PRIOR-BASED SEGMENTATION OF MR IMAGES USING GRAPH CUTS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jie Zhu

May 2008

PRIOR-BASED SEGMENTATION OF MR IMAGES USING GRAPH CUTS

Jie Zhu, Ph.D.

Cornell University 2008

The development of faster and higher resolution MR imaging devices has made accessible mass quantities of image data. Much information can be extracted by analyzing these high spatial resolution images. For instance, tissue volumes, which can be measured through MR images, are used as an indicator in many clinical applications and research studies. Research studies involving tissue volume analysis often require the processing of a vast amount of data hence manual segmentation of the images by experts is very time-consuming. Data segmented by human experts are also likely to show inter- and intra-observer inconsistency. For these reasons, automated segmentation of MR images is of great importance and interest.

MR images present many challenges for automated segmentation. In addition to noise, MR images are also corrupted by problems specific to MR imaging such as intensity inhomogeneity. Furthermore, poor contrast at tissue boundaries due to multiple tissues having similar MR intensities also present problems. On the other hand, strong contours may exist where boundaries are not desired, because multiple tissues with very different MR intensities may be present within the same anatomical structure. Prior information, such as spatial atlases and shape priors can be very powerful in these cases.

In this thesis, we developed highly accurate and robust graph cuts-based method that automatically segments MR images. The images we are most interested in are those that cannot be correctly segmented using intensity information alone. We developed models for robustly incorporating prior information such as

spatial atlas and geometric or statistical shape priors into the efficient graph cuts segmentation framework. Specifically, we developed methods to incorporate spatial atlas, statistical and geometric shape priors with graph cuts for MR image segmentation. We tested our methods on MR brain, abdomen and cardiac images with intensity inhomogeneity, poor contrast at desired boundaries and/or strong contrast at undesired boundaries and obtained encouraging results. Finally, we proposed a way of dealing with objects with curvy boundaries.

BIOGRAPHICAL SKETCH

Jie Zhu was born the daughter of a political science professor and a university administrator. She received primary and secondary education from six schools in three different countries spreading over two continents. Some time during this period, she discovered that anything is possible with a little bit of luck, a lot of determination and a moderate amount of nagging from her mother, when, after learning Japanese for three months, she scored the highest on several classical Japanese exams amongst students who were native Japanese speakers.

Jie entered the University of Tokyo in 1998, where she earned the B.Eng. degree and was proudly one of the four female students in the department of electrical engineering studying alongside 115 male students. She did her undergraduate research project in the Harashima/Naemura Lab, where she was introduced to the infinite possibilities of engineering. Her project was on developing a virtual mirror based on computer vision and image-based rendering techniques using a camera array and a display.

She entered the Ph.D. program at Cornell University in August 2002 and defended her thesis proposal in January 2007. During her years at Cornell, she occupied her time with medical image analysis, computer vision, machine learning and cross-country skiing. With this thesis, she will complete her Ph.D. requirements and is very excited about what the future holds in store for her.

ACKNOWLEDGEMENTS

I would like to, first and foremost, express my gratitude to my advisor Prof. Ramin Zabih. I feel very fortunate for this opportunity to study under a person who is extremely knowledgable and insightful as a researcher, supportive and encouraging as a mentor, and above all fair and kind as a human being.

To Prof. Ashish Raj and Dr. Susanne Muller at the University of California, San Francisco for their collaboration and helpful discussions. I am also in debt to Susanne for providing the medical images, verifying the segmentations and lending me her medical expertise on many occasions.

To Dr. Mike Weiner, Prof. Norbert Schuff, Prof. Colin Studholme, Prof. Lara Stables, Peter J. Lorenzen, Ph.D., Stathis Hadjidemetriou, Ph.D., and others at the Center for Imaging of Neurodegenerative Diseases at UCSF for their patience, advice and discussions. To Prof. Yi Wang and Dr. Martin Prince of the Weill Medical College of Cornell University for constructing a wonder medical imaging research environment.

I thank my committee members Prof. Dan Huttenlocher and Prof. Kavita Bala for their insightful questions during the examinations and their patience for reading this thesis. I am also thankful to Prof. Huttenlocher for introducing me to the fascinating field of computer vision. I am grateful to Ms. Cindy Robinson for making life easier for me on many occasions.

To my fellow students and friends Gurmeet Singh and Bryan Kressler for their companionship. To all my friends at Cornell whom I met through the ECE and CS departments, at Schuyler House, Hasbrouck Apartments and the Gamma Alpha Science Co-op, and through Chinese and Japanese connections for reminding me how exciting graduate school is, thank you!

To my mother for showing me by example that the Chinese saying of, there are

three sexes: males, females and female Ph.D.s, is utterly untrue and for sparking my interest in knowledge acquisition with her constant academic discussions in our living room throughout my childhood. To her students for caring for my parents for the past 10 years as if they were their own. To my father for showing me that being kind and content is the key to happiness. To my late grandfather, who remained honorable and respectable after losing everything (including his mother) to the revolution, for teaching me that knowledge is the only asset you can carry with you.

Last but far from the least, to my husband, Blake Charles Jacquot, the kindest, warmest, most wonderful human being I've ever known in my life, for going through the grad school adventure together with me and for being caring and understanding even in very difficult circumstance.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

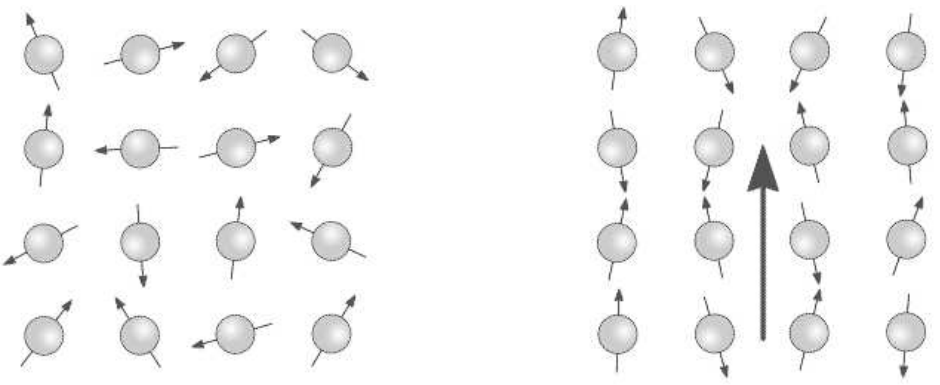
Chapter 1
Introduction

1.1 Background and Motivation

The first MR image, produced over 30 years ago, brought a new era of modern medicine [25]. Today, there are thousands of MR scanners in use across the United States, generating high quality internal images of human bodies for diagnostic and research purposes. Unlike its earlier counterparts such as X-ray and computed tomography (CT) scans, MR imaging does not involve exposure to imaging radiation which greatly limits its side effects [52, 49]. Imaging in 3D is also easy with MR scanners. The goal of this thesis is to develop robust method with high accuracy that automatically divides MR images into anatomical structures. In this section, we introduce the basics of MR imaging technologies and discuss the importance and challenges of MR image segmentation.

1.1.1 Magnetic Resonance (MR) Imaging Technologies

Discussing the full scale of quantum physics behind MR imaging is beyond the scope of this thesis and we will give only a brief summary. Hydrogen nuclei in water and lipids play an integral part in medical MR imaging. Atomic nuclei are made of protons and neutrons. Protons and neutrons both have spins (or angular momentum intrinsic to the body) of 1/2, and therefore possess magnetic moments. Most elements have nuclei made up of an even number of protons and neutrons and the magnetic moments are canceled out. However, the most common isotope of hydrogen nuclei (with an abundance of more than 99.98%) is made up of one proton and no neutrons possessing a magnetic moment and tends to align with a strong magnetic field (Fig. 1.1). The hydrogen protons in a magnetic field have



(a) In their normal states, the magnetic moments of hydrogen nuclei point in all directions.

(b) In a strong magnetic field the nuclei tend to align with the field in one direction or the other. (The arrow in the middle shows the direction of the magnetic field.)

Figure 1.1: Magnetic moments of hydrogen nuclei.

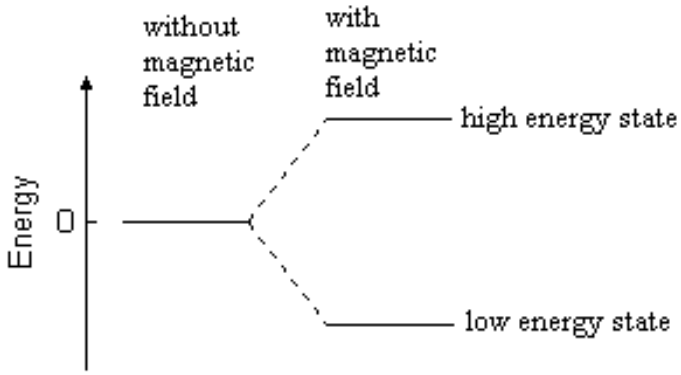


Figure 1.2: Protons in a magnetic field have two potential energy levels.

two potential energy levels (Fig. 1.2) [55]. It is possible to excite a proton in the lower energy state to the higher energy state with an electromagnetic wave possessing exactly the transition energy (energy difference between the two states) of the proton. We call the frequency of this particular electromagnetic wave the Larmor frequency. The transition energy (ΔE) and Larmor frequency (f_{Larmor}) are determined by the strength of the magnetic field:

$$\Delta E = \frac{\gamma h B}{2\pi}, \qquad f_{Larmor} = \frac{\gamma B}{2\pi}, \tag{1.1}$$

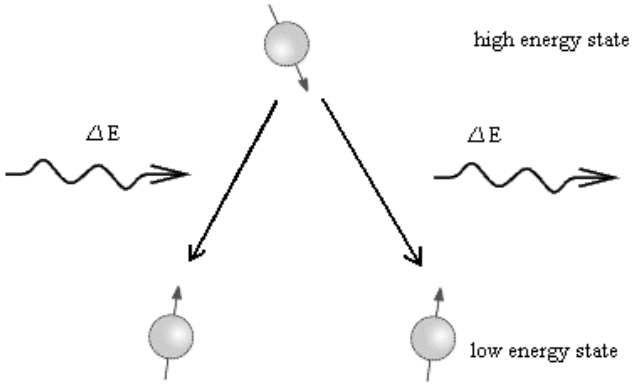


Figure 1.3: In a magnetic field, a proton in the lower energy state can be excited to the higher energy state by an electromagnetic wave with the proton's transition energy. The proton eventually comes back to the lower energy state releasing the transition energy.

where γ is the gyromagnetic ratio of protons, h is the Planck's constant and B is the strength of the magnetic field. Excited protons stay in the higher energy state for a while, but eventually come back to the lower energy state, releasing the transition energy (Fig. 1.3). The time lapse between the excitement of the proton and the energy release is called relaxation time [55].

Medical MR machines image the relaxation times of excited hydrogen protons since the relaxation times differ from tissue to tissue. The scanner first applies a gradient magnetic field, which varies in three orthogonal directions. This gives each location in space a magnetic field with unique strength, hence a unique Larmor frequency. The relaxation time at each location is then measured by exciting the protons through electromagnetic energy (RF pulses) with their Larmor frequency and receiving the released energy through a receiving coil [57].

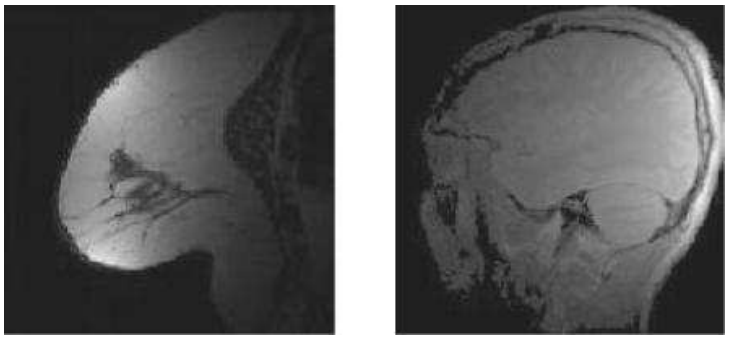
1.1.2 Why Magnetic Resonance (MR) Image Segmentation?

The development of faster and higher resolution MR imaging devices have made accessible mass quantity of image data. Much information can be extracted by analyzing these high spatial resolution images. For instance, tissue volumes, which can be measured through MR images, are used as an indicator in many clinical applications and research studies. For example, medical studies are underway to investigate the relationship between brain volume loss and Alzheimer's disease [56, 71]. In autosomal dominant polycystic kidney disease (ADPKD) patients, the rate of kidney enlargement have been shown to characterize the rapidness of renal function decline [32]. Thickening of the myocardium of the left ventricle is often used as a marker in clinical diagnosis for cardiovascular diseases [14]. Dividing medical images into different anatomical structures, i.e. medical image segmentation, is the key to tissue volume analysis. Medical image segmentation also leads to many other useful applications such as visualization of 3D anatomical structures for the purpose of surgical planning and simulation [81].

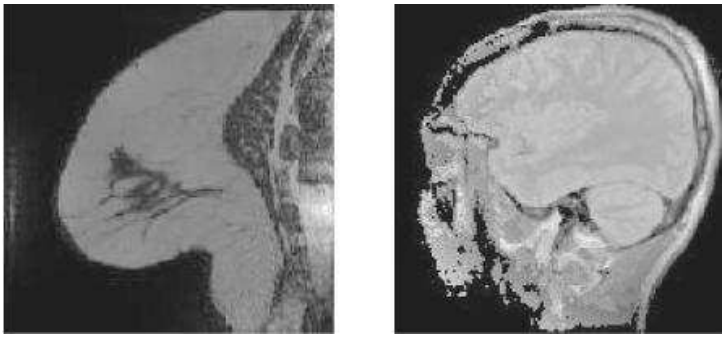
Research studies involving tissue volume analysis often require the processing of a vast amount of data and hence manual segmentation of the images by experts is very time-consuming. Data segmented by human experts are also likely to show inter- and intra-observer inconsistency [43, 40]. For these reasons, automated segmentation of MR images is of great importance and interest.

1.1.3 Challenges of MR Image Segmentation

MR images present many challenges for automated segmentation. In addition to noise, MR images are also corrupted by problems specific to MR imaging such as



(a) MR images with bias field.

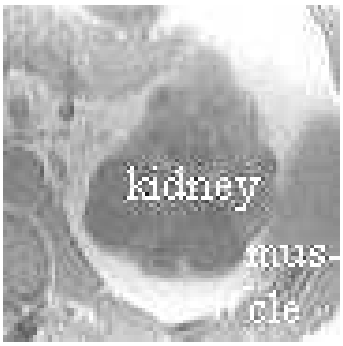


(b) MR images with bias field correction.

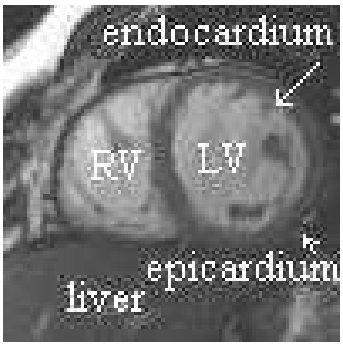
Figure 1.4: Effects of intensity inhomogeneity in MR images. Images based on [36].

intensity inhomogeneity. Intensity inhomogeneity (also called bias field or intensity non-uniformity) refers to the non-anatomic spatial variation in intensities within an MR image caused in part by the difference in distance between image voxels (the smallest unit in MR images) and the receiving coil [36]. Examples of images corrupted by intensity inhomogeneity are shown in Fig.1.4. For comparison, the same images without intensity inhomogeneity are also shown.

Furthermore, poor contrast at tissue boundaries due to multiple tissues having similar MR intensities also present problems. For example, the kidney tissues and neighboring psoas muscle tissues have similar intensities in Fig. 1.5(a). Fig. 1.5(b) shows a cardiac image with weak contrast at boundaries between the epicardium and the liver and those between the endocardium and the blood pool. On the other hand, strong contours may exist where boundaries are not desired, because



(a) Poor contrast exists between kidney tissues and muscle tissues.



(b) Poor contrast exists between the epicardium and liver tissues and between the endocardium and the blood pool inside the left ventricle (LV).

Figure 1.5: MR images that are hard to segment because of poor contrast at desired boundaries.

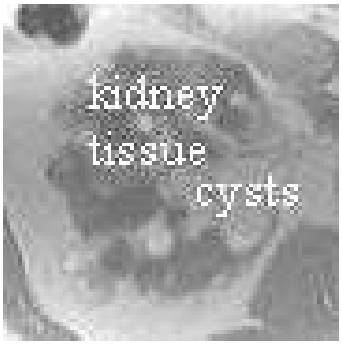


Figure 1.6: An MR image that is hard to segment because of strong contrast at undesired boundaries. Here, cysts with very different intensities than normal kidney tissues exist within the kidney.

multiple tissues with very different MR intensities may be present within the same anatomical structure. For example, Kidneys of ADPKD patients also include a large amount of liquid, which have very different intensity levels than normal kidney tissues, creating strong contours within the kidney as shown in Fig. 1.6.

Traditional image segmentation methods make general assumptions such as objects generally have smooth boundaries [70]. Some anatomical structures stray from such assumptions, which make their segmentations using conventional image segmentation method difficult. Fig. 1.7 shows a brain MR image with many curvy

(a) Brain image

(b) White matter tissue in the brain with curvy boundaries

Figure 1.7: MR brain images have many curvy boundaries.

boundaries which cannot be correctly segmented using traditional segmentation methods assuming smooth boundaries.

1.2 Medical Image Segmentation using Prior Information

As mentioned above, automated MR image segmentation is a challenging problem. Intensity inhomogeneity, poor contrast at tissue boundaries and strong contours within anatomical structures make segmentation using intensity information alone very difficult. Prior information, such as spatial atlases and shape priors can be very powerful in these cases. Atlases can help segment images using spatial information where intensity information alone cannot classify voxels accurately. Shape priors allow global information to be incorporated into segmentations. In this section, we first briefly summarize recent image segmentation methods. We then discuss representations and generations of spatial atlases and shape priors and their incorporation with recent segmentation techniques.

1.2.1 Image Segmentation Methods

Computer vision is a decades old field of study and a large number of image segmentation methods have been proposed in the literature. In image segmentation, one wishes to divide an image into different regions. This task can be accomplished directly by defining contours representing object boundaries or indirectly by assigning each image pixel a label representing the region to which it belongs. Early segmentation techniques depended on basic image features including intensity and edge information. Such techniques include: region growing, clustering, thresholding and edge detection-based approaches [70].

In recent years, elegant approaches involving energy minimization have gained the most interest. These segmentation methods can be roughly divided into two categories. One involves continuous boundary representations and the other discrete boundary representations. The first category includes active contour methods such as "snake" [41] and level set-based approaches [15, 17]. The level set-based methods embed object boundaries as the zero level set in a level set function and use curve evolution to find the boundaries with the lowest energy for the cost function. Within the second category of energy-based segmentation methods, graph cuts-based approaches have become popular because they allow for globally optimal efficient solutions in an N-dimensional setting [9, 33, 91]. These methods represent images using weighted undirected graphs. Image pixels are nodes on the graph and each pair of pixel neighbors are connected through an edge. The weights of edges depend on the similarity of the pixels. Image segmentation is performed by separating the nodes into disjoint sets through the removal of some edges in the graph. Developments in this thesis on based on the graph cuts image segmentation method and we will give a detailed review in the next chapter.

1.2.2 Spatial Atlases-Guided Brain MR Image Segmentation

With medical images, it is often possible to produce anatomical probabilistic maps based on manual labeling. These atlases can help when voxels cannot be easily classified using intensities by classifying them using spatial information. In this thesis, we focus on atlas-guided brain image segmentation since brain atlas is the most prevalent type of spatial atlas in the literature. Spatial atlases are generated using manually segmented data [27, 62]. The first atlas-based brain segmentation approaches registered the atlas with the image and segmented the image based on the atlas alone [67], [38]. Although these methods are very fast, their accuracy is highly dependent on the quality of the atlas.

Later techniques incorporate a brain atlas into a segmentation framework. Such methods generally include two steps: registering the atlas to the MR brain image and segmenting the brain image using the registered atlas. Sequential approaches were first proposed [21, 51, 64, 19, 47]. In these methods, the segmentation is susceptible to inaccuracies in the atlas registration. There have been recent attempts to unify the medical image segmentation and atlas registration. Joint segmentation and rigid registration has been studied in [5], [87], [84], [61] and [29]. In [5], an entropy-based framework is developed for CT images. [87] and [61] use variational approaches while [84] and [29] adopt Bayesian methods for segmentation. Deformation of brain images cannot be adequately represented using rigid motion and so non-rigid transformation is necessary.

In [3] and [65], methods integrating non-rigid registration with mixture model-based segmentation are proposed. Here, intensity probability distributions of different tissues in the brain are modeled using Gaussian Mixture Models (GMMs) [7] and the GMM parameters and segmentations are calculated through expectation

maximization (EM)-based approaches [26, 53]. [3] is a voxel-based method where voxels are classified independently and labels of neighboring voxels are not taken into consideration whereas [65] uses spatial neighboring constraints although not full scale Markov random field model. Classifications of image voxels generally exhibit spatial dependency and segmentation methods that take this into consideration work better. Coupled segmentation and registration methods of misaligned medical images using Markov Random Field models has been studied in [85]. [80] presents a level set-based approach, which unifies the image segmentation and atlas registration problems.

1.2.3 Segmentation of Anatomical Structures using Shape Priors

Many shape models have been explored for segmentation purposes. [23] used a point-based shape model and [42] represented object surfaces using spherical harmonics. Both of these methods used principle component analysis (PCA) to reduce the dimensionality of the shape statistics. Medial shape representation and non-parametric shape models were studied in [63] and [24] respectively.

Algorithms for incorporating statistical shape information with level set-based segmentation methods have been studied extensively. In [68, 18], variational level set-based approaches were proposed. In [76, 48, 66], signed distance maps were used to represent the statistical shape information. During each step, [48] estimated the shape based on shape information from the previous step and intensity information of the image. The segmentation was then found using the shape at the current step and the intensity information. [76] calculated the segmentation directly in the shape space. Expectation maximization [26] was used for optimizing the bias field/shape parameters and the segmentation iteratively in [66].

There have been recent attempts to add a shape prior to the graph cuts segmentation technique. [72] proposed the usage of an elliptical prior. [30] presented a method that uses a fixed shape template aligned with the image by the user input.

1.3 Validation of Medical Image Segmentation Methods

Validating medical image segmentation methods is generally hard because of lack of ground truth. For real images, even if manual expert segmentations exist, these cannot be treated as ground truth automatically because inter- and intra-observer variabilities are always present for manual segmentation. There have been algorithms developed for generating an accurate ground truth segmentation when multiple expert manual segmentations exists [82]. However, very few medical images and their ground truth segmentations are publicly available. Simulated data based on phantoms are very popular for validation purposes. In addition to having ground truth, one can also generate synthetic images with different level of intensity non-uniformity and noise, making it easy to analyze the effect of these factors on the segmentation method.

Possible quantitative measures for validation include the percentage of misclassified voxels and the Dice Similarity Measure (DSM) [90] w.r.t. the ground truth. DSM is defined as

$$\mathrm{DSM}^t_{a,b} = \frac{2 \cdot N^t_{a,b}}{N^t_a + N^t_b}, \tag{1.2}$$

where N^t_a and $N^t_b$are the voxels classified as tissue t with methods a and b, respectively, and $N^t_{a,b}$ is the number of voxels classified as tissue t by both methods. This measure takes into account both false positives and false negatives.

1.4 Contributions of the Thesis

In this thesis, we are interested in developing highly accurate and robust graph cuts-based method that automatically segment MR images. The images in which we are most interested are those that cannot be correctly segmented using intensity information alone. We developed models for robustly incorporating prior information such as spatial atlas and geometric or statistical shape priors into the graph cuts segmentation framework. The accuracy of the alignment of the atlas with the image and that of the shape prior are important to the accuracy of the segmentation. Therefore, we use an iterative approach and alternately calculate the atlas registration or the shape prior and the segmentation of the image, so that atlas registration and the shape prior could also be updated as more information becomes available. We also designed our models so that updates of the atlas registration and the shape prior are not merely based on the segmentation, in order to avoid a bad segmentation from sending the methods into vicious cycles of inaccurate shape priors and segmentations. Unlike some previous approaches, our final segmentation is not limited by the variabilities presented in the spatial atlas or by the shape priors. Finally, we also proposed a way of dealing with objects with curvy boundaries.

The major contributions of this thesis can be summarized as the development of models to

- simultaneously perform graph-based segmentation, non-rigid atlas registration and intensity non-uniformity correction for MR brain images (algorithm 1).

- incorporate statistical shape priors with graph cuts segmentation for MR images (algorithm 2).

- incorporate geometric shape priors with graph cuts segmentation for MR

images (algorithm 3).

We discuss, in more detail, the major contributions with the rest of the section.

1.4.1 Simultaneous Graph-Based Segmentation, Atlas Registration and Intensity Inhomogeneity Correction

Curvy boundaries, intensity inhomogeneity and poor contrast between tissues, these are all challenges faced by automated brain MR image segmentation. We propose an approach that simultaneously performs graph-based brain MR image segmentation, non-rigid atlas registration, and intensity inhomogeneity correction. We propose a graph-based segmentation approach that can deal with fine brain structures. Our segmentation accounts for spatial dependency and fine brain structures with efficient calculations. We address the intensity inhomogeneity by dividing the brain image into small blocks and modeling the intensity probability distributions of brain tissues with local GMMs. To capture both the global and local effects, our non-rigid registration uses two sets of parameters. We model the global registration of the atlas with the brain image using an affine transformation and use a B-spline representation for the deformable transformation. We consider both atlas/segmentation and atlas/brain image fit during registration to increase robustness and accuracy.

We validate our method on both synthetic and real T1 brain images. We compare the performance of our proposed method and its equivalent sequential method and show the effectiveness of conducting brain image segmentation and atlas registration simultaneously. We compare the results from our method with those from a well known brain image segmentation method, to evaluate the objective

performance of our approach.

1.4.2 Graph Cuts Segmentation with Statistical Shape Priors

Many techniques combining shape priors with level set-based segmentations have been reported in the literature. Although graph cuts segmentation method can also benefit from shape priors greatly, the area has yet to be explored. In this thesis, we studied the incorporation of two types of shape priors with graph-based segmentation.

We propose a segmentation method that incorporates statistical shape priors to the graph cuts technique for robust and accurate segmentations of medical images. We adopt the implicit shape representation proposed in [58]. The statistical shape information is obtained from a training set of segmented images. We take a unified approach and solve these two problems through one objective function. Two novel terms accounting for shape/image fit and shape/segmentation fit are introduced to the graph cuts energy function. The first term prevents an initial inaccurate segmentation from producing a vicious cycle of inaccurate shape priors and segmentations while the second allows the segmentation and shape fitting problems to interact with one another. Our proposed method is able to deal with complex shapes and shape variations while taking advantage of the globally efficient optimization by graph cuts. We demonstrated the effectiveness of our method on kidney images that cannot be segmented correctly using intensity information alone due to weak boundaries. The results show that our proposed method is able to handle images that cannot be correctly segmented without using any priors. Our segmentations are also not restricted by the shape variabilities represented by the training shapes.

1.4.3 Graph Cuts Segmentation with Geometric Shape Priors

Finally, we propose a novel segmentation method that incorporates geometric shape priors with the graph cuts technique for robust segmentations of MR images. The geometric shape priors do not require training, making the method fast and efficient. Again, we add terms accounting for shape prior/segmentation and shape prior/image fit to the graph cuts representation. We demonstrate the effectiveness of our method by correctly segmenting the left ventricle using concentric circles as shape priors and by segmenting the kidney using an ellipse as a shape prior. These images are hard to segment without priors because of reasons discussed previously in this chapter.

1.4.4 Comparison of the Three Algorithms

The priors we use in the three algorithms are respectively spatial atlases, statistical shape priors and geometric shape priors. The first two are constructed using registering and averaging hand-segmented brains and aligning segmented organs then applying Principle Component Analysis to reduce dimensionality. The last method does not require prior construction thus works faster, but the constraints applied here are less accurate. Which method among the three is applicable is determined by the type of available prior information. Since the accuracy of the segmentations are dependent on the match between the prior and the medical image of interest, the methods cannot segment accurately organs largely different from those used in the prior constructions.

Each of the algorithms uses graph cuts as the basis for segmentation, although in the first algorithm, we alter the standard graph cuts to address the issue of

fine structures prevalent in brains. Additional terms are introduced to the energy functions including a term measuring the match between the prior and the segmentation and a grounding term preventing vicious cycles. The former is calculated based on the spatial probabilistic information in the registered atlas in the first method and distances to the shape prior in the second and third method. The grounding term is measured through mutual information between the registered atlas and the image in the first algorithm and entropies of intensity distributions inside and outside the shape prior in the second and third algorithms.

Chapter 2
Preliminaries

In this chapter, we summarize the algorithms which make the basis of this thesis namely the graph cuts algorithm and the expectation-maximization method.

2.1 Graph Cuts Segmentation

2.1.1 Markov Random Field Modeling for Computer Vision

Let $\mathbf{F} = \{f_1, ..., f_N\}$ be a set of random variables at sites $S = \{1, ..., N\}$ w.r.t. a neighboring system N where $f_n \in L$. If $\mathbf{F}$ has the following properties, then $\mathbf{F}$ is said to be a Markov Random Field (MRF):

- Positivity: $P(\mathbf{f}) > 0 \qquad \forall \mathbf{f} \in \mathbf{F}$,
- -Markovianity: $P(f_n|f_{S-\{n\}}) = P(f_n|f_{N_n})$,

where $f_{N_n} = \{f_q | q \in N_n\}$ and N_n are neighboring sites of n. The Markovianity property says that each random variable f_n depends on other random variables through its neighbors. According to the Hammersley-Clifford theorem [34], the joint probability of f can be written as

$$P(\mathbf{f}) \propto \exp(-\Sigma_{c \in C} V_c(\mathbf{f})), \tag{2.1}$$

where $V_c(\mathbf{f})$ is the clique potential for c.

Many early vision problems such as image segmentation can be formulated in a Bayesian framework using Markov Random Field, where each image pixel is considered a site and a pair of neighboring pixels are considered a clique [31]. In

the case of image segmentation, the random variable at each site represents the classification of the pixel. If we denote the pixel intensities as $\mathbf{I} = \{I_1, ..., I_N\}$, then we could solve the pixel labeling problem by maximum a posteriori (MAP) estimation $P(\mathbf{f}|\mathbf{I}) \propto P(\mathbf{I}|\mathbf{f})P(\mathbf{f})$, where

$$P(\mathbf{f}) \propto \exp(-\Sigma_{n,q\in N} V_c(f_n, f_q)), \tag{2.2}$$

$$P(\mathbf{I}|\mathbf{f}) = \Sigma_n P(I_n|f_n). \tag{2.3}$$

We can also formulate this into an energy minimization problem with the following energy function:

$$\mathbf{E}(\mathbf{f}) = -\Sigma_n \log P(I_n|f_n) + \Sigma_{n,q\in N} V_c(f_n, f_q). \tag{2.4}$$

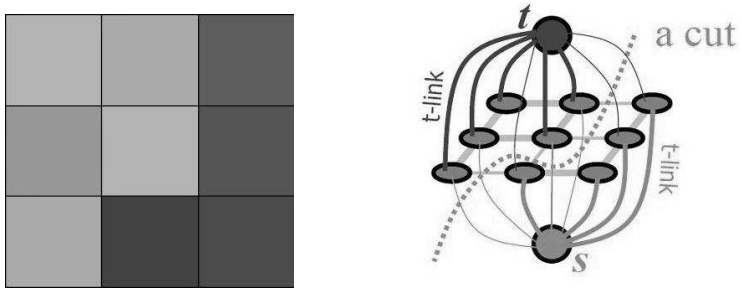
This problem is computational intractable and many approximations for the solution of MRF have been proposed including those using simulated annealing [35, 59], iterated conditional modes [29, 89], Newton descent [51] and mean field approximation [88, 39]. Although, the most popular solution is based on the graph cuts framework described in the rest of this section.

2.1.2 Efficient MRF Optimization using Graphs

The basic graph cuts image segmentation framework is developed in [12]. Following the MRF formulation, the standard form for the energy function in graph cuts segmentation is:

$$\mathbf{E}(\mathbf{f}) = \sum_n D_n(f_n) + \sum_{n,q\in C, f_n \neq f_q} V_c(f_n, f_q), \tag{2.5}$$

where $\mathbf{E}$ is the energy for label configuration $\mathbf{f}$ and $n, q \in C$ denotes that pixel n and m are neighbors. $D_n(f_n)$ measures the cost of assigning label f_n to pixel n (assignment cost) while $V_{n,q}(f_n, f_q)$ measures the cost of assigning neighboring



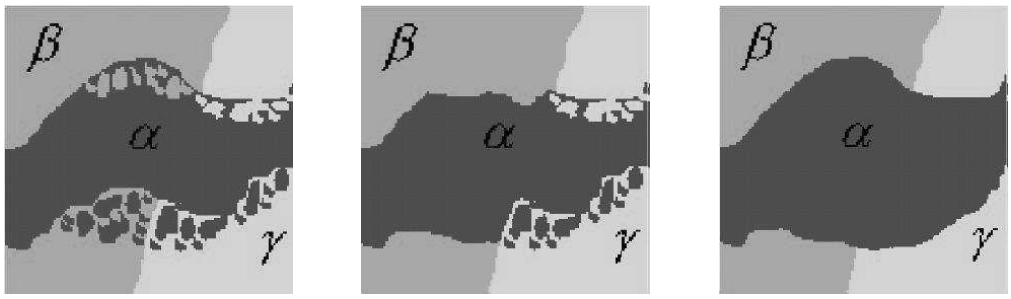
(a) (b)

Figure 2.1: A simple 2D 3x3 image (a) and its corresponding graph/the minimum cut on the graph (b). Figure based on [13].

pixels n and q different labels f_n and f_q (separation cost). The graph cuts method has become popular because it provides globally optimal efficient solutions in N-dimensional settings for binary segmentations when the separation costs satisfy the regularity condition [44].

The idea for binary segmentation using graph cuts is as follows. An image is mapped onto a weighted undirected graph where each pixel is represented as a node and each pair of neighboring pixels is linked by an edge. Two additional "terminal" nodes, the source s and the sink t, represent the object and the background. Every non-terminal node is linked to s and t through edges called t-links. A cut on the graph divides the nodes into two sets: one that is connected to the source s and one that is connected to the sink t, hence producing a binary segmentation. The cost of a cut is the sum of weights of all the edges severed by the cut. Edge weights on the graph are formulated so that the minimum cut (i.e. the cut with the minimum cost) on the graph produces the label configuration that minimizes the energy function in equation (2.5). There are numerous algorithms that can solve the minimum cut problem in polynomial time [2, 11]. An example of an image and its corresponding graph/the minimum cut on the graph are shown in Fig. 2.1(a) and Fig. 2.1(b).

Graph cuts can be extended to multiple labels using $\alpha-\beta$ swap or α-expansion



(a) Initial labeling (b) $\alpha - \beta$-swap (c) α-expansion

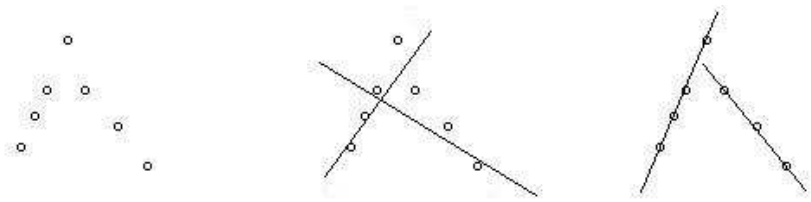
Figure 2.2: $\alpha - \beta$-swap and α-expansion moves. Figure based on [78].

[13]. The $\alpha - \beta$ swap move exchanges the labels of an arbitrary set of pixels labeled α and another arbitrary set labeled β (Fig. 2.2(b)). The α-expansion move assigns label α to an arbitrary set of pixels (Fig. 2.2(c)). In the case of multiple labels, exact minimum on the graph can be found efficiently for a specific separation cost $V_{n,q}(f_n, f_q) = |f_n - f_q|$ [37], but the problem is generally NP-hard [78]. Both $\alpha - \beta$ swap and α-expansion moves guarantee that the energy decreases constantly. In addition, the α-expansion move generates a local minimum that is within a known factor of the global minimum when the separation cost is a metric [78].

The specific separation cost for which an exact minimum is possible in the case of multi-labels ($|f_n - f_q|$) does not preserve discontinuity, since the penalty is allowed to grow arbitrarily large. Adding a cap to the maximum value the penalty can take helps to preserve discontinuity. Generalized Potts models are useful for this purpose and can be defined as:

$$V_{n,q}(f_n, f_q) = \lambda_{n,q}(1 - \delta(f_n - f_q)), \tag{2.6}$$

where $\lambda_{n,q}$s are non-negative coefficients and $\delta()$ is a delta function. Graph cuts can find approximate solutions within the factor of 2 of the global optimums for multi-label energy functions with generalized Potts models as separation costs



(a) Data sampled from two lines (b) Initial guess (c) Final guess

Figure 2.3: "Line fitting" using an EM-style method.

2.2 Expectation-Maximization in Image Analysis

Many problems in computer vision, including image segmentation, can be framed as estimation problems where we want to find the underlying parameters when the observed data is incomplete. A simple example of such a problem is the "line fitting" problem, where we sample data from two lines without knowing which data points came from which line. In the "line fitting" problem, our goal is to calculate the parameters for the underlying lines with the assignment of points to lines as hidden data (or unobserved data). An intuitive way to solve this problem would be to make an estimation of the assignment of points to lines, then calculate the line parameters based on the estimations. Once we have the line parameters, we can improve our estimations of the assignment of points to lines and based on the new estimations recalculate the line parameters. By repeating these two processes, we should be able to find a reasonable solution (Fig. 2.3). Expectation-Maximization (EM) proposed in [26] formalizes this iterative method and proves that in certain cases, the algorithm is guaranteed to converge. The rest of this section discusses the EM method more formally and shows how it can be applied to calculating parameters for Gaussian mixtures with incomplete data, a problem that is closely related to MR image segmentation.

2.2.1 The Expectation-Maximization Algorithm

Suppose we have a set of observable random variables $\mathbf{X} = \{X_1, ..., X_N\}$ and a set of hidden random variables $\mathbf{H} = \{H_1, ..., H_N\}$ drawn from an unknown distribution $\underline{\theta}$, and X_n and H_n are dependent. We can find the underlying distribution $\underline{\theta}$ by computing the maximum likelihood:

$$\underline{\theta}^* = \arg\max \log P(\mathbf{X}|\underline{\theta}), \tag{2.7}$$

where

$$P(\mathbf{X}|\underline{\theta}) = \Sigma_{\mathbf{h}} P(\mathbf{X}, \mathbf{h}|\underline{\theta}) = \Sigma_{\mathbf{h}} P(\mathbf{X}|\mathbf{h}, \underline{\theta}) P(\mathbf{h}|\underline{\theta}), \tag{2.8}$$

and $\mathbf{h}$ are the realizations of $\mathbf{H}$. This computation is not always possible and EM produces a good approximation. As mentioned previously, EM is an iterative method. If we can define a lower bound function $l(\underline{\theta}|\underline{\theta}^t)$ at each iteration t for the likelihood function $L(\underline{\theta}) = \log P(\mathbf{X}|\underline{\theta})$, where $l(\underline{\theta}|\underline{\theta}^t) = L(\underline{\theta})$ for $\theta = \theta^t$, and maximize it as shown in Fig. 2.2.1, then we can gradually close in on the global maximum of $L(\underline{\theta})$. This process is strictly non-decreasing, since

$$L(\underline{\theta}^{t+1}) \geq l(\underline{\theta}^{t+1}|\underline{\theta}^t) \geq l(\underline{\theta}^t|\underline{\theta}^t) = L(\underline{\theta}^t). \tag{2.9}$$

$$E_{\mathbf{H}|\mathbf{X},\underline{\theta}^t}[\log P(\mathbf{X}, \mathbf{H}|\underline{\theta})] = \Sigma_{\mathbf{h}} P(\mathbf{h}|\mathbf{X}, \underline{\theta}^t) \log P(\mathbf{X}, \mathbf{h}|\underline{\theta}) \tag{2.10}$$

has been proven to be such a function and is used in EM. Skipping all the details, the two steps in EM work as follows and together maximize $l(\underline{\theta}|\underline{\theta}^t) = E_{\mathbf{H}|\mathbf{X},\underline{\theta}^t}[\log P(\mathbf{X}, \mathbf{H}|\underline{\theta})]$ during each iteration.

1) **E(stimation) step**: estimate the likelihood for each realization of the hidden parameters based on the current set of parameters $P(\mathbf{h}|\mathbf{X}, \underline{\theta}^t)$,

2) **M(aximization) step**: maximize the lower bound function to find the best parameters $\arg\max_{\underline{\theta}} E_{\mathbf{H}|\mathbf{X},\underline{\theta}^t}(\log P(\mathbf{X}, \mathbf{H}|\underline{\theta}))$.

Detail derivations of the EM method can be found in [53, 54]. In the "line fitting" example mentioned previously, we assigned data points to lines during each

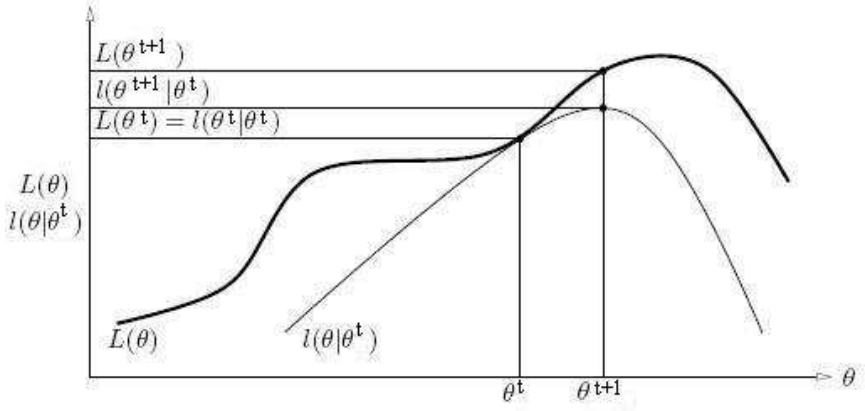


Figure 2.4: A graphical interpretation of the convergence of EM. Image based on [8].

iteration. This does not strictly follow the formulation of EM, since with EM, we need to estimate the probability of the data points belonging to each of the lines. Methods that do not strictly follow the mathematical framework of EM cannot be guaranteed convergence, although they are still popular because they generally work. The segmentation proposed in this thesis follows the style of EM, although we assign each voxel one label and do not estimate the probability of the voxels belonging to each of the labels. As such, it is not strict EM.

2.2.2 Gaussian Mixture Models and the EM algorithm

Let us think about the problem where we have a Gaussian mixture model (GMM) made up of multiple Gaussians and we sample data from these Gaussians to calculate their parameters. We denote the Gaussian mixture as $\underline{\theta} = \{\mu_d, \Sigma_d, \pi_d\}, d \in \{1...D\}$ where D is the number of Gaussians, μ_d, Σ_d, π_d are respectively the mean, variance, and prior probability of the d-th Gaussian. Say we observed N data points: $\mathbf{X} = \{X_1, ..., X_N\}$ and we do not know which points came from which Gaussian. The EM method works beautifully in this case. First, we introduce a

variable:

$$\mathbf{H} = \left(\begin{pmatrix} H_1^1 \\ \vdots \\ H_1^d \\ \vdots \\ H_1^D \end{pmatrix}, \ldots, \begin{pmatrix} H_n^1 \\ \vdots \\ H_n^d \\ \vdots \\ H_n^D \end{pmatrix}, \ldots, \begin{pmatrix} H_N^1 \\ \vdots \\ H_N^d \\ \vdots \\ H_N^D \end{pmatrix} \right), \tag{2.11}$$

where $H_n^d = 0$ when data point n does not come from Gaussian d and $H_n^d = 1$ when it does. During each iteration of EM, the E-step estimates the hidden data $\mathbf{H}$ according to

$$< H_n^d >= 1 \times P(H_n^d = 1 | X_n, \underline{\theta}^t) + 0 \times P(H_n^d = 0 | X_n, \underline{\theta}^t). \tag{2.12}$$

The M-step calculates the underlying parameters of the Gaussian by maximizing

$$E_{\mathbf{H}|\mathbf{X},\underline{\theta}^t}(\log P(\mathbf{X}, \mathbf{H}|\underline{\theta})) = \Sigma_n \Sigma_d < H_n^d > \log P(X_n, H_n^d|\underline{\theta}). \tag{2.13}$$

The M-step can be solved analytically by taking the first order partial derivative w.r.t. each model parameter and have a closed form solution:

$$\frac{\partial E_{\mathbf{H}|\mathbf{X},\underline{\theta}^t}(\log P(\mathbf{X}, \mathbf{H}|\underline{\theta}))}{\partial \mu_d} = \frac{\partial}{\partial \mu_d} \Sigma_n \Sigma_d < H_n^d > \log P(X_n, H_n^d|\underline{\theta}) = 0, \tag{2.14}$$

where $P(X_n, H_n^d|\underline{\theta}) = P(X_n|\mu_d, \Sigma_d) \cdot \pi_d$ and $P(X_n|\mu_d, \Sigma_d)$ is Gaussian probability. After some derivation:

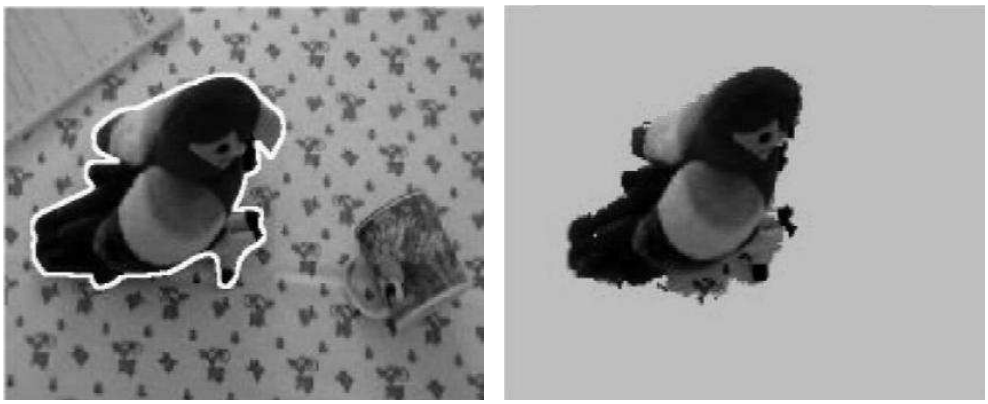
$$\frac{\partial}{\partial \mu_d} \Sigma_n \Sigma_d < H_n^d > \log P(X_n, H_n^d|\underline{\theta}) \propto \Sigma_n \Sigma_d < H_n^d > (X_n - \mu_d) = 0, \tag{2.15}$$

giving us the updated equation:

$$\mu_d = \frac{\sum_n < H_n^d > \cdot X_n}{\sum_n < H_n^d >}. \tag{2.16}$$

Similarly, the updated equations for Σ_d and π_d are respectively:

$$\Sigma_d = \frac{\sum_n < H_n^d > \cdot (X_n - \mu_d)^2}{\sum_n < H_n^d >}, \pi_d = \frac{\sum_n < H_n^d >}{\sum_n 1}. \tag{2.17}$$



(a) Original image (b) Segmentation based on GMM

Figure 2.5: Segmentation of an image using GMM. Image based on [75].

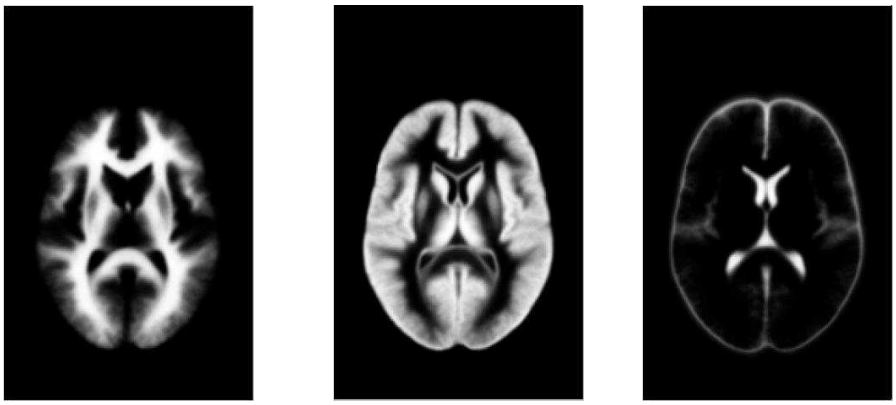
Gaussian mixture models can be applied to image segmentation problems by assuming a number of Gaussians for the intensity probability distribution of each label. Then, the assignment of pixels to Gaussians would also determine the classification of the pixels. Fig. 2.5 shows a segmentation of an image using GMM. Gaussian mixture models have been used in segmentation and tracking algorithms in [28, 50, 75]. GMM-EM has been used in medical image segmentation in [47, 83].

Chapter 3
Simultaneous Geo-Cuts Segmentation, Non-Rigid Atlas Registration and Intensity Non-Uniformity Correction for MR Brain Images

3.1 Chapter Overview

Segmentation of brain MR images leads to many clinical and research applications. Factors such as intensity inhomogeneity and poor contrast at tissue boundaries make automated brain MR image segmentation difficult. In the medical imaging community, spatial atlas-aided segmentation techniques have been proposed [21, 51, 64, 19, 47, 29, 5, 87, 84, 61, 3, 65]. Early atlas-based techniques performed atlas registration and brain image segmentation sequentially. The accuracy of these methods is highly dependent on the accuracy of the atlas registration. Recent methods focus on the unification of registration and segmentation. However, a method has not been suggested which precisely models spatial relations between voxel neighbors during image segmentation. Spatial neighboring relations are especially hard to model for brain images because they contain many thin structures.

In this chapter, we propose a novel approach that simultaneously performs graph-based brain MR image segmentation, non-rigid atlas registration, and intensity inhomogeneity correction. For segmentation, we define a Riemannian space in which the ideal segmentation boundaries, including those around fine brain structures, have the smallest Riemannian length. The geodesic contours in this space are found efficiently using graph cuts. Our segmentation accounts for spatial dependency and fine brain structures with efficient calculations. To increase robustness



(a) white matter (b) gray matter (c) CSF

Figure 3.1: MNI spatial brain atlas.

and accuracy, we consider both atlas/segmentation and atlas/brain image fit during registration. We address the intensity inhomogeneity by dividing the brain image into blocks and modeling the intensity probability distributions of brain tissues with local Gaussian mixture models (GMMs). By finding the segmentation and optimizing registration and GMM parameters through one objective function, we unify the three problems and allow them to interact with one another.

We use the widely available MNI brain atlas in the Statistical Parametric Mapping (SPM) package [1]. This atlas is constructed from the average of hundreds of normal brain scans. The brain scans are registered with each other through affine transformation [73]. The spatial atlas contains probabilistic information of the major brain tissues including white matter (WM), gray matter (GM) and cerebrospinal fluid (CSF). Fig. 3.1 shows the spatial atlas for WM, GM and CSF in the brain.

We validate our method on both synthetic and real T1 brain images. The proposed method correctly segments fine structures in our test images. In order to find out whether unified segmentation/atlas registration is effective, we also implement a sequential version of our proposed method. We compare the performance of our proposed method and its equivalent sequential method and show the effectiveness

of conducting brain image segmentation and atlas registration simultaneously.

The rest of the chapter is organized as follows. In section 3.2, we first summarize our graph-based segmentation method and our models for non-rigid registration and intensity non-uniformity correction. We then describe our novel objective function, which links the three problems and allows them to be solved simultaneously. In section 3.3, we show how our energy function is optimized through an expectation maximization (EM)-style approach. We describe in section 3.4, our datasets and methods for evaluating our proposed algorithm and show the experimental results in section 3.5.

3.2 Objective Function

In this section, we first give a summary of our graph-based segmentation methods. We then discuss our models for non-rigid registration and intensity inhomogeneity correction. We end by introducing our novel energy function. In the next section, we will show how this function is optimized.

We represent brain images as a 1D array $\mathbf{I} = (I_1, \ldots, I_N)$, where I_n is the intensity of voxel n and N is the total number of voxels. We assign each voxel one of three labels: white matter (WM), gray matter (GM), and cerebrospinal fluid (CSF). Our segmentations are represented as label configuration of all voxels, $\mathbf{f} = (f_1, \ldots, f_N)$, where $f_n \in \{\mathrm{WM}, \mathrm{GM}, \mathrm{CSF}\}$.

Geodesics via Graph Cuts

Graph cuts-based segmentation methods using standard separation costs determined by intensity differences between neighboring voxels such as the one in equation (3.1) do not work well in preserving fine structures prevalent in brain images. For example, Fig. 3.2(a) shows a slice of brain MRI with many fine structures

and Fig. 3.2(b) shows its segmentation using graph cuts with equation (3.1) as separation cost. Fig. 3.2(b) fails to capture the thin structures of white matter and CSF in Fig. 3.2(a).

$$V_c(f_n, f_q) = \frac{1}{1 + (I_n - I_q)^2} \tag{3.1}$$

We propose incorporating the geo-cuts method to overcome the problem mentioned above. The geo-cuts method developed in [10] combines the benefits of graph cuts and geodesic active contours [16, 86, 60]. The idea is to first define an image induced Riemannian space so that ideal segmentation boundary in the image is equivalent to the geodesic curve in this space. We will give a brief review of Riemannian geometry later on in this paper but the Riemannian metric replace the second term in equation (2.5) as separation cost. Boykov showed in [10] how to build a grid graph and set its edge weights so that lengths of contours in an anisotropic Riemannian space are arbitrarily close to the costs of cuts on this graph, as in equation (3.2):

$$w_k(n) = C_{grid} \cdot \frac{|e_k^n|^3 \det D(n)}{({e_k^n}^T D(n) e_k^n)^2}, \tag{3.2}$$

where $w_k(n)$ is the edge weight for edge e_k^n at voxel n, $D(n)$ is the Riemannian metrics at voxel n, and C_{grid} is a constant determined by the grid graph. The geodesic curve representing the ideal segmentation boundary is then equivalent to the minimum cut on its corresponding graph, which can be found efficiently through graph cuts.

Riemannian Geometry for Image Segmentation

In a Euclidean space, the length of a vector τ can be calculated as

$$L^E(\tau) = \sqrt{\tau^T \tau}, \tag{3.3}$$

and does not depend on the location or orientation of the vector. In a Riemannian space, this is no longer true. In order to measure length in a Riemannian space, a metric or a collection of symmetric and positive definite matrices $D(n)$ at each point n is given. The length of a vector in the local neighborhood of n can then be calculated as

$$L_n^R(\tau) = \sqrt{\tau^T D(n)\tau}. \tag{3.4}$$

A Euclidean space can be considered as a special type of Riemannian space and the Riemannian metric in this case is the identity matrix for all points.

In order for the geodesic contour to represent the ideal segmentation, we need our Riemannian metric to satisfy the following:

1) In an area where gradient magnitude of the image is low, Riemannian lengths of vectors should not depend on the orientations of the vectors and should be large for all orientations, since it is equally unlikely for boundaries to exist in this area in all directions.

2) In an area with strong contours that are straight, Riemannian lengths of vectors should depend on the orientations of the vectors. The contours are likely to correspond with object boundaries in the image whereas their normals very unlikely. Therefore, vectors in the direction of the contours should be short in the Riemannian space and in the normal directions long. We can preserve fine structures and prevent undesirable smoothing by setting the Riemannian length of the former to be very short and the latter very long.

3) In an area with strong contours that have large curvature, Riemannian lengths of vectors should again not depend on the orientations of the vectors and should be short for all orientations, since in this case, it is possible for boundaries to exist in all directions and not just the direction of the contour.

The specific Riemannian metric $D(n)$ we use is as follows:

$$\begin{aligned} D(n) = \; & (1 - g(|\nabla I_n|))\mathbf{I}_{ident} \\ & + \; g(|\nabla I_n|) \cdot (f(K_n)\varepsilon \mathbf{I}_{ident} + (1 - f(K_n))|\nabla I_n|\mathbf{u}\mathbf{u}^T), \end{aligned} \tag{3.5}$$

where g and f are strictly increasing functions between 0 and 1. $\mathbf{I}_{ident}$ is the identity matrix, and ε is a small number that is less than 1. $D(n)$, I_n, $\mathbf{u}$ and K_n are respectively the Riemannian metric, the image intensity, the unit vector in the direction of the image gradient and the curvature at point n. Fig. 3.2(c) shows the segmentation result of the brain MRI in Fig. 3.2(a) using our set up. Compared to Fig. 3.2(b), fine structures are better preserved in Fig. 3.2(c).

3.2.1 Non-Rigid Registration of the Spatial Atlas to the Brain Image

We denote the atlas as $\mathbf{A} = (\mathbf{A}_1, \ldots, \mathbf{A}_m, \ldots, \mathbf{A}_M)$, where $\mathbf{A}_m = (A_m^{\mathrm{WM}}, A_m^{\mathrm{GM}}, A_m^{\mathrm{CSF}})^{\mathrm{T}}$ and $A_m^{\mathrm{WM}}, A_m^{\mathrm{GM}}$, and A_m^{CSF} are respectively the probability for voxel m to take the label WM, GM and CSF. The atlas could also be written as $\mathbf{A} = (\mathbf{A}^{WM}, \mathbf{A}^{\mathrm{GM}}, \mathbf{A}^{\mathrm{CSF}})^T$, where

$$\mathbf{A}^{\mathrm{WM}} = (A_1^{\mathrm{WM}}, ..., A_m^{\mathrm{WM}}, ..., A_M^{\mathrm{WM}}),$$

$$\mathbf{A}^{\mathrm{GM}} = (A_1^{\mathrm{GM}}, ..., A_m^{\mathrm{GM}}, ..., A_M^{\mathrm{GM}}),$$

$$\mathbf{A}^{\mathrm{CSF}} = (A_1^{\mathrm{CSF}}, ..., A_m^{\mathrm{CSF}}, ..., A_M^{\mathrm{CSF}}).$$

Following the approach in [69], we use an affine transformation and a free form deformation (FFD) based on B-splines [45, 46] to represent our registration. The affine transformation models the global registration between the atlas and the brain image with rotation, translation, scaling and shearing. The FFD models local

deformation by manipulating a mesh of control points. FFD has been previously applied to the analysis of cardiac images [6].

Let $\mathbf{R} = \{R_l,\ \phi_{i,j,k},\ 1 \le l \le 12,\ 1 \le i \le n_x,\ 1 \le j \le n_y,\ 1 \le k \le n_z\}$ denote the registration parameters, where R_ls are the affine transformation parameters and $\phi_{i,j,k}$s are a mesh of $n_x \times n_y \times n_z$ control points with uniform spacing for the FFD. Transformation $\mathbf{R}$ maps a point (x, y, z) in the original atlas $\mathbf{A}$ to $(\tilde{x}, \tilde{y}, \tilde{z})$ in the registered atlas $\tilde{\mathbf{A}}(\mathbf{R})$ according to:

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{pmatrix} = T_{\text{affine}}(x, y, z) + T_{\text{FFD}}(x, y, z), (\tilde{x}, \tilde{y}, \tilde{z})^T = T_{\text{affine}}(x, y, z) + T_{\text{FFD}}(x, y, z), \tag{3.6}$$

where T_{affine} is the affine transformation and can be written as

$$T_{\text{affine}}(x, y, z) = \begin{pmatrix} R_1 & R_2 & R_3 \\ R_4 & R_5 & R_6 \\ R_7 & R_8 & R_9 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} R_{10} \\ R_{11} \\ R_{12} \end{pmatrix}, \tag{3.7}$$

and T_{FFD} is the B-spline-based FFD and can be written as

$$T_{\text{FFD}}(x, y, z) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(r) B_m(s) B_n(t) \phi_{i+l, j+m, k+n}, \tag{3.8}$$

where $i = \lfloor x/n_x \rfloor - 1,\ j = \lfloor y/n_y \rfloor - 1,\ k = \lfloor z/n_z \rfloor - 1,\ r = x/n_x - \lfloor x/n_x \rfloor,\ s = y/n_y - \lfloor y/n_y \rfloor$, and $t = z/n_z - \lfloor z/n_z \rfloor$ and $B_l(\bullet)$s are the l-th basis function of the cubic B-spline [77]:

$$\begin{aligned} B_0(r) &= (1 - r)^3/6, \\ B_1(r) &= (3r^3 - 6r^2 + 4)/6, \\ B_2(r) &= (-3r^3 + 3r^2 + 3r + 1)/6, \\ B_3(r) &= r^3/6. \end{aligned}$$

As the values of the basis functions at each point are fixed, the control points

serve as parameters for the local deformation and the transformations of points are governed by the 4x4x4 control points in their local areas.

3.2.2 Mixture of Gaussians

To address intensity non-uniformity in MR images, we use a straightforward approach and divide the image into small blocks, modeling the intensity probability distributions of brain tissues locally for each block with a GMM. We use one Gaussian for each brain tissue within a block. The size of the blocks is predetermined. Let $\underline{\theta} = \{\mu_k^b,\ \Sigma_k^b,\ \pi_k^b,\ k \in \mathrm{WM}, \mathrm{GM}, \mathrm{CSF},\ b \in 1, \ldots, B\}$ denote the GMM parameters. Here, $\mu_k^b,\ \Sigma_k^b,\ \pi_k^b$ are respectively the mean, variance, and prior probability of the Gaussian associated with tissue type k in block b . B is the total number of blocks. The prior probabilities within each block add up to 1 ($\sum_k \pi_k^b = 1$).

3.2.3 Energy Function

Our problem can be formulated as: given a brain image $\mathbf{I}$ and a spatial brain atlas $\mathbf{A}$, we wish to simultaneously find the segmentation of the brain image $\mathbf{f}$ and the parameters that register the atlas with the image $\mathbf{R}$. We define the energy functional in equation (3.9) to guide the image segmentation and atlas registration.

$$\mathbf{E}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{R}) = \mathbf{E}_{\mathrm{I}}(\mathbf{I}, \underline{\theta}, \mathbf{f}) + \mathbf{E}_{\mathrm{S}}(\mathbf{I}, \mathbf{f}) + \mathbf{E}_{\mathrm{P}}(\mathbf{f}, \mathbf{R}) + \mathbf{E}_{\mathrm{A}}(\mathbf{I}, \mathbf{R}). \tag{3.9}$$

The first term measures how well the voxel labels and the GMM parameters fit together given the voxel intensities and can be written as

$$\mathbf{E}_{\mathrm{I}}(\mathbf{I}, \underline{\theta}, \mathbf{f}) = -\sum_n \log \mathrm{P}(I_n | f_n, \underline{\theta}) = -\sum_n \log \mathrm{P}(I_n | \mu_{f_n}^b, \Sigma_{f_n}^b, \pi_{f_n}^b), \tag{3.10}$$

where $n \in b$. The second term is the Riemannian length of the segmentation boundary. The third term denotes the correctness of the non-rigid registration

between the atlas and the current segmentation and can be written as

$$\mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R}) = -\sum_{n} \log \mathrm{P}(f_n|\tilde{\mathbf{A}}(\mathbf{R}))) = -\sum_{n} \log(\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})). \tag{3.11}$$

The fourth term measures how well the atlas itself fits with the brain image through measuring mutual information [20, 79]:

$$\mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R}) = -\sum_{i}\sum_{\mathbf{a}} P(i,\mathbf{a}) \log(\frac{P(i,\mathbf{a})}{P_I(i)P_A(\mathbf{a})}), \tag{3.12}$$

where $P_I(i)$ is the probability for a voxel to have intensity i in the brain image, $P_A(\mathbf{a})$ the probability for a voxel to have intensity $\mathbf{a} = (a^{\mathrm{WM}}, a^{\mathrm{GM}}, a^{\mathrm{CSF}})^T$ in the atlas and $P(i,\mathbf{a})$ the joint probability of a voxel to have intensity i in the brain image and $\mathbf{a}$ in the registered atlas. $a^{\mathrm{WM}}, a^{\mathrm{GM}},$ and a^{CSF} are respectively the probabilities for a voxel to have the label WM, GM and CSF.

3.3 Energy Minimization

As mentioned previously, we use an EM style approach to minimize the energy function presented in equation (3.9) and alternately update the GMM and registration parameters while fixing the segmentation (maximization step) and use the GMMs and the registered atlas to facilitate the image segmentation (estimation step).

3.3.1 M(aximization) Step

Differentiating equation (3.10) w.r.t. μ_k^b, we get:

$$\frac{\partial \mathbf{E}(\mathbf{I},\underline{\theta},\mathbf{f},\mathbf{R})}{\partial \mu_k^b} = -\frac{\partial}{\partial \mu_k^b} \sum_{n\in b, f_n=k} \log \mathrm{P}(I_n|\mu_k^b,\Sigma_k^b,\pi_k^b) = 0, \tag{3.13}$$

giving us the updated equation:

$$\mu_k^b = \frac{\sum_{n\in b, f_n=k} I_n}{\sum_{n\in b. f_n=k} 1}. \tag{3.14}$$

Similarly, the updated equations for Σ_k^b and π_k^b are respectively:

$$\Sigma_k^b = \frac{\sum_{n\in b, f_n=k}(I_n-\mu_k^b)^2}{\sum_{n\in b, f_n=k}1}, \pi_k^b = \frac{\sum_{n\in b, f_n=k}1}{\sum_{n\in b}1}. \tag{3.15}$$

We use the regular step gradient descent method to update the registration parameters. The update equation can be written as

$$\begin{aligned} R_l^{t+1} &= R_l^t - \lambda\left(\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial R_l} + \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial R_l}\right) \\ \phi_{i,j,k}^{t+1} &= \phi_{i,j,k}^t - \lambda\left(\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial \phi_{i,j,k}} + \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial \phi_{i,j,k}}\right) \end{aligned}, \tag{3.16}$$

where λ is the step size for the gradient descent optimization.$R_l^t, \phi_{i,j,k}^t$ and $R_l^{t+1}, \phi_{i,j,k}^{t+1}$ are the registration parameters at time t and t+1. The terms $\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial R_l}$ and $\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial \phi_{i,j,k}}$ are updated according to

$$\begin{aligned} \frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial R_l} &= -\sum_n \frac{\partial \log(\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R}))}{\partial R_l} = -\sum_n \frac{1}{\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}\cdot\frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial R_l} \\ \frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial \phi_{i,j,k}} &= -\sum_n \frac{\partial \log(\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R}))}{\partial \phi_{i,j,k}} = -\sum_n \frac{1}{\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}\cdot\frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \phi_{i,j,k}} \end{aligned}, \tag{3.17}$$

where

$$\begin{aligned} \frac{\partial \tilde{\mathbf{A}}_n^{F_{fn}}(\mathbf{R})}{\partial R_l} &= \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{x}}\cdot\frac{\partial \tilde{x}}{\partial R_l} + \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{y}}\cdot\frac{\partial \tilde{y}}{\partial R_l} + \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{z}}\cdot\frac{\partial \tilde{z}}{\partial R_l} \\ \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \phi_{i,j,k}} &= \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{x}}\cdot\frac{\partial \tilde{x}}{\partial \phi_{i,j,k}} + \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{y}}\cdot\frac{\partial \tilde{y}}{\partial \phi_{i,j,k}} + \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{z}}\cdot\frac{\partial \tilde{z}}{\partial \phi_{i,j,k}} \end{aligned}. \tag{3.18}$$

Here, $\left(\frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{x}}, \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{y}}, \frac{\partial \tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})}{\partial \tilde{z}}\right)$ is the image gradient for $\tilde{\mathbf{A}}_n^{f_n}(\mathbf{R})$ whereas $\left(\frac{\partial \tilde{x}}{\partial R_l}, \frac{\partial \tilde{y}}{\partial R_l}, \frac{\partial \tilde{z}}{\partial R_l}\right)$ and $\left(\frac{\partial \tilde{x}}{\partial \phi_{i,j,k}}, \frac{\partial \tilde{y}}{\partial \phi_{i,j,k}}, \frac{\partial \tilde{z}}{\partial \phi_{i,j,k}}\right)$ can be found easily by differentiating equations (3.7) and (3.8). The terms $\frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial R_l}$ and $\frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial \phi_{i,j,k}}$ are updated according to

$$\begin{aligned} \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial R_l} &= -\sum_i\sum_{\mathbf{a}} \frac{\partial P(i,\mathbf{a})}{\partial R_l}\log(\frac{P(i,\mathbf{a})}{P_I(i)}) \\ \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial \phi_{i,j,k}} &= -\sum_i\sum_{\mathbf{a}} \frac{\partial P(i,\mathbf{a})}{\partial \phi_{i,j,k}}\log(\frac{P(i,\mathbf{a})}{P_I(i)}) \end{aligned}. \tag{3.19}$$

For derivations of equation (3.19) and details on the calculation of $\frac{\partial P(i,\mathbf{a})}{\partial R_l}$ and $\frac{\partial P(i,\mathbf{a})}{\partial \phi_{i,j,k}}$, please refer to [74].

3.3.2 E(stimation) Step

Since we have three labels, segmentation of the image consists of three stages based on the α-expansion. In the first stage, voxels previously classified as gray matter or CSF are either re-labeled as white matter or remain with the same label while voxels previously labeled as white matter do not change labels. The second stage re-classifies a set of the voxels previously labeled white matter or CSF as gray matter and the third stage a set of those previously labeled white matter or gray matter as CSF.

During each step, we create a graph with nodes corresponding to voxels and two additional terminal nodes. The first and third terms in our energy function are applied to the graph as t-links. We follow the derivation in [10] to set the Riemannian lengths as n-links in the graph. The fourth term does not apply. The minimum cut on the graph is then computed using max-flow [22].

3.4 Evaluation

The simulated data in this study come from the brain phantom from McConnell Brain Imaging Center [19]. The brain phantom consists of a brain model and a MRI simulator. The brain model comes from classification of a normal subject and contains probabilities for each voxel to belong to WM, GM and CSF. The MR simulator uses this model to generate images with different RF non-uniformity and noise levels. Since these images are widely used by brain image segmentation methods in the literature and quantitative performance by existing methods are readily available, we use these images for the quantitative evaluation of our method. To verify the effectiveness of jointly segmenting the brain image and registering the brain atlas with the image, we compare the performance of our proposed method

SUMMARY OF ALGORITHM

Initialization

The GMM parameters are initialized using pre-defined values. We initialize the rigid motion by registering the brain image and the atlas. The initial non-linear deformation is set to zero.

Iterative Minimization

REPEAT

E(stimation) Step

Find segmentation using graph cuts:

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} E(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{R}).$$

M(inimization) Step

Learn Gaussian mixture model parameters:

$$\mu_k^b = \frac{\sum_{n\in b, f_n=k} I_n}{\sum_{n\in b. f_n=k} 1},\ \Sigma_k^b = \frac{\sum_{n\in b, f_n=k} (I_n-\mu_k^b)^2}{\sum_{n\in b, f_n=k} 1}\ ,\ \pi_k^b = \frac{\sum_{n\in b, f_n=k} 1}{\sum_{n\in b} 1}.$$

Update registration parameters through gradient descent optimization until convergence:

$$R_l^{t+1} = R_l^t - \lambda \left(\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial R_l} + \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial R_l} \right),$$
$$\phi_{i,j,k}^{t+1} = \phi_{i,j,k}^t - \lambda \left(\frac{\partial \mathbf{E}_{\mathrm{P}}(\mathbf{f},\mathbf{R})}{\partial \phi_{i,j,k}} + \frac{\partial \mathbf{E}_{\mathrm{A}}(\mathbf{I},\mathbf{R})}{\partial \phi_{i,j,k}} \right),$$

UNTIL GMM/REGISTRATION PARAMETERS CONVERGE

and the performance of the same method without the simultaneous component. We also compare our method with an existing brain MRI segmentation method - expectation-maximization segmentation (EMS) presented in [47], to evaluate the effectiveness of our method objectively. We use as quantitative measures, the percentage of misclassified voxels over all tissues and the Dice Similarity Measure over each tissue w.r.t. the ground truth.

Real images are much more complex than the synthetic ones from a brain phantom, therefore validating our method on real brain images is necessary. 18 real volumetric T1-weighted brain images and their hand segmentations are available

from the Internet Brain Segmentation Repository (IBSR) at the Massachusetts General Hospital. We test our proposed method on these images and inspect its performance against the hand segmentations.

3.4.1 Equivalent Sequential Method

We implemented a sequential method that is equivalent to our proposed unified method, in order to validate the effectiveness of unified segmentation and atlas registration. In this sequential method, we first register the image with the atlas through minimizing:

$$\mathbf{E}_{\mathrm{A}}(\mathbf{I}, \mathbf{R}) = -\sum_{i} \sum_{\mathbf{a}} P(i, \mathbf{a}) \log(\frac{P(i, \mathbf{a})}{P_I(i) P_A(\mathbf{a})}). \tag{3.20}$$

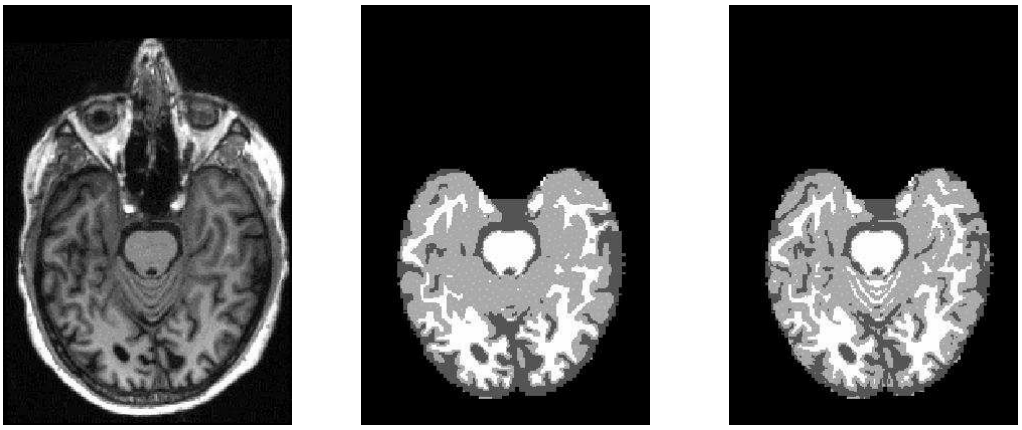
Using the registered atlas, we then perform image segmentation through minimizing the energy function:

$$\mathbf{E}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{R}) = \mathbf{E}_{\mathrm{I}}(\mathbf{I}, \underline{\theta}, \mathbf{f}) + \mathbf{E}_{\mathrm{S}}(\mathbf{I}, \mathbf{f}) + \mathbf{E}_{\mathrm{P}}(\mathbf{f}, \mathbf{R}), \tag{3.21}$$

where $\mathbf{E}_{\mathrm{I}}(\mathbf{I}, \underline{\theta}, \mathbf{f})$, $\mathbf{E}_{\mathrm{S}}(\mathbf{I}, \mathbf{f})$ and $\mathbf{E}_{\mathrm{P}}(\mathbf{f}, \mathbf{R})$ are defined previously in section 3.2. The minimization of equation (3.21) occurs through the iterations of two steps: image segmentation based on graph cuts and GMM parameters calculation. The GMM parameters can still be updated using Eq. 3.14 and Eq. 3.15 whereas during segmentation, the first and third terms in equation (3.21) could still be applied as t-links and the second term n-links.

3.5 Results

First, we show the effectiveness of our proposed method in preserving fine brain structures with Fig. 3.2. Our proposed method does a much better job preserving



(a) (b) (c)

Figure 3.2: Segmentation results by standard graph cuts (b) and our proposed geo-cuts-based method (c) of a slice of real brain MRI data (a). (c) does a better job preserving fine structures in white matter and CSF.

fine structures of white matter and CSF than the standard graph cuts method in this case.

The percentages of voxels misclassified by our method and its sequential counterpart (sequential method) on synthetic images with different level of noise and intensity inhomogeneity are shown in Table 3.1. Table 3.2 and Table 3.3 show the Dice Similarity Measure for these two methods w.r.t. the ground truth on the same images. The results indicate that our proposed method shows improvement over the sequential method on images with all levels of noise and bias field for both overall performance and individual tissues. Note that the standard level of noise for real MR images is 3%. The Z-scores are ∞, 33.12 and ∞ according to the misclassifications, DSM for WM and DSM for GM, which make the joint method better at the 99% confidence level in all cases.

Fig. 3.3 and Fig. 3.4 demonstrate two sample segmentations by our proposed joint method and its sequential counterpart on the synthetic images. As shown in Fig. 3.3 and Fig. 3.4, the joint method performs better than the sequential method especially in areas where intensities of the voxels themselves cannot accurately

Table 3.1: Percentages of misclassified voxels by our method and its sequential counterpart on synthetic images with different level of noise (N) and bias field (BF). The results indicate that our proposed method shows improvement over the sequential method on image with all levels of noise and bias field.

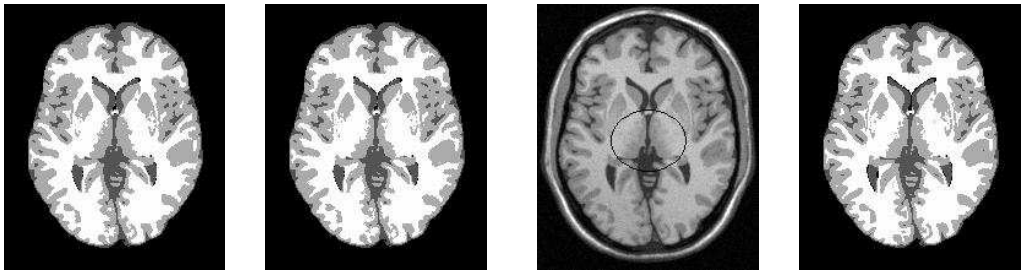
	Sequential Method	Joint Method
N=0,BF=0	5.08085%	3.66684%
N=1,BF=0	3.59626%	2.47611%
N=3,BF=0	4.32948%	3.34663%
N=5,BF=0	5.72533%	4.67929%
N=7,BF=0	7.55655%	5.70425%
N=0,BF=40	3.08586%	2.19323%
N=1,BF=40	3.21537%	2.35018%
N=3,BF=40	4.27003%	3.37058%
N=5,BF=40	6.43218%	5.11434%
N=7,BF=40	7.75894%	7.51967%

Table 3.2: Dice Similarity Measure (DSM) for WM w.r.t. the ground truth are shown for our proposed joint method and its sequential counterpart. The results show that our proposed method works better than the sequential method on for images with all levels of noise and bias field.

	DSM-WM	
	Sequential Method	Joint Method
N=0,BF=0	0.955394	0.977121
N=1,BF=0	0.967582	0.983641
N=3,BF=0	0.958922	0.972563
N=5,BF=0	0.948609	0.964309
N=7,BF=0	0.927565	0.953107
N=0,BF=40	0.973651	0.986247
N=1,BF=40	0.972185	0.984362
N=3,BF=40	0.960478	0.972953
N=5,BF=40	0.937137	0.955087
N=7,BF=40	0.942725	0.938707

Table 3.3: Dice Similarity Measure (DSM) of GM w.r.t. the ground truth are shown for our proposed joint method and its sequential counterpart. The results show that our proposed method works better than the sequential method for images with all levels of noise and bias field.

	DSM-GM	
	Sequential Method	Joint Method
N=0,BF=0	0.945581	0.960133
N=1,BF=0	0.961172	0.973404
N=3,BF=0	0.952958	0.964020
N=5,BF=0	0.938320	0.949318
N=7,BF=0	0.917758	0.938611
N=0,BF=40	0.966672	0.976551
N=1,BF=40	0.965269	0.974874
N=3,BF=40	0.953787	0.963913
N=5,BF=40	0.929783	0.94481
N=7,BF=40	0.917564	0.919515

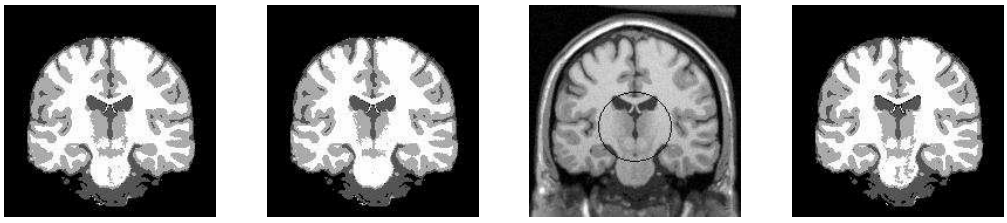


(a) (b) (c) (d)

Figure 3.3: Segmentation by our proposed joint method (Fig. 3.3(a)) and its sequential counterpart (Fig. 3.3(b)) on a slice of synthetic MR image (Fig. 3.3(c)) with 3% noise and 40% intensity inhomogeneity. The ground truth segmentation is shown in Fig. 3.3(d). The joint method shows improvement over the sequential method for areas where intensities alone cannot classify the voxels correctly as indicated by the circle in Fig. 3.3(c).

indicate to which tissue types the voxels belong and the classifications depend heavily on the atlas. With the joint method, the overall atlas alignment is improved by the correct segmentation of other areas, which in turn benefits the segmentation of the problematic areas.

One interesting aspect of the results by our proposed method one can see from



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

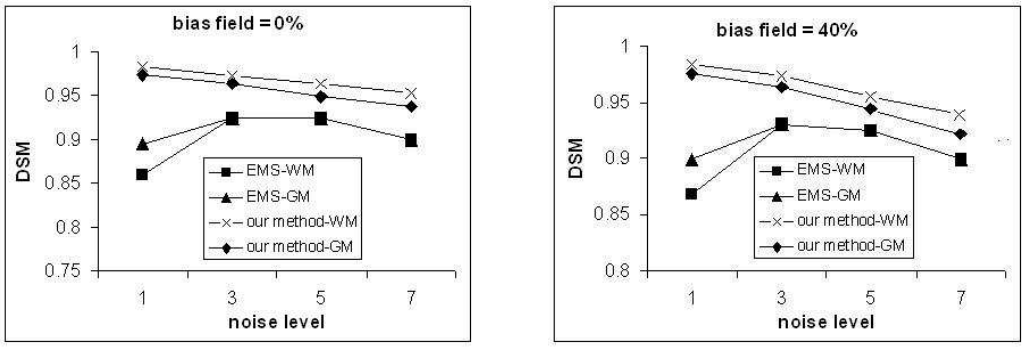
Figure 3.4: Segmentation by our proposed joint method (Fig. 3.4(a)) and its sequential counterpart (Fig. 3.4(b)) on a slice of synthetic MR image (Fig. 3.4(c)) with 3% noise and 40% intensity inhomogeneity. The ground truth segmentation is shown in Fig. 3.4(d). The joint method shows improvement over the sequential method for areas where intensities alone cannot classify the voxels correctly as indicated by the circle in Fig. 3.4(c).

Table 3.1 and Table 3.2, Table 3.3 is that the images with a moderate amount of noise are better segmented than images with no noise. At the same time, images with intensity inhomogeneity are better segmented than images with no bias field when the level of noise is low. For a closer examination, we look at table 3.4. Table 3.4 shows the percentages of white matter voxels classified as white matter or gray matter and the percentages of gray matter voxels classified as white matter or gray matter by our proposed method. As one can see from this table, a large percentage of white matter voxels are classified as gray matter for the image with no noise or bias field, whereas the segmentation of white matter is significantly improved when either noise or bias field is added. This effect can be explained by the difference in the GMM parameters. For brain MRI images where no noise or bias field is present, the voxel intensities for gray matter exhibit a much bigger variance than that for white matter. Therefore, partial voxels consisting of both gray matter and white matter have a large tendency to be classified as gray matter. When noise or bias field is present, the variance of white matter intensities also becomes larger, forcing the partial voxels to be classified more evenly.

Fig. 3.5 compares the performance of our proposed method against EMS for

Table 3.4: Percentages of WM voxels labeled as WM and GM and percentages of GM voxels labeled as WM and GM by our proposed method.

	WM voxel	
	labeled as WM	labeled as GM
N=0,BF=0	91.6743%	8.31934%
N=1,BF=0	96.8939%	3.10503%
N=0,BF=40	98.2076%	1.78666%
	GM voxel	
	labeled as WM	labeled as GM
N=0,BF=0	0.175322 %	95.2741%
N=1,BF=0	2.53092 %	96.1036%
N=0,BF=40	2.63104 %	96.5441%

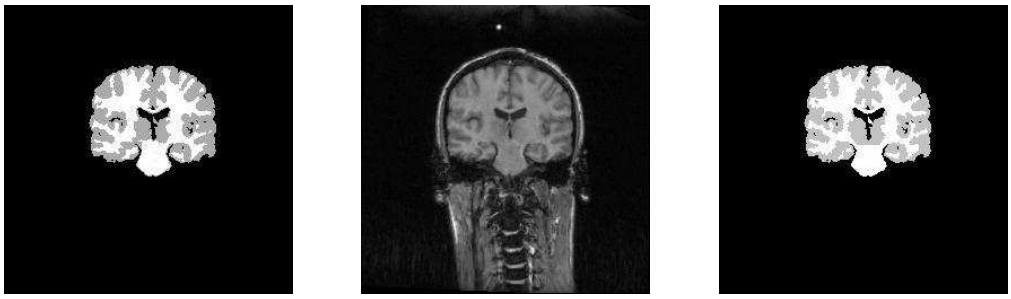


(a) (b)

Figure 3.5: DSM for our proposed method and EMS on synthetic images with 0% intensity non-uniformnities (Fig. 3.5(a)) and 40% intensity non-uniformnities (Fig. 3.5(b)) and various levels of noise.

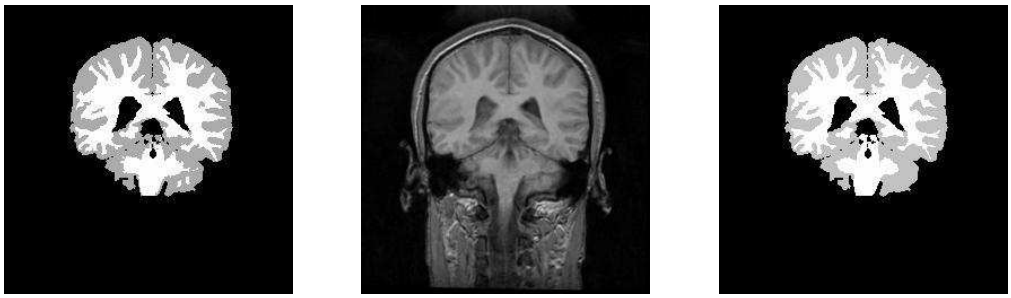
synthetic images with different levels of noise and bias field. As indicated by these graphs, our method performs significantly better than EMS. The Z-scores are ∞ and ∞ according to the DSM for WM and DSM for GM, which make our proposed method better at the 99% confidence level in both cases.

Fig. 3.6 and Fig. 3.7 show the segmentations by our method and the hand segmentations for the real images from IBSR. As shown by the images, although the real brain MRI data from IBSR have low contrast, large intensity inhomogeneity and contain imaging artifacts, our method is able to obtain segmentations similar



(a) (b) (c)

Figure 3.6: Segmentation of WM and GM tissues by our proposed method (Fig. 3.6(a)) on a real T1-weighted MRI image from ISBR (Fig. 3.6(b)) and its expert hand segmentation (Fig. 3.6(c)).



(a) (b) (c)

Figure 3.7: Segmentation of WM and GM tissues by our proposed method (Fig. 3.7(a)) on a real T1-weighted MRI image from ISBR (Fig. 3.7(b)) and its expert hand segmentation (Fig. 3.7(c)).

to the expert segmentations.

3.6 Discussions

In this paper, we proposed a novel approach that simultaneously performs graph-based brain MR image segmentation, non-rigid atlas registration, and intensity inhomogeneity correction. We define a Riemannian space in which the ideal segmentation boundaries, including those around fine brain structures, have the smallest Riemannian length. "Geo-cuts" is incorporated to find the geodesic contours in this space. Our segmentation accounts for spatial dependency and fine brain

structures with efficient calculations. We addressed the intensity inhomogeneites by dividing the brain image into small blocks and modeling the intensity probability distributions of brain tissues with local GMMs. To capture both the global and local effects, our non-rigid registration uses two sets of parameters. We model the global registration of the atlas with the brain image using an affine transformation and use a B-spline representation for the deformable transformation. To increase robustness and accuracy, we consider both the atlas/segmentation and the atlas/brain image fit during registration. By finding the segmentation and optimizing the registration and the GMM parameters through one objective function, we unified the three problems and allowed them to interact with one another. We validated our method on both synthetic and real T1 brain images. We compared the performance of our proposed method and its equivalent sequential method and showed the effectiveness of conducting brain image segmentation and atlas registration simultaneously.

Although we used local GMMs to compensate for intensity non-uniformity, many other methods have been proposed to address this issue in literature. It is certainly interesting to validate our approach against other methods and will be one of our future works. In our current work, we used a brain atlas constructed from normal brains. This atlas will not work for brains of Alzheimer's patients or brains with tumors. Disease specific atlases are required in these cases. Whether an atlas is useful for the segmentation of brains with tumors is also an interesting topic for future studies.

Chapter 4
Graph Cuts Segmentation with Statistical Shape Priors For Medical Images

4.1 Chapter Overview

As mentioned, medical images present many challenges for automated segmentation and traditional segmentation methods based solely on intensity information. In the previous chapter, we studied spatial atlas guided image segmentation. Spatial atlases are very powerful for brain images. Although, for the MR image segmentation of organs other than the brain, priors capturing shape variabilities and constraints are more applicable.

In this chapter, we propose a graph cuts-based segmentation method for medical images that incorporates statistical shape priors to increase robustness. Many shape representations for segmentation purposes have been proposed. such as point-based shape models [23], spherical harmonics [42], medial shape representations [63] and non-parametric shape models [24]. Here, we adopt the implicit shape representation with signed distance maps. The statistical shape information is obtained from a training set of segmented images. We solved the segmentation and shape parameters calculation simultaneously. In our proposed method, an initial inaccurate segmentation is prevented from producing a vicious cycle of inaccurate shape priors and segmentations. Our method is able to deal with complex shapes and shape variations while taking advantage of the globally efficient optimization by graph cuts. We demonstrate the effectiveness of our method on kidney images without strong boundaries.

In section 4.2, we first give a brief summary of our shape representation us-

ing signed distance maps. We then describe our novel objective function, which links the shape fitting and image segmentation problems. In section 4.3, we show how our energy function is optimized through an expectation maximization-style approach. We evaluate in section 4.4, our proposed algorithm and show the experimental results.

4.2 Segmentation Method

In this section, we first discuss our models for shape representation and intensity probability distribution. We end by presenting our novel objective function and its optimization. In the next section, we show how the energy function is optimized.

We represent images as a 1D array $\mathbf{I} = (I_1, \ldots, I_N)$, where I_n is the intensity of voxel n and N is the total number of voxels. We assign each voxel one of two labels: object (O) or background (B). Our segmentations are represented as label configuration of all voxels, $\mathbf{f} = (f_1, \ldots, f_N)$, where $f_n \in \{\mathrm{O}, \mathrm{B}\}$.

4.2.1 Shape Representation

We adopt the implicit representation and model shapes with signed distance functions, where shape boundary voxels have the value of zero and inside and outside voxels are assigned negative and positive distances respectively.

The training shapes are first aligned with one another through linear transformation. We then follow [48, 76] and reduce the redundancy from the statistical information stored in the training shapes using principle component analysis (PCA). PCA is a method for dimensionality reduction by finding the modes of variation with significance. It is applicable when random variables we observe at a higher dimension come from a smaller set of random variables at a lower dimension. A

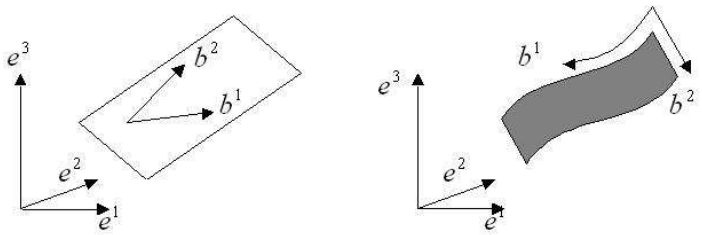


Figure 4.1: PCA is useful in these cases for dimensionality reduction. Although the data are observed in 3-D (e^1, e^2, e^3), they really come from 2-D (b^1, b^2).

simple example of when PCA can be used is shown in Fig. 4.1. In our case of dimensionality reduction for shape representations, we treat each element in the signed distance function as a random variable. Given the signed distance functions of C aligned training shapes $\{\mathbf{v}_1, ..., \mathbf{v}_C\}$ in 1-D array forms, we first calculate the mean shape:

$$\overline{\mathbf{U}} = \frac{1}{C} \sum_c \mathbf{v}_c. \tag{4.1}$$

To calcuate the shape variabilities using PCA, we need to find the covariance matrix for elements in the signed distance function. This is calculated by first subtracting the mean shape $\overline{\mathbf{U}}$ from the signed distance functions creating the mean-offset functions $\{\hat{\mathbf{v}}_1, ..., \hat{\mathbf{v}}_C\}$. A $N \times C$ shape variability matrix $\mathbf{V}$, whose c-th column is $\hat{\mathbf{v}_c}$, is then constructed. The covariance matrix is $\frac{1}{C}\mathbf{V}\mathbf{V}^T$. Performing singular value decomposition (SVD) on this covariance matrix, we get:

$$\frac{1}{C}\mathbf{V}\mathbf{V}^T = \mathbf{W}\mathbf{S}\mathbf{W}^T, \tag{4.2}$$

where the columns of $\mathbf{W}$ give us the eigenshapes $\mathbf{U}^k$ and the diagonal elements of $\mathbf{S}$ their corresponding eigenvalues. Eigenshapes with larger eigenvalues are more important and represent the major shape variabilities. New shapes $\mathbf{U}$ can then be represented using the mean shape and the K eigenshapes with the K largest eigenvalues:

$$\mathbf{U} = \overline{\mathbf{U}} + \sum_{k=1}^{K} w_k \mathbf{U}^k, \tag{4.3}$$

where $\mathbf{w} = \{w_1, ..., w_k, ..., w_K\}$ are the weights for the eigenshapes and each $\mathbf{w}$ represents a different shape.

A rigid transformation $\mathbf{T} = \{T_0, ..., T_8\}$ describing rotation, translation and scaling is necessary to align the shape information with the image to be segmented. $\mathbf{T}$ maps a point (x, y, z) to point $(\tilde{x}, \tilde{y}, \tilde{z})$ according to

$$\begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \\ 1 \end{pmatrix} = T(T_0, T_1, T_2) \cdot S(T_3, T_4, T_5) \cdot R_x(T_6) \cdot R_y(T_7) \cdot R_z(T_8) \cdot \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}, \tag{4.4}$$

where

$$T(T_0, T_1, T_2) = \begin{pmatrix} 1 & 0 & 0 & T_0 \\ 0 & 1 & 0 & T_1 \\ 0 & 0 & 1 & T_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.5}$$

$$S(T_3, T_4, T_5) = \begin{pmatrix} T_3 & 0 & 0 & 0 \\ 0 & T_4 & 0 & 0 \\ 0 & 0 & T_5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.6}$$

$$R(T_6) = \begin{pmatrix} \cos(T_6) & -\sin(T_6) & 0 & 0 \\ \sin(T_6) & \cos(T_6) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.7}$$

$$R(T_7) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(T_7) & -\sin(T_7) & 0 \\ 0 & \sin(T_7) & \cos(T_7) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{4.8}$$

$$R(T_8) = \begin{pmatrix} \cos(T_8) & 0 & \sin(T_8) & 0 \\ 0 & 1 & 0 & 0 \\ -\sin(T_8) & 0 & \cos(T_8) & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{4.9}$$

Here, $T(T_0, T_1, T_2)$ represents transformation, $S(T_3, T_4, T_5)$ scaling, and $R(T_5)$, $R(T_6)$ and $R(T_7)$ rotation. We denote the shape prior, the mean and eigenshapes after the transformation as $\mathbf{U}^\star$, $\overline{\mathbf{U}}^\star$ and $\mathbf{U}^{k\star}$.

4.2.2 Mixture of Gaussians

We model the intensity probability distribution of the image with a Gaussian mixture model (GMM) using one Gaussian for the object and D Gaussians for the background. Let $\underline{\theta} = \{\mu_d, \Sigma_d, \pi_d, d \in 0, ..., D\}$ denote the parameters for the GMM, where μ_d, Σ_d, π_d are respectively the mean, variance, and prior probability of Gaussian d with d=0 being for the object (O) and d=1,...,D for the background (B). We define a parameter $\mathbf{b} = \left(\begin{pmatrix} b_1^{\mathrm{O}} \\ b_1^{\mathrm{B}} \end{pmatrix}, \ldots, \begin{pmatrix} b_n^{\mathrm{O}} \\ b_n^{\mathrm{B}} \end{pmatrix}, \ldots, \begin{pmatrix} b_N^{\mathrm{O}} \\ b_N^{\mathrm{B}} \end{pmatrix} \right)$ (where $b_n^{\mathrm{O}} \in \{0\}$ and $b_n^{\mathrm{B}} \in \{1, ..., D\}$) that assigns each voxel to one Gaussian belonging to the object and one to the background.

4.2.3 Objective Function

Our problem can be formulated as: given a medical image $\mathbf{I} = (I_1, \ldots, I_N)$ with N voxels and shape information of the target object, we wish to

(1) identify the parameters $(\mathbf{w}, \mathbf{T})$ of the shape that best matches the object in the image and

(2) segment the image into object and background.

The segmentation and shape fitting problems are inter-related and we solve them in an iterative manner in that we use the shape prior to aid the segmentation and use the segmentation to calculate the shape parameters.

We define the energy functional in equation (4.10) to guide the image segmentation and shape prior calculation.

$$\begin{aligned} E(\mathbf{I}, \underline{\theta}, \mathbf{w}, \mathbf{T}, \mathbf{f}, \mathbf{b}) = E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b}) + E_{\mathrm{N}}(\mathbf{I}, \mathbf{f}) \\ + E_{\mathrm{P}}(\mathbf{f}, \mathbf{w}, \mathbf{T}) + E_{\mathrm{S}}(\mathbf{w}, \mathbf{T}, \mathbf{I}). \end{aligned} \tag{4.10}$$

The first term measures how well the voxel labels and the GMM parameters fit the image given its intensities and can be written as

$$\begin{aligned} E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b}) = & -\sum_{n} E_{\mathrm{D}}(I_n, \underline{\theta}, f_n, b_n) \\ = & -\sum_{n} \log \mathrm{P}(I_n, b_n | \underline{\theta}), \end{aligned} \tag{4.11}$$

where

$$b_n = \begin{cases} b_n^{\mathrm{O}} & (f_n = \mathrm{O}) \\ b_n^{\mathrm{B}} & (f_n = \mathrm{B}) \end{cases}. \tag{4.12}$$

The second term measures the smoothness of the label configuration and follows the standard graph cuts formulation:

$$E_N(\mathbf{I}, \mathbf{f}) = \sum_{f_n \neq f_q, q \in N_n} \frac{1}{1 + (I_n - I_q)^2}, \tag{4.13}$$

where $, q \in N_n$ denotes that voxel n and q are neighbors.

The third term denotes the fitness between the current shape prior and the current segmentation and can be written as

$$E_{\mathrm{P}}(\mathbf{f}, \mathbf{w}, \mathbf{T}) = \sum_{n} (M_n - U_n^{\star})^2, \tag{4.14}$$

where $M_n = const$ when $f_n = \mathrm{B}$ and $M_n = -const$ when $f_n = \mathrm{O}$ and $const$ is a positive constant. A penalty is applied for voxels classified differently by the segmentation and the shape prior through equation (4.14) and by adding this term, we allow the segmentation and the shape prior to interact with each other.

The last term measures how well the shape prior itself fits with the image through calculating the entropy of intensity distributions inside and outside the shape prior. This term can be written as

$$E_{\mathrm{S}}(\mathbf{w}, \mathbf{T}, \mathbf{I}) = -\sum_{i} (p_{\mathrm{S}}(i) \log p_{\mathrm{S}}(i) + p_{\mathrm{NS}}(i) \log p_{\mathrm{NS}}(i)), \tag{4.15}$$

where $p_{\mathrm{S}}(i)$ and $p_{\mathrm{NS}}(i)$ are respectively the probability for voxels inside and outside the shape prior to have intensity i. $p_{\mathrm{S}}(i)$ and $p_{\mathrm{NS}}(i)$ are calculated as

$$p_{\mathrm{S}}(i) = \frac{N_{\mathrm{S}}(i)}{A_{\mathrm{S}}}, \qquad p_{\mathrm{NS}}(i) = \frac{N_{\mathrm{NS}}(i)}{A_{\mathrm{NS}}}, \tag{4.16}$$

where $N_{\mathrm{S}}(i)$ and A_{S} are the number of voxels with intensity i and the area inside shape prior, and $N_{\mathrm{NS}}(i)$ and A_{NS} are those for outside the shape prior. As in the previous chapter, we denote a heavyside step function with $H(\bullet)$ and a delta function with $\delta(\bullet)$, and

$$\begin{aligned} A_{\mathrm{S}} &= \sum_{n} H(-U_n^{\star}), \\ A_{\mathrm{NS}} &= \sum_{n} H(U_n^{\star}), \\ N_{\mathrm{S}}(i) &= \sum_{n} H(-U_n^{\star}) \delta(I_n - i), \\ N_{\mathrm{NS}}(i) &= \sum_{n} H(U_n^{\star}) \delta(I_n - i). \end{aligned} \tag{4.17}$$

By adding the last term in equation (4.10), we prevent an initial inaccurate segmentation from producing inaccurate shape priors.

4.3 Energy Minimization

As mentioned above, we use an expectation maximization-style approach to minimize the energy function presented in equation (4.10) and alternately update the GMM and shape parameters while fixing the segmentation (maximization step) and use the GMMs and the shape prior to facilitate the image segmentation (estimation step).

4.3.1 M(aximization) Step

We start by assigning two Gaussians to each voxel, one which minimizes the first term in equation (4.10) when the voxel belongs to the object and one when the voxel belongs to the background.

Differentiating equation (4.11) w.r.t. μ_d, we get:

$$\frac{\partial E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b})}{\partial \mu_d} = -\frac{\partial}{\partial \mu_d} \sum_{b_n = d} \log \mathrm{P}(I_n | \mu_d, \Sigma_d) = 0, \tag{4.18}$$

giving us the update equation

$$\mu^d = \frac{\sum_{b_n=d} I_n}{\sum_{b_n=d} 1}. \tag{4.19}$$

Similarly, the updated equations for Σ_d and π_d are respectively:

$$\begin{aligned} \Sigma_d &= \frac{\sum_{b_n=d} (I_n - \mu_d)^2}{\sum_{b_n=d} 1}, \\ \pi_d &= \begin{cases} \frac{\sum_{b_n=d} 1}{\sum_{b_n=0} 1} & (f_n = \mathrm{O}) \\ \frac{\sum_{b_n=d} 1}{\sum_{b_n=1,\ldots,D} 1} & (f_n = \mathrm{B}) \end{cases}. \end{aligned} \tag{4.20}$$

We use a gradient descent optimization to update the shape parameters. The update equation for $\mathbf{w}$ is

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha(\nabla_{\mathbf{w}} E_{\mathrm{P}} + \nabla_{\mathbf{w}} E_{\mathrm{S}}), \tag{4.21}$$

where α is the step size for the gradient descent optimization and $\mathbf{w}_t$ and $\mathbf{w}_{t+1}$ are the weights at time t and $t+1$ respectively. $\nabla_{\mathbf{w}} E_{\mathrm{P}}$ is calculated by differentiating (4.14):

$$\frac{\partial E_{\mathrm{P}}}{\partial w_k} = \sum_n 2(U_n^\star - M_n) U_n^{k\star}. \tag{4.22}$$

$\nabla_{\mathbf{w}} E_{\mathrm{S}}$ is calculated by differentiating (4.15):

$$\begin{aligned} \frac{\partial E_{\mathrm{S}}}{\partial w_k} = & -\sum_i [\frac{1}{A_S} Q_S - \frac{1}{A_{NS}} Q_{NS}] \frac{\partial N_S(i)}{\partial w_k} \\ & + \sum_i [\frac{N_S(i)}{A_S^2} Q_S - \frac{N_{NS}(i)}{A_{NS}^2} Q_{NS}] \frac{\partial A_S(i)}{\partial w_k}, \end{aligned} \tag{4.23}$$

where

$$\begin{aligned} Q_S = & \ (\log \frac{N_S(i)}{A_S} + 1), \\ Q_{NS} = & \ (\log \frac{N_{NS}(i)}{A_{NS}} + 1), \end{aligned} \tag{4.24}$$

and

$$\begin{aligned} \frac{\partial A_S}{\partial w_k} = & \ -\sum_n \delta(-U_n^\star) U_n^{k\star}, \\ \frac{\partial N_S(i)}{\partial w_k} = & \ -\sum_n \delta(-U_n^\star) U_n^{k\star} \delta(I_n - i). \end{aligned} \tag{4.25}$$

The alignment parameters $\mathbf{T}$ can be updated in a similar manner:

$$\mathbf{T}_{t+1} = \mathbf{T}_t - \beta(\nabla_{\mathbf{T}} E_{\mathrm{P}} + \nabla_{\mathbf{T}} E_{\mathrm{S}}), \tag{4.26}$$

where β is the step size and $\mathbf{T}_t$ and $\mathbf{T}_{t+1}$ are the alignment parameters at time t and $t+1$ respectively and

$$\frac{\partial E_{\mathrm{P}}}{\partial T_k} = \sum_n 2(U_n^\star - M_n) \frac{\partial U_n{}^\star}{\partial T_k}, \tag{4.27}$$

and

$$\begin{aligned} \frac{\partial E_{\mathrm{S}}}{\partial T_k} = & -\sum_i [\frac{1}{A_S} Q_S - \frac{1}{A_{NS}} Q_{NS}] \frac{\partial N_S(i)}{\partial T_k} \\ & + \sum_i [\frac{N_S(i)}{A_S^2} Q_S - \frac{N_{NS}(i)}{A_{NS}^2} Q_{NS}] \frac{\partial A_S(i)}{\partial T_k}. \end{aligned} \tag{4.28}$$

Here,

$$\begin{aligned}\frac{\partial A_S}{\partial T_k} &= -\sum_n \delta(-U_n^\star)\frac{\partial U_n^\star}{\partial T_k},\\ \frac{\partial N_S(i)}{\partial T_k} &= -\sum_n \delta(-U_n^\star)\frac{\partial U_n^\star}{\partial T_k}\delta(I_n - i),\end{aligned} \tag{4.29}$$

and

$$\frac{\partial {U_n}^\star}{\partial T_k} = \frac{\partial U_n^\star}{\partial \tilde{x}}\cdot\frac{\partial \tilde{x}}{\partial T_k} + \frac{\partial U_n^\star}{\partial \tilde{y}}\cdot\frac{\partial \tilde{y}}{\partial T_k} + \frac{\partial U_n^\star}{\partial \tilde{z}}\cdot\frac{\partial \tilde{z}}{\partial T_k}, \tag{4.30}$$

where $(\frac{\partial U_n^\star}{\partial \tilde{x}}, \frac{\partial U_n^\star}{\partial \tilde{y}}, \frac{\partial U_n^\star}{\partial \tilde{z}})$ is the image gradient for $U_n^\star$ and $(\frac{\partial \tilde{x}}{\partial T_k}, \frac{\partial \tilde{y}}{\partial T_k}, \frac{\partial \tilde{z}}{\partial T_k})$ can be easily calculated by differentiating equation (4.4).

4.3.2 E(stimation) Step

We follow the graph cuts framework for segmentation and create a graph with nodes corresponding to voxels and two additional terminal nodes representing the object and the background. The first and third terms in our energy function are applied to the graph as t-links. The second term is added as n-links between neighboring nodes. We use the max-flow [13] algorithm to find the minimum cut on this graph.

4.4 Experimental Results

We tested our algorithm on MR kidney images of patients with ADPKD. These images are hard to segment using intensity information alone because the kidney tissues and neighboring psoas muscle tissues have similar intensities, resulting in poor contrast at desired tissue boundaries. Kidneys of ADPKD patients also include a large amount of liquid, which have very different intensity levels than normal kidney tissues, creating strong contours within the kidney. As shown in Fig. 4.2, the segmentation produced by the graph cuts method without any pri-

Summary of Algorithm

Initialization

The GMM parameters are initialized using pre-defined values. We use the average shape as the initial shape prior. The initial transformation of the shape prior is set to zero.

Iterative Minimization

REPEAT

E(stimation) Step

Find segmentation using graph cuts:

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} E(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{w}, \mathbf{b}).$$

M(inimization) Step

Assign two Gaussians to each voxel:

$$\hat{b}_n^{\mathrm{O}} = \operatorname*{argmin}_{b_n^{\mathrm{O}}} E_{\mathrm{D}}(I_n, \underline{\theta}, \mathrm{O}, b_n),\ \hat{b}_n^{\mathrm{B}} = \operatorname*{argmin}_{b_n^{\mathrm{B}}} E_{\mathrm{D}}(I_n, \underline{\theta}, \mathrm{B}, b_n).$$

Learn Gaussian mixture model parameters:

$$\mu_d = \frac{\sum_{b_n=d} I_n}{\sum_{b_n=d} 1}, \Sigma_d = \frac{\sum_{b_n=d}(I_n-\mu_d)^2}{\sum_{b_n=d} 1},\ \pi_d = \begin{cases} \frac{\sum_{b_n=d} 1}{\sum_{b_n=0} 1} & (f_n = \mathrm{O}) \\ \frac{\sum_{b_n=d} 1}{\sum_{b_n=1,\ldots,D} 1} & (f_n = \mathrm{B}) \end{cases}.$$

Update shape parameters through gradient descent until Convergence:

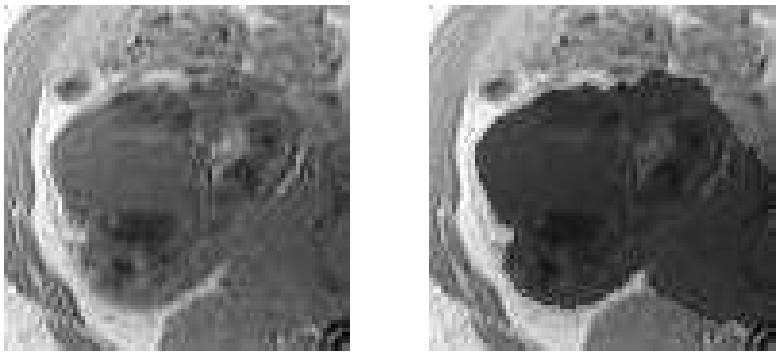
$$\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha(\nabla_{\mathbf{w}} E_{\mathrm{P}} + \nabla_{\mathbf{w}} E_{\mathrm{S}}),\ \mathbf{T}_{t+1} = \mathbf{T}_t - \alpha(\nabla_{\mathbf{T}} E_{\mathrm{P}} + \nabla_{\mathbf{T}} E_{\mathrm{S}}).$$

UNTIL GMM/SHAPE PARAMETERS CONVERGE

ors "leaks" into tissues that do not belong to, but have similar intensities as, the kidney.

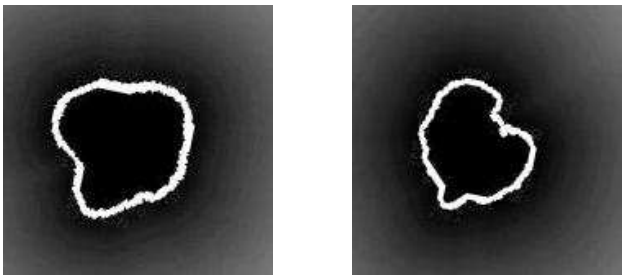
Our training set contains 30 segmented 2-D images from three different patients. Fig. 4.3 shows examples of training kidney shapes and signed distance functions associated with them. Fig. 4.4 shows the mean shape and the first three primary modes of variance obtained through the training set.

Fig. 4.5 shows the shape prior and final segmentation obtained by our method



(a) (b)

Figure 4.2: On the right is the segmentation produced by the graph cuts method without any priors for the image on the left. The voxels classified as kidney are shown in green. Without any priors, the segmentation "leaks" into tissues that do not belong to, but have similar intensities as, the kidney.



(a) (b)

Figure 4.3: Examples of the training shapes and their signed distance maps.

for a patient whose kidney images were used in the training set (although these images themselves were not a part of the training data). Fig. 4.6 shows the shape prior and final segmentation obtained by our method for patients whose kidney images were not used in the training set. As visible from Fig. 4.5 and Fig. 4.6, our method overcomes the "leakage" problem encountered by the graph cuts method without priors and produce accurate segmentations. The shape prior in Fig. 4.6 is more accurate than the ones in Fig. 4.5, because the shape of the kidney in Fig. 4.5 is closer to those used for training the eigenshapes. Even though the shape priors in Fig. 4.6 are not very accurate, our method is able to produce good final segmentations showing that our results are not restricted by the shape variabilities

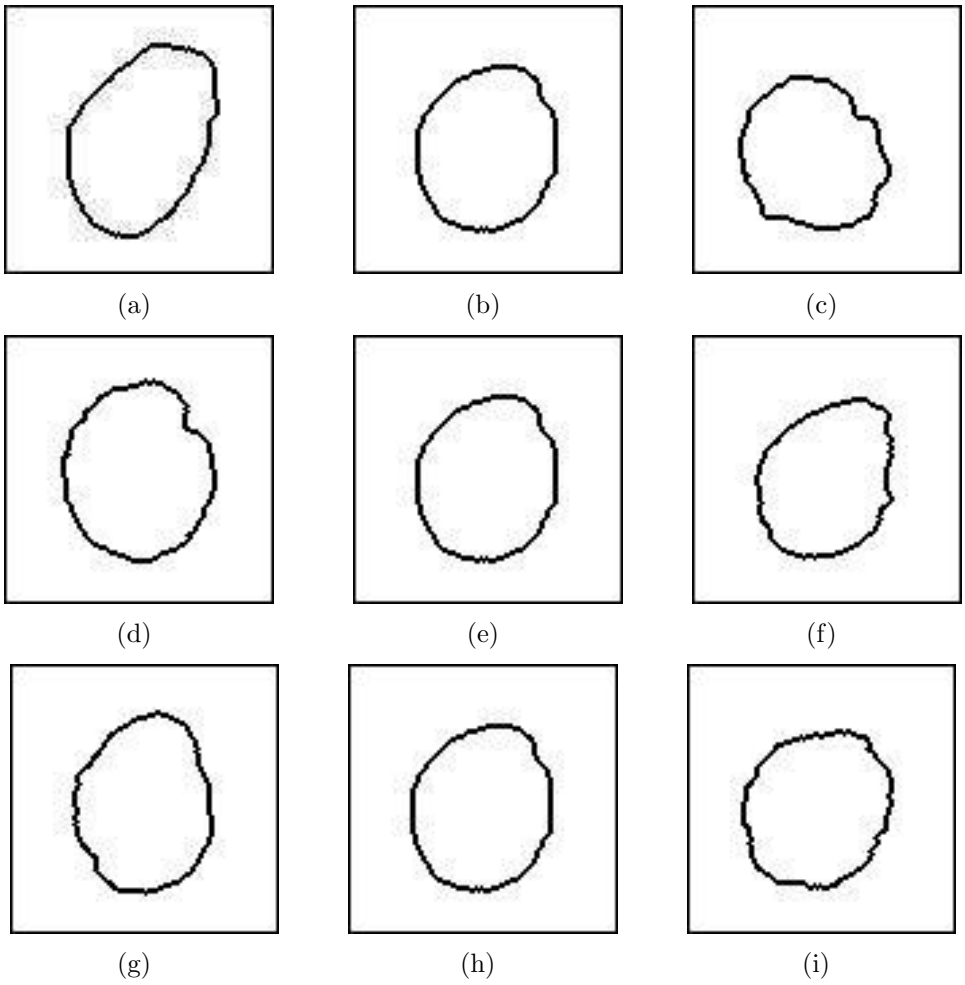
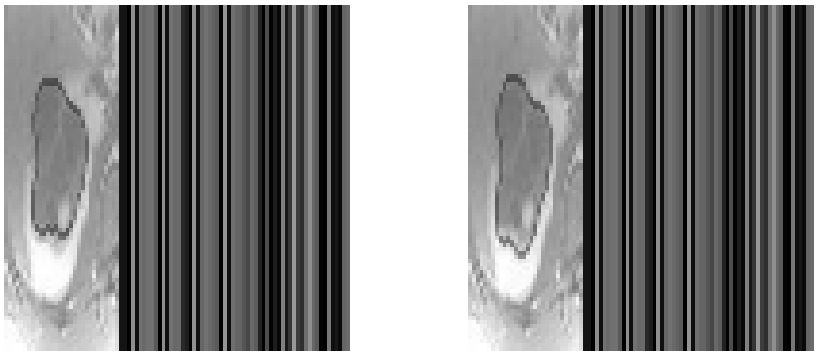


Figure 4.4: The mean shape and shape variabilities obtained from the training dataset. The middle column shows the mean shape. The left column shows $+1\sigma$ variation and the right column -1σ variation. The first, second and third row represents the first, second and third principle modes respectively. σ^2 is the eigenvalue associated with each principle mode.

represented in the training dataset.

4.5 Discussions

In this chapter, we proposed a graph cuts-based method for medical image segmentation that incorporates statistical shape priors to increase segmentation robustness in cases of poor contrast. We adopted the implicit shape representation pro-



(a) (b)

Figure 4.5: Shape prior (a) and segmentation (b) produced by our proposed method for a kidney image from a patient whose kidney images were used in the training set (although these images themselves were not a part of the training data). Our method was able to overcome the "leakage" problem encountered by the graph cuts method without priors as shown in Fig. 4.2.

posed in [58]. The statistical shape information was obtained from a training set of segmented images and extracted through PCA. We took a unified approach and solved the image segmentation and shape fitting problems through one objective function. Two novel terms accounting for shape/image fit and shape/segmentation fit were introduced. The former prevents an initial inaccurate segmentation from producing a vicious cycle of inaccurate shape priors and segmentations while the latter allows the segmentation and shape fitting problems to interact with one another. Our objective function is minimized through an EM-style approach where the segmentation and shape prior were calculated iteratively. The proposed method is able to deal with complex shapes and shape variations while taking advantage of the globally efficient optimization by graph cuts.

We demonstrated the effectiveness of our method on kidney images that cannot be segmented correctly using intensity information alone due to weak boundaries. The results show that our proposed method is able to handle images that cannot be correctly segmented without using any priors. Our segmentations are also not restricted by the shape variabilities represented by the training shapes. In the

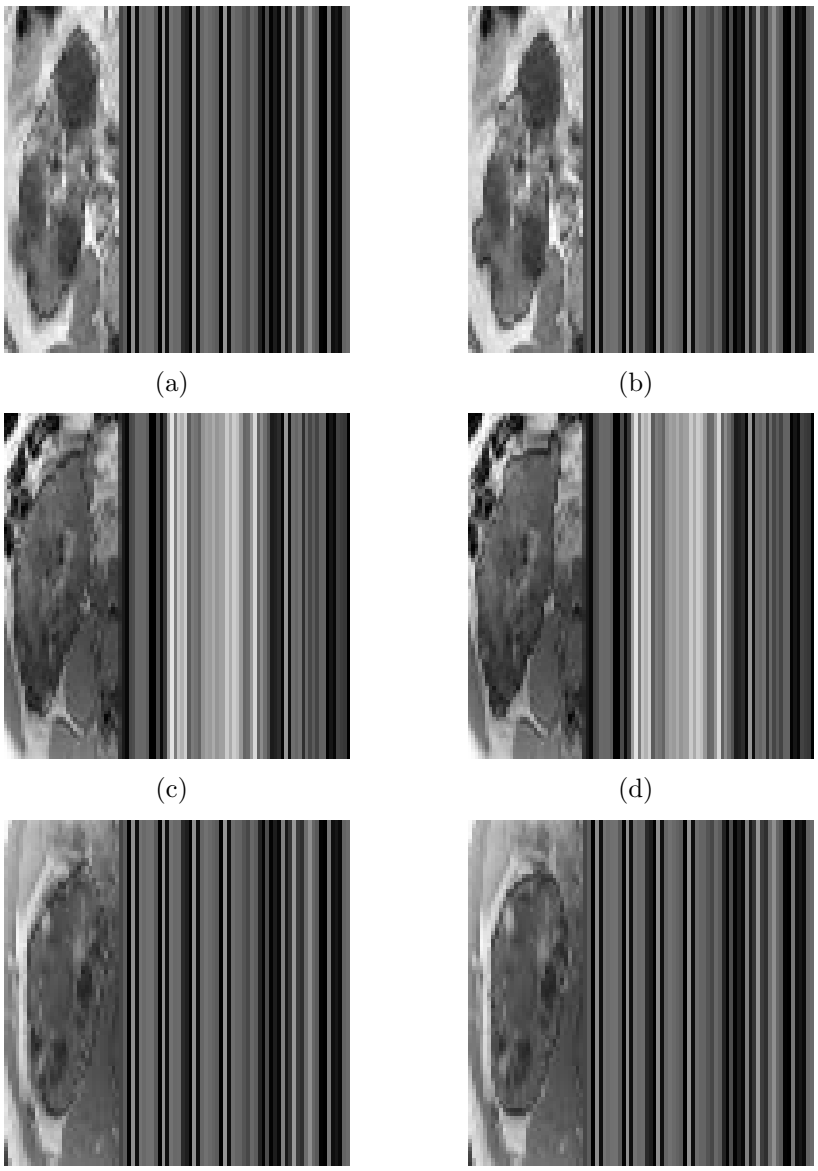


Figure 4.6: Shape priors (left column) and segmentations (right column) produced by our proposed method for kidney images from patients whose kidney images were not used in the training set. Our method was able to overcome the "leakage" problem encountered by the graph cuts method without priors as shown in Fig. 4.2.

future, we need to focus on quantitative verification of results by our proposed method.

Chapter 5
Graph Cuts Segmentation with Geometric Shape Priors for Medical Images

5.1 Chapter Overview

In the previous section, we studied statistical shape priors with graph cuts for MR image segmentation. Statistical shape priors require manual training set, which is not always available. Some anatomical structures resemble basic geometric shapes, which make geometric priors that do not require statistical training attractive. In this chapter, we propose a novel segmentation method that incorporates geometric shape priors with the graph cuts technique for medical image segmentation. We demonstrate the effectiveness of our method on cardiac images and kidney images without strong boundaries using respectively concentric circles and ellipses as shape priors. Since statistical training is not involved, the method is very efficient.

One of the major differences between our method and [72], which proposed elliptical priors for graph cuts segmentation, is that the shape in [72] is estimated using the segmentation from the previous step alone whereas we estimate the shape priors using both the segmentation from the previous step and information from the image. Therefore, an initial bad segmentation may produce a vicious cycle of inaccurate shape priors and segmentations for the method in [72], while it would not for our proposed method.

In section 5.2, we first give a brief summary of the graph cuts image segmentation framework. We then discuss our representation of intensity probability distribution and model for shape priors. We end by presenting our novel objective function and its optimization. We demonstrate our method on left ventricle

segmentation using concentric circles as shape priors in section 5.4 and on kidney image segmentation using elliptical shape priors in 5.3.

5.2 Segmentation Method

5.2.1 Mixture of Gaussians

We model the intensity probability distribution of the image with a Gaussian mixture model (GMM). Let us say we wish to segment the image into M regions. We denote G_m as the number of Gaussians we use to model the intensity distribution in the m-th region. Let $\underline{\theta} = \{\mu_d^m, \Sigma_d^m, \pi_d^m, m \in 1, ..., G_m\}$ denote the parameters for the GMM associated with the m-th region, where $\mu_d^m, \Sigma_d^m, \pi_d^m$ are respectively the mean, variance, and prior probability of the d-th Gaussian in the m-th region. The image is represented in an array form: $\mathbf{I} = (I_1, ..., I_N)$. We define a parameter $\mathbf{b} = \left(\begin{pmatrix} b_1^1 \\ \vdots \\ b_1^m \\ \vdots \\ b_1^M \end{pmatrix}, \ldots, \begin{pmatrix} b_n^1 \\ \vdots \\ b_n^m \\ \vdots \\ b_n^M \end{pmatrix}, \ldots, \begin{pmatrix} b_N^1 \\ \vdots \\ b_N^m \\ \vdots \\ b_N^M \end{pmatrix} \right)$ (where $b_n^m \in \{1, ..., G_m\}$) that assigns all N voxels in the image to one Gaussian belonging to each of the M regions.

5.2.2 Shape Representation

We model shapes using basic geometric shapes. The type and number of shapes are determined by the structure we wish to segment. We denote the shape parameters as $\mathbf{w} = \{w_0, ..., w_K - 1\}$, where K is the number of parameters necessary to express our shape prior. The shape prior divides the image into M regions. fs_n denotes

the region the shape prior assigns voxel n to. We show the specific shape priors for left ventricle segmentation in section 5.4 and those for kidney segmentation in section 5.3.

5.2.3 Objective Function

Our problem can be formulated as: given a medical image and basic shape information of the desired biological structure, we wish to

(1) identify the shape parameters $\mathbf{w}$ that best match the image and

(2) segment the image into M regions.

Our segmentations are represented as label configurations $\mathbf{f} = (f_1, \ldots, f_N)$, where each voxel n is assigned a label $f_n \in \{0, ..., M-1\}$. The segmentation and shape fitting problems are inter-related and we solve them in an iterative manner in that we use the shape prior to aid the segmentation and use the segmentation to calculate the shape parameters.

We define the energy functional in equation (5.1) to guide the image segmentation and shape parameters calculation.

$$\begin{aligned} E(\mathbf{I}, \underline{\theta}, \mathbf{w}, \mathbf{f}, \mathbf{b}) = E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b}) + E_{\mathrm{N}}(\mathbf{I}, \mathbf{f}) \\ + E_{\mathrm{P}}(\mathbf{f}, \mathbf{w}) + E_{\mathrm{S}}(\mathbf{w}, \mathbf{I}). \end{aligned} \tag{5.1}$$

The first term measures how well the voxel labels and the GMM parameters fit the image given its intensities and can be written as

$$\begin{aligned} E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b}) &= -\sum_n E_{\mathrm{D}}(I_n, \underline{\theta}, f_n, b_n) \\ &= -\sum_n \log \mathrm{P}(I_n, b_n | \underline{\theta}), \end{aligned} \tag{5.2}$$

where

$$b_n = b_n^m \qquad (m = f_n). \tag{5.3}$$

The second term measures the smoothness of the label configuration and follows the standard graph cuts formulation:

$$E_N(\mathbf{I}, \mathbf{f}) = \sum_{f_n \neq f_q, n,q \in C} \frac{1}{1 + (I_n - I_q)^2}, \tag{5.4}$$

where $n, q \in C$ denotes that voxel n and q are neighbors.

The third term denotes the fitness between the current shape prior and the current segmentation and can be written as

$$E_{\mathrm{P}}(\mathbf{f}, \mathbf{w}) = \sum_n \sum_m [(CS^m - U_n^m)^2], \tag{5.5}$$

where $CS^m = c_m$ when $f_n = m$ and $CS^m = -c_m$ when $f_n \neq m$, and c is a positive constant. Here, U_n^m has a positive value when $fs_n = m$ and a negative value when $fs_n \neq m$. The value of U_n^m is determined by the distance to the region boundaries. A penalty is applied for voxels classified differently by the segmentation and the shape prior through equation (5.5). By adding this term, we allow the segmentation and the shape prior to interact with each other.

The last term measures how well the shape prior itself fits the image through calculating the entropy of intensity distributions inside each of the M regions divided by the shape prior. This term can be written as

$$E_{\mathrm{S}}(\mathbf{w}, \mathbf{I}) = -\sum_m \sum_i (p_m(i) \log p_m(i)), \tag{5.6}$$

where $p_m(i)$ is the probability for voxels where $fs_n = m$ to have intensity i and is calculated as

$$p_m(i) = \frac{N_m(i)}{A_m}, \tag{5.7}$$

where $N_m(i)$ and A_m are respectively the number of voxels with intensity i and the total number of voxels where $fs_n = m$. By adding the last term in equation (5.1), we prevent inaccurate segmentations from producing a loop of inaccurate shape priors and segmentations.

5.2.4 Energy Minimization

As mentioned previously, we use an EM-style approach to minimize the energy function presented in equation (5.1) and alternately update the GMM and shape parameters while fixing the segmentation (maximization step) and use the GMMs and the shape prior to facilitate the image segmentation (estimation step).

M(aximization) Step

We start by assigning M Gaussians to each voxel, one which minimizes the first term in equation (5.1) when the voxel belongs to each of the M regions.

Differentiating equation (5.2) w.r.t. μ_d^m, we get:

$$\frac{\partial E_{\mathrm{D}}(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{b})}{\partial \mu_d^m} = -\frac{\partial}{\partial \mu_d^m} \sum_{f_n=m, b_n^m=d} \log \mathrm{P}(I_n | \mu_d^m, \Sigma_d^m) = 0, \tag{5.8}$$

giving us the update equation

$$\mu_d^m = \frac{\sum_{f_n=m, b_n^m=d} I_n}{\sum_{f_n=m, b_n^m=d} 1}. \tag{5.9}$$

Similarly, the updated equations for Σ_d^m and π_d^m are respectively:

$$\Sigma_d^m = \frac{\sum_{f_n=m, b_n^m=d} (I_n - \mu_d^m)^2}{\sum_{f_n=m, b_n^m=d} 1}, \pi_d^m = \frac{\sum_{f_n=m, b_n^m=d} 1}{\sum_{f_n=m} 1}. \tag{5.10}$$

We use a gradient descent optimization to update the shape parameters. The update equation for $\mathbf{w}$ is

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha(\nabla_{\mathbf{w}} E_{\mathrm{P}} + \nabla_{\mathbf{w}} E_{\mathrm{S}}), \tag{5.11}$$

where α is the step size for the gradient descent optimization and $\mathbf{w}_t$ and $\mathbf{w}_{t+1}$ are the shape parameters at time t and $t+1$ respectively. $\nabla_{\mathbf{w}} E_{\mathrm{P}}$ is calculated by differentiating (5.5):

$$\frac{\partial E_{\mathrm{P}}}{\partial w_k} = \sum_n \sum_m [2(U_n^m - CS_n^m) \frac{\partial U_n^m}{\partial w_k}]. \tag{5.12}$$

$\nabla_{\mathbf{w}} E_{\mathrm{S}}$ is calculated by differentiating (5.6):

$$\frac{\partial E_{\mathrm{S}}}{\partial w_k} = -\sum_m \sum_i [\frac{1}{A_m} Q_m \frac{\partial N_m(i)}{\partial w_k} - \frac{N_m(i)}{A_m^2} Q_m \frac{\partial A_m(i)}{\partial w_k}], \tag{5.13}$$

where

$$Q_m = (\log \frac{N_m(i)}{A_m} + 1). \tag{5.14}$$

$\nabla_{\mathbf{w}} \mathbf{U}$, $\frac{\partial A_m}{\partial w_k}$ and $\frac{\partial N_m}{\partial w_k}$ are necessary to update $\mathbf{w}$ and we show their calculations in later chapters.

E(stimation) Step

We follow the graph cuts framework for segmentation and create a graph with nodes corresponding to voxels and two additional terminal nodes. The first and third terms in our energy function are applied to the graph as t-links. The second term is added as n-links between neighboring nodes. We use the max-flow [13] algorithm to find the minimum cut on this graph. When there are more than two labels, we use the α expansion.

Summary of Algorithm

Initialization

The GMM and shape parameters are initialized using pre-defined values.

Iterative Minimization

REPEAT

E(stimation) Step

Find segmentation using graph cuts:

$$\hat{\mathbf{f}} = \min_{\mathbf{f}} E(\mathbf{I}, \underline{\theta}, \mathbf{f}, \mathbf{w}, \mathbf{b}).$$

M(inimization) Step

Assign M Gaussian to each voxel:

$$\hat{b}_n^m = \underset{b_n^m}{\operatorname{argmin}}\, E_{\mathrm{D}}(I_n, \underline{\theta}, f_n = m, b_n).$$

Learn Gaussian mixture model parameters:

$$\mu_d = \frac{\sum_{f_n=m, b_n^m=d} I_n}{\sum_{f_n=m, b_n^m=d} 1},\ \Sigma_d = \frac{\sum_{f_n=m, b_n^m=d} (I_n-\mu_d)^2}{\sum_{f_n=m, b_n^m=d} 1},\ \pi_d = \frac{\sum_{f_n=m, b_n^m=d} 1}{\sum_{f_n=m} 1}.$$

Update shape parameters through gradient descent until convergence:

$$\mathbf{w}_{t+1} = \mathbf{w}_t - \alpha(\nabla_{\mathbf{w}} E_{\mathrm{P}} + \nabla_{\mathbf{w}} E_{\mathrm{S}}).$$

UNTIL GMM/SHAPE PARAMETERS CONVERGE

5.3 Elliptical Priors For Kidney Segmentation

We model the kidney using an ellipse. We denote the shape parameters as $\mathbf{w} = \{w_0, w_1, w_2, w_3, w_4\}$, where (w_0, w_1) is the center of the ellipse, w_2 and w_3 its radii

for the major and minor axes and w_4 its rotation. We denote the ellipse using its signed distance function $\mathbf{V} = (V_1, ..., V_N)$.

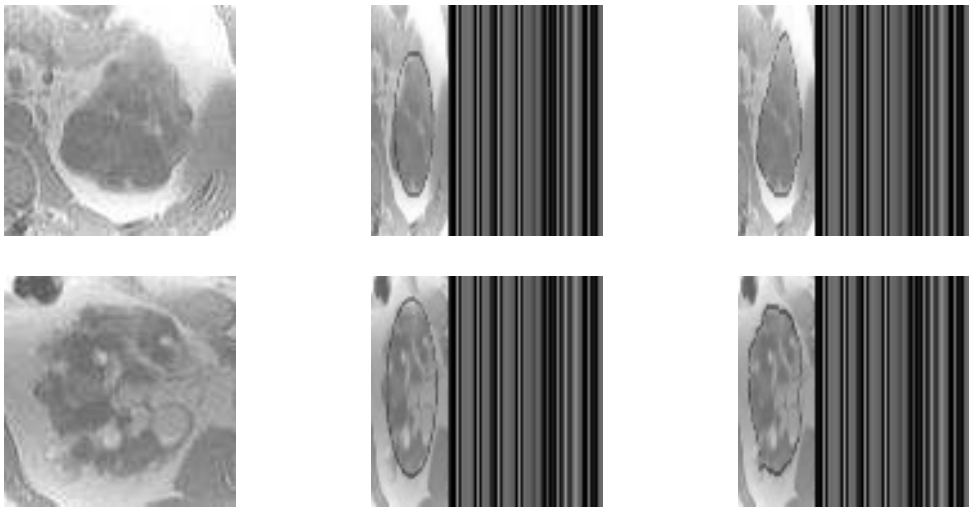
In this case, we have two regions we want to divide the image into: the kidney ($f_n = 0$) and the background ($f_n = 1$). We use one Gaussian for the kidney and three Gaussians for the background. For U_n^m in the third term of equation (5.1), we use $U_n^0 = -V_n$ and $U_n^1 = V_n$. $N_w(i)$ and A_w can be calculated as:

$$\begin{aligned} N_0(i) &= \sum_n H(-V_n)\delta(I_n - i), \\ N_1(i) &= \sum_n H(V_n)\delta(I_n - i), \\ A_0 &= \sum_n H(-V_n), \\ A_1 &= \sum_n H(V_n). \end{aligned} \tag{5.15}$$

$\nabla_{\mathbf{w}}\mathbf{U}$ can again be updated easily using geometry whereas

$$\begin{aligned} \frac{\partial N_0(i)}{\partial w_k} &= -\sum_n \delta(-V_n)\frac{\partial V_n}{\partial w_k}\delta(I_n - i), \\ \frac{\partial N_1(i)}{\partial w_k} &= \frac{\partial N_0(i)}{\partial w_k}, \\ \frac{\partial A_0}{\partial w_k} &= -\sum_n \delta(-V_n)\frac{\partial V_n}{\partial w_k}, \\ \frac{\partial A_1}{\partial w_k} &= \frac{\partial A_0}{\partial w_k}. \end{aligned} \tag{5.16}$$

Fig. 5.1 shows kidneys segmented from MR images by our proposed method. Although these images exhibit many challenges for segmentation, such as having weak contrast at boundaries between the kidney and the muscles and having strong contours within the kidney, our method was able to segmentation them accurately.



(a) Original MR images (b) Shape priors (c) Segmentations

Figure 5.1: Two kidney segmentations obtained by our method are shown in each row. The left column shows the original MR images, the middle and right column shows the shape priors and segmentations obtained by our method. Although these images have weak contrast at boundaries and strong contrast within the kidney, our method was able to segmentation them accurately.

5.4 Concentric Circle Priors For Left Ventricle Segmentation

Since in short-axis MR cardiac images, the left ventricle roughly resembles the shape of a donut, we model the left ventricle using two concentric circles. We denote the shape parameters as $\mathbf{w} = \{w_0, w_1, w_2, w_3\}$, where (w_0, w_1) is the center of the circles and w_2 and w_3 are respectively the radii for the inner circle (representing the endocardium boundary) and outer circle (representing the epicardium boundary). We denote the signed distance functions of the concentric circles as $\mathbf{V} = (\mathbf{V}^0, \mathbf{V}^1)^T$, where $\mathbf{V}^0 = (V_1^0, ..., V_N^0)$ and $\mathbf{V}^1 = (V_1^1, ..., V_N^1)$ are respectively the signed distance functions of the inner and outer circles. These functions are assigned zero value for boundary voxels and negative and positive distances respectively for voxels inside and outside the boundaries.

Here, we have three regions we want to divide the image into: the blood pool ($f_n = 0$), the myocardium ($f_n = 1$) and the background ($f_n = 2$). We use one Gaussian each for the first two regions and four Gaussians for the background. For U_n^m in the third term of equation (5.1), we use $U_n^0 = -V_n^0$, $U_n^2 = V_n^1$ and

$$U_n^1 = \begin{cases} V_n^0 & (V_n^0 < 0) \\ \text{less}(V_n^0, |V_n^1|) & (V_n^0 > 0, V_n^1 < 0) \\ -|V_n^1| & (V_n^1 > 0), \end{cases} . \tag{5.17}$$

If we denote a heavyside step function with $H(\bullet)$ and a delta function with $\delta(\bullet)$, then

$$\begin{aligned} N_0(i) &= \sum_n H(-V_n^0)\delta(I_n - i), \\ N_1(i) &= \sum_n H(V_n^0)H(-V_n^1)\delta(I_n - i), \\ N_2(i) &= \sum_n H(V_n^1)\delta(I_n - i), \\ A_0 &= \sum_n H(-V_n^0), \\ A_1 &= \sum_n H(V_n^0)H(-V_n^1), \\ A_2 &= \sum_n H(V_n^1). \end{aligned} \tag{5.18}$$

$\nabla_{\mathbf{w}}\mathbf{U}$, $\frac{\partial N_m}{\partial w_k}$ and $\frac{\partial A_m}{\partial w_k}$ are necessary to update $\mathbf{w}$ using equation (5.11). $\nabla_{\mathbf{w}}\mathbf{U}$ can

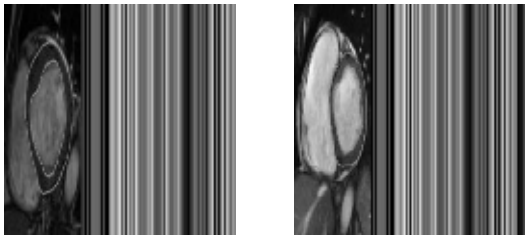
be updated easily using geometry whereas

$$\begin{aligned}
\frac{\partial N_0(i)}{\partial w_k} &= -\sum_n \delta(-V_n^0)\frac{\partial V_n^0}{\partial w_k}\delta(I_n - i) \\
\frac{\partial N_2(i)}{\partial w_k} &= \sum_n \delta(V_n^1)\frac{\partial V_n^1}{\partial w_k}\delta(I_n - i) \\
\frac{\partial N_1(i)}{\partial w_k} &= -\frac{\partial N_0(i)}{\partial w_k} - \frac{\partial N_2(i)}{\partial w_k} \\
\frac{\partial A_0}{\partial w_k} &= -\sum_n \delta(-V_n^0)\frac{\partial V_n^0}{\partial w_k} \\
\frac{\partial A_2}{\partial w_k} &= \sum_n \delta(V_n^1)\frac{\partial V_n^1}{\partial w_k} \\
\frac{\partial A_1}{\partial w_k} &= -\frac{\partial A_0}{\partial w_k} - \frac{\partial A_2}{\partial w_k}
\end{aligned} \quad . \tag{5.19}$$

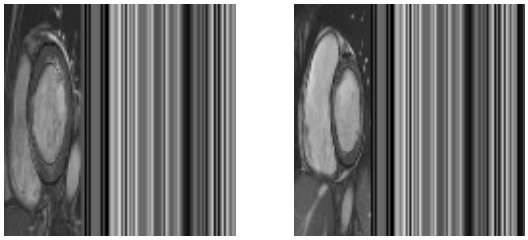
Fig. 5.2 shows the segmentations and the shape priors obtained by our method for left ventricles. Although these images have weak contrast at boundaries between the epicardium and the liver and those between the endocardium and the blood pool, our method was able to segmentation them accurately. The average Dice Similarity Measure between our segmentation and the hand segmentation for the entire dataset was 0.81, whereas over 0.7 indicates a good agreement.

5.5 Discussions

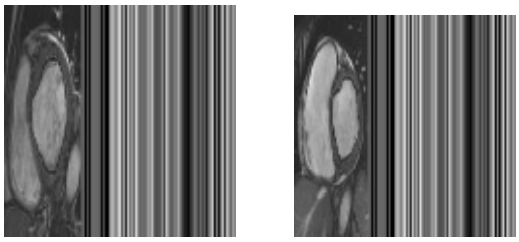
In this chapter, we proposed a novel segmentation method that incorporates geometric shape priors, which do not require statistical training, with the graph cuts technique for robust and efficient segmentations medical images. We demonstrated the effectiveness of our method by segmenting the left ventricle using concentric circles as shape priors and by segmenting the kidney using an ellipse as a shape prior.



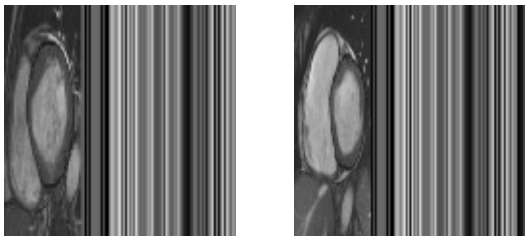
(a) Hand segmentations



(b) Geometric priors calculated by our method



(c) Endocardium boundaries obtained by our method



(d) Epicardium boundaries obtained by our method

Figure 5.2: Two left ventricle segmentations obtained by our method are shown in each column. The top row shows the original cardiac images. The second row shows the shape priors and the third and fourth row show the endocardium and epicardium boundaries obtained by our method.

Chapter 6
Conclusions

6.1 Conclusions

In this thesis, we developed highly accurate and robust graph cuts-based methods that automatically segment MR images. The images in which we are most interested are those that cannot be correctly segmented using intensity information alone. We developed models for robustly incorporating prior information such as spatial atlas and geometric or statistical shape priors into the efficient graph cuts segmentation framework. The accuracy of the alignment of the atlas with the image and that of the shape prior are important to the accuracy of the segmentation. In order to improve these, we use an EM-based approach and alternately calculate the alignment of the atlas/the shape prior and the segmentation of the image. We also designed our models so that updates of the atlas registration and the shape prior are not merely based on the segmentation, in order to avoid a bad segmentation from sending the methods into vicious cycles. Finally, we proposed a way of dealing with objects with curvy boundaries.

Specifically, we developed methods to incorporate spatial atlas with graph-based segmentation for MR brain images. Another kind of prior information we incorporated with graph cuts for MR image segmentation is shape priors. We studied statistical shape priors and the simpler geometric shape priors. As shown in chapter 3, an EM-based iterative method for segmentation and prior calculation is better than a sequential method because correctly classified voxels in the current segmentation make the match between the prior information and the image more accurate. This could also be extended to methods described in chapter 4 and 5. We tested our methods on MR brain, abdomen and cardiac images with inten-

sity inhomogeneity, poor contrast at desired boundaries and/or strong contrast at undesired boundaries and obtained encouraging results.

6.2 Future Works

Many immediate possible future works come to mind. Our proposed methods for incorporating shape priors require quantitative validation on images with ground truth segmentation. A comparison of our proposed methods and their equivalent sequential methods would also be interesting in this case. The spatial brain atlas we used was constructed through linear registration. Non-rigid registration is better at capturing variabilities within the training data, and an atlas based on it might be able to improve our algorithm. Although we used local GMMs to compensate for intensity non-uniformity and the results are quite good, many other methods have been proposed to address this issue in literature dividing voxels into homogenous and inhomogeneous [4], using low pass filters [83] and modeling with spline functions [47]. It is certainly interesting to validate our approach against other methods. The EM-based approach provides a good framework for incorporating prior information. Prior information can be used to improve basic image segmentation techniques such as region-growing.

BIBLIOGRAPHY

[1] "Statistical parameter mapping," http://www.fil.ion.ucl.ac.uk/spm.

[2] R. Ahuja, T. Magnanti, and J. Orlin, *Network Flows: Theory, Algorithms, and Applications*. Prentice Hall, 1993.

[3] J. Ashburner and K. Friston, "Unified segmentation," *NeuroImage*, pp. 839 – 851, 2005.

[4] S. Aylward and J. Coggins, "Spatially invariant classification of tissues in mr images," in *Conference on Visualization in Biomedical Computing*, 1994.

[5] R. Bansal, L. Staib, Z. Chen, A. Rangarajan, J. Knisely, R. Nath, and J. Duncan, "Entropy-based, multiple-portal-to-3DCT registration for prostate radiotherapy using iteratively estimated segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, 1999, pp. 567 – 578.

[6] E. Bardinet, L. Cohen, and N. Ayache, "Tracking and motion analysis of the left ventricle with deformable superquadrics," *Medical Image Analysis*, pp. 129 – 149, 1996.

[7] C. Bishop, *Neural Networks for Pattern Recognition*. Oxford University Press, 1995.

[8] S. Borman, "The expectation maximization algorithm – a short tutorial," July 2004.

[9] Y. Boykov and M. Jolly, "Interactive organ segmentation using graph cuts," in *Medical Image Computing and Computer-Assisted Intervention*, 2000, pp. 276 – 286.

[10] Y. Boykov and V. Kolmogorov, "Computing geodesics and minimal surfaces via graph cuts," in *IEEE International Conference on Computer Vision*, 2003, pp. 26 – 34.

[11] ——, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1124 – 1137, 2004.

[12] Y. Boykov, O. Veksler, and R. Zabih, "Markov random fields with efficient approximations," in *IEEE Conference on Computer Vision and Pattern Recognition*, 1998, pp. 648 – 655.

[13] ——, "Fast approximate energy minimization via graph cuts," in *IEEE International Conference on Computer Vision*, 1999, pp. 377–384.

[14] P. Casale, R. Devereux, D. Alonso, E. Campo, and P. Kligfield, "Improved sex-specific criteria of left ventricular hypertrophy for clinical and computer interpretation of electrocardiograms: validation with autopsy findings," *Circulation*, pp. 565–572, 1987.

[15] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," in *IEEE International Conference on Computer Vision*, 1995, pp. 694–699.

[16] ——, "Geodesic active contours," *International Journal of Computer Vision*, pp. 61 – 79, 1997.

[17] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, pp. 266–277, 2001.

[18] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser, "Using prior shapes in geometric active contours in a variational framework," *International Journal of Computer Vision*, pp. 315–328, 2002.

[19] C. Cocosco, A. Zijdenbos, and A. Evans, "A fully automatic and robust brain MRI tissue classification method," *Medical Image Analysis*, pp. 513 – 527, 2003.

[20] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multi-modality image registration based on information theory," *Information Processing in Medical Imaging*, pp. 263 – 274, 1995.

[21] D. Collins, A. Zijdenbos, W. Barre, and A. Evans, "ANIMAL + INSECT: improved cortical structure segmentation," *Inf. Process. Med.*, p. 1613, 1999.

[22] W. Cook, W. Cunningham, W. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*. John Wiley and Sons, 1998.

[23] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical imaging," *Imaging and Vision Computing*, p. 335C366, 1994.

[24] D. Cremers, T. Kohlberger, and C. Schnorr, "Nonlinear shape statistics in mumfordcshah based segmentation," in *ECCV*, 2002, p. 93C108.

[25] R. Damadian, M. Goldsmith, and L. Minkoff, "NMR in cancer: XVI. fonar image of the live human body," *Physiol. Chem. and Phys.*, pp. 97–100, 1977.

[26] A. Dempster, N. Laird, and D. Rubin, "Maximal likelihood form incomplete data via the EM algorithm," *Proceedings of the Royal Statistical Society*, pp. 1 – 38, 1977.

[27] A. Evans, D. Collins, P. Neelin, D. MacDonald, M. Kamber, and T. Marrett, "Three-dimensional correlative imaging: applications in human brain mapping," *Functional neuroimaging: technical foundations*, p. 145C162, 1994.

[28] M. Everingham and B. Thomas, "Supervised segmentation and tracking of non-rigid objects using a mixture of histograms model," in *IEEE International Conference on Image Processing*, 2001, pp. 62–65.

[29] B. Fischl, D. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. van der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, A. Montillo, N. Makris, B. Rosen, and A. Dale, "Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain," *Neuron*, p. 33, 2002.

[30] D. Freedman and T. Zhang, "Interactive graph cut based segmentation with shape priors," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 755–762.

[31] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 721 – 741, 1984.

[32] J. Grantham, V. Torres, A. Chapman, L. Guay-Woodford, K. Bae, B. King, L. Wetzel, D. Baumgarten, P. Kenney, P. Harris, S. Klahr, W. Bennett, G. Hirschman, C. Meyers, X. Zhang, F. Zhu, and J. Miller., "Volume progression in polycystic kidney disease," *The New England Journal of Medicine*, pp. 2122–2130, 2006.

[33] D. Greig, B. Porteous, and A. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society*, pp. 271 – 279, 1989.

[34] J. Hammersley and P. Clifford, "Markov field on finite graphs and lattices," 1971.

[35] K. Held, E. Kops, B. Krause, W. Wells, R. Kikinis, and H. Muller-Gartner, "Markov random field segmentation of brain mr images," *IEEE Transactions on Medical Imaging*, vol. 16, p. 878C886, 1997.

[36] Z. Hou, "A review on mr image intensity inhomogeneity correction," *International Journal of Biomedical Imaging*, pp. 1–11, 2006.

[37] H. Ishikawa and D. Geiger, "Segmentation by grouping junctions," in *IEEE Conference on Computer Vision and Pattern Recognitionf*, 1998, pp. 125 – 131.

[38] M. Kamber, R. Shinghal, D. L. Collins, G. S. Francis, and A. C. Evans, "Model-based 3-D segmentation of multiple sclerosis lesions in magnetic resonance brain images," *IEEE Transactions on Medical Imaging*, vol. 14, no. 3, pp. 442–453, 1995.

[39] T. Kapur, "Model based three dimensional medical imaging segmentation," Ph.D. dissertation, Massachusetts Institute of Technology, 1999.

[40] T. Kapur, W. Grimson, W. Wells, and R. Kikinis, "Segmentation of brain tissue from magnetic resonance images," *Medical Image Analysis*, pp. 109–127, 1996.

[41] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour model," *International Journal of Computer Vision*, p. 321C331, 1988.

[42] A. Kelemen, G. Szekely, and G. Gerig, "Elastic model-based segmentation of 3-D neuroradiological data sets medical imaging," *IEEE Transactions on Medical Imaging*, p. 828 C 839, 1999.

[43] R. Kikinis, M. Shenton, G. Gering, M. Anderson, J. Martin, D. Metcalf, C. Guttmann, R. McCarley, W. Lorensen, H. Line, and F. Jolesz, "Routine quantitative analysis of brain and cerebrospinal fluid spaces with MR imaging," *Journal of Magnetic Resonance Imaging*, pp. 619–629, 1992.

[44] V. Kolmogorov and R. Zabih, "What energy functions can be minimized via graph cuts?" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 147–159, 2004.

[45] S. Lee, G. Wolberg, K. Chwa, and S. Shin, "Image metamorphosis with scattered feature constraints," *IEEE Trans. Visualization Comput. Graph.*, pp. 337 – 354, 1996.

[46] S. Lee, G. Wolberg, and S. Shin, "Scattered data interpolation with multilevel B-splines," *IEEE Trans. Visualization Comput. Graph.*, p. 228 C 244, 1997.

[47] K. Leemput, F. Maes, D. VanDermeulen, and P. Suetens, "Automated model-based bias field correction of MR images of the brain," *IEEE Transactions on Medical Imaging*, pp. 885 – 895, 1999.

[48] M. Leventon, W. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2000, pp. 1316–1323.

[49] J. Mallard, "Magnetic resonance imaging : the Aberdeen perspective on developments in the early years," *Phys Med Biol*, vol. 51, pp. 45–60, 2006.

[50] S. Marlow and N. Connor, "Supervised object segmentation and tracking for MPEG-4 VOP generation," in *International Conference on Pattern Recognition*, 2000, pp. 1125–1128.

[51] J. Marroquin, E. Santana, and S. Botello, "Hidden Markov measure field models for image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 1380–1387, 2003.

[52] J. Mattson and M. Simon, *The Pioneers of NMR and Magnetic Resonance in Medicine: The Story of MRI.* Jericho and New York: Bar-Ilan University Press, 1996.

[53] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions.* John Wiley and Sons, Inc., 1997.

[54] G. McLachlan and D. Peel, *Finite Mixture Models.* John Wiley and Sons, 2000.

[55] A. Messiah, *Quantum Mechanics.* Courier Dover Publications, 1999.

[56] S. Mueller, M. Weiner, L. Thal, R. Petersen, C. Jack, W. Jagust, J. Trojanowski, A. Toga, and L. Beckett, "Ways toward an early diagnosis in Alzheimers disease: The Alzheimers disease neuroimaging initiative (ADNI)," *Alzheimers Dement*, pp. 55–66, 2005.

[57] W. Oldendorf, *Basics of Magnetic Resonance Imaging.* Springer, 1988.

[58] S. Osher and J. Sethian, "Fronts propagation with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, pp. 12–49, 1988.

[59] C. Pachai, Y. Zhu, C. Guttmann, R. Kikinis, F. Jolesz, G. Gimenez, J. Froment, C. Confavreux, and S. Warfield, "Unsupervised and adaptive segmentation of multispectral 3D magnetic resonance images of human brain: a generic approach," in *Medical Image Computing and Computer-Assisted Intervention*, 2001, p. 1067C1074.

[60] N. Paragios and R. Deriche, "Geodesic active regions: A new framework to deal with frame partition problems in computer vision," *Journal of Visual Communication and Image Representation*, pp. 249 – 268, 2002.

[61] N. Paragios, M. Rousson, and V. Ramesh, "Knowledge-based registration and segmentation of the left ventricle: A level set approach," in *WACV*, 2002, pp. 37 – 42.

[62] T. Paus, F. Tomaiuolo, N. Otaky, D. MacDonald, M. Petrides, J. Atlas, R. Morris, and A. Evans, "Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map," *Cereb Cortex*, p. 207C214, 1996.

[63] S. Pizer, G. Gerig, S. Joshi, and S. Aylward, "Multiscale medial shape-based analysis of image objects," *IEEE Special Issue on: Emerging Medical Imaging Technology*, p. 670 C 679, 2003.

[64] K. Pohl, S. Bouix, R. Kikinis, and W. Grimson, "Anatomical guided segmentation with non-stationary tissue class distributions in an expectation-maximization framework," in *IEEE International Symposium on Biomedical Imaging.*, 2004, p. 81C 84.

[65] K. Pohl, J. Fisher, W. Grimson, R. Kikinis, and W. Wells, "A Bayesian model for joint segmentation and registration," *NeuroImage*, pp. 228 – 239, 2006.

[66] K. Pohl, S. Warfield, R. Kikinis, W. Grimson, and W. Wells, "Coupling statistical segmentation and PCA shape modeling," in *Medical Image Computing and Computer-Assisted Intervention*, 2004, pp. 151–159.

[67] T. Rohlfing and C. Maurer, "Multi-classifier framework for atlas-based image segmentation," *Pattern Recognition Letters*, vol. 26, no. 13, pp. 2070–2079, 2005.

[68] M. Rousson and N. Paragios, "Shape priors for level set representations," in *ECCV*, 2002, pp. 78–92.

[69] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, "Non-rigid registration using free-form deformations: Application to breast MR images," *IEEE Transactions on Medical Imaging*, pp. 712 – 722, 1999.

[70] L. Shapiro and G. Stockman, *Computer Vision*. Prentice-Hall, 2001.

[71] L. Silbert, J. Quinn, M. Moore, E. Corbridge, M. Ball, G. Murdoch, G. Sexton, and J.Kaye, "Changes in premorbid brain volume predict Alzheimers disease pathology," *Neurology*, pp. 487–492, 1999.

[72] G. Slabaugh and G. Unal, "Graph cuts segmentation using an elliptical shape prior," in *ICIP*, 2005, pp. 1222–1225.

[73] C. Studholme, "Simultaneous population based image alignment for template free spatial normalisation of brain anatomy. 2nd international workshop on biomedical image registration," in *Workshop on Biomedical Image Registration*, 2003, pp. 81–90.

[74] P. Thevenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Transactions of Image Processing*, pp. 2083 – 2099, 2000.

[75] D. Thirde and G. Jones, "Hierarchical probabilistic models for video object segmentation and tracking," in *IEEE International Conference on Pattern Recognition*, 2004, pp. 636–639.

[76] A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, W. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Transactions on Medical Imaging*, pp. 137–154, 2003.

[77] M. Unser, A. Aldroubi, and M. Eden, "B-spline signal processing: Part I - Theory," *IEEE Transactions on Signal Processing*, pp. 821 – 832, 1993.

[78] O. Veksler, "Efficient graph-based energy minimization methods in computer vision," Ph.D. dissertation, Cornell University, 1999.

[79] P. Viola and W. Wells, "Alignment by maximization of mutual information," in *IEEE International Conference on Computer Vision*, 1995, pp. 16 – 23.

[80] F. Wang and B. Vemuri, "Simultaneous registration and segmentation of anatomical structures from brain MRI," in *Medical Image Computing and Computer-Assisted Intervention*, 2005, pp. 17 – 25.

[81] S. Warfield, F. Jolesz, and R. Kikinis, "Real-time image segmentation for image-guided surgery," in *IEEE/ACM Conference on Supercomputing*, 1998, pp. 42 – 42.

[82] S. Warfield, K. Zou, and W. Wells, "Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation," *IEEE Transactions on Medical Imaging*, pp. 903–921, 2004.

[83] W. Wells, W. Grimson, R. Kikinis, and F. Jolesz, "Adaptive segmentation of MRI data," *IEEE Transactions on Medical Imaging*, vol. 15, p. 429C442, 1996.

[84] P. Wyatt and J. Noble, "MRF-MAP joint segmentation and registration," in *Medical Image Computing and Computer-Assisted Intervention*, 2002, p. 580C587.

[85] C. Xiaohua, M. Brady, J. Lo, and N. More, "Simultaneous segmentation and registration of contrast-enhanced breast MRI," *Information Processing in Medical Imaging*, pp. 126 – 137, 2005.

[86] A. Yezzi, S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, "A geometric snake model for segmentation of medical imagery," *IEEE Transactions on Medical Imaging*, pp. 199 – 209, 1997.

[87] A. Yezzi, L. Zollei, and T. Kapur, "A variational framework for joint segmentation and registration," in *IEEE CVPR - MMBIA*, 2001, p. 388C400.

[88] J. Zhang, "The mean field theory in EM procedures for blind Markov random field image restoration," *IEEE Transactions on Image Processing*, vol. 2, no. 1, pp. 27–40, 1993.

[89] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm," *IEEE Transactions on Medical Imaging*, vol. 20, p. 45C57, 2001.

[90] A. Zijdenbos, B. Dawant, R. Margolin, and A. Palmer, "Morphometric anal-

ysis of white matter lesions in MR images: method and validation," *IEEE Transactions on Medical Imaging*, pp. 716 – 724, 1994.

[91] Z.Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1101 – 1113, 1993.