of 3D constituents. In this case, the surface equation can be computed by either generalized control-point-based methods or through a weighted sum of 3D interpolants.

On the other hand, multiscale methods may be applied: having detected a structure of interest in a coarser scale propagate the results through finer scales until the finest one. Behind such approach, there is the hypothesis that at the lower resolution, small details become less significantly relative to the structure of interest.

All these methods and algorithms must be incorporated in a software to provide surgeons with a computational tool for prosthesis design and surgical planning. Modeling and visualization systems have revolutionized many scientific disciplines by providing novel methods to visualize complex data structures and by offering the means to manipulate this data in real time. With these systems, the surgeons are able to: navigate through the anatomy, practice both established and new procedures, learn how to use new surgical tools, and assess their progress. Digital prosthesis design is a beautiful example of such application. Softwares such as Mimics, InVesalius, and 3-Matic [1, 5] incorporate image processing and geometric design methods to allow medical teams to customize the prosthesis before the implant. In this context, user interaction becomes an important issue.

The techniques mentioned above are applied with the hope to automatically get the geometric model of the prosthesis. However, only the medical team knows exactly the requirements for each specific case. Besides, our algorithms always have a percentage of error. Therefore, we must provide user interaction strategies to the surgeons. We can use traditional mouse-controlled interaction with the three-dimensional model on a standard graphics monitor or virtual reality devices combined with stereo viewing. We can also provide the user with an intuitive sculpting system such that user should be able to freely deform, add, and remove material in the digital model.

This chapter covers these issues and is organized as follows. We review related works in Sect. 8.2. Section 8.3 presents the background methods. In Sect. 8.4, we describe the proposed models to address the mentioned issues. Along this section experimental results are also presented. Next, in Sect. 8.6, we offer a discussion by presenting some perspectives and issues related to the models described on previous sections. Finally, we offer the conclusions of our work (Sect. 8.7).

## 8.2 Related Works

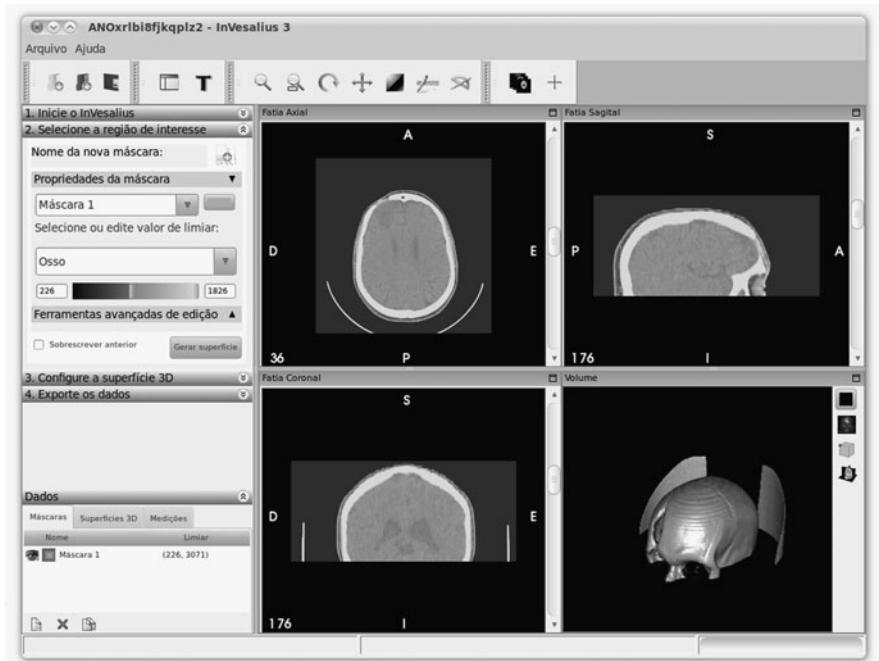
In older surgical procedures, cranioplasty implants were manufactured directly in the operating-theater where the surgeon modeled the suitable material by hand, namely polymethylmethacrylate (PMMA), to build the prosthesis. The PMMA takes a few minutes to acquire shape and become ready to be implanted in the patient. But, in order for this procedure to be successful, the surgeon needs to have splendid manual ability and still there is an increased chance of infection and rejection [15].

In the early 1980s, the introduction of 3D CT scanners and surface reconstruction methods provided a new effective tool for surgical planning. The possibility of generating a geometric (digital) model of the lost part combined with stereolithography rapid prototyping technologies (RP) can make skull prostheses manufacture more accurate and customized for each case [28].

Besides image acquisition, the avenue in this area includes material science, algorithms, and computational systems. To be ideal for cranioplasty, the material must have properties such as capable of growth, resistant to infection, noncorrosive, stable, inert, among many others (see for a recent study [19]). Up to now, no currently available materials satisfy all the requirements and the research of new materials still remains an active area.

The primary objective of the algorithms in this field is the virtual rebuild of the skull imperfection. It has been performed using (whenever possible) a reflection of the cranium as a guide to fill the holes or by drawing new pixels manually [15].

In order to evaluate the 3D morphology of the bone, thresholding methods were used to exclude the soft tissues. Afterward, marching methods may be used to acquire the bone's polygonal representation. From the marching result, the



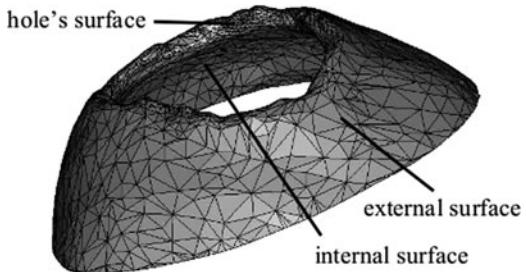
**Fig. 8.1** InVesalius software and surface reconstruction from CT images

creation of the digital model of the implant may begin. This model may also be a polygonal surface or a parametric one (NURBS), according to the fabrication method applied [10, 26]. Mimics, 3-Matic and the free software in Vesalius (Fig. 8.1) are examples of computational systems which are accessible and render these properties [1, 3, 5].

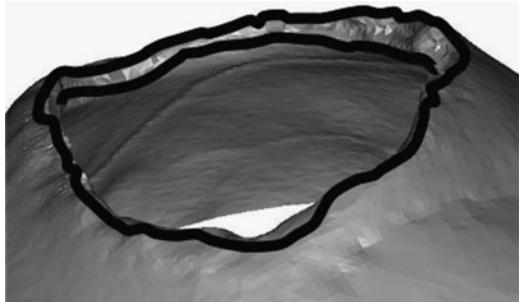
Despite of the simplicity of these functionalities, the above software, have been applied in various surgical planning cases found in the medical literature [15]. This demonstrates that the development of innovative algorithms is a promising research area.

For instance, let us see the method presented in [49], which, to our knowledge, is a state-of-the-art method for digital prosthesis design. At first, user chooses a segment of the image to preserve memory space. Afterward, each voxel is separated into five tetrahedral parts and the marching cubes method is used to acquire a polygonal picture of the imperfect cranium (see Fig. 8.2). Obeying the nomenclature in [49], the respective model has the internal, external and hole surfaces (Fig. 8.2). Next, the boundary link is removed, meaning, the pathway made of edges which is shared by the internal (external) and hole surfaces. In [49] the authors affirm that the angle of the two triangles along the boundary link is smaller. This is the dihedral angle criterion which the algorithms are based upon. Only the edges which meet the three rules obey this criterion:

**Fig. 8.2** Defective skull geometry recovered by Marching Cubes (reprinted from [48])



**Fig. 8.3** External and internal boundary links (reprinted from [48])



1. Two vertices of edge are inside the region link.
2. The direction of the edge is closest to the direction of a guiding vector, defined through two guiding points specified by the user.
3. The dihedral angle on the edge is smallest.

Two sets of guiding points must be selected, one for the outside and other for the inside boundary link, to guide the searching algorithm (details in [49]). The result, shown in Fig. 8.3, is used to extract the hole's surface. Now, each boundary link is filled by triangulating [9] the corresponding 3D polygons. So, let a polygon  $P = (v_0, v_1, \dots, v_{n-1}, v_n = v_0)$ , with  $v_i \in V = \mathbb{R}^3, i = 0, 1, 2, \dots, n$ , and a weight function  $\Omega : V^3 \rightarrow L$ , where  $L$  is a weight set and  $\Omega$  assigns a weight (area, for instance) to each triangle whose vertices are in  $P$ . For  $0 \leq i < j \leq n - 1$ , let  $W_{i,j}$  be the weight of the minimum-weight triangulation of the sub-polygon  $(v_i, v_j)$ , computed as follows:

1. For  $i = 0, 1, \dots, n - 2$ , let  $W_{i,i+1} \leftarrow 0$ . For  $i = 0, 1, \dots, n - 3$ , let  $W_{i,i+2} \leftarrow \Omega(v_i, v_{i+1}, v_{i+2})$ . Set  $j \leftarrow 2$ .
2. Put  $j \leftarrow j + 1$ . For  $i = 0, 1, \dots, n - j - 1$  and  $k \leftarrow i + j$ , compute:

$$W_{i,k} \leftarrow \min_{i < m < k} [W_{i,m} + W_{m,k} + \Omega(v_i, v_m, v_k)] \quad (8.1)$$

Let  $O_{i,k}$  be the index  $m$  where the minimum is achieved.

3. If  $j < n - 1$ , go back to step 2. Otherwise,  $W_{0,n-1}$  holds the weight of the minimum-weight triangulation.

4. Let  $S \leftarrow \emptyset$ . Call the recursive function *Trace* with parameters  $(0, n - 1)$ .

Function *Trace* (i, k)

if  $i + 2 = k$  then

$S \leftarrow S \cup \Delta v_i v_{i+1} v_k$

else

Let  $o \leftarrow O_i, k$

if  $o \neq i + 1$  then *Trace* (i, o)

$S \leftarrow S \cup \Delta v_i v_o v_k$

if  $o \neq k - 1$  then *Trace* (o, k)

endif

In Fig. 8.5 the result is demonstrated after fulfilling the outside boundary link with the recorded triangles. This triangulation must be made better, by adding new vertices and performing edge relaxation in order to acquire a Delaunay-like triangulation (Fig. 8.4) [34]. Fundamentally, given two triangles  $T_1$  and  $T_2$ , along side the edge  $E$ , we compute the circum-sphere of the triangle  $T_1$  and see if the vertex  $v \in T_2$ , on the other edge side, is inside this circum-sphere. The edge is switched in this case. The algorithm also uses a density control factor as a stopping criterion.

Conclusively, this smoothing procedure is used. Let  $\omega : V^2 \rightarrow \mathfrak{R}$  be a weighting function defined on the surface edges. Also, define the weighted umbrella-operator  $U_\omega : V \rightarrow E$  from  $V$  to the Euclidean space  $E$ :

$$U_\omega(v) = -v + \left[ \frac{1}{\sum_i \omega(v, v_i)} \right] \sum_{v_i \in B} \omega(v, v_i) v_i, \quad (8.2)$$

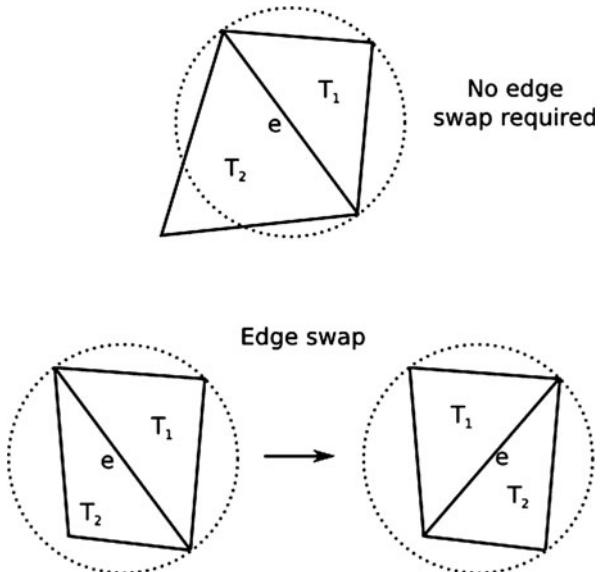
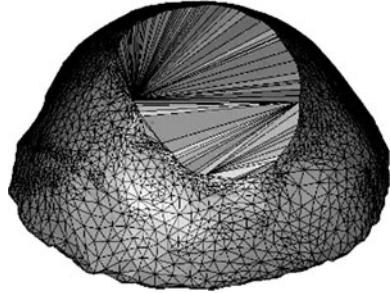


Fig. 8.4 Triangulated hole  
(reprinted from [48])

**Fig. 8.5** Process of edge relaxation (reprinted from [48])



where  $B$  is the set of direct neighbors of  $v$ . So, given a vertex  $v$ , we compute expression (8.2) and replace  $v$  with  $v + U_\omega(v)$  to get a smoother version of the surface. Given an edge  $(v_i, v_j)$ , the weights  $\omega$  can be computed by the expression:

$$\omega(v_i, v_j) = \cot(\angle(v_i, v_{k1}, v_j)) + \cot(\angle(v_i, v_{k2}, v_j)),$$

where  $(v_i, v_{k1}, v_j)$  and  $(v_i, v_{k2}, v_j)$  are the two adjacent triangles of  $(v_i, v_j)$ .

### 8.3 Background

The field of digital prosthesis modeling includes methods for segmentation, surface reconstruction, and computational geometry. Also, multiscale and geometric modeling methods can be used. Besides, user interaction must be considered. So, we present the background in the area by reviewing methods in isosurface extraction, deformable models, wavelets, and subdivision surfaces. These methods are the background for the next sections.

#### 8.3.1 Mass-Spring System and T-Surfaces

In the mass-spring system, the surface nodes work as masses and the edges define the linear springs with damping. So, given a particle  $i$  with mass  $m_i$  and position vector  $x_i$ , the force system is composed by the elastic ( $f_{\text{elastic}}^i$ ), gravitational ( $f_{\text{grav}}^i$ ), and damping ( $f_{\text{damp}}^i$ ) forces, defined, respectively, by [21, 45]:

$$\mathbf{f}_{\text{elastic}}^i = \sum_{j \in V} c_{ij} (l_{ij} - \| \mathbf{x}_i - \mathbf{x}_j \|) \frac{(\mathbf{x}_i - \mathbf{x}_j)}{\| \mathbf{x}_i - \mathbf{x}_j \|}, \quad (8.3)$$

where  $V$  is the set of nodes linked to  $\mathbf{x}_i$ ,  $c_{ij}$  is the stiffness of the spring linking the nodes  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , and  $l_{ij}$  the spring rest length;

$$\mathbf{f}_{\text{grav}}^i = m_i \mathbf{g}, \quad (8.4)$$

$$\mathbf{f}_{\text{damp}}^i = \gamma_i \dot{x}_i, \quad (8.5)$$

$\mathbf{g}$  is the gravity field and  $\gamma_i$  is the damping factor. Following Newton's Laws, we get the following evolution equation:

$$m_i \ddot{\mathbf{x}}_i = \mathbf{f}_{\text{elastic}}^i + \mathbf{f}_{\text{damp}}^i + \mathbf{f}_{\text{grav}}^i. \quad (8.6)$$

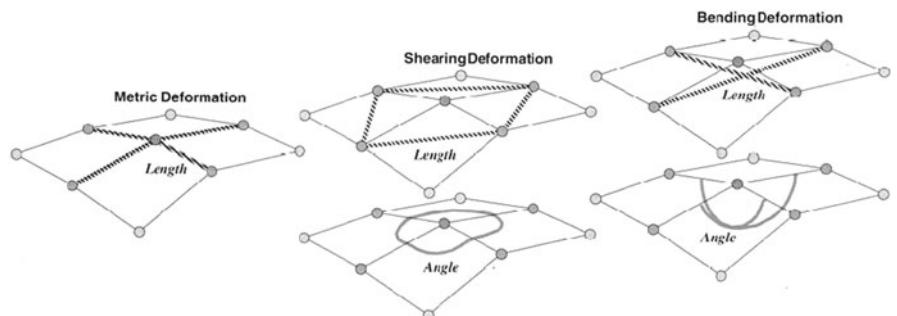
In this development, there is a one-to-one relationship between the mesh topology and the springs. We can relax such coupling by allowing the interaction between particles that are not connected by any edge. Following this idea, we can include shearing and bending besides material tension without including extra edges. Figure 8.6 represents this idea by showing a regular mesh, in which the particles are connected by structural springs to account for tension, as well as additional springs that do not correspond to edges. They are diagonal springs for shearing and interleaving springs for bending [45]. Each spring is governed by Hook's Law and all the corresponding forces can be computed by expressions analogous to (8.3)–(8.5).

The T-Surfaces approach is a topologically adaptable deformable model which is composed basically by three components [33]: (1) a tetrahedral decomposition (CF-Triangulation) of the image domain  $D \subset \mathbb{R}^3$ ; (2) a mass-spring system; (3) a *Characteristic Function*  $\chi$  defined on the grid nodes which distinguishes the interior ( $\text{Int}(S)$ ) from the exterior ( $\text{Ext}(S)$ ) of a surface  $S$ :

$$\chi : D \subset \mathbb{R}^3 \rightarrow \{0, 1\}, \quad (8.7)$$

where  $X(p) = 1$  if  $p \in \text{Int}(S)$  and  $X(p) = 0$ , otherwise, where  $p$  is a node of the grid.

Following the classical nomenclature, a vertex of a tetrahedron is called a *node* and the collection of nodes and triangle edges is called the grid  $\Gamma_s$ . A tetrahedron



**Fig. 8.6** Square mesh with additional links to model bending and shearing forces. There are no edge supporting the additional links (reprinted from [45])

(also called a simplex)  $\sigma$  is a *transverse* one if the characteristic function  $\chi$  in (8.7) changes its value in  $\sigma$ , analogously, for an edge.

In this framework, the reparameterization of a surface is done by [33]: (1) taking the intersections points of the surface with the grid; (2) find the set of transverse tetrahedra (*Combinatorial Manifold*); (3) for each transverse edge choose an intersection point belonging to it; (4) connect these points properly.

In this reparameterization process, the transverse simplices play a central role. Given such a simplex, we choose in each transverse edge an intersection point to generate the new surface patch. In general, we will have three or four transverse edges in each transverse tetrahedron (Fig. 8.8). The former gives a triangular patch and the later defines two triangles. So, at the end of the step (4) we have a triangular mesh. Each triangle is called a *triangular element* [33].

As an example for 2D, consider the characteristic functions ( $\chi_1$  and  $\chi_2$ ) relative to the two contours pictured in Figure 8.7. The functions are defined on the vertices of a CF-triangulation of the plane. The vertices marked are those where  $\max\{X_1, X_2\} = 1$ . Observe that they are enclosed by a merge of the contours. This merge can be approximated by a curve belonging to the region obtained by tracing the transverse triangles. The same would be true for more than two contours (and obviously for only one).

Besides the forces (8.3) and (8.5), the model also has a normal force which can be weighted as follows [33]:

$$F_i = k(\text{sign}_i)n_i, \quad (8.8)$$

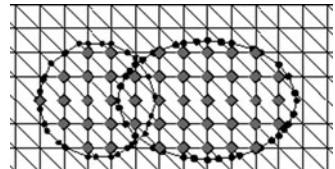
where  $n_i$  is the normal vector at node  $i$ ,  $k$  is a scale factor, and  $\text{sign}_i = +1$  if  $I(v_i) > T$  and  $\text{sign}_i = -1$  otherwise ( $T$  is a threshold for image  $I$ ). This force is used to push the model toward image edges until it is opposed by external image forces.

The forces given in expressions (8.3) and (8.8) are internal forces. The external force is defined as a function of the image data, according to the features we seek. One possibility is:

$$\text{Image} :: \text{Force} :: f_i^t = -\gamma_i \nabla P; \quad P = \|\nabla I\|^2. \quad (8.9)$$

The evolution of the surface is governed by the following dynamical system:

$$v_i^{(t+\Delta t)} = v_i^t + h_i \left( \overrightarrow{\alpha}_i^t + \overrightarrow{F}_i^t + \overrightarrow{f}_i^t \right), \quad (8.10)$$



**Fig. 8.7** Two snakes colliding with the inside grid nodes and snaxels marked

where  $h_i$  is an evolution step. During the T-Surfaces evolution, some grid nodes become interior to a surface. Such nodes are called *burnt nodes* and its identification is fundamental to update the characteristic function [33]. To deal with self-intersections of the surface, the T-Surfaces model incorporates an entropy condition: *once a node is burnt it stays burnt*. A termination condition is obtained based on the number of deformations steps that a simplex has remained a transverse one.

### 8.3.2 Shape Model

We consider an object to be described by points, referred to as landmark points. These points may represent the boundary of the object or some structures of interest in the images of a database.

In general, the landmark points are (manually) determined in a set of  $s$  training images. From these collections of landmark points, a point distribution model [16] is constructed as follows. The landmark points are stacked in shape vectors:

$$\mathbf{x}^i = (x_1^i, y_1^i, x_2^i, y_2^i, \dots, x_n^i, y_n^i)^T, \quad i = 1, 2, \dots, s, \quad (8.11)$$

where  $(x_j^i, y_j^i)$  refers to the landmark  $j$  for the image  $i$ .

Principal component analysis (PCA), or KL-Transform [16, 18, 22], is applied to the shape vectors by computing the mean shape:

$$\bar{\mathbf{x}} = \frac{1}{s} \sum_{i=1}^s \mathbf{x}^i, \quad (8.12)$$

the covariance matrix:

$$\mathbf{S} = \frac{\mathbf{1}}{s-1} \sum_{i=1}^s (\mathbf{x}^i - \bar{\mathbf{x}})(\mathbf{x}^i - \bar{\mathbf{x}})^T, \quad (8.13)$$

and the eigenvectors of the covariance matrix sorted in decreasing order of the corresponding eigenvalues:

$$\{\phi_1, \phi_2, \dots, \phi_{2n}\}. \quad (8.14)$$

The eigenvectors corresponding to the largest  $t$  eigenvalues  $\lambda_i$  are retained in a matrix  $\Phi$ . A shape  $\mathbf{x}$  can be now approximated by:

$$\min_{\mathbf{b}} \|\mathbf{x} - (\bar{\mathbf{x}} + \Phi \cdot \mathbf{b})\|, \quad (8.15)$$

where  $\mathbf{b}$  is a vector of  $t$  elements containing the model parameters.

### 8.3.3 Surface Reconstruction

Isosurface extraction is one of the most used techniques for surface reconstruction in 3D data sets. Depending on the data type (time-varying or stationary) and the data size, many works have been done to improve the basic methods in this area [44]. In this chapter, we consider two kinds of isosurface generation methods: the marching ones and continuation ones.

In Marching Cubes, each surface-finding phase visits all cells of the volume, normally by varying coordinate values in a triple “for” loop [32]. As each cell that intersects the isosurface is found, the necessary polygon(s) to represent the portion of the isosurface within the cell is generated. There is no attempt to trace the surface into neighboring cells. Space subdivision schemes (like Octree and k-d-tree) have been used to avoid the computational cost of visiting cells that the surface do not cut [13, 44].

Once the T-Surfaces grid is a CF one, the Tetra-Cubes is a natural choice [11]. Like in the marching cubes, its search is linear: each cell of the volume is visited and its simplices (tetrahedron) are searched to find surfaces patches. Following marching cubes, its implementation uses auxiliary structures based on the fact that the topology of the intersections between a plane and a tetrahedron can be reduced to three basic configurations pictured in Fig. 8.8.

Unlike marching methods, continuation algorithms attempt to trace the surface into neighboring simplices [8]. Thus, given a transverse simplex, the algorithm searches its neighbors to continue the surface reconstruction. The key idea is to generate the combinatorial manifold (set of transverse simplices) that holds the isosurface.

The following definition will be useful. Let us suppose two simplices  $\sigma_0, \sigma_1$ , which have a common face and the vertices,  $v \in \sigma_0$  and  $v' \in \sigma_1$ , both opposite to the common face. The process of obtaining  $v'$  from  $v$  is called *pivoting*. Let us present the basic continuation algorithm [8].

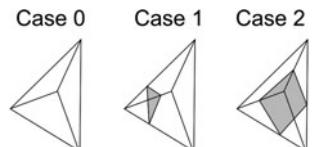
#### PL Generation Algorithm:

*Find a transverse triangle  $\sigma_0$ ;*

$\sum = \{\sigma_0\}; V(\sigma) = \text{set of vertices of } \sigma;$

*while  $V(\sigma) \neq 0$  for some  $\sigma \in \sum$*

- *get  $\sigma \in \Sigma$  such that  $V(\sigma) \neq 0$ ;*
- *get  $v \in V(\sigma)$ ;*
- *obtain  $\sigma'$  from  $\sigma$  by pivoting  $v$  into  $v'$*



**Fig. 8.8** Basic types of intersections between a plane and a simplex in 3D

- if  $\sigma'$  is not transverse
- then drop  $v$  from  $V(\sigma)$ ;
- else
- if  $\sigma' \in \Sigma$  then
- drop  $v$  from  $V(\sigma), v'$  from  $V(\sigma')$
- else
- $\Sigma \leftarrow \Sigma + \sigma';$
- $V(\sigma') \leftarrow \text{set of vertices of } \sigma';$
- drop  $v$  from  $V(\sigma), v'$  from  $V(\sigma')$

Differently from Tetra-Cubes, once started the generation of a component, the algorithm runs until it is completed. However, the algorithm needs a set of seed simplices to be able to generate all the components of an isosurface. This is an important point when comparing continuation and marching methods.

If we do not have guesses about seeds, every simplex should be visited. Thus, the computational complexity of both methods are the same ( $O(N)$  where  $N$  is the number of simplices). However, if we know in advance that the target boundary is connected, we do not need to search for inside components. Thus, the computational cost is reduced if compared with Tetra-Cubes. That is why we use continuation methods in [17, 41] to get the initial surfaces.

### 8.3.4 Wavelet Theory and Multiscale Analysis

Wavelets can be introduced in the context of basis functions and scale-varying basis functions. Historically, such theory received a great impact in the 1930s, when several groups working independently had established some foundations of such theory and its applications. For instance, using the scale-varying Haar basis function, the physicist Paul Levy investigated at that time the Brownian motion. He found the Haar basis function superior to the Fourier basis functions for that study [20]. In 1980, Grossman and Morlet, a physicist and an engineer, broadly defined wavelets in the context of quantum physics. In 1985, Stephane Mallat discovered some relationships between quadrature mirror filters, pyramid algorithms, and orthonormal wavelet bases. A couple of years later, Ingrid Daubechies used Mallat's work to perform perhaps the most elegant development in this field. The set of wavelet orthonormal basis functions constructed has become a very remarkable work for wavelet applications since then [20].

*Definition 5:* A multiscale analysis (MSA) of  $L^2(\mathfrak{R})$  is an increasing sequence of closed subspaces, called **scale spaces**,  $V_m \subset L^2(\mathfrak{R})$  :

$$\{0\} \subset \cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \subset L^2(\mathfrak{R}) \quad (8.16)$$

such that the following are true:

$$\overline{\bigcup_{m \in \mathbb{Z}} V_m} = L^2(\mathfrak{R}), \quad (8.17)$$

$$\bigcap_{m \in \mathbb{Z}} V_m = \{0\}, \quad (8.18)$$

$$f(x) \in V_m \iff f(2^m x) \in V_0. \quad (8.19)$$

There is a function  $\varphi \in L^2(\mathfrak{R})$  whose integer translates generate a Riesz basis of  $V_0$  (that mean,  $V_0$  is the closure of the set  $\text{span}\{\varphi_{m,k} | k \in \mathbb{Z}\}$ ) :

$$V_0 = \overline{\text{span}\{\varphi(x - k) | k \in \mathbb{Z}\}}$$

and

$$A \sum_{k \in \mathbb{Z}} c_k^2 \leq \left\| \sum_{k \in \mathbb{Z}} c_k \varphi(\cdot - k) \right\|_{L^2}^2 \leq B \sum_{k \in \mathbb{Z}} c_k^2 \quad (8.20)$$

for all  $\{c_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$ .  $A$  and  $B$  are positive constants.

It is important to emphasize the following remarks:

(a) Conditions (8.17) and (8.18) are satisfied by many families  $\{V_m\}_{m \in \mathbb{Z}}$ . Property (8.19) is the special feature of an MSA: the spaces  $V_m$  are scaled versions of the basic space  $V_0$ , which is spanned by the translations of  $\varphi$ , the scaling function.

For  $m \rightarrow \infty$ , the functions in  $V_m$  are dilated, i.e., their details enlarged. If  $m$  tends to  $-\infty$ , then the spaces  $V_m$  contain smaller and smaller structures. The limits

$$\lim_{m \rightarrow +\infty} \|P_m f\|_{L^2} = 0, \quad (8.21)$$

$$\lim_{m \rightarrow -\infty} \|P_m f - f\|_{L^2} = 0, \quad (8.22)$$

make this interpretation precise.  $P_m$  denotes the orthogonal projector onto  $V_m$ . The following phrase has been adopted, but it is more suggestive than exact:  $P_m f$  is the representation of  $f$  on the “scale”  $V_m$  and contains all details of  $f$  up to the size  $2^m$ .

(b) The relation (8.20) implies that  $V_0$  is invariant under integer translation

$$f \in V_0 \iff f(x - k) \in V_0, \text{ for } k \in \mathbb{Z}. \quad (8.23)$$

With (8.19), it follows that:

$$f \in V_m \iff f(x - 2^m k) \in V_m, \text{ for } k \in \mathbb{Z}. \quad (8.24)$$

(c) The space  $V_m$  is spanned by the functions:

$$\varphi_{m,k}(x) := 2^{-m/2} \varphi(2^{-m}x - k), \quad (8.25)$$

$$V_m = \overline{\text{span}\{\varphi_{m,k} \mid k \in \mathbb{Z}\}}. \quad (8.26)$$

This is based on (8.19) and (8.20). The functions in (8.25) all have the same  $L^2$ -norm  $\|\varphi_{m,k}\|_{L^2} = \|\varphi\|_{L^2}$ .

*Scaling Equation:* The scaling function  $\varphi$  satisfies a scaling equation, i.e., there is a sequence  $\{h_k\}_{k \in \mathbb{Z}}$  of real numbers such that:

$$\varphi(x) = \sqrt[2]{2} \sum_{k \in \mathbb{Z}} h_k \varphi(2x - k). \quad (8.27)$$

The key to the construction of both orthogonal wavelet bases (wavelet spaces) and fast algorithms lies in (8.27). In this way, we denote by  $W_m$  the orthogonal complement of  $V_m$  in  $V_{m-1}$ , that means:

$$V_{m-1} = W_m \oplus V_m, V_m \perp W_m. \quad (8.28)$$

If  $Q_m$  is the orthogonal projector of  $L^2(\mathfrak{R})$  in  $W_m$  and  $P_{m-1}$  denotes the orthogonal projector onto  $V_{m-1}$  (likewise in expression (8.21)–(8.22)), then decomposition (8.28) means:

$$P_{m-1} = Q_m + P_m. \quad (8.29)$$

From (8.28), it is clear that:

$$V_{m-1} = W_m \oplus (W_{m+1} \oplus V_{m+1}), V_{m+1} \perp W_{m+1}, \quad (8.30)$$

and so on. Therefore, using the MSA definition, it follows that:

$$V_{m-1} = \bigoplus_{j \geq m} W_j \quad (8.31)$$

and so:

$$L^2(\mathfrak{R}) = \bigoplus_{j \in \mathbb{Z}} W_j. \quad (8.32)$$

The spaces  $W_m$ , with  $m \in \mathbb{Z}$ , are named **wavelet spaces**.

Let  $\{V_m\}_{m \in \mathbb{Z}}$  be an MSA generated by the orthogonal scaling function  $\varphi \in V_0$  and the function  $\psi \in V_{-1}$ , defined by

$$\psi(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k \varphi(2x - k) = \sum_{k \in \mathbb{Z}} g_k \varphi_{-1,k}(x), \quad (8.33)$$

$$g_k = (-1)^k h_{1-k}, \quad (8.34)$$

where  $\{h_k\}_{k \in \mathbb{Z}}$  are the coefficients of the scaling equation (8.27).

*Theorem 1:* The function  $\psi$  defined by expressions (8.33)–(8.34) has the following properties:

- (i)  $\{\psi_{m,k}(\cdot) = 2^{-m/2}\psi(2^{-m}x - k) \mid k \in \mathbb{Z}\}$  is an orthonormal basis for  $W_m$ ,
- (ii)  $\{\psi_{m,k} \mid m, k \in \mathbb{Z}\}$  is an orthonormal basis for  $L^2(\mathfrak{R})$ ,
- (iii)  $\psi$  is a wavelet; that means:  $c_\psi = 2\pi \int_{-\infty}^{\mathfrak{R}} |\omega|^{-1} |\psi(\omega)|^2 d\omega = 2 \ln 2 < \infty$ .

In Sect. 8.5.1, we will discuss the application of this approach do help the definition of the boundary of the hole pictured in Fig. 8.3.

### 8.3.5 2D Deformable Model Balloon

Deformable models are very useful for shape recovery in 2D/3D images. In our case, we will apply a 2D deformable model, a balloon-like one [14]. Geometrically, this model is described by a parametric contour  $c$  embedded in a domain  $D \subset \mathfrak{R}^2$ :

$$\begin{aligned} c : [0, 1] &\rightarrow D \subset \mathfrak{R}^2, \\ c(s) &= (x(s), y(s)). \end{aligned} \quad (8.35)$$

From the dynamic viewpoint, we have a deformable contour which is viewed as a time-varying curve:

$$c(s, t) = (x(s, t), y(s, t)). \quad (8.36)$$

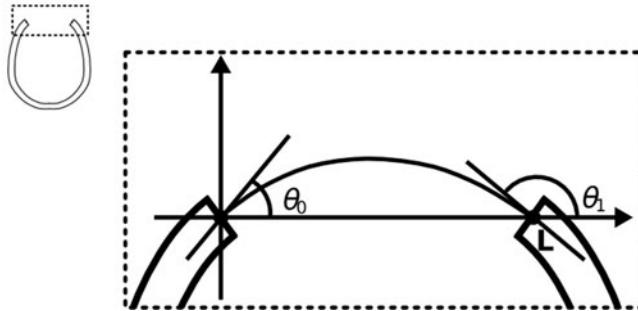
In this formulation, the Lagrange equations of motion can be expressed as:

$$\dot{c} - (\omega_1 c')' + (\omega_2 c'')'' = \vec{F}_{ext} + k \vec{n}(c), \quad (8.37)$$

subject to the specification of the following boundary conditions:

$$c(0, t), c(1, t), c'(0, t), c'(1, t), \quad (8.38)$$

where  $\dot{c} \equiv \frac{\partial c}{\partial t}$ ,  $c' \equiv \frac{\partial c}{\partial s}$ ,  $c'' \equiv \frac{\partial^2 c}{\partial s^2}$ ,  $\vec{n}(c)$  is the unit normal over the curve, and  $k$  is a scale factor. The field  $\vec{F}_{ext}$  means an external force, which depends on image elements or constraints.



**Fig. 8.9** Reference frame for balloon initialization

In our case, the curve is an open one, fixed in the end points of the lesion following the tangents at that points. Besides these boundary conditions, we must set the initial curve to complete the initialization of the balloon.

So, to include the initial curve in the numerical method, we must give a sequence of points  $C^0 = c(i\Delta s, 0) = c_i^0$ , com  $i = 0, \dots, N$ , which is the discrete version of the initial curve. Such curve must satisfy the boundary conditions also. Therefore, using a reference frame as pictured in Fig. 8.9, which has both the end points on the horizontal axis (one in the origin and the other one at  $x = L$ ), we can write the constraints as:

$$\begin{aligned} C(0) &= 0, \\ C(L) &= 0, \\ C'(0) &= \tan \theta_0, \\ C'(L) &= \tan \theta_1 \end{aligned} \quad (8.39)$$

Besides, the initial shape should be as closer as possible to the target to reduce time computation. We observe that a cubic curve represented by  $C(x) = ax^3 + bx^2 + cx + d$  fits all of these requirements. With a simple algebra, we can demonstrate that:

$$a = \frac{\tan \theta_0 + \tan \theta_1}{L^2}, \quad (8.40)$$

$$b = -\frac{2 \tan \theta_0 + \tan \theta_1}{L}, \quad (8.41)$$

$$c = \tan \theta_0, \quad (8.42)$$

$$d = 0. \quad (8.43)$$

By considering that  $s \in [0, 1]$  in the balloon definition, we can represent the initial curve as:

$$c(s) = (Ls, L(\tan \theta_0 + \tan \theta_1)s^3 - L(2 \tan \theta_0 + \tan \theta_1)s^2 + \tan \theta_0 s). \quad (8.44)$$

We can approximate the derivatives in the expression (8.37) through finite differences, with discretization steps of  $\Delta s = 1/N$  and  $\Delta t = 1/M$  for space and time, respectively, and by considering  $\omega_1, \omega_2$  constants:

$$\dot{c} \approx \frac{c_i^\tau - c_i^{\tau-1}}{\Delta t},$$

$$c'' \approx \frac{1}{\Delta s^2} [c_{i-1}^\tau - 2c_i^\tau + c_{i+1}^\tau],$$

$$c''' \approx \frac{1}{\Delta s^4} [c_{i-2}^\tau - 4c_{i-1}^\tau + 6c_i^\tau - 4c_{i+1}^\tau + c_{i+2}^\tau],$$

where  $c_i^\tau = c(i\Delta s, \tau\Delta t)$ , with  $i = 0, \dots, N$  and  $\tau = 0, \dots, M$ .

We shall observe that once the *snake* is an open curve with fixed endpoints (say,  $P_0 \in P_N$ ), then  $c_0^\tau = P_0$  and  $c_N^\tau = P_N$ , for any  $\tau$ . Also, using the expressions:

$$c'(0, t) \approx \frac{c_1^\tau - c_0^\tau}{\Delta s}, \quad (8.45)$$

$$c'(1, t) \approx \frac{c_N^\tau - c_{N-1}^\tau}{\Delta s}, \quad (8.46)$$

and considering that  $c'(0, t) = \mathbf{u}_0$  and  $c'(1, t) = \mathbf{u}_1$ , we get  $c_1^\tau = P_0 + \Delta s \mathbf{u}_0$  and  $c_{N-1}^\tau = P_N - \Delta s \mathbf{u}_1$  for any  $\tau$ . Therefore, the points  $c_0^\tau, c_1^\tau, c_{N-1}^\tau$  and  $c_N^\tau$  are known along the whole balloon evolution.

From this fact and substituting the derivative expressions in (8.37), we get [14]:

$$(I + \Delta t A) C^\tau = C^{\tau-1} + \Delta t (kF - L), \quad (8.47)$$

where  $I$  is the identity matrix of order  $(N - 3)$ ,  $C^\tau = [c_2^\tau \dots c_{N-2}^\tau]^T$ ,

$$F = [\mathbf{n}(c_2^{\tau-1}) \dots \mathbf{n}(c_{N-2}^{\tau-1})]^T,$$

$$L = [l_1 \ l_2 \ 0 \ \dots \ 0 \ l_3 \ l_4]^T,$$

$$A = \begin{bmatrix} a_0 & a_1 & a_2 & & & \\ a_1 & a_0 & a_1 & a_2 & & \\ a_2 & a_1 & a_0 & a_1 & a_2 & \\ & a_2 & a_1 & a_0 & a_1 & a_2 \\ & & & \ddots & & \\ & a_2 & a_1 & a_0 & a_1 & a_2 \\ & & a_2 & a_1 & a_0 & a_1 & a_2 \\ & & & a_2 & a_1 & a_0 & a_1 \\ & & & & a_2 & a_1 & a_0 \end{bmatrix},$$

where:

$$\begin{aligned} a_2 &= \frac{\omega_2}{\Delta s^4} \\ a_1 &= -\frac{\omega_1}{\Delta s^2} - 4 \frac{\omega_2}{\Delta s^4} \\ a_0 &= 2 \frac{\omega_1}{\Delta s^2} + 6 \frac{\omega_2}{\Delta s^4} \\ l_1 &= a_2 P_0 + a_1 (P_0 + \Delta s \mathbf{u}_0) \\ l_2 &= a_2 (P_0 + \Delta s \mathbf{u}_0) \\ l_3 &= a_2 (P_N - \Delta s \mathbf{u}_1) \\ l_4 &= a_2 P_N + a_1 (P_N - \Delta s \mathbf{u}_1) \end{aligned}$$

In this way, the system (8.47) gives a method to compute  $C^{t+1}$  once we have  $C^{t-1}$ . Besides, we shall notice that  $(I + \Delta t A)$  is a symmetric pentadiagonal banded matrix. So, it is computationally efficient to solve the system through LU decomposition of the matrix  $A$  [37]. The balloon comes to rest when the internal and external forces balance, which implies that:

$$\|C^{t+\Delta t} - C^t\| \leq \varepsilon, \quad (8.48)$$

where  $C^{t+\Delta t}$  and  $C^t$  are the curves at time  $t + \Delta t$  and  $t$ , respectively, and  $\varepsilon$  is a predefined parameter.

The method is sensitive to the choice of the coefficients  $\omega_1$  and  $\omega_2$ . Following [14], we set these parameters such that  $\omega_1$  and  $\omega_2$  are proportional to  $\Delta s^2$  and  $\Delta s^4$  respectively.

### 8.3.6 Subdivision Scheme

In the last decade, subdivision surfaces have found their way into wide applications in geometric modeling and animation. One reason for this fact is that subdivision is

intricately linked to multiresolution. Besides, constructing surfaces through subdivision elegantly addresses issues of covering arbitrary topology, level-of-detail modeling, numerical requirements for further application in finite element, elegant formulation through mathematical tools such as wavelets, etc.

For instance, let us consider the Sabin's algorithm [38] which is a variant of the traditional Catmull–Clark's method [12]. The input for Sabin's algorithm is a closed net  $N = (V, E)$ , where  $V$  is the set of vertices connected according to a topology  $E$  (a net is closed if each edge is shared by exactly two faces). Then, the subdivision algorithm will compute a new net  $N' = (V', E')$  by applying the following rules over the input net  $N$  (see [38] for details):

1. For each old face  $f$ , make a new face-vertex  $v_f^*$  as the weighted average of the old vertices of  $f$ , with weights  $W_n$  that depend on the valency  $n$  of each vertex.
2. For each old edge  $e$ , make a new edge-vertex  $v_e^*$  as the weighted average of the old vertices of  $e$  and the new face vertices associated with the two faces originally sharing  $e$ . The weights  $W_n$  (which are the same as that used in rule 1) depend on the valency  $n$  of each vertex (the valency of new vertices is equal to the one of the corresponding faces).
3. For each old vertex  $v$ , make a new vertex–vertex  $v_v^*$  at the point given by the following linear combination, whose coefficients  $\alpha_n$ ,  $\beta_n$ , and  $\gamma_n$  depend on the valency  $n$  of  $v$ :  

$$\alpha_n \cdot (\text{the centroid of the new edge vertices of the edges meeting at } v) + \beta_n \cdot (\text{the centroid of the new face vertices of the faces sharing those edges}) + \gamma_n \cdot v.$$

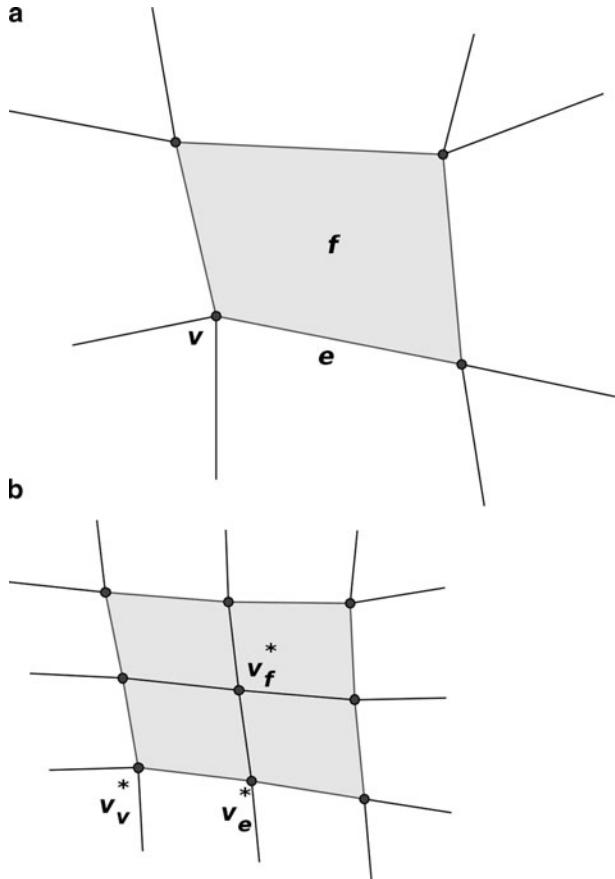
The topology  $E'$  of the new net is calculated by the following rule:

For each old face  $f$  and for each vertex  $v$  of  $f$ , make a new quadrilateral face whose edges join  $v_f^*$  and  $v_v^*$  to the edge vertices of the edges of  $f$  sharing  $v$  (see Fig. 8.10).

The weights  $W_n$  are computed as follows. Let  $n > 2$  be the valency of a vertex and  $k = \cos(\pi/n)$ . Compute the real roots of  $x^3 + (4k^2 - 3)x - 2k = 0$  and take the only one satisfying  $x > 1$ . So, the weights are calculated by:

$$\begin{aligned} W_n &= x^2 + 2kx - 3, \\ \alpha_n &= 1, \\ \gamma_n &= \frac{kx + 2k^2 - 1}{x^2(kx + 1)}, \\ \beta_n &= -\gamma_n \end{aligned} \tag{8.49}$$

In the context of polygonal surfaces, digital prosthesis design can be cast in the problem of filling  $n$ -sided holes in the surface. Therefore, algorithms that generate a subdivision surface which connects smoothly with the surface around the hole can be useful for prosthesis modeling.



**Fig. 8.10** (a) Initial net and highlighted face (reprinted from [29]). (b) New net obtained with the Sabin's algorithm (reprinted from [29])

## 8.4 Computer-Aided Design Methods and Systems

In this section, we discuss some proposals for prosthesis modeling. We start with a balloon model that is applied slice-by-slice for reconstruction of the lost part (Sect. 8.4.1). Then, a marching cubes technique is used to get the final geometry. On the other hand, we can perform all the tasks in 3D space. The T-Surfaces plus isosurface generation models offer a general framework for surface reconstruction and user interaction [42]. Therefore, we can explore this framework for digital prosthesis design, as discussed in Sect. 8.5. The wavelet theory can be applied to simplify the detection of the boundary points of the hole. Besides, subdivision surfaces have been used for filling holes methods. So, we discuss the application of these methods for prosthesis modeling in Sect. 8.5.1. We describe also experimental

results. In Sect. 8.5.2, we present a computational tool that we are developing for image processing and visualization which can be used for digital prosthesis design.

### 8.4.1 Balloon-Based Model

In this section, we propose a new methodology for digital prosthesis generation, which can be roughly divided into five stages:

1. Segmentation: Extract the bone from the other tissues.
2. Feature Extraction: Get geometric information about the frontier of the lesion (terminal points and their tangents).
3. Deformable Model: With the information from step 2, find the appropriate curve, using a deformable model of type *balloon*.
4. Lesion Reconstruction: Each patch obtained in the step above is dilated to complete the prosthesis volume.
5. Surface Reconstruction: Marching cubes is applied to generate the digital version of the prosthesis geometry.

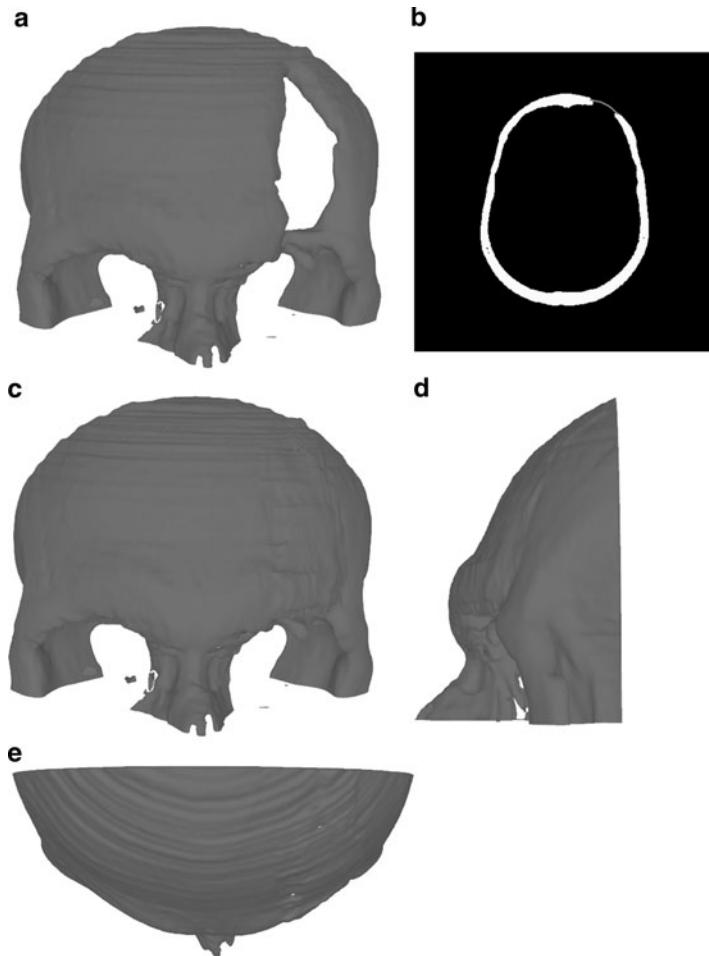
To accomplish the first step, we apply a simple thresholding technique based on image inspection. In the actual implementation, the obtained result is manually processed, frame-by-frame, to get the end points and their tangents (second stage). Next, we apply a deformable model, a balloon-like model [14], to compute the patch of the prosthesis in each image frame. We take the end points and tangents computed in the step 2 as boundary conditions for the deformable model. In the final stages, we recover the lost part of the bone, following the balloon result in each frame and, finally, reconstruct the geometry of the prosthesis. We demonstrate the advantages of our technique by comparing with the one presented in [48].

The standard format for CT images is DICOM (Digital Imaging and Communications in Medicine) [35]. CT images of the head show bone and soft tissues (brain, skin etc.). In a DICOM image, the data matrix has real values in the range  $[-1000, 1000]$ , called TC numbers. A simple binarization technique may then be used to bring out just the structure elicited since the TC number for bone is in the  $[400, 1000]$  range.

Now, once the bone is segmented, we take each frame and get the end points of the lesion and their tangents (see Fig. 8.9) through an user interaction procedure.

Now, we apply the proposed model to generate the prosthesis for repairing the defective skull pictured in Fig. 8.11a. In this case, we also could use a reflection technique, based on the assumption of skull symmetry, to reconstruct the skull. The defective region intersects 36 slices of the whole image volume.

In the experiments of this section, we set null the external force in expression (8.37). Following [14], the parameters  $\omega_0$  and  $\omega_1$  are set to  $(\Delta s)^2$  and  $(\Delta s)^4$ , where  $\Delta s$  is the discretization step applied to get the numerical solution of equation (8.37). The value of the normal force scale parameter is  $k = -1/16$ . The stopping



**Fig. 8.11** (a) Original defective skull. (b) Volume slice with the balloon result. (c)–(e) Skull surface reconstructed

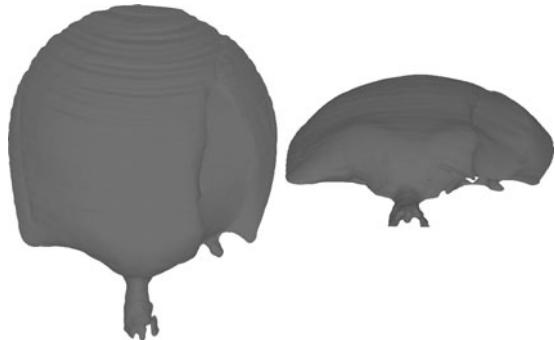
**Table 8.1** Statistics for initialization parameters ( $\theta_0$ ,  $\theta_1$ ) and number of interactions of balloon

Parameter	Min	Max
$\theta_0$	6 °	10 °
$\theta_1$	170 °	174 °
No. interactions	100	189

criterium is based on expression (8.48), with  $10^{-3}$ . Table 8.1 reports some statistics about the initialization (expression (8.44)) and evolution of the balloon model for the 36 slices that intersect the defective region.

Considering that the mean number of interactions of the numerical scheme defined by expression (8.47) is 123 and we have 21 snaxels for each instance of

**Fig. 8.12** Viewpoints of the solution obtained by the method described in [48]



the balloon model, we can say that the computational cost is not expensive. Figure 8.11a shows the defective skull and Fig. 8.11b pictures the balloon result for one of the slices. Once the balloon stops evolution, we dilate the result using a  $5 \times 5$  mask centered in the snixels. Figures 8.11c–e shows three viewpoints of the repaired skull. A visual inspection indicates that the prosthesis geometry suitably reconstructs the defective region. However, the method is sensitive to the parameters choice. So, we must be careful about this point during the setup of the balloon technique.

It is worthwhile to compare our method with the one proposed in [48], summarized in Sect. 8.2. Figure 8.12 pictures the solution obtained with that method. The method proposed in [48] does not take into account the tangent directions at the boundary of the hole. Therefore, the obtained prosthesis may not fit the curvature of the skull. In fact, the method has a bias toward planar shapes due to the fact that there is no any constraint related to local curvature. We can check this problem in the result pictured in Fig. 8.12.

We have also some considerations about the algorithm that searches the boundary link, described in Sect. 8.2. In our implementation of that algorithm, we get suitable results without testing whether the vertices are interior to the polygonal frontier (rule 1), as done in [30]. It is important to notice that this polygon is a curve in the 3D space which makes the determination of inner points a nontrivial task. Besides, we define an upper bound for the number of triangles generated in the mesh refinement algorithm to avoid too many triangles in low curvature regions.

## 8.5 Isosurface Methods and T-Surfaces

In this section, we describe an approach which integrates the T-Surfaces model and isosurface generation methods in a general framework for surface reconstruction in 3D medical images (see [42] for details). The reparameterization of T-Surfaces gives the link between the above isosurface generation methods and T-Surfaces model. To explain this, let us define an *Object Characteristic Function* as:

$$\begin{aligned}\chi(p) &= 1, & \text{if } I(p) < T, \\ \chi(p) &= 0, & \text{otherwise,}\end{aligned}\tag{8.50}$$

where  $p$  is a node of the triangulation.

If we apply tetra-cubes or continuation method to this field, we get a set of piecewise linear (PL) surfaces that involve the structures of interest. From the way the PL surfaces are generated, each connected component  $\widehat{M}$  so obtained has the following properties: (1) the intersection  $\sigma_1 \cap \sigma_2$  of two triangles  $\sigma_1, \sigma_2 \in \widehat{M}$  is empty, a common vertex or edge of both triangles; (2) an edge  $\tau \in \widehat{M}$  is common to at most two triangles of  $\widehat{M}$ ; (3)  $\widehat{M}$  is locally finite, that is, any compact subset of  $\Re^3$  meets only finitely many cells of  $\widehat{M}$ .

A polygonal surface with such property is called a Piecewise Linear Manifold (*PL Manifold*). From the reparameterization process of Sect. 8.3.1, we can see that a T-Surface is also a PL Manifold. Thus, the isosurface extraction methods can be used straightforward to initialize T-Surfaces with the object characteristic function as the initial characteristic function.

But, what kind of isosurface method should be used? Based on the discussion in Sect. 8.3.3, about tetra-cubes and continuation methods, we can conclude that if we do not have topological and scale restrictions, marching methods are more appropriate to initialize the T-Surfaces. In this case, it is not worthwhile to attempt to reconstruct the surface into neighboring simplices because all simplices should be visited to find surface patches.

However, for the T-Surfaces reparameterization (steps (1)–(4) of Sect. 8.3.1), the situation is different. Now, each connected component is evolved at a time. Thus, it is interesting a method which generates only the connected component being evolved, that is, the PL Generation algorithm of Sect. 8.3.3.

The segmentation/surface reconstruction method that we propose in [42] is based on the following steps: (1) extract region-based statistics; (2) coarser image resolution; (3) define the object characteristic function; (4) PL manifold extraction by the tetra-cubes; (5) if needed, increase the resolution. Return to step (3); (6) apply T-Surfaces model; (7) user interaction if need.

It is important to highlight that T-Surfaces model can deal naturally with the self-intersections that may happen during the evolution of the surfaces obtained by step (4). This is an important advantage of T-Surfaces.

Among the surfaces extracted in step (4), there may be open surfaces which starts and end in the image frontiers, small surfaces corresponding to artifacts or noise in the background. The former is discarded by a simple automatic inspection. To discard the later, we need a set of predefined features (volume, surface area, etc) and corresponding lower bounds. For instance, it is straightforward to set the volume lower bound as  $8(r^3)$ , where  $r$  is the dimension of the grid cells. Besides, some polygonal surfaces may contain more than one object of interest. Now, we can use upper bounds for the features. These upper bounds are application dependent (anatomical elements can be used).

The surfaces whose interior have volumes larger than the upper bound will be processed in a finer resolution. When the grid resolution of T-Surfaces is increased,

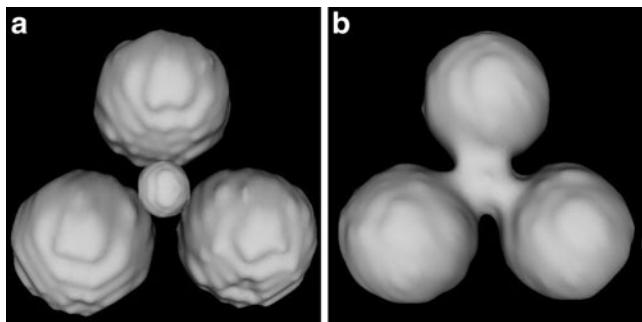
we just reparameterize the model through the finer grid and evolve the corresponding T-Surfaces.

Besides, due to inhomogeneities of the image field (supposed gray level), some objects may be split in the step (4). Sometimes, T-Surfaces model cannot merge them again. To correct these problems, the user can manually burn some grid nodes to force merges or splits. From the entropy condition, these nodes remain burnt until the end of the process. This functionality can be implemented by selecting grid nodes with a pointer (e.g., mouse), through implicitly defined surfaces, or through a *virtual scalpel*.

In order to demonstrate the potential of the whole framework, let us consider the following example. We have three spheres of intensity 50 placed in a noisy  $150 \times 150 \times 150$  image volume. The spheres were previously extracted by the proposed method with the following parameters: grid  $5 \times 5 \times 5$ , freezing point = 10,  $\gamma = 0.01$ ,  $k = 1.222$ ;  $c = 0.750$ . Every sphere has radius 30 (pixels).

In this case, the merge is forced through an implicit-defined surface placed between the spheres (Fig. 8.13a). Grid nodes inside the surface are easily detected by its equation and then burnt. During the evolution, the four surfaces will merge and the final result is a connected surface (Fig. 8.13b). Other possibility would be to burn manually a set of grid nodes linking the spheres. The idea in this case is that the new set of burnt grid nodes generates a connected combinatorial manifold. A similar idea can be implemented for manual split.

This framework is useful for prosthesis design due to the following aspects. When working in a coarser image resolution, we are discarding details in the reconstructed surfaces which may simplify the drawing of the boundary links in Fig. 8.3. On the other hand, once resolved the segmentation in a coarser resolution, we can update the result for a finer image resolution and to get a segmentation with more details. So, we do not lose precision in the whole process. Besides, once completed the geometry reconstruction of the defective skull, we can turn off the T-Surfaces reparameterization (T-Surfaces becomes a mass-spring system). Then, we apply the algorithm summarized in Sect. 8.2 to obtain a triangulation of the target surface but using a smoothing process based on the mass-spring system instead of the umbrella-operator. Finally, if user interaction is required to improve



**Fig. 8.13** (a) Original objects. (b) Merge through the user interaction method

the prosthesis geometry, we can turn on simplicial domain decomposition framework and allow the user to burn (or unburn) some grid nodes to mimic material addition or removal.

### 8.5.1 Multiscale and Subdivision Schemes

The application of the multiscale schemes for digital prosthesis design follows the idea that performing tasks in coarser scales may be easier than in finer ones. For instance, let us return to the Fig. 8.3 and to the problem of finding the internal and external boundary links of the hole. It is not a trivial task specially if the portion of the surface nearby the hole has too many details due to the surface reconstruction process. So, if we apply a multiscale scheme, we can reduce details and simplify the definition of these curves.

To be more formal, let us consider parametrically defined curves:

$$\mathbf{x}(t) = (f^1(t), f^2(t), f^3(t)); \quad t \in \mathfrak{R}.$$

Then the wavelet decomposition of the curve is obtained by the decomposition of each coordinate function:

$$f^k(t) = \sum_{i,j} \langle f^k, \psi_{i,j} \rangle \psi_{i,j}(x).$$

The wavelet coefficients of the curve are given by:

$$c_{i,j}^k = \langle f^k, \psi_{i,j} \rangle.$$

We can similarly define the scaling decomposition and coefficients by:

$$f^k(t) = \sum_{i,j} \langle f^k, \varphi_{i,j} \rangle \varphi_{i,j}(x),$$

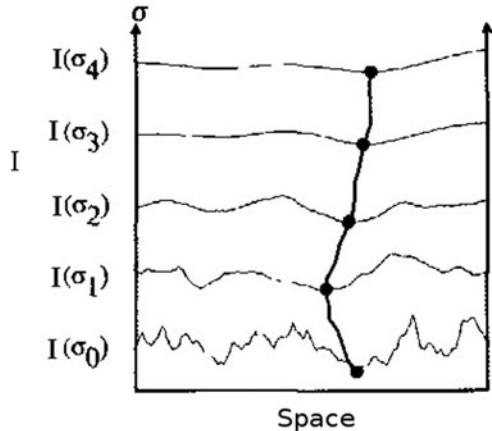
$$d_{i,j}^k = \langle f^k, \varphi_{i,j} \rangle.$$

From expression (8.29), it is clear that:

$$P_{m-1}(f^k) = Q_m(f^k) + P_m(f^k). \quad (8.51)$$

This expression is the starting point for a multiscale approximation of the curve. The scaling coefficients obtained by the projection  $P_m(f^k)$  define the multiscale approximation of the curve at level  $m$ , and the error between the two approximation at scale level  $m$  and  $m - 1$  is given by the wavelet coefficients computed through

**Fig. 8.14** Scale space representation (reprinted from [47])



the projection  $Q_m(f^k)$ . A similar process can be defined for parametrically defined surfaces through tensor product of scaling functions [34]. The fact that  $P_m(f^k)$  is an approximation of  $f^k$  in a coarser scale; that means, with less details, can be explored to simplify the search for structures of interest.

Just to clarify the ideas, let us consider the Fig. 8.14 which pictures a (generic) multiscale representation of a function  $I$  at scales  $\sigma_0, \sigma_1, \dots, \sigma_4$ . We observe that the singular points (local maxima and minima) in coarser scales can be easily identified than the corresponding points in the finer scales. So, following expression (8.51), the key idea when using wavelet representation is to get the target points in the coarser scale and then to correct its position, using the error  $Q_m(f^k)$ .

Once the boundaries of the hole are drawn (Fig. 8.3), the subdivision scheme described in Sect. 8.3.6 can be also applied to generate the two surfaces that, when combined with the hole's surface, complete the prosthesis model. Such process will generate two polygonal surfaces with quadrilateral patches instead of the triangular ones generated by the algorithm described in [48] (Sect. 8.2). Another aspect is that the user can place guiding points to control the geometry of the input net. In the case of the algorithm described in [48], such control is not possible unless we introduce extra machinery.

### 8.5.2 The PyImageVis Computer Systems

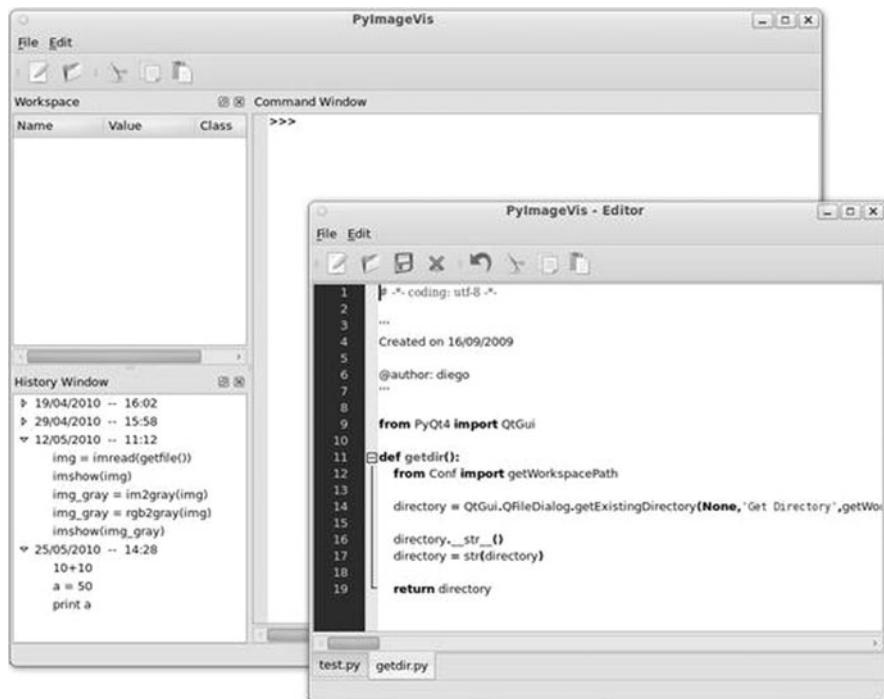
This section describes the PyImageVis system, an open-source software implemented in Python [6], for processing and visualization of medical images. The main goal of PyImageVis is to provide a computational system for the research of new algorithms for medical image processing and visualization.

In Brazil, the research on visualization and medical image processing is concentrated in universities and research institutes. The INCT-MACC, a virtual institute

for applications of scientific computing in medicine, is an example of this fact [2]. The INCT-MACC team need a computational platform which facilitates the validation of new algorithms in image processing and visualization. Softwares such as *Matlow* and *Octave* can be used for this purpose. However, they are limited for visualization of three-dimensional images. In order to meet this requirement, we are developing the *PyImageVis* platform. This software has been implemented in Python language and incorporates functions and methods of existing open source Python libraries, like *Bumpy* and *Scion* [4, 7]. It provides friendly graphical interfaces, which follows the *Matlow* layout (Fig. 8.15) and is user extendable.

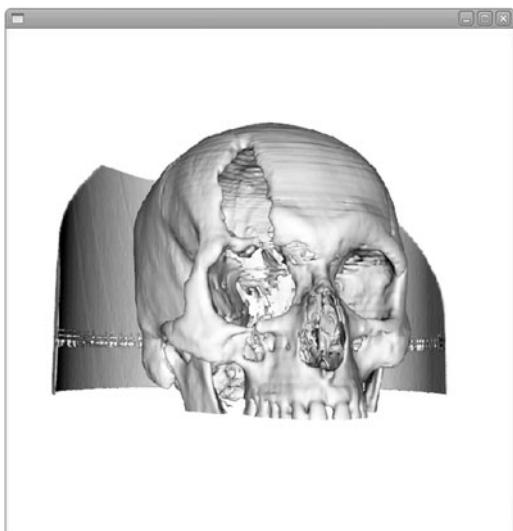
We are incorporating computational resources for digital prosthesis design in the *PyImageVis* system. Figure 8.16 pictures the main software interface and an application example for isosurface visualization in the case of a defective skull. The image volume is composed by 108 slices, in the *DICOM* format,  $512 \times 512$  pixel resolution [5].

Figure 8.17 describes the capabilities of the *PyImageVis* to reconstruct the skull slice-by-slice in the image space. Figure 8.17b shows the result obtained by reflecting the hole image respect to axis of symmetry of the image pictured in Fig. 8.17a. Besides, we also demonstrate the *PyImageVis* potential with the execution of another slice-by-slice processing task conducted by user interaction. In this

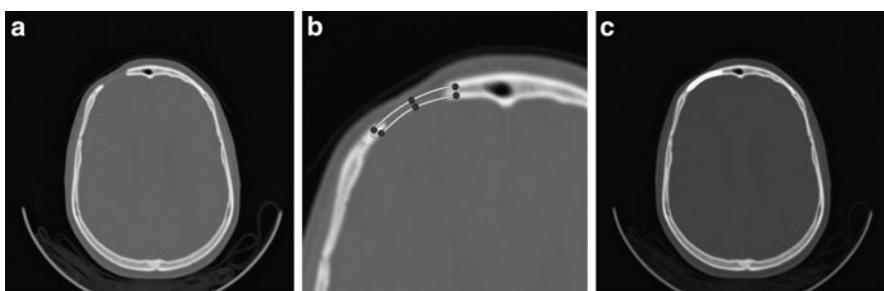
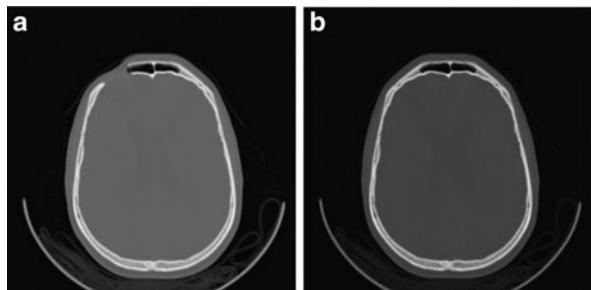


**Fig. 8.15** Graphical user interface components of *PyImageVis*: *Command Window*, *History Window*, *Workspace*, and *Python Editor*

**Fig. 8.16** Surface reconstruction through marching cubes in the *PyImageVis* software



**Fig. 8.17** (a) Original slice of defective skull. (b) Result after reflection



**Fig. 8.18** (a) Original slice with lost part. (b) User-defined ribbon. (c) Repaired figure

case, the user first places paths linking terminal points of the lesion in Fig. 8.18a, obtaining the boundary of a ribbon like the one in the Fig. 8.18b. Then, the system applies a region filling to generate the ribbon that completes the lost part of the bone (Fig. 8.18c).

## 8.6 Discussion and Perspectives

A question in this area is how to incorporate prior knowledge about the skull shape in each slice. This question needs some kind of learning process to be addressed. The theory presented in Sect. 8.3.2, based on PCA, is a possibility in this direction.

First, 40 slices of a normal cranium are used to compose a data set. Figure 8.19 shows the first, middle, and last (segmented) slices. Then, we compute the skeleton for each slice and fit a cubic spline for each obtained skeleton. Then, we could calculate a set of 100 uniformly spaced landmarks for each curve composing a database of 40 curves. Then, we align each curve with the middle one (curve generated from Fig. 8.19b) to discard variations due to rigid transformations (rotation, scale, and translation).

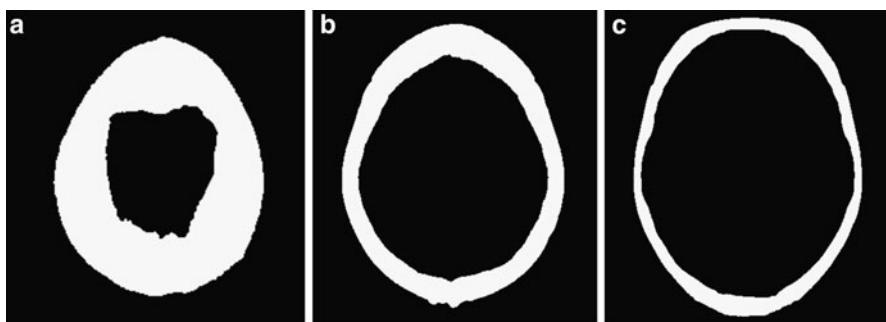
The aligned curves are used to compute the mean shape and covariance matrix in expressions (8.12) and (8.13), respectively. The spectrum of the covariance matrix is pictured in Fig. 8.20. This figure shows that as the dimension of the PCA most expressive subspace increases, there is an exponential decrease in the amount of total variance explained by the first principal components with the largest eigenvalues.

This is a well-known behavior of the dimensionality reduction provided by the standard PCA [23]. More specifically, the first 20 principal components have eigenvalues  $\lambda$  in the range  $[2 \leq \lambda \leq 7,800]$  and for the first five principal components this range is  $[184 \leq \lambda \leq 7,800]$ . Figure 8.21 shows the mean curve  $\bar{x}$  and the curve obtained by the expression:

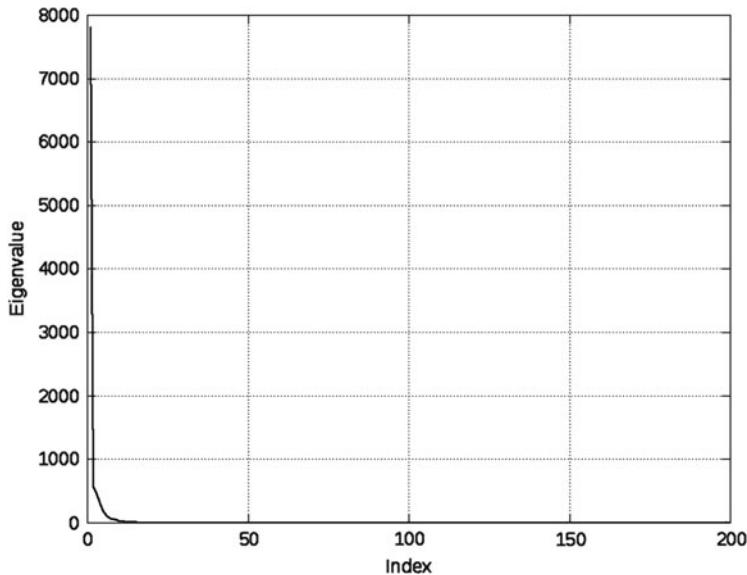
$$\mathbf{x} = \bar{\mathbf{x}} + \alpha \cdot \phi_1, \quad (8.52)$$

with  $\alpha = 50.0$ .

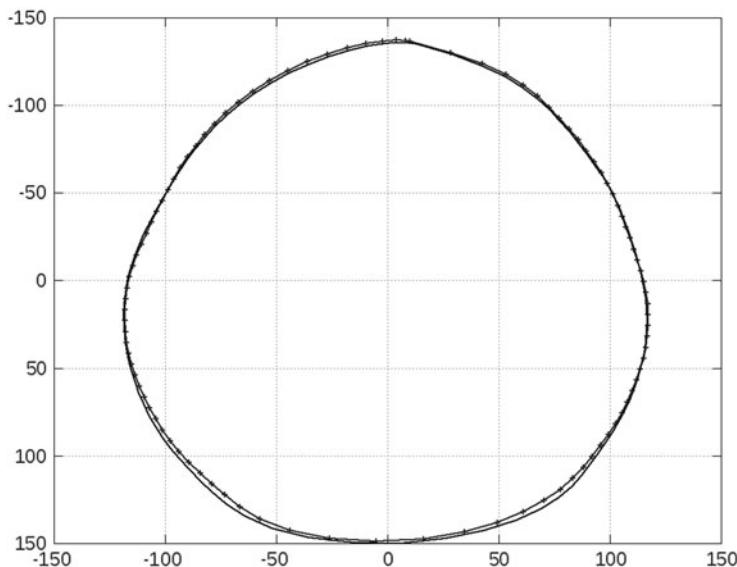
We observe that the total variance explained by the PCA most expressive principal component captures the basic shape of the database. This can be further used to implement the previous knowledge to the digital prosthesis design process.



**Fig. 8.19** Bone segmentation results for: (a) first image of the data set; (b) the middle one; (c) the last image



**Fig. 8.20** Exponential decrease of the eigenvalues of the covariance matrix



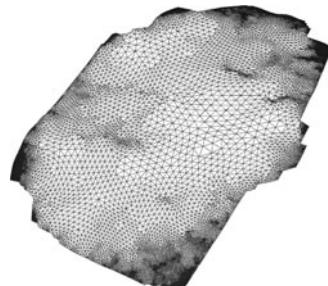
**Fig. 8.21** Mean curve (*marked one*) and the curve generated by expression (8.52)

Fundamentally, we may enforce a limitation that the skeleton of the rebuilt format has to belong to the CA subspace spanned, for example, the first five main components.

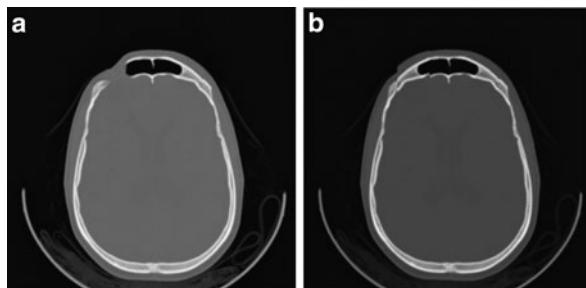
Section 8.5.1 proposes the application of subdivision methods as an alternative to generate the prosthesis representation. Such process generates surfaces with quadrilateral patches instead of the triangular ones generated by the algorithm described in [48]. From Fig. 8.22, we observe that the triangle density of the latter is very high and not uniform over the surface. This feature may be a problem if we need user interaction to fix the curvature in some places. We must consider this intervention because it is very hard to account for all the clinical aspects when customizing a prosthesis. Specifically, we intend to apply the mass-spring system to implement this task. Such variation in the density of triangles implies that the system may be more rigid in some places which is not interesting for user interaction.

The reflection is a very useful method but has some issues also. Figure 8.23a shows the original image and Fig. 8.23b shows the result of the reflection. We observe some defects that must be fixed. This can be accomplished by user interaction through mouse-guided intervention. Other possibility is to give the user a tool to construct a ribbon to fix the lesion, like in the Fig. 8.18. However, this process may tedious and imprecise even if we apply the result obtained for a slice, say slice  $k$ , and use it to initialize the process in the slice  $n + 1$ .

On the other hand, the method presented in section balloon depends on the boundary conditions to set up the deformable model. The straightforward way to get this information is through the skeleton of the bone region. Figure 8.24 shows one example of skeleton that helps us to understand that this process has also some drawbacks. In this case, if we follow the tangent at the end points of the skeleton, we obtain a balloon solution far from the target.

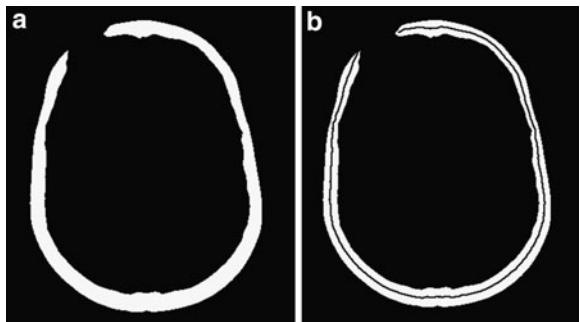


**Fig. 8.22** Triangulated surface generated by the algorithm proposed in [48]



**Fig. 8.23** (a) Original image. (b) Result of reflection

**Fig. 8.24** (a) Segmented bone region. (b) Skeleton with end points unsuitable to get boundary conditions



The application of implicit formulations for deformable models is an interesting perspective in this area. As pointed out in [42], we can implement the framework described in Sect. 8.5 using the level set model [43]. This alternative is interesting not only due to the capabilities of level set but also because its implementation is simpler than for T-Surfaces. In fact, the embedding function for the level set plays a similar role of the characteristic function for T-Surfaces (see [40] for details). However, the update of the latter is based on exhaustive tests [33] while the former is just a consequence of the governing equation of the level set method. As a practical consequence, we observed that it is easier to implement the level set than the T-Surfaces which is a motivation for further applications of level set in the prosthesis modeling problem.

## 8.7 Conclusions

In this chapter, we review approaches to reconstruct the defective region of a skull. The main goal is to automatically construct the prosthesis model for the defective region.

Re-building a new cranium after injury may be achieved through surgical methods by implanting a customized prosthesis. Lately, the creation of digital prosthesis and surgical planning has been done by using image processing, surface rebuilding, and geometric techniques.

The field of prosthesis modeling may include methods for segmentation and surface reconstruction, geometric modeling, multiscale methods, and user interaction approaches. We show a promising result based on the balloon model and compare our technique with a state-of-the-art one, showing that our method can generate a more suitable prosthesis geometry. Besides, we discuss perspectives in this area to perform all the tasks in 3D space as well as to incorporate prior knowledge in the process.

**Acknowledgements** The authors would like to thank CNPq (grant 133865/2009-6) and INCT-MACC for financial support.

## References

1. 3-Matic for biomedical research. <http://www.materialise.com/BiomedicalRnD/3-matic>
2. INCT-MACC – Brazilian Virtual Institute for Applications of Scientific Computing in Medicine. URL <http://www.lncc.br/>
3. Mimics. <http://www.materialise.com/mimics>
4. NumPy. URL <http://numpy.scipy.org/>. Acessado em: 15 dez. 2009
5. Portal of brazilian public software. URL <http://www.softwarepublico.gov.br>
6. Python programming language. URL <http://www.python.org>. Acessado em: 18 dez. 2009
7. Scipy. URL <http://www.scipy.org/>. Acessado em: 15 dez. 2009
8. Allgower EL, Georg K (1990) Numerical continuation methods: an introduction. Springer, Berlin/Heidelberg
9. Barequet G, Sharir M (1995) Filling gaps in the boundary of a polyhedron. Comput Aided Geom Des 12(2):207–229
10. Bhargava D, Bartlett P, Russell J, Liddington M, Tyagi A, Chumas P (2010) Construction of titanium cranioplasty plate using craniectomy bone flap as template. Acta Neurochir 152(1): 173–176
11. Carneiro BP, Silva CT, Kaufman AE (1996) Tetra-cubes: an algorithm to generate 3D isosurfaces based upon tetrahedra. In: International symposium on computer graphics, image processing and vision (SIBGRAPI'96)
12. Catmull E, Clark J (1978) Recursively generated b-spline surfaces on arbitrary topological meshes. Comput Aided Des 10:350–355
13. Chiang YJ, Silva C, Schroeder W (1998) Interactive out-of-core isosurface extraction. In: Proceedings of IEEE visualization 1998, pp 167–174
14. Cohen LD (1991) On active contour models and balloons. CVGIP:Image Underst 53(2): 211–218
15. Di Lorenzo N, Pirillo V (2008) From computerized tomography data processing to rapid manufacturing of custom-made prostheses for cranioplasty case report comment. J Neurosurg Sci 52(4):116
16. van Ginneken B, Frangi A, Staal J, ter Haar Romeny B, Viergever M (2002) Active shape model segmentation with optimal features. IEEE Trans Med Imaging 21(8):924–933
17. Giraldi G, Strauss E, Apolinario A, Oliveira AF (2001) An initialization method for deformable models. In: 5th World multiconference on systemics, cybernetics and informatics (SCI 2001)
18. Giraldi GA, Oliveira AAF, Carvalho L (2005) Mining cellular automata databases through pca models. <http://arxiv.org/abs/cs/0511052>
19. Goiato MC, Anchieti RB, Pita MS, dos Santos DM (2009) Reconstruction of skull defects: currently available materials. J Craniofac Surg 20(5):1512–1518
20. Graps A (2004) Historical perspective. <http://www.amara.com/IEEEwave/IW\history.html>
21. House DH, Breen DE (eds) (2000) Cloth modeling and animation. A K Peters, Natick, MA
22. Jain AK (1989) Fundamentals of digital image processing. Prentice-Hall, Upper Saddle River, NJ
23. Johnson R, Wichern D (1998) Applied multivariate statistical analysis. Prentice-Hall, Englewood Cliffs, NJ
24. Kai CC, Meng CS, Ching LS, Hoe EK, Fah LK (2009) Cranioplasty using polymethyl methacrylate prostheses. J Clin Neurosci 16(1):56–63
25. Laeven P, Poukens J, Beerens MMA (2003) 3-Matic assists in true rapid design and manufacturing of a custom skull implant. Technical report, Materialize, [www.materialise.com/download/en/2706472/file](http://www.materialise.com/download/en/2706472/file)
26. Lauric R, Frisken S (2007) *Soft segmentation of CT Brain Data*. Technical Report TR-2007-3, Department of Computer Science – Tufts University, Medford, MA
27. Lee SC, Wu CT, Lee ST, Chen PJ (1998) Rapid prototyping assisted surgery planning. Int J Adv Manuf Technol 14(9):624–630

28. Lee MY, Chang CC, Lin CC, Lo LJ, Chen YR (2002) Custom implant design for patients with cranial defects. *IEEE Eng Med Biol Mag* 21(2):38–44
29. Levin A (1999) Filling n-sided holes using combined subdivision schemes. In: Laurent PJ, Sabinonniere P, Schumaker LL (eds) *Curve and surface design*. Vanderbilt University Press, Nashville, TN, pp 221–228
30. Lin L, Zhang J, Fang M (2008) Modelling the bio-scaffold for repairing symmetrical and unsymmetrical defective skull. In: The 2nd International conference on bioinformatics and biomedical engineering, 2008 (ICBBE 2008). pp 905–908, doi:10.1109/ICBBE.2008.222
31. Liulan L, Hanqing L, Qingxi H, Limin L, Minglun F (2006) Research of the method of reconstructing the repair bionic scaffold based on tissue engineering. *IET Conference Publications* 2006(CP524):1275–1279
32. Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. *Comput Graph* 21(4):163–169
33. McInerney T, Terzopoulos D (1999) Topology adaptive deformable surfaces for medical image volume segmentation. *IEEE Trans Med Imaging* 18(10):840–850
34. Meyer Y (1992) *Wavelets and operators*. Cambridge University Press, Cambridge
35. Pfeifle R, Seidel HP (1996) Triangular b-splines for blending & filling of polygonal holes. In: Proceedings of the conference on Graphics interface'96, pp 186–193
36. Pianykh OS (2008) *Digital imaging and communications in medicine (DICOM): a practical introduction and survival guide*. Springer, Berlin/Heidelberg
37. Quarteroni A, Sacco R, Saleri F (2000) *Numerical mathematics*. Springer, Berlin
38. Sabin M (1991) Cubic recursive division with bounded curvature. In: Laurent PJ, le Mehaute A, Schumaker LL (eds) *Curves and surfaces*, Academic, Boston, pp 411–414
39. Sanan AM, Haines SJM (1997) Repairing holes in the head: a history of cranioplasty. *Neurosurgery* 40(3):588–603
40. Sethian JA (1996) *Level set methods: evolving interfaces in geometry, fluid mechanics, computer vision and materials sciences*. Cambridge University Press, Cambridge
41. Strauss E, Jimenez W, Giraldo G, Silva R, Oliveira A (2002) A surface extraction approach based on multi-resolution methods and T-surfaces framework. Technical report, National Laboratory for Scientific Computing, <ftp://ftp.lncc.br/pub/report/rep02/rep1002.ps.Z>
42. Strauss E, Jimenez W, Giraldo GA, Silva R, Oliveira AF (2002) A semi-automatic surface reconstruction framework based on T-surfaces and isosurface extraction methods. In: International symposium on computer graphics, image processing and vision (SIBGRAPI'2002)
43. Suri JS, Liu K, Singh S, Laxminarayanan S, Zeng X, Reden L (2002) Shape recovery algorithms using level sets in 2-d/3-d medical imagery: a state-of-the-art review. *IEEE Trans Inf Technol Biomed* 6(1):8–28
44. Sutton P, Hansen C (2000) Accelerated isosurface extraction in time-varying fields. *IEEE Trans Vis Comp Graph* 6(2):98–107
45. Volino P, Magnenat-Thalmann N (2000) *Virtual clothing: theory and practice*. Springer, Berlin
46. Wang Y (2005) The study of three-dimensional reconstructed bioceramics in the successional defect of mandible. The Second Military Medical University, Shanghai
47. Witten G (1993) Scale space tracking and deformable sheet models for computational vision. *IEEE Trans Pattern Anal Mach Intell* 15(7):697–706
48. You F, Hu Q, Yao Y, Lu Q (2009) A new modeling method on skull defect repair. In: International conference on Measuring technology and mechatronics automation, 2009 (ICMTMA'09) 1(11–12):568–572
49. Zhao Y, Zhou J (2006) The analysis of clinical effect of individualized cranioplasty with different materials for 75 skull defect patients. *Med J Chin People's Liberation Army* 31:354–355

## Biography



**Gilson A. Giraldi** received his PhD degree in Computer Graphics from Federal University of Rio de Janeiro, Brazil, in 2000. Since then he has been with the National Laboratory for Scientific Computing, Brazil, where he is responsible for academic research projects in the areas of image segmentation and data visualization.



**Luiz C. M. de Aquino** received his Diploma in Mathematics from Feira de Santana State University in 2004 and he is taking his MSc degree from National Laboratory for Scientific Computing.



**Paulo S. S. Rodrigues** received his PhD degree in Computer Science from Federal University of Minas Gerais 2003. From 2003 to 2006 he worked in National Laboratory for Scientific Computing. Since 2007 he is Professor of FEI University, São Bernardo do Campo, São Paulo. His main interest areas are pattern recognition and medical image analysis.



**Antonio L. Apolinrio Jr** received the Doctor in Science degree from COPPE/UFRJ. He is currently a Professor in the Technology Department of State University of Feira de Santana, Bahia. His research areas are: Computer Graphics, Visualization, Computer Animation, Augmented Reality, GPU programming and Computer Games.



**Jaime S. Cardoso** received his PhD in Computer Vision from Universidade do Porto in 2006. Cardoso is currently a Project Leader in the Telecommunications and Multimedia Unit of INESC Porto, where he leads a research team on machine learning, computer vision and digital signal processing. He is also Assistant Professor at the Faculty of Engineering, University of Porto.



**Jasjit S. Suri**, PhD, MBA, Fellow AIMBE, is an innovator, scientist and an internationally known world leader in Biomedical Imaging Sciences and Devices applied to Diagnostics and Therapeutics. He has published over 300 publications, 22 books, 60 innovations (patents), and 4 FDA clearances covering several fields of medicine. He has been Senior Director and Chief Technology Officer of Fischer Imaging, Eigen Inc. and then for Biomedical Technologies. His main interests are in the area of biomedical engineering and its management.

# Chapter 9

## Medical Image Registration

Usaf E. Aladl and Terry Peters

**Abstract** In this chapter, we cover the necessary background information required to understand medical image registration, the basic tools required to implement registration algorithms, and demonstrate a complete application for various types of registration between different modalities using freely available and maintained software.

### 9.1 Introduction

Over the past 40 years, medical imaging technology has evolved from two-dimensional X-ray projections, to sophisticated three-dimensional digital medical imaging modalities such as CT, MRI, fMRI, and ultrasound. In addition, histology images can now be provided digitally, and are an additional rich source of medical imaging information. These tools have now become part of standard daily diagnostic procedures for many areas of medicine, and there are extensive computer-aided diagnostic techniques that assist the physician in their interpretation. While these modalities record valuable information about the patient, each provides information about a different physical quantity (electron density, magnetic behavior of hydrogen nuclei, reflection of ultrasound from tissue interfaces, etc.). These datasets are often complementary to one another. Because these different datasets each tell their own story, the problem is to consolidate these individual pieces of information into a single narrative to better understand the disease process. This is where medical image registration plays a crucial role.

Image registration is the process of aligning different images or 3D volumes together in a single space. One of the most important applications of image registration is for matching an image acquired at one time point during a treatment to an image from a different time-point, or to an image acquired for planning the

---

U.E. Aladl (✉)

Imaging Research Laboratories, Robarts Research Institute, 100 Perth Drive, P.O. Box 5015, London, ON, Canada N6A 5K8

e-mail: ualadl@imaging.robarts.ca

procedure. Such procedures are known as intra-patient registration, which is often used to monitor progress during treatment. Matching images from different patients can be used for generating a model of organ motion across a population of patients, or to match an image to an atlas (an average of images acquired from a number of patients), thereby allowing information to be transferred from the atlas to the newly acquired image.

Image registration is the process of transforming different images into a single coordinate space so that the data obtained from the different image modalities can be compared on a pixel-by-pixel or a voxel-by-voxel basis. In this case, we often refer to one image as the source and the second as the target. The image registration process involves spatially transforming the source image so that it is accurately aligned with the target image, and can be performed as a single operation between images from a single-modality or from multiple modalities. Single-modality methods tend to register images of the same modality acquired by the same scanner, while multi-modality registration methods register images acquired by different scanners. For example, the registration of brain CT/MRI images, whole body PET/CT images for tumor localization, registration of ultrasound and CT images for prostate localization in radiotherapy, and registration of contrast-enhanced CT images against non-contrast-enhanced CT images for segmentation of specific parts of the anatomy. This process is carried out by mapping the source image to the target, thereby establishing point-by-point correspondence between the source and target [1]. Several techniques have been proposed previously for automatic medical image registration, and comprehensive reviews of these methods have been presented by Maintz et al. [1], Hill et al. [2], and Makela et al. [3]. In particular, voxel-based image registration methods, such as mutual information [4] and normalized mutual information [5], have been considered robust algorithms. These voxel-based algorithms have been utilized mostly for the registration of brain images from various modalities [6, 7] and have demonstrated superior performance compared to surface based or landmark-based methods [8]. Some routine studies, (gated cardiac MRI or SPECT for example) provide another dimension of time added to the three spatial dimensions yielding the need for 4D registration algorithms [9].

In this chapter, we cover the necessary background information required to understand medical image registration, the basic tools required to implement registration algorithms, and demonstrate a complete application for various types of registration between different modalities using freely available and maintained software.

## 9.2 Medical Images

Medical imaging modalities employ noninvasive means to retrieve images from inside the human body for diagnosis, to study of normal and abnormal anatomy, and plan therapeutic procedures. Images are produced through the solution of mathematical inverse problems, where the properties of living tissue are inferred from the observed signal. In ultrasonography, the probe consists of ultrasonic pressure waves, and echoes from inside the tissue show the internal structure. In projection

radiography, the probe is an X-ray beam, which is absorbed at different rates in different tissue types such as bone, muscle, and fat. In the following subsection, briefly we describe some types of medical image modalities:

### ***9.2.1 Single Photon Emission Computed Tomography (SPECT)***

SPECT imaging is performed by using a gamma camera to acquire multiple projection 2D images from multiple angles, from which a 3D volume data set is reconstructed using a computer algorithm from multiple projections. SPECT uses radioactive tracer material to detect gamma rays. The tracer used in SPECT emits gamma radiation that is typically measured with a detector scintillation detector. When triggered by electrocardiogram, a SPECT system can collect gated cardiac images to obtain differential information about the heart in various parts of the cardiac cycle. Gated myocardial SPECT can be used to obtain quantitative information about myocardial perfusion, thickness, and contractility of the myocardium throughout the cardiac cycle from which important parameters such as the left ventricular ejection fraction, stroke volume, and cardiac output may be calculated.

### ***9.2.2 Positron Emission Tomography (PET)***

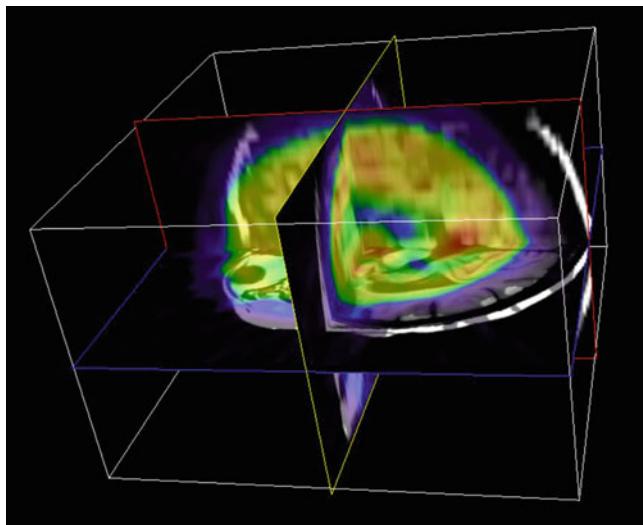
PET scanning is similar to SPECT, except that the isotopes emit positrons (rather than gamma rays). A pair of gamma-rays is subsequently generated when the positron is annihilated by an electron. These gamma rays are detected in a manner similar to that for the SPECT scanner, with both PET and SPECT delivering functional information about organs. To be properly interpreted, this information must be fused with anatomical data from CT or MRI.

### ***9.2.3 Computed Tomography (CT)***

Also known as CAT (computerized axial tomography), CT reveals both bone and soft tissues. By collecting a large series of two-dimensional X-ray images taken around a single axis of rotation and then using computer reconstruction software, the data from multiple cross-sections can be assembled as a 3D volume. CT can be a helpful aid for diagnosis and surgery or other treatment, including radiation therapy.

### ***9.2.4 Magnetic Resonance Imaging (MRI)***

MRI is based on the principles of nuclear magnetic resonance (NMR), used to obtain microscopic chemical and physical information about molecules. It uses a powerful magnetic field to align the nuclear magnetization of (usually) hydrogen



**Fig. 9.1** Three orthogonal slices of the 3D volume fused from two modalities

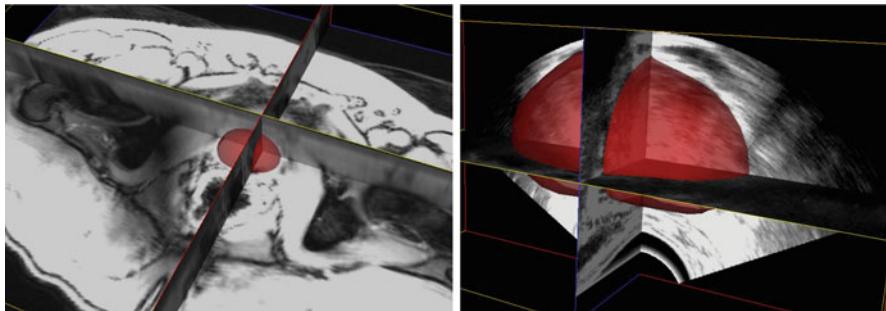
nuclei of water in the body. Radio frequency (RF) fields are used to systematically alter the alignment of this magnetization, causing the hydrogen nuclei to produce a radiofrequency signal that is detected by the scanner. This signal can be manipulated by additional magnetic fields to build up enough information to construct an image of the body [10].

### 9.2.5 3D Volume Visualization

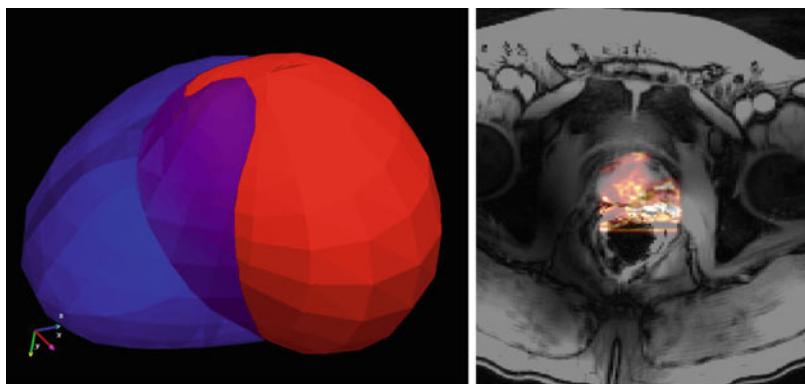
One of the most commonly used tools for visual assessment of 3D to 3D volume registration is a platform that creates a 3D scene rendered as three orthogonal slices from source and target volumes. By assigning different lookup tables for both modalities, along with the capability of adjusting the opacity in real time, one can evaluate the registration accuracy (Fig. 9.1). This platform can also be used to initialize other registration algorithms by adding a bounding box to control the orientation of the source volume interactively.

## 9.3 Surface-Guided Registration

In many examples of multimodality registration, an automated registration algorithm fails to align the image and the search is trapped in some local minima. To overcome this problem, the user must initialize the transformation parameter manually by aligning the segmented surfaces to initialize the registration (Fig. 9.2). In the example of a registration between ultrasound and MRI images of the prostate,



**Fig. 9.2** Prostate surface segmented from MRI (*left*) and ultrasound (*right*)



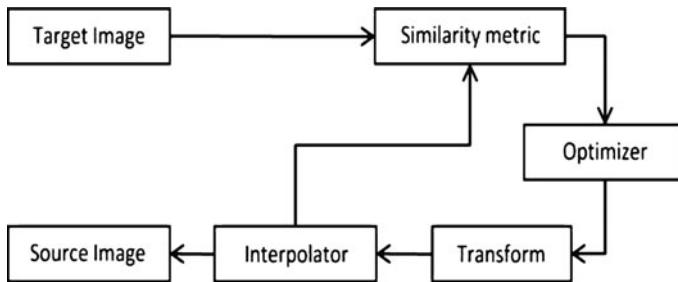
**Fig. 9.3** Prostate surfaces from MRI and US used for the image registration

the different modalities provide different representations of the information about the same organ. By segmenting both surfaces from the two modalities, one can easily align the surfaces using manual or automated techniques within other external applications (Fig. 9.3).

A comprehensive review of surface to surface registration methods is presented by Audette et al. [11].

## 9.4 Registration Framework Pipeline

In a typical image registration framework, we define the source image as  $f(X)$  and the target image as  $g(X)$ , where  $X$  is a position in  $n$ -dimensional space. By considering the registration as an optimization problem, our task is to find the optimal transformation, defined as  $T(X)$ , which aligns the source image to the target image (Fig. 9.4). A similarity metric is employed to measure how well the transformed source image  $Tf(X)$  aligns to the target image  $g(X)$ . This measure represents the value of the cost function value to be optimized by the optimizer over the search in the parameter



**Fig. 9.4** Outline of the main components of an image registration framework

space defined by the parameters of the transformation  $T(X)$ . Since most images have different sizes in discrete pixel space, all computation during the searching is performed in continuous space and requires some kind of interpolation.

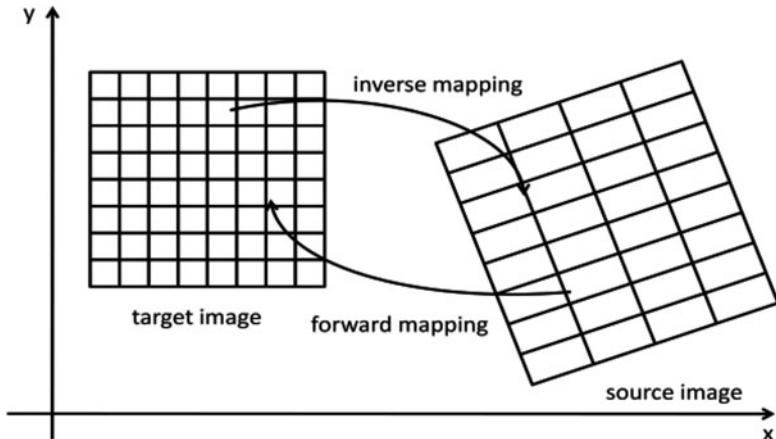
## 9.5 Coordinate Systems

Coordinate systems are an important part of any medical image registration procedure. Medical scanners acquire their data as a regular, “rectangular” array of points. The topology of these points is implicitly defined in this representation. The geometric location of each point is also implicit. In the DICOM standard special tags that define the image in world coordinates are included in the file and can be parsed from the header. Examples of these tags are:

- Image Orientation (Patient) (0020, 0037). This tag defines the direction cosines of the first row and the first column with respect to the patient in the world space.
- Image Position (Patient) (0020, 0032). The  $x$ ,  $y$ , and  $z$  coordinate of the upper left corner of the image, in millimeters. The image position specifies the  $x$ ,  $y$ , and  $z$  coordinates of the upper left hand corner of the image; it is the center of the first voxel transmitted.

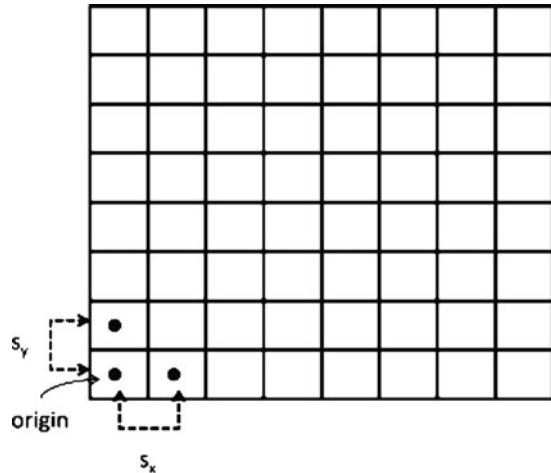
## 9.6 Registration in Physical Space

Even though medical images are represented by points on a 2D or 3D rectilinear grid, it is important to understand that in any registration framework the registration is performed in the context of physical or continuous space and not in the discrete space of the image grid (Fig. 9.5). One important consequence of this fact is that having the correct image origin and image pixel size is fundamental to the success of the registration process. Users must therefore make sure that they provide correct values for the origin and spacing of both the fixed (target) and moving (source) images. A typical case that helps understand this issue is to consider the registration of two images where one has a pixel size/spacing different from the other.



**Fig. 9.5** The process of mapping the source image to the target image is done in physical space

**Fig. 9.6** Mapping between discrete space and continuous space



Precise conversions between index space and physical point space are critical for the performance of medical image algorithms. For example, in 3D images, there are three pieces of information required to fully define the image in world space. The origin ( $o_x, o_y, o_z$ ) specifies the world coordinate of the first point in the index space. The Image spacing ( $s_x, s_y, s_z$ ) specifies the distance between points along each axis. Using the origin and spacing, one can map between  $(i, j, k)$  and  $(x, y, z)$  using the following formulas (Fig. 9.6):

$$x = o_x + i * S_x, \quad x = o_y + j * S_y, \quad x = o_z + k * S_z$$

$$i = \frac{x - o_x}{S_x}, \quad j = \frac{y - o_y}{S_y}, \quad k = \frac{z - o_z}{S_z}$$

## 9.7 Transformation

The operation that describes the mapping of points from one space to another is defined as a transform. Transforms employed in medical imaging range from simple translation, rotation, and scaling, to general affine. Before specific transforms are discussed, it is important to know how the transform is applied to the source image. The direct method, known as forward mapping, maps coordinates from the source image to the target image. This process has many problems pertaining to interpolation and pixel assignment, especially if the transform modifies the scale of the source image. Since we are mapping the source image onto the target image's field of view, we are only concerned with values that are defined within the domain of the source image. In reality, therefore, an inverse mapping is normally performed, which creates a mapping from the target to source image. Since some of these parameters have different units, for example, translation is usually defined in millimeters and rotation angle is defined in degrees or radians, one has to consider some kind of normalization for fair treatment within the search range and to inform the optimizer about any flexibility within that range and step size, that leads to the appropriate convergences.

### 9.7.1 3D Rigid Transformation

A typical 3D rigid transform is composed of three parts: translation, rotation, and linear scaling. Since the 3D image samples in world coordinates are represented on regular rectangular grids in discrete space, the process of aligning one image to another becomes a matter of mapping one point  $p = (x, y, z)$  to a new point  $p' = (x', y', z')$  generated through the transform.

#### 9.7.1.1 Translation

To translate or shift a point  $p(x, y, z)$  by a length in millimeter  $T = (t_x, t_y, t_z)$ , the new translated point  $p' = (x', y', z')$  is the result of adding each component in  $x, y, z$  axes to the component of the source point  $p$  as  $x' = x + t_x$ ,  $y' = y + t_y$  and  $z' = z + t_z$  or in matrix form:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

### 9.7.1.2 Scaling

To scale a point  $p = (x, y, z)$  with  $S = (s_x, s_y, s_z)$  in the direction of each component in  $x, y, z$  axes, we multiply each value by the original point component, leading to the result in following matrix form:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

### 9.7.1.3 Rotation

Rotation of an arbitrary point  $p = (x, y, z)$  with angles  $\theta, \alpha, \beta$  about  $x, y, z$  axes can be represented by the following matrix equation:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta & 0 & 0 \\ -\sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

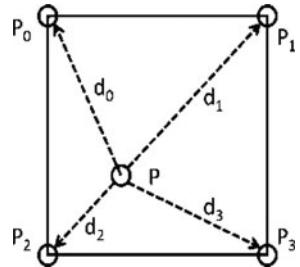
$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos \beta & 0 & -\sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

The several transformations that describe the registration process can be concatenated into a single transformation. For example, translation  $t$ , scaling  $s$ , and rotation  $r$  transformations can be combined as single  $4 \times 4$  matrix  $a = r * s * t$ .

## 9.8 Interpolation

In the process of aligning the source image to the target, every voxel within the source image is transformed using the current transformation and mapped to the target image. Since both images have discrete domains, the transform maps point from the source image to locations in the target image that are not defined on

**Fig. 9.7** Intensity value at  $P$  is computed as weighted sum of its neighboring points



a regular grid. In this case, the interpolator must estimate the value of the image at grid points in the target. There are several strategies for the interpolation of image values, each varying in computational complexity and relative quality of the estimate. The simplest among these is the nearest neighbor approach while others include linear interpolation, trilinear interpolation, and cubic interpolation.

### 9.8.1 Nearest Neighbor Interpolation

This is the simplest and fastest interpolator in terms of computation cost. With this approach, the intensity value of the mapped pixel in the source image that is nearest to the current grid-point in the target image is chosen. This approach is simple to implement, but lacks precision.

### 9.8.2 Trilinear Interpolation

Trilinear interpolation estimates the intensity value of the mapped point  $P$  as a weighted sum of the neighboring points' intensities. The normalized weight is proportional to their distance to that point. Let  $d_i$  represent the normalized distance between the neighboring points and point  $P$  as shown in Fig. 9.7, and then the interpolated intensity value at point  $P$  is given by the formula:

$$I(P) = \sum_i^4 d_i * P_i$$

## 9.9 Similarity Measures

The main component of any registration process is the similarity measure. This metric plays a crucial role in the success or failure in the search for the optimum transform that will correctly align the source image to the target image. The selection of the similarity metric is highly dependent on the registration problem

to be solved, and will often be different in single-modality or multimodality cases. For example, the mean squares and normalized correlation metrics are simple and inexpensive computationally, and can be used when images from the same modality need to be registered. For multimodality registration, we need other suitable metrics that are based on information theory. One such metric is known as mutual information, which is based on the assumption that the objects under consideration exhibit similar information properties, therefore the alignment is performed by maximizing the similarity (or minimizing the difference) between images. Although there are many similarity measures employed in medical image registration, we limit our discussion to only the basic and most often used.

### **9.9.1 Mean-Squares Metric**

The mean-squares metric is restricted to mono-modality applications. For the mean squares metric, in general, we look for the smallest difference (or error) between image intensity of the source and target images. We then square the error to make all values positive and then calculate the average (mean square) as presented in this formula:

$$s = \frac{1}{N} \sum_i^N (g(x_i) - f(x'_i))^2$$

### **9.9.2 Normalized Correlation Metric**

For image registration applications in which the intensity of the images can vary, the images should be first normalized. This is typically done at every step by subtracting the mean and dividing by the standard deviation. The cross-correlation  $s$  of a source image with the target image is given by:

$$s = \frac{\sum_i^N g(x_i)f(x'_i)}{\sqrt{\sum_i^N g^2(x_i) \sum_i^N f^2(x'_i)}}$$

### **9.9.3 Mutual Information Metric**

Registration of images from different modalities is the most challenging, especially when using a similarity metric measure, which is based on direct comparison of intensity values. In these cases, metrics based on the evaluation of mutual information are considered the most suitable option for overcoming the difficulties of

multimodality registration, where the contrast relationships in the different images may be very dissimilar.

The concept of mutual information is derived from information theory and its application to image registration has been proposed by Maes et al. [4] and Wells et al. [12]. Mutual information may be defined as:

$$M(u, v) = H(u) + H(v) - H(u, v)$$

where  $H(u)$ ,  $H(v)$  are the marginal probability distributions of  $u$ ,  $v$  and are defined as

$$H(w) = \int f(w) \log(f(w)) dw$$

where  $f(w)$  is the probability density function of a random variable  $w$ , and  $H(u, v)$  is the joint probability distribution, which can be calculated from the joint histogram of  $u$ ,  $v$ :

$$H(u, v) = - \iint f(u, v) \log(f(u, v)) du dv$$

Several variations of mutual information have been proposed and implemented to suit various requirements; among these are normalized mutual information [13] and another implementation proposed by Mattes et al. [14].

## 9.10 Optimization

Optimization is a genetic mathematical technique used to solve many minimization (maximization) problems including medical image registration. The basic requirement of an optimizer is a cost function. The optimizer searches for the cost function's parameters that produce the optimal values that would lead to a minimum function value. In terms of a registration framework, the similarity metric provides and calculates the value of the cost function for the parameter range supplied by the optimizer. In the context of image registration, the optimization parameters are representations of the transformation parameters. Transformation parameters can be defined over different domains, i.e., rotation angle and translation. A unit change in angle unit may have a much greater impact on an image registration than a unit change in translation. While this difference in the scale in the search space makes the optimization problem more difficult, rescaling the translation parameters can help to overcome this problem. There are some types of optimizers (so-called multiple valued optimizers) that require and deal with a cost function that returns multiple values at each search step. The following lists some single valued optimizers:

### 9.10.1 Amoeba Optimizer

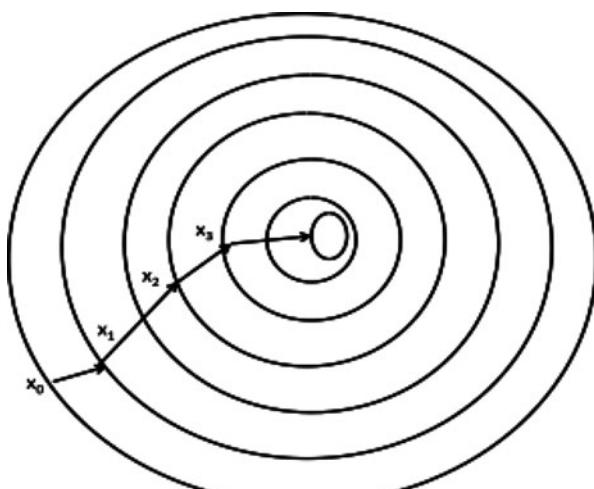
This algorithm is also known as a downhill simplex as proposed by Nelder and Mead [15]. This method is simple and requires only function evaluations, but not the derivatives of the function.

### 9.10.2 Conjugate Gradient Optimizer

If  $C(P)$  is a function and  $P$  is a  $N$ -dimensional point, **additionally** the gradient of the function  $\nabla C(P)$  can be computed (vector of first partial derivatives). The conjugate gradient method is an algorithm to find the nearest local minimum of the function.

### 9.10.3 Gradient Descent Optimizer

This approach is also known as steepest descent. The optimizer uses gradient descent in the process of finding a local minimum of the cost function. Gradient descent is based on the assumption that if the function  $C(X)$  is defined and differentiable in a neighborhood of the current point  $X_i$ , then the function  $C(X)$  decreases when the next step point  $X_{i+1}$  goes in the direction of negative gradient of  $C(X)$ , which is  $-\nabla C(X_i)$ . This means  $X_{i+1} = X_i - \gamma_i \nabla C(X_i)$  for small value  $\gamma_i > 0$  (Fig. 9.8).



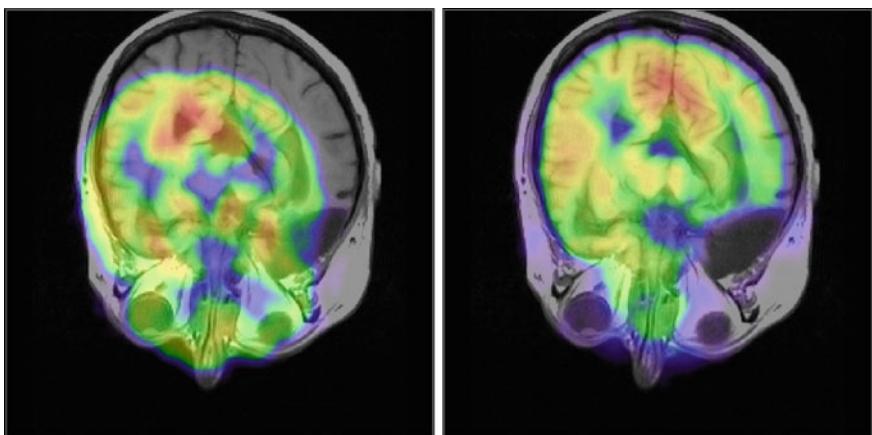
**Fig. 9.8** Gradient descent optimizer's search path

## 9.11 Image Visualization

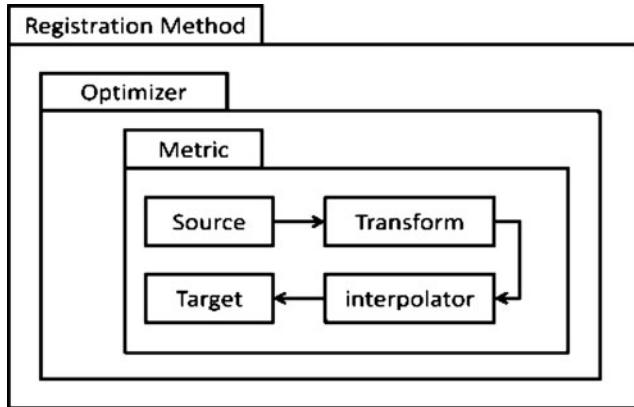
To evaluate the registration result and view the combined images, we need the source and target images to be displayed in some integrated manner, and for this purpose there are many visualization techniques that can be utilized, the most common of which is to overlay the source image on top of the target. Both images are displayed with different lookup tables and an opacity value (between [0, 1]) is applied to the foreground image's lookup table during the rendering. Figure 9.9 shows this technique using a slice from a PET brain case registered to an MR image of the same patient, left (before registration) and right (after the registration).

## 9.12 Open Source Software

There are several free open source software packages available for performing analysis of medical images (segmentation and registration) such as ITK (Insight Segmentation and Registration Toolkit) [16]. ITK is a cross-platform, open-source application development framework widely used for development of image segmentation and image registration programs. ITK was developed with funding from the National Library of Medicine (U.S.). The software is implemented in C++ and is wrapped for other languages like Tcl, Python, and Java. The toolkit provides a large collection of segmentation and registration algorithms in arbitrary dimensions. ITK uses the CMake build environment to manage the configuration and build process.



**Fig. 9.9** Example shows slice from PET brain image overly on MRI image, *left* (before registration) and *right* (after the registration)



**Fig. 9.10** The building blocks of ITK registration algorithm

### 9.12.1 ITK Pipeline

The ITK toolkit is organized around a data-flow architecture where the data are represented using data objects, which are then processed by process objects (filters). Data objects and process objects are connected together into pipelines. Pipelines are capable of processing the data in pieces according to user-specified selected filters (Fig. 9.10).

#### 9.12.1.1 ITK Implementation of Medical Image Registration

In this section, we use a simple example to demonstrate how to use the ITK library for registering two 2D images. This example illustrates the use of the image registration framework in the Insight Toolkit, and can be used as an introduction to the elements that are typically involved in solving an image registration problem. The first step in the registration task implementation is to load the source and the target images into memory. ITK supports an extensive list of various image formats and can read and load the data in the internal representation “itkImage”. The second step is to include the headers required and to define the appropriate type for object, using `typedef` statements as shown below:

```

// Transform Type
typedef itk::QuaternionRigidTransform<double> TransformType;
// Optimizer Type
typedef itk::QuaternionRigidTransformGradientDescentOptimizer OptimizerType;
// Metric Type
typedef itk::MutualInformationImageToImageMetric<ImageType, ImageType> MetricType;
// Interpolation technique
typedef itk::LinearInterpolateImageFunction<ImageType, double> InterpolatorType;
// Registration Method
typedef itk::ImageRegistrationMethod<ImageType, ImageType> RegistrationType;

```

Now we need to create the necessary objects for each of the components of the registration such as the metric, transform and interpolator, then to connect the pipeline as illustrated below:

```

MetricType::Pointer metric=MetricType::New();
TransformType::Pointer transform =TransformType::New();
OptimizerType::Pointer optimizer = OptimizerType::New();
InterpolatorType::Pointer interpolator = InterpolatorType::New();
RegistrationType::Pointer registration = RegistrationType::New();

// Connect up the components
registration->SetMetric(metric);
registration->SetOptimizer(optimizer);
registration->SetTransform(transform);
registration->SetInterpolator(interpolator);
registration->SetFixedImage(fixed);
registration->SetMovingImage(moving);
optimizer->AddObserver( itk::IterationEvent(), observer);
// Start registration
registration->StartRegistration();

```

ITK provides a mechanism to observe any event, and the base class of all ITK objects maintains a list of event observers. The typical observers register themselves to an object and request the type of events in which they are interested. To observe the registration progress or terminate the registration, one has to connect the observer to the optimizer. ITK optimizers execute an iterative process and invoke an iteration event at the end of each iteration. By executing the start function, the registration object will connect all the parts needed and will trigger the optimizer to start searching. The optimizer runs iteratively and in every step sends a call-back to the observer with the current metric and the parameters values. Once the minimum value is found or the maximum number of iteration is reached, the optimizer returns and exits the loop.

## 9.13 Registration Validation

Because there is a lack of ground truth or a gold standard for many registration problems in general, there is no easy way to validate a registration algorithm, especially when using nonrigid registration techniques. To validate a rigid registration algorithm, one can apply some known transform to an image and resample a new image. The resampled image can then be used as a source image and the

original image as a target image. The output transform from the registration algorithm can then be compared with the original applied gold standard.

## 9.14 Conclusions

In this chapter, readers at all levels should be able to gain a better understanding of medical image registrations, as well as the when, the how, and the why of applying these techniques. It provides a roadmap for how to reduce the complexity of such a pipeline, and simplify and reveal the basic structure requirement to build medical image registration applications.

We often do not know which measurements best reflect the dynamics of our problem in question. Furthermore, we sometimes have to deal with the real-world problem of “noise”. In an image registration problem, for example, such noise can be due to deformation of the body organ caused by breathing and patient movement, and by variation in the pixel intensities in the final image due to beam hardening (CT) [17] or RF penetration (MRI) effect. Such nonideal image characteristics contaminate our data set and render the problem of matching two images an ill-posed problem, where the metric tries to compare two completely different characteristics, or the optimizer becomes trapped in some local minima. For medical image registration, to accomplish any successful registration, we must select a suitable metric/optimizer that is tolerated to such organ/modality-specific problems.

## References

1. Maintz JB, Viergever MA (1998) A survey of medical image registration. *Med Image Anal* 2:1–36
2. Hill DL, Batchelor PG, Holden M, Hawkes DJ (2001) Medical image registration. *Phys Med Biol* 46:1–45
3. Mäkelä T et al (2002) A review of cardiac image registration methods. *IEEE Trans Med Imaging* 21:1011–1021
4. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16:187–198
5. Studholme C, Hawkes D, Hill D (1998) Normalized entropy measure for multimodality image alignment. In: Proceedings of SPIE Medical Imaging, vol 3338, pp 132–143
6. Hawkes DJ (1998) Algorithms for radiological image registration and their clinical application. *J Anat* 193(Pt 3):347–361
7. West J et al (1997) Comparison and evaluation of retrospective intermodality brain image registration techniques. *J Comput Assist Tomogr* 21:554–566
8. West J et al (1999) Retrospective intermodality registration techniques for images of the head: surface-based versus volume-based. *IEEE Trans Med Imaging* 18:144–150
9. Aladl UE et al (2004) Automated image registration of gated cardiac single-photon emission computed tomography and magnetic resonance imaging. *J Magn Reson Imaging* 19:283–290

10. Novelline RA, Squire LF (1997) Squire's fundamentals of radiology, 5th edn. Harvard University Press, Cambridge, MA
11. Audette MA, Ferrie FP, Peters TM (2000) An algorithmic overview of surface registration techniques for medical imaging. *Med Image Anal* 4:201–217
12. Wells WM, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1:35–51
13. Hawkes DJ, Hill D, Hajnal J (2001) Medical image registration. CRC Press, London, UK
14. Haynor DR, Vesselle H, Lewellen TK, Eubank W, Mattes D (2001) Non-rigid multimodality image registration. *Medical Imaging 2001: Image Processing*, pp 1609–1620
15. Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput J* 7:308–313
16. Ibanez L, Schroeder W, Ng L, Cates J (2005) The ITK software guide manual. Kitware Inc, Clifton Park, NY
17. Brooks RA, Di Chiro G (1976) Beam hardening in x-ray reconstructive tomography. *Phys Med Biol* 21:390–398

## Biography



**Dr. Usaf E. Aladl**, Ph.D., is a senior analyst in the Imaging Research Laboratories at Robarts Research Institute, Schulich School of Medicine & Dentistry at the University of Western Ontario, Canada. He received his B.Sc. degree from Mathematics Department, Al-Fateh University, Tripoli, Libya, his M.Sc. degree from Department of Mathematics, Middle East Technical University, Ankara, Turkey, and his Ph.D. degree from Department of Applied Mathematics, University of Western ON, London, Ontario, Canada. His research interests include numerical analysis, scattering wave problems in unbounded domains, visualization and analysis of medical images, multimodality medical image registration and segmentation, and histology image processing.



**Dr. Terry Peters**, B.E. (Hons), Ph.D., FCCPM, FAAPM, FIEEE, FACPSEM, FMICCAI, F Inst P, C Phys, is a scientist in the Imaging Research Laboratories at the Robarts Research Institute, London, ON, Canada, and professor in the Departments of Medical Imaging and Medical Biophysics at the University of Western Ontario. He holds B.Eng. and Ph.D. degrees in electrical engineering from the University of Canterbury, Christchurch, New Zealand. For the past 30 years, his research area has been in the application of computational hardware and software advances for medical imaging modalities in surgery and therapy. Starting in 1978 at the Montreal Neurological Institute (MNI), Dr. Peters' lab began developing techniques for image-guided applications in the surgical treatment of epilepsy and Parkinson's disease. Since 1997, at Robarts, he has extended his focus to image-guided surgery and therapy of the heart, brain, and abdominal organs.

# Chapter 10

## Robust Image Registration Based on Learning Prior Appearance Model

Ayman El-Baz and Georgy Gimel’farb

**Abstract** In this chapter, a novel approach to align an image of a textured object with a given prototype will be proposed. Visual appearance of the images, after equalizing their signals, is modeled with a Markov–Gibbs random field (MGRF) with pairwise interaction. Similarity to the prototype is measured by a Gibbs energy of signal co-occurrences in a characteristic subset of pixel pairs derived automatically from the prototype. An object is aligned by an affine transformation maximizing the similarity by using an automatic initialization followed by gradient search. Experiments confirm that the proposed approach aligns complex objects better than the conventional algorithms used in alignment.

### 10.1 Introduction

Image registration aligns two or more images of similar objects taken at different times, from different viewpoints, and/or by different sensors. The images are geometrically transformed to ensure their close similarity. Registration is a crucial step in many applied image analysis tasks, e.g., to fuse various data sources (such as computer tomography (CT) and MRI data in medical imaging) for image fusion, change detection, or multichannel image restoration; form and classify multi-band images in remote sensing; update maps in cartography, perform automatic quality control in industrial vision, and so forth. Coregistered medical images provide more complete information about the patient, help to monitor tumor growth, verify treatment, and allow for comparing the patient’s data to anatomical atlases.

Most of the known registration methods fall into two main categories: feature-based and area-based techniques [1]. Feature based techniques rely on salient local structures extracted from images, e.g., specific areas such as water reservoirs and lakes [2, 3], buildings [4], forests [5], or urban areas [6], specific lines such as

---

A. El-Baz (✉)

BioImaging Laboratory, Department of Bioengineering, University of Louisville, 423 Lutz Hall Building, Louisville, KY, USA  
e-mail: [aselba01@louisville.edu](mailto:aselba01@louisville.edu)

straight segments [7–9], object contours [10–12], coast lines [13, 14], rivers, or roads [15, 16], and specific points, e.g., road crossings [17], centroids of water areas, or oil and gas pads [18]. Recently, scale invariant feature transform (SIFT) proposed by Lowe [19] is the most commonly used one because of its reliability in determining a number of point-wise correspondences between two images differing by affine transformation and local contrast/offset signal deviations. These methods are limited to the applications in which objects have distinctive and nonrepetitive local features.

Area-based methods such as the classical least square correlation match directly image signals to avoid feature extraction [20], where the objects are assumed identical to within spatially uniform signal deviations and hence the correlation is too sensitive to nonuniform and spatially interdependent deviations of the corresponding signals due to sensor noise, illumination variations, and/or different sensor types. Alternative phase correlation and spectral-domain (e.g., Fourier–Mellin transform based) methods [21–24] are more robust with respect to the correlated and frequency-dependent noise and nonuniform time varying illumination. Unfortunately, these methods typically allow for only very limited geometric transformations.

More powerful mutual information (MI) based image registration [25, 26] exploits a probabilistic similarity measure that accounts for more general types of signal deviations than correlation. The statistical dependency between two data sets is measured by comparing a joint empirical distribution of the corresponding signals in the two images to the joint distribution of the independent signals. Because the MI-based registration performs the best with multimodal images [26], it is used in many of medical imaging applications. The joint distribution is estimated using Parzen windows [27, 28] or discrete histograms [29]. The main advantage of the MI is insensitivity to monotone variations of correspondence between the object and prototype signals, but the objects should be of almost identical shape apart from their affine geometrical and monotone signal transformations. The MI allows also for some nonmonotone signal correspondence variations, although they may change the visual appearance too much and hinder registration accuracy.

A more general case of registering a textured object to a prototype with similar, but not necessarily identical visual appearance under their relative 2D affine transformations and monotone variations of signal correspondences. The variations are suppressed by equalizing signals in the images will be presented in this chapter. The coregistered equalized images are described with a characteristic subset of signal co-occurrence statistics. The description implicitly “homogenizes” the images, i.e., considers them as spatially homogeneous patterns with the same statistics. In contrast to the conventional area-based techniques, similarities between the statistics rather than pixel-to-pixel correspondences are measured. Section I.B represents the equalized object and prototype images as samples of a generic Markov–Gibbs random field (MGRF) with pairwise pixel interaction. Gibbs potentials are analytically estimated from co-occurrence statistics for the prototype. Similarity between an affinely transformed object and the prototype is measured with a total Gibbs energy for a characteristic pixel neighborhood. A new algorithm for selecting the neighborhood for the MGRF model is introduced. After an automatic initialization, the affine transformation aligning the object with the

prototype is found by the gradient search for the maximum Gibbs energy of the transformed object. Experiments in Section I.C confirm that our method is more efficient for complex textured objects than more conventional SIFT and MI based registration techniques.

## 10.2 MGRF-Based Image Registration

### 10.2.1 Basic Notation

Let  $\mathbf{Q} = \{0, \dots, Q-1\}$ ;  $\mathbf{S} = \{(i, j) : i = 1, \dots, I; j = 1, \dots, J\}$ , and  $\mathbf{S}_p \subset \mathbf{S}$  denote a finite set of scalar image signals (e.g., gray levels), a rectangular arithmetic lattice supporting digital images  $\mathbf{Y}: \mathbf{S} \rightarrow \mathbf{Q}$ , and its arbitrary-shaped part occupied by the prototype, respectively. A finite set  $\mathbf{N} = \{(\zeta_1, \eta_1), \dots, (\zeta_n, \eta_n)\}$  of  $(i, j)$ -coordinate offsets defines neighbors  $\{(i + \zeta, j + \eta), (i - \zeta, j - \eta) : (\zeta, \eta) \in \mathbf{N}\} \setminus \mathbf{S}_p$  interacting with each pixel  $(i, j) \setminus \mathbf{S}_p$ . The set  $\mathbf{N}$  yields a neighborhood graph on  $\mathbf{S}_p$  to specify translation invariant pairwise interactions with  $n$  families  $C_{\zeta, \eta}$  of cliques  $c_{\zeta, \eta}(i, j) = ((i, j), (i + \zeta, j + \eta))$ . Interaction strengths are given by a vector  $\mathbf{V}^T = [\mathbf{V}_{\zeta, \eta}^T : (\zeta, \eta) \in \mathbf{N}]$  of potentials  $\mathbf{V}^T = [\mathbf{V}_{\zeta, \eta}(q, q') : (q, q') \in \mathbf{Q}^2]$  depending on signal co-occurrences; here T indicates transposition.

### 10.2.2 Image Normalization

The prototype and object images are equalized using the cumulative empirical probability distributions of their signals on  $\mathbf{S}_p$  to account for monotone (order-preserving) changes of signals (e.g., due to different illumination or sensor characteristics).

### 10.2.3 MGRF-Based Appearance Model

In this chapter, the general MGRF with multiple pairwise interaction will be used in formulating the solution. To make the proposed approach for alignment clear, the main equations of general MGRF that clarify the main idea of alignment will be rewritten. The Gibbs probability  $P(\mathbf{Y}) \propto \exp(E(\mathbf{Y}))$  of an object  $\mathbf{Y}$  aligned with the prototype  $\mathbf{Y}^o$  on  $\mathbf{S}_p$  is specified with the Gibbs energy

$$E(\mathbf{Y}) = |\mathbf{S}_p| \mathbf{V}^T \mathbf{F}(\mathbf{Y})$$

where  $\mathbf{F}^T(\mathbf{Y})$  is the vector of scaled empirical probability distributions of signal co-occurrences over each clique family:  $\mathbf{F}^T(\mathbf{Y}) = [\rho_{\zeta,\eta} \mathbf{F}_{\zeta,\eta}^T(\mathbf{Y}) : (\zeta, \eta) \in \mathbf{N}]$  where  $\rho_{\zeta,\eta} = (|C_{\zeta,\eta}|)/(|S_p|)$  is the relative size of the family and  $\mathbf{F}_{\zeta,\eta}(\mathbf{Y}) = [f_{\zeta,\eta}(q, q'|\mathbf{Y}) : (q, q') \in \mathbf{Q}^2]^T$ ; here,  $f_{\zeta,\eta}(q, q'|\mathbf{Y}) = (|C_{\zeta,\eta;q,q'}(\mathbf{Y})|)/(|C_{\zeta,\eta}|)$  are empirical probabilities of signal co-occurrences, and  $C_{\zeta,\eta;q,q'}(\mathbf{Y}) \subseteq C_{\zeta,\eta}$  is a subfamily of the cliques  $c_{\zeta,\eta;q,q'}(i, j)$  supporting the co-occurrence ( $Y_{i,j} = q$ ,  $Y_{i+\zeta,j+\eta} = q'$ ) in  $\mathbf{Y}$ . The co-occurrence distributions and the Gibbs energy for the object are determined over  $S_p$ , i.e., within the prototype boundary after an object is affinely aligned with the prototype. The initial image has been resampled to the back-projected  $S_p$  by interpolation, in order to compensate for the affine transformation.

The appearance model consists of the neighborhood  $\mathbf{N}$  and the potential  $\mathbf{V}$  to be learned from the prototype.

#### 10.2.4 Learning the Potentials

The MLE of  $\mathbf{V}$  is proportional in the first approximation to the scaled centered empirical co-occurrence distributions for the prototype [30]:

$$\mathbf{V}_{\zeta,\eta} = \lambda \rho_{\zeta,\eta} \left( \mathbf{F}_{\zeta,\eta}(\mathbf{Y}^o) - \frac{1}{Q^2} \mathbf{U} \right); \quad (\zeta, \eta) \in \mathbf{N}$$

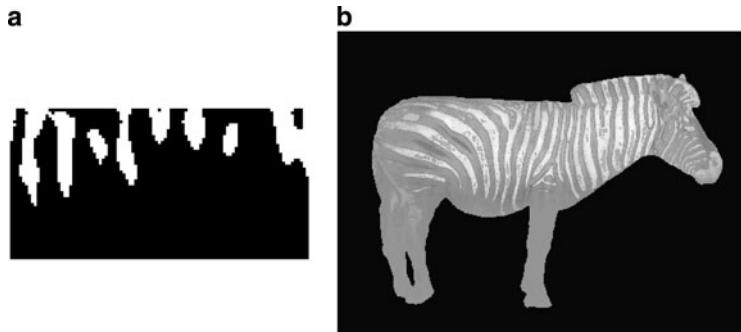
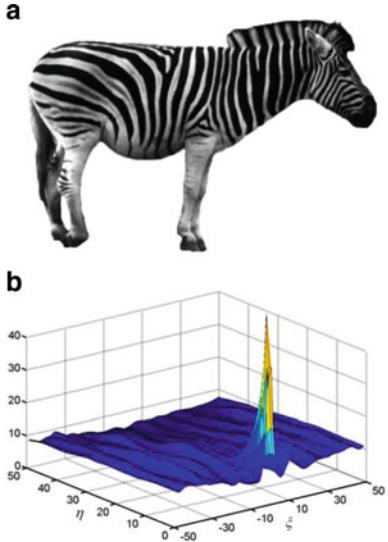
where  $\mathbf{U}$  is the vector with unit components. The common scaling factor  $\lambda$ , is also computed analytically; it is approximately equal to  $Q^2$  if  $Q >> 1$  and  $\rho_{\zeta,\mu} \approx 1$  for all  $(\zeta, \eta) \in \mathbf{N}$ . In the case of interest, it can be set to  $\lambda = 1$  because the registration uses only relative potential values and energies.

#### 10.2.5 Learning the Characteristic Neighbors

To find the characteristic neighborhood set  $\mathbf{N}$ , the relative energies  $E_{\zeta,\eta}(\mathbf{Y}^o) = \{\rho_{\zeta,\eta} \mathbf{V}_{\zeta,\eta}^T \mathbf{F}_{\zeta,\eta}(\mathbf{Y}^o)\}$  for the clique families, i.e., the scaled variances of the corresponding empirical co-occurrence distributions, are compared for a large number of possible candidates. Figure 10.1 shows a zebra prototype and its Gibbs energies  $E_{\zeta,\eta}(\mathbf{Y}^o)$  for 5,000 clique families with the inter-pixel offsets  $|\zeta| \leq 50; 0 \leq \eta \leq 50$ .

To automatically select the characteristic neighbors, we consider an empirical probability distribution of the energies as a mixture of a large “non-characteristic” low-energy component and a considerably smaller characteristic high-energy component:  $P(E) = wP_{lo}(E) + (1-w)P_{hi}(E)$ . Both the components  $P_{lo}(E), P_{hi}(E)$  are of arbitrary shape and thus are approximated with linear combinations of positive and negative discrete Gaussians [31].

**Fig. 10.1** Zebra prototype  
**(a)** and relative interaction energies  
**(b)** for the clique families in function of the  
offsets  $(\zeta, \eta)$



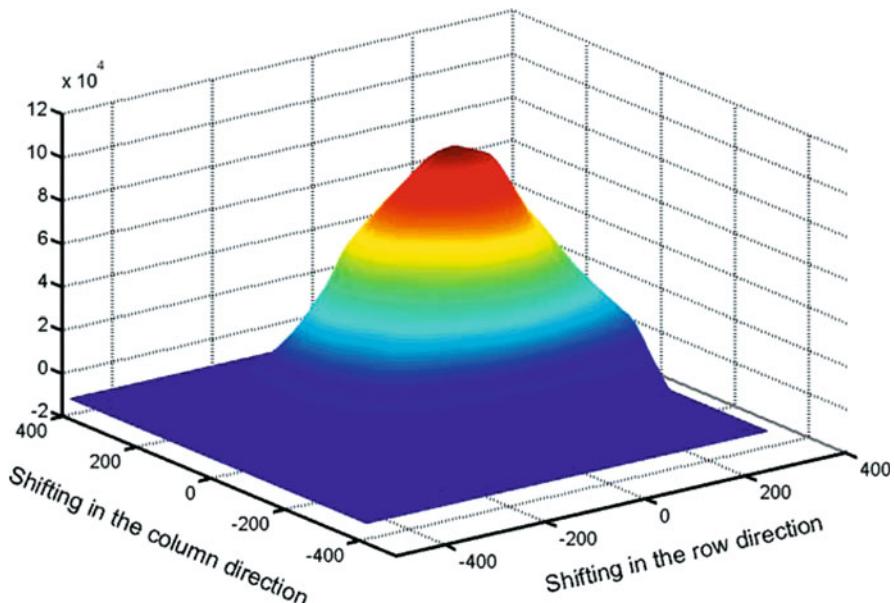
**Fig. 10.2** **(a)** Characteristic 168 neighbors among the 5,000 candidates **(a;** in white) and the pixel-wise Gibbs energies **(b)** for the prototype under the estimated neighborhood

The intersection of the approximate mixture components gives an energy threshold  $t$  for selecting the characteristic neighbors:  $N = \{(\zeta, \eta) : E_{\zeta, \eta}(g^o) \geq t\}$ , where  $P_{hi}(E) \geq P_{lo}(E)w/(1-w)$ . The above example results in the threshold  $t = 28$  producing 168 characteristic neighbors shown in Fig. 10.2 together with the corresponding relative pixel-wise energies  $e_{i,j}(\mathbf{Y}^o)$  over the prototype:

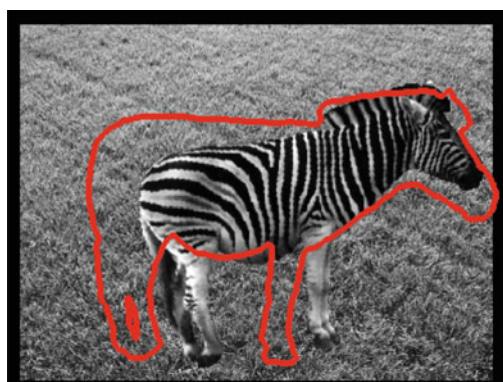
$$e_{i,j}(\mathbf{Y}^o) = \sum_{(\zeta, \eta) \in N} \mathbf{V}_{\zeta, \eta}(\mathbf{Y}_{i,j}^o, Y_{i+\zeta, j+\eta}^o)$$

### 10.2.6 Appearance-Based Registration

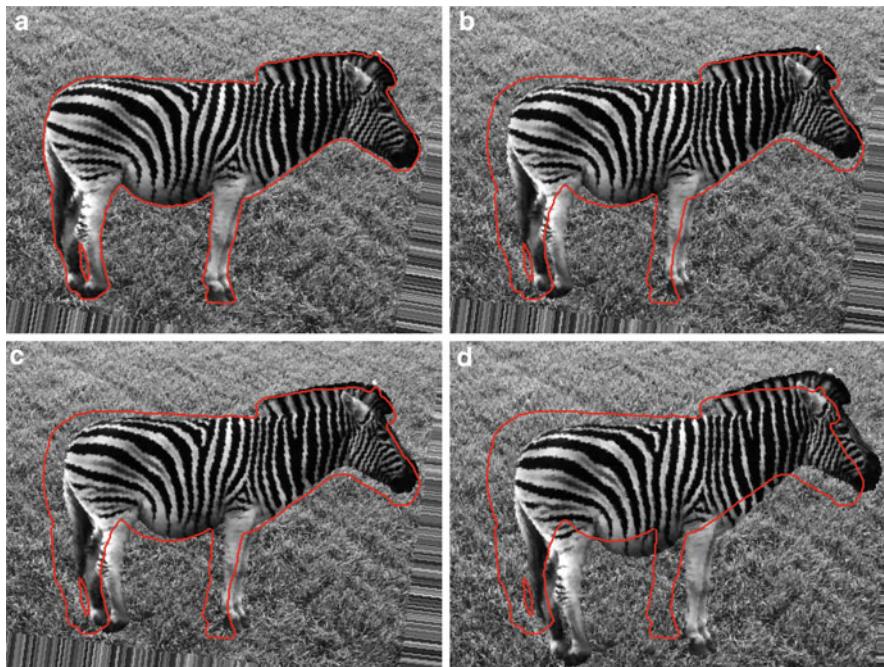
The object  $Y$  is affinely transformed to (locally) maximize its relative energy  $E(g_a) = \mathbf{V}^T \mathbf{F}(Y_a)$  under the learned appearance model  $[N; V]$  where,  $Y_a$  is the part of the object image reduced to  $S_p$  by the affine transformation  $a = [a_{11}, \dots, a_{23}] : i' = a_{11}i + a_{12}j + a_{13}; j' = a_{21}i + a_{22}j + a_{23}$ . The initial transformation is a pure translation with  $a_{11} = a_{22} = 1$  and  $a_{12} = a_{21} = 0$ , ensuring the most “energetic” overlap between the object and prototype. The energy for different translations  $(a_{13}, a_{23})$  of the object relative to the prototype is shown in Fig. 10.3; the chosen initial position  $(a_{13}^*, a_{23}^*)$  in Fig. 10.4 maximizes this energy. Then the



**Fig. 10.3** Gibbs energies for translations of the object with respect to the prototype



**Fig. 10.4** Initial position of the object with respect to the prototype



**Fig. 10.5** Our (a), MI-based (b), NMI-based (c), and SIFT-based (d) registration

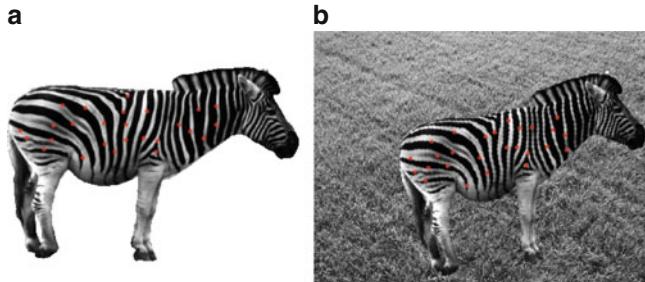
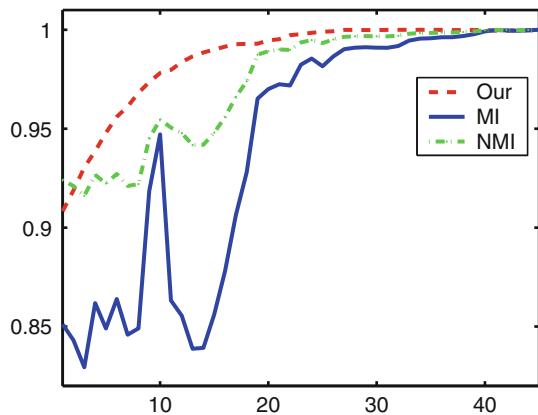
gradient search for the local energy maximum closest to the initialization selects the six parameters  $a$ ; Fig. 10.5a shows the final transformation aligning the prototype contour to the object.

### 10.3 Experimental Results

The proposed approach has been applied on three examples to test its validity and evaluate its performance. These three examples are zebra photos, dynamic contrast enhanced magnetic resonance imaging (DCE-MRI) of a human kidney, and low dose computed tomography (LDCT) of a human lung commonly perceived as difficult for both the area- and feature-based registration. The results are obtained for several other types of complex objects (e.g., starfish photos or brain images). The proposed approach has been compared to three of the most commonly used registration techniques, namely, to the area-based registration using MI [26] or normalized MI [29] as well as the feature-based registration establishing correspondences between the images with SIFT [19]. Results are shown in Fig. 10.5.

Figure 10.6 clarifies the reason behind the lack of accuracy of the MI- or NMI-based alignment. It compares the MI/NMI and Gibbs energy values for the affine parameters that appear at successive steps of the gradient search for the maximum

**Fig. 10.6** Gibbs energy, MI, and NMI values at the successive steps of the gradient search

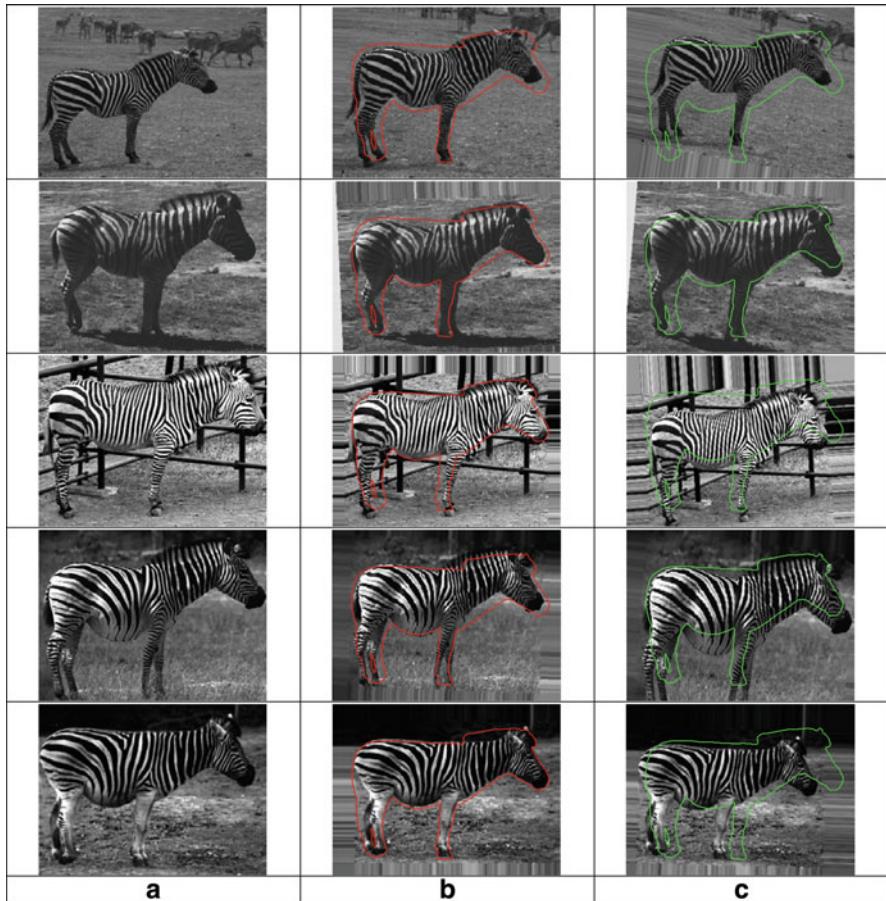


**Fig. 10.7** Corresponding points by SIFT

energy. Both the MI and NMI have many local maxima that potentially hinder the search, whereas the energy is practically unimodal in these experiments. The SIFT-based alignment fails because it cannot establish accurate correspondences between similar zebra stripes (Fig. 10.7).

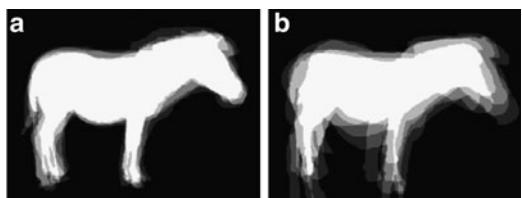
In Fig. 10.5, the object aligned with the prototype has mainly different orientation and scale. Figure 10.8 shows more diverse zebras and their Markov–Gibbs appearance-based and MI-based alignment with the prototype in Fig. 10.1a. Visually, the back-projection of the prototype contour onto the objects emphasizes the better performance of the proposed approach. To quantitatively evaluate the accuracy, masks of the coaligned objects obtained by manual segmentation are averaged in Fig. 10.9. The common matching area is notably larger for our approach (91.6%) than for the MI-based registration (70.3%).

Similar results were obtained for the kidney images as shown in Figs. 10.10–10.14. The common matching area is 90.2% for the proposed approach vs. 62.6% for the MI-based one. The approach has been applied to a more challenging category of images; LDCT lung images. The results of LDCT lung images are shown in Figs. 10.15–10.19; the common matching area is 96.8% for the proposed approach vs. 54.2% for the MI-based one.



**Fig. 10.8** Original zebras (a) aligned with our (b) and the MI-based (c) approaches

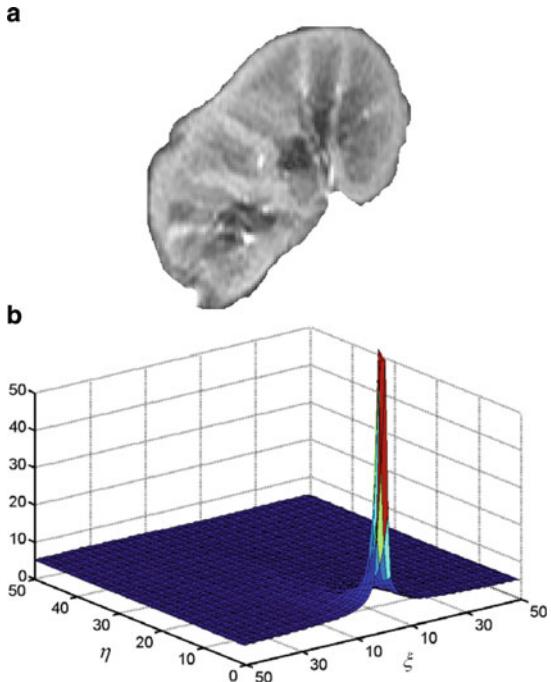
**Fig. 10.9** Overlap between the object masks aligned with our (a; 91.6%) and the MI-based approaches (b; 70.3%)



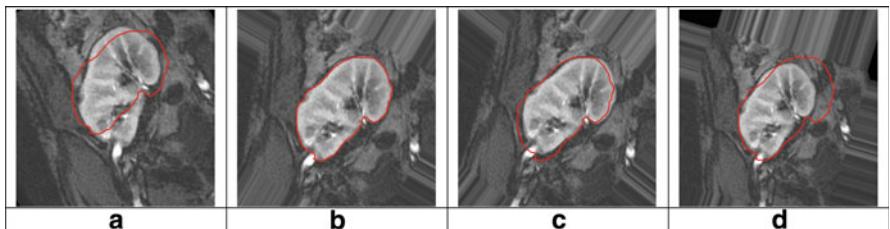
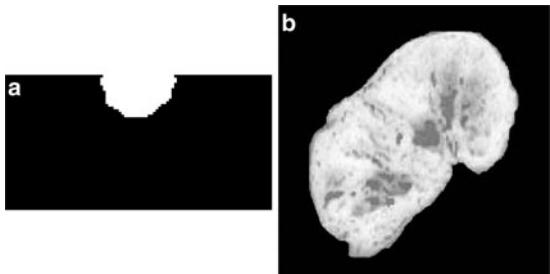
## 10.4 Summary

In this chapter, a novel approach to align an image of a textured object with a given prototype whose appearance is modeled with a Markov–Gibbs random field with pairwise interaction has been proposed. Experimental results confirm that image

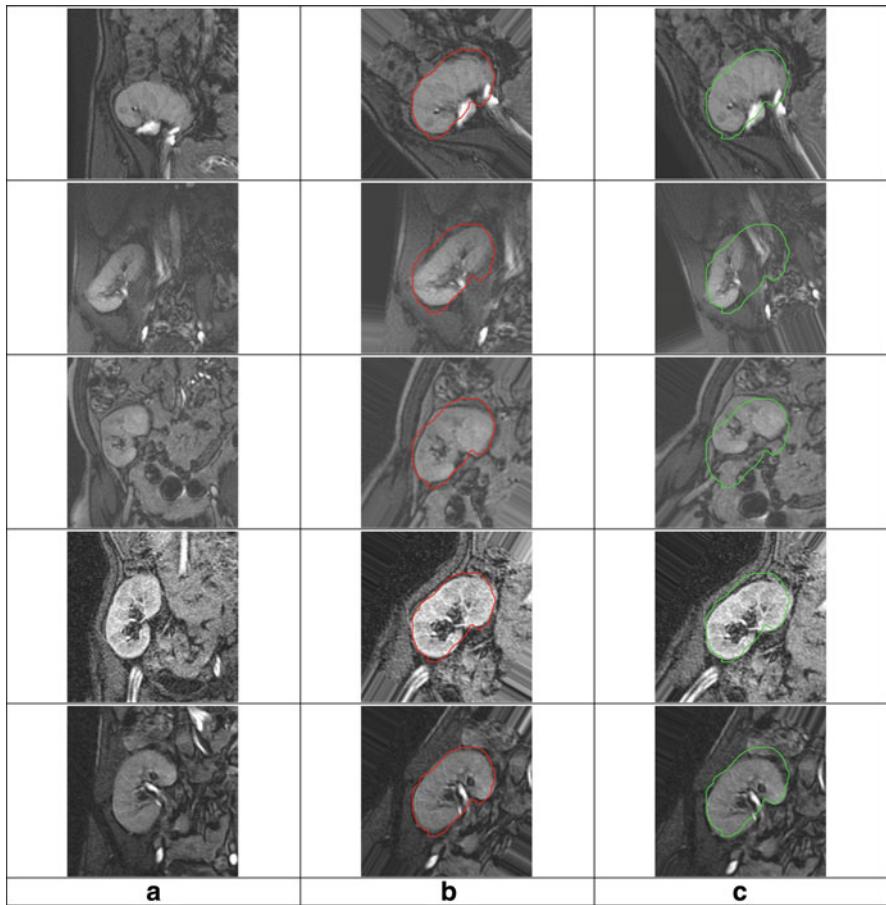
**Fig. 10.10** Kidney image (a) and relative interaction energies (b) for the clique families in function of the offsets  $(\zeta, \eta)$



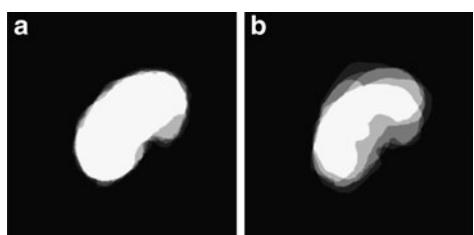
**Fig. 10.11** (a) Most characteristic 76 neighbors among the 5,000 candidates (a; in white) and the pixel-wise Gibbs energies (b) for the prototype under the estimated neighborhood



**Fig. 10.12** Initialization (a) and our (b), MI-based (c), and SIFT-based (d) registration

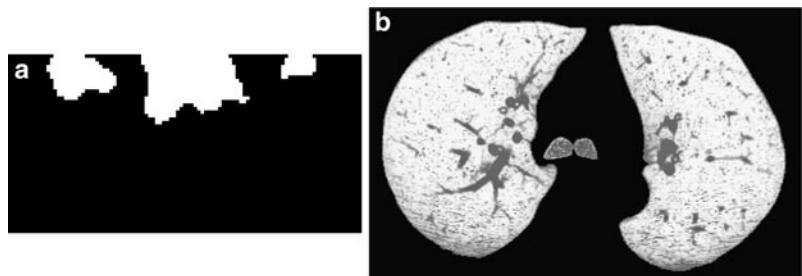
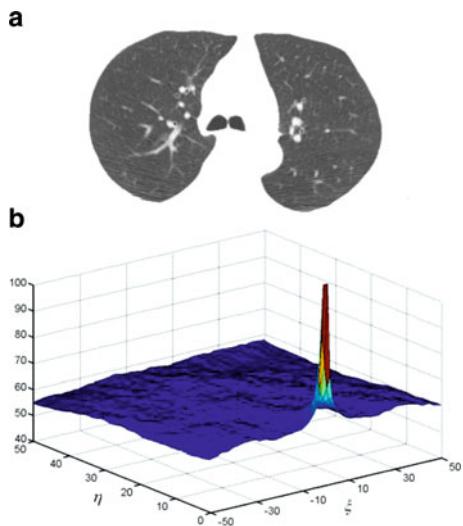


**Fig. 10.13** Original kidneys (a) aligned with our (b) and the MI-based (c) approaches

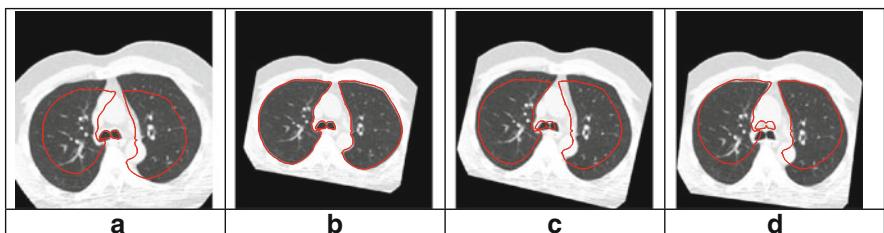


**Fig. 10.14** Overlap between the object masks aligned with our (a; 90.2%) and the MI-based approaches (b; 62.6%)

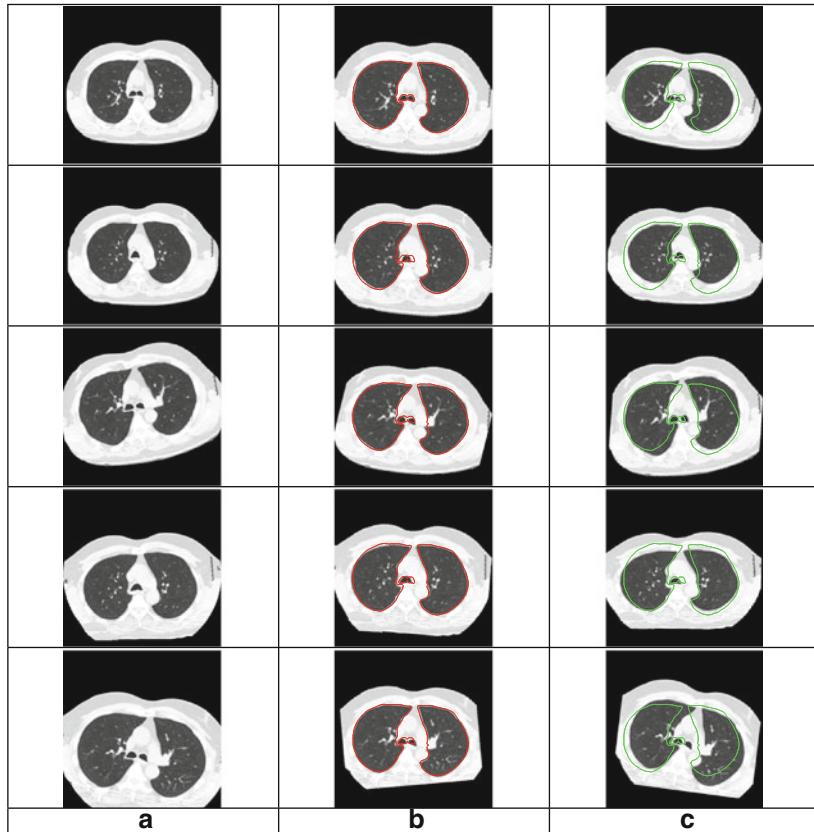
**Fig. 10.15** Lung image (**a**) and relative interaction energies (**b**) for the clique families in function of the offsets  $(\eta, \zeta)$



**Fig. 10.16** (**a**) Most characteristic 173 neighbors among the 5,000 candidates ( $\alpha$ ; in white) and the pixel-wise Gibbs energies (**b**) for the prototype under the estimated neighborhood

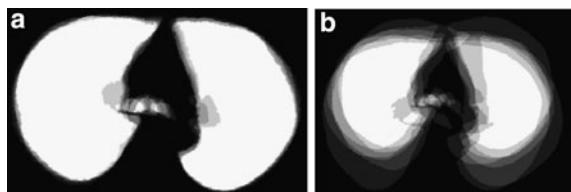


**Fig. 10.17** Initialization (**a**) and our (**b**), MI-based (**c**), and SIFT-based (**d**) registration



**Fig. 10.18** Original lungs (a) aligned with our (b) and the MI-based (c) approaches

**Fig. 10.19** Overlap between the object masks aligned with our (a; 96.8%) and the MI-based approaches (b; 54.2%)



registration based on the proposed Markov–Gibbs appearance model is more robust and accurate than popular conventional algorithms. Due to the reduced variations between the coaligned objects, the proposed approach results in more accurate average shape models that are useful, e.g., in image segmentation based on shape priors. The proposed approach is not only limited to zebra photos, kidney, and lung images, but it is general and suitable to register other types of images such as starfish photos and brain images.

## References

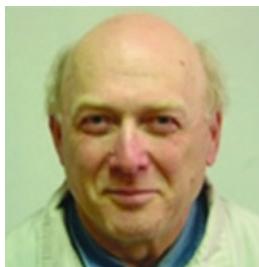
1. Zitova B, Flusser J (2003) Image registration methods: a survey. *Image Vis Comput* 21:977–1000
2. Goshtasby A, Stockman GC (1985) Point pattern matching using convex hull edges. *IEEE Trans Syst Man Cybern* 15:631–637
3. Holm M (1991) Towards automatic rectification of satellite images using feature based matching. In: Proceedings international geoscience and remote sensing symposium IGARSS'91, Espoo, Finland, pp 2439–2442
4. Hsieh JW, Liao HYM, Fan KC, Ko MT, Hung YP (1997) Image registration using a new edge-based approach. *Comput Vis Image Underst* 67:112–130
5. Sester M, Hild H, Fritsch D (1998) Definition of ground control features for image registration using GIS data. In: Proceedings symposium on object recognition and scene classification from multispectral and multisensor pixels, CD-ROM, Columbus, OH
6. Roux M (1996) Automatic registration of spot images and digitized maps. In: Proceedings IEEE international conference on image processing ICIP'96, Lausanne, Switzerland, pp 625–628
7. Hsieh YC, McKeown DM, Perlant FP (1992) Performance evaluation of scene registration and stereo matching for cartographic feature extraction. *IEEE Trans Pattern Anal Mach Intell* 14:214–237
8. Moss S, Hancock ER (1997) Multiple line-template matching with EM algorithm. *Pattern Recognit Lett* 18:1283–1292
9. Wang WH, Chen YC (1997) Image registration by control points pairing using the invariant properties of line segments. *Pattern Recognit Lett* 18:269–281
10. Dai X, Khorram S (1997) Development of a feature-based approach to automated image registration for multitemporal and multisensor remotely sensed imagery. In: Proceedings international geoscience and remote sensing symposium IGARSS'97, Singapore, pp 243–245
11. Govindu V, Shekhar C, Chellappa R (1998) Using geometric properties for correspondenceless image alignment. In: Proceedings international conference on pattern recognition (ICPR'98), Brisbane, Australia, pp 37–41
12. Li H, Manjunath BS, Mitra SK (1995) A contour-based approach to multisensory image registration. *IEEE Trans Image Process* 4:320–334
13. Maitre H, Wu Y (1987) Improving dynamic programming to solve image registration. *Pattern Recognit* 20:443–462
14. Shin D, Pollard JK, Muller JP (1997) Accurate geometric correction of atsrimages. *IEEE Trans Geosci Remote Sens* 35:997–1006
15. Li SZ, Kittler J, Petrou M (1992) Matching and recognition of road networks from aerial images. In: Proceedings 2nd European conference on computer vision ECCV'92, St Margherita, Italy, pp 857–861
16. Mendoza EH, Santos JR, Rosa ANCS, Silva NC (2004) Land use/land cover mapping in brazilian amazon using neural network with aster/terra data. In: Proceedings geo-imaging bridging continents, Istanbul, Turkey, pp 857–861
17. Growe S, Tonjes R (1997) A knowledge based approach to automatic image registration. In: Proceedings IEEE international conference on image processing ICIP'97, Santa Barbara, California, pp 228–231
18. Ton J, Jain AK (1989) Registering landsat images by point matching. *IEEE Trans Geosci Remote Sens* 27:642–651
19. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60:91–110
20. Pope P, Theiler J (2003) Automated image registration (AIR) of MTI imagery. In: Proceedings SPIE 5093, Santa Barbara, CA, pp 294–300
21. Anuta PE (1970) Spatial registration of multispectral and multitemporal digital imagery using fast fourier transform. *IEEE Trans Geosci Electron* 8:353–368

22. Chen Q, Defrise M, Deconinck F (1994) Symmetric phase-only matched filtering of Fourier-Mellin transform for image registration and recognition. *IEEE Trans Pattern Anal Mach Intell* 16:1156–1168
23. Reddy BS, Chatterji BN (1994) An fft-based technique for translation, rotation and scale-invariant image registration. *IEEE Trans Image Process* 16:1266–1271
24. Foroosh H, Zerubia JB, Berthod M (2002) Extension of phase correlation to subpixel registration. *IEEE Trans Image Process* 11:188–200
25. Inglada J, Giros A (2004) On the possibility of automatic multisensor image registration. *IEEE Trans Geosci Remote Sens* 42(10):2104–2120
26. Viola P (1995) Alignment by maximization of mutual information. PhD Thesis, MIT, Cambridge, MA
27. Thevenaz P, Unser M (1998) Alignment an efficient mutual information optimizer for multi-resolution image registration. In: Proceedings IEEE international conference on image processing ICIP'98, Chicago, IL, pp 833–837
28. Thevenaz P, Unser M (1996) A pyramid approach to sub-pixel image fusion based on mutual information. In: Proceedings IEEE international conference on image processing ICIP'96, Lausanne, Switzerland, pp 265–268
29. Studholme C, Hill DLG, Hawkes DJ (1999) An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recognit* 32(10):71–86
30. Gimel'farb G, Farag AA (2005) Texture analysis by accurate identification of simple markovian models. *Cybern Syst Anal* 41(1):37–49
31. El-Baz A, Gimel'farb G (2007) EM based approximation of empirical distributions with linear combinations of discrete Gaussians. In: Proceedings of IEEE international conference on image processing (ICIP'07), vol IV, pp 373–376, San Antonio, TX, USA, 16–19 September 2007

## Biography



**Ayman S. El-Baz** received B.Sc. and M.S. degrees in electrical engineering from Mansoura University, Egypt, in 1997 and 2000, respectively and the Ph.D. degree in electrical engineering from University of Louisville, Louisville, KY. He joined the Bioengineering Department, University of Louisville, in August 2006. His current research includes developing new computer assisted diagnosis systems, image modeling, image segmentation, 2D and 3D registration, visualization, and surgical simulation including finite element analysis, about which he has authored or coauthored more than 70 technical articles.



**Georgy Gimel'farb** graduated from Kiev Polytechnic Institute, Kiev, Ukraine, and received the Ph.D. degree in engineering cybernetics from the Institute of Cybernetics, Academy of Sciences of the Ukraine, and the D.Sc. (Eng) degree in control in engineering from the Higher Certifying Commission of the USSR, Moscow, Russia. After a long time with the Institute of Cybernetics, he joined the University of Auckland, Auckland, New Zealand, in July 1997, where he is currently an Associate Professor of computer science. His research is focused on image analysis, computer vision, and statistical pattern recognition, with main contributions in the areas of computational stereovision and probabilistic texture modeling and analysis.

# Chapter 11

## Image Registration in Medical Imaging: Applications, Methods, and Clinical Evaluation

**Gerardo Hermosillo, Christophe Chefd’hotel, K.-H. Herrmann, Guillaume Bousquet, Luca Bogoni, Kallol Chaudhuri, Dorothee R. Fischer, Christian Geppert, Rolf Janka, Arun Krishnan, Berthold Kiefer, Ines Krumbein, Werner Kaiser, Michael Middleton, Wanmei Ou, Jürgen R. Reichenbach, Marcos Salganicoff, Melanie Schmitt, Evelyn Wenkel, Susanne Wurdinger, and Li Zhang**

**Abstracts** In this chapter, we describe in detail algorithms that tackle specific clinical problems and that have been extensively validated in close collaboration with radiologists. Our hope is that its contents will be a useful overview of some of the problems that present themselves in real clinical practice and of some of the techniques that have proven to fulfill the challenging requirements of these problems in terms of speed and robustness, and that have become properly validated products.

**Keywords** Image registration · Motion compensation · Multi-resolution approach

### 11.1 Introduction

Image registration refers to the ability to establish correspondences between locations in two or more images. The need for this ability is ubiquitous in medical imaging. Automatic or semiautomatic methods that bring images into a common reference frame are keystone components of many clinical applications such as analysis of contrast agents in perfusion imaging, lesion localization and comparison in prior studies, atlas based methods for learning normal and abnormal anatomy, analysis of therapy effectiveness on tumors, alignment of the imaging device depending on the region of interest, evaluation of stenosis in arteries and veins, motion analysis of heart walls and indexing methods that group images of similar characteristics. The large number of modalities and protocols that are used in the evaluation of patients nowadays makes it challenging to design automatic methods to put their complementary data into correspondence. At the same time, the images being acquired are of ever increasing resolution and size and the fast-paced

---

M. Salganicoff (✉)  
Siemens Computer Aided Diagnosis & Therapy, Malvern, PA, USA  
e-mail: marcos.salganicoff@siemens.com

environment created by large volume of procedures and patients, or by emergency situations, create considerable challenges for automatic registration algorithms in terms of performance and reliability. In clinical applications, having access to sophisticated automatic methods for registration may improve the patient outcome and the physician's effectiveness and efficiency. In summary, registration is a challenging and, in most cases, ill-posed problem that nevertheless is part of everyday clinical problems in need for solutions.

Mathematically, the registration problem for three-dimensional scalar images can be stated in a very concise fashion: Find  $\phi$  such that:

$$\begin{cases} I_1(x) = I_2(\phi(x)) & \forall x \in \Omega_1 \subset \mathbf{R}^3 \\ I_2(y) = I_1(\phi^{-1}(y)) & \forall y \in \Omega_2 \subset \mathbf{R}^3 \end{cases}$$

where  $\sim$  denotes “correspondence”. In this setting, the unknown is a function  $\phi : \Omega_1 \rightarrow \Omega_2$  and its inverse  $\phi^{-1}$  that provide correspondences between locations in both images. The images to be registered are viewed as functions  $I_1 : \Omega_1 \rightarrow \mathbf{R}$  and  $I_2 : \Omega_2 \rightarrow \mathbf{R}$ .

Within this simple formulation of the registration problem, many open issues allow for almost infinite variations of the details and components that may go into a particular engineering solution to a registration problem. For example:

- How is “correspondence” determined? This is usually based on some implicit or explicit local similarity measure.
- What kind of transformation is assumed? This could be rigid, affine, elastic, etc.
- Is the transformation constrained by landmarks?
- Are landmarks detected automatically?
- Is the similarity of local regions based on multiple image features?
- Are those features automatically learned from a large family or are they carefully designed and chosen?
- How is the similarity optimized? In many cases, a global optimum of the similarity criterion is not possible and perhaps not even desirable. In that case, the method used to optimize plays a significant role both in the results that should be expected as well as in the algorithm performance.
- Are explicit deformation models used? These can include physical models trying to simulate the way a specific tissue deforms. Once such models are available, their parameters can be optimized with respect to image similarity.
- Is the registration constrained by either segmentation of local regions or landmarks?
- Is the transformation invertible? Is the inverse computed simultaneously?
- Is the problem solved in a symmetric way with respect to both images?
- Given a set of registration problems that are all part of one application, should they be solved separately or simultaneously?
- How is the solution validated?

In this chapter, we describe in detail algorithms that tackle specific clinical problems and that have been extensively validated in close collaboration with

radiologists. Our hope is that its contents will offer a useful overview of some of the problems that present themselves in real clinical practice and of some of the techniques that have proven to address the challenging requirements of these problems in terms of speed and robustness, and that have become properly validated products. The remainder of the chapter is structured as follows. In Sect. 11.2 we describe a method for motion compensation in temporal sequences such as those encountered in breast MR or other perfusion studies. In Sect. 11.3, we dwell into methods producing diffeomorphisms and provide innovative insights on the subject. In Sect. 11.4, we present a method that models the relationships between intensities in two modalities explicitly, making it possible to apply fast optimization methods when dealing with multimodal image registration. In Sect. 11.5, complementing the subject of registering a sequence of images, we present a method dealing with the possibility of optimizing all deformations simultaneously. In Sect. 11.6, we present example studies on the evaluation of these methods within a clinical environment. We present some concluding remarks in Sect. 11.7.

## 11.2 A System for Motion Compensation in a Temporal Sequence of Images

This section describes a system for performing motion compensation in a temporal sequence of images. The compensation is done by performing a conjugate gradient maximization of a similarity measure between two images, based on the local cross-correlation of corresponding regions around each point. The nonsingularity of the deformation is ensured by a special composition technique when building the solution. Methods producing diffeomorphisms play an important role in problems that require a guaranty that structures are not overly shrunk or expanded. The system is targeted at solving the motion correction problem in Breast MR. Potential tumors are detected considering the rapid intake (wash in) of contrast agent, as well as a rapid wash out, in breast tissue. In order to determine which regions in the acquired MR volume are suspect, one needs to track a single voxel in a temporal sequence of such volumes. However, since the patient moves between consecutive acquisitions then a single point in space can no longer be tracked, unless motion correction is performed.

Many of the existing methods solve the image registration problem by looking for a displacement  $u(x)$  at each point  $x \in \mathbf{R}^3$  such that  $\phi(x) = x + u(x)$  where one of the images is chosen as reference. This problem can be dealt with in the context of the calculus of variations. Optimization of a given similarity function can be accomplished in several ways. One of these is performed by computing a gradient for the similarity function and applying a gradient-ascent algorithm to optimize it starting from an initial deformation, usually the null deformation or identity mapping. Another component of automatic registration algorithms besides the similarity function

is the particular model assumed for the deformation. Common models include rigid deformation, affine deformation, and the sum of a set interpolating functions based on control points. In turn, there are usually also multiple choices for the interpolating function.

The problem of estimating the geometric deformation between two images has a long history in the scientific literature. Techniques for computing the optic flow can be traced back to papers like [1] and references therein. The use of the cross-correlation as similarity measure can be found in [2–6]. Other related similarity measures have been proposed, like the correlation ratio [7] and the mutual information [8–11], among others [12–15]. Conjugate Gradient minimization is described in [16]. The type (or family) of deformation which is assumed is the second key component of any motion correction algorithm. Parametric transformations are the most commonly used [8–10, 17, 18]. When the deformation is not defined parametrically, the family is often constrained by requiring some smoothness of the displacement field, possibly preserving discontinuities [19–27]. Some regularizing approaches are based on explicit smoothing of the field, as in [19, 28], while others consider an additive term in the error criterion, yielding (possibly anisotropic) diffusion terms [29–33]. Fluid methods fix the amount of desired smoothness or fluidness of the deformation using a single parameter [28, 34, 35]. Multi-resolution approaches have also being previously investigated. In [22], a scale-space focusing strategy is used, while in [34], the authors look for a continuously invertible mapping, which is obtained by the composition of small displacements. This approach has been generalized using Lie group ideas on sets of diffeomorphisms in [35]. Under a similar formalism, a very general framework, which also allows for changes in the intensity values is proposed in [36]. Most of the existing methods either do not account for intensity variations or are limited to parametric transformations. Extensions to more complex transformations, which account for intensity variations include approaches relying on block-matching strategies [37–39], or parametric intensity corrections [40]. Some recent approaches rely on the computation of the gradient of the local cross-correlation [4, 6, 28, 31–33].

### ***11.2.1 Design Constraints***

The following design constraints were identified as appropriate for the registration module:

- Invertibility – The algorithm should provide both forward and reverse maps, these approximates inverses of one another. Notice that this requires that the transformation found by the algorithm be invertible.
- Locality and format of transformations – The algorithm should be able to run in an arbitrarily defined region of interest and provide its result in the form of a

vector map defined over this region. Each vector is to represent the amount of displacement in millimeters.

- Composition of transformations – The algorithm should be able to present its result taking into account previous deformations, presented as local deformation maps (at most ten of them). These deformation maps are not constrained to be applied on the same region or to have the same resolution.
- To provide interactivity – It should be possible to interrupt the registration at any point and still have an estimated deformation map. The allowed interruptions should be spaced by about 10 s.
- Computational requirements – the initial rigid/first order computation should be completed within 20–30 s. Any subsequent refinements and chain compositions should be completed within 20–30 s each, such that a reasonable result is produced in less than 2 min.
- Memory footprint – algorithmic requirements (exclusion made for deformation maps) should not exceed a predefined size.

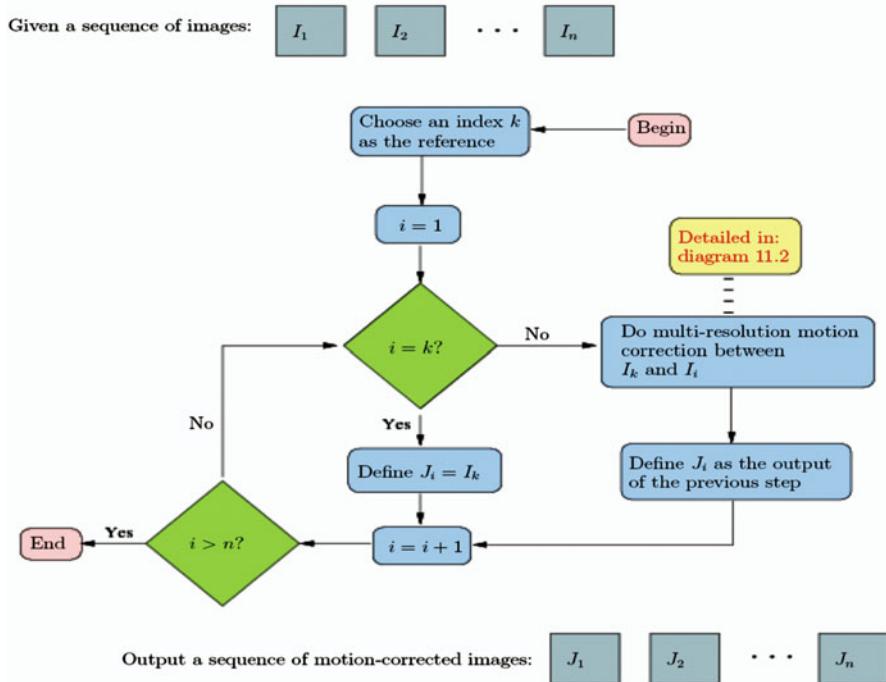
### ***11.2.2 Global Handling of the Sequence and Multi-Resolution Approach***

The system is designed to work with a set of images as input, which is assumed to be a temporal sequence of similar images, meaning for instance images of the same patient acquired at different points in time. For breast MR in particular, 6–14 images are acquired. The interval between acquisitions is typically 2–3 min. Each image is a three-dimensional array of scalar values covering roughly the chest area of the patient. The output of the system is again a set of images (one for each input image) which are “motion-corrected”. The input images are similar to one another but differ mainly because of three factors:

- Motion of the patient between acquisitions
- Intensity modifications due to blood intake of an injected contrast agent.
- Noise

The output sequence is obtained by choosing a reference image from the input sequence and finding, for each of the remaining ones, a spatial nonrigid deformation that, applied to the particular image considered, compensates for the motion that has occurred with respect to the reference. This overall procedure is described in Diagram 11.1.

The input of the motion compensation module is: a pair of images, one being the reference and the other, which we call floating image, being the one to which the compensating deformation is to be applied. The computations are performed using a multi-resolution scheme, which allows for large, global motions to be quickly recovered at low resolutions. The deformation (or displacement) obtained at low resolution is used to initialize the search in the next finer resolution. The deformation obtained at the finest level is applied to the floating image to yield



**Diagram 11.1** Motion correction given a temporal sequence of images

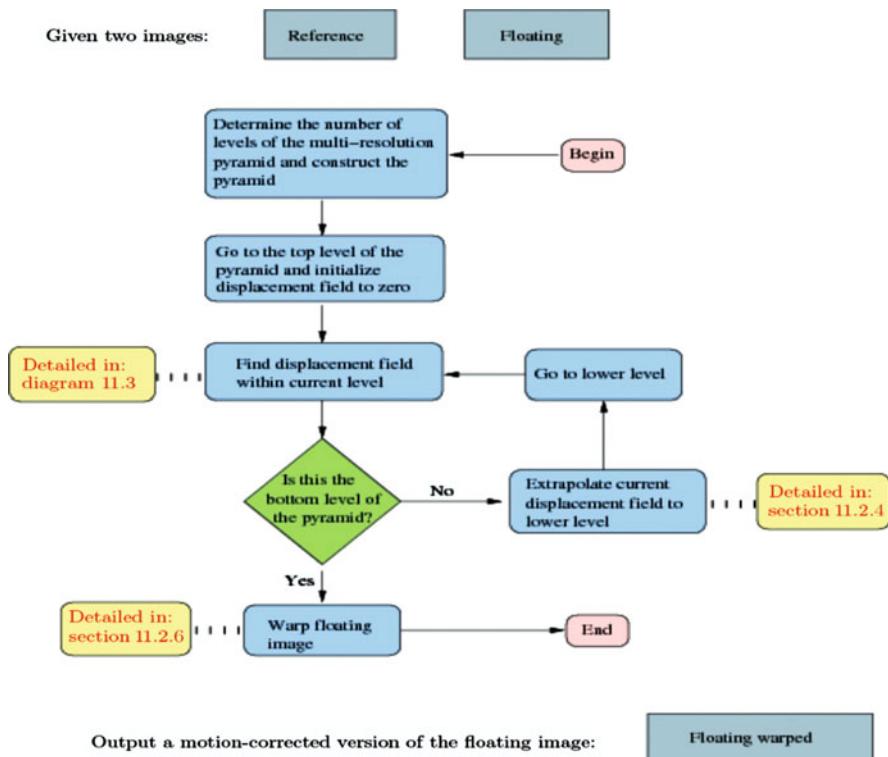
the output, which is the motion corrected image. The overall procedure is described in Diagram 11.2.

### 11.2.3 Construction of the Multi-Resolution Pyramid

Since the number of planes is small with respect to the number of columns and rows of each plane, only the size of the planes is reduced in the construction of the multi-resolution pyramid. This is performed using a low-pass filter specially designed for reducing in half the sampling frequency with minimal loss of information.

### 11.2.4 Extrapolating Results from Low Resolution Levels

Given the displacement field computed at a given resolution, trilinear interpolation is used to produce an initial displacement field at a higher resolution. If the resolution has doubled, all displacements are multiplied by two accordingly.



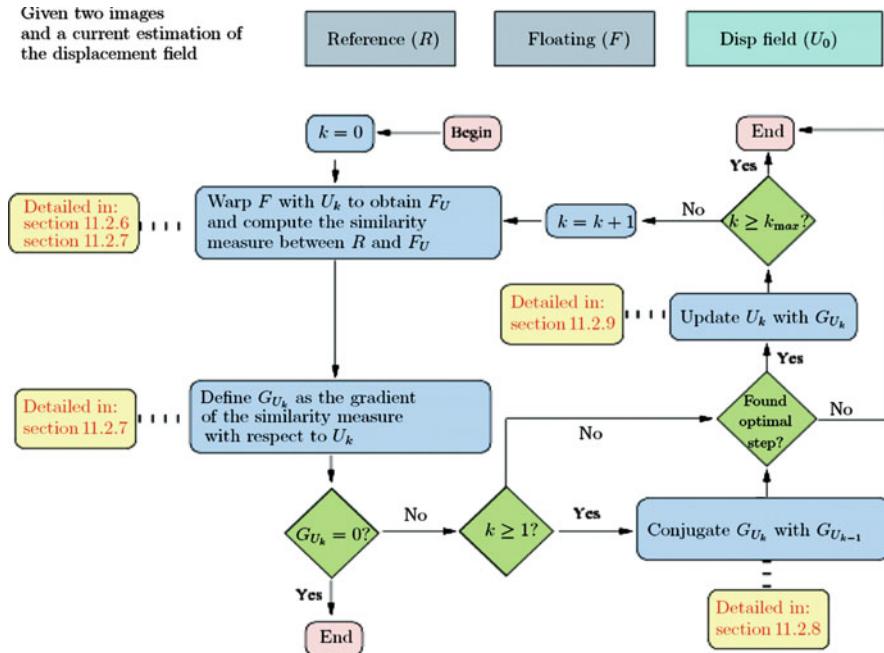
**Diagram 11.2** Multi-resolution motion correction procedural flow

### 11.2.5 Single Resolution Motion Estimation

The estimation of the compensating deformation between a reference and a floating image within a given resolution level is the main component of the system. It is done by performing a conjugate gradient maximization of a similarity measure between the two images, based on the local cross-correlation of corresponding regions around each point. The non-singularity of the deformation is ensured by a special composition technique when building the solution. The flow diagram of the algorithm is depicted in Diagram 11.3. The following subsections provide details on specific modules and operations used by the algorithm.

### 11.2.6 Warping

Warping refers to the application of the deformation applied to the current estimate of the displacement field to the deforming image. In essence, denoting  $I_2$ , the moving image and  $\text{id}$  the identity mapping, it is the process for defining  $I'_2$ :



**Diagram 11.3** Estimation of the displacement field between two images

$$I'_2 \equiv I_2 \circ \phi = I_2 \circ (\text{id} + U_k)$$

at all voxel locations of the fixed image  $I_1$ , using trilinear interpolation to evaluate  $I_2$  at locations not corresponding to a voxel.

### 11.2.7 Computation of the Local Cross-Correlation and Its Gradient

Given a current estimate of a displacement field  $\mathbf{h}(x)$ , a deforming image  $I_2$  and a fixed or reference image  $I_1$ , we can compute the local cross-correlation similarity measure  $J_{CC}$  from an estimate of their respective local mean functions  $\mu_1$  and  $\mu_2$ , their respective local variance estimate functions  $v_1$  and  $v_2$ , as well as the estimate of the local covariance of intensities  $v_{1,2}$  (see [33] for details):

$$J_{CC} = \int_{\Omega} J_{CC}(x) dx = - \int_{\Omega} \frac{v_{1,2}(x)^2}{v_1(x)v_2(x)} dx$$

Its first-order variation is well-defined and defines a gradient given by

$$\nabla_{\mathbf{H}} J_{CC} = -f_{CC}(\mathbf{I}(x), x) \nabla I_2(x + \mathbf{h}(x))$$

Where,  $f_{CC}(\mathbf{i}, x) = G_\gamma * L_{CC}(\mathbf{i}, x)$  and

$$L_{CC}(\mathbf{i}, x) = \frac{v_{1,2}(x)}{v_2(x)} \left( \frac{i_1 - \mu_1(x)}{v_1(x)} \right) + J_{CC}(x) \left( \frac{i_2 - \mu_2(x)}{v_2(x)} \right) \quad (11.1)$$

The function  $L_{CC}$  is estimated as

$$L_{CC}(\mathbf{i}, x) = (G_\gamma * f_1)(x)i_1 + (G_\gamma * f_2)(x)i_2 + (G_\gamma * f_3)(x)$$

where,

$$f_1(x) = \frac{v_{1,2}(x)}{v_1(x)v_2(x)}, \quad f_2(x) = \frac{J_{CC}(x)}{v_2(x)}$$

and

$$f_3(x) = f_1(x)\mu_1(x) + f_2(x)\mu_2(x)$$

All the required space dependent quantities like  $\mu_1(\mathbf{x})$  are computed through recursive spatial smoothing (see Sect. 11.2.10).

### 11.2.8 Conjugate Gradient Optimization

The explicit time discretization using a fixed time step corresponds to a steepest descent method without line search, which is generally quite inefficient. Our system performs line searching and uses a Fletcher-Reeves conjugate gradient minimization routine as described in [16]. The conjugate gradient method allows about one order of magnitude reduction in the total number of iterations required. The gain in speed is much higher since the number of iterations at the finest level is very small, despite the fact that each iteration is slightly more costly.

In essence, the conjugate gradient method changes the direction in which each new estimate is incremented. The standard gradient ascent method defines a sequence of gradient directions at successive “points”, depending on the incremental step size. If  $g_0, \dots, g_\infty$  denotes this sequence, the conjugate gradient defines another sequence  $\tilde{g}_0, \dots, \tilde{g}_\infty$  that takes into account the gradients at the current and previous points, as follows:

$$\begin{cases} \tilde{g}_0 = g_0 \\ \tilde{g}_k = g_k + \frac{|g_k|^2}{|\tilde{g}_{k-1}|^2} \tilde{g}_{k-1} \quad \text{for } k = 1, \dots, \infty \end{cases} \quad (11.2)$$

The algorithm would of course stop in the unlikely event that  $\tilde{g}_k = 0$  for some  $k$ .

### 11.2.9 Updating Procedure

For clarity, we assume in this section a standard gradient ascent method with fixed step. The effect of the conjugate gradient method of Sect. 11.2.8 is only the modification of the updating “direction” as described in (11.2). Given a current estimate of the displacement field  $U_k$  and the gradient of the similarity measure with respect to it,  $G_{U_k}$ , a new estimate is obtained by defining:

$$\phi_k = \text{id} + U_k$$

and performing a small “update” in the direction  $G_{U_k}$ :

$$\phi_{k+1} = \phi_k \circ [\text{id} + \varepsilon G_{U_k} * G_\sigma]$$

where,  $\varepsilon$  is a sufficiently small scalar value and “ $* G_\sigma$ ” denotes convolution with a Gaussian kernel. This is done using trilinear vector interpolation as implementation of the composition operator “ $\circ$ ”. This scheme ensures the invertibility of  $\phi$  at all steps. We analyze in more detail this method of producing diffeomorphisms in Sect. 11.3.

### 11.2.10 Convolutions

The convolutions by a Gaussian kernel are approximated by recursive filtering using the smoothing operator introduced in [41]. Given a discrete 1D input sequence  $x(n), n = 1, \dots, M$ , its convolution by the smoothing operator  $S_\alpha(n) = k(\alpha|n| + 1)e^{-\alpha|n|}$  is calculated efficiently as:

$$y(n) = (S_\alpha * x)(n) = y_1(n) + y_2(n)$$

where

$$\begin{cases} y_1(n) = k(x(n) + e^{-\alpha}(\alpha - 1)x(n-1)) + 2e^{-\alpha}y_1(n-1) - e^{-2\alpha}y_1(n-2) \\ y_2(n) = k(e^{-\alpha}(\alpha + 1)x(n+1) - e^{-2\alpha}x(n+2)) + 2e^{-\alpha}y_2(n+1) - e^{-2\alpha}y_2(n+2) \end{cases}$$

The normalization constant  $k$  is chosen by requiring that  $\int_{\mathbb{R}} S_\alpha(t) dt = 1$ , which yields  $k = \alpha/4$ . This scheme is very efficient since the number of operations required is independent of the smoothing parameter  $\alpha$ . The smoothing filter can be readily generalized to  $n$  dimensions by defining the separable filter  $T_\alpha(x) = \prod_{i=1}^n S_\alpha(x_i)$ .

### 11.3 Building Diffeomorphisms

At each iteration of the process described in Diagram 11.3, a gradient is computed and used to obtain a new estimate of the deformation using the updating procedure described in Sect. 11.2.9. This procedure guarantees that the deformation being constructed remains invertible at all times, provided that each of the steps  $[\text{id} + \varepsilon G_{U_k} * G_\sigma]$  is itself invertible. A good way of ensuring this property is to make  $\varepsilon$  sufficiently small. The smoothing with a Gaussian kernel ensures that the step is differentiable, and making  $\varepsilon$  sufficiently small ensures its invertibility. In this section, we present an alternative updating scheme that gives a more efficient algorithm, and also justify the use of the smoothing operation to build a proper gradient under a global deformation model.

#### 11.3.1 Using a Fixed Vector Map During the Optimization Process

The gradient-based method for registration can be summarized by the following algorithm, starting with a reference image  $I_1$ , a deformed image  $I_2$ , and an initial deformation  $\phi_0$ :

$$\begin{cases} \phi^* = \phi_0 \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_\infty \\ \phi_k = \text{id} + \varepsilon \frac{\partial}{\partial \phi} S(I_1, I_2 \circ \phi) \end{cases} \Big|_{\phi=\phi_0 \circ \phi_1 \circ \phi_2 \circ \cdots \circ \phi_{k-1}} \quad (11.3)$$

where  $\varepsilon > 0$  is small enough to guarantee the invertibility of  $\phi_k$ ,  $\text{id}$  denotes the identity mapping, and  $\phi_\infty$  represents  $\phi_n$  for some  $n$  large enough to attain convergence criteria. This algorithm can also be stated using pseudocode as follows:

- Start.
- Set  $i = 0$ .
- Set  $\phi^* = \phi_0$ .
- While a stopping criterion is not reached:
  - Apply  $\phi_i$  to  $I_2$  in order to produce  $I_2^i = I_2 \circ \phi_i$ .
  - Compute the similarity function between  $I_1$  and  $I_2^i$ .
  - Compute the gradient of  $S$  with respect to the deformation of  $I_2$  at  $\phi_i$ .
  - Add a fraction of this gradient to  $\text{id}$  in order to form  $\phi_{i+1}$ .
  - Set  $\phi^* = \phi_i \circ \phi_{i+1}$ .
  - Increment  $i$  by one.
- End.

Instead, we propose a slightly different, more efficient algorithm:

$$\left\{ \begin{array}{l} \phi^* = \phi_0 \circ \delta_1^{-1} \circ \delta_2^{-1} \circ \cdots \circ \delta_\infty^{-1} \\ \delta_k = \text{id} + \varepsilon \frac{\partial}{\partial \phi} S(I_1 \circ \delta_1, I_2 \circ \phi_0 \circ \delta_1^{-1} \circ \delta_2^{-1} \circ \cdots \circ \delta_{k-1}^{-1}) \end{array} \right|_{\delta=\text{id}} \quad (11.4)$$

where  $\varepsilon > 0$  is small enough to guarantee the invertibility of  $\delta_k$ ,  $\text{id}$  denotes the identity mapping, and  $\delta_\infty$  represents  $\delta_n$  for some  $n$  large enough to attain convergence criteria. This algorithm can also be stated using pseudocode as follows:

- Start.
- Set  $i = 0$ .
- Set  $\phi^* = \phi_0$ .
- While a stopping criterion is not reached:
  - Apply  $\phi_i$  to  $I_2$  in order to produce  $I_2^i = I_2 \circ \phi_i$ .
  - Compute the similarity function between  $I_1$  and  $I_2^i$ .
  - Compute the gradient of  $S$  with respect to the deformation of  $I_2$  at  $\text{id}$ .
  - Add a fraction of this gradient to  $\text{id}$  in order to form  $\delta_{i+1}$ .
  - Invert  $\delta_{i+1}$  to produce  $\delta_{i+1}^{-1}$ .
  - Set  $\phi^* = \phi_i \circ \delta_{i+1}^{-1}$ .
  - Increment  $i$  by one.
- End.

The advantage of this algorithm is that the gradient of  $S$  with respect to a small deformation of  $I_1$  at  $\text{id}$  involves only estimating the gradient of the fixed reference image  $I_1$ . This vector map is the same during the whole optimization process and therefore a lot of computations are saved in comparison with the original algorithm. Also, the gradient of  $I_1$  is computed only at the sampling grid points. No interpolation is necessary, contrary to the original algorithm. Finally, although it may seem that the additional step of inverting  $\delta_{i+1}$  could increase the computational cost, in fact,  $\delta_{i+1}$  is always a very small deformation and therefore the inverse of  $\delta_{i+1}$  is well approximated by  $\text{id} - (\delta_{i+1} - \text{id})$ .

### 11.3.2 A Proper Gradient Using Fast Convolutions

We now establish the link between smoothing of the gradient vector map and the gradient of a global parametric model, giving proper justification to a fast algorithm for computing the similarity gradient using convolutions. Our method applies these convolutions using recursive filters. This property significantly increases the computational speed of the algorithm without disrupting its robustness or precision. We next describe this property more in detail.

### 11.3.2.1 Gaussian Smoothing as Gradient of a Global Model

As mentioned before, there are several options for the similarity function. A commonly used one is the sum of absolute differences between the intensities of the two images, also sometimes called the mean square error. It is defined as:

$$E_u(\phi) = \frac{1}{2} \int_{\mathbf{R}^3} (I_1(x) - I_2(\phi(x)))^2 dx \quad (11.5)$$

However, optimizing  $\phi$  to minimize  $E_u$  is usually not satisfactory because some regularity needs to be enforced in  $\phi$ . This is usually achieved by adding a regularization term to the energy function as follows:

$$E_c(\phi) = \frac{1}{2} \int_{\mathbf{R}^3} (I_1(x) - I_2(\phi(x)))^2 dx + \lambda \int_{\mathbf{R}^3} \psi(D(\phi))^2 dx \quad (11.6)$$

where  $D(\phi)$  is the Jacobian of  $\phi$  and  $\psi$  is a convex function. The parameter  $\lambda$  is most of the time problem-dependent and manually adjusted. In some other methods, regularization is imposed more empirically by explicitly smoothing the gradient of the unconstrained similarity measure  $E_u$  at each step of the minimization process. We show that the latter strategy is equivalent to having an explicit model for the deformation as a function  $\alpha : \mathbf{R}^3 \rightarrow \mathbf{R}^3$  convolved with a smoothing kernel. To fix the ideas, we assume the smoothing kernel to be Gaussian. Indeed, assume the unknown of the problem to be  $\alpha$  and define the following functional:

$$E_g(\alpha) = \frac{1}{2} \int_{\mathbf{R}^3} (I_1(x) - I_2(\phi(x)))^2 dx \quad (11.7)$$

with

$$\phi(x) = \int_{\mathbf{R}^3} \alpha(y) G_\sigma(x - y) dy \quad \text{and} \quad G_\sigma(x - y) \equiv \exp\left(-\frac{|x - y|^2}{2\sigma}\right) \quad (11.8)$$

To do a gradient descent on  $E_g$ , one must find a function  $\nabla E_g : \mathbf{R}^3 \rightarrow \mathbf{R}^3$  satisfying:

$$\lim_{\varepsilon \rightarrow 0} \frac{E_g(\alpha + \varepsilon\tau) - E_g(\alpha)}{\varepsilon} = \int_{\mathbf{R}^3} \nabla E_g(x) \cdot \tau(x) dx. \quad \text{for any } \alpha\tau : \mathbf{R}^3 \rightarrow \mathbf{R}^3 \quad (11.9)$$

we have

$$\lim_{\varepsilon \rightarrow 0} \frac{E_g(\alpha + \varepsilon\tau) - E_g(\alpha)}{\varepsilon} \equiv \left. \frac{dE_g(\alpha + h\tau)}{dh} \right|_{h=0} \quad (11.10)$$

$$E_g(\alpha + h\tau) = \frac{1}{2} \int_{\mathbf{R}^3} \left[ I_1(x) - I_2 \left( \int_{\mathbf{R}^3} (\alpha(y) + h\tau(y)) G_\sigma(x-y) dy \right) \right]^2 dx \quad (11.11)$$

so that

$$\begin{aligned} \frac{dE_g(\alpha + h\tau)}{dh} \Big|_{h=0} &= \int_{\mathbf{R}^3} [I_1(x) - I_2(\phi(x))] \nabla I_2(\phi(x)) \cdot \left( \int_{\mathbf{R}^3} \tau(y) G_\sigma(x-y) dy \right) dx \\ &= \int_{\mathbf{R}^3} \left( \int_{\mathbf{R}^3} [I_1(x) - I_2(\phi(x))] \nabla I_2(\phi(x)) G_\sigma(x-y) dx \right) \cdot \tau(y) dy \end{aligned}$$

Therefore a gradient operator can be defined as:

$$\nabla E_g(y) = \int_{\mathbf{R}^3} ([I_1(x) - I_2(\phi(x))] \nabla I_2(\phi(x))) G_\sigma(x-y) dx \quad (11.12)$$

We see that the gradient obtained with the constrained model is nothing but the convolution with a Gaussian kernel of the gradient obtained with the unconstrained model.

## 11.4 Polynomial Intensity Correction for Multimodal Image Registration

A fast and robust multimodal image registration algorithm can significantly enhance a clinician's ability to take advantage of multiple sources of information. However, the complex relationship in intensity values for images obtained from different types of sensors presents a serious challenge for registration methods [42]. To avoid modeling the complex relationship between intensity values, feature-based registration algorithms extract and match corresponding anatomical structures as landmarks [43–45]. Nevertheless, their performance is highly sensitive to the choice of features. On the other hand, intensity-based algorithms take into account intensity variations of corresponding anatomical structures in different imaging modalities. However, they are often computationally intensive. As part of this work, we developed a fast and robust intensity-based registration algorithm.

The most popular intensity-based algorithms for multimodal registration are based on the correlation coefficient (CC) [46], correlation ratio (CR) [47], and mutual information (MI) [48, 49]. CC assumes a linear relationship between the intensity values of two images. CR relaxes this linear relationship to any forms of functional relationship. MI removes the functional assumption, and uses information theoretical measures to evaluate the level of statistical dependency between intensity values.

Strengthening CR's assumption, Guimond et al. [50] explicitly incorporated a functional intensity correction between a pair of images in the registration process. Their algorithm alternates between estimating the intensity correction parameters and minimizing the energy function with respect to the transformation parameters. In fact, the assumption of a functional dependence between the intensity values of multimodal images was first discussed by Feldmar et al. [51].

Based on this idea, we model the relationship of the intensity values using a low-order polynomial instead of two polynomials as in [50]. We will refer to the proposed algorithm as the polynomial-intensity-corrected sum of squared differences algorithm (PIC-SSD). Although the proposed functional dependency is not valid for every pair of modalities, our experiments demonstrate that the algorithm can efficiently register intrasubject T1 and T2 brain images, positron emission tomography (PET) to T1, and PET to T2.

PIC-SSD gives rise to an energy function that can be written as a sum of squares, enabling the use of the efficient Gauss-Newton method. In contrast to the gradient descent method as employed in [50], the Gauss-Newton method provides an approximation of the Hessian matrix to speedup the rate of convergence in the optimization process. For example, PIC-SSD registers a pair of  $181 \times 217 \times 181$  T1-T2 images with affine transformation in less than half of a second with a standard PC, when they are initially 15-degree rotation apart. Furthermore, the local curvature information computed in each Gauss-Newton iteration reduces the sensitivity of the algorithm to the parameterization of the transformation space. The Gauss-Newton method has also been used in [52] to minimize the local cross-correlation criterion for multimodal image registration.

Due to the special form of the polynomial intensity correction model, the Gauss-Newton method can be applied in two different ways. First, we can directly optimize the energy function with respect to both the transformation and the intensity correction parameters. Alternatively, we can express the intensity correction parameters via the linear least squares estimates, and then employ the Gauss-Newton method to optimize a nonlinear least squares energy function solely with respect to the transformation parameters. We will refer to the first approach as joint optimization and the second approach as constrained optimization.

In the following, we first present the model underlying PIC-SSD and two methods to optimize the corresponding energy functions. We then present the experimental comparisons among PIC-SSD, CC, and MI using simulated data. Finally, we demonstrate an application of PIC-SSD to MR-PET registration.

## 11.5 Methods

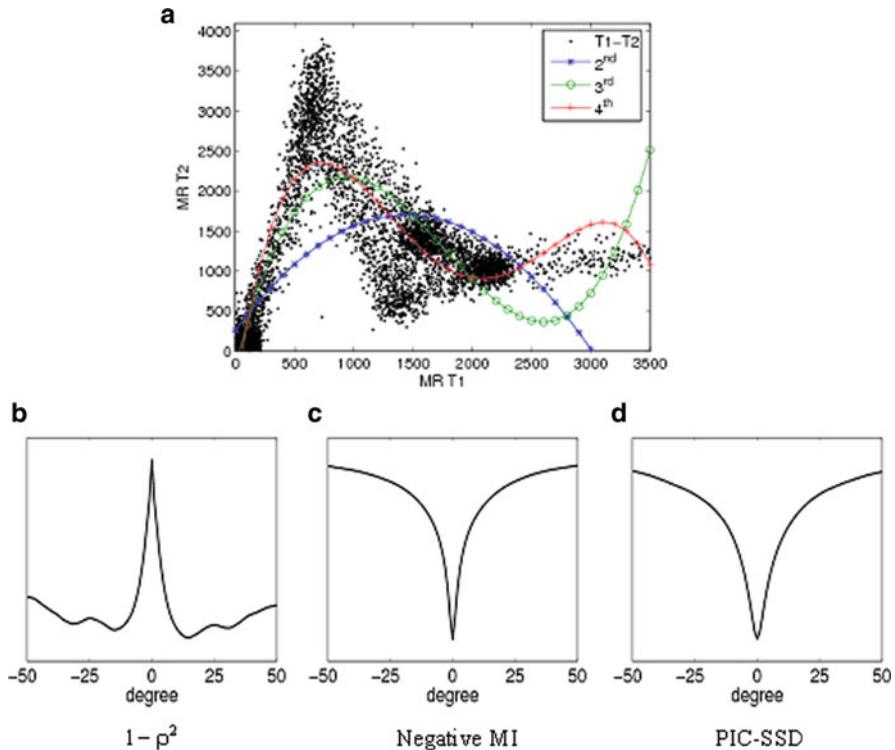
One typically uses the SSD criterion under the assumption that, when two images are aligned, the intensity difference between them is small. However, this assumption is not valid for images from different modalities. Assuming a functional relationship  $\Phi$  between the intensity values of the reference and moving images,

we modify the SSD formulation by incorporating intensity correction for the reference image. Hence, the energy function becomes

$$E(\mathbf{a}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^N (\Phi(f(x_i); \mathbf{a}) - g(T(x_i; \mathbf{p})))^2 \quad (11.13)$$

where  $x_i$  and  $f(x_i)$  denote the location and the reference image intensity of voxel  $i$ , respectively.  $g(T(x_i; \mathbf{p}))$  is the intensity value of voxel  $i$  in the moving image with transformation  $T$  applied. There are two sets of parameters in the energy function: the transformation parameters,  $\mathbf{p}$ , and the intensity correction parameters,  $\mathbf{a}$ . In this chapter, we focus on intrasubject brain registration, and  $T$  is chosen to be affine or rigid transformation.

In general, this functional relationship is difficult to describe, or may not even exist. Our experiments using simulated MR images show that the intensity values of a T2 image can be reasonably approximated by a third-order polynomial of the intensity values of the corresponding T1 image (Fig. 11.1a). In the following, we model  $\Phi$  as an  $m$ th-order polynomial. Because of the special form of the polynomial, we can write (11.13) in matrix notation:



**Fig. 11.1** Joint intensity histogram of MR T1 and T2 images (a), and energy profiles of CC (b), negative MI (c), and PIC-SSD (d)

$$E(\mathbf{a}, \mathbf{p}) = \frac{1}{2} \|\mathbf{W}\mathbf{a} - \mathbf{y}(\mathbf{p})\|^2 \quad (11.14)$$

where the  $\{i,j\}$  element of  $\mathbf{W}$ ,  $w_{ij} = f^{j-1}(x_i)$  is the  $(j-1)$  power of the intensity value in the reference image. Vector  $\mathbf{a} = [a_0, a_1, \dots, a_m]$  is the set of polynomial coefficients, and vector

$$\mathbf{y}(\mathbf{p}) = [g(T(x_1; \mathbf{p})), g(T(x_2; \mathbf{p})), \dots, g(T(x_N; \mathbf{p}))]$$

is the stack of voxel intensity values of the moving image with transformation  $T$  applied.

### 11.5.1 Joint Optimization

Since the energy function is a sum of squares, it is possible to estimate the optimal parameters through the Gauss-Newton method. To simplify our derivations, we combine the parameters into a single vector  $\mathbf{q} = [\mathbf{p}^T, \mathbf{a}^T]^T$ , so that the energy function becomes

$$E(\mathbf{q}) = \frac{1}{2} \|\tilde{\mathbf{W}}\mathbf{q} - \tilde{\mathbf{y}}(\mathbf{q})\|^2 \quad (11.15)$$

where  $\tilde{\mathbf{W}} = [0, \dots, 0, \mathbf{W}]$ . For instance,  $\tilde{\mathbf{W}}$  begins with six and twelve columns of 0's in 2D and 3D affine cases, respectively.  $\tilde{T}$  is a  $K \times K$  matrix, where  $K$  is the number of elements in  $\mathbf{q}$ . The upper left submatrix of  $\tilde{T}$  is equal to  $T$ ; all the remaining elements equal zero. Hence, vector

$$\tilde{\mathbf{y}}(\mathbf{q}) = \left[ g(\tilde{T}(x_1; \mathbf{q})), g(\tilde{T}(x_2; \mathbf{q})), \dots, g(\tilde{T}(x_N; \mathbf{q})) \right]$$

still contains the moving image intensities.

The Gauss-Newton method [53] is based on the Taylor expansion of  $\mathbf{q} = \mathbf{q}_0 + \delta_{\mathbf{q}}$ , where  $\mathbf{q}_0$  is the current estimate of the parameters and  $\delta_{\mathbf{q}}$  is a perturbation to  $\mathbf{q}_0$ :

$$E(\mathbf{q}) \approx \frac{1}{2} \left\| E(\mathbf{q}_0) + \frac{\partial E(\mathbf{q}_0)}{\partial \mathbf{q}} \delta_{\mathbf{q}} \right\|^2 \quad (11.16)$$

where  $\partial E(\mathbf{q}_0)/\partial \mathbf{q}$  is the value of  $\partial E/\partial \mathbf{q}$  at  $\mathbf{q}_0$ . Hence, the optimal perturbation for reducing the objective function can be obtained by setting  $dE/d\delta_{\mathbf{q}} = 0$  and we get

$$\delta_{\mathbf{q}} = -\hat{H}^{-1} \left[ \tilde{\mathbf{W}} - \frac{\partial \tilde{\mathbf{y}}(\mathbf{q}_0)}{\partial \mathbf{q}} \right] (\tilde{\mathbf{W}}\mathbf{q}_0 - \tilde{\mathbf{y}}(\mathbf{q}_0))$$

where

$$\hat{H} = \left[ \tilde{\mathbf{W}} - \frac{\partial \tilde{\mathbf{y}}(\mathbf{q}_0)}{\partial \mathbf{q}} \right]^T \left[ \tilde{\mathbf{W}} - \frac{\partial \tilde{\mathbf{y}}(\mathbf{q}_0)}{\partial \mathbf{q}} \right]$$

and

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \mathbf{q}} = \left[ \nabla g \frac{\partial \tilde{T}(x_1; \mathbf{q})}{\partial \mathbf{q}}, \nabla g \frac{\partial \tilde{T}(x_2; \mathbf{q})}{\partial \mathbf{q}}, \dots, \nabla g \frac{\partial \tilde{T}(x_N; \mathbf{q})}{\partial \mathbf{q}} \right]$$

The above derivations assume that the intensity correction is applied to the reference image. It is also possible to apply the correction to the moving image. The derivation becomes slightly more complex, but the energy function remains a sum of squares form; hence, the Gauss-Newton method is applicable.

### 11.5.2 Constrained Optimization

Since the energy function in (11.14) is a quadratic function of the intensity correction parameter  $\mathbf{a}$ , it is possible to reduce the search space in the optimization procedure by substituting the least square estimate of  $\mathbf{a}$  into the original energy function:

$$E(\mathbf{p}) = \frac{1}{2} \|(\mathbf{I} - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T) \mathbf{y}(\mathbf{p})\|^2 = \frac{1}{2} \|R \mathbf{y}(\mathbf{p})\|^2 \quad (11.17)$$

where  $\mathbf{I}$  is the  $N \times N$  identity matrix. Matrix  $R = (\mathbf{I} - \mathbf{W}(\mathbf{W}^T \mathbf{W})^{-1} \mathbf{W}^T)$  is symmetric and singular.  $R$  is also idempotent, i.e.,  $RR = R$ . We can again use the Gauss-Newton method to obtain the update step,  $\delta_{\mathbf{p}}$ .

$$\begin{aligned} \delta_{\mathbf{p}} &= -\hat{H}^{-1} \mathbf{y}^T R \left[ \frac{\partial \tilde{\mathbf{y}}(\mathbf{p}_0)}{\partial \mathbf{p}} \right] \\ \hat{H} &= \left[ \frac{\partial \tilde{\mathbf{y}}(\mathbf{p}_0)}{\partial \mathbf{p}} \right]^T R^T R \left[ \frac{\partial \tilde{\mathbf{y}}(\mathbf{p}_0)}{\partial \mathbf{p}} \right] = \left[ \frac{\partial \tilde{\mathbf{y}}(\mathbf{p}_0)}{\partial \mathbf{p}} \right]^T R \left[ \frac{\partial \tilde{\mathbf{y}}(\mathbf{p}_0)}{\partial \mathbf{p}} \right] \\ \frac{\partial \mathbf{y}}{\partial \mathbf{p}} &= \left[ \nabla g \frac{\partial T(x_1; \mathbf{p})}{\partial \mathbf{p}}, \nabla g \frac{\partial T(x_2; \mathbf{p})}{\partial \mathbf{p}}, \dots, \nabla g \frac{\partial T(x_N; \mathbf{p})}{\partial \mathbf{p}} \right] \end{aligned}$$

The above derivations are general for any order polynomials in the intensity correction model  $\Phi$ . If we restrict  $\Phi$  to be a first-order polynomial, the energy in (11.17) reduces to a monotonic function of the correlation coefficient  $\rho^2$ :

$$E(\mathbf{p}) = \frac{N}{2}(1 - \rho^2) \text{Var}(Y) \quad (11.18)$$

where each element in  $\mathbf{y}$  is a sample of the random variable  $Y$ . Since  $\text{Var}(Y)$  is fixed for a given image, minimizing (11.18) is equivalent to maximizing the correlation coefficient between the two images. In a more general case,  $\Phi$  is equal to the conditional expectation of the moving image intensity values given the reference image intensity values. The resulting energy is a monotonically decreasing function of the correlation ratio  $\eta^2$ :

$$E(\mathbf{p}) = \frac{N}{2}(1 - \eta^2) \text{Var}(Y)$$

Therefore, the PIC-SSD algorithm makes a weaker assumption compared to CC, but a stronger assumption compared to CR.

### 11.5.3 Joint Optimization vs. Constrained Optimization

The optimal estimate for the intensity correction parameters  $\mathbf{a}$  must be the least squares estimate for any given transformation parameters  $\mathbf{p}$ , including  $\mathbf{p}^*$  for which  $E$  reaches a global minimum. Thus, a global minimum obtained by the joint approach,  $[\mathbf{p}^*, \mathbf{a}^*]$ , must lie inside the search space of the constrained approach. Theoretically, the constrained approach is better because of a smaller search space, and our experiments show that this approach is more robust in a high noise setting. However, the constrained approach is only feasible for certain forms of  $\Phi$ , such as the polynomial model.

In general, the energy function is not convex and may contain local minima. The constrained approach avoids local minima that lie outside of its search space. Therefore, it encounters no more local minima than the joint approach. When local minima are present in the constrained search space, either approach might converge to a local minimum.

In terms of computation time, the joint approach requires a constant factor of operations fewer than the constrained approach for each iteration. Since the number of iterations for either approach before convergence is problem specific, the joint approach is not always faster than the constrained approach. Nevertheless, as the number of samples  $N$  becomes large, the joint approach often converges faster than the constrained approach.

### 11.5.4 Results

In this section, we first examine PIC-SSD using simulated MR T1 and T2 images. Then, we illustrate the application of our method for MR-PET registration.

### 11.5.4.1 MR T1 and T2 Registration

MR T1 and T2 images were generated from an identical template using the Brain Web software [54] ( $181 \times 217 \times 181$ ,  $1 \times 1 \times 1 \text{ mm}^3$ , 1% noise, 20% inhomogeneity). We set the T1 image as reference and the T2 data as moving image.

*Joint intensity histogram study.* The joint intensity histogram (Fig. 11.1a) of the T1 and T2 images shows that it is possible to model the T2 image intensity values as a third-order polynomial of the T1 image intensity values. The joint intensity histogram suggests a more suitable intensity correction model could be a piecewise linear function; we are currently exploring this direction.

*Capture range.* We compared energy profiles for CC, MI, and PIC-SSD, created by applying various amounts of initial rotation along the  $z$ -axis<sup>1</sup> to the moving image. As shown in Fig. 11.1b–d, MI and PICSSD have similar energy profiles, covering up to  $50^\circ$ . However, in the profile for CC, a minimum for  $1 - \rho^2$  does not appear when two images are aligned. This reveals that a simple linear relationship fails to describe the intensity dependency between T1 and T2 images. We tested both the joint and the constrained methods. The constrained method appears to have a slightly larger capture range, and it is more robust to noise present in the data compared to the joint approach. Additional experiments are required for a thorough validation. On the other hand, the joint approach takes 0.4–3 s, depending on the initial displacement; the constrained method takes twice that time.

### 11.5.4.2 PET-MR Registration

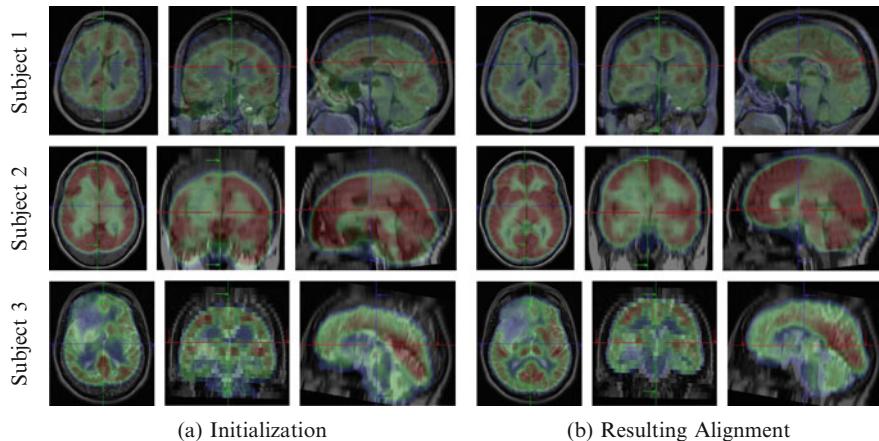
This section presents PIC-SSD registration results using PET and MR scans from three patients. The anatomical images were acquired with a 1.5T Siemens SONATA machine (T1-weighted,  $512 \times 448 \times 25$ ,  $0.45 \times 0.45 \times 6 \text{ mm}^3$  and T2-weighted,  $256 \times 256 \times 96$ ,  $0.98 \times 0.98 \times 2 \text{ mm}^3$ ). The PET images were acquired using a Siemens ECAT machine ( $128 \times 128 \times 47$ ,  $2 \times 2 \times 2.38 \text{ mm}^3$ ).

We chose rigid transformation to perform the intra-subject PET-MR registration. We treated the high resolution MR images as reference images, and low resolution PET images as moving images. No preprocessing was applied to the data. Fig. 11.2a presents the initial position of the MR and PET images with center alignment; Fig. 11.2b illustrates the registration results. We obtained similar registration results using the joint and the constrained optimization approaches. The results from the joint approach are presented here.

In Fig. 11.2, we can see a significant displacement gap between the reference and the moving images in the initialization. PIC-SSD significantly improves the

---

<sup>1</sup> $x$ ,  $y$ , and  $z$  directions are defined as directions from posterior to anterior, from left to right, and from inferior to superior, respectively.



**Fig. 11.2** Overlayed MR and PET images before (a) and after (b) registration. Both T1 (top two rows) and T2 (bottom row) MR images are used in this experiment. A grayscale is used for the MR data and selected colors for PET scan

alignment quality. Due to geometric distortions specific to different modalities, there still exists 0.5–1 mm displacement along the edge of the cortical surface between the registered images. Nevertheless, by incorporating the polynomial intensity correction model, PIC-SSD achieves satisfactory registration results, with run times between 3 and 15 s with a 3.2 GHz (1 GB RAM) machine.

### 11.5.5 Conclusions

Under the assumption of a functional relationship between the intensity values of two modalities, we can cast the image registration problem as a minimization of a nonlinear least squares criterion. We explored the joint and the constrained approaches for optimizing the PIC-SSD energy function. Both approaches benefit from the efficient Gauss-Newton method. The T1–T2 and MR-PET brain registration experiments demonstrate the potential of this algorithm. The fast convergence and simplicity of the optimization scheme (no time-step selection and robustness to changes in the parameterization of the problem) make this algorithm highly practical.

## 11.6 Groupwise Motion Correction of Brain Perfusion Images

Registration of a group of images has traditionally been tackled by repeatedly applying a pairwise registration [55, 56]. Over the past few years, increasing amount of research has been aimed on atlas construction using a groupwise

registration for its advantage of unbiased information representation [57–59]. A similar technique was also used for color image demosaicing [60]. In this work, we performed a groupwise motion correction of MR brain perfusion images. This intrasubject registration has its unique challenges due to high accuracy and reliability requirements. The groupwise registration evaluates all the image information in the group as a whole and optimizes the registration parameters simultaneously for the group, and is thus more likely to achieve a higher accuracy and robustness with computational efficiency.

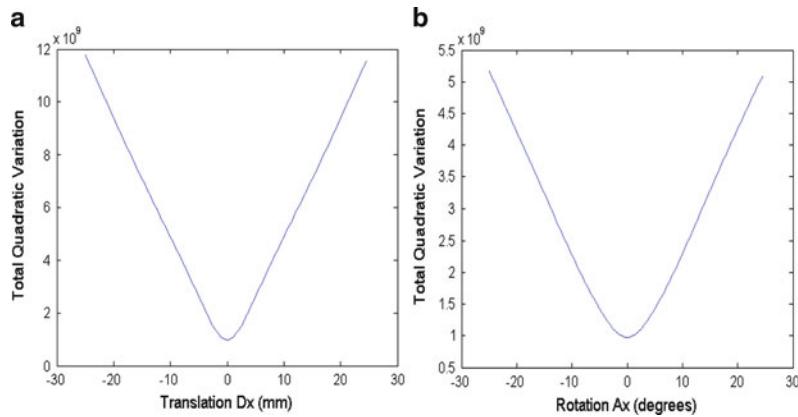
In a brain perfusion study, the intensity profile for a fixed location in the image domain would be much more smooth over the time course if all the images in the series are well aligned. One way to measure the smoothness of an intensity profile is to calculate the voxel-wise total variation, which is defined as the sum of the image intensity difference between two adjacent time points. In our work, the total quadratic variation is used so that the cost function optimization becomes a nonlinear least squares fitting problem and can be solved efficiently by the Levenberg–Marquardt method, a variation of the Gauss-Newton method. A rigid body registration was used in this work because misalignments in brain perfusion images are only caused by patient motion with no anatomy deformation. In the following sections, we first present the cost function of the total quadratic variation, and then describe the Gauss-Newton equations and the Hessian matrix modification of the Levenberg–Marquardt method. Finally, we demonstrate the efficacy of the groupwise registration by two series of MR brain perfusion images.

## 11.6.1 Methods

Like a pairwise registration problem, there are two main components for a groupwise registration: cost function and optimization. The cost function in our work is derived from the total quadratic variation of voxel-wise intensity and evaluated using tricubic interpolation; this cost function is optimized over registration parameters by the Levenberg–Marquardt method in a multi-resolution pyramid.

### 11.6.1.1 Total quadratic variation

For a perfusion study, if we model a voxel-wise intensity profile over time by a real valued 1D function  $f(t)$ , defined on an interval  $[0, t_n]$ , its smoothness can be measured by the total variation  $\int_0^{t_n} f'(t)dt$ , assuming  $f(t)$  is continuously differentiable. The total variation is not differentiable at  $f'(t) = 0$  for numerical stability, and more importantly, to use the efficient Gauss-Newton style optimization (see Sect. 11.5.1), the quadratic variation  $\int_0^{t_n} (f'(t))^2 dt$  is used as the cost function in the groupwise registration. In Fig. 11.3, the cost function was plotted by translating an image in a group of 50 images with known transformations (the



**Fig. 11.3** The cost function profile of the total quadratic variation for a group of 50 perfusion brain scans. (a) translation  $D_x$  along the X axis for one volume; (b) rotation  $A_x$  about the X axis for one volume.  $D_x = 0$  and  $A_x = 0$  indicate the ground truth transformation, which is the minimum of the total quadratic variation

ground truths) along the  $X$  axis by  $D_x$ ; the image is also rotated about the  $X$  axis by  $A_x$ . The smooth profiles shown in Fig. 11.3 demonstrate the agreeable optimization behavior of the total quadratic variation.

Let  $X$  denote an image domain, then a random sample can be expressed as  $\mathbf{x}_m \in X, m = 1, \dots, M$  where  $M$  is the total number of samples in the image domain. Let  $I_n, n = 1, \dots, N$  denote the  $n$ th image in a group, where  $N$  is the total number of images in the group. In a registration process,  $I_n$  is deformed under a rigid transformation  $T_n$  and the total quadratic variation under  $T_n$  is

$$E = \sum_{n=1}^N \sum_{m=1}^M [I_n(T_n(\mathbf{x}_m)) - I_{n'}(T_{n'}(\mathbf{x}_m))]^2$$

where  $n' = n - 1$  or  $n' = N$  for  $n = 1$ . Note that the indexes defined for  $E$  indicate a looped version of the total quadratic variation with an additional term  $[I_1(T_1(\mathbf{x}_m)) - I_N(T_N(\mathbf{x}_m))]^2$  in the summation so that all the images in the group would contribute equally to the cost function. Without this additional term, the first and last image are only evaluated once in the cost function while the other images are evaluated twice, resulting in less accurate transformations for the first and last image.

Let  $Z(x)$  denote the matrix representation of the intensity difference between images, then we have  $Z(x) = [Z_2^T, Z_3^T, \dots, Z_N^T, Z_1^T]^T$ , where  $Z_n$  is an  $M$ -element vector indexed by  $m$ . An element in  $Z_n$  with index  $\{m\}$  is the intensity difference between image  $I_n$  and  $I_{n'}$  at location  $x_m$ , that is,

$$Z_n\{m\} = I_n(T_n(x_m)) - I_{n'}(T_{n'}(x_m))$$

Thus the total quadratic variation  $E$  can also be given by a matrix representation:

$$E = \frac{1}{2} Z^T(x) Z(x)$$

### 11.6.1.2 Gauss-Newton Optimization

The image difference  $Z(x)$  can be considered as a residual function, then minimizing the total quadratic variation  $E = \frac{1}{2} Z^T(x) Z(x)$  is a least squares problem and can be efficiently solved by the Gauss-Newton method. To find the transformations for a group of images, the total quadratic variation  $E$  is minimized iteratively over registration parameters  $\mathbf{P} = \{P_n, n = 1, \dots, N\}$ , where  $P_n = \{p_{nk}, k = 1, \dots, 6\}$  is the rigid registration parameters for transformation  $T_n$ .

For an iteration  $i$ , the approximate Newton step  $\delta\mathbf{P}$  is given by solving the linear system

$$J_Z^T(x) J_Z(x) \delta\mathbf{P} = -J_Z^T(x) Z(x) \quad (11.19)$$

where  $J_Z^T(x)$  is the transpose of  $J_Z(x)$ , and  $J_Z(x)$  is the Jacobian matrix of the residual function  $Z(x)$  over the registration parameters  $\mathbf{P}$ . The Jacobian matrix  $J_Z(x)$  can be viewed as a  $N \times N$  matrix with each element  $[\partial Z_{nm}/\partial p_{n'k}]$  being a  $M \times 6$  submatrix indexed by  $m$  and  $k$ . When  $n = n'$ , an element  $\{m, k\}$  in  $[\partial Z_{nm}/\partial p_{n'k}]$  is  $\frac{\partial Z_{nm}}{\partial p_{nk}} = \nabla I_n(T_n(x_m)) \frac{\partial T_n(x_m)}{\partial p_{nk}}$ ; when  $n \neq n'$  we have

$$\frac{\partial Z_{nm}}{\partial p_{n'k}} = \begin{cases} -\frac{\partial Z_{nm}}{\partial p_{n'k}} & \text{for } n' = n - 1 \text{ and } n \geq 2, \\ 0 & \text{or } n' = N \text{ and } n = 1 \\ & \text{for } n' \neq n - 1 \end{cases} \quad (11.20)$$

Using (11.20), the Jacobian matrix  $J_Z(x)$  can be written as a block matrix:

$$J_Z(x) = \begin{bmatrix} -\left[ \frac{\partial Z_{1m}}{\partial p_{1k}} \right] & \left[ \frac{\partial Z_{2m}}{\partial p_{2k}} \right] & \cdots & 0 \\ 0 & -\left[ \frac{\partial Z_{2m}}{\partial p_{2k}} \right] & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \left[ \frac{\partial Z_{Nm}}{\partial p_{Nk}} \right] & 0 & \cdots & -\left[ \frac{\partial Z_{Nm}}{\partial p_{Nk}} \right] \end{bmatrix} \quad (11.21)$$

Note that in some applications, for example, a brain perfusion study, the first volume may be treated as a fixed reference, thus the registration parameters in  $T_1$

will be excluded from the optimizing parameter set  $\mathbf{P}$ . Once the linear system in (11.19) is solved, the registration parameter estimate for next iteration  $i + 1$  can be calculated by  $\mathbf{P}_{i+1} = \mathbf{P}_i + \delta\mathbf{P}$ .

### 11.6.1.3 Levenberg–Marquardt Method

When a Gauss-Newton step is far from the optimal point, the points for quadratic fitting might be linear or nearly linear. In such cases, quadratic fitting would overshoot to a distant nonoptimal point, leading to erroneous results. A damping factor  $\lambda$  is thus added to introduce a gradient factor in order to cushion the quadratic convergence step:

$$(J_Z^T(x)J_Z(x) + \lambda \mathbf{diag}(J_Z^T(x)J_Z(x)))\delta\mathbf{P} = -J_Z^T(x)Z(x) \quad (11.22)$$

The damping factor  $\lambda$  needs to be balanced between the convergence speed and the risk of overshooting. The larger the damping factor, the smaller the convergence step, but lower possibility of overshooting. In our approach,  $\lambda$  is determined adaptively in the optimization process. Let  $\lambda_i$  be the damping factor chosen for iteration  $i$ , then for iteration  $i + 1$ , the damping factor would be chosen among  $\Lambda = \{2^c \lambda_i, c = -1, 0, 1, \dots, c_o\}$  where  $c_o \leq 10$  is a predefined non-negative integer constant. Let  $\delta\mathbf{P}(\lambda)$  be the solution of (11.22) using  $\lambda$  at iteration  $i + 1$ , then  $\delta\mathbf{P}$  for iteration  $i + 1$  is given by

$$\delta\mathbf{P}_{i+1} = \underset{\delta\mathbf{P}(\lambda), \lambda \in \Lambda}{\operatorname{argmin}} [E(\mathbf{P}(\lambda))] \quad (11.23)$$

Equation (11.23) is used only when  $\min_{\lambda \in \Lambda} [E(\mathbf{P}(\lambda))] < E(\mathbf{P}(\lambda_i))$ . Otherwise, the iteration process has reached the convergence and should stop.

## 11.6.2 Experimental Results

To quantify the accuracy of the groupwise registration, we selected a series of 50 MR scans from a real brain perfusion study with no apparent patient motion, but still featuring the changing intensity distribution induced by the injection of contrast agents. Each perfusion scan consisted of 19 slices with slice spacing 6:50 mm, the dimension for each slice was  $128 \times 128$  with pixel spacing 1.80 mm. For each of the three translation parameters  $D_{(x,y,z)}$ , a set of 49 artificial translations was generated in the range of  $[-12.0, 12.0]$  mm with an incremental step of 0.5 mm; similarly, a set of 49 artificial rotations was also generated for each rotation parameter  $A_{(x,y,z)}$  in the range of  $[-12^\circ, 12^\circ]$  with an incremental

**Table 11.1** Error comparison between the groupwise registrations and the pairwise registration

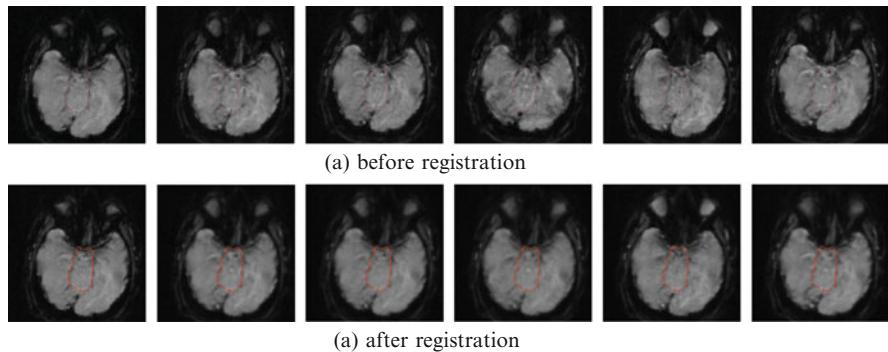
Error $\Delta_p$ (mm)	Groupwise I		Groupwise II		Pairwise	
	Average $\pm$ standard deviation	Maximum	Average $\pm$ standard deviation	Maximum	Average $\pm$ standard deviation	Maximum
$I_{Dx}$	0.55 $\pm$ 0.33	1.61	0.57 $\pm$ 0.29	1.15	0.55 $\pm$ 0.32	1.78
$I_{Dy}$	1.16 $\pm$ 0.65	2.55	1.18 $\pm$ 0.64	2.67	1.22 $\pm$ 0.67	2.46
$I_{Dz}$	1.24 $\pm$ 0.71	2.81	1.27 $\pm$ 0.72	2.66	1.29 $\pm$ 0.68	2.83
$I_{Ax}$	0.46 $\pm$ 0.11	1.42	0.54 $\pm$ 0.12	1.53	0.48 $\pm$ 0.18	1.55
$I_{Ay}$	0.94 $\pm$ 0.51	3.55	0.96 $\pm$ 0.52	3.54	0.99 $\pm$ 0.49	3.39
$I_{Az}$	0.85 $\pm$ 0.47	3.42	0.86 $\pm$ 0.46	3.41	0.91 $\pm$ 0.45	3.58
overall	0.87 $\pm$ 0.58	3.55	0.90 $\pm$ 0.57	3.54	0.9 $\pm$ 0.58	3.58
avg. sec	9.4		1.4		5.5	

step of  $0.5^\circ$ . The ground truth series  $I_{D_{(x,y,z)}}$  and  $I_{A_{(x,y,z)}}$  were obtained by respectively applying the generated six sets of the rigid artificial motions to the scans starting from the second time point, leaving the first time point scan intact as the reference.

To compute the registration error, 400 points were randomly sampled in the region of interest, i.e., the head in our case. The region of interest was segmented by an optimal thresholding method that maximized the inter-class variation. A sample point  $\mathbf{p}$  was first transformed by the ground truth registration matrix  $T_{gt}$  to get  $\mathbf{p}_{gt} = T_{gt}\mathbf{p}$ , then the groupwise registration matrix  $T_{gw}$  was also applied to get  $\mathbf{p}_{gw} = T_{gw}\mathbf{p}$ . The distance  $\Delta_p = |\mathbf{p}_{gw} - \mathbf{p}_{gt}|$  was calculated as errors.

The groupwise registration errors were listed in Table 11.1. Groupwise registration I and II in Table 11.1 differed on the sub-sampling rate. Groupwise registration I sampled on every pixel; Groupwise registration II sampled on every pixel in the lowest resolution image, every two pixels in the second lowest resolution image, and every three pixels in the third lowest and higher resolution images (recall that the registration parameters were optimized via a multi-resolution pyramid). For comparison purpose, Table 11.1 also lists results from a pairwise registration method using the normalized mutual information with tricubic interpolation and the hill climbing optimization. The mutual information metric is an only effective pairwise registration cost function for a perfusion study with changing intensity. The groupwise registrations tackle the changing intensity issue by gradually adding up the total variation of consecutive images over the time course. The computation time given in Table 11.1 was obtained on a 2.66 GHz CPU computer.

The capture range and robustness of the groupwise registration was also validated by a brain perfusion study with large motions. Each perfusion scan in this study consisted of 12 slices with slice spacing 7.8 mm; the dimension for each slice was  $128 \times 128$  with pixel spacing 1.875 mm. The images with a fixed region of interest (ROI) were compared before and after registration (see Fig. 11.4 for example results).



**Fig. 11.4** ROI comparison of an MR brain perfusion study. For each image, ROI is marked by a highlighted polygon intensity issue by gradually adding up the total variation of consecutive images over the time course

### 11.6.3 Discussion

The experimental results showed that the groupwise registration was robust for a large capture range and could get an accuracy at a sub-pixel level. Comparing with the pairwise registration with an average error of 0.905 mm, groupwise registration *I* gave a higher accuracy: average error 0.869 mm; groupwise registration *II* was almost four times faster than the pairwise registration while attaining the same high level of accuracy: average error 0.896 mm.

The Levenberg–Marquardt method used in the groupwise registrations converged very fast – typically in less than five iterations it could reach a highly accurate result. Although for each iteration, the computational effort was high for the groupwise registration due to the large parameter space and costly gradient evaluation, the speed comparison between groupwise registration *II* and the pairwise registration demonstrated the computational efficiency of the groupwise approach.

In a group of images, the existence of one or several low quality images is typical. These low quality images may result in erroneous or less accurate outcomes for many pairwise registration methods, leading to significant measurement errors. A groupwise registration method is more likely to achieve better results in such cases because the contribution from the high quality images in the group can outweigh the negative influence from the few low quality images, as all the image information is evaluated at the same time. More experiments are to be performed to demonstrate this potential capability of a groupwise registration.

While the efficacy of the groupwise registration was demonstrated by MR images in this work, it can also be applied to other modalities. The total variation cost function can be further improved using a polynomial function to model the intensity relationship among the images. In that case, the quadratic residual can be used:  $E = \sum_n \sum_m [I_n(T_n(\mathbf{x}_m)) - f(a, x_m, \mathbf{P}))]^2$ , where  $f(a, x_m, \mathbf{P})$  is an a valued

polynomial that is a least-squares fitting function for voxel-wise intensity at location  $x_m$  under parameters  $\mathbf{P}$ . The total quadratic residual can also be efficiently optimized by the Levenberg–Marquardt method over registration and polynomial parameters using either a joint or constrained optimization [61].

## 11.7 Clinical Evaluation

In this section, we present three studies done in close collaboration with radiologists. The studies were aimed at assessing the performance of registration algorithms and their impact on the radiologist workflow and accuracy. Although the following subsections are primarily concerned with presenting the scope and results of studies, the process of defining and executing clinical evaluations with multiple radiologists at one or more institutions can be exceedingly complex. While a detailed description of the process that enables a successful execution of a study is beyond the scope of this chapter and a subject matter in its own right, it is worthwhile to identify some of the main aspects

- contractual and ethical approvals need to be obtained;
- principal investigator(s), specialist in the domain area, as well as readers, certified radiologists, need to be recruited;
- a study protocol needs to be designed;
- tools that enable the specific evaluation need to be created;
- logistical aspects of scheduling readers, study coordination, and monitoring need to be undertaken;
- tools and processes for gold standard generation (ground truth) need to be built;
- statistical analysis, often FROC, of the collected data needs to be performed.

Additionally, when an evaluation must satisfy regulatory requirements for submission to the FDA, the cost and effort to implement a clinical study may exceed the actual research and development effort combined. This is required whenever tools are to be used in clinical practice.

### 11.7.1 Application and Assessment of a Robust Elastic Motion Correction Algorithm to Dynamic MRI

The purpose of this study was to assess the performance of the motion correction algorithm described in Sect. 11.2, called BRACE. 25 dynamic MR mammography (MRM) data sets and 25 contrast-enhanced 3D peripheral MR angiographic (MRA) data sets which were affected by patient motion of varying severeness were selected retrospectively from routine examinations. Anonymized data were registered by BRACE.

Images before and after correction were visually evaluated and scored by experienced radiologists with respect to reduction of motion, improvement of image quality, disappearance of existing lesions, or creation of artifactual lesions. It was found that the correction improves image quality (76% for MRM and 96% for MRA) and diagnosability (60% for MRM and 96% for MRA).

### 11.7.1.1 Motivation

Dynamic contrast-enhanced (CE) MRI is an important and valuable diagnostic tool, especially for the investigation of vessel and tissue status as is the case in MR mammography (MRM) or MR angiography (MRA). For the analysis of a dynamic CE MR study it is essential that no changes in the patient's position occur during the whole acquisition. However, this requirement is often not met and may thus affect interpretation of the images or even prevent accurate diagnosis [62, 63].

Regardless of the many efforts undertaken to reduce or suppress motion artifacts, there remains quite a substantial number of examinations which suffer from motion induced artifacts. To take into account the highly nonlinear elastic properties of human tissue (e.g., breast, muscle), several elastic deformation image registration algorithms have been published [64–69]. One of the dangers with elastic algorithms, however, is that some parts of the original image might vanish because of overlays or singular distortion fields [70].

An additional problem encountered in CE dynamic MRI is the actual change of the image signal due to the presence of a contrast agent and any useful registration algorithm has to be robust against signal enhancing areas which may be highly localized, but can also be – especially in MRM – multi focal, diffuse, or even cover the whole breast. Simultaneously, the image registration has to preserve the signal changes due to a contrast agent.

## 11.8 Materials and Methods

*Patients:* 25 contrast-enhanced MRM data sets (patients' age range: 33–73 years) and 25 dynamic peripheral MRA data sets (age range: 49–84 years) affected by motion artifacts of varying severeness were retrospectively selected from routine examinations for further analysis. The medically indicated CE MRI examinations were conducted with informed consent by the patients. Selection of the data sets was based on the presence of visible motion artifacts in the images, regardless of the actual diagnosis. Data sets were selected from routine examinations during a time period of 3 months. The percentage of cases with visible motion artifacts was approx. 12% for MRM and 30% for MRA.

*MR Sequences:* Dynamic MR mammography was performed by using a 2D multislice, T1-weighted FLASH sequence on a 1.5T scanner (Magnetom Symphony, Siemens Medical Solutions, Erlangen, Germany) and a double breast coil with the following scan parameters:  $T_R/T_E/\alpha = 113\text{ ms}/4.76\text{ ms}/80^\circ$ , matrix

$384 \times 384$ , FoV = 350 mm, slice thickness 3 mm, inter slice distance 10–30% of slice thickness depending on volume coverage requirements, GRAPPA partial parallel imaging acceleration factor 2, scan time 60 s. The dynamic contrast examination was performed with one native scan followed by a series of seven scans, one each minute, after CA injection.

The contrast-enhanced peripheral angiography data were acquired with a 3D FLASH sequence on a 1.5T scanner (Magnetom Symphony or Sonata, Siemens Medical Solutions, Erlangen, Germany). Typical parameters were  $T_R/T_E/\alpha = 4$  ms/1.5 ms/40, slice thickness 1 mm, matrix of 384–512 and 80–104 partitions, GRAPPA acceleration factor 3 and FoV of up to 500 mm. Two data sets were acquired, one native and one after administration of a contrast agent.

*Evaluation:* To assess the algorithm's performance, the anonymized MR data were analyzed and subtraction images were calculated. Since routine diagnosis of peripheral MRA includes maximum intensity projections (MIP) of the subtracted images, MIPs were calculated for the MRA data.

The typical duration for correcting a complete dynamic breast MRI examination with a matrix of  $384 \times 384$ , 33 slices, and all eight acquired volumes was 1 min. Depending on the number of slices and resolution motion, correction of the MRA data required typically 1–3 min.

All MR images were visually inspected and evaluated retrospectively by two experienced radiologists in consensus. Based on the uncorrected subtraction images, the motion induced artifacts were visually classified into two classes “moderate” and “severe”.

To evaluate MRM data, the uncorrected (UC) and corrected (CC) images were compared at 1, 2, and 7 min after contrast agent administration and rated on a “better – same – worse” scale with respect to improvement of image quality and diagnosability. Careful attention was given to the possible disappearance of existing lesions or creation of artifactual lesions. Together with the uncorrected and corrected subtraction images, the original, unsubtracted images and the T2-weighted images were provided to the radiologist as in routine diagnosis.

For the evaluation of the 3D MRA data, uncorrected and corrected images were compared. The algorithm performance was also rated on a “better – same – worse” scale and careful attention was given to the visibility of tiny vessels.

## 11.9 Results

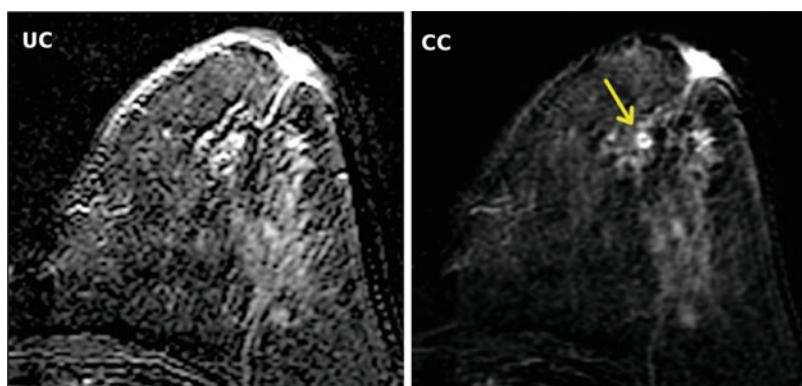
The results of the evaluation are summarized in Table 11.2. For both applications, i.e., MRA and MRM, the algorithm was successful in most cases. Some of the results are described in greater detail below.

*MR mammography* Although many dynamic breast MRI examinations reveal some minimal motion artifacts, most of them do not affect diagnosis. The correction algorithm unavoidably causes light blurring due to the interpolation of the data set and does not improve image quality necessarily for such minimal motion artifacts. Therefore, we did not consider these artifacts in our study.

**Table 11.2** Summary of the radiological evaluation of the corrected MR data sets in comparison to the uncorrected images

	Image quality			Diagnostic usability		
	Better	Same	Worse	Better	Same	Worse
<b>Breast MRI</b>						
Moderate motion	14 (78%)	1 (5.6%)	3 (17%)	11 (61%)	4 (22%)	3 (17%)
Severe motion	5 (72%)	1 (14%)	1 (14%)	4 (57%)	2 (29%)	1 (14%)
<b>Peripheral angiography</b>						
Moderate motion	15 (94%)	0	1 (6%)	15 (94%)	1 (6%)	0
Severe motion	9 (100%)	0	0	9 (100%)	0	0

Image quality was rated based on subjective impression and diagnostic usability was rated on clarity of lesions (MRM) or vessel delineation (MRA). The percent values are calculated separately for MRM and MRA and for each class of motion severeness

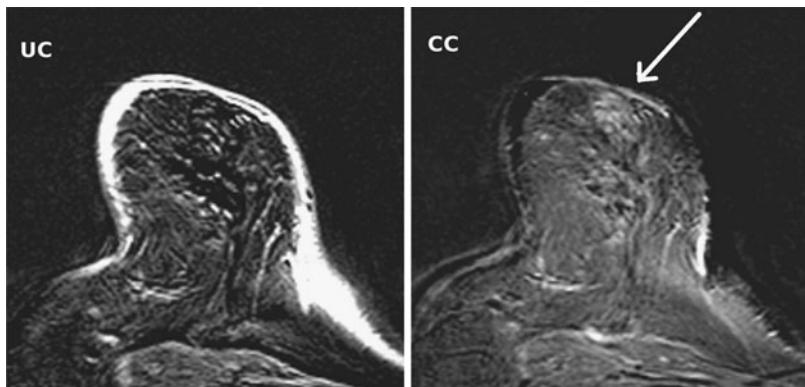


**Fig. 11.5** MR mammography (2D) subtraction images of a female patient, age 41 years. (a) The lesion is very difficult to detect in this uncorrected image among all artifacts. (b) The lesion is clearly visible after correction

For moderate and severe motion artifacts, however, the algorithm improved image quality in most cases (19 better vs. 2 same, 4 worse, see Table 11.2). The improvement of diagnostic usability was rated slightly less efficient (15 better vs. 6 same and 3 worse). Examples are shown in Figs. 11.5 and 11.6.

One case, where the software facilitated diagnosis, is shown in Fig. 11.5. Although in this case patient motion was not particularly strong, the resulting artifacts are severe (Fig. 11.5a), making the lesion almost invisible among other bright artifacts. In the corrected image (Fig. 11.5b), the lesion is well delineated and the artifacts are completely eliminated.

While the software was very helpful in most cases of motion artifacts, there were a few cases with suspicious areas after correction. The uncorrected image in Fig. 11.6 is strongly affected by patient motion. While the corrected image (Fig. 11.6b) shows improvement in most areas, there is a light enhancement in the marked area (arrow). The unsubtracted original data (not shown) did not support any existence of this enhancement.

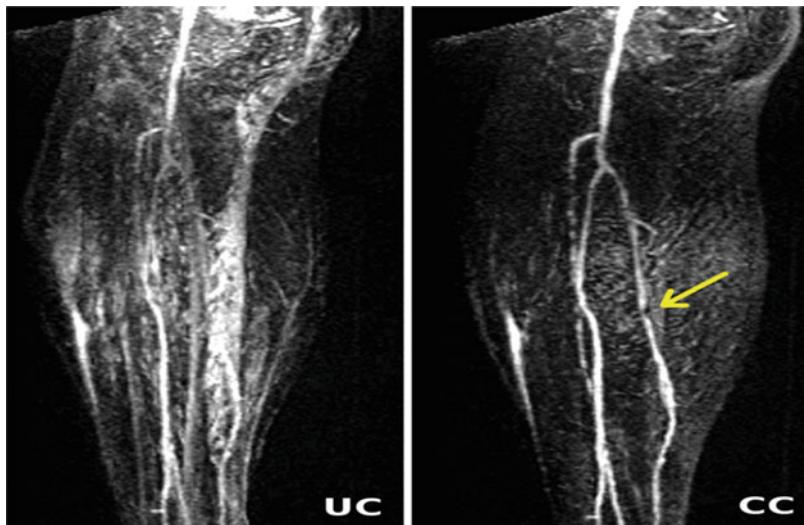


**Fig. 11.6** MR mammography (2D) subtraction images of a female patient, age 61 years. (a) The uncorrected image shows strong artifacts. (b) After correction: the artifacts are eliminated, but a suspicious patchy area (arrow) has been produced by the correction algorithm. There is no indication in the uncorrected images for signal enhancement in that area



**Fig. 11.7** Peripheral angiogram of the lower legs of a female patient, aged 84 years, with severe motion artifacts. (a) MIP image of the uncorrected data. Strong surface artifacts obscure large areas. The vena saphena magna (arrows) is displayed with erroneously “enhanced” signal. (b) MIP of the corrected data. All surface artifacts in (a) are removed and the vein is not visible in the MIP image. The arteries are not suppressed

**MR angiography** The algorithm performed very well on MRA data sets which were affected by motion (see Table 11.2). Most of the artifacts were completely removed in the MIPs. No additional subtraction artifacts were introduced in all but one case where a very light, diffuse artifact was generated. Figure 11.7 shows one example of an MRA of the upper leg and knee area with very severe motion artifacts. The uncorrected MIP (Fig. 11.7a) suffers from large, bright artifacts.



**Fig. 11.8** Contrast-enhanced peripheral MR angiogram of a female patient, aged 64 years, with a stenosis of the right arteria tibialis posterior in the corrected MIP (*right*). The stenosis is not visible in the uncorrected images (*left*)

In addition, it shows an anatomically plausible vena saphena magna (“vein”) implying that the contrast agent has already reached this venous vessel which, however, is an erroneous result due to patient motion. The correction algorithm not only eliminates all the irritating artifacts at tissue boundaries (Fig. 11.7b), but also eliminates the fake CA signal of the vein while conserving the real CA signal of the arteries.

Figure 11.8 shows a MIP where a diagnostically relevant part of the image is affected. The stenosis of the arteria tibialis posterior is completely hidden by the band-like artifact in the uncorrected image (left), but clearly visible in the corrected image (right, see arrow).

## 11.10 Discussion

Data corruption resulting from patient motion can be a serious limitation in many MR examinations. Patient movement can lead to false enhancing regions, causing longer diagnostic reading times, and less clear diagnoses. On the other hand, motion correction by registration can correct for artifacts and potentially eliminates the need for a second scan if patient movement is severe. Although several different techniques have been developed to correct for patient motion, each technique has specific capabilities and limitations.

Apart from nonimage-based techniques, such as immobilization, and imaging methods at the time of the MRI examination, like navigator-based motion

compensation or real time tracking, there is a growing number of publications on image based post-processing techniques to reduce motion artifacts. One of the problems of correcting MR breast data is that breast tissue deforms in a nonrigid, nonlinear manner and does not contain unambiguous internal landmarks. Although there have been several approaches to the problem of matching MR data of the breast [64–69, 71–73], many degrees of freedom are required to perform these nonrigid elastic transformations, which makes the registration problem more challenging and computationally more burdensome [64–66].

Our registration algorithm overcomes this limitation by working on a reduced set of parameters and efficiently computing a dense deformation map from these parameters. The algorithm ensures that the deformation map stays invertible, which is important to ensure that structures in the uncorrected image are not shrunk to a point in the corrected image at which they become invisible. Additionally, the new algorithm uses a robust similarity measure that avoids erroneous interpretation of signal enhancement due to contrast-agent intake as motion.

For dynamic MR mammography, the software improved image quality. It also assisted and enabled successfully the diagnosis based on subtraction images in most cases. In none of the evaluated data sets, obscuration or disappearance of lesions was observed after the correction. Nevertheless, it should be pointed out that only part of the data sets contained lesions and the algorithm should be validated further on pathological cases. Although the algorithm is very capable in correcting motion artifacts for focal lesions, it might prove more difficult to correct motion artifacts in the presence of diffuse contrast enhancements, as commonly seen in, e.g., DCIS patients. As demonstrated in Fig. 11.6, light, diffuse enhancements can be created by the software. Therefore, further patient studies, especially with a diagnosis of DCIS or mastitis, are necessary.

For contrast-enhanced peripheral MR angiography the correction software was found to be highly useful as it substantially improved the subtraction images and the commonly used MIPs. This shortens reading times of angiography images and data sets and facilitates fast and accurate diagnosis. In comparison, the correction algorithm performed better for MRA than for MRM. One of the reasons for this finding is the presence of easily identified internal landmarks (bones, fat, muscle boundaries) in MRA, which are missing in breast MRI. Furthermore, many breast MR exams show only very light motion artifacts. Since radiologists are well familiar with these artifacts, diagnosis is not hampered at all. On the other hand, even light artifacts can affect the quick survey of the vessel status and may make MIPs of MRA exams quite useless. Here, the algorithm helped to improve image quality even in cases with artifacts.

## 11.11 Conclusions

In summary, the proposed algorithm has been shown to improve the quality of subtracted MR images and computed data such as MIPs. Using the software may help to overcome the need for a second examination if patient movement has

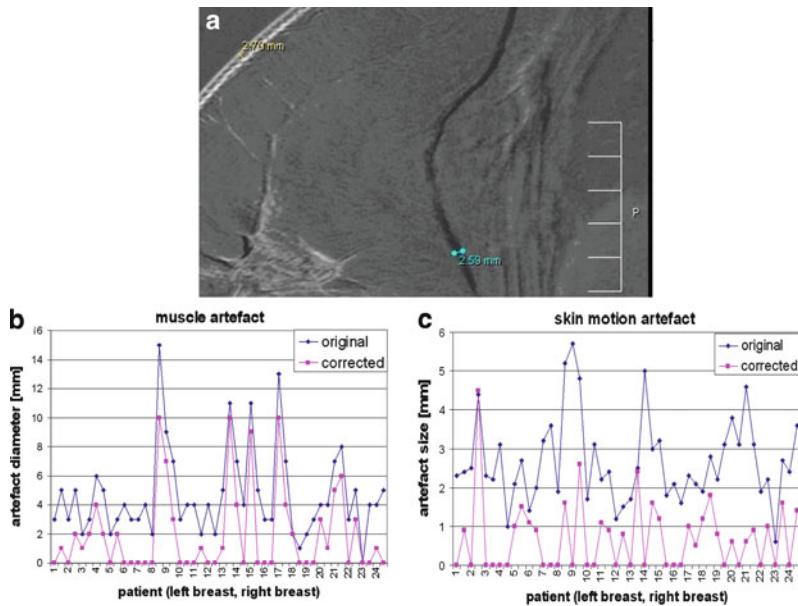
caused artifacts that obscure the images. In addition, the time required to read and analyze the data may be reduced, since many of the falsely enhancing areas were eliminated. The corrected images are not a replacement of the uncorrected images, but represent a valuable adjunct. Examining the corrected subtractions can help to detect otherwise possibly missed lesions (Fig. 11.5).

### 11.11.1 Quantitative Analysis of Motion Artifacts

Motion artifacts can occur in MR breast imaging even with very cooperative patients and highest care in patient positioning, for example due to progressive relaxation of the pectoral muscles. Moreover, as subtractions are routinely calculated for data sets acquired several minutes apart from each other, several motion correction algorithms have been introduced, e.g., [74–76]. In this work, the performance of a 3D elastic motion correction [31] to high resolution breast MR data is described. As opposed to previous publications [77, 78], we have tried to identify quantitative measures on high resolution data sets rather than criteria such as image quality scores that have been applied to the time resolved dynamic measurement.

#### 11.11.1.1 Methods

All experiments were performed on 5T Siemens Magnetom scanners (Avanto or Espree) using a Siemens breast array coil. As part of the routine clinical breast exam, two high-resolution sagittal T1-weighted FLASH data sets are acquired using two excitation slabs,  $0.4 \times 0.4 \times 2 \text{ mm}^3$  voxel size,  $25^\circ$  flip angle, TR 21 ms and 4.7 ms echo time. These scans are performed in the beginning and at the end of the exam, i.e., before and after the transversal dynamic protocol and are used to additionally assess the lesion morphology. These data sets in particular are prone to significant motion artifacts because of their large time interval (approx 15 min). The data sets were subtracted and the motion correction was performed using the algorithm described in Sect. 11.2 (BRACE). The algorithm computes the local cross correlation of corresponding areas around each point. Intensity variations caused by the uptake of a contrast agent are taken into account by a similarity measure across the complete volume, which is maximized using a conjugate gradient method. As a measure of the motion, the maximum apparent skin thickening artifact (due to the subtraction of translated volumes) was evaluated in the central sagittal plane as well as the largest diameter of the subtraction artifact of the greater pectoral muscle in the image that showed the largest subtraction artifact. Independent values were taken for both breasts for 24 out of 31 consecutive patients.



**Fig. 11.9** (a) measurement of skin and muscle motion artifact. (b, c): muscle and skin artefacts before (blue) and after motion correction. Connecting lines have been added between data points for better visibility

### 11.11.1.2 Results and Discussion

Figure 11.9 displays an example for the evaluation of motion artifacts in a subtracted image (before motion correction). Seven data sets did not pass the algorithms quality check. For the remaining data sets, the mean values of the skin motion artifact were originally 2.7 mm (std deviation 1.1 mm) and 0.7 (0.89) mm after performing the correction. The averaged values for the muscle motion artifact was 4.8 (2.95) mm before and 2.04 (2.18) mm after the correction. Overall, a complete correction i.e., no remaining measureable artifact was achieved in 23 out of 47 times for skin motion respective 23 out of 48 times for muscle motion artifact. Further analyses, e.g., a comparison with the motion correction on the dynamic data sets are in preparation. We assume that the uncorrectable data sets were rejected because of too much motion out of or into the imaging volume, so that no continuous solution could be found.

### 11.11.2 Effect of Motion Artifacts on Kinetic Analysis of Breast MR Lesions

The purpose of this study is to investigate whether 3D motion correction can reduce false categorization of lesion kinetics in breast MR imaging.

### 11.11.2.1 Methods and Materials

Thirty-nine breast MR imaging studies showing subtraction motion artifact at clinical read were evaluated to determine whether through-plane motion (TPM) contributed to artifact. Original axial images were registered using the BRACE registration algorithm (Sect. 11.2) in 2D (in-plane only) and 3D (in-plane and throughplane) modes, and corrected subtraction images were generated. Motion artifact was assessed for uncorrected, 2D-corrected, and 3D-corrected subtraction images on a four-point scale (none to minimal, mild, moderate, and marked). TPM was assessed by analyzing sagittal reconstructions of original axial images as an animation of time points, also on a 4-point scale. Forty-three discrete enhancing objects were identified for analysis. Objects were categorized as upgraded, downgraded, or not changed in late enhancement pattern (progressive, plateau, or washout). Truth was determined from original images.

### 11.11.2.2 Results

Misregistration led to false kinetic categorization in 15/43 objects, all upgraded. Correct kinetic categorization was obtained from 3D correction for all 15 objects. 2D correction provided correct categorization for 6/15 objects, 3 from cases with mild or no TPM. Of the 9 objects correctly classified from 3D correction alone, 7 were in cases with moderate or marked TPM. The correlation coefficient between motion artifact and TPM was 0.59 ( $p < 0.001$ ), suggesting that strong subtraction artifact is an indicator of TPM, and that there is a need for 3D registration.

### 11.11.2.3 Conclusion

Misregistration can lead to false kinetic categorization of lesions, and hence to possible lesion misinterpretation and incorrect diagnosis. For cases of strong misregistration in our study, 2D correction often failed leading to false characterization of kinetics. 3D correction consistently provided improved registration without false kinetic categorization, demonstrating the benefit of incorporating 3D correction. TPM may be present in a large proportion of cases that have marked motion artifact. A larger study is required to validate these preliminary results.

## 11.12 Concluding Remarks

In this chapter, we have presented three registration methods that address some of the technical choices that are to be considered as part of an engineering solution to a registration problem, as identified at the beginning of the chapter. For example, the system described in Sect. 11.2 provides specific answers to questions such as how

correspondence is determined and how similarity is optimized. Section 11.3 emphasizes the construction of diffeomorphisms as a direct answer to the invertibility issue. The algorithm described in Sect. 11.4 specifically focuses on the question of how to define correspondence for multimodal data and the algorithm of Sect. 11.5 dwells on the question of whether various registration problems should be tackled simultaneously. Concerning the important question of how a solution is validated, we have also presented three studies designed and performed in close collaboration with radiologists. These provided an insight as to the kind of questions that are of importance in a clinical setting, given any registration method. They focused on exploring qualitative morphology, quantitative morphology, and kinetics of structures of interest in breast MR images, respectively. More specifically, they answered the following:

- Section 11.6.1: Does registration affect the overall image quality and conspicuity of lesions?
- Section 11.6.1: Does registration introduce image artifacts?
- Section 11.6.2: Can quantitative measurements be defined that correlate with the quality of a registration algorithm?
- Section 11.6.3: Can 3D motion correction reduce false categorization of lesion kinetics in breast MR imaging?

Answers to these and other questions, expressed through empirical evaluation with human readers and weighted by statistical analysis, provide better understanding of clinical issues, benefits and limitations of the tools, and how these tools may impact clinical practice. While the algorithms presented in this chapter were driven by concrete clinical needs, they may generalize well to similar problems in other domains that require image processing. The process of implementing an algorithm and validating it in clinical practice, as demonstrated in this chapter, is the vehicle whereby advanced technologies can have the greatest benefit to patients. Thanks to clinical validations, these algorithms have found application in the highly regulated arena of medical industry. We hope that the reader may have found the technical aspects of this chapter useful and the clinical insights motivational toward the development of advanced technologies in medicine.

## References

1. Horn BK, Schunk BG (1981) Determining optical flow. *Artif Intell* 17:185–203
2. Faugeras O, Hotz B, Mathieu H, Vieville T, Zhang Z, Fua P, Theron E, Moll L, Berry G, Vuillemin J, Bertin P, Proy C (1993) Real time correlation based stereo: algorithm implementations and applications. Technical Report 2013, INRIA Sophia-Antipolis, France
3. Faugeras O, Keriven R (1998) Variational principles, surface evolution, PDE's, level set methods and the stereo problem. *IEEE Trans Image Process* 7(3):336–344
4. Cachier P, Pennec X (2000) 3D non-rigid registration by gradient descent on a Gaussian weighted similarity measure using convolutions. In: Proceedings of MMBIA, June 2000, pp 182–189

5. Bride J, Hermosillo G (2001) Recalage rigide sans contrainte de preservation d'intensite par regression heteroscedastique. In: TAIMA, Hammamet, Tunisie, October 2001
6. Netsch T, Rosch P, van Muiswinkel A, Weese J. Towards real-time multi-modality 3D medical image registration
7. Roche A, Malandain G, Pennec X, Ayache N. The correlation ratio as new similarity metric for multimodal image registration. pp 1115–1124
8. Viola P (1995) Alignment by maximization of mutual information, Ph.D. Thesis, MIT
9. Wells WM III, Viola P, Atsumi H, Nakajima S, Kikinis R (1996) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1(1):35–51
10. Viola P, Wells WM III (1997) Alignment by maximization of mutual information. *Int J Comput Vis* 24(2):137–154
11. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16(2):187–198
12. Woods RP, Mazziotta JC, Cherry SR (1993) MRI-PET registration with automated algorithm. *J Comput Assist Tomogr* 17(4):536–546
13. Hill D (1993) Combination of 3D medical images from multiple modalities, Ph.D. Thesis, University of London
14. Penney G, Weese J, Little JA, Desmedt P, Hill DLG, Hawkes DJ (1998) A comparison of similarity measures for use in 2D-3D medical image registration. In: van Leeuwen J, Goos G, Hartmanis J (eds) First international conference on medical image computing and computer-assisted intervention (Lecture notes in computer science), vol 1496. Springer
15. Leventon ME, Grimson WEL. Multi-modal volume registration using joint intensity distributions
16. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1988) Numerical recipes in C. Cambridge University Press, New York
17. Meyer C, Boes J, Kim B, Bland P (1998) Evaluation of control point selection in automatic, mutual information driven, 3D warping. In: van Leeuwen J, Goos G, Hartmanis J (eds) Proceedings of the first international conference on medical image computing and computer-assisted intervention (Lecture Notes in Computer Science), vol 1496
18. Rueckert D, Hayes C, Studholme C, Summers P, Leach M, Hawkes DJ. Non-rigid registration of breast MR images using mutual information
19. Thirion JP (1998) Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal* 2(3):243–260
20. Alvarez L, Deriche R, Weickert J, Sanchez J (2000) Dense disparity map estimation respecting image discontinuities: A PDE and scale-space based approach. International journal of visual communication and image representation, special issue on partial differential equations in image processing, computer vision and computer graphics
21. Proesmans M, Van Gool L, Pauwels E, Oosterlinck A (1994) Determination of optical flow and its discontinuities using non-linear diffusion. In: Proceedings of the 3rd ECCV, II, 1994, number 801, pp 295–304
22. Alvarez L, Weickert J, J Sánchez (2000) Reliable estimation of dense optical flow fields with large displacements. Technical Report, Cuadernos del Instituto Universitario de Ciencias y Tecnologias Ciberneticas, 2000, A revised version has appeared at IJCV 39(1):41–56.
23. Memin E, Perez P (1998) A multigrid approach for hierarchical motion estimation. pp 933–938
24. Memin E, Perez P (1999) Dense parametric estimation of fluid flows. In: IEEE international conference on image processing, ICIP'99, Kobe, Japan, October 1999
25. Aubert G, Kornprobst P (1999) A mathematical study of the relaxed optical flow problem in the space BV. *SIAM J Math Anal* 30(6):1282–1308
26. Aubert G, Deriche R, Kornprobst P (1999) Computing optical flow via variational techniques. *SIAM J Appl Math* 60(1):156–182
27. Deriche R, Kornprobst P, Aubert G (1995) Optical flow estimation while preserving its discontinuities: a variational approach. vol 2, pp 71–80

28. Chefd'hotel C, Hermosillo G, Faugeras O (2002) Flows of diffeomorphisms for multimodal image registration. In: IEEE international symposium on biomedical imaging
29. Aubert G, Kornprobst P (2002) Mathematical problems in image processing: partial differential equations and the calculus of variations, 147. Springer-Verlag, Applied Mathematical Sciences
30. Weickert J, Schnörr C (2001) A theoretical framework for convex regularizers in PDE-based computation of image motion. *Int J Comput Vis* 45(3):245–264
31. Hermosillo G, Chefd'hotel C, Faugeras O (2002) Variational methods for multimodal image matching. *Int J Comput Vis* 50(3):329–343
32. Hermosillo G, Faugeras O (2001) Dense image matching with global and local statistical criteria: a variational approach
33. Hermosillo G (2002) Variational methods for multimodal image matching. Ph.D. Thesis, Institut National de Recherche en Informatique et en Automatique (INRIA), Nice, France
34. Christensen G, Miller MI, Vannier MW (1994) A 3D deformable magnetic resonance textbook based on elasticity. In: Proceedings of the American association for artificial intelligence, symposium: applications of computer vision in medical image processing
35. Trouve A (1998) Diffeomorphisms groups and pattern matching in image analysis. *Int J Comput Vis* 28(3):213–221
36. Miller M, Younes L (2001) Group actions, homeomorphisms, and matching: a general framework. *Int J Comput Vis* 41(1/2):61–84
37. Maintz JBA, Meijering HW, Viergever MA (1998) General multimodal elastic registration based on mutual information. In: Medical imaging 1998 – image processing. SPIE, vol 3338, pp 144–154
38. Hata N, Dohi T, Warfield S, Wells W III, Kikinis R, Jolesz FA (1998) Multi-modality deformable registration of pre-and intra-operative images for MRI-guided brain surgery. In: van Leeuwen J, Goos G, Hartmanis J (eds) First international conference on medical image computing and computer-assisted intervention (Lecture notes in computer science), vol 1496. Springer
39. Gaens T, Maes F, Vandermeulen D, Suetens P (1998) Non-rigid multimodal image registration using mutual information. In: van Leeuwen J Goos G, Hartmanis J (eds) First international conference on medical image computing and computer-assisted intervention (Lecture notes in computer science), vol 1496. Springer
40. Roche A, Guimond A, Meunier J, Ayache N. Multimodal elastic matching of brain images
41. Deriche R (1990) Fast algorithms for low-level vision. *IEEE Trans Pattern Anal Mach Intell* 12(12):78–88
42. Hajnal JV, Hill DLG, Hawkes DJ (2001) Medical image registration, the biomedical engineering series. CRC Press, Boca Raton, FL
43. Evans AC, Beil C, Marrett S, Thompson CJ, Hakim A (1988) Anatomical-functional correlation using an adjustable MRI-based region of interest atlas with positron emission tomography. *J Cereb Blood Flow Metab* 8:513–530
44. Hill DLG, Hawkes DJ, Crossman JE, Gleeson MJ, Cox TCS, Bracey EE, Strong AJ, Graves P (1991) Registration of MR and CT images for skull base surgery using point-like anatomical features. *Br J Radiol* 64:1030–1035
45. West J, Fitzpatrick JM, Wang MY, Dawant BM, Maurer CR Jr, Kessler RM, Maciunas RJ, Barillot C, Lemoine D, Collignon A, Maes F, Suetens P, Vandermeulen D, van den Elsen PA, Napel S, Sumanaweera TS, Harkness B, Hemler PF, Hill DLG, Hawkes DJ, Studholme C, Maintz JBA, Viergever MA, Malandain G, Pennec X, Noz ME, Maguire GQ Jr, Pollack M, Pelizzari CA, Robb RA, Hanson D, Woods RP (1997) Comparison and evaluation of retrospective intermodality brain image registration techniques. *J Comput Assist Tomogr* 21(4):554–566
46. Gonzales R, Woods R (2002) Digital image processing, 2nd edn. Prentice-Hall, Upper Saddle River, NJ

47. Roche A, Malandain G, Pennec X, Ayache N (1998) The correlation ratio as a new similarity measure for multimodal image registration. In: MICCAI, pp 1115–1124
48. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P (1997) Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* 16:187–198
49. Viola P, Wells WM III (1997) Alignment by maximization of mutual information. *Int J Comput Vis* 24(2):137–54
50. Guimond A, Roche A, Ayache N, Meunier J (2001) Three-dimensional multimodal brain warping using the demons algorithm and adaptive intensity corrections. *IEEE Trans Med Imaging* 20(1):58–69
51. Feldmar J, Declerck J, Malandain G, Ayache N (1997) Extension of the ICP algorithm to nonrigid intensity-based registration of 3D volumes. *Comput Vis Image Underst* 66:193–206
52. Netsch T, Roesch P, van Muiswinkel A, Weese J (2001) Towards real-time multi-modality 3-D medical image registration. In: ICCV, pp 718–725
53. Nocedal J, Wright S (1999) Numerical optimization, 2nd edn. Springer, New York
54. Brainweb simulated brain database, Available: <http://www.bic.mni.mcgill.ca/brainweb>
55. Pluim JPW, Maintz JBA, Viergever MA (2003) Mutual information based registration of medical images: a survey. *IEEE Trans Med Imaging* 22(8):986–1004
56. Jenkinson M, Bannister P, Brady M, Smith S (2006) Improved optimization for the robust and accurate linear registration and motion correction of brain images. *IEEE Trans Pattern Anal Mach Intell* 28(2):236–250
57. Zollei L, Learned-Miller E, Grimson E, Wells W (2005) Efficient population registration of 3D data. In: LNCS computer vision for biomedical image applications (ICCV), vol 3765, pp 291–301
58. De Craene MS, Macq B, Marques F, Salembier P, Warfield SK (2008) Unbiased group-wise alignment by iterative central tendency estimation. *Math Model Nat Phenom* 3(6):2–32
59. Marsland S, Twining CJ, Taylor CJ (2008) A minimum description length objective function for groupwise non-rigid image registration. *Image Vis Comput* 26:333–346
60. Pires B, Aguiar P (2005) Featureless global alignment of multiple images. In: IEEE ICIP Proceedings, vol I, pp 57–60
61. Ou W, Chefd'hotel C (2009) Polynomial intensity correction for multimodal image registration. In: Proceedings of ISBI, pp 939–942
62. Heywang-Köbrunner SH, Viehweg P, Heinig A, Küchler C (1997) Contrast-enhanced MRI of the breast: accuracy, value, controversies, solutions. *Eur J Radiol* 24(2):94–108
63. Teifke A, Hlawatsch A, Beier T, Vomweg TW, Schadmand S, Schmidt M, Lehr HA, Thelen M (Sep 2002) Undetected malignancies of the breast: dynamic contrast-enhanced MR imaging at 1.0T. *Radiology* 224(3):881–888
64. Fischer H, Otte M, Ehritt-Braun C, Laubenberger J, Hennig J (1999) Local elastic matching and pattern recognition in MR mammography. *Int J Imaging Syst Technol* 10:199–206
65. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ (1999) Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imaging* 18:712–721
66. Lucht L, Knopp MV, Brix G (2000) Elastic matching of dynamic MR mammographic images. *Magn Reson Med* 43:9–16
67. Tanner C, Schnabel JA, Chung D, Clarkson MJ, Rueckert D, Hill DLG, Hawkes DJ (2000) Volume and shape preservation of enhancing lesions when applying non-rigid registration to a time series of contrast enhancing MR breast images. In: Third international conference on medical image computing and computer-assisted intervention (Lecture notes in computer science), vol 1935. Springer Verlag, pp 327–337
68. Schnabel JA, Tanner C, Castellano-Smith AD, Degenhard A, Leach MO, Hose DR, Hill DL, Hawkes DJ (2003) Validation of nonrigid image registration using finite-element methods: application to breast MR images. *IEEE Trans Med Imaging* 22(2):238–247
69. Lu W, Chen ML, Olivera GH, Ruchala KJ, Mackie TR (2004) Fast free-form deformable registration via calculus of variations. *Phys Med Biol* 49(14):3067–3087

70. He J, Christensen GE (2003) Large deformation inverse consistent elastic image registration. *Inf Process Med Imaging* 18:438–449
71. Reichenbach JR, Hopfe J, Bellemann ME, Kaiser WA (2002) Development and validation of an algorithm for registration of serial 3D MR breast data sets. *MAGMA* 14(3):249–57
72. Rohlfing T, Maurer CR Jr, Bluemke DA, Jacobs MA (2003) Volume-preserving nonrigid registration of MR breast images using free-form deformation with an incompressibility constraint. *IEEE Trans Med Imaging* 22(6):730–741
73. Hopfe J, Herrmann KH, Lucht R, Bellemann ME, Kaiser WA, Reichenbach JR (2005) Validation of an entropy-based algorithm for registration of serial 3D MR mammography data. *Z Med Phys* 15(2):107–114
74. Zuo CS, Jiang A, Buff BL, Mahon TG, Wong TZ (1996) Automatic motion correction for breast MR imaging. *Radiology* 198:903–906
75. Buurman J, Kose U, Gerritsen FA (2007) Fast registration for motion correction of breast MR images. In: ISMRM, 2799
76. Martel AL, Froh MS, Brock KK, Plewes DB, Barber DC (2007) Evaluating an optical-flow-based registration algorithm for contrast-enhanced magnetic resonance imaging of the breast. *Phys Med Biol* 52(13):3802–3816
77. Hermann K-H, Wurdinger S, Fischer DR, Krumbein I, Schmitt M, Hermosillo G, Chaudhuri K, Krishnan A, Salganicoff M, Kaiser WA, Reichenbach JR (2007) Application and assessment of a robust elastic motion correction algorithm to dynamic MRI. *Eur Radiol* 17:259–64
78. Moy L, Hermosillo G, Bogoni L, Periswamy S, Mercado C, Toth HB (2006) Evaluation of a motion correction (MC) algorithm for breast MR. *Eur Radiol* 16(Suppl 5).

## Biography



**Gerardo Hermosillo** received his Ph.D. in 2002, working with Prof. Olivier Faugeras (INRIA, France) on variational methods for multi-modal image registration. In 2003, he joined the Computer Aided Diagnosis group of Siemens Healthcare as deformable matching expert. Since then, he has contributed and led the design and development of algorithms for automatic image analysis and processing related to computer aided detection and diagnosis.



**Christophe Chefd'hotel** graduated from Ecole Supérieure d’Informatique-Electrotechnique-Automatique, Paris, France, and obtained a Ph.D. in applied mathematics from Ecole Normale Supérieure, Cachan, France. His main research interest is the development of medical image registration methods for diagnostic and interventional applications. He currently heads the Interventional Imaging research program at Siemens Corporate Research in Princeton, NJ.



**K.-H. Herrmann** holds a Degree in Physics at Friedrich-Alexander-University Erlangen-Nürnberg, Germany (1997) In 2001 he received a Ph.D. in Geophysics at research center Jülich and University Bonn,Germany; Subsequently he was Post doctoral research at research center Jülich and Department of Atomic Energy, IGCAR, Kalpakkam, Tamil Nadu, India; Since 2004 he has been a member of the Medical Physics Group at the Jena University Hospital.



**Guillaume Bousquet** graduated from EPITA (Paris, France) in 2006 as valedictorian. His main research focus is the optimization and validation of image registration methods. He is now a Research Engineer at Siemens Corporate Research in Princeton, NJ.



**Luca Bogoni** received a B.A. in Mathematics and Computer Science (1986) and a M.S in Computer Science (1988) from Columbia University (NY, NY). He was

awarded his Ph.D. degree from the Moore School of the University of Pennsylvania's GRASP Lab in 1995 (Philadelphia, PA). Worked until 1998 as a Computer Vision researcher at Siemens Corporate Research (Princeton,NJ). In 1998 joined Sarnoff Corporation (Princeton N.J.) and participated and lead various advanced application areas in computer vision and robotics including AVS systems, Biometrics. In 2000 championed medical research activities in Ophthalmology, Breast MR and PET/CT registration. Joined Siemens Healthcare (Malvern, PA) in 2002, where he managed various CAD projects; he is currently the director of Collaboration for CAD.



**Kallol Chaudhuri** works as a Sr. Project Manager at Siemens Healthcare in the Computer Aided Diagnosis Group. His special interest is in new product development and launch. He is also an MBA student at Columbia Business School.



**Dr. Dorothee R. Fischer** holds a degree in Medical Studies from Univ. Freiburg and completed her Residency and Fellowship; Institut für diagnostische, interventionelle und pädiatrische Radiologie Bern, University Hospital Berne, Switzerland, Since 2008 she has been at the Klinik für Nuklearmedizin, University Hospital Zurich, Switzerland.



**Christian Geppert** works as an application developer for Siemens Healthcare in MR Oncology and Interventions. His primary research topic is MR imaging of the breast.



**Rolf Janka** is an assistant medical director at the University Hospital of Erlangen. His research focus is on MR angiography, kidney perfusion and image-guided interventions.



**Arun Krishnan** received the BTech degree in Electrical (Electronics) engineering from Indian Institute of Technology, Chennai in 1987, and the MS and PhD degrees in electrical and computer engineering from the Carnegie Mellon University, Pittsburgh, in 1989 and University of Illinois, Urbana-Champaign in 1997, respectively. From 1995 until 2002, he was at Siemens Corporate Research in Princeton. In 2002, he co-founded the Computer Aided Diagnosis and Therapy (CAD) group in Siemens Medical Solutions in Malvern. He received the Siemens Inventor of the Year award in 2009. Currently, he is the head of the CAD R&D group. His research

interests include computer vision, medical image processing, and graphics. Dr. Krishnan is a member of the RSNA.



**Berthold Kiefer** works as senior scientist for Siemens Healthcare in MR Oncology and Interventions with special interests in whole body and abdominal imaging.



**Ines Krumbein** education 1980–1983. Since 1983 employed by Institute of Diagnostic and Interventional Radiology, Jena University Hospital, Germany. Since 1994 she has worked exclusively in MR Imaging Department as a radiographer.



**Dr. Werner Kaiser** holds degrees in both chemistry and medicine from the University of Freiburg in Germany. Dr. Kaiser has served as a Professor of

Radiology at the University of Wurzburg, as an Assistant Professor of Radiology at the University of Bonn, Chairman of Magnetic Resonance at the Hospital of Nuremberg in Germany, and visiting professor at Harvard University. He is a past recipient of the European Magnetic Resonance Award and the Gödecke Scientific Award. Currently, he serves as Professor and Chairman Chairman of the Institute of Diagnostic and Interventional Radiology at the Friedrich Schiller University in Jena.



**Dr. Michael Middleton** is an Assistant Professor of Radiology at the University of California, San Diego. He received his PhD in physics from the University of Cambridge in England, his MD degree from UC Irvine, and his residency and fellowship training in Radiology at UCSD. He has been imaging breast cancer clinically using MRI for the last 10 years, and partly in collaboration with Siemens. His research interests have included breast MRI registration, artifact characterization, and computer-aided detection.



**Wanmei Ou** obtained a PhD degree in electrical engineering and computer science at the Massachusetts Institute of Technology. Her research focused on biomedical image analysis, ranging from brain anatomy to brain activation detection. Wanmei interned at Siemens Corporate Research where she developed a multimodal image registration algorithm. She is currently working in the product strategy division of Oracle Health Sciences Global Business Unit.



**Jürgen R. Reichenbach** received his MS and PhD in physics from the University of Karlsruhe in 1988 and 1992, respectively. In 1993, he was postdoctoral fellow at the University of Montpellier in France and in 1994 at the Heinrich-Heine University Düsseldorf, Germany. Following 2 years as a visiting associate scientist at the Mallinckrodt Institute of Radiology in St. Louis, he joined the Friedrich-Schiller-University in Jena, Germany in 1997. Since 2007 he is Professor of Medical Physics/Imaging in Jena. He heads a research team working on the development of new measurement techniques in the field of magnetic resonance imaging.



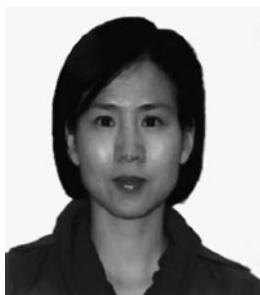
**Marcos Salganicoff** received a BS in electrical engineering from Case Western Reserve University in 1985. From 1985–1987 he was an Member of Technical of Staff at the Caltech’s Jet Propulsion Lab in Pasadena in the applied robotics group. He received the MS and PhD degrees from the Moore School of the University of Pennsylvania’s GRASP Lab in 1992. From 1992 to 1994 he was a National Science Foundation Postdoctoral Associate at University of Pennsylvania in the domain of Machine Learning and Robotics; from 1994 to 1996 he was with the University of Delaware in Rehabilitation Robotics lab. Since 1996 he has held positions of industrial R&D leadership within start-up and well-established companies, focusing on advanced application areas, including biometric security and medical imaging, most recently at Siemens Healthcare in Malvern, Pennsylvania, where he is currently director of CAD R&D.



**Evelyn Wenkel** is an Assistant Medical Director at the University Hospital of Erlangen. Her main research interests are multimodal breast imaging and multi-parametric MR imaging of the breast.



**Susanne Wurdinger, MD**, received her degree in medicine at the University of Jena 1994 and interned from 1995 to 2002 at the Institute for Diagnostic and Interventional Radiology, University of Jena; from 2002 to 2005 she was senior house officer at the Institute for Diagnostic and Interventional Radiology at the University of Jena. In 2005, she established her own private practice specializing in Breast Cancer Screening.



**Li Zhang** received the BEng degree in biomedical engineering from Tsinghua University, Beijing, in 1995, and the PhD degree in biomedical engineering from

the University of Iowa, Iowa City, in 2002. She has been a research scientist at Siemens Corporate Research since 2002. Her primary research interests are medical image registration, motion correction in CT and MRI time series, and statistical model-based segmentation.

# **Chapter 12**

## **The Applications of Feature-Based Image Metamorphosis and Eyelashes Removal in the Investigations of Ocular Thermographic Sequences**

**Jen Hong Tan, E.Y.K. Ng, Rajendra Acharya U, and Jasjit S. Suri**

**Abstract** Ocular surface temperature relates deeply with ocular physiology. It is nowadays looked into mainly by the use of infrared thermography, and occasionally functional infrared imaging. In this study, we recorded ocular thermographic sequences of 60 normal subjects; 30 of them were young and the rest were the elderly. Ocular region in the thermographic sequences was warped to a standardized form, and the eyelashes in each thermographic image were identified and avoided for further analysis. Average horizontal, vertical temperature profiles were unambiguously defined in the warped thermographic sequences. They were extracted, and comparisons were made between the young and the elderly. It was found that the elder subjects exhibited a cooler temperature profile, but their temperature differences between corneal center and limbus were generally higher.

### **12.1 Introduction**

Infrared (IR) thermography is a noninvasive temperature measurement technique capable of real-time capturing and displaying temperature distribution over some surfaces. Nowadays its relevant applications can be found in the field of thermo-fluid dynamics [1], agriculture [2], environmental monitoring [3], and polygraph testing [4]. In medical field, researches have been conducted with this technology to manage neuropathic pain [5, 6], to diagnose impotence [7], breast cancer [8, 9], and thyroid gland disease [10]. For ophthalmology, this method was employed to record ocular surface temperature (OST), either as a single image or a sequence, for the study of ocular physiology and pathology [11].

---

U.R. Acharya (✉)

School of Mechanical and Aerospace Engineering, College of Engineering, Nanyang Technological University, 50, Nanyang Avenue, 639798, Singapore

e-mail: aru@np.edu.sg

Ocular surface temperature deeply relates to many other physiologies and research. The lower corneal surface temperature in comparison to body temperature drives the fluid flow in anterior chamber [12], and is part of the factors in the determination of tear film thinning [13]. The meibomian lipid that forms the lipid layer, and sits on the ocular surface, is temperature-dependent on its fluidity [14]. Ocular models built for finite element analysis were examined and refined according to reported corneal surface temperature [15]. More importantly, under a number of circumstances, a warmer ocular surface temperature is often a manifestation of sick eye [11].

Long before thermography was available to the medical field, corneal temperature was primarily looked into by using invasive measuring techniques, such as needle probe. These methods had a number of disadvantages. They were traumatic and topical anesthesia was often required during measurement. Furthermore, unnecessary blood flow was induced by the act of measurement in the eye and as a consequence, surface temperature was altered. The resultant discrepancies were determined to be up to  $6^{\circ}\text{C}$  in some cases [16].

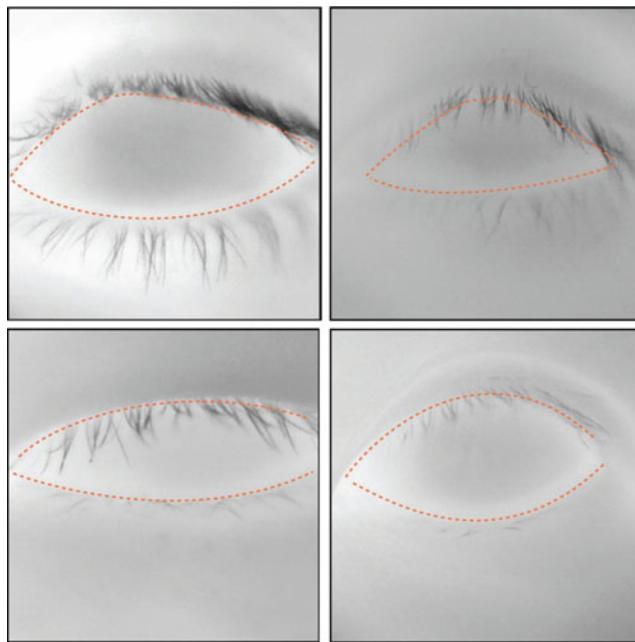
More importantly, invasive measuring techniques measure only the temperature of a single point on the ocular surface, on the other hand infrared thermography images the entire ocular surface temperature. Such advancement has allowed medical researchers to depict the topographical variation in temperature over the ocular surface in detail [17]. In most of the eyes, temperature apex (the coolest point) was observed to be slightly inferior to the geometric center of the cornea, and encircled by a number of ellipsoidal isotherms in most of the OST [17]; vertical OST profile was illustrated to mimic the shape of spoon [18].

Other than the above, infrared thermography has also enabled investigations on the relationship between OST and several physical parameters of anterior chamber [19], human lacrimal drainage system and the effect of non-contact corneal esthesiometer (NCCA) air stimulus [20]. The surface temperature of some ocular diseases such as dry eye [21], Graves' ophthalmopathy [22], glaucoma [23], were also assessed in conjunction with this technique.

For example, in Grave's ophthalmopathy patients, higher temperature difference between lateral orbit and other target areas such as caruncle, medial conjunctiva, lateral conjunctiva, and lower eyelid were observed in comparison to normal subjects [23]. Retrobulbar hemodynamics was determined to not only positively correlate to the resistivity index of left and right eyes respectively, but also the ipsilateral values of end diastolic velocity (both left and right eyes) [23].

The ocular surface of dry eye was observed to cool faster, but found to be warmer [21]. The treatment effect of acupuncture on this ocular disease was looked into and found that that acupuncture was able to affect the temperature of the precorneal tear film [24]. Besides, the infrared technology was applied to diagnose dry eye disease. Chiang et al. attempted the diagnosis of the dry eye patients with a sensitivity of 79% and a specificity of 75%, using dynamic infrared imaging [25].

Despite the above investigations, in the realm of ocular surface temperature, two issues remained unresolved. First, the focus of investigation was restricted on corneal region, and the relatively simple and incomprehensive analytical methods.

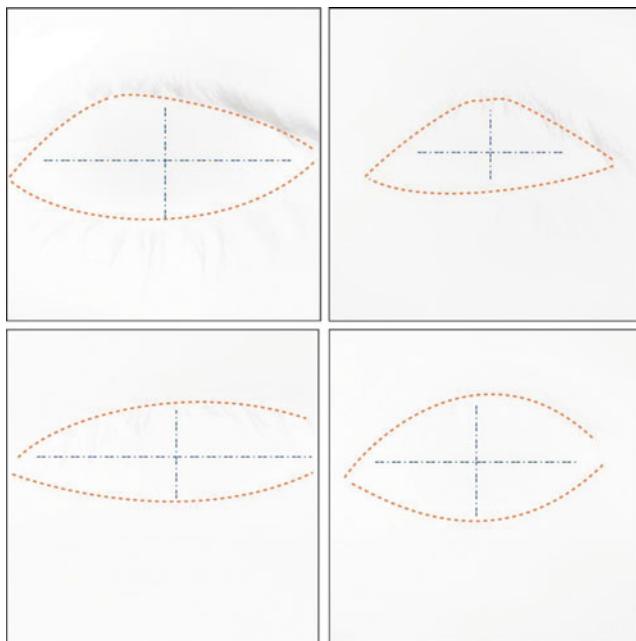


**Fig. 12.1** Myriad shapes in palpebral fissure, and the indiscernible corneal limbus. *Dotted lines* in the subfigures indicate the shape profile of each palpebral fissure

Ocular surface temperature has hardly been preprocessed and analyzed as whole, but analyzed only portion of it. This is partly due to the indiscernible corneal limbus, and also the myriad shapes of palpebral fissure as illustrated in Fig. 12.1.

Second, the topographical pattern and both the horizontal and vertical temperature profile among age groups were scarcely looked into. Although there were studies on temperature profile, the difficulty in marking out a standardized, consistent horizontal profile in myriad ocular shapes led the previous investigations reveal pitiful information. As observed in Fig. 12.2, which draws out the lines where the usual horizontal and vertical temperature profiles are derived in each thermogram, these profiles are rather poorly defined except their coincidence with the corneal geometric center. As for vertical temperature profile, the presence of eyelashes, which exhibit lower temperature, often obstructs the acquisition of actual temperature values.

Hence in this chapter, we attempt to address these issues and provide satisfactory solutions to the problems. First, we recorded ocular thermographic sequences of the young and the elderly. Second, for each thermographic sequence, the temperature data within the subject's palpebral fissure is warped into a standard form. Such transformation allows consistent investigations on the entire ocular surface, and the myriad shapes in palpebral fissure are no longer a hindrance to the comprehensive analysis of temperature in that region. Subsequently, in each thermographic frame,



**Fig. 12.2** The lines of horizontal and vertical temperature profiles in each thermogram

we remove the eyelashes so that their far lower temperature would not be included in latter analysis. Finally, we examined the average horizontal and vertical temperature profile in each age group.

The chapter is laid out as follows: details in data acquisition are elaborated in Sect. 12.2. Section 12.3 illustrates the methodology of image warping and removal of eyelashes. Results are reported and discussed in Sect. 12.4. Finally, the chapter ends in Sect. 12.5.

## 12.2 Data Acquisition

In this study, we recorded ocular thermographic sequences of 60 normal subjects (mean age of  $33 \pm 15.3$ ), as tabulated in Table 12.1. Subjects were asked to relax for 20 min prior to thermographic recording and participants who had the following conditions were excluded:

1. Hot drinks, rubbing of eyes, use of tear drops
2. Poor general health
3. History of serious eye disease
4. History of ocular or facial surgery

**Table 12.1** Details of the subject group

	No. of subjects	(Age) Mean $\pm$ Std
Young subjects	30	18.9 $\pm$ 0.8
Elder subjects	30	47 $\pm$ 8.4

5. History of serious eye disease
6. Recent usage of tear drop or medications.

Four thermographic sequences were recorded for each subject, two for the right eye and another two for the left, using the VarioCAM, JENOPTIK Laser (Germany) (URL: <http://www.jenoptik-ir.com/>), which has a measuring accuracy of  $\pm 0.2\%$  and thermal resolution power of less than 0.1 K. In each thermographic sequence, 25 sequential shots were captured in every second, and the recording lasted for 20 s. During recording, subjects were allowed to blink at will, placed their head on a chin rest, and the instrument was supported by a tripod, kept 50 cm away from the chin rest. After the recording, the corresponding sequences were stored in *irbis* format and exported to *jpeg* format for latter processing. In each sequential thermal image, the darkest and the brightest image intensity level correspond to 28°C and 36°C, respectively.

The data was collected at Biomedical Centre, Ngee Ann Polytechnic, Singapore. Environment conditions of the room where this study took place were maintained, with room temperature controlled at  $22.5 \pm 1.0^\circ\text{C}$ , and humidity at  $70 \pm 4\%$ .

## 12.3 Methods

After the recording of thermographic sequence and the export to jpeg images, the ocular regions in all the sequential thermal images of sequence is warped into a standard form (except frames in which subject performs blinking). This allows the comparisons of ocular thermal data of various ocular sizes and shapes, leading to better analysis and comparison. Followed by this is the identification of eyelashes in the above thermal images using a novel algorithm, and the corresponding temperature data are excluded in the subsequent analysis.

### 12.3.1 *Image Warping*

Generally speaking, there are two sorts of image warping: forward mapping and reverse mapping [26]. The method used in this study uses reverse mapping. It was originally proposed by Beier and Neely [27] and similarly adopted in the work by Tan et al. [28]. The warping algorithm starts with a pair of lines, one defined on a source image and another on destination image, it scans through destination image

pixel by pixel, and samples the relevant pixels of source image [27, 28]. For each line  $\mathbf{Q}_{q_i}\mathbf{Q}_{q_{i+1}}$ ,

$$\mathbf{u} = \frac{(\mathbf{D} - \mathbf{Q}_{q_i}) \cdot (\mathbf{Q}_{q_{i+1}} - \mathbf{Q}_{q_i})}{\|\mathbf{Q}_{q_{i+1}} - \mathbf{Q}_{q_i}\|^2} \quad (12.1)$$

$$\mathbf{v} = \frac{(\mathbf{D} - \mathbf{Q}_{q_i}) \cdot \text{perpendicular } (\mathbf{Q}_{q_{i+1}} - \mathbf{Q}_{q_i})}{\|\mathbf{Q}_{q_{i+1}} - \mathbf{Q}_{q_i}\|} \quad (12.2)$$

$$\mathbf{D}' = \mathbf{Q}'_{q_i} + \mathbf{u} \cdot (\mathbf{Q}'_{q_{i+1}} - \mathbf{Q}'_{q_i}) + \frac{\mathbf{v} \cdot \text{perpendicular } (\mathbf{Q}'_{q_{i+1}} - \mathbf{Q}'_{q_i})}{\|\mathbf{Q}'_{q_{i+1}} - \mathbf{Q}'_{q_i}\|} \quad (12.3)$$

where  $\mathbf{D}$  and  $\mathbf{D}'$  denote the coordinates of destination image pixel and source image pixel respectively, as shown in Fig. 12.3; the function perpendicular( ) produces the vector perpendicular to and equal in magnitude to the input vector. For the case of multiline, the algorithm is as follows:

For each pixel  $\mathbf{D}$  in destination image

**DSUM = 0**

**weightsum = 0**

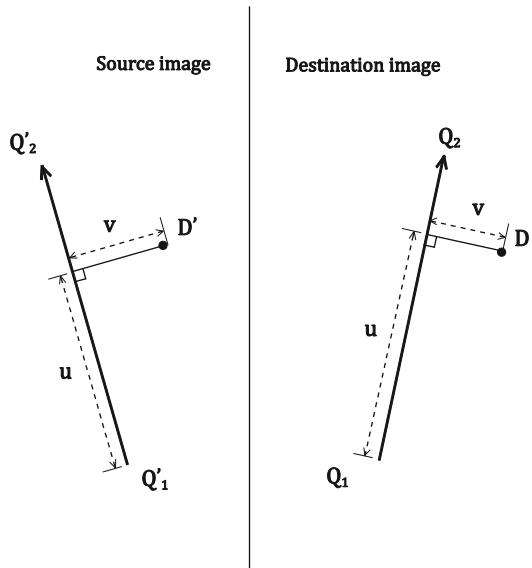
For each line  $\mathbf{Q}_{q_i}\mathbf{Q}_{q_{i+1}}$

Get  $\mathbf{u}, \mathbf{v}$  determined on  $\mathbf{Q}_{q_i}\mathbf{Q}_{q_{i+1}}$

Get  $\mathbf{D}'_{q_i}$

Calculate  $\mathbf{Ds}_{q_i} = \mathbf{D}'_{q_i} - \mathbf{D}$

**Fig. 12.3** A single line pair between the source and destination image

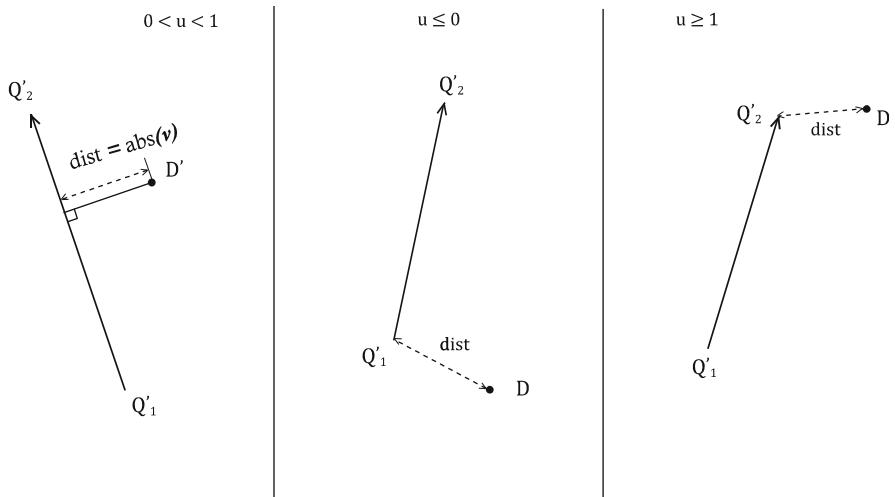


$$\begin{aligned}
 \mathbf{dist} &= \text{shortest distance from } \mathbf{D} \text{ to } \mathbf{Q}_{q_i} \mathbf{Q}_{q_i+1} \\
 \mathbf{weight} &= \left( \|Q_{\downarrow}(\uparrow q)i Q_{\downarrow}(\uparrow q)i + 1\|^{\uparrow} (c_{\downarrow} 1) / ((c_{\downarrow} 2 + \mathbf{dist})) \right)^{\uparrow} (c_{\downarrow} 3) \\
 \mathbf{DSUM+} &= \mathbf{Ds}_{q_i} \cdot \mathbf{weight} \\
 \mathbf{weightsum+} &= \mathbf{weight} \\
 \mathbf{D}'_{q_i} &= \mathbf{D}_{q_i} + \mathbf{DSUM}/\mathbf{weightsum} \\
 \mathbf{destinationimage}(\mathbf{D}) &= \mathbf{sourceimage}(\mathbf{D})
 \end{aligned}$$

In which the shortest distance from  $\mathbf{D}$  to  $\mathbf{Q}_{q_i} \mathbf{Q}_{q_i+1}$  is  $abs(v)$  if  $0 < u < 1$ ; for  $u \leq 0$ , and that value equals to the straight-line distance between  $\mathbf{D}$  and  $\mathbf{Q}_{q_i}$ ; if  $u \geq 1$ , it is the straight-line distance between  $\mathbf{D}$  and  $\mathbf{Q}_{q_i+1}$  (Fig. 12.4).

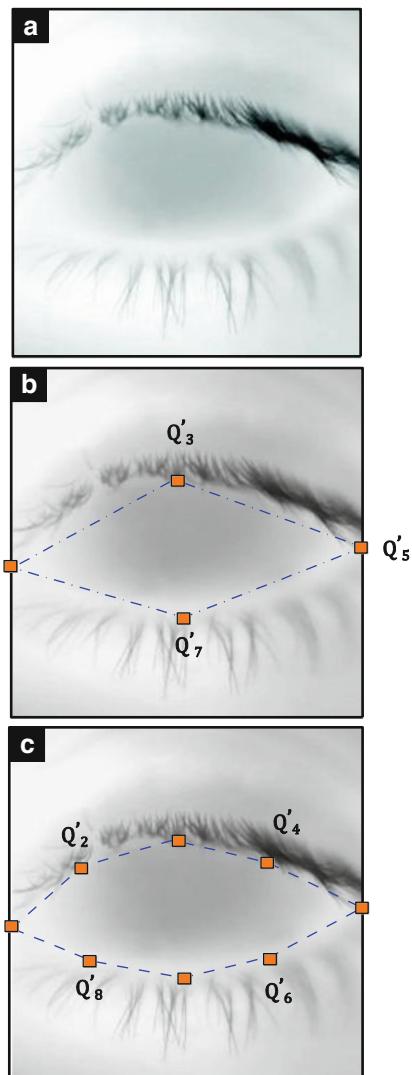
In this investigation, we warp the ocular thermal images by the following procedures. First, determine eight points, forming a polygon that approximates the shape of an eye, on the first thermal image of any length of thermographic sequence. As shown in Fig. 12.5b, points  $Q'_1, Q'_3, Q'_5$ , and  $Q'_7$  are selected at first and followed by  $Q'_2, Q'_4, Q'_6$ , and  $Q'_8$ , as seen in Fig. 12.5c. Points  $Q'_1, Q'_3, Q'_5$ , and  $Q'_7$  correspond to the medial canthus, middle upper edge, lateral canthus, and middle lower edge of a left eye (and lateral canthus, middle upper edge, medial canthus and middle lower edge of a right eye).  $Q'_2$  is on the location somewhere middle in between  $Q'_1$  and  $Q'_3$  along the edge of an eye, and  $Q'_4$  somewhere middle in between  $Q'_3$  and  $Q'_5$ ; similarly for  $Q'_6$  and  $Q'_8$ . Figure 12.6 illustrates the corresponding points and lines in destination image, the size of which is  $75 \times 150$  pixels.

After the above point selection, every thermal image in the thermographic sequence is warped according to the above reverse mapping algorithm, based on the selected polygon. Analysis is performed on the temperature values within the



**Fig. 12.4** Three cases on the shortest distance from  $\mathbf{D}$  to  $\mathbf{Q}_{q_i} \mathbf{Q}_{q_i+1}$

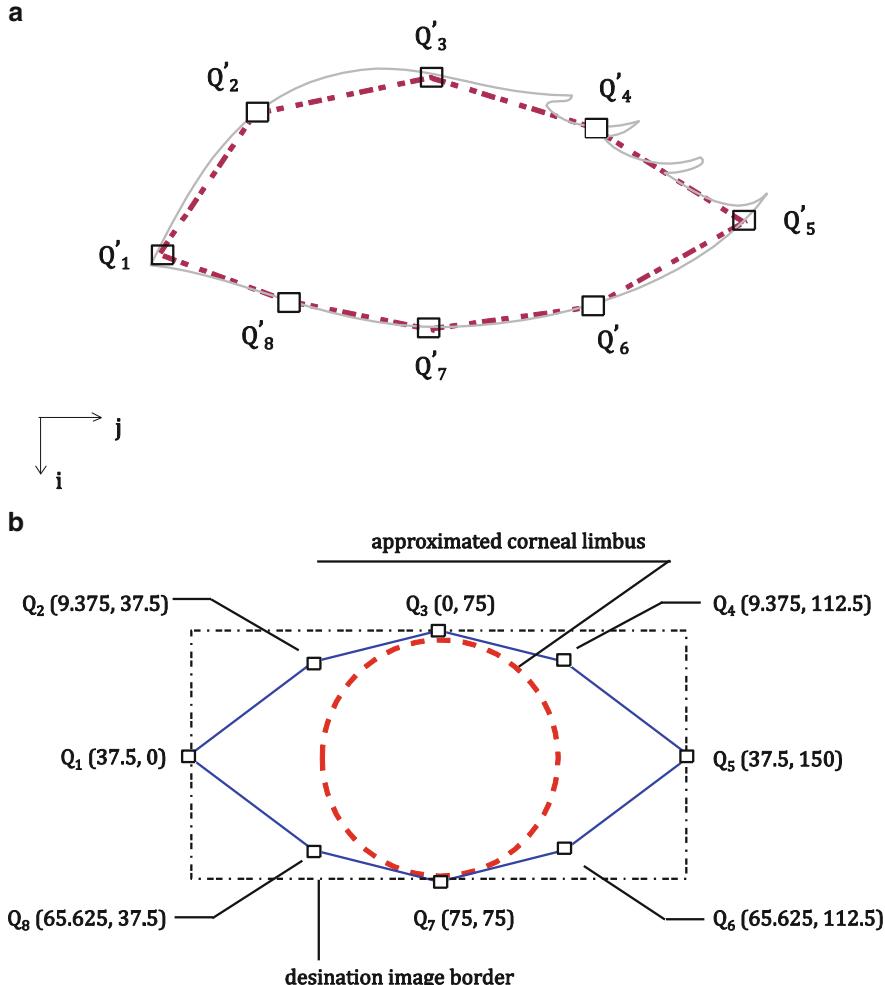
**Fig. 12.5** (a) Original thermal image (b)  $Q'_1, Q'_3, Q'_5$ , and  $Q'_7$  are determined at first, dotted brown lines connecting these points follow to appear to facilitate the selection of points  $Q'_2, Q'_4, Q'_6$  and  $Q'_8$ . (c) The determination of points  $Q'_1, Q'_4, Q'_6$ , and  $Q'_8$



region enclosed by  $Q_1$  to  $Q_8$  in each destination image. An example is given in Fig. 12.7, with isothermal lines marked out to illustrate the thermal distribution across the region.

### 12.3.2 Removal of Eyelashes

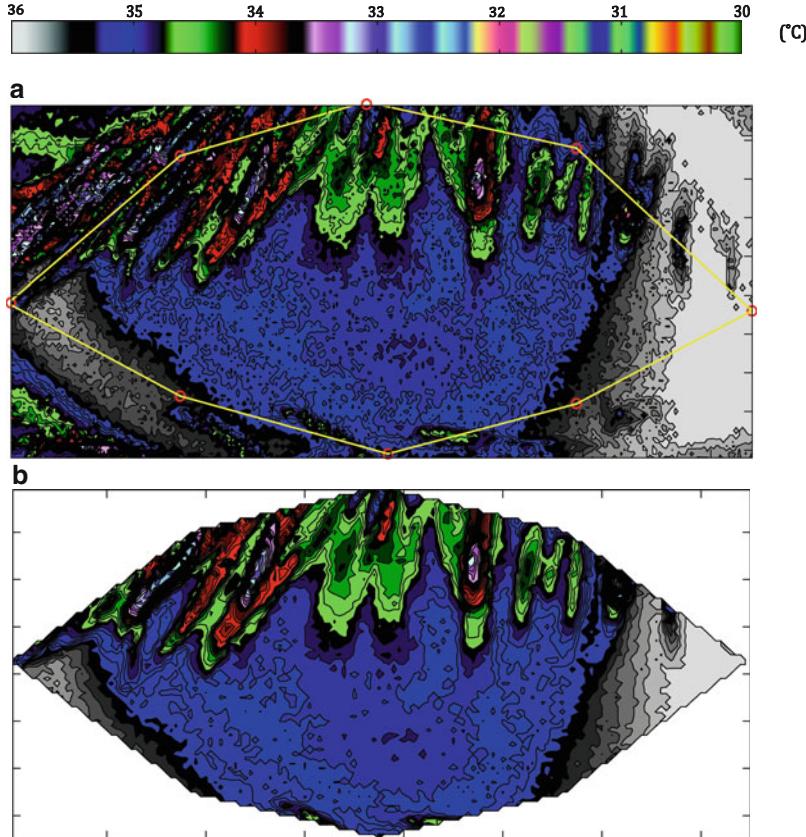
Let  $I(i, j, f)$  denote a thermographic sequence in which the frames, where the subject does a blink are removed; the  $i$  and  $j$  refer to the image spatial domain



**Fig. 12.6** Points and lines defined in (a) source and (b) destination images

whereby  $f$  refers to the number of a particular frame. Such a thermographic sequence, assume, has  $n_F$  of frames, and is further divided into a number of subsequences, each of them having  $n_{S_q}$  of frames (except the last subsequence if  $n_F/n_{S_q}$  is not an integer). The differential image [29] for any subsequence is given by

$${}^nI_\varepsilon(i,j) = \frac{I_\varepsilon(i,j) - \min I_\varepsilon(i,j)}{\max I_\varepsilon(i,j) - \min I_\varepsilon(i,j)} \cdot 255 \quad (12.4)$$



**Fig. 12.7** Image metamorphosis: (a) original image and (b) destination image

and

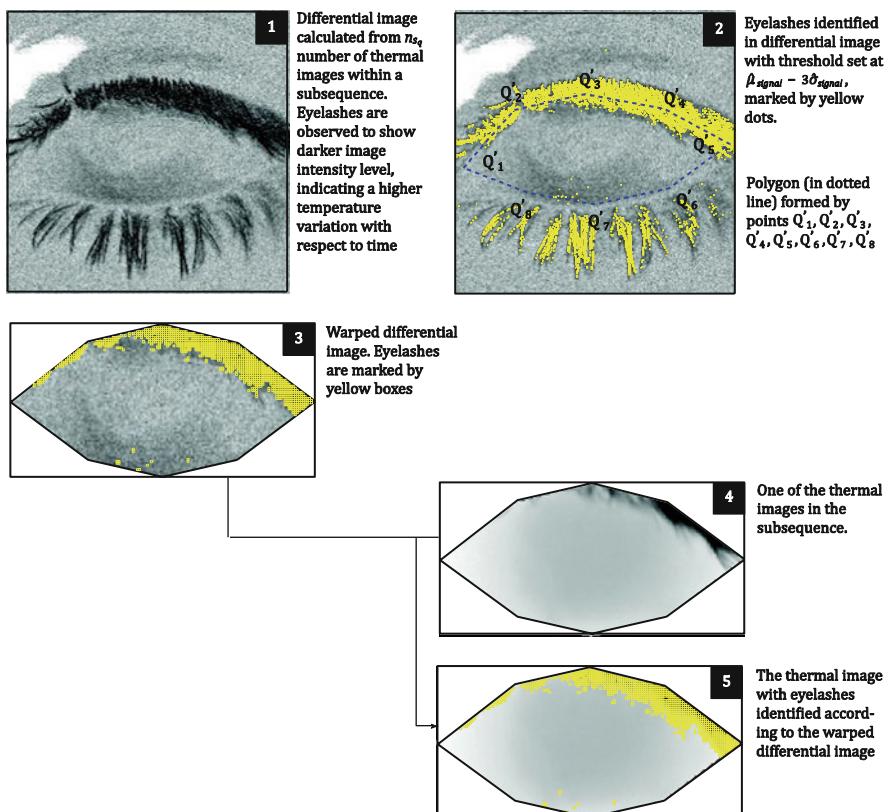
$$I_e(i,j) = \frac{1}{n_{S_q}} \sum_{i_e} \frac{1}{1 + \|I(i,j, i_e) - I(i,j, i_e + i_{df})\|}, \quad i_e = f_{df}, f_{df} + 1, f_{df} + 2, \dots, n_{S_q} \quad (12.5)$$

where  $\|\cdot\|$  denote absolute value and  $f_{df}$  the number of the beginning frame for the subsequence. In the differential image, it can be observed that, the larger the temporal variation in temperature results in dark image pixel [29]. In general, temperature of eyelashes varies greater compared to ocular surface and orbital skin. More importantly, the resultant differential image has an image histogram showing approximately a bimodal distribution (with image intensity level 255 excluded) which can be modeled by a Gaussian mixture

$$\phi(x[n]; \varepsilon) = (1 - \varepsilon)\phi_{\text{signal}}(X) + \varepsilon\phi_{\text{noise}}(X) \quad (12.6)$$

where  $\varepsilon$  is the mixture parameter, and  $0 < \varepsilon < 1$ ;  $\phi_{\text{signal}}$  and  $\phi_{\text{noise}}$  correspond to the Gaussian probability density functions of the non-eyelashes and eyelashes regions.

To identify eyelashes, the Gaussian mixture parameter  $\hat{\mu}_{\text{signal}}, \hat{\sigma}_{\text{signal}}, \hat{\mu}_{\text{noise}}, \hat{\sigma}_{\text{noise}}$  are first estimated using Expectation-Maximization (EM) algorithm. Then set a threshold value to be  $\hat{\mu}_{\text{signal}} - 3\hat{\sigma}_{\text{signal}}$ , those pixels having image intensity level lower than the threshold in the differential image are identified as eyelashes, and these pixels are given a new but very high intensity value (e.g.,  $1 \times 10^6$ ). After that, the modified differential image is warped according to the reverse mapping algorithm (Sect. 3.1), and eyelashes in the corresponding destination image would be manifested as pixels with image intensity level beyond value 255. For each thermal image in the subsequence, from which the differential image was derived, the eyelashes in their corresponding destination images are picked out according to the warped differential image. Temperature values of those pixels are excluded in later analysis. The above procedures are illustrated in Fig. 12.8.



**Fig. 12.8** Removal of eyelashes

### 12.3.3 Temperature Profile

Let  $t_{i,j,f}$  denote the temperature of a point  $(i,j)$  on the particular frame  $f$  in a warped thermographic sequence  $I(i,j,f)$ , which meets the definition laid down in Sect. 3.2. The point  $(i,j)$  belongs to neither any part of eyelashes nor to regions other than the ocular surface and in frame  $f$ . By this, the horizontal temperature profile at the frame  $f$  (Fig. 12.9) is obtained by

$$\mathcal{H}_f = [t_{38,1,f} \quad t_{38,2,f} \quad t_{38,3,f} \quad \cdots \quad t_{38,150,f}] \quad (12.7)$$

since the size of the destination image for any warped thermal images is  $75 \times 150$ . From (12.7), we get the horizontal temperature profile of the thermographic sequence by

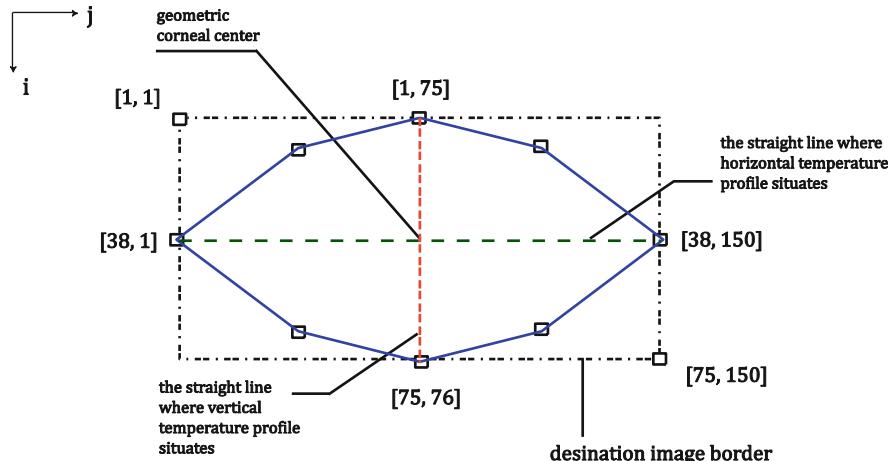
$$\mathcal{H} = \frac{\mathcal{H}_1 + \mathcal{H}_2 + \mathcal{H}_3 + \cdots + \mathcal{H}_f}{n_F} \quad (12.8)$$

And the average horizontal temperature profile for the subject group is

$${^{\mu}\mathcal{H}} = \frac{1}{m_s} \sum_{k=1}^{m_s} {}_k\mathcal{H} = [{}^{\mu}_{\mathcal{H}}t_1 \quad {}^{\mu}_{\mathcal{H}}t_2 \quad {}^{\mu}_{\mathcal{H}}t_3 \quad \cdots \quad {}^{\mu}_{\mathcal{H}}t_{150}] \quad (12.9)$$

where  ${}_k\mathcal{H}$  denote the horizontal temperature profile of the  $k$ th thermographic sequence in the subject group, which in total has  $m_s$  thermographic sequences. Besides, we can get the mean ocular surface temperature which is defined on  ${}^{\mu}\mathcal{H}$  by

$${}_{\mathcal{H}}\mathcal{D} = \frac{({}^{\mu}_{\mathcal{H}}t_1 + {}^{\mu}_{\mathcal{H}}t_2 + {}^{\mu}_{\mathcal{H}}t_3 + \cdots + {}^{\mu}_{\mathcal{H}}t_{150})}{150} \quad (12.10)$$



**Fig. 12.9** Horizontal and vertical temperature profiles in destination image

Similarly, we can get the vertical temperature profile in the same spirit as (12.7–12.10). Hence,

$$v_f = [t_{1,76,f} \quad t_{2,76,f} \quad t_{3,76,f} \quad \cdots \quad t_{75,76,f}] \quad (12.11)$$

$$v = \frac{v_1 + v_2 + v_3 + \cdots + v_f}{n_F} \quad (12.12)$$

$${^{\mu}}v = \frac{1}{m_s} \sum_{k=1}^{m_s} k v = [{^{\mu}}v_1 \quad {^{\mu}}v_2 \quad {^{\mu}}v_3 \quad \cdots \quad {^{\mu}}v_{75}] \quad (12.13)$$

$${_v}\mathcal{D} = \frac{({^{\mu}}v_1 + {^{\mu}}v_2 + {^{\mu}}v_3 + \cdots + {^{\mu}}v_{75})}{75} \quad (12.14)$$

Lastly, we obtain the mean ocular surface temperature that is derived from both the  ${}^{\mu}\mathcal{H}$  and  ${}^{\mu}v$ :

$${}_{\mathcal{H},v}\mathcal{D} = \frac{({^{\mu}}\mathcal{H}_1 + {^{\mu}}\mathcal{H}_2 + {^{\mu}}\mathcal{H}_3 + \cdots + {^{\mu}}\mathcal{H}_{150}) + ({^{\mu}}v_1 + {^{\mu}}v_2 + {^{\mu}}v_3 + \cdots + {^{\mu}}v_{75})}{150 + 75} \quad (12.15)$$

## 12.4 Results and Discussion

In this investigation, younger subjects were observed to have warmer ocular surface. In this age group, we recorded temperature value of 34.42°C, 34.09°C, and 34.31°C for the statistics  ${}_{\mathcal{H}}\mathcal{D}$ ,  ${}_v\mathcal{D}$ , and  ${}_{\mathcal{H},v}\mathcal{D}$  respectively, whereas in the elders, the calculated values for  ${}_{\mathcal{H}}\mathcal{D}$ ,  ${}_v\mathcal{D}$ , and  ${}_{\mathcal{H},v}\mathcal{D}$  were 33.88°C, 33.54°C, and 33.77°C. The above findings were statistically significant ( $p < 0.001$ ).

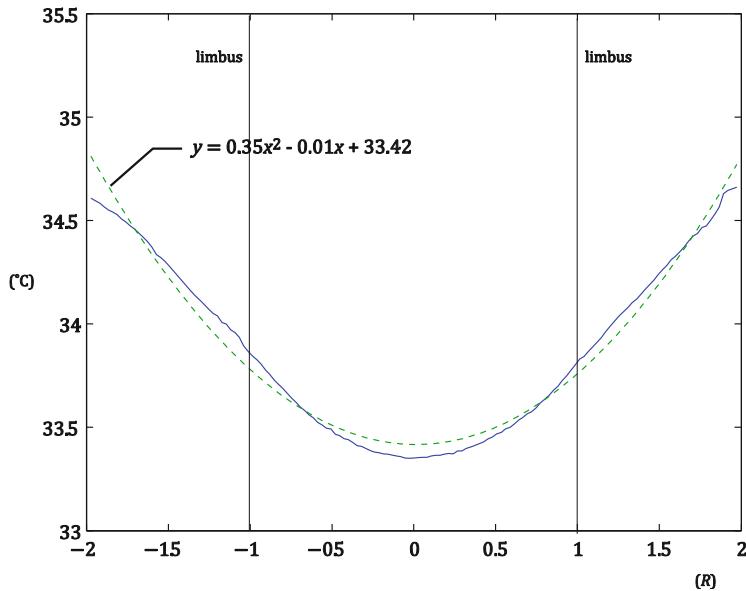
Figures 12.10 and 12.11 illustrate the average horizontal and vertical temperature profile for the age group above 35, whereas Figs. 12.12 and 12.13 for the age group below 35. In each average temperature profile, a polynomial regression was approximated in the form of

$$T = a_h x^2 + b_h x + c_h \quad (12.16)$$

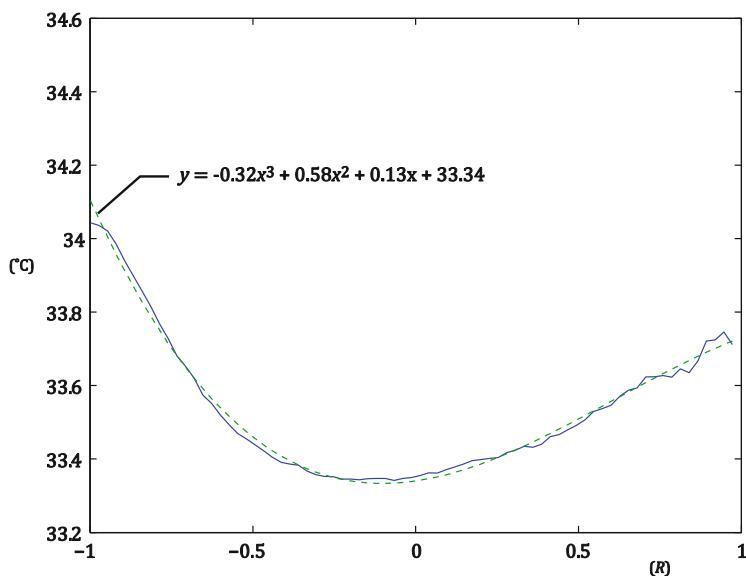
for horizontal temperature profile, and

$$T = a_v x^3 + b_v x^2 + c_v x + d_v \quad (12.17)$$

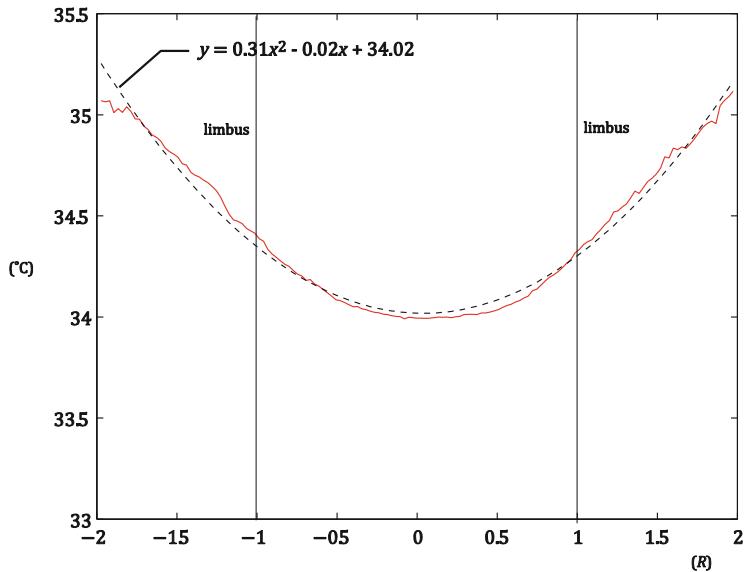
for vertical temperature profile.  $T$  denotes the temperature at the location  $x$ , which bears the unit  $R$ , the corneal radius, and  $x = 0$  is the corneal center. The respective polynomial coefficients are tabulated in Tables 12.2 and 12.3.



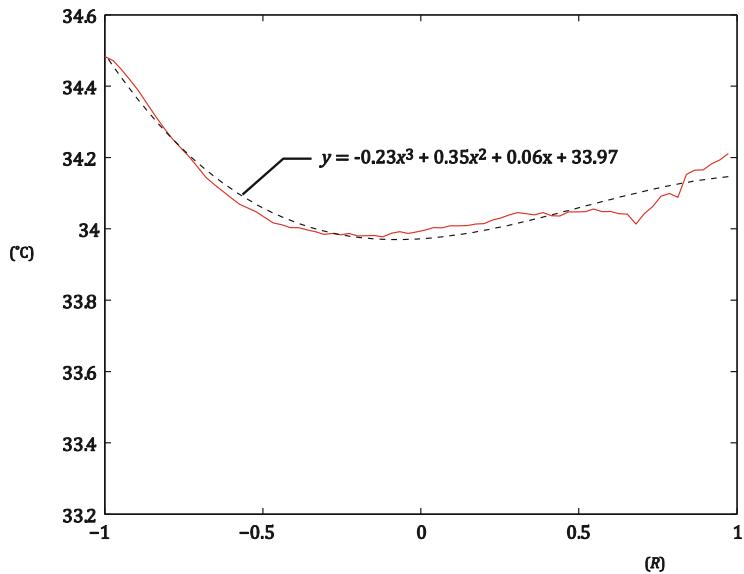
**Fig. 12.10** Average horizontal temperature profile for age group above 35. *Dotted line* denotes the profile generated by the corresponding polynomial regression



**Fig. 12.11** Average vertical temperature profile for age group above 35. *Dotted line* denotes the profile generated by the corresponding polynomial regression



**Fig. 12.12** Average horizontal temperature profile for age group below 35. *Dotted line* denotes the profile generated by the corresponding polynomial regression



**Fig. 12.13** Average vertical temperature profile for age group below 35. *Dotted line* denotes the profile generated by the corresponding polynomial regression

For horizontal temperature profile, the coolest point almost fell at the corneal center in both age groups, having the temperature value of 33.35°C in the elders and 33.99°C for the other. In the case of vertical temperature profile, the lowest

**Table 12.2** Polynomial coefficients for horizontal temperature profile

	$a_h$	$b_h$	$c_h$
Age group above 35	0.35	-0.01	33.42
Age group below 35	0.31	-0.02	34.02

**Table 12.3** Polynomial coefficients for vertical temperature profile

	$a_v$	$b_v$	$c_v$	$d_v$
Age group above 35	-0.32	0.58	0.13	33.34
Age group below 35	-0.23	0.35	0.06	33.97

temperature was observed at some point inferior to the corneal center, regardless of age group. The coolest point (vertical temperature profile) for the younger subjects recorded was 33.98°C and 33.34°C for their elder peers.

Previously, Efron et al. measured surface temperature at every 0.5 mm increments on either side of corneal surface horizontally. In total there were 11 points running across the anterior surface, with the sixth point coinciding with the geometric corneal center [17]. They reported a temperature of 34.3°C at corneal center, and derived a polynomial regression for the horizontal temperature profile written as

$$\Delta T = 0.01x_e^2 + 0.003x_e + 0.01 \quad (18)$$

however,  $\Delta T$  denotes the temperature difference between a point  $x_e$  and the geometric corneal center, and  $x_e$  the distance from the geometric center of cornea in unit millimeter.

On the other hand, though no polynomial regression was provided, Tan et al. reported a similar horizontal temperature profile and a vertical temperature profile, which had a shape like a spoon [18]. However, during data collection, they made use of cotton bud to lift up the subject's eyelid to get rid of the eyelashes falling on the ocular surface. Such practice could induce undesired reflex tearing in subject, as reported by Mapstone [30]. In our case, this issue was overcome by the eyelashes removal algorithm.

The temperature profile illustrated in this study, either of horizontal or vertical, was well-defined. As detailed in Fig. 12.9, horizontal temperature profile was inferred from the straight line connecting medial canthus, lateral canthus, and the corneal center. This was only possible in the warped destination image, since in actual palpebral fissure, these three anatomical points rarely form a straight line, as shown in Fig. 12.2. As for the vertical counterpart, it is a straight line profile which links the middle upper eyelid, middle lower eyelid, and the center of cornea.

The variations in temperature profile between the age groups, both of horizontal and vertical, were clearly visible. Regardless of directionality, the temperature of every points in the profile was higher in the younger subjects. On the other hand the parabola in the elder subjects' average horizontal temperature profile was observed to be slightly steeper. The limbal temperature, determined on horizontal temperature

profile, was 33.83°C for the elderly and 34.36°C for the young. In other words, corneal limbus was 0.48°C warmer than the corneal center in the elders, and their younger peers' in average had a corneal center 0.37°C cooler compared to their limbus. These values were close to the one observed by Efron et al., who reported a temperature difference between geometric corneal center and limbus of 0.45°C, also the figure of 0.37°C by Morgan et al [31]. Tan et al. did not report the mean for this parameter, but indicated that the value ranged from 0.23°C to 0.43°C. They speculated that such a lower figure of this parameter in their finding was a consequence of the recruited Chinese subjects, who were found to generally exhibit lower tear volume and tear stability compared to Caucasians [18]. However, we think that this might be more due to the age of the recruited subjects: the participants in their investigations aged 17–21, with a mean of 19, and in our investigations, the younger age group (mean age 18.9) had a lower limbus-center temperature difference (0.37°C). For the elders, our finding reported a difference of 0.48°C, which was 0.03°C warmer than the one by Efron et al.

## 12.5 Conclusion

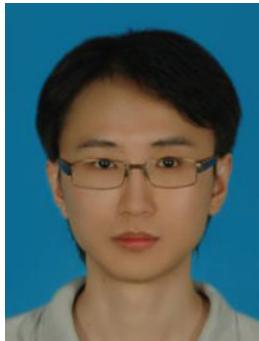
In this chapter, we have demonstrated the application of warping (image metamorphosis) and eyelashes removal algorithm to the ocular thermographic sequences. The use of these two computer methods allowed a better defined horizontal and vertical temperature profiles. Our finding indicated a cooler temperature profile in the elders, but a higher temperature difference between corneal center and limbus.

## References

1. Meola C, Carlomagno GM (2004) Recent advances in the use of infrared thermography. *Meas Sci Technol* 15:R27–R58
2. Wisniewski M, Lindow S, Ashworth E (1997) Observations of ice nucleation and propagation in plants using infrared video thermography. *Plant Physiol* 113:327–334
3. Cehlin M, Moshfegh B, Sandberg M (2000) Visualization and measuring of air temperatures based on infrared thermography. Presented at the proceedings of the 7th international conference on air distribution in rooms ROOMVENT, Reading, UK
4. Pavlidis I, Levine J (2002) Thermal image analysis for polygraph testing. *IEEE Eng Med Biol Mag* 21:56–64
5. Hooshmand H, Hashmi M, Phillips EM (2001) Infrared thermal imaging as a tool in pain management an 11 year study: I. *Thermol Int* 11:53–65
6. Hooshmand H, Hashmi M, Phillips EM (2001) Infrared thermal imaging as a tool in pain management an 11 year study: II. Clinical applications. *Thermol Int* 11:119–129
7. Merla A, Romani GL, Tangherlini A, Romualdo SD, Proietti M, Rosato E, Aversa A, Salsano F (2007) Penile cutaneous temperature in systemic sclerosis: a thermal imaging study. *Int J Immunopathol Pharmacol* 20:139–144
8. Ng EYK, Chen Y (2006) Segmentation of breast thermogram: improved boundary detection with modified snake algorithm. *J Mech Med Biol* 6:123–136

9. Ng EYK (2008) A review of thermography as promising non-invasive detection modality for breast tumor. *Int J Therm Sci* 48:849–859
10. Helmy A, Holdmann M, Rizkalla M (2008) Application of thermography for non-invasive diagnosis of thyroid gland disease. *IEEE Trans Biomed Eng* 55:1168–1175
11. Tan JH, Ng EYK, Acharya UR, Chee C (2009) Infrared thermography on ocular surface temperature: a review. *Infrared Phys Technol* 52:97–108
12. Kumar S, Acharya S, Beuerman R, Palkama A (2006) Numerical solution of ocular fluid dynamics in a rabbit eye: parametric effects. *Ann Biomed Eng* 34:530–544
13. Braun RJ, Fitt AD (2003) Modelling drainage of the precorneal tear film after a blink. *Math Med Biol* 20:1–28
14. Bron AJ, Tiffany JM, Gouveia SM, Yokoi N, Voon LW (2004) Functional aspects of the tear film lipid layer. *Exp Eye Res* 78:347–60
15. Ng EYK, Ooi EH (2007) Ocular surface temperature: a 3D FEM prediction using bioheat equation. *Comput Biol Med* 37:829–835
16. Rosenbluth RF, Fatt I (1977) Temperature measurements in the eye. *Exp Eye Res* 25:325–341
17. Efron N, Young G, Brennan N (1989) Ocular surface temperature. *Curr Eye Res* 8:901–906
18. Tan L, Cai Z-Q, Lai N-S (2009) Accuracy and sensitivity of the dynamic ocular thermography and inter-subjects ocular surface temperature (OST) in Chinese young adults. *Cont Lens Anterior Eye* 32:78–83
19. Purslow C, Wolffsohn JS (2007) The relation between physical properties of the anterior eye and ocular surface temperature. *Optom Vis Sci* 84:197–201
20. Murphy PJ, Morgan PB, Patel S, Marshall J (1999) Corneal surface temperature change as the mode of stimulation of the non-contact corneal aesthesiometer. *Cornea* 18:333–342
21. Morgan PB, Tullo AB, Efron N (1995) Infrared thermography of the tear film in dry eye. *Eye* 9:615–618
22. Chang T-C, Hsiao Y-L, Liao S-L (2008) Application of digital infrared thermal imaging in determining inflammatory state and follow-up effect of methylprednisolone pulse therapy in patients with Graves' ophthalmopathy. *Graefes Arch Clin Exp Ophthalmol* 246:45–49
23. Galassi F, Giambene B, Corvi A, Falaschi G (2007) Evaluation of ocular surface temperature and retrobulbar haemodynamics by infrared thermography and colour Doppler imaging in patients with glaucoma. *Br J Ophthalmol* 91:878–881
24. Nepp J, Tsubota K, Goto E, Schauersberger J, Schild G, Jandrasits K, Abela C, Wedrich A (2002) The effect of acupuncture on the temperature of the ocular surface in conjunctivitis sicca measured by non-contact thermography: preliminary results. *Adv Exp Med Biol* 506:723–726
25. Chiang HK, ChenCY, Cheng HY, Chen K-H, Chang DO (2006) Development of infrared thermal imager for dry eye diagnosis. In: Proceedings of SPIE – the international society for optical engineering, San Diego, CA, USA
26. Wolberg G (1990) Digital image warping. IEEE Computer Society Press, New York, NY
27. Beier T, Neely S (1992) Feature-based image metamorphosis. *Comput Graph* 26:35–42
28. Tan JH, Ng EYK, Acharya UR (2010) Evaluation of tear evaporation from ocular surface by functional infrared thermography. *Med Phys* 37:6022
29. Tan JH, EYK Ng, Acharya UR (2010) An efficient automated algorithm to detect ocular surface temperature on sequence of thermograms using snake and target tracing function. *J Med Syst* doi: [10.1007/s10916-010-9552-6](https://doi.org/10.1007/s10916-010-9552-6)
30. Mapstone R (1968) Determinants of corneal temperature. *Br J Ophthalmol* 52:729–741
31. Morgan PB, Soh MP, Efron N, Tullo AB (1993) Potential applications of ocular thermography. *Optom Vis Sci* 70:568–576

## Biography



**Jen-Hong Tan**, is currently an application research engineer in Ngee Ann Polytechnic. He did his Ph.D. in Nanyang Technological University, Singapore, under Asso. Prof. E. Y. K. Ng since 2006. He received his Bachelor of Engineering from Nanyang Technological University. His major interest is in biomedical signal processing.



**Eddie Y.K. Ng** (Researcher-ID: A-1375-2011) received Ph.D. at Cambridge University, UK. His main interest is thermal imaging and physiology, breast cancer detection, biomedical engineering; computational turbomachinery aerodynamics, and CFD-CHT. He is associate professor in NTU and published more than 300 technical articles and six books. He is the Editor-in-Chief for both Journal of Mechanics in Medicine and Biology & Journal of Medical Imaging and Health Informatics; as well as assoc. editor for various international journals such as: WJCO, IJBC, CMJ, CFDJ, IJRM, OMIIJ, ONMJ, etc.



**Rajendra Acharya U**, Ph.D., D.Eng. is a visiting faculty in Ngee Ann Polytechnic, Singapore and adjunct faculty in Manipal Institute of Technology, Manipal, India. He received his Ph.D. from *National Institute of Technology Karnataka, Surathkal, India* and D.Engg. from *Chiba University, Japan*. He has published more than 120 papers in refereed international journals (94), international conference proceedings (30), and books (9) over the last years. Currently, he is on the editorial board of the following journals: Journal of Medical Imaging and Health Informatics, Open Access Medical Informatics Journal, Open Access Artificial Intelligence Journal, Journal of Biomedical Nanotechnology, and Journal of Mechanics in Medicine and Biology and Biomedical Imaging and Intervention Journal. He served as Guest Editor for many journals.



**Jasjit S. Suri**, Ph.D. is an innovator, scientist, a visionary, an industrialist, and an internationally known world leader in Biomedical Engineering/Imaging Sciences and its Management and has spent over 20 years in this field. He received his Doctorate from the *University of Washington, Seattle* and MBA from *Ivy League Weatherhead School of Mgt., Case Western Reserve University, Cleveland, OH*. Dr. Suri has published over 300 technical articles. Dr. Suri was crowned with the President's Gold Medal in 1980 and elected as a Fellow of the American Institute for Medical and Biological Engineering and won over 50 awards during his career.

# Chapter 13

## Segmentation-Assisted Registration for Brain MR Images

Jundong Liu

**Abstract** Tracking brain morphological changes in magnetic resonance images (MRIs) requires high accurate segmentation or registration of brain structures. In this chapter, we propose a robust solution for this problem, based on a unified variation framework where prior segmentation information can be seamlessly integrated into a nonrigid registration procedure. Under this framework, in addition to the force arising from the similarity minimization in seeking for detailed correspondence, atlas contours provide an extra guidance to assist the alignment procedure in achieving a more meaningful, stable, and noise-tolerant result. Sum of local correlation (LC) is used as the underlying similarity metric. Our approach can very well handle intensity scaling, contrast reversal, as well as image noise. Experimental results on 2D/3D synthetic and real data demonstrate the improvement made by our algorithm.

### 13.1 Introduction

Several recent studies show that brain atrophy is a normal physiological process observable in a number of brain diseases, including Alzheimer's disease (AD), frontotemporal dementia (FTD), vascular dementia (VaD), normal pressure hydrocephalus (NPH), primary lateral sclerosis (PLS), and multiple sclerosis (MS). Accurate assessment of subtle longitudinal brain change in morphology and volume therefore can provide an objective measure of the pathology for these diseases.

Currently, the solutions used for measuring brain atrophy in patients with MS or AD can be classified into two groups: segmentation-based and registration-based methods. Examples of segmentation-based methods used to measure brain volumes include fuzzy connectedness [30], the Alfano method [1] and SIENAX [40]. Brain atrophy is estimated based on calculating the volume difference of the segmented areas/tissues in serial MRI images. As spatial correspondence between the image

---

J. Liu

School of Electrical Engineering and Computer Science, Ohio University, Athens, OH 45701, USA  
e-mail: liu@cs.ohio.edu

components is not possible to be retrieved via the segmentation procedure, these methods can only capture the volumetric, but not the morphological changes occurred in patients' brains. Registration-based methods, however, can catch both morphological and volumetric changes through an accurate spatial position matching followed by digital subtraction, and therefore are more desirable for detailed brain analysis.

## 13.2 Automated Medical Image Registration Methods

Generally speaking, image registration methods can be classified into two main classes [29], namely, feature-based and direct methods. In the former, features (e.g., contours or surfaces) are extracted from the images and matched to estimate the transformation, and hence the accuracy of registration is dictated by the accuracy of the feature detector. In the latter, the coordinate transformation between the images are determined directly from the image data or a derived image-like representation of the same.

### 13.2.1 Feature-Based Methods

Pellizari et al. [23] present a method for the registration of brain surfaces identified from two different data sets by minimizing the distance between the surfaces measured from a single central point. The surfaces were formed by delineating the external skin surface in each slice of the MR and PET scans. The minimization is carried out numerically using Powell's method. In order to find the correspondence in each iteration, they made an assumption that the two surfaces are predominantly spherical. The task of extracting corresponding shapes in their method is quiet easy, but the accuracy of the rigid alignment is not guaranteed due to the skin surface deformation and so is the robustness of the matching due to the symmetries and lack of the fine structure in the skin surface.

Evans et al. [28] developed a registration scheme based on approximating the 3D warp between the model and target image by a 3D thin plate spline fitted to landmarks. This allowed them to deal with smooth deformations between the corresponding landmarks in the two data sets. Szeliski et al. [16] developed a fast matching algorithm for registration of 3D anatomical surfaces found in 3D medical image data. They use a distance metric minimization to find the optimal transformation between the surfaces. The key difference in their scheme from earlier methods is the use of tensor product splines to represent different registration transformations, namely rigid, affine, trilinear, quadratic, etc. In addition, they introduced the novel concept of octree splines for very fast computation of distance between surfaces. Davatzikos et al. [12] introduced a two-stage brain image registration algorithm. The first stage involved using active contours to establish

a one-to-one mapping between cortical and ventricular boundaries in the brain, and in the second stage, an elastic deformation transformation that determines the best correspondence between the identified contours in the two data sets is determined. Feldmar et al. developed a novel surface to surface nonrigid registration scheme, using local affine transformations. For more on feature-based methods, see the survey by Maintz et al. [18]. *All of these approaches have one commonality, i.e., they need to detect features/surface/contours in the image hence the accuracy of registration is dictated by the accuracy of the feature detector.*

### 13.2.2 Direct Methods

Direct methods subsume the approaches operating directly on the image grey values, without prior feature extraction. One straightforward approach is the optical flow formulation, in which the problem of registering two images is treated as equivalent to computing the flow between the data sets. There are numerous techniques for computing the optical flow from a pair of images [2, 15]. Most of the direct methods for image registration reported in literature use a region-based optical flow methods described in literature [2].

Bajscy and Kovacic [3] studied registration under nonrigid deformations using volumetric deformations based on elasticity theory of solids. Other direct approaches have used the so-called fluid registration approach introduced by Christensen et al. [8]. In this approach, the registration transformation is modeled by a viscous fluid flow model expressed as a nonlinear PDE. In all the fluid-based registration methods, the reported computational times for registration of 3D data sets are very large for implementation on uniprocessor workstations.

In Thirion [13], an optical flow-based nonrigid registration method is described, which conceptualizes image matching in a manner similar to the nineteenth century thought experiment of James C. Maxwell. Given two image to register, the demons algorithm takes the object boundaries in the fixed image as semi-permeable membrane and to let the floating image diffuse through these interfaces. Three nonrigid matching algorithms were derived based on different diffusion models, one using all the intensity levels in the fixed image, one using only contour points, and the last one operating on already segmented images. Due to the simplicity of its formulation as well as its success in achieving a fast and accurate registration, the demons algorithm has gained tremendous popularity in recent years. A similar intensity-based deformable registration was introduced by Chen et al. [7], which consists of three hierarchical levels. A global transformation estimation is first applied to correct the image acquisition variations, followed by a smooth deformation to approximately match the corresponding anatomical structures in the image volumes. The smooth deformation was computed at the control points only and interpolated elsewhere using B-spline basis. Finally, a fine-tuning deformation, which allows each voxel to move independently,

refines the matching of corresponding anatomical features. Sum of squared difference (SSD) is used as the matching metric. No prior segmentation is assumed in this work.

In [33], Vemuri et al. proposed a curve evolution approach which can achieve both image intensity morphing and nonrigid image registration. Atlas-based segmentation of hippocampal shape from several MR brain scan was shown as an application of the nonrigid registration scheme. Although this approach has many attractive features, such as simplicity and high efficiency, the available atlas information, however, has essentially been neglected in the coordinate warping procedure.

The main problem associated with above-mentioned nonrigid registration methods is that they are all based on a *brightness constancy hypothesis*, which assumes that the intensity values at the corresponding voxels of the two data sets are identical. Due to a variety of causes, such as MR bias field, patient motion effects, and partial volume artifacts, this assumption is commonly violated in longitudinal brain MRI data.

### 13.2.3 Intensity-Variation-Tolerant Measures

A group of similarity measures that do not assume *brightness constancy* are based on the concept of maximizing mutual information reported in Viola and Wells [34], Collignon et al. [9] and Studholme et al. [27]. One noticeable drawback of MI-based registration is its proneness of being trapped in local maxima, as no spatial information has been taken into consideration in the original formulation [34, 27]. Although some studies, such as combining MI with gradient information [24] and using regional MI [11], showed impressive improvements for boosting the robustness for MI-based elastic registration, no systematic solution has been proposed up to date.

Recently, the *sum of Local Correlation (LC)* [36, 21], derived from image statistics (mean and variance), started to show great successfulness in registering mono-modal and multi-modal images with impressive accuracy and efficiency. Due to the fact that it can capture both the statistical correspondence and the spatial coherence existing in the input images, LC is fairly responsive to local spatial changes, and therefore has a great potential to reveal very detailed nonrigid motions. In addition, LC can be formulated [36] to be relatively invariant to intensity scaling and reversal, which makes it an ideal similarity metric for handling the registration problems where “subtle spatial changes” and “presence of intensity variations” are two major characteristics. For more on other direct methods, we refer the reader to the recent survey by Maintz et al. [19].

We should note that, though the above-mentioned direct methods have enjoyed great popularity in solving the rigid or nonrigid registration problems, none of them has attempted to integrate any prior segmentation information, which is often available for a variety of applications including brain atrophy estimation and atlas-based brain segmentation, into the registration procedure. This negligence can be the cause of the failures for registering certain images, especially when the inputs are noisy.

### 13.2.4 Joint Segmentation + Registration Models

Registration and segmentation are the two most fundamental problems in the field of medical image analysis. Traditionally, they were treated as separate problems, each with numerous solutions proposed in literature. In recent years, the notion of integrating segmentation and registration into a unified procedure has gained great popularity, partially due to that more and more practical problems, e.g., atlas-based segmentation, subsume both segmentation and registration components.

In [35], Wang et al. proposed two atlas-based nonrigid registration methods that use the combined power of physical and statistical shape models. Two physical models, deformable elastic solids in the first method and viscous fluids in the second, are used to maintain smoothness and continuity of the transformation field. In addition to the dense set of forces generated by the intensity similarity measure to accommodate complex anatomical details, a sparse set of forces derived from the statistical shapes of a training set is used to lead a smoother and more consistent registration process. A number of experiments performed on both synthetic and real images show the beneficial effects of the statistical boundary shape information in improving registration robustness. However, it is foreseeable that these methods would produce less smooth results for noisy inputs if the boundary points are not sampled densely enough.

Vemuri et al. [41] propose a segmentation + registration model to solve the atlas-based image segmentation problem where target image is segmented through the registration of the atlas to the target. A novel variational formulation was presented, which put segmentation and registration processes under a unified variational framework. Optimization is achieved by solving a coupled set of nonlinear PDEs.

Another segmentation + registration model proposed by Noble et al. [22] seeks for the best possible segmentation and registration from the *maximum a posteriori* point of view. Improvements in accuracy and robustness for both registration and segmentation have been shown, and potential applications were identified. This model is primarily designed for combining segmentation and *rigid* registration. While nonrigid algorithm was also implemented, the motion field estimation is based on block-matching of size  $(7 \times 7)$ , which is not dense enough for most nonrigid registration applications.

Freedman et al. [14] proposed a tracking method where prior segmentation is also successfully incorporated. Under this system, tracking decision is based on a density matching between the intensity distribution of an estimation image region and a model distribution, and prior segmentation is integrated into the density matching procedure in the form of a series of level set. The performance of resulting tracker shows the helpfulness of prior segmentation for maintaining tracking success rate in complicated scenario. We should note that this model does not produce or rely on dense motion field for making tracking decisions.

Inspired by the above-mentioned approaches, the work presented in this chapter is aimed to establish a segmentation-assisted framework to boost the robustness of

nonrigid image registration. Another component of our method, aiming to achieve the same goal, is the choosing of the *sum of Local Correlation (LC)* as the underlying similarity measure. LC measure is invariant to intensity scaling and contrast, and this property is very crucial for registration applications where input data have substantial intensity variations. The following is a brief review on some of the related similarity measures that can handle intensity variations.

### 13.2.5 Proposed Registration Method

In this chapter, we propose to develop a robust segmentation-guided registration framework, with LC as the underlying matching metric. Under this framework, prior shape information can be fully integrated into the registration procedure, as an extra guiding force to lead a more meaningful, stable, and noise-tolerant image alignment process.

Our segmentation-guided model is inspired by the works mentioned in the previous section, and it differs from other models in that (1) our method is a fully nonrigid dense deformation estimation model; (2) it uses a unified segmentation + registration energy minimization formulation; and (3) the optimization is carried out under a natural, parameterization-free, and numerically stable level set framework. *A salient feature of our model is its robustness against input image noise.*

With different similarity measure embedded into our framework, our model can handle both single-modality and multi-modality image registrations. In this chapter, we propose a modified LC as the underlying similarity metric. Comparing with the existing LC formulations, our LC formula has the advantage (over [21]) of being able to handle both global and local motion, and advantage (over [36]) of being able to handle both intensity scaling and reversal.

The outline of the paper is as follows. In the next section, we introduce our segmentation guided registration model as an energy minimization and formulate it under level set framework. Associated Euler-Lagrange equations will be provided and discussed. In Sect. 13.3, we give experiment results to demonstrate the performance of our algorithm on several 2D images. We conclude this paper in Sect. 13.4.

## 13.3 Segmentation-Guided Registration Model

Commonly, the basic input data to a registration process are two images: one is defined as *fixed* (or *target*) image  $I_1(X)$  and the other as the *moving* (or *source*) image  $I_2(X)$ . A typical solution to the nonrigid registration problem is to look for a deformation function  $V$  assigned to each point  $X$ . The function is searched by minimizing an energy function  $E$  of the form

$$E(V) = S(V) + R(V) \quad (13.1)$$

The term  $S(V)$  is designed to measure the dissimilarity between the input image  $I_1$  and  $I_2$ . The term  $R(V)$  is designed to penalize fast variations of the deformation function  $V$ .

In addition to these two image, our model requires a segmentation of the fixed image, indicating a studying area of  $I_1(X)$ , as another input component. Let  $C$  be the boundary curve of the segmentation. We denote by  $C_{\text{in}}$  and  $C_{\text{out}}$  representing the inside and outside areas of the curve  $C$ . Let  $C_1$  and  $C_2$  be the average values for  $C_{\text{in}}$  and  $C_{\text{out}}$ , respectively.

The contour  $C$  can be either input by user or derived from a training set. We assume that the region captured by  $C$  contains a single object of the fixed image; therefore the intensity profiles of both inside and outside of the region should be able to be characterized by certain property. Examples of the property include “being relatively homogenous”, or “conforming to certain distribution”. Suppose the fixed and moving images are well corresponded, then, at the time a perfect alignment is achieved, the intensities in the warped moving image should also have a similar property within both  $C_{\text{in}}$  and  $C_{\text{out}}$ . This observation provides the justification for our model, which is designed based on the following considerations:

- In addition to the set of forces generated by intensity similarity measure (e.g., SSD, LC or MI) to warp the moving image toward the target, another set of forces, derived from the region property constraint, should be utilized to pull the moving image toward the correct alignment. This set of forces can provide an extra guidance for the registration process to avoid local energy optima, which is especially helpful when input images are noisy.

Our solution to the segmentation-guided registration can be formulated as the minimization of a new energy, which integrates the available segmentation information:

$$E(V) = S(V) + H(V) + R(V) \quad (13.2)$$

where  $H(V)$  is designed to penalize the deformations that would result in inhomogeneous intensity profiles within  $C_{\text{in}}$  and  $C_{\text{out}}$ .

For different applications, we can assign different forms for the terms  $S(V)$ ,  $H(V)$ , and  $R(V)$ .  $S(V)$  term can take sum of squared difference (SSD) for single-modality applications and mutual information (MI) and correlation ratio (CR) for multimodal registration. For  $H(V)$ , piecewise constant function [6] and piecewise linear function [26] are among the most popular choices. Gaussian diffusion model, elastic model, and viscous fluid model have been widely used as the regularization options for term  $R(V)$ .

### 13.3.1 Frameworks Based on Intensity Homogeneity

To handle the intensity variations existing in the input images, our solution to the robust segmentation-guided registration is formulated as the minimization an

energy, which relies on LC to measure the image similarity, with the available segmentation information being used as a homogeneity constraint,

$$\begin{aligned}
 E(V) = & \int_{\Omega} LC(I_1(X), I_2(X + V(X))) dX \\
 & + \lambda_1 \int_{C_{in}} [I_2(X + V(X)) - C_1]^2 dX \\
 & + \lambda_2 \int_{C_{out}} [I_2(X + V(X)) - C_2]^2 dX \\
 & + \lambda_3 \int_{\Omega} |\nabla V(X)|^2 dX
 \end{aligned} \tag{13.3}$$

where  $LC(I_1, I_2)$  is the LC similarity between  $I_1$  and transformed  $I_2$ .  $\Omega$  is the image domain, and  $V(X)$  denotes the deformation field.  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are three constant parameters that weight the importance of each term in the optimization energy.

Several variations of LC measures have been investigated in Cashier et al. [5] and Netsch et al. [21]. In our implementation, we use a different customized form of local correlation and it has the advantage (over [21]) of being able to handle both global and local motion, and advantage (over [36]) of being able to handle both intensity scaling and reversal. We formulate the local correlation measure as follows:

$$LC(I_1, I_2) = \sum_b \frac{\sum_{a \in n(b)} (i_{1a} - \bar{i}_{1b})^2 (i_{2a} - \bar{i}_{2b})^2}{\left[ \sum_{a \in n(b)} (i_{1a} - \bar{i}_{1b})^2 \sum_{a \in n(b)} (i_{2a} - \bar{i}_{2b})^2 \right]}$$

where  $i_1 = I_1(X)$  and  $i_2 = I_2(X + V(X))$  with  $a$  representing the pixels in the neighborhood  $n(b)$  around pixel  $b$  in the image. In the energy function  $E(V)$ , the first term  $LC(I_1, I_2)$  in the energy function provides the main force for matching two images, while  $\int_{C_{in}} [I_2(X + V(X)) - C_1]^2 dX$  and  $\int_{C_{out}} [I_2(X + V(X)) - C_2]^2 dX$  terms allow the a priori segmentation to exert its influence, aiming to enforce the homogeneity constraints.  $\int_{\Omega} |\nabla V(X)|^2 dX$  is a diffusion term to smooth the deformation field.

### 13.3.2 Level Set Formulation of the LC-Based Model

The energy function  $E(V)$  can be minimized under the level set framework. Introduce a continuous function  $\phi : \Omega \rightarrow R$ , so  $C = \{(X) \in \Omega : \phi(X) = 0\}$ , and we choose  $\phi$  to be positive in  $C_{in}$  and negative in  $C_{out}$ . We adopt the model presented in Chan et al. [6] and we have the following functional:

$$\begin{aligned}
E(V) = & \int_{\Omega} LC(I_1(X), I_2(X + V(X))) dX \\
& + \lambda_1 \int_{\Omega} [I_2(X + V(X)) - C_1]^2 H(\phi(X)) dX \\
& + \lambda_2 \int_{\Omega} [I_2(X + V(X)) - C_2]^2 (1 - H(\phi(X))) dX \\
& + \lambda_3 \int_{\Omega} |\nabla V(X)|^2 dX
\end{aligned} \tag{13.4}$$

where  $H$  is the Heaviside function. The Euler–Lagrange differential equation of this functional is given by:

$$\begin{aligned}
\frac{dE}{dV} = & \frac{d(LC)}{dV} \\
& + 2\lambda_1(I_2(X + V) - C_1) \nabla I_2(X + V) \cdot H(\phi(X)) \\
& + 2\lambda_2(I_2(X + V) - C_2) \nabla I_2(X + V) \cdot (1 - H(\phi(X))) \\
& + \lambda_3 \nabla^2 V
\end{aligned}$$

where

$$\begin{aligned}
\frac{d(LC)}{dV} = & \sum_b \frac{2}{\sum_{a \in n(b)} (i_{1a} - \bar{i}_{1b})^2} \left[ \frac{\sum_{a \in n(b)} (i_{1a} - \bar{i}_{1b})^2 (i_{2a} - \bar{i}_{2b})}{\sum_{a \in n(b)} (i_{2a} - \bar{i}_{2b})^2} \right. \\
& \left. - \frac{\sum_{a \in n(b)} (i_{1a} - \bar{i}_{1b})^2 (i_{2a} - \bar{i}_{2b})^2 \sum_{a \in n(b)} (i_{2a} - \bar{i}_{2b})}{\sum_{a \in n(b)} (i_{2a} - \bar{i}_{2b})^2} \right] \nabla i_2
\end{aligned}$$

and

$$\begin{aligned}
C_1 &= \frac{\int_{\Omega} I_2(X + V) H(\phi(X + V)) dX}{\int_{\Omega} H(\phi(X + V)) dx dy} \\
C_2 &= \frac{\int_{\Omega} I_2(X + V) (1 - H(\phi(X + V))) dX}{\int_{\Omega} (1 - H(\phi(X + V))) dx dy}
\end{aligned}$$

To estimate the deformation field between  $I_1$  and  $I_2$ , we initialize the deformation field as  $X(V) = 0$  at each pixel, and use the Euler equation as a gradient descent process that eventually leads to the convergence of the alignment process. The level set function being used in this chapter is  $\phi(X, 0) = D(X)$ , where  $D(X)$  is the signed distance from each grid point to the zero level set  $C$ . This procedure is standard, and we refer the reader to [39] for details.

### 13.3.3 A SSD-Based Segmentation + Registration Model

For comparison purpose, we also provide the SSD-based segmentation-guided model, which is to minimize the following energy:

$$E(V) = \int_{\Omega} [I_1(X) - I_2(X + V(X))]^2 dX + H(V) + R(V)$$

where the homogeneity and regularization parts are identical to those in (13.4).

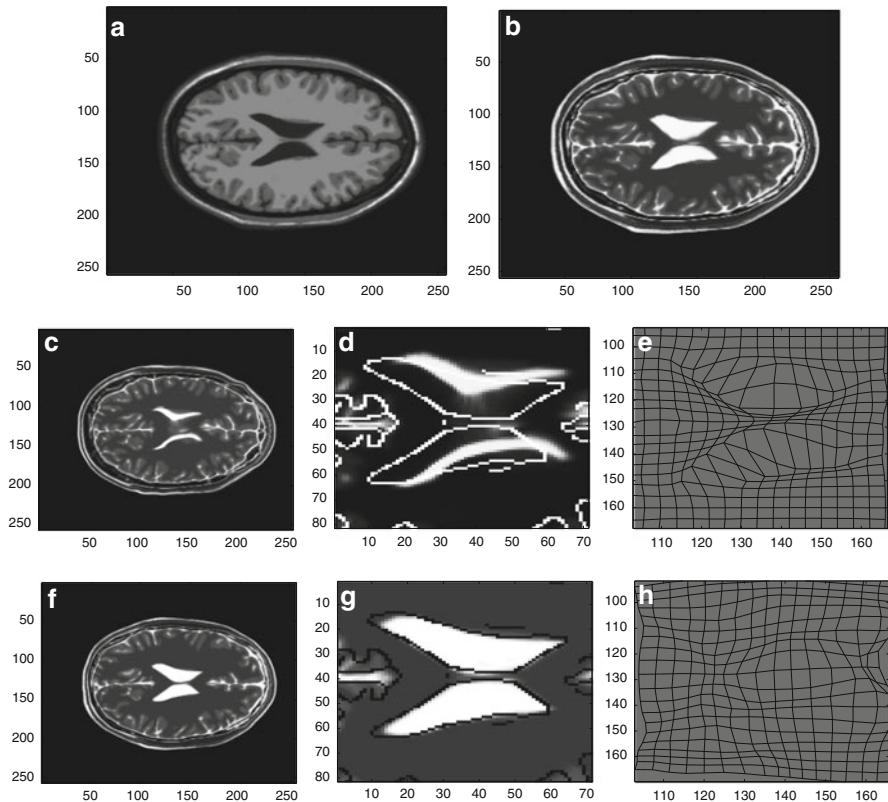
## 13.4 Experimental Results

In this section, we present three sets of experiments to demonstrate the improvement made by our proposed registration method. To test the functionality of the two components of our approach: *local correlation* and segmentation-guide framework, we first compare *local correlation* with a *brightness constancy assumption* measure – SSD to demonstrate the capability of LC in handling intensity variations. We then compare our “segmentation + registration” model with a popular “registration-only” *Demons* algorithm to show the usefulness of integrating segmentation information in leading to a stable and noise-tolerant registration process. The performance of our complete *segmentation + registration + LC* model is evaluated in the third experiment.

### 13.4.1 Registrations Based on the Local Correlation Component

One example is used to demonstrate the robustness of the sum of *local correlation* in handing intensity variations. SSD is used as a comparison metric. The specific matching algorithm being used is the famous *Demons* algorithm, which is based on SSD. The data used in this example are obtained from the Montreal Brainweb reference database [1]. We extract a pair of T1/T2 slices, where substantial disparity exists in shape of the ventricle area. Figure 13.1 shows the images and results.

Figure 13.1a, b shows the fixed and moving images, respectively. The second and third rows depict the results from the SSD model (Fig. 13.1c) (*Demons* algorithm) and that from our LC model (Fig. 13.1f), respectively. Figure 13.1d, g is the corresponding zoom-in versions. The estimated deformation fields are displayed in Fig. 13.1e, h. In order to test the functionality of LC, we turned off the effect of the segmentation input (by setting the weighting factor of the segmentation component to zero) in our model. The advantage of LC measure over SSD is its ability in handling intensity variations, especially intensity scaling and reversal, which are commonly existing in T1/T2 image pairs. As evident, *Demons* algorithm,



**Fig. 13.1** Comparison between SSD and LC. First row: (a) the fixed image and (b) the moving image. Second row: registration result (c) using SSD model (*Demons* algorithm). (d) is the zoom-in version of (c), and (e) is the estimated deformation field. Third row: the registration result (f), zoom-in version (g) and the deformation field (h) of using our LC model

due to its reliance on the *brightness constancy assumption*, failed to handle the intensity reversal, transforming the ventricle area into a totally wrong direction, while our LC model can push the ventricle in the moving image (white area in Fig. 13.1d, g) into an accurate alignment.

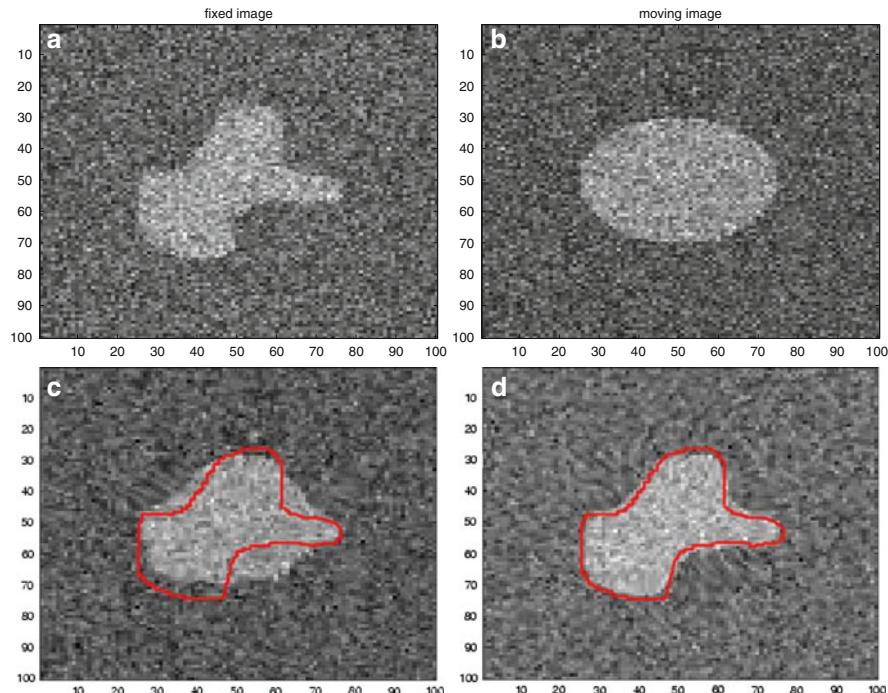
### 13.4.2 Registration Based on the Segmentation + Registration Component

Three examples are used to test the segmentation + registration component. In all cases, we still compare the results using our model with that of using the *Demons* algorithm. This time, *Demons* algorithm is counted as a representative of those “registration-only” approaches. Considering that the *Demons* algorithm is

a SSD-based method, to make the comparison more meaningful, our model being used for these three examples is a *segmentation + registration + SSD* version. The goal of these examples is to demonstrate the helpfulness of integrating segmentation information into the registration procedure, especially in handling image noise.

The first example contains a pair of synthetically generated images, where the fixed image was generated from the moving by a known nonrigid field. Zero-mean Gaussian noise was then added to each image. The standard deviation is 20. Figure 13.1a, b shows the two images. In the following examples, we chose the constants  $\lambda_1 = 0.1$ ,  $\lambda_2 = 0.1$ , and  $\lambda_3 = 1$ , respectively. The segmentation of the fixed image was manually obtained, as superimposed on the moving image in Fig. 13.2c. Two registration approaches, the *Demons algorithm* and our *segmentation-guided registration* model, are then applied. We should note that the *Demons* algorithm relies on intensity alone for registration.

The results are shown in the second row of Fig. 13.2. Figure 13.2c is the transformed moving image from the *Demons* algorithm, after the registration is finished. Figure 13.2d shows the result of our model. As evident, the *Demons* algorithm had trouble in warping the moving image to a perfect matching, which is partially due to the numerous local energy minima resulted from the huge amount

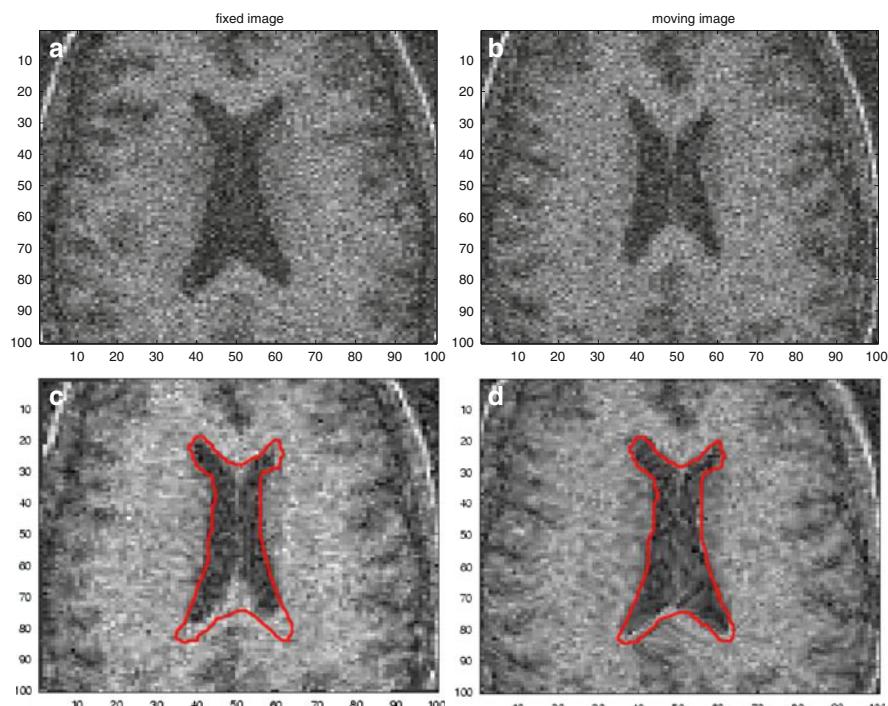


**Fig. 13.2** Registration results for a pair of synthetic images. First row: (a) the fixed image and (b) the moving image. Second row: the registration result of (c) using the *Demons* algorithm, and (d) using our segmentation guided registration model. The edge map from the fixed image is superimposed

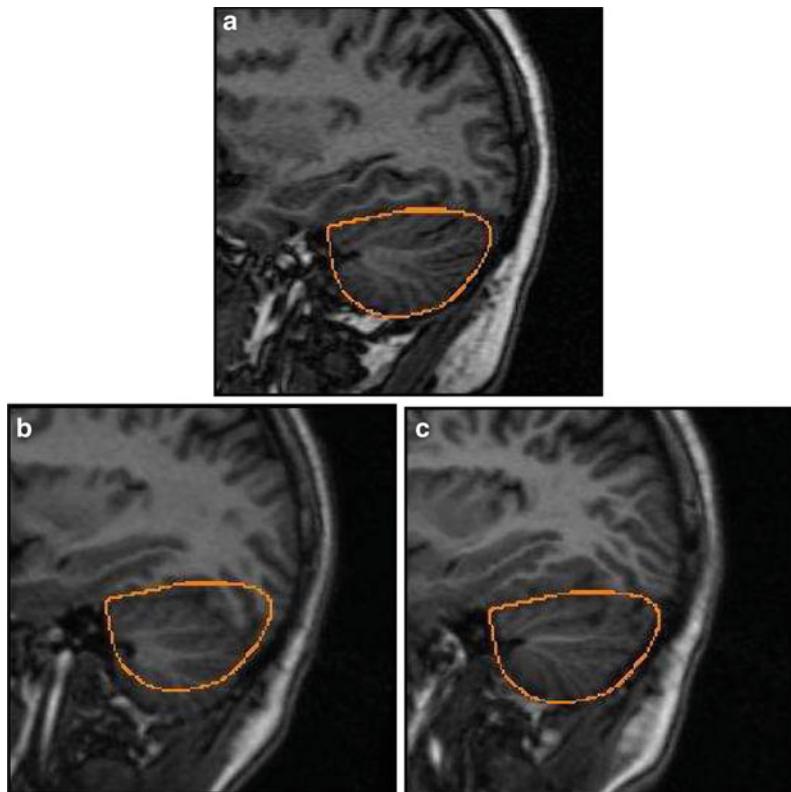
of noise existing in the images. However, the registration result generated from our model is quite accurate, which indicates that the integrated segmentation information is very helpful in pulling the moving image toward a correct matching.

We designed and carried out a similar experiment on a pair of MRI brain slices. The two slices have substantial disparity in shape of the ventricles, which is the region of interest. Figure 13.3 shows the images and results. Figure 13.3a, b shows the fixed and moving images, respectively. Figure 13.3c, d depicts the results from the *Demons* algorithm and our segmentation-guided registration model, respectively. As evident, the former model fails to transform the ventricle area into a desired position, while the latter accurately achieves the registration goal.

A 3D MRI experiment is conducted with similar fashion. The cerebellum area of the fixed image is the segmented. Figure 13.4 shows the images and results. Figure 13.4a is the fixed image, with the boundary contours of the available segmentation superimposed. There is quite amount of visible difference within the segmented cerebellum area. Figure 13.4b, c depicts the results from the *Demons* algorithm and our segmentation-guided registration model, respectively. As evident, our model can lift the cerebellum area to match the ROI of the fixed image.



**Fig. 13.3** Registration results for a pair of 2D MR images. First row: (a) the fixed image and (b) the moving image. Second row: the registration result of (c) using the *Demons* algorithm, and (d) using our segmentation-guided registration model

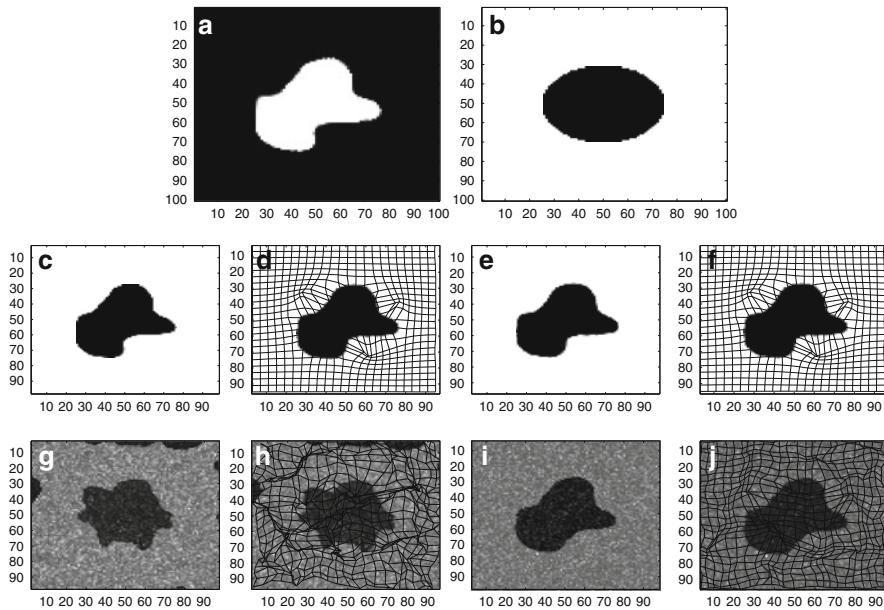


**Fig. 13.4** Registration results for a pair of 3D MR images. First row: (a) the fixed image, superimposed with the available segmentation contour. Second row: (b) the registration result of using the *Demons* algorithm, and (c) using our segmentation-guided registration model. The edge map from the segmentation is superimposed

### 13.4.3 Registrations of Our Segmentation + Registration LC Model

In this section, we demonstrate the performance of our algorithm with synthetic as well as real images. In both cases, we compare the results using our *registration + segmentation LC* model with that of using the *registration-only LC* method, as in Sect. C.1.1.

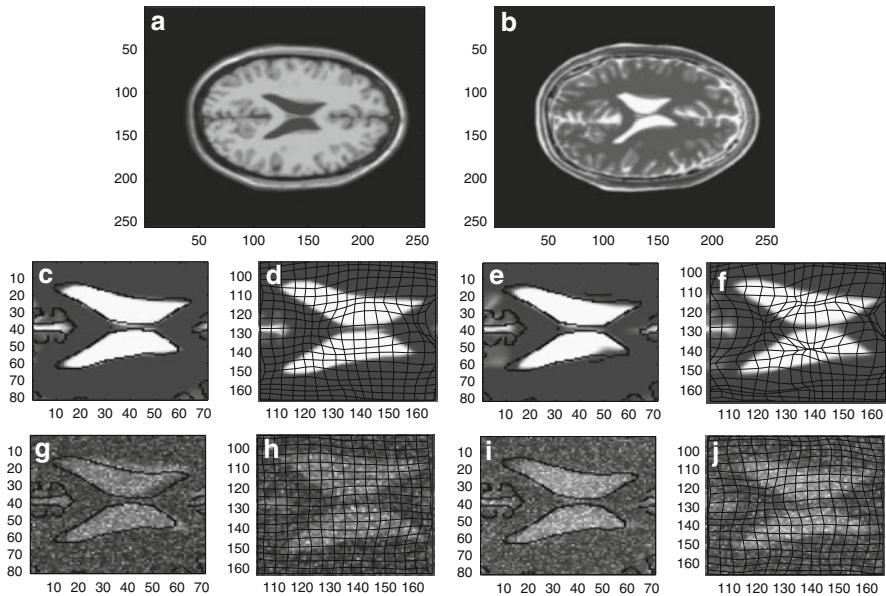
The synthetic data example contains a pair of synthetically generated images, where the fixed image was generated from the moving image (ellipse) by applying a known nonrigid field. In order to demonstrate LC's ability of handling multimodality images, we intentionally set the pixel values of the two images into a contrast (i.e., black vs. white and white vs. black) where similarity measures such as



**Fig. 13.5** Registration results for a synthetic image set. First row: (a) and (b) are the fixed and moving images respectively. Second row: (c) is the deformed moving image using the *registration only LC* model and (d) is the estimated deformation field from (c). (e) is the deformed moving image using our *registration + segmentation LC* model and (f) is the deformation field for (e). Third row: same experiment with noisy image inputs. Results from the *registration only* (g and h) and our *registration + segmentation LC* model (i and j)

SSD will fail. Figure 13.5a, b shows the fixed image and moving image, respectively. Results from the *registration only LC* and our *registration + segmentation LC* models are shown in Fig. 13.5c, e. The estimated deformation fields are shown in Fig. 13.5d, f. As evident, both models can obtain fairly accurate matching. To demonstrate the ability of the *segmentation + registration* component in handling noisy image data, we applied a zero-mean Gaussian noise with standard deviation of 10 onto both input images. Figure 13.5g, i shows the results from the two models. It is clearly shown that, in contrast to the immediate failure of the *registration only* method, our model can still perfectly find the matching.

A similar experiment is conducted on a pair of T1/T2 brain slices. The region of interest is around the ventricle, where two slices have substantial disparity. Figure 13.6 shows the images and results. Figure 13.6a, b shows the fixed and moving images, respectively. Figure 13.6c, e depicts the results from the *registration only LC* model and the *registration + segmentation LC* model, respectively. Similar to the previous example, *registration-only LC* is fair enough to handle this clean image pair. But when noise is present, the *segmentation + registration LC* model starts showing its robustness property over the other model, as shown in the third row of Fig. 13.6. A zero-mean Gaussian noise with standard deviation of



**Fig. 13.6** Registration results for a pair of T1/T2 images. First row: (a) the fixed image and (b) the moving image. Second row: (c) is the deformed moving image using the *registration only LC* model and (d) is the estimated deformation field from (c). (e) is the deformed moving image using our *registration + segmentation LC* model and (f) is the deformation field for (e). Third row: same experiment with noisy image inputs. Results from the *registration only* (g and h) and our *registration + segmentation LC* model (i and j)

10 has been added onto both input images. Figure 13.6g, i shows the results from “registration-only” and “segmentation + registration” models, respectively. It is clearly shown that the former fails to transform the ventricle area into a desired position, while the latter accurately achieves the registration goal.

In summary, our *segmentation + registration LC* model has the desired property of being insensitive to intensity reversal, scaling as well as image noise; therefore, it has the great potential to be used to accurately and robustly capture the volumetrical and morphological brain changes demonstrated in serial MRI images, in which intensity variations are inevitably present.

## 13.5 Conclusions

In this chapter, we present a segmentation-guided nonrigid registration algorithm, which integrates the available prior shape information as an extra forces to lead to a noise-tolerant registration procedure. Our model differs from other methods in that we use a unified segmentation + registration energy minimization formulation, and

the optimization is carried out under level-set framework. *Local Correlation* has been used as the similarity measure to handle intensity variations. We showed the improvement made with our model by comparing the results with that of the Demons algorithm. To explore other similarity metrics under the same framework to handle more complicated inputs will be the focus of our future work.

## References

1. Simulated brain database [Online]. <http://www.bic.mni.mcgill.ca/brainweb/>
2. Alfano B, Quarantelli M, Brunetti A et al (1998) Reproducibility of intracranial volume measurement by unsupervised multispectral brain segmentation. *Magn Reson Med* 39:497–499
3. Bajscy R, Kovacic S (1989) Multiresolution elastic matching. *Comput Vis Graph Image Process* 46(1):1–21
4. Barron JL, Fleet DJ, Beauchemin SS (1994) Performance of optical flow techniques. *Int J Comput Vis* 1(12):43–77
5. Cachier P, Pennec X (2000) Non-rigid registration by gradient descent on a Gaussian-windowed similarity measure using convolutions. In: Proceedings of the IEEE workshop on mathematical methods in biomedical image analysis, p 182
6. Chan T, Vese L An active contour model without edges. *IEEE Trans Image Process* 10 (2):266–277
7. Chen M, Kanade T, Rowley H, Pomerleau D (1998) Anomaly detection through registration. In: IEEE CVPR, pp 304–310
8. Christensen GE, Miller MI, Vannier M (1996) Individualizing neuroanatomical atlases using a massively parallel computer. *IEEE Comput* 1(29):32–38
9. Collignon A, Maes F, Delaere D, Vandermeulen D, Suetens P, Marchal G (1995) Automated multimodality image registration using information theory. In: Proceedings of IPMI, pp 263–274
10. Davatzikos C, Prince JL (1994) Brain image registration based on curve mapping. In: IEEE WBIA, Seattle, WA, pp 245–254
11. Evans AC (1991) Warping of computerized 3D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis. In: Proceedings of SPIE Medical Imaging V, vol 1445:236–246
12. Lai SH, Vemuri BC (1995) Robust and efficient computation of optical flow. In: IEEE Symposium on CVS, Miami, FL, pp 455–460
13. Lavallee S, Szeliski R, Brunie L (1991) Matching 3D smooth surfaces with their 2D projections using 3D distance maps. In: Proceedings of SPIE, vol 1570, pp 322–336
14. Leventon M, Grimson WEL (1999) Multi-modal volume registration using joint intensity distributions. In: MICCAI 1999
15. Maintz JB, Viergever MA (1998) A survey of medical image registration. In: MedIA, vol 2, pp 1–36
16. Maintz JBA, van den Elsen PA, Viergever MA (1996) Comparison of edge-based and ridge-based registration of CT and MR brain images. *Med Image Anal* 1(2):151–161
17. Meyer C et al (1997) Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline-warped geometric deformation. *Med Image Anal* 1:195–206
18. Netsch T et al (2001) Towards real-time multi-modality 3D medical image registration. In: ICCV 2001, pp 718–725
19. Pelizarri CA, Chen GTY, Spelbring DR (1989) Accurate three-dimensional registration of CT, PET and MR images of the brain. *J Comput Assist Tomogr* 13(1):20–26

20. Pluim JP, Maintz JB, Viergever MA (2000) Image registration by maximization of combined mutual information and gradient information. *IEEE Trans Med Imag* 19(8):809–814
21. Roche A, Malandain G, Pennec X, Ayache N (1998) Multimodal image registration by maximization of the correlation ratio. Research Report RR-3378, INRIA, August 1998
22. Rueckert D, Sonoda LI, Hayes C, Hill DLG, Leach MO, Hawkes DJ (1999) Non-rigid registration using free-form deformations: application to breast MR images. *IEEE Trans Med Imag* 18(8):712–721
23. Russakoff DB, Tomasi C, Rohlfing T, Maurer CR Jr (2004) Image similarity using mutual information of regions. In: *ECCV 2004*, LNCS 3023, pp 596–607
24. Studholme C, Hill D, Hawkes D (1999) An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn* 32:71–86
25. Sussman M, Smereka P, Osher S (1994) A level set approach for computing solutions to incompressible two-phase flow. *J Comput Phys* 114:146–159
26. Thirion JP (1998) Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal* 2(3):243–260
27. Tsai A, Yezzi AJ, Willsky A (2000) A curve evolution approach to smoothing and segmentation using the Mumford-Shah functional. In: *CVPR 2000*, pp 1119–1124
28. Udupa JK, Samarasekera S (1996) Fuzzy connectedness and object definition: theory, algorithms and applications in image segmentation. *Graph Models Image Process* 58:246–261
29. Unal G, Slabaugh G (2005) Coupled PDEs for non-rigid registration and segmentation. In: *CVPR 2005*, San Diego, CA
30. Unal G, Slabaugh G, Yezzi A, Tyan J (2004) Joint segmentation and non-rigid registration without shape priors. Siemens Technical Report SCR-04-TR-7495
31. Vemuri BC, Huang S et al. (1998) An efficient motion estimator with application to medical imaging. *Med Image Anal* 2(1):79–98
32. Vemuri B, Chen Y, Wang Z (2002) Registration assisted image smoothing and segmentation. In: *ECCV*, vol 4, pp 546–559
33. Vemuri B, Ye J, Chen Y, Leonard C (2003) Image registration via level-set motion: applications to atlas-based segmentation. *Med Image Anal* 7(1):1–20
34. Viola PA, Wells WM (1995) Alignment by maximization of mutual information. In: *Fifth ICCV*, MIT, Cambridge, MA, pp 16–23
35. Wang Y, Staib LH (2000) Physical model-based non-rigid registration incorporating statistical shape information. *Med Image Anal* 4(1):7–20
36. Weese J, Rosch P, Netsch T, Blaffert T, Quist M (1999) Gray-value based registration of CT and MR images by maximization of local correlation. In: *MICCAI 99*, pp 656–663
37. Wells WM III, Viola P, Atsumi H (1997) Multi-modal volume registration by maximization of mutual information. *Med Image Anal* 1(1):35–51
38. Wyatt PP, Noble JA (2003) MAP MRF joint segmentation and registration of medical images. *Med Image Anal* 7:539–552
39. Yezzi A, Zollei L, Kapur T (2003) A variational framework for joint segmentation and registration. *Med Image Anal* 7:171–185
40. Zhang T, Freedman D (2003) Tracking objects using density matching and shape priors. In: *ICCV 2003*, vol 2, pp 1056–1062
41. Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR image through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging* 20:45–57

## Biography



**Jundong Liu** is currently an Associate Professor in the School of Electrical Engineering and Computer Science at Ohio University. He received the PhD degree in Computer Information Science and Engineering from the University of Florida in 2002. His current research interests include segmentation-assisted registration for medical images, shape representation and analysis for neurodegenerative diseases, and longitudinal studies for Alzheimer's disease (AD).

## About the Editors



**Ayman S. El-Baz** is currently an assistant professor of bioengineering in the Department of Bioengineering at the University of Louisville (UofL). Dr. El-Baz is an expert in the fields of bioimaging modeling and computer-assisted diagnosis systems. He has extensive experience in automated detection, segmentation, and diagnosis of lung nodules in LDCT images. He received his doctorate from the University of Louisville, Louisville, KY.



**Rajendra Acharya U, Ph.D., D.Eng.** is a leader in the field of data mining and medical devices. He received two doctorates: one from National Institute of Technology Karnataka, Surathkal, India and the other from Chiba University, Japan. He is a Senior IEEE member and on the editorial board of several journals. His major interests are in signal and image processing, visualization and biophysics for better healthcare design, delivery, and therapy. Currently, he is a visiting faculty at Ngee Ann Polytechnic, Singapore.



**Andrew F. Laine** is a director of the Heffner Biomedical Imaging Laboratory in the Department of Biomedical Engineering at Columbia University in New York City and is professor of Biomedical Engineering and Radiology (Physics). His research interests include quantitative image analysis; cardiac functional imaging: ultrasound and MRI, retinal imaging, intravascular imaging, and biosignal processing. He is a fellow of AIMBE and IEEE.



**Dr. Jasjit S. Suri** is an innovator, a scientist, a visionary, an industrialist, and an internationally known world leader in Biomedical Engineering. Dr. Suri has spent over 20 years in the field of biomedical engineering/devices and its management. He received his doctorate from the University of Washington, Seattle and Business Management Sciences from Weatherhead, Case Western Reserve University, Cleveland, OH. Dr. Suri was crowned with President's Gold medal in 1980 and was the fellow of American Institute of Medical and Biological Engineering for his outstanding contributions.

# Index

## A

- Abell, F., 164  
Abu El-Ghar, M., 25  
Acharya, R.U., 137  
Acharya, U.R., 315, 319  
Active shape model (ASM), 8  
Adrien, K.T., 164  
Aladl, U.E., 227  
Amini, A.A., 23  
Amoeba optimizer, 239  
Appearance-based registration techniques, 252–253  
Appearance prior model  
    central-symmetric neighborhood sets, 106  
    joint probability distributions, 107  
    relative energies, 108  
Area-based registration techniques, 248  
Artiges, E., 164  
Ashburner, J., 164  
Atkins, M.S., 22  
Atlas-based segmentation  
    multi models, 19–20  
    single models, 18–19  
Audette, M.A., 231  
Autism  
    classifications, 162  
    3-D binary representation, brain, 169  
    early detection, 163  
    flowchart, proposed analysis system, 168  
    Gyral window index (GWI) and volume analysis, 166  
    incidence and prevalence, 162  
    M-CHAT, 164–165  
    mesh generation  
        mesh render, 171  
        procedure, 171–172  
    MRI image, 168  
    neuropathology, 162

- vs. non-autistic brain, 161, 178  
shape analysis, 166–167  
spherical deformation  
    algorithm, 174  
    attraction–repulsion algorithm, 174–176  
Cartesian mesh, 174  
Laplacian smoothing Matlab algorithm, 171–173  
spherical inflations, 176  
spherical harmonic analysis  
    convergence curves, smoothing, 178  
    Euclidean distance, 178  
    probability density functions, error threshold, 179–180  
    reconstruction method, 177–178  
    weighted-SPHARM representation, 176  
three dimensional volume construction  
    procedure, 170  
VBM approach, 164  
Autism spectrum disorder (ASD)  
    M-CHAT, 164  
    prevalence estimates, 162–163  
Automated image registration, brain MR images  
    direct methods, 337–338  
    feature-based methods, 336–337  
    intensity-variation-tolerant measures, 338  
    joint segmentation registration models, 339–340  
    segmentation-guided registration, 340  
Axel, L., 70  
Ayache, N., 277  
Aylward, S., 100

## B

- Baillard, C., 26  
Bajscy, R., 337

- Bakardjieva, A., 164  
 Balloon based model, 208–210  
 Barillot, C., 26  
 Barthelemy, C., 164  
 Batchelor, P.G., 228  
 Bayesian classification statistical model, 14  
 Beach, G.M., 69  
 Beier, T., 319  
 Berman, D.S., 72  
 Bernard, O., 71  
 Bick, U., 120  
 Blaffert, T., 342  
 Bland, P.H., 20  
 Boas, D.A., 171  
 Boddaert, N., 164  
 Bogoni, L., 263  
 Bosch, H., 70  
 Botello, S., 20  
 Bourgeois, M., 164  
 Bousquet, G., 263  
 Boykov, Y., 24, 42, 44  
 Brain atrophy, 335  
 Brain MR images  
     atrophy, 335  
     automated image registration  
         direct methods, 337–338  
         feature-based methods, 336–337  
         intensity-variation-tolerant  
             measures, 338  
         joint segmentation registration models, 339–340  
         segmentation-guided registration, 340  
     registration based on segmentation plus  
         registration component, 345–348  
     registrations based on local correlation  
         component, 344–345  
     registrations of segmentation plus  
         registration LC model, 348–350  
 segmentation guided registration model  
     design considerations, 341  
     intensity homogeneity, 341–342  
     level set formulation, LC-based model,  
         342–343  
     SSD-based segmentation plus  
         registration model, 344  
 Brain perfusion images, groupwise motion  
     correction of  
     experimental results, 287–289  
     methods  
         Gauss–Newton optimization, 286–287  
         Levenberg–Marquardt, 287  
         total quadratic variation, 284–286  
     vs. pairwise registration, 289–290  
 Breadth First Search (BFS) technique, 145–146  
 Brechbuhler, C., 22  
 Brennan, N., 330, 331  
 Briggs, R.W., 24  
 Brooks, D.H., 171  
 Brunelle, F., 164  
 Brunie, L., 336  
 Burnt nodes, 197
- C**
- Cachier, P., 342  
 Cai, Z-Q., 330, 331  
 Calderon, F., 20  
 Canny edge detector, 9–10  
 Cappelli, C., 138  
 Cardoso, J.S., 187  
 Carp, S.A., 171  
 Casanova, M.F., 161, 166  
 Castro-Pareja, C.R., 26  
 Caviness, V.S. Jr., 164  
 Cerebral white matter segmentation  
     atlas-based methods, 44  
     challenging factors, 42  
     deformable models, 44  
     ground truth graphs, 55–61  
     probabilistic graph cut algorithm  
         Box plot of performance, 61–62  
         dice coefficients, 56  
         2D segmented surfaces, 55–58  
         3D segmented surfaces, 58–60  
         parameter selection strategy, 51–52  
         procedure, 45–46  
         shape prior imposition, 52–55  
         vs. traditional graph cut, 47–50  
     statistical models, 44  
     traditional graph cut algorithm  
         advantage, 42  
         construction, 45  
         criteria, 45  
         2D segmentation, 55–58  
         3D segmentation, 59–61  
         EMS algorithm, 43  
         vs. probabilistic graph cut, 47–50  
         T1-weighted MRI input data, 42
- Chabane, N., 164  
 Chakraborty, A., 100  
 Challenges, 2–3  
 Chang, D.O., 316  
 Chan, H., 120  
 Chan, T., 342  
 Chaudhuri, B., 6  
 Chaudhuri, K., 263  
 Chef'd'hotel, C., 263

- Chen, C.Y., 316  
Cheng, H.Y., 316  
Cheng, K., 24, 74  
Chen, K-H., 316  
Chen, M., 337  
Chen, Q., 71  
Chen, T., 101  
Chen, W., 71  
Chen, Y., 8, 24, 26, 74, 100  
Chiang, H.K., 316  
Cho, P.S., 8, 21  
Chowdhury, A.S., 41  
Christensen, G.E., 337  
Cignoni, P., 173  
Cine cardiac magnetic resonance image segmentation  
active shape model (ASM), 70  
global index  
cavity area, 83  
ejection fraction, 82  
ventricular function curves, 84–85  
hybrid active appearance model (AAM), 70  
joint MGRF model, 75  
left ventricle, shape statistics  
Geodesic Active Region (GAR) model, 71  
Gradient Vector Flow (GVF) model, 71  
reliability challenges, 71–72  
local index  
Laplace’s equation, 86  
wall thickness estimation, 87–92  
MRI data, 73  
myocardial borders  
algorithm, 79  
anatomical structures, accuracy challenges, 73  
inner contour shape prior, 77  
intensity model, 77  
joint MGRF model, 75  
level set evolution, 74–75  
manual segmentation, 81  
outer contour shape prior, 75–77  
performance evaluation, 79  
segmentation errors, 81, 82  
spatial interaction model, 78  
stochastic speed function, 78–79  
test data sets, 80  
parametric shape model limitations, 70  
proposed framework, 73  
wall thickness measurements  
borders and streamlines, 87, 90  
color-coded maps, GGMRF  
smoothing, 91  
3-D GGMRF model, 89  
electric field vectors, 91  
ground truth vs. point-to-point correspondences, 88  
limitations, 72  
point-to-point correspondences, 87–88  
values, 88, 90  
ventricular function curve, 89  
Coggins, J.M., 5  
Colchester, A., 43  
Collignon, A., 238, 338  
Color Doppler ultrasound imaging, goiter nodule, 140, 149  
Computed tomography (CT), 229  
Computer aided breast detection systems  
image dataset, 129  
mass classification  
compactness, 124  
experiment results, 129–132  
Fourier descriptor, 125  
linear discriminant analysis (LDA), 128  
methods, 128  
moments, 124–125  
morphological features, 123–126  
NRL, 125–126  
receiver operating characteristic (ROC) analysis, 131  
RGO, 126  
support vector machine (SVM)  
classification method, 131–132  
texture features, 126–128  
mass segmentation  
algorithm performance, 121–123  
experiment results, 129  
Computer-aided design methods, 207–210  
Conjugate gradient optimizer, 239  
Constrained optimization, 280–281, 376–277  
Contrast-enhanced ultrasound imaging (CEUS)  
lesion vasculature strategy, 152–153  
micro-bubble detection, 142  
preprocessing, 141–142  
steps, 140  
threshold process and 3-D reconstruction, 142  
Cooper, D.H., 8, 21, 100  
Coordinate systems, 232  
Cootes, T.F., 8, 21, 100, 101  
Cortex variability, shape-based detection  
ASD  
M-CHAT, 164  
prevalence estimates, 162–163  
autism  
classifications, 162

- 3-D binary representation, brain, 169  
 early detection, 163  
 flowchart, proposed analysis system, 168  
**G**  
 Gyral window index (GWI) and volume analysis, 166  
 incidence and prevalence, 162  
**M**-CHAT, 164–165  
 mesh generation, 170–172  
 MRI image, 168  
 neuropathology, 162  
 vs. non-autistic brain, 161, 178  
 shape analysis, 166–167  
 spherical deformation, 172–175  
 spherical harmonic analysis, 175–180  
 three dimensional volume construction procedure, 170  
 VBM approach, 164
- D**  
 Davatzikos, C.A., 8, 22, 74, 100, 336  
 Davoudi, Y., 152  
 Deandrea, M., 137  
 de Aquino, L.C.M., 187  
 de Bruijne, M., 71  
 Declerck, J., 277  
 Deformable segmentation models  
   geometric level-set based models, 24–26  
   parametric deformable models, 20–24  
 Delaere, D., 338  
 Delbarba, A., 138  
 De Martino, E., 138  
 Demons algorithm, 344–347  
 Deriche, R., 101  
 Desautels, J., 120  
 Deutsch, C.K., 164  
 de Weert, T.T., 26  
 Diffeomorphisms  
   fast convolutions, 274–276  
   fixed vector map, 273–274  
 Digital prosthesis design  
   bone segmentation, 188  
   computer-aided design methods and systems, 207–210  
   continuation and marching methods, 189  
   cranioplasty, 190  
   2D deformable model balloon, 202–205  
   defective skull geometry, 191–192  
   edge relaxation, 193, 194  
   external and internal boundary links, 192  
   image processing and geometric modeling, 188  
 InVesalius software, 191
- isosurface methods and T-surfaces  
   multiscale and subdivision schemes, 213–214  
 object characteristic function, 210–211  
 piecewise linear manifold, 211  
 PyImageVis computer systems, 214–216  
 T-surfaces reparameterization, 211  
   virtual scalpel, 212  
 mass-spring system, 194–197  
 PCA, 217  
 polymethylmethacrylate (PMMA), 190  
 reflection method, 188, 219  
 segmented bone region, 219–220  
 shape model, 197  
 soft tissue differentiation, 189  
 subdivision scheme, 205–207  
 surface reconstruction, 198–199  
 triangulated hole, 193  
 triangulated surface, 219  
 T-surfaces approach, 195–197  
 wavelet theory and multiscale analysis, 199–202
- Ding, G., 71  
 Ding, Z., 121  
 Duncan, J.S., 9, 23, 26, 71, 100  
 3D vascular tortuosity. *See* Sum of angles metric parameter
- E**  
 Ecabert, O., 22  
 Eden, M., 5  
 Efron, N., 330, 331  
 Elastic motion correction algorithm, dynamic MRI, 290–291  
 El-Baz, A., 1, 25, 41, 69, 99, 161, 166, 247  
 Elnakib, A., 1, 41, 166  
 EMS algorithm. *See* Expectation maximization segmentation algorithm  
 Eubank, W., 238  
 Evans, A.C., 336  
 Expectation maximization segmentation (EMS) algorithm  
   3D segmented surfaces, 59  
   statistical method, 43  
 Extrapolation, 268–269  
 Eyelashes, removal of, 322–325
- F**  
 Fan, A., 24  
 Fang, Q., 171  
 Faugeras, O., 24, 74, 100  
 Feature-based image metamorphosis

- ocular thermographic sequences  
average horizontal temperature profile,  
327–329  
average vertical temperature profile,  
327–329  
data acquisition, 318–319  
eyelashes, removal of, 322–325  
Grave’s ophthalmopathy, 316  
horizontal and vertical temperature  
profiles, 317–318  
image warping, 319–322  
infrared thermography, 315  
myriad shapes of palpebral fissure, 317  
ocular surface temperature, 315–316  
polynomial coefficients, 330  
temperature profile, 326–327
- Feature-based image registration methods,  
336–337
- Feature-based registration techniques, 247–248
- Feldmar, J., 277
- Feng, Q., 71
- Fernandez-Bouzas, A., 20
- Ferrie, F.P., 231
- Fischer, D.R., 263
- Follicular adenoma, 138
- Follicular carcinoma, 139
- Fourier descriptor (FD), 125
- Fox, M., 124
- Frackowiak, R.S.J., 164
- Fractal models, 6
- Freedman, D., 339
- Frequency domain filtering, 5–6
- Friston, K., 164
- Frith, C., 164
- Frith, U., 164
- Fu, K.S., 22
- G**
- Gandossi, E., 138
- Garberoglio, R., 137
- Gauss-Newton optimization, 286–287
- Geiser, E.A., 8, 24, 74
- Geometric level-set based deformable models,  
24–26
- Geppert, C., 263
- Gerig, G., 22, 100
- Gerlach, J., 72
- Germano, G., 72
- Gervais, H., 164
- Giger, M., 120
- Gimel’farb, G., 1, 25, 69, 99, 166, 247
- Giraldi, G.A., 187
- Giridharan, G.A., 69
- Gobbetti, E., 173
- Gong, L., 8, 21
- Good, C.D., 164
- Gopinath, K.S., 24
- Gore, J., 121
- Gradient ascent method, 272
- Gradient descent optimizer, 239
- Graham, J., 8, 21, 100
- Grave’s ophthalmopathy, 316
- Gray Level Co-occurrence Matrix (GLCM),  
126–128
- Greig, D.M., 18
- Grimson, E., 100
- Grimson, W.E.L., 24
- Grimson, W.L., 74
- Groupwise motion correction, brain perfusion  
images  
experimental results, 287–289  
methods  
Gauss-Newton optimization, 286–287  
Levenberg–Marquardt, 287  
total quadratic variation, 284–286  
*vs.* pairwise registration, 289–290
- Grzeszczuk, R.P., 24
- Guimond, A., 277
- Gu, L., 24, 74
- Gurcan, M., 120
- Gyral window index (GWI) analysis, 166
- H**
- Hadjiiiski, L., 120
- Hameeteman, R., 26
- Happe, F., 164
- Harris, G.J., 164
- Hawkes, D.J., 228, 338
- Haynor, D.R., 8, 21, 238
- Hellier, P., 26
- Helvie, M., 120
- Heng, P.A., 71
- Herbert, M.R., 164
- Hermosillo, G., 263
- Herrmann, K.-H., 263
- High-resolution ultrasonography (HRUS),  
139, 149
- Hill, D.L., 228, 338
- Ho, C., 71
- Hodgson, J., 164
- Hoffman, E.A., 20
- Holden, M., 228
- Huang, F., 8, 24, 74
- Huang, S., 339
- Huang, X., 101
- Huo, Z., 120

**I**

Image features  
 appearance  
     intensity, 5  
     spatial interaction, 5–7  
 edges  
     detectors, 9  
     texture and image smoothing, 10  
     watershed transformation, 10–11  
 scale-space, 11  
 shape, 8–9  
 Image visualization, 240  
 Image warping, 319–322  
 INCT-MACC, 214–215  
 Infrared (IR) thermography, 315  
 Intensity-variation-tolerant measures, 338  
 Interpolation, 235–236  
 Isgum, I., 19  
 Isosurface methods and T-surfaces  
     multiscale and subdivision schemes,  
         213–214  
     object characteristic function, 210–211  
     piecewise linear manifold, 211  
 PyImageVis computer systems, 214–216  
 T-surfaces reparameterization, 211  
 virtual scalpel, 212  
 ITK toolkit, 241–242

**J**

Jain, A.K., 5  
 Jain, R.C., 23  
 Janka, R., 263  
 Jia, X., 71  
 Joint optimization, 279–280  
     vs. constrained optimization, 281  
 Joint segmentation registration models,  
     339–340  
 Jolly, M.P., 42, 44  
 Jordan, D., 166  
 Joshi, S., 100, 101  
 Junior, A.L.A., 187

**K**

Kaiser, W., 263  
 Kanade, T., 337  
 Kao, C., 121  
 Kapur, T., 26  
 Karl, W.C., 8, 74  
 Kass, M., 20, 100  
 Kavanagh, P., 72  
 Kelemen, A., 22  
 Kennedy, D., 164  
 Khalifa, F., 25, 69, 161

Kiefer, B., 263

Kilday, J., 124  
 Kim, Y., 8, 21  
 Kolmogorov, V., 24  
 Kopans, D.B., 171  
 Kovacic, S., 337  
 Krams, M., 164  
 Krishnan, A., 263  
 Kuba, A., 143  
 Kundu, P., 6

**L**

Lai, N-S., 330, 331  
 Lange, N., 164  
 Lavallee, S., 336  
 Lavi, G., 22  
 Layegh, P., 152  
 LCDG model. *See* Linear combination of  
     discrete Gaussians probability  
     model  
 Learning prior appearance model  
     experimental results, 253–255  
     MGRF-based image registration  
         appearance-based registration, 252–253  
         appearance model, 249–250  
         basic notation, 249  
         characteristic neighborhood set,  
             250–251  
         image normalization, 249  
         learning the potentials, 250  
         mutual information, 248  
         registration methods, 247–248  
 Leclerc, X., 71  
 Leemput, K.V., 19  
 Lelieveldt, B.F., 70  
 Leo, D.J.E., 125  
 Leonard, C.M., 26  
 Level set segmentation. *See* Computer aided  
     breast detection systems  
 Levenberg–Marquardt methods, 287  
 Leventon, M.E., 24, 74, 100  
 Levin, D.N., 24  
 Lewellen, T.K., 238  
 Leyegh, P., 152  
 Liang, B., 71  
 Liang, J., 71  
 Liang, S., 125  
 Li, C., 71, 121  
 Li, H., 120  
 Limone, P., 137  
 Linear combination of discrete Gaussians  
     (LCDG) probability model  
     equation and restriction, 109

- evolution, 109–110  
segmentation algorithm, 110
- Linear discriminant analysis (LDA), 128
- Lin, X., 70
- Litvin, A., 8, 74
- Liu, J., 335
- Liu, X., 119
- Li, W., 24, 74
- LOG edge detector, 9–10
- Lorenz, C., 22
- Lowe, D.G., 101, 104
- Lu, P., 120
- M**
- Mackiewich, B.T., 22
- Madabhushi, A., 22
- Maes, F., 43, 238, 338
- Magnetic resonance imaging (MRI), 229–230
- brain atrophy, 335
- brain, automated image registration
- direct methods, 337–338
- feature-based methods, 336–337
- intensity-variation-tolerant measures, 338
- joint segmentation registration models, 339–340
- segmentation-guided registration, 340
- breast imaging
- effect of motion artifacts on, 298–299
- quantitative analysis of, 297–298
- dynamic MR angiography, 294–296
- dynamic MR mammography, 291–294
- elastic motion correction algorithm of dynamic, 290–291
- registration based on segmentation plus registration component, 345–348
- registrations based on local correlation component, 344–345
- registrations of segmentation plus registration LC model, 348–350
- segmentation guided registration model
- design considerations, 341
- intensity homogeneity, 341–342
- level set formulation, LC-based model, 342–343
- SSD-based segmentation plus registration model, 344
- T1 and T2 registration, 282
- Maintz, J.B.A., 228, 337, 338
- Mäkelä, T., 228
- Makris, N., 164
- Malandain, G., 277
- Malik, J., 5
- Ma, M.C., 143, 152
- Manniesing, R., 26
- Mantovani, A., 137
- Mapstone, R., 330
- Marchal, G., 238, 338
- Marching cubes method, 188, 189, 191, 198, 207, 208, 216
- Markov–Gibbs random field (MGRF) image model. *See also* Appearance prior model
- appearance-based registration, 252–253
- appearance model, 249–250
- basic notation, 249
- Bayesian maximum a posteriori decision rule, 16
- characteristic neighborhood set, 250–251
- image normalization, 249
- iterative graph-cut optimization, 18
- learning the potentials, 250
- local energy minimum, 17–18
- spatial 2-D and 3-D interaction, 7
- two-level segmentation, 15–17
- Markov random field (MRF) model.
- See* Markov–Gibbs random field image model
- Marroquin, J.L., 20
- Mattes, D., 238
- Maurer, C.R. Jr., 19
- Mead, R., 239
- Mean squares metric, 237
- Medical image registration
- applications and methods
- constrained optimization, 280–281
- diffeomorphisms, 273–276
- groupwise motion correction, brain perfusion images, 283–289
- joint intensity histogram, MR images, 278–279
- joint optimization, 279–280
- joint optimization *vs.* constrained optimization, 281
- motion compensation, temporal image sequence, 265–272
- MR T1 and T2 registration, 282
- PET-MR registration, 282–283
- polynomial intensity correction, 276–277
- SSD criterion, 277–278
- applications of, 227–228
- clinical evaluation, 290–291
- contrast-enhanced peripheral MR angiography, 296

- Medical image registration (*cont.*)  
 coordinate systems, 232  
 CT scan, 229  
 3D volume visualization, 230  
 dynamic MR mammography, 296  
 framework, components of, 231–232  
 image visualization, 240  
 interpolation, 235–236  
 ITK implementation of, 241–242  
 learning prior appearance model  
   experimental results, 253–255  
   MGRF-based image registration, 249–253  
   mutual information, 248  
   registration methods, 247–248  
 materials and methods, 291–292  
 mathematical inverse problems, 228–229  
 mean squares metric, 237  
 motion artifacts, breast MR lesions  
   kinetic analysis of, 298–299  
   quantitative analysis of, 297–298  
 MR angiography, 294–295  
 MRI, 229–230  
 MR mammography, 292–294  
 mutual information metric, 237–238  
 normalized correlation metric, 237  
 open source software, 240–242  
 optimization, 238–239  
 PET, 229  
 in physical space, 232–233  
 similarity measure, 236–238  
 SPECT, 229  
 surface guided registration, 230–231  
 transformation, 234–235  
 validation of, 242–243
- Medical image transformation, 234–235
- Mehrab, M., 152
- Metaxas, D.N., 22, 70, 101
- Meunier, J., 277
- Meyer, C.R., 20
- Michel, T.S., 71
- Middleton, M., 263
- Miller, E.L., 171
- Miller, M.I., 337
- Mirsadaee, S., 152
- Mitchell, S., 70
- Modified Checklist for Autism in Toddlers (M-CHAT), 164–165
- Mohanna, F., 22
- Mokhtarian, F., 22
- Molinari, F., 137
- Moore, R., 171
- Morgan, P.B., 331
- Motion compensation, temporal image sequence  
 conjugate gradient optimization, 271  
 convolutions, 272  
 design constraints, 266–267  
 extrapolation, 268–269  
 global handling and multi-resolution approach, 267–268  
 gradient ascent method, 272  
 local cross-correlation, 270–271  
 methods, 265–266  
 multi-resolution pyramid, 268  
 single resolution motion estimation, 269  
 warping, 269–270
- Moura, J.M.F., 71
- Mouren, M.C., 164
- MR angiography, 294–296
- MR mammography, 291–294
- Mudigonda, N., 120
- Multi atlas-based segmentation, 19–20
- Multiscale analysis (MA)  
 definition, 199–200  
 digital prosthesis design, 213–214  
 scaling equation, 201–202  
 theorem, 202
- Mutual information (MI) based image registration, 247–248
- Mutual information metric, 237–238
- Myocardial borders segmentation  
 algorithm, 79  
 anatomical structures, accuracy challenges, 73  
 inner contour shape prior, 77  
 intensity model, 77  
 joint MGRF model, 75  
 level set evolution, 74–75  
 manual segmentation, 81  
 outer contour shape prior, 75–77  
 performance evaluation, 79  
 segmentation errors, 81, 82  
 spatial interaction model, 78  
 stochastic speed function, 78–79  
 test data sets, 80
- N**
- Nandi, A., 120
- Nearest neighbor interpolation, 236
- Neely, S., 319
- Nelder, J.A., 239
- Netsch, T., 342
- Ng, E.Y.K., 315, 319
- Nielsen, M., 71
- Niessen, W.J., 26

- Nitzken, M., 161  
Noble, J.A., 339  
Non toxic goiter nodule, 138  
Normalized correlation metric, 237  
Normalized radial length (NRL), 125–126  
NRL. *See* Normalized radial length
- O**  
Object characteristic function, 210–211  
O’Brien, L.M., 164  
Ocular surface temperature, 315–316  
Ocular thermographic sequences  
    average horizontal temperature profile, 327–329  
    average vertical temperature profile, 327–329  
    data acquisition, 318–319  
    eyelashes, removal of, 322–325  
    Grave’s ophthalmopathy, 316  
    horizontal and vertical temperature profiles, 317–318  
    image warping, 319–322  
    infrared (IR) thermography, 315  
    myriad shapes of palpebral fissure, 317  
    ocular surface temperature, 315–316  
    polynomial coefficients, 330  
    temperature profile, 326–327
- Olszewski, M.E., 22  
Optimal statistical inference segmentation  
    MGRF model, 15–18  
    probability density models, 14
- Osher, S., 24  
Ousephand, R., 25  
Ou, U., 263
- P**  
Palágyi, K., 143  
Palmieri, F., 124  
Pang, B., 25  
Papademetris, X., 71  
Papillary carcinoma, 138  
Paragios, N., 71, 101  
Paramagul, C., 120  
Parametric deformable model  
    appearance prior model  
        central-symmetric neighborhood sets, 106  
        joint probability distributions, 107  
        relative energies, 108  
    DCE-MRI kidney image segmentation  
        LCDG estimates, 113  
        mutually aligned training images, 113  
        segmentation error, 113–114
- LCDG model  
    equation and restriction, 109  
    evolution, 109–110  
    segmentation algorithm, 110  
segmentation techniques, 20–24  
shape prior model  
    mutually aligned training  
        boundaries, 103  
        SIFT descriptors, 104–106  
starfish image segmentation  
    aligned images, 111, 112  
    LCDG estimates, 112  
    segmentation errors *vs.* signal-to-noise ratio, 112
- Park, H., 20  
Parzen window lung/chest models, 14  
Passingham, R., 164  
Pathak, S.D., 8, 21  
Pennec, X., 342  
Pentland, A., 100  
Perona, P., 5  
Persoon, E., 22  
Peters, J., 22  
Peters, T.M., 227, 231  
PET-MR registration, 282–283  
Petrick, N., 120  
Piecewise linear (PL) generation algorithm, 198–199  
Pintus, R., 173  
Pirola, I., 138  
Pizer, S., 100  
Pluempiwiriyawej, C., 71  
Plumet, M.H., 164  
Polymethylmethacrylate (PMMA), 190  
Polynomial intensity correction, 276–277  
Pomerleau, D., 337  
Porteous, B.T., 18  
Positron emission tomography (PET), 229  
Prasad, M., 72  
Prewitt edge detector, 9  
Prince, J.L., 22, 71, 336  
Probabilistic Gaussian models, 6  
Probabilistic graph cut algorithm  
    Box plot of performance, 61–62  
    dice coefficients, 56  
    3D segmentation  
        axial view, 60  
        coronal and sagittal view, 60  
        surfaces, 59  
    2D segmented surfaces, 55–58  
        axial view, 56  
        coronal view, 57  
        sagittal view, 58

- Probabilistic graph cut algorithm (*cont.*)  
 parameter selection strategy, 51–52  
 procedure, 45–46  
 shape prior imposition  
   DC variation, 52  
   2D segmentation, 53, 54  
   ellipse, 53  
   subgraph, advantages, 54  
 vs. traditional graph cut  
   edge capacities and scaling functions, 49–50  
   maxflow-mincut algorithm, 47–49
- Prokop, M., 19, 20
- PyImageVis computer systems, 214–216
- Q**
- Qian, Z., 70
- Qin, A., 71
- Qin, J.X., 26
- Quist, M., 342
- R**
- Rad, M.P., 152
- Rainey, S., 166
- Ramesh, A., 72
- Rangayyan, R.M., 120, 125
- Rapin, I., 162
- Receiver operating characteristic (ROC)  
 analysis, 131
- Reiber, J.C., 70
- Reichenbach, J.R., 263
- Reinhardt, J.M., 20
- Reiver, J., 70
- Relative gradient orientation (RGO), 126
- Roberts edge detector, 9
- Roche, A., 277
- Rodrigues, P.S.S., 187
- Rohlfing, T., 19
- Rojas Dominguez, A., 120
- Rosch, P., 342
- Rowley, H., 337
- Rozie, S., 26
- Rudra, A.K., 41
- Rule-based segmentation method  
 region growing method, 12–13  
 region split-and-merge techniques, 13
- Rutten, A., 19
- S**
- Sahiner, B., 120
- Salganicoff, M., 263
- Samson, Y., 164
- Sarker, N., 6
- Scale invariant feature transform (SIFT)  
 technique  
   image alignment, 105  
   steps, 104
- Scale spaces, 199
- Schmidt, R., 120
- Schmitt, M., 263
- Schramm, H., 22
- Sclaroff, S., 100
- Scopigno, R., 173
- Segmentation guided registration model  
 design considerations, 341  
 intensity homeogeneity, 341–342  
 level set formulation, LC-based model, 342–343  
 SSD-based segmentation plus registration model, 344
- Seheult, A.H., 18
- Selb, J., 171
- Sen, M., 41
- Sennett, C., 120
- Sethian, J.A., 24
- Shah, M., 23
- Shape prior  
 cine cardiac magnetic resonance image  
 segmentation  
   inner contour, 77  
   outer contour, 75–77
- estimating steps, 8
- parametric deformable model  
 mutually aligned training boundaries, 103
- SIFT descriptors, 104–106
- probabilistic graph cut algorithm  
 DC variation, 52  
 2D segmentation, 53, 54  
 ellipse, 53  
 subgraph, advantages, 54
- Shekhar, R., 26
- Shen and Davatzikos, 24
- Shen, D., 8, 74, 100
- SIFT. *See* Scale invariant feature transform  
 technique
- Simultaneous autoregressive models, 6–7
- Single atlas-based segmentation, 18–19
- Single photon emission computed tomography  
 (SPECT), 229
- Sinusas, A.J., 71
- Slomka, P.J., 72
- Sluimer, I., 20
- Smith, G.M., 6
- Sobel edge detector, 9
- Soh, M.P., 331

- Sokhadze, G., 161  
Song, J.-M., 26  
Sonka, M., 70, 143, 152  
Spatial domain filtering, 5  
Staib, L.H., 23, 26, 100, 339  
Staring, M., 19  
Steele, S., 164  
Studholme, C., 338  
Styner, M., 100  
Subdivision scheme, digital prosthesis design, 205–207  
Subramanyan, K., 22  
Suetens, P., 43, 238, 338  
Sum of angles metric (SOAM) parameter, 146–147  
Sum of local correlation, 338, 344  
Sum of squared difference (SSD), 338, 341, 344  
Sun, Y., 71  
Support vector machine (SVM) classification method, 131–132  
Surface guided registration, 230–231  
Suri, J.S., 1, 137, 187, 315  
Suzuki, K., 120  
Switala, A., 166  
Szekely, G., 22  
Szeliski, R., 336
- T**  
Tagare, H.D., 8, 24, 74, 100  
Taleb-Ahmed, A., 71  
Tang, J., 119  
Tang, M., 71  
Tan, J.H., 315, 319  
Tan, L., 330, 331  
Tao, X., 8  
Taylor, C.J., 8, 21, 100, 101  
Techniques  
atlas-based segmentation  
  multi models, 19–20  
  single models, 18–19  
deformable models  
  geometric level-set based models, 24–26  
  parametric deformable models, 20–24  
optimal statistical inference  
  MGRF model, 15–18  
  probability density models, 14  
rule-based method  
  region growing method, 12–13  
  region split-and-merge techniques, 13
- Tempany, C., 24  
Terzopoulos, D., 20, 100
- Tetra-Cubes, 198, 199, 211  
Thirion, J.P., 337  
Thiruvenkadam, S., 8, 24, 74, 100  
3D volume visualization, 230  
Thyroid nodule  
  CEUS image processing strategy  
    lesion vasculature strategy, 152–153  
  micro-bubble detection, 142  
  preprocessing, 141–142  
  steps, 140  
  threshold process and 3-D reconstruction, 142  
Color Doppler image, goiter nodule, 149  
computer-aided diagnosis, 150–153  
diagnosis, 139–140  
HRUS and CEUS images, 149  
pathologies, 138–139  
selection, 147–150  
skeletonization techniques  
  3D skeleton, 150  
  parameters, malignant nodules, 151  
  procedure, 143–145  
  sum of angles metric (SOAM), 146–147  
Total quadratic variation, 284–286  
Trilinear interpolation, 236  
Tsai, A., 8, 24, 74, 101  
T-surfaces approach, 195–197  
Tuchman, R.F., 162  
Tucker, D., 24  
Tullo, A.B., 331  
Tutar, I.B., 8
- U**  
Unser, M., 5  
Uzümçü, M., 70
- V**  
van den Elsen, P.A., 34, 337, 338  
van der Geest, R.J., 70  
van der Lugt, A., 26  
Vandermeulen, D., 43, 238, 338  
van Ginneken, B., 19, 20  
Van Leemput, K., 43  
Vannier, M., 337  
van Walsum, T., 26  
VBM. *See* Voxel-based-morphometry  
Vembar, M., 22  
Vemuri, B.C., 20, 26, 339  
Ventriculometrics segmentation. *See* Cine cardiac magnetic resonance image segmentation  
Vese, L., 342  
Vesselle, H., 238

- Viergever, M.A., 19, 228, 337, 338  
Viola, P.A., 338  
von Berg, J., 22  
Voxel-based-morphometry (VBM), 164  
Vukadinovic, D., 26  
Vyborny, C., 120
- W**  
Walimbe, V., 26  
Walker, M.J., 22  
Wallner, K., 8  
Wang, K., 25  
Wang, Y., 339  
Warfield, S.K., 8  
Warping, 269–270  
Wavelet spaces, 201  
Wavelet theory, 199–202  
Weese, J., 22, 342  
Wells, W.M., 8, 24, 338  
Wenkel, E., 263  
Weymouth, T.E., 23  
Williams, D.J., 23  
Willsky, A., 8, 24  
Wilson, D., 8, 24, 74  
Witkin, A., 20, 100  
Wolverton, D., 120  
Wu, J., 24, 74  
Wurdinger, S., 263
- Wu, Y., 71  
Wu, Y.-J.L., 71  
Wyatt, P.P., 339
- X**  
Xia, D.-S., 71  
Xu, C., 22, 71  
Xu, J., 24, 74
- Y**  
Yang, J., 9, 26  
Ye, J., 26  
Yezzi, A., 26  
Yezzi, A. Jr., 24  
Young, G., 330, 331  
Yuan, Y., 120
- Z**  
Zagrodsky, V., 26  
Zhang, D., 25  
Zhang, L., 20, 263  
Zhang, T., 339  
Zhou, S., 71  
Zhou, Z.M., 71  
Zhu, Y., 71  
Ziegler, D.A., 164  
Zilbovicius, M., 164  
Zollei, L., 26