An MRI Segmentation Framework for Brains with Anatomical Deviations

Marcelinus Prastawa

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Computer Science.

Chapel Hill
2007

Approved by:

Guido Gerig, Ph.D.

Stephen R. Aylward, Ph.D.

Elizabeth Bullitt, M.D.

Sarang C. Joshi, D.Sc.

J. Stephen Marron, Ph.D.

Stephen M. Pizer, Ph.D.

ABSTRACT

MARCELINUS PRASTAWA: An MRI Segmentation Framework for Brains with

Anatomical Deviations

(Under the direction of Guido Gerig, Ph.D.)

The segmentation of brain Magnetic Resonance (MR) images, where the brain is partitioned into anatomical regions of interest, is a notoriously difficult problem when the underlying brain structures are influenced by pathology or are undergoing rapid development. This dissertation proposes a new automatic segmentation method for brain MRI that makes use of a model of a homogeneous population to detect anatomical deviations. The chosen population model is a *brain atlas* created by averaging a set of MR images and the corresponding segmentations. The segmentation method is an integration of robust parameter estimation techniques and the Expectation-Maximization algorithm.

In clinical applications, the segmentation of brains with anatomical deviations from those commonly observed within a homogeneous population is of particular interest. One example is provided by brain tumors, since delineation of the tumor and of any surrounding edema is often critical for treatment planning. A second example is provided by the dynamic brain changes that occur in newborns, since study of these changes may generate insights into regional growth trajectories and maturation patterns. Brain tumor and edema can be considered as anatomical deviations from a healthy adult population, whereas the rapid growth of newborn brains can be considered as an anatomical deviation from a population of fully developed infant brains.

A fundamental task associated with image segmentation is the validation of segmentation accuracy. In cases in which the brain deviates from standard anatomy, validation is often an ill-defined task since there is no knowledge of the ground truth (information

about the actual structures observed through MRI). This dissertation presents a new method of simulating ground truth with pathology that facilitates objective validation of brain tumor segmentations. The simulation method generates realistic-appearing tumors within the MRI of a healthy subject. Since the location, shape, and volume of the synthetic tumors are known with certainty, the simulated MRI can be used to objectively evaluate the accuracy of any brain tumor segmentation method.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

1.1. Motivation

Medical image segmentation is the task of classifying image components (pixels or voxels) into relevant anatomical components or describing the structural and intensity changes in terms of the underlying functional process. The knowledge of the location, size, and shape of different anatomical structures is a fundamental step in understanding and analyzing medical images. Explicit knowledge of the segmented structures in medical images allows us to do more than qualitative visual assessment, as in the following examples:

- The location of a pathology relative to healthy anatomical structures is useful in planning radiological treatments and surgeries.
- Growth patterns can be determined by analyzing changes of segmented structures of a population group over time.
- Analysis of the shape of the segmented brain structures can be used to find characteristics or markers of neurological disorders.

Magnetic Resonance Imaging (MRI) [102] is currently one of the most powerful imaging techniques available to obtain in-vivo anatomical and functional brain data. The most widely used and trusted method of obtaining segmentation of MR images is the manual labeling of image pixels by human experts. In an ideal setting, manual segmentation by trained experts provides the good results with low to moderate probability of errors. However, the task of segmenting or annotating 3D Magnetic Resonance (MR) images is
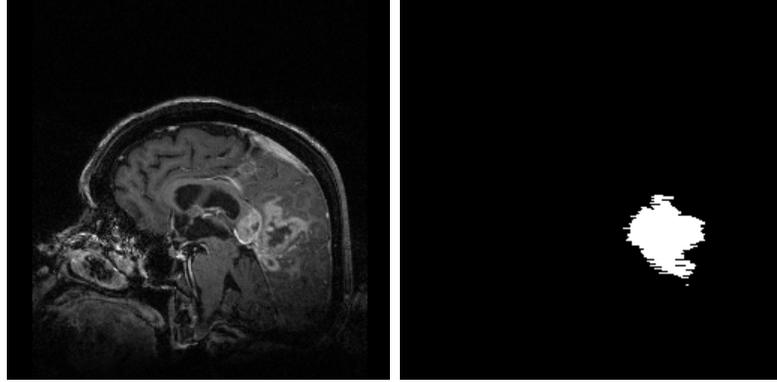
FIGURE 1.1. Gadolinium contrast enhanced T1-weighted MR image (sagittal view) and the manual segmentation result. Note the ragged outline in the segmentation that can be attributed to the slice-by-slice 2D painting in the axial direction (Tumor031 dataset).

generally time consuming and challenging. Furthermore, manual segmentations are difficult to reproduce in a reliable and objective manner, even by the same human expert. The task is mostly performed by drawing image regions slice-by-slice, limiting the human rater's view and generating suboptimal outlines with limited consistency across slices. An example of a manual segmentation of brain tumor from MRI is shown in Figure 1.1.

Due to the limitations of manual segmentation methods, an automatic segmentation framework is crucial for the study of medical phenomena, especially when it involves a large set of images. An automatic segmentation method is desirable because it reduces the work load of human experts and generates fully reproducible segmentations. A computer program also has the advantage of being able to process large amounts of information as typically presented within 3D multi-modal MR images in a more consistent manner compared to human raters.

1.2. Automatic MRI Segmentation

The task of automatically segmenting medical images, as opposed to natural scenes, has the significant advantage that structural and intensity characteristics are well-known up to a natural biological variability and presence of other factors such as pathology. The typical adult brain tissue can be divided into three main categories: white matter
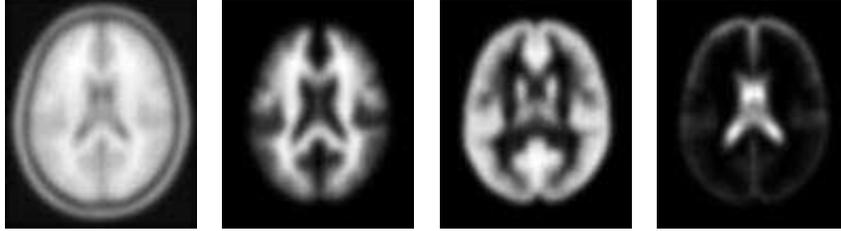
FIGURE 1.2. The ICBM (International Consortium of Brain Mapping) brain atlas that functions as a spatial probabilistic model for a healthy adult population. From left to right: T1 weighted template image and probability values of white matter, gray matter, and csf.

(information transmitters), gray matter (information processors), and cerebrospinal fluid. The central gray matter region can be divided further into other structures such as the caudate, hippocampus, etc. [37]. A widely used model for the general adult brain population is the probabilistic brain atlas [20], which is created by averaging MR images and the corresponding sets of segmentations. The ICBM (International Consortium of Brain Mapping) brain atlas (Figure 1.2) provides the spatial probabilities of a brain location being white matter, gray matter, or cerebrospinal fluid (csf). The atlas was created by registering subject images using affine transformation. Due to the limited degrees of freedom for affine transformation, most of the subject variability is retained and the atlas appears blurry. A sharper atlas can be created by using a deformable registration with more degrees of freedom, such as the fluid image warping [55]. A mesh-based atlas generation scheme that automatically determines the degree of warping and blurring was proposed by van Leemput [94].

Existing automatic MRI segmentation methods make use of the brain atlas as spatial priors [93] or as sampling constraints [17]. These methods provide good results for healthy brain MR images that have similar structure to the one described by the brain atlas. In clinical applications however, there is strong interest in analysis of MR images that show deviations from the typical population, which implies deviations from the reference population model (brain atlas). These deviations can be caused by pathology or natural growth patterns, as shown in Figure 1.3. The standard atlas-based segmentation methods fail in detecting anatomical deviations because they typically do not take into
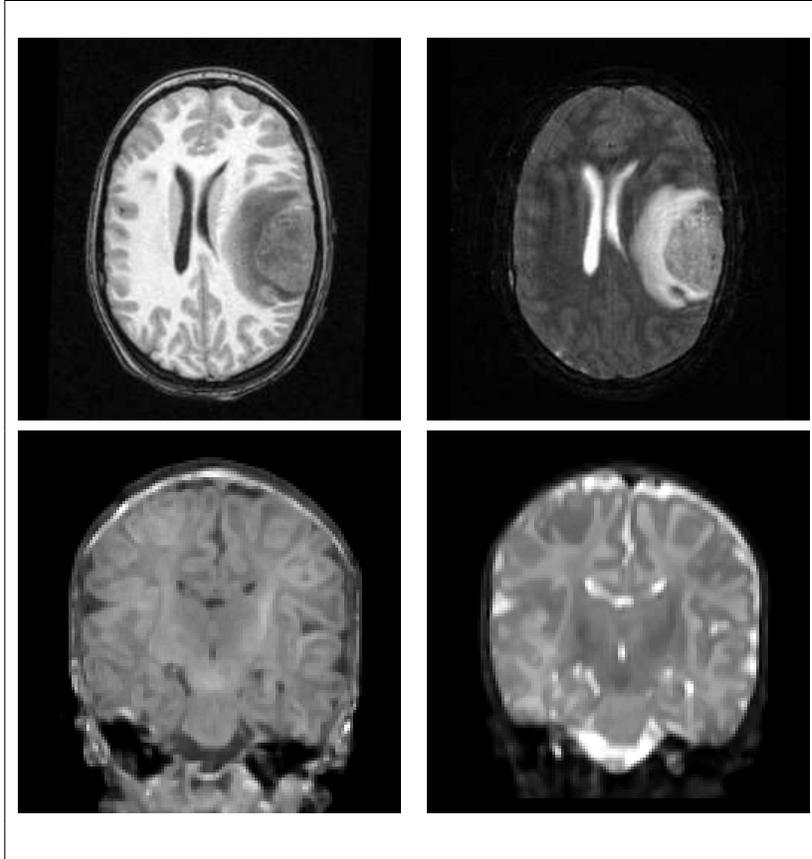
FIGURE 1.3. Example MR images that exhibit deviations from a reference population. Top: MRI of an adult with tumor and edema (T1w and T2w) which show deformation due to tumor mass effect and infiltration of brain tissue by edema. Bottom: MRI of a newborn infant (T1w and T2w) which shows the presence of two different types of white matter due to myelination.

account strucutural and intensity changes not modelled by the atlas. Figure 1.4 shows an example result of applying the method proposed by van Leemput *et al.* [93] to a brain tumor MRI. The method uses the normal brain atlas as spatial priors and computes the anatomical label assignment using the Expectation-Maximization (EM) algorithm [25]. The tumor region is incorrectly labeled as a fluid region since tumor appearance is similar to appearance of fluid in the standard T1w and T2w MRI scans.

Atlas-based segmentation approaches, such as the one proposed by van Leemput *et al.* , have been shown to perform well on MRI of healthy subjects. However, they do not take into account the deviations in intensity and structure caused by pathology and
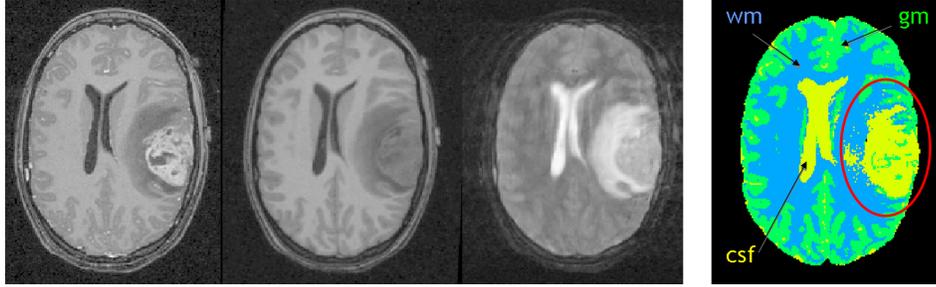
FIGURE 1.4. Failure of atlas-based segmentation of brain tumor MRI when tumor structure is not taken into account. From left to right: contrast enhanced T1w, T1w, and T2w images; followed by the segmented label image. The tumor and edema regions (circled) are mostly considered to be part of the cerebrospinal fluid structure.

thus fail to estimate the proper anatomical assignments. I propose to extend the atlas-based segmentation approach using combinatorial robust parameter estimation methods that can handle significant proportions of outlier data due to noise, pathology, growth changes, or other deviations from the normal population model. The new robust approach is shown to be suitable for two interesting clinical problems: automatic segmentation of adult brain MRI with tumor and of newborn brain MRI with rapid myelination changes. In adult brains with tumor, tumor causes significant deformation due to mass effect while surrounding healthy tissue can be infiltrated by tumor cells and edema. These changes result in significant deviations from the atlas with regard to structure and appearance. Healthy infant brains undergo rapid growth during the first year, where the white matter fibers are being covered in myelin sheaths. The myelin sheath is a crucial component for the transmission of neural signals. Since white matter is not fully developed at birth, the structure does not appear homogeneous in newborn brain MRI [**81**]. The myelination process results in changes in appearance when compared to the standard atlas where white matter is modeled as a single tissue category.

1.3. Thesis and Contributions

Thesis: *Reference population models and priors on the possible deviations can be effectively combined in a robust maximum likelihood segmentation*

5
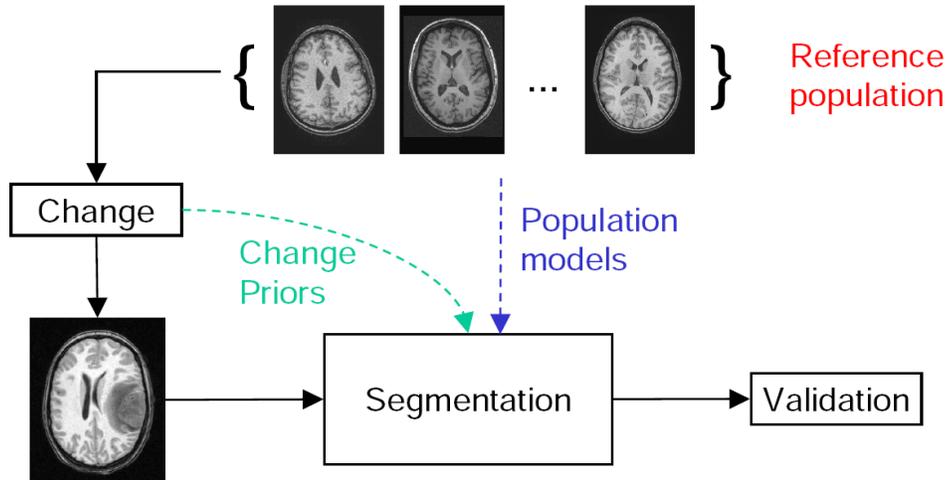
FIGURE 1.5. A conceptual overview of the proposed dissertation topic. It involves a segmentation framework which treats healthy adults as the reference population and tumor as a change process, and a validation framework for the segmentation results. The segmentation framework makes use of reference population models and priors on the change processes. The validation framework simulates the change processes to obtain known ground truth.

framework for brain MR images, extending applications to image data presenting pathology or neurodevelopmental and neurodegenerative changes. Considering the lack of a reliable ground truth, the reference population model and deterministic models of the anatomical deviations can also be combined to create synthetic ground truth. This facilitates objective evaluation of the performance of different segmentation methods.

A graphical overview of the topics covered in this dissertation is shown in Figure 1.5. The practical aim of the work described in this dissertation is the creation of a system that performs automatic segmentation of images with clinically interesting anatomical deviations, along with objective validation of segmentation results using well defined ground truth. Such a system has significant potential value for clinical studies involving large populations since it provides fully reproducible segmentations with reliable measures of segmentation performance for quality control.

Segmentation of MR images with pathological deviations, such as brain tumor, has been approached in different ways [13, 31, 47, 33]. The previously available tumor segmentation methods are not fully automatic and do not provide segmentations of healthy tissue and edema. Detailed review of other segmentation approaches for brain tumor MRI is covered in Section 3.1. Review of other segmentation approaches for newborn brain MRI is presented in Section 4.1. The segmentation framework proposed in this dissertation is fully automatic and provides a complete description of the 3D brain anatomy.

With regard to the simulation of MR images with pathological deviations, there is a lack of models that make use of relevant biological models. For example, the brain tumor MRI simulator proposed by Rexilius *et al.* [79] determines edema regions by using the white matter mask and distances to tumor boundary, and restricts contrast enhancement to the brain tumor regions. The brain tumor MRI simulation framework proposed in this dissertation uses a model of local diffusion properties for edema and computes contrast enhancement in both tumor and blood vessel regions.

My work as presented in this dissertation expands the previous work done by others in the field of Bayesian image segmentation and simulation of brain pathologies. The contributions of this dissertation are as follows:

(1) Image segmentation using a modified Expectation-Maximization (EM) algorithm: the novelties of this approach are its use of robust parameter estimation techniques and its automatic detection of the feature space clusters for the mixture model.

(2) Generation of an augmented feature space for image segmentation through the use of spatial constraints such as location, curvature, and adjacency.

(3) Application of the proposed segmentation framework for healthy brains as well as images that exhibit deviations due to pathology (brain tumor) and growth (newborn brains).

(4) A method for generating pathological ground truth (tumor and edema) from image data with known healthy ground truth by combining a linear elastic biomechanical model with random surface tractions and a reaction-diffusion process guided by diffusion tensor imaging (DTI). The simulation of a new pathological ground truth is guided by the underlying biological processes.

(5) Simulation of the accumulation of contrast agent for a brain tumor subject to generate contrast enhanced T1w MRI, which is the standard diagnostic imaging modality. The accumulation model is guided by the underlying biological processes.

(6) Simulation of MR images with brain tumor and edema using textures synthesized from real tumor MRI samples. The synthetic MR images and the associated ground truth provides the means for objective evaluation of different segmentation schemes.

1.4. Overview of Chapters

The remainder of this dissertation is organized as follows:

Chapter 2 presents the background material for Bayesian image segmentation with the maximum likelihood approach. This chapter also proposes a modification of the standard EM algorithm for computing the maximum likelihood estimate using robust parameter estimation techniques to detect deviations or noise.

Chapter 3 presents the application of the robust maximum likelihood image segmentation framework described in Chapter 2 for segmenting MRI of adult brains with tumor. The anatomical deviations from the adult brain atlas involve the deformation of healthy tissue due to tumor mass effect and the infiltration of the regions surrounding tumor by edema.

Chapter 4 presents the application of the robust maximum likelihood image segmentation framework described in Chapter 2 for segmenting MRI of newborn brains. The

early growth pattern is treated as an anatomical deviation compared to the child brain atlas, where white matter appears as two distinct regions.

Chapter 5 describes the challenges and goals of validating segmentation results where the ground truth is difficult to obtain. In this chapter, I will develop a framework for generating synthetic brain tumor MR images with the associated ground truth based on the simulation of tumor and edema growth processes.

Chapter 6 concludes with a summary of the contributions and discussion of possible future work.

CHAPTER 2

Maximum Likelihood Image Segmentation

This chapter describes the image segmentation framework that forms the main contribution of this dissertation. Section 2.1 discusses the basic concepts for Bayesian image segmentation. Section 2.2 describes the segmentation process by maximizing the image likelihood using the Expectation-Maximization (EM) algorithm. Methods for estimating model parameters from noisy data with outliers are discussed in Section 2.3. Finally, extensions to the EM algorithm for segmenting images with deviations from an expected model is presented in Section 2.4.

2.1. Background

An image $I = (I_k)$ is a collection of values arranged in a regular lattice Λ. In this dissertation, Λ refers to the 3D image lattice, where the gaps in the lattice configuration can be different for each dimension. For every $k \in \Lambda$, I_k represents the feature values associated with the voxel location k. In the case of multi-modal images, such as color images, I_k is a vector containing a fixed number of scalar component values. Image segmentation is the task of assigning labels to each image location based on the feature values. This results in what is typically called the segmented image or the label image $S = (S_k)$. The value S_k is an assignment label drawn from a finite set of classes or labels $\mathcal{C}$ or $S_k \in \mathcal{C}$.

The complexities of the real world configuration and the image acquisition process generally makes it impossible to have accurate deterministic models for image content. This leads to the development of probabilistic models, where an image I is considered as an observation drawn from a probability distribution. Bayesian image segmentation

provides a framework for estimating a map from an image I into a label image S, while balancing the prior knowledge information and the observed data.

There are three essential probability distributions functions involved in the Bayesian framework. The *prior* distribution embodies the knowledge of likely configurations before an actual image is observed. In contrast, the probability distribution that is derived after an observation has been made is called the *posterior* distribution. The *likelihood* is defined as the probability of obtaining a particular observation given a set of model parameters (a conditional probability).

The theorem proposed by Bayes [7] describes the relation between the posterior probability $p(B|A)$, prior probability $Pr(B)$, and likelihood $p(A|B)$.

THEOREM 2.1. *Bayes' theorem states that the posterior probability $p(B|A)$ is proportional to the likelihood $p(A|B)$ multiplied by the prior $Pr(B)$.*

$$p(B|A) = \frac{p(A|B)Pr(B)}{p(A)} \tag{2.1.1}$$

Within a Bayesian framework, segmentation of a given image is performed by estimating the label assignments and the parameters describing the image appearance and/or geometry. The label image S and the model parameters θ form a tuple that describes the world state that generates the image I,

$$W = (S, \theta).$$

Bayes estimators that map the image I into the segmentation S can be constructed using the joint probability $p(I, W)$. When the joint probability function is known, a simple estimator is to choose the most likely estimate $\hat{W}$ that maximizes $p(I, W)$. In practice however, the full joint probability is typically very complex thus most estimators use the associated conditional probabilities instead.

2.2. Image Segmentation using Expectation-Maximization

One can consider segmentation as a problem of finding the hidden label assignments from the observed image data. The Maximum Likelihood (ML) segmentation estimate is the one that maximizes the likelihood of observing the complete data $J = (I, S)$ given the model parameters θ: $p(J|\theta)$, which is an associated conditional probability of $p(I, W)$. When the parameters θ are known the estimate of the segmentation S is straightforward. On the other hand, when the hidden segmentation S is known we can infer the model parameters θ. This leads to the development of the well known Expectation-Maximization (EM) algorithm [25] for ML estimation.

The Expectation-Maximization algorithm can be considered as an optimization technique where we maximize the lower bound for the image likelihood function [67]. The lower bound is derived from Jensen's inequality:

$$\sum_j g(j) a_j \geq \prod_j g(j)^{a_j} \tag{2.2.1}$$

given $\sum_j a_j = 1$. In image segmentation, the EM algorithm finds the parameter which maximizes the probability of a configuration over all the possible values for the hidden label assignment. The function that is maximized is:

$$f(\theta) = p(I|\theta) = \sum_S p(I, S|\theta). \tag{2.2.2}$$

We can create a lower bound for $f(\theta)$ by using a probability function $q(S)$ and applying Jensen's inequality (Equation 2.2.1)

$$f(\theta) = \sum_S \frac{p(I, S|\theta)}{q(S)} q(S) \geq g(\theta, q) = \prod_S \left(\frac{p(I, S|\theta)}{q(S)} \right)^{q(S)}. \tag{2.2.3}$$

The function $q(S)$ needs to be chosen so that we obtain a good bound on the image likelihood function. The bound can be maximized using the log of the bound, which

yields

$$
\begin{aligned}
log\, g(\theta, q) &= \sum_S log\left(\frac{p(I, S|\theta)}{q(S)}\right)^{q(S)} \\
&= \sum_S q(S) log\frac{p(I, S|\theta)}{q(S)} \\
&= \sum_S q(S) log\, p(I, S|\theta) - q(S) log\, q(S). \qquad (2.2.4)
\end{aligned}
$$

The probability ratio within the log term can be rewritten as follows:

$$
\frac{p(I, S|\theta)}{q(S)} = \frac{p(S|I, \theta)}{p(S|I, \theta)}\frac{p(I, S|\theta)}{q(S)} = \frac{p(S|I, \theta)}{q(S)}\frac{p(I, \theta)}{p(\theta)} = \frac{p(S|I, \theta)}{q(S)}p(I|\theta).
$$

Thus, obtaining the following interpretation for the log of the bound that defines the optimal $q(S)$ [11]:

$$
\begin{aligned}
log\, g(\theta, q) &= \sum_S q(S) log\frac{p(I, S|\theta)}{q(S)} \\
&= \mathcal{E}_{q(S)}\{log\frac{p(S|I, \theta)}{q(S)}\} + log\, p(I|\theta) \\
&= -\mathcal{E}_{q(S)}\{log\frac{q(S)}{p(S|I, \theta)}\} + log\, p(I|\theta) \\
&= -D_{KL}(q(S)\,||\,p(S|I, \theta)) + log\, p(I|\theta) \qquad (2.2.5)
\end{aligned}
$$

where $\mathcal{E}$ is the expectation function and D_{KL} is the (non-negative) relative entropy or Kullback-Leibler divergence [49]. To obtain the optimal $q(S)$ for the bound, we need to minimize the relative entropy between $q(S)$ and the label posterior probability. This can be achieved when this divergence is made zero by using the same probability function, $q(S) = p(S|I, \theta)$. Computing q to obtain a good bound on the expected likelihood is called the E-step, while maximizing the bound over θ is called the M-step. With regard to image segmentation, the estimation of the hidden segmentation labels is called the E-step and the estimation of the best parameters θ from the complete labeled image data is called the M-step.

13

The Kullback-Leibler divergence has also been demonstrated to be a useful metric for defining average anatomies, as proposed by Lorenzen *et al.* [56, 55]. In this case, the divergence is minimized in order to maximize a lower bound on the Bayes probability of error between the average anatomy and the set of subject anatomies. This is similar to the maximization of a lower bound on complete data likelihood in the EM algorithm. Maximizing the Bayes probability of error ensure that the average anatomy resembles the actual observed anatomies.

The EM algorithm computes the ML estimate through an iterative process. In one iteration it performs the estimation of the $q(S)$ function given the current model parameters and the calculation of the model parameters θ that maximizes the complete data likelihood $p(J|\theta)$. During the nth iteration the algorithm proceeds as follows:

- E-step: Given the observation I and the current model parameter $\theta^{(n)}$, compute the conditional expectation of the complete data likelihood defined as $Q(\theta|\theta^{(n)})$,

$$Q(\theta|\theta^{(n)}) = \mathcal{E}_{p(S|I,\theta^{(n)})}\{log\, p(J|\theta)\} = \mathcal{E}_{p(S|I,\theta^{(n)})}\{log\, p(I,S|\theta)\}. \qquad (2.2.6)$$

  The goodness-of-fit function Q is derived by extracting the relevant term from Equation 2.2.4 and substituting $q(S)$ with the optimal probability function $p(S|I,\theta)$.

- M-step: Update the model parameters for the next iteration so that the expected complete data likelihood is maximized,

$$\theta^{(n+1)} = \arg\max_{\theta} Q(\theta|\theta^{(n)}). \qquad (2.2.7)$$

When the true maximization of the expected data likelihood is difficult to do, an alternative is to simply choose the model parameters that generate a higher value of data likelihood. The selection of higher likelihood values as opposed to a maximum gives rise to the Generalized EM (GEM) algorithm, which may be particularly attractive when the likelihood is modeled using a parametric distribution and the gradient of the expected data likelihood is available in closed form. Each iteration of the EM or GEM algorithm

guarantees that the log-likelihood $log\ p(J|\theta)$ is increased [103]. The GEM algorithm in practice converges to a local maxima given a particular initialization, which makes it particularly sensitive with regard to its initial values.

Here, I will develop an EM image segmentation algorithm using a Gaussian mixture model for multi-modal images with voxel data in $\mathbb{R}^d$. This is a standard approach that has been used for medical image segmentation by Wells *et al.* [100], van Leemput *et al.* [93], and Zhang *et al.* [106]. Image segmentation using EM is an iterative process where we begin with a model parameters $\theta^{(0)}$. Each segmentation iteration involves estimating the segmentation labels and updating the model parameters. The complete image data at different locations are assumed to be statistically independent, $J_u \perp\!\!\!\perp J_v \ \forall u \neq v$. For each location k, the appearance model or the individual image likelihood for a specific label or class c is

$$p(I_k|S_k = c, \theta) = \mathcal{N}_{\mu_c, \Sigma_c}(I_k) \tag{2.2.8}$$

where $\mathcal{N}_{\mu_c, \Sigma_c}$ is the multivariate normal distribution associated with class c with mean μ_c and covariance matrix Σ_c. The model parameters θ are composed of parameter values for the class-specific multivariate normals, $\theta = \{(\mu_i, \Sigma_i) \ \forall i \in \mathcal{C}\}$. With the voxelwise independence assumption the image likelihood becomes

$$p(I|W) = p(I|S, \theta) = \prod_k p(I_k|S_k, \theta). \tag{2.2.9}$$

In the case of natural images, the prior probabilities at specific locations $Pr(S_k)$ are not known. A common approach is to choose a global value ρ_c for a class label c based on experiments or prior knowledge, $Pr(S_k = c) = \rho_c \ \forall k$. In the case of anatomical images, the prior probabilities for specific structures are known to some degree. For example, the likely spatial configuration for the brain can be described reliably using the spatial priors $Pr(S_k = c)$ which describes the probability of observing an anatomical component of the brain (labeled by c) at location k. In Chapters 3 and 4, the spatial priors from the brain atlas will be used to segment real brain MRI. The use of an explicit, predefined spatial priors is relevant for medical images since we have known structures or anatomy.

The probabilistic models of $p(I_k|S_k = c, \theta)$ and $Pr(S_k)$ are the main components for ML estimation. Combining the image likelihood function (Equation 2.2.8) with the voxelwise-independence assumption leads to the following complete data likelihood function:

$$
\begin{aligned}
Q(\theta|\theta^{(n)}) &= \sum_c \sum_k log\left[p(I_k, S_k = c|\theta)\right] p(S_k = c|I_k, \theta^{(n)}) \\
&= \sum_k \sum_c log\left[p(I_k|S_k = c, \theta)Pr(S_k = c)\right] p(S_k = c|I_k, \theta^{(n)}) \\
&= \sum_k \sum_c log\left[\mathcal{N}_{\mu_c, \Sigma_c}(I_k)Pr(S_k = c)\right] p(S_k = c|I_k, \theta^{(n)}). \quad (2.2.10)
\end{aligned}
$$

The best estimate for the model parameters θ can be found through explicit maximization, i.e. determining the μ_c and Σ_c values that satisfy the conditions $\frac{\partial Q(\theta|\theta^{(n)})}{\partial \mu_c} = 0$ and $\frac{\partial Q(\theta|\theta^{(n)})}{\partial \Sigma_c} = 0$. This leads to the following update equations for the M-step:

$$
w_{k,c}^{(n+1)} = \frac{p(I_k|S_k = c, \theta^{(n)})Pr(S_k)}{\sum_{c'} p(I_k|S_k = c', \theta^{(n)})Pr(S_k)} \quad (2.2.11)
$$

$$
\mu_c^{(n+1)} = \frac{\sum_k w_{k,c}^{(n+1)} I_k}{\sum_k w_{k,c}^{(n+1)}} \quad (2.2.12)
$$

$$
\Sigma_c^{(n+1)} = \frac{\sum_k w_{k,c}^{(n+1)} (I_k - \mu_c^{(n+1)})^t (I_k - \mu_c^{(n+1)})}{\sum_k w_{k,c}^{(n+1)}} \quad (2.2.13)
$$

Equation 2.2.11 describes the weight values for the expected likelihood in the E-step, $w_{k,c}^{(n)} = p(S_k = c|I, \theta^{(n)})$. The update equations for the image model with voxelwise-independent normal distributions show that model parameters for the next iteration are the the mean and covariance values of the image intensities, weighted by the normalized class posterior probabilities. In this example, image segmentation can be reduced to the following iterative steps:

(1) Estimating the initial model parameters $\theta^{(0)}$. This can be achieved through unsupervised clustering [**100, 106**] or using a spatial prior such as the brain atlas [**93, 17**].

(2) Computing the EM weights $w_{k,c}$ for each location k and class label c.

(3) Updating the mean and covariance parameter values using the observed image intensity values.

(4) Repeat the second and third step until convergence to a local maxima.

(5) Obtaining final segmentation labels from the class posterior probability. For discrete labels, the classification label $\mathcal{L}_k$ for location k is obtained as follows:

$$\mathcal{L}_k = \arg\max_c p(S_k = c | I_k, \theta^{(n_{final})}). \qquad (2.2.14)$$

The ML image segmentation framework does not yield a true Bayes estimator for the world state W. Although the ML framework makes use of the class prior probabilities $Pr(S_k)$, it does not include the modeling of $Pr(W) = Pr(S, \theta)$ or even $Pr(\theta)$. For the example framework described above, this is not a major issue since the model parameters θ are simply the means and covariances for the image likelihood described using normal distributions. However, when the model parameters θ become more sophisticated the use of prior knowledge of θ can improve the segmentation performance. Appendix B discusses a framework using a true Bayes estimator, where one can make use of the prior $Pr(W)$.

2.3. Robust Parameter Estimation

Estimating the model parameters θ in the M-step is a critical part of the EM algorithm. The original EM formulation computes the means and covariances of the Gaussian mixture model without taking account of possible outliers in the image data. I propose to improve the EM algorithm by using robust parameter estimation in the M-step in place of the standard approach. There has been some limited previous work for this type of approach, where robust estimation is combined with the EM algorithm. Malyutov and Lu [59] used the robust least median of squares and the M-estimator for estimating object trajectories from frame sequences that are corrupted with noise and clutter (correlated noise). With regard to medical image segmentation, van Leemput *et al.* applied the M-estimator for detecting brain lesions from MRI [95]. Their approach uses a multiplier

17

for the E-step weights, effectively deweighting based on intensity thresholding through the following equation:

$$q'(S_k) = \frac{p}{p + \mathcal{I}\{I_k > threshold\}\kappa} \, q(S_k) \tag{2.3.1}$$

where p is the Gaussian likelihood for healthy brain tissue, $\mathcal{I}$ is the indicator function, and κ is the class-specific deweighting parameter based on the covariance determinant. The intensity thresholding is done relative to the class intensity means. The means are computed as the weighted averages of the relevant image data values, which may be skewed by outliers. Equation 2.3.1 assigns lower probablity of healthy tissue in regions that fit the criteria of being abnormal (in this case lesions).

Estimators using weight functions such as the M-estimator typically cannot handle large proportion of outliers in data since they have a low breakdown point [12]. The breakdown point is defined as the fraction of data that must be moved to infinity (i.e., become outliers) for the method to generate inaccurate results. In fact, the classic un-bounded M-estimators has a breakdown point of 0 [86] since a single outlier with feature values that are significantly different from most of the data can lead to a local minima. In practice, the M-estimator typically have an approximate breakdown point of 0.15 to 0.2. The method proposed by van Leemput *et al.*, while performing well on brain lesion data, can fail when used to segment images with significant proportion of outliers due to malignant pathology or growth changes. In this dissertation, I propose the use of methods with a high breakdown point such as a combinatorial robust parameter estimator (Section 2.3.1) and a graph-based parameter estimator (Section 2.3.2). These methods isolate the clusters within data while taking account of outliers. They can be used in place of the standard M-step to generate parameters that are not unduly influenced by data outliers.

2.3.1. Minimum Covariance Determinant Estimator. The Minimum Covariance Determinant (MCD) estimator is an algorithm that computes the mean and covariance of a Gaussian model that represents an ellipsoid covering at least half of the data with

the lowest determinant of covariance. The generated estimate is highly robust with a high breakdown point. The MCD estimator has a breakdown point of 0.5, more than half of the data needs to be contaminated to make the results be unreasonable.

Rousseeuw and van Driesen [**80**] proposed a fast algorithm that computes an approximation of the MCD estimate. The algorithm first creates several initial subsets, where the elements are chosen randomly. From each subset, the algorithm determines different initial estimates of the robust mean and covariance. The estimates are then refined by performing a number of C-step operations on each initial selections. A single C-step operation consists of the following steps:

(1) Given a subset of the data, compute the mean and covariance of the elements in the subset.

(2) Compute Mahalanobis distances of the data elements in the whole set.

(3) Sort points based on distances, smallest to largest.

(4) Select a new subset where the distances are minimized (i.e. the first half of the whole set of the sorted data points).

An illustration of a single C-step iteration is shown in Figure 2.1. A C-step operation will result in a subset selection that yields a determinant of covariance less or equal to the one obtained from the previous subset. The iterative applications of C-steps yield final estimates with the smallest determinant of covariance. From all the final estimates computed with different initial selections, the mean and covariance estimate with the smallest determinant of covariance is chosen as the robust estimate.

2.3.2. Minimum Spanning Tree Clustering. The Minimum Spanning Tree (MST) clustering algorithm is a graph based techniques that yields a predefined number of data clusters while ignoring outliers. Unlike the MCD estimator that only provides a single data cluster, MST clustering can yield two or more data clusters.

Given a graph $G = (V, E)$ with vertices V and edges E, the Minimum Spanning Tree [**21**] is a graph $G_{MST} = (V', E')$ where all the vertices are connected such that the total
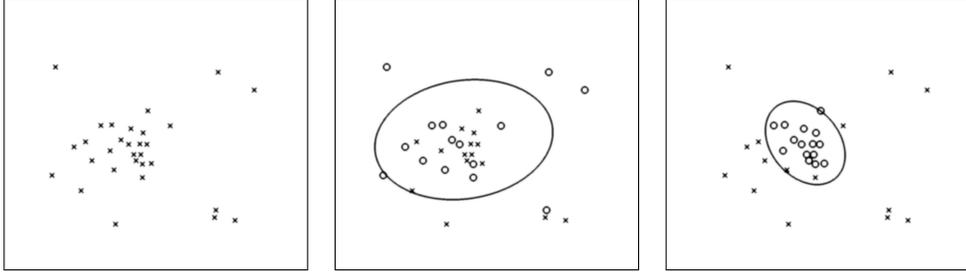
FIGURE 2.1. An illustration of a single C-step iteration, a key component of the MCD robust estimation algorithm. Left: original 2-D data with M points. Center: random selection of a subset of the data (marked with circles). Right: the selection after a C-step iteration, where the first $M/2$ closest points to the previous mean and covariance estimates are selected. The ellipsoidal curves in the center and right plots show the locations one standard deviation away from the mean and covariance estimates, which are computed from the selected points.

edge lengths are minimized. The MST graph is a subset of the original graph G, $V' \subseteq V$ and $E' \subseteq E$, where E' does not contain any closed loops (cycles).

The MST-based clustering technique proceeds by first creating the MST graph from the sampled data (image intensities) and then iteratively breaking the long edges to form connected clusters. At each iteration, it breaks an undirected edge $e(v, w)$ that connects vertices v and w if its length is larger than $A(v) \times T$ or $A(w) \times T$. $A(v)$ is the average length of edges incident on vertex v. Given $N(v)$ as the set of neighboring vertices for a vertex v, $A(v) = \frac{1}{|N(v)|} \sum_q |e(v, q)|$, where $q \in N(v)$. T is a scalar distance multiplier that determines which edges are considered to have significant deviation from the edges in the neighborhood.

Figure 2.2 shows the results of applying an iteration of the edge breaking step to an example dataset. As seen in the figure, the main clusters generated by the method do not contain the isolated data points that are located far away from the clusters. The edge breaking results in subtrees where each subtree forms an intensity cluster. The clustering algorithm terminates when it detects a predefined number of large clusters that satisfy certain intensity characteristics. For example, in the segmentation of an image into foreground and background regions, the algorithm should only terminate when it detects
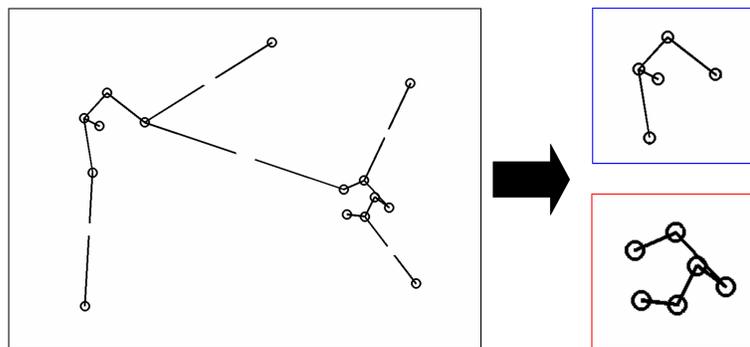
FIGURE 2.2. Clustering by breaking the long edges of a Minimum Spanning Tree. Left: the MST created from the input data, with the long edges broken. Right: the generated clusters, note that the isolated far away points are treated as outliers and are not included.

at least two large data clusters where one of them is very dark and another is relatively bright.

The MST clustering algorithm uses the local property $A(v)$ for a vertex v to determine the clusters, as opposed to the combinatorial MCD scheme which involves large subsets of data. This can cause problems when the global configuration of the data in feature space is not optimal. The breakdown point can be as low as zero when the actual data samples and the outlier samples are connected by a series of short, equal-length edges. Figure 2.3 shows an example of a configuration that leads to a low breakdown point. This configuration will cause the algorithm to fail in detecting the separation between data and outliers. However, this type of configuration is not generic and rarely occurs with real data. Some heuristics can also be applied to avoid having a sequence of short edges, such as the one presented in Chapter 4. Assuming typical sample configurations for the input and picking the largest data cluster as inliers, the MST clustering algorithm can have a breakdown point as high as 0.5.

Details on the properties of the MST graph structure can be found in the book by Cormen *et al.* [**21**]. Applications of graph-based clustering for pattern recogniction is described in the book by Duda *et al.* [**27**]. Of relevance to this dissertation, a description
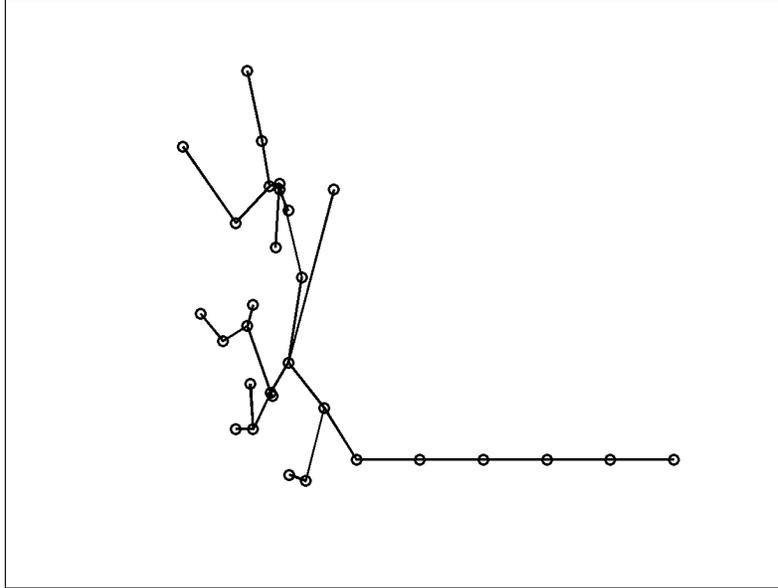
FIGURE 2.3. An example configuration where the Minimum Spanning Tree clustering cannot properly separate outliers from data. The outlier points on the lower right are connected by relatively short and equal-length edges. These edges will only be broken when some of the edges connecting the inlier samples are also broken, which leads to a skewed estimate of the inlier distribution.

of the robust MST clustering method applied to the segmentation of medical images can be found in the paper by Cocosco *et al.* [**17**].

2.4. Robust EM Segmentation Framework

The generative model or the image likelihood $p(I|W)$ is an important component for the EM based approach. However, in practice it can be difficult to obtain the generative model from the data, particularly in cases where there are unknown deviations from prior knowledge. This problem is compounded by the high level of variability in image acquisitions and the amount of noise in the image data. Therefore, it is crucial to have the ability to draw a reasonable estimate of the generative model from noisy data that may have outliers. The computation of the parameter θ that forms the image likelihood in the standard M-step does not take outliers into account. When outliers are present, this presents a serious problem since they may skew the parameter estimation results and generate improper segmentations.

I propose the modification of the EM algorithm into a new robust EM algorithm, where the parameter update in the M-step is replaced with a robust parameter estimator described in the previous section. This makes it suitable for detecting anatomical deviations in medical images, as they can alter significant proportions of the image data. Since real images can have specific and complex appearance properties, the Gaussian mixture model in the EM image segmentation may not always be appropriate. To improve segmentation results in the general case, I propose the use of non-parametric density functions for the image likelihood. Instead of computing the Gaussian distributions for $p(I|W)$, the algorithm computes non-parametric kernel density estimates.

In the M-step, the robust data driven parameter estimation methods (the Minimum Covariance Determinant estimator and the Minimum Spanning Tree clustering) are used to generate data clusters together with the knowledge of data outliers. The knowledge of the relevant clusters allows the method to robustly determine the proper intensity ranges for the major structures. The generative model $p(I|W)$ is estimated by fitting a probability distribution to the data clusters, performed by using the data within each cluster as training data for the nonparametric kernel densities.

The new robust EM algorithm is an iterative process that replaces the parameter update equations (Equations 2.2.12 and 2.2.13) with the combination of robust data clustering and non-parametric density estimation. The modified algorithm proceeds as follows in the nth iteration:

- E-step: This part of the EM algorithm is unchanged. The goodness-of-fit function Q is computed according to Equation 2.2.6 using the class posteriors $p(S|I, \theta^{(n)})$ to form the expectation.
- M-step: Computation the updated parameters $\theta^{(n+1)}$ is done through clustering of image data samples. The samples are obtained using the parameters from the current iteration n, which determines the class posteriors $p(S|I, \theta^{(n)})$. The class posteriors describe the likely regions for obtaining data samples. Good regions

23

for sample selection are determined by thresholding the current class posteriors or using a Monte Carlo approach (e.g., the Metropolis-Hastings algorithm).

Once the clusters are identified, the individual mixture components $p(I|S = c, \theta)$ are computed by fitting a kernel density function to the intensity data within the clusters. The samples form the training data for classification using non-parametric density mixtures. For a class c and voxel location k the updated voxel image likelihood becomes

$$p(I_k|S_k = c, \theta^{(n+1)}) = \frac{1}{M} \sum_{i=1}^{M} K_\lambda(I_k - Y_i^{(n)}) \qquad (2.4.1)$$

where K_λ is the multivariate Gaussian kernel with standard deviation λ and zero mean, and $Y^{(n)}$ is the set of M training samples obtained by applying robust clustering to samples from $p(S|I, \theta^{(n)})$. The choice of the standard deviation or kernel width λ is crucial in determining the classification decision boundaries. The kernel width can be estimated through heuristics, for example by choosing a fraction of the expected image intensity range. The kernel width can also be estimated using a bootstrap technique.

Initialization of the robust EM algorithm is performed by using the spatial priors provided by the atlas, $p(S_k = c|I_k, \theta^{(0)}) = Pr(S_k = c)$. Each iteration of the robust algorithm refines the class posterior probability $p(S_k = c|I_k, \theta^{(n)})$ and tends to result in sharper spatial posterior probabilities.

The modification of the M-step is mainly done by excluding or deweighting samples that do not follow the expected data distribution in the image intensity feature space. In addition to explicitly identifying outliers in the feature space, another approach is to identify samples that violate prior knowledge on the spatial or geometric features. This approach is used in brain tumor segmentation to isolate tumor samples by removing small spurious regions and to isolate edema samples by constraining it to regions adjacent to tumor (Section 3.2). In newborn brain segmentation, the method avoids white matter samples near the decision boundary by restricting sampling to regions with low

image gradient magnitude values (Section 4.2). This approach samples smooth image regions, which results in coherent data clusters that are easier to detect and to separate. Using knowledge on the geometric configuration within the images avoid contaminating the samples with ambiguous data or data which induces a certain bias in parameter estimation.

The classic EM algorithm with mixture models restricts the number of mixture components to a predetermined value. This model is inappropriate for many real images, as objects may enter and leave the scene or pathological structures can be formed or can disappear, for example. I propose a modification of the classic EM algorithm where new mixture components are detected through explicit detection of multiple clusters within the data. The detection of multiple clusters can be achieved in a robust manner that discards outliers by using the robust MST clustering proposed in the previous section. The existence of new mixture components can be tested by evaluating the amount of overlap between the detected clusters from the MST graph. This detection of new mixture components is applied to the determination of possible edema surrounding tumor in Section 3.2.

The robust image segmentation framework includes modifications that make it no longer a true EM algorithm. The log likelihood Q is no longer guaranteed to increase at every iteration. Therefore, the iterative process is terminated when the change is smaller than a threshold:

$$|\frac{Q(\theta^{(n+1)}|\theta^n) - Q(\theta^n|\theta^{(n-1)})}{Q(\theta^n|\theta^{(n-1)})}| < \epsilon_Q \qquad (2.4.2)$$

The robust EM image segmentation framework performs classification mainly based on the image intensity feature space data. If available, the spatial information from the brain atlas prior $Pr(S)$ is combined with the class likelihood using the Bayes rule. The spatial priors $Pr(S)$ may be suboptimal since there may exist some deviations between the data and the prior model. In cases with small to moderate deviations, classification using the intensity feature space can still be reliable due to the use of the robust parameter estimation techniques. However, in cases with large deviations, the use of the spatial prior

$Pr(S)$ may result in samples that have a significant proportion of outliers. When the proportion of outliers exceed the breakdown point of the robust parameter estimator, the algorithm would fail to estimate the proper image likelihood and the segmentation would fail. To deal with this issue, it will be necessary to explicitly account for possible changes to $Pr(S)$ to form an appropriate model for the observed images. As an example, spatial priors coded in the healthy adult atlas need to account for brain tumors that can deform surrounding structures.

Another limitation of the algorithm is related to the voxel-based approach, where it is assumed that each voxel is independent to other voxels. This approach can lead to spurious segmentations where small collections of voxels within some regions are incorrectly labeled as distinct from the surrounding labels. Noisy results can be avoided by extending the framework via the application of a Markov Random Field (MRF) model to $Pr(S_k)$, similar to the approach proposed by van Leemput *et al.* [**93**] and Zhang *et al.* [**106**]. However, the MRF model tends to smoothen the segmented structures and thus may not be appropriate for some applications where we require segmentations with fine details. An alternative to the MRF approach would be the use of an extended image model with a modeling of image region coherence and the more complete prior $Pr(S, \theta)$ rather than only $Pr(S)$. Such an extended model is proposed in a Maximum a Posteriori (MAP) image segmentation framework discussed in Appendix B.

Applications of the new robust EM image segmentation framework are described in Chapter 3 for brain tumors and Chapter 4 for newborn brains. Possible extensions to this framework are discussed in the future work section of Chapter 6.

CHAPTER 3

Brain Tumor MRI Segmentation

In this chapter, I will describe the application of the robust EM framework described in Chapter 2 to automatic brain tumor segmentation from MRI. Section 3.1 provides a discussion of the motivation and challenges in segmenting brain tumor MR images. The adaption of the robust EM image segmentation framework to accomodate pathology is described in Section 3.2. Deviations in intensity and appearance due to tumor and edema are detected as outliers from the expected image model of healthy tissue. Existence of edema is determined automatically by testing for a bimodal distribution of the outlier samples. In Section 3.3, I present the results of the automatic segmentation framework applied to three different types of tumor, along with the validation of those results compared to manual segmentations.

3.1. Background

The segmentation of brain tumor and edema from MR images is of particular interest to clinicians. The knowledge of tumor and edema extent combined with the location of healthy structures can be used to provide a reference for surgical and radiological treatment planning, for monitoring tumor growth, and for evaluating efficacy of treatment. However, the presence of brain tumor and edema in a brain MRI gives rise to many issues and challenges related to quantitative analysis of such images. Brain tumors may be of any size, may have a variety of shapes, may appear at any location, and may appear in different image intensities. Some tumors also deform other structures and appear together with edema that changes intensity properties of the nearby region.

For many human experts, manual segmentation of brain tumor from MRI is a difficult and time consuming task. Therefore, an automated brain tumor segmentation method is desirable. There are many potential applications of automated segmentation. The changes of healthy tissue due to tumor and edema need to be identified for determining surgical and radiological treatment planning. The knowledge of affected regions is also vital for studying possible diagnostic markers for brain tumors. For example, the shape of blood vessels within tumor regions is observed to have significant correlation with tumor classification into benign and malignant types [10].

The multiple challenges associated with automatic brain tumor segmentation have given rise to many different approaches. Automated segmentation methods that combine fuzzy clustering and knowledge-based classification were proposed in [13, 31]. The two methods do not rely on intensity enhancements provided by the use of contrast agents. A particular limitation of the two methods is the use of specific classification rules which restricts the image modalities to the T1, T2, and PD weighted MR image channels. Additionally, the methods require a manually guided training phase prior to segmenting a set of images. Other methods are based on statistical pattern recognition techniques, for example the method proposed by Kaus *et al.* [48]. This method combines the information from a registered atlas template and user input to supervise training of a the classifier, demonstrating the strength of combining voxel-intensity with geometric brain atlas information. This method was validated against meningiomas and low-grade gliomas. Gering *et al.* [33] proposed a method that detects deviations from normal brains using a multi-layer Markov random field framework. The information layers include voxel intensities, structural coherence, spatial locations, and user input. Cuadra *et al.* presented high-dimensional warping to study deformation of brain tissue due to tumor growth [22]. Their technique relies on a prior definition of the tumor boundary whereas the method proposed in this chapter focuses on automatically finding tumor regions.

Previous work on automatic brain tumor segmentation generally uses the enhancement provided by the gadolinium contrast agent in the T1 channel or constrained to

28

blobby shaped tumors with uniform intensity. Even though the intensity enhancement can aid the segmentation process, it is not always necessary to obtain good results. In fact, requiring the use of contrast enhancement to segment tumors can be problematic. Typically, tumors are only partially enhanced and some tumors are not enhanced at all. Blood vessels also generally appear enhanced by the contrast agent. These inconsistencies create an ambiguity in the image interpretation, which makes the T1-enhanced image channel a less than ideal feature for tumor segmentation in general. Methods that primarily use the contrast enhancement to drive the tumor segmentation will have difficulties in isolating tumors.

Edema surrounding tumors and infiltrating mostly white matter was most often not considered as important for tumor segmentation. It has been shown previously [**71, 77**] that edema can be segmented using a prior for edema intensity and restriction to the white matter region. The extraction of the edema region is essential for diagnosis, therapy planning, and surgery. It is also essential for efforts that involve modeling the brain deformation due to tumor growth. The swelling produced by infiltrating edema usually has distinctly different tissue property characteristics than space occupying tumor. The segmentation strategy presented in this chapter is based on the detection of *changes from normal* and will thus systematically include segmentation of edema. Differential identification of the two abnormal regions tumor and edema is clinically highly relevant. Even though the primary therapeutic focus will be on the tumor region, the edema region may require secondary analysis and treatment.

The proposed method combines a model of normal tissues and the geometric and spatial model of tumor and edema. It relies on the information provided in the T2 weighted image channel for identifying edema, and it can make use of additional image channels to aid the segmentation. Here, I have chosen to use only the T1 and T2 image channels. Tumor and edema are treated as intensity abnormalities or outliers. After identifying the abnormalities, an unsupervised clustering technique is applied to the intensity features before utilizing geometric and spatial constraints. I will show that this
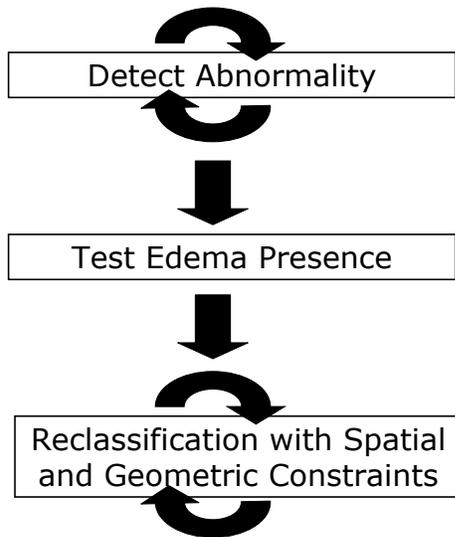
FIGURE 3.1. The three major stages of the brain tumor MRI segmentation method.

method can segment tumors with or without intensity enhancements and automatically detects the presence of edema. This approach offers a means of approaching lesions of multiple types and of different image intensities, and, with a single method, lesions that enhance or do not, and that may or may not be surrounded by edema.

3.2. Method

The automated segmentation method for brain tumor MRI is composed of three major stages, as shown in Figure 3.1. First, it detects abnormal regions, where the intensity characteristics deviate from the expectation. In the second stage, it determines whether these regions are composed of both tumor and edema. Finally, once the estimates for tumor and edema intensity parameters are obtained, the spatial and geometric properties are used for determining proper sample locations. The details of each stage are discussed in the following subsections.

3.2.1. Detection of Abnormality. Before identifying tumor and edema, it is necessary to first detect regions that have properties that deviate from the expected properties of a normal, healthy brain. This involves finding the intensity parameters for healthy classes and the abnormal class. The initial parameters for the healthy brain classes are
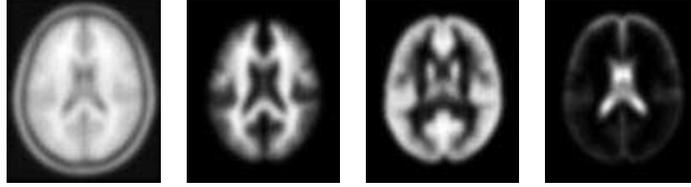
FIGURE 3.2. The brain atlas that acts as the model of the healthy adult population, provided by the International Consortium for Brain Mapping (ICBM). From left to right: the T1 template image and probability values of white matter, gray matter, and cerebrospinal fluid. The atlas does not account for brain tumor and edema, thus its use for segmentation of images presenting pathology requires a new approach.

obtained by sampling specific regions based on the probabilistic brain atlas for healthy adults shown in Figure 3.2 [28].

The atlas is aligned with the subject image data by registering the atlas template image with the subject image. The registration is performed using affine transformation with the mutual information image match measure [58]. After alignment, the samples for each healthy class (white matter, gray matter, and cerebrospinal fluid) are obtained by randomly selecting the voxels with high atlas probability values. The set of training samples is constrained to be the voxels with probabilities higher than a threshold $\tau = 0.85$ [17].

The training data for the healthy classes generally contain unwanted samples due to contamination with samples from other tissue types, particularly tumor and edema. The pathological regions are not accounted for in the brain atlas and they therefore occupy regions that are marked as healthy. The *contaminants are considered data outliers*, and they need to be removed so that the training samples for the healthy classes are representative. The samples are known to be contaminated if their characteristics differ from prior knowledge. The intensities for healthy classes are known to be well clustered and can be well approximated using Gaussians (Figure 3.3).

Handling data outliers is a crucial step for atlas based image segmentation. Cocosco *et al.* [17] developed a segmentation method for healthy brains that builds the Minimum Spanning Tree from the training samples and iteratively breaks the edges to remove false
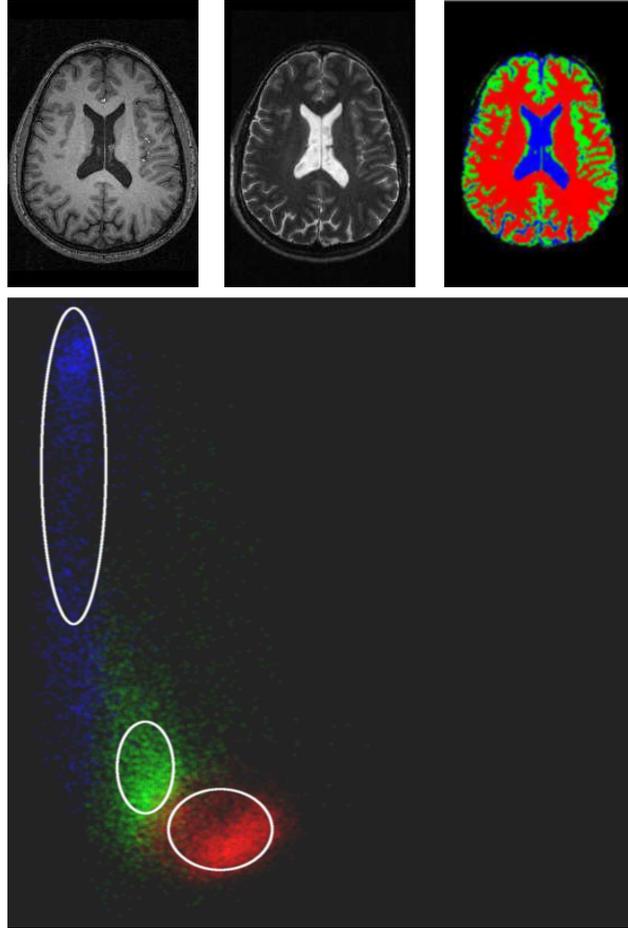
FIGURE 3.3. Example healthy adult MRI dataset. Top, from left to right: T1 image, T2 image, and segmentation labels (red is white matter, green is gray matter, and blue is cerebrospinal fluid). Bottom: the intensity histogram for the three classes, the horizontal axis represents T1 intensities and the vertical axis represents T2 intensities. The intensity features for each class is tightly clustered and can be approximated with a Gaussian.

positives (pruning). They showed that pruning the training samples results in significant improvement of the segmentation quality, particularly for image data presenting enlarged ventricles. In my method, robust estimate of the mean and covariance of the training data is used to determine the outliers to be removed.

The robust estimator used for detecting abnormality is the Minimum Covariance Determinant (MCD) estimator [80] discussed in Section 2.3.1. Given the robust mean and covariance from the MCD estimator, image intensity samples that are further than three standard deviations are considered as outliers. This outlier removal process is
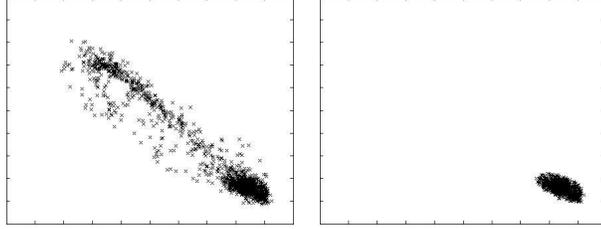
FIGURE 3.4. The white matter training data for a subject with tumor and edema, the horizontal axis represents the T1 intensities and the vertical axis represents the T2 intensities. Left: original samples obtained by atlas-guided sampling which is contaminated with samples from other distributions. Right: remaining samples after trimming using the robust MCD estimate, representing the feature distribution of healthy white matter.

shown in Figure 3.4 for white matter samples. The inliers of the healthy brain tissue class samples are used as training samples for estimating the corresponding density functions.

The specific aim at this stage is to compute the density estimates and posterior probabilities for the set of class labels $\mathcal{C} = \{$white matter, gray matter, csf, abnormal, non-brain$\}$. A parametric density function is not ideal for the case of tumor segmentation. Tumors do not always appear with uniform intensities, particularly in the case where some tissues inside the tumor are necrotic tissues. Therefore, no assumption can be made regarding the intensity distributions and thus I use a non-parametric model for the probability density functions. The density functions are approximated using kernel expansion or Parzen windowing [27]. Given the vector of intensity features I_k at location k, the probability density function on intensity for class label $S_k = c$ is

$$p(I_k|S_k = c, \theta = (\lambda, Y)) = \frac{1}{N} \sum_{i=1}^{N} K_\lambda(I_k - Y_i) \tag{3.2.1}$$

where K_λ is the multivariate Gaussian kernel with standard deviation λ, and Y is the set of class training samples. The kernel bandwidth λ chosen is 4% of the intensity range for each channel, determined using empirical tests on multiple images.

The posterior probability at location k is computed using the class prior probability from the atlas $Pr(S_k)$ by applying the Bayes rule

$$p(S_k|I_k, \theta) = \frac{p(I_k|S_k, \theta)Pr(S_k)}{p(I_k)}. \qquad (3.2.2)$$

The spatial priors for white matter, gray matter, csf, and non-brain classes are the corresponding atlas probabilities. For the abnormal class, a fraction of the sum of white matter and gray matter atlas probabilities is used since tumor and edema usually appear in these regions and not in the csf regions. With the kernel density estimate as the likelihood, the image appearance parameters θ for tumor segmentation is composed of the set of training samples for each class and the kernel width.

An issue with MR images is the presence of the image inhomogeneity or the bias field. This is dealt by interleaving the segmentation process with bias correction, following the spirit of [**100**]. The entire process of detecting the abnormal regions is shown in Figure 3.5, a loop that is composed of the following five stages:

(1) Threshold the class posterior probabilities and sample the high confidence regions. The posterior probabilities are initialized using the prior probabilities obtained from the brain atlas.

(2) Remove the samples for normal tissues that exceed a distance threshold based on the MCD estimate.

(3) Estimate the non-parametric density for each class likelihood using kernel expansions. The initial density for the abnormal class is set to be uniform, which makes this class act as a rejection class. The brain voxels with intensity features that are different from those of healthy classes or not located in the expected spatial coordinates will be assigned to this class.

(4) Update the class posterior probabilities using the new class likelihoods.

(5) Estimate bias field from white matter and gray matter probabilities. Apply correction using the estimated bias field.
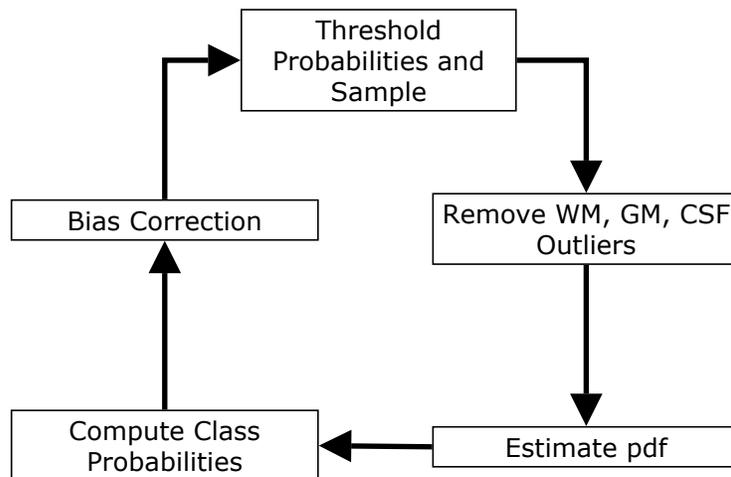
FIGURE 3.5. The process of detecting abnormal regions, the first stage of the brain tumor MRI segmentation method.

| Initialization | Iteration 1 | Iteration 2 | Iteration 3 |
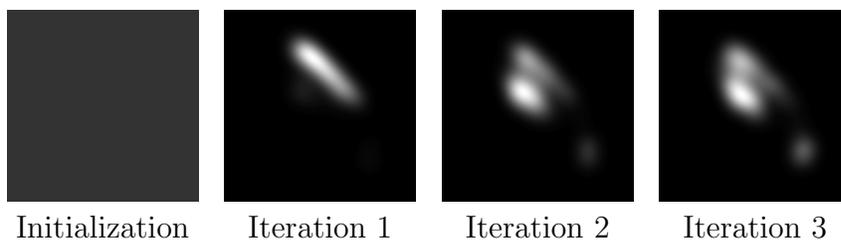
FIGURE 3.6. Snapshots of the estimated probability density function of the abnormal class for the Tumor020 data. Each image shows the result of different iterations of the loop shown in the previous figure. The density is initialized so that all intensities are equally likely. The horizontal axis represents the T1 intensities and the vertical axis represents the T2 intensities. The two high density regions visible at the final iteration are the tumor and edema densities, which have a significant separation along the dimension of the T2 intensities.

The first major segmentation stage detects the abnormal regions by executing the loop for several iterations, obtaining the intensity descriptions for each class. The abnormal class density at different iterations for the Tumor020 data is shown in Figure 3.6.

The bias correction method is based on the one developed by van Leemput *et al.* [**92**]. The method uses the posterior probabilities to estimate the homogeneous image. It then computes the bias field estimate, as the log-difference between the homogeneous images and the real subject images. The bias field is modeled as a polynomial, and the

coefficients of the polynomial is determined through least squares fitting. The method assumes that the class intensity distributions are approximately Gaussians. Only the white matter and gray matter probabilities are used for estimating the parameters for the bias field correction, as they generally can be approximated by Gaussians without significant errors. Additionally, the combination of white and gray matter probabilities represents a large connected region covering the major part of the brain.

3.2.2. Tumor and Edema Separation. The densities and posterior probabilities computed for the abnormal class in the previous stage give us a rough estimate of how likely it is that some voxels are part of tumor or edema. I make the assumption that the detected abnormal voxels are composed mostly of tumor and possibly edema. Edema is not always present when tumor is present, therefore it is necessary to specifically test the presence of edema. This is done by first obtaining the intensity samples for the abnormal region, which is performed by thresholding the posterior probabilities and selecting a subset of the high probability regions. The samples are then clustered and a test is done to determine whether there exist separate clusters for tumor and edema. The density estimate for tumor (and edema, if present) is obtained by performing kernel expansion on the samples.

Tumor and edema are generally separable given the information in the T2 weighted image (Figures 3.6 and 3.7). Edema has high fluid content and therefore appears brighter than tumor in this image channel. To separate the densities, unsupervised clustering is applied to the samples obtained by thresholding. The method I have chosen is k-means clustering with $k = 2$ [**27**]. For dealing with outliers, the robust MST clustering described in Section 2.3 can also be used as an alternative. Once the clusters are obtained, the tumor cluster can be identified as the cluster with the lower T2 mean value, making use of prior domain knowledge.

To determine the validity of the clustering, the method tests for the overlap using the Davies-Bouldin index [**24**]. This measure is the ratio of the average within cluster distances and the between cluster distance. Given m candidate tumor samples τ_i with
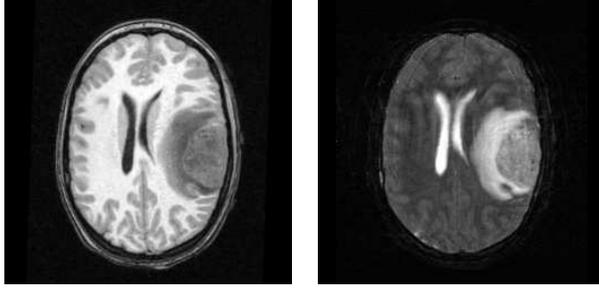
FIGURE 3.7. The T1 image (left) and the T2 image (right) from the Tumor020 data. The tumor and edema on the right part of the brain can be clearly differentiated based on the T2 intensities. As observed in the T2 image, the tumor region (rightmost) is darker than the surrounding edema region, as edema is composed mostly of fluid.

the mean value μ_{tumor}, and n candidate edema samples ϵ_i with the mean value μ_{edema}, the overlap measure is:

$$\frac{1}{2}\left(\frac{\frac{1}{m}\sum_{i=1}^{m}||\tau_i - \mu_{tumor}|| + \frac{1}{n}\sum_{i=1}^{n}||\epsilon_i - \mu_{edema}||}{||\mu_{tumor} - \mu_{edema}||}\right) \qquad (3.2.3)$$

The T2 channel contains most of the information needed for differentiating tumor and edema. Therefore, the overlap is measured only for the T2 data dimension of each cluster. If the amount of overlap is larger than a specified threshold, then the tumor density is set to be the density for the abnormal class and the edema density is set to zero.

3.2.3. Application of Spatial and Geometric Constraints. Once this stage is reached, tumor and edema are already segmented based on atlas priors and intensity characteristics. However, voxel-based processing does not consider geometric and spatial properties and this generally leads to noisy segmentation results. Since there is no model for the intensity distributions of tumor and edema, it is necessary to use geometric and spatial heuristics to prune the samples that are used for estimating the densities. The prior knowledge used in this stage is the fact that tumor is mostly blobby. For edema, the applied constraint is that each edema region needs to be connected to a nearby tumor region. Some edema voxels can be located far away from tumor regions, but they must be spatially connected to a tumor region.

Tumor structures generally appear as blobby lumps, and this shape constraint is enforced through region competition snakes [**84, 88, 89, 108**]. The tumor posterior probabilities are used as the input for the snake, which is represented as the zero level set of the implicit function ϕ. The level set evolution is governed by the following equation [**40**]:

$$\frac{\partial \phi}{\partial t} = \alpha(p(S_k = tumor|I_k, \theta) - p(S_k = \overline{tumor}|I_k, \theta))|\nabla\phi| + \beta\nabla \cdot \left(\frac{\nabla\phi}{|\nabla\phi|}\right)|\nabla\phi| \quad (3.2.4)$$

The propagation term is represented by α. It is modulated by the difference of the posterior probabilities for the tumor class and the non-tumor class ($p(S_k = tumor|I_k, \theta)$ and $p(\overline{tumor}|I_k, \theta)$), so that the direction of the propagation is determined by the sign of the difference. The probability that a voxel is part of brain and not part of tumor is represented by $p(\overline{tumor}|I_k, \theta)$, more explicitly:

$$
\begin{aligned}
p(\overline{tumor}|I_k, \theta) &= p(\{white\ matter\}|I_k, \theta) + p(\{gray\ matter\}|I_k, \theta) + p(\{csf\}|I_k, \theta) \\
&+ p(\{edema\}|I_k, \theta) \quad (3.2.5)
\end{aligned}
$$

The snake shrinks when the boundary encloses part of the regions not part of tumor and expands when the boundary is inside the tumor region. Smoothing on the snake contour is applied using mean curvature flow, and the strength of this smoothing is controlled by the β term. The initial level set function is obtained by performing a signed distance transform on the segmented tumor objects.

Edema, if present, is always contiguous with the tumor. With this prior knowledge, it is therefore assumed that edema is located near tumor structures. Each segmented edema object must have a voxel that is adjacent and no further than some small distance from tumor regions. Adjacency can be tested efficiently by dilating the neighboring structures and by determining if intersections exist. First, the method generates a binary image representing the segmented edema region. Then, this image is used as an input for the connected component algorithm to determine the individual edema objects. Each object is dilated and then compared against the segmented tumor regions. The objects that

share some voxels with a tumor region are considered valid. Edema samples from these regions are kept, while other edema samples are discarded.

The final segmentation is obtained by reclassifying the image using the iterative steps similar to the one described in Section 3.2.1, with some modifications (Figure 3.8). The outlier removal stage is removed and there are additional steps where these geometric and spatial constraints are enforced. The entire loop is performed several times, after going through each loop the tumor and edema probabilities at the voxel locations that do not pass the tests are set to zero. This way, the segmentation for these locations are determined based on the next best candidate class. The tumor shape constraint is disabled at the last fitting stage. This is done to obtain the proper boundary for the tumor structures, which may not be entirely smooth. For instance, gliomas typically have a general blobby shape and ragged boundaries.

The application of geometric and spatial constraints modifies the M-step of the standard EM algorithm so that it ignores the data samples obtained from inappropriate locations. This is a geometric-based approach to robust parameter estimation, where we make use of prior knowledge of the application domain with regard to spatial and geometric properties. This modification makes sure that the method can exclude subtle outliers (outliers located close to actual data clusters) by using the augmented features that include location and shape features.
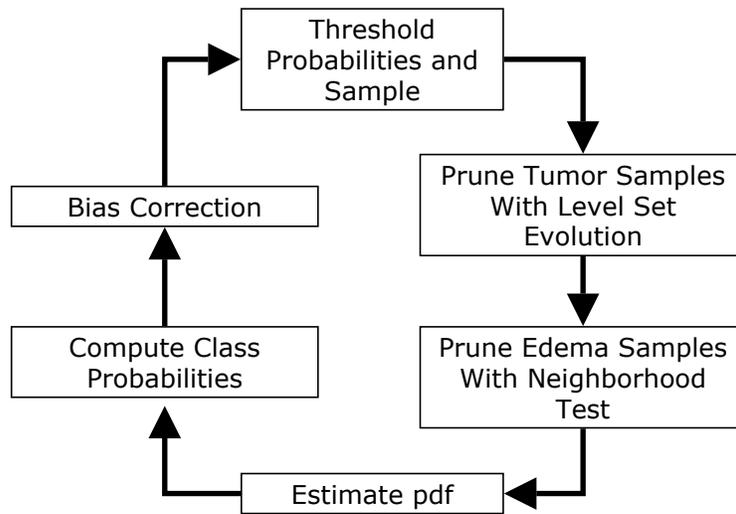
FIGURE 3.8. The third stage of the method where the image is reclassified using tumor geometric properties and edema spatial relation.

3.3. Results and Validation

Validation on the brain tumor segmentation method described in this chapter has been performed on three real datasets. The datasets along with the results of the automatic segmentation method are shown in Figure 3.9. Each dataset represents a different tumor shape, location, size, image intensity or appearance, and enhancement. Tumor020 has a partially enhancing tumor that causes a large deformation of the normal structures. Tumor025 contains a large, partially enhancing tumor inside the brain stem. Tumor033 contains a low grade tumor which is not highlighted in the contrast enhanced T1w channel.

For validation purposes, two sets of segmentations are done manually by one human rater at different times. The volumes of the manually segmented structures are shown in Table 3.1. The first set of manual segmentations is considered to be the gold standard for validating the automatic segmentation method. The VALMET segmentation validation tool [**32**] is used to generate three validation metrics (described in Appendix A). The first measure is a volume overlap measure, the Jaccard similarity coefficient. The other metrics are the symmetric Hausdorff surface distance and the average surface distance.

The intra-rater variability is shown in Table 3.2. The surface distance values indicate that the manual segmentations considered as reliable (a real test would need statistical testing). The Jaccard Similarity Coefficient (JSC) values that measure volumetric overlap are also high, with the exception of the Tumor033 segmentation. This is likely due to the small size of the tumor. The quantitative validation of the automatic segmentation method is shown in Table 3.3. The level of agreement based on surface distances is similar for all tumors. However, the varying overlap values demonstrate that the overlap metric is sensitive to the size and complexity of the segmented objects (Figure 3.9), a fact that is not sufficiently discussed in existing literature. The level of agreement with the manual result for edema is lower than that for tumor. This is mainly due to the ambiguity in determining the edema boundary, especially the tumor-edema boundary. The intra-rater reliability (Table 3.2) is higher than inter-method agreement (Table 3.3).

41

| | T1 | T2 | Tumor | Edema | 3D View |
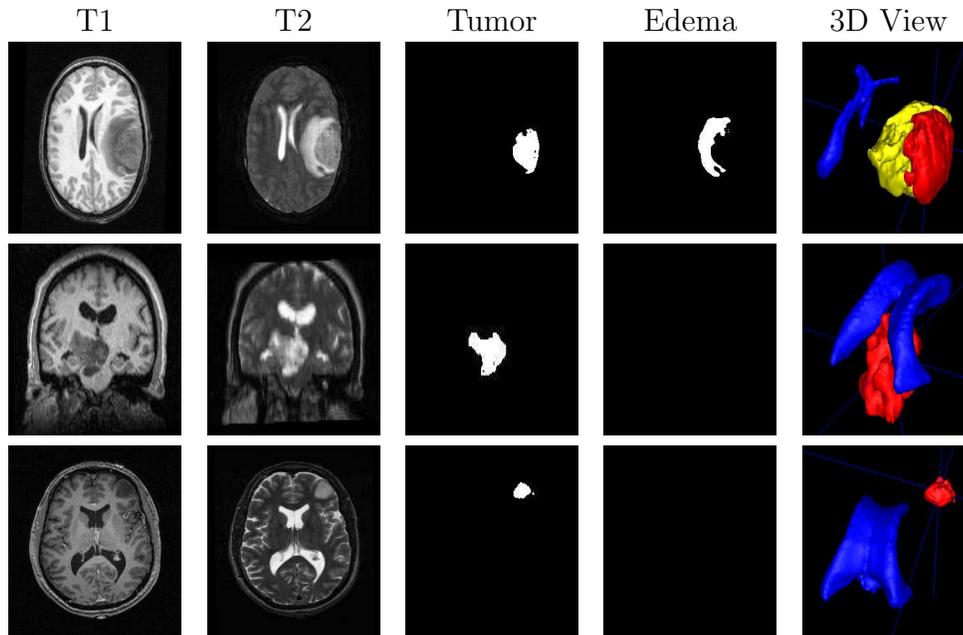
FIGURE 3.9. The datasets and the generated segmentation results. The last column shows the 3D views of the segmented structures: red represents tumor, yellow represents edema, and blue represents ventricles. From top to bottom: Tumor020, Tumor025, Tumor033. These results illustrate that the proposed method provides the differential segmentation for tumor and edema, which works also in cases where no edema is present.

Dataset	Tissue Type	Volume (mm³)
Tumor020	Tumor	35578.6
Tumor020	Edema	64860.6
Tumor025	Tumor	24742.4
Tumor033	Tumor	3661.5

TABLE 3.1. Volumes of the segmented tumor and edema structures, from the results of the first set of manual segmentation results.

However, it would be interesting to compare automatic segmentation to segmentations of multiple experts, which is often shown to yield significant differences.

Dataset	Tissue Type	JSC (%)	Symm Hausdorff (mm)	Average distance (mm)
Tumor020	Tumor	89.0	3.98	0.54
Tumor020	Edema	75.5	13.1	0.75
Tumor025	Tumor	81.2	4.1	0.73
Tumor033	Tumor	59.4	5.22	1.51

TABLE 3.2. Validation metrics comparing the two sets of manual segmentation results done by the same human rater, demonstrating the intra-rater variability of the manual segmentations.

Dataset	Tissue Type	JSC (%)	Symm Hausdorff (mm)	Average distance (mm)
Tumor020	Tumor	80.0	16.79	1.64
Tumor020	Edema	68.2	12.80	1.75
Tumor025	Tumor	79.2	17.85	1.44
Tumor033	Tumor	70.6	8.60	1.85

TABLE 3.3. Validation measures of the automatic tumor segmentation results against the first set of manual segmentation results.

3.4. Conclusions

This chapter presented a new pproach for automatic segmentation of brain tumors and adjoining edema from non-enhancing multichannel MRI. Most methods so far have been applicable only to enhancing tumors with homogeneous appearance. Furthermore, they require user-guidance in training a supervised classifier or to obtain a rough outline of the region of interest. The presented technique automatically identifies the presence of edema. This is a clinically relevant feature as the edema region often requires secondary treatment and analysis after a primary focus to the treatment region.

The proposed automatic segmentation framework uses a concept that detects difference from normal and uses non-parametric kernel densities in place of the traditional Gaussian mixture model. In addition to the explicit detection of intensity outliers, the method uses geometric constraints such as the shape of the brain tumor and the location of edema. The segmentation results of three tumor datasets with different appearances illustrate that using robust parameter estimation in an EM framework might be a promising new approach for segmenting brains with pathological deviations. To my knowledge, this is the the first fully automatic segmentation system for whole brain tissue, tumor, and edema.

The brain atlas is a vital component of the proposed method, where it acts as a spatial prior and sampling constraint. A potential issue that is not handled by the proposed method is the large deformation of brain structures. Severe deformations can increase the number of intensity outliers and can exceed the breakdown point of the robust MCD parameter estimator. A further discussion on other potential extensions for the method can be found in Chapter 6.

CHAPTER 4

Newborn Brain MRI Segmentation

This chapter describes an adaptation of the robust EM framework described in Chapter 2 for automatic segmentation of newborn (neonatal) brains from MRI. The motivation and challenges for automatic segmentation of newborn (neonatal) brains are presented in Section 4.1. Section 4.2 provides a detailed description of the modified segmentation framework for this task. The framework makes use of the graph-based robust parameter estimation technique (Section 2.3) to determine relevant clusters that form the optimal Gaussian mixtures. Finally, segmentation results of five different cases are presented in Section 4.3, including limited validation on selected 2D slices.

4.1. Background

The segmentation of newborn brain structures from magnetic resonance images (MRI) is crucial for the study of normal development and comparison to neurodevelopmental disorders at early stages. The development of new segmentation methods for this age group is driven by the increasing use of MRI to study newborns and infants, for example the ongoing studies of early brain development in normal and high risk children at UNC [**105, 34**] and the lack of appropriate segmentation methodology. Manual segmentation of newborn brains is tedious, time consuming, lacks reproducibility, and limits clinical studies to very small number of subjects. Therefore, it is necessary to use automatic segmentation methods for clinical studies with a large population size, which is required to obtain statistical significance. Identification of the growth patterns during the critical natural development phase may yield vital clues about the origin and nature of neurodevelopmental diseases. This task is considerably more challenging compared
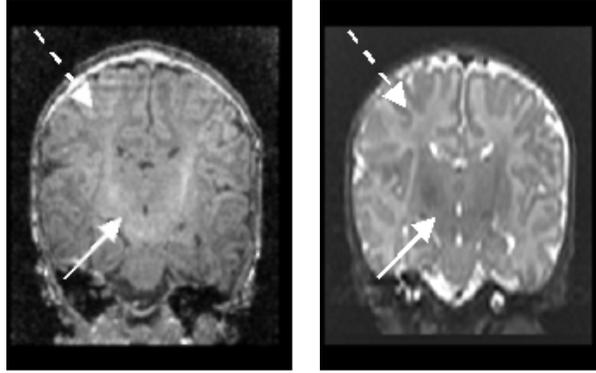
FIGURE 4.1. MR images of a newborn brain (subject 0096, coronal view). Left: T1w image, right: T2w image. The arrows show the white matter structure. The arrow with the solid line indicates myelinated white matter, the arrow with the dashed line indicates non-myelinated white matter. Early myelination in white matter is shown as bright regions in the T1w image and dark regions in the T2w image.

to automatic segmentation of adult brain MRI due to the early development process. Rutherford *et al.* [81] provide an excellent description of newborn MRI and the dynamic changes seen over the early development period, illustrating the significant challenges for reliable image segmentation.

As described in Chapter 1, the white matter structure in newborn infant brains undergoes myelination, where the fibers are being covered in myelin sheaths. At birth, the white matter of the brain stem and the posterior limbs of the internal capsule are myelinated and have white to gray matter contrast similar to that of adults (white matter is brighter than gray matter in the T1w image). Most regions of white matter such as the centrum semi-ovale, corpus callosum, and off-center regions are not myelinated and the white to gray matter contrast is inverted (white matter is darker than gray matter in the T1w image). As the child ages from birth to one year, myelination progresses through the anterior limbs of the internal capsule, the occipital radiations, and then to the frontal white matter. As this happens, the MR relaxation times of these regions change with the new myelinated fibers consequently changing the MRI signal. By age 1.5 years, the MR image contrast is almost adult-like. Figure 4.1 shows an example of a newborn MR image with the myelinated and non-myelinated white matter regions.

Automatic segmentation methods for healthy adult brain MRI typically fail in segmenting the different structures apparent in newborn brain MRI, particularly the myelinated white matter regions. Methods that use probabilistic brain atlases [93, 19] or templates [96] cannot be directly applied to newborn brain MRI since the spatial prior information for rapidly changing myelination property would be very difficult to define. Warfield *et al.* [96] uses a specific template for newborn brains with predefined classifications for myelinated and non-myelinated white matter. Methods that are driven by image intensities [100, 17] would have difficulties in the initialization phase. The MR image intensities for newborn brains are significantly affected by both low contrast and RF inhomogeneity, which can be difficult to overcome without spatial prior information.

Matsuzawa *et al.* [62] presented a segmentation method for infant brain MRI, as part of a study of early brain development. Their method does not identify myelinated white matter and non-myelinated white matter separately. The results show that their method has difficulties dealing with proper tissue separation. Hüppi *et al.* [42] and Inder *et al.* [43] showed segmentation results of newborn infants, using the method of Warfield *et al.* [96]. They study both prematurely born infants and normal infants. The prematurely born infants tend to have simpler cortical folding compared to normal newborns. The segmentation method identifies non-myelinated and myelinated white matter. Boardman *et al.* [9] used image deformation for detecting regions of major development.

Automatic segmentation of newborn brain MRI is significantly more challenging than the segmentation of healthy adult brain MRI. This is mainly due to the biology and the rapid growth process. The specific challenges are:

(1) The white matter and gray matter contrast to noise ratio (CNR) for newborn MRI can be as low as half of the one for adult brain MRI. Factors that reduce CNR are the small size of the infant brains, the short scan time, and the low contrast for gray and white matter. The small head size requires subjects to be scanned at higher resolution, which leads to higher noise levels. The infants

need to be scanned in very short time since they are not sedated or constrained. The low CNR further causes difficulty in segmenting the partial volume regions.

(2) Typically, newborn brain MRI exhibits subtle motion artifacts even with very short scan sequences. The infants may not stay motionless during the scan period even while sleeping. This problem can be difficult to solve since the infants are not mentally aware, and healthy infants cannot be sedated or restrained due to ethical reasons. Potential solutions are further reduction of scan time by parallel imaging and on-scanner motion correction.

(3) The process of myelination separates white matter tissue into two types: myelinated and not myelinated. Myelination is treated as a fractional property because the MR image intensities reflect the degree of myelination and partial voluming. The dividing boundaries between regions that are fully myelinated and non-myelinated are generally ambiguous [81]. The myelinated white matter regions are mostly distributed near the spine (central posterior) and parts of the internal capsule. The presence of myelinated white matter around the regions associated with the sensory and motor cortex is also observed.

(4) Each tissue type in newborn brain MRI exhibits significant levels of intensity inhomogeneity and variability, which may be due to a combination of RF inhomogeneity and biological properties of the developing tissue [46].

(5) The different tissues have large overlaps in their intensity characteristics, as shown in Figure 4.2. The decision boundaries for intensity-based classification are typically ambiguous and complex.

I propose adaption of the robust EM segmentation framework (Chapter 2) for segmentation of newborn brain MRI that addresses the challenges listed above. The method incorporates the robust clustering method proposed by Cocosco *et al.* [17] and the robust parameter estimation method presented by Rousseeuw and van Driesen [80] to deal with noisy data. It also uses the intensity inhomogeneity estimation scheme from spatial classification proposed by van Leemput *et al.* [92]. The complex decision boundaries
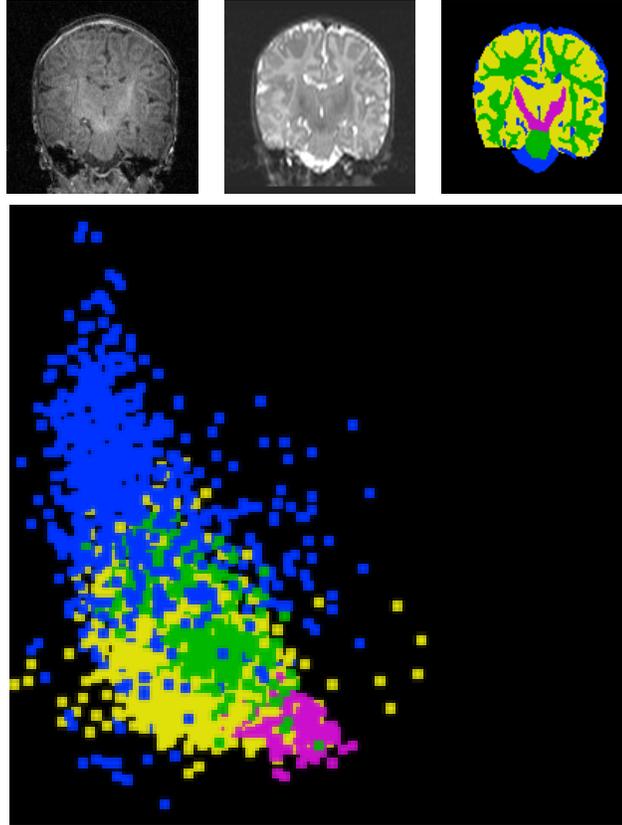
FIGURE 4.2. Intensity characteristics of one coronal slice of a newborn brain MRI dataset (subject 0096). Top, from left to right: T1w image, T2w image, and the manually assigned labels. Purple is myelinated white matter, green is non-myelinated white matter, yellow is gray matter, and blue is cerebrospinal fluid. Bottom: the scatterplot of the tissue intensities, the horizontal axis represents T1w intensities and the vertical axis represents T2w intensities. There is significant overlap between the intensities of different tissues, and there are ambiguities in the decision boundaries.

are modeled using non-parametric kernel density estimates, using the efficient method of Girolami *et al.* [**36**]. Probabilistic atlas of the newborn brain is used as a spatial prior in the classification process as proposed by [**92**]. All these procedures represent components of the new, robust EM segmentation as described in Chapter 2.

Due to the large overlap in the tissue intensity distributions, it becomes necessary to use spatial priors for the segmentation. The spatial priors chosen are part of a probabilistic brain atlas of newborn MRI developed for this application, shown in Figure 4.3. The atlas provides voxel prior probabilities for white matter, gray matter, and cerebrospinal fluid (csf). Myelinated white matter and non-myelinated white matter are combined as
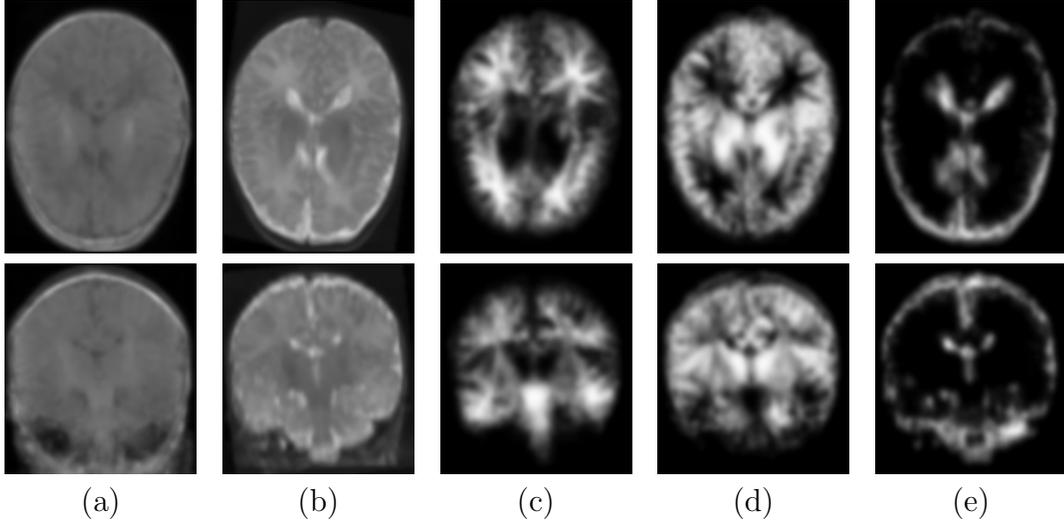
FIGURE 4.3. The probabilistic brain atlas of a newborn brain. From left to right: (a) the T1w average image, (b) the T2w average image, and the spatial prior probability values for (c) white matter (either myelinated or non-myelinated), (d) gray matter, and (e) csf. Top: axial view. Bottom: coronal view.

one white matter class in the atlas. This is necessary because it is difficult to model the different dynamic growth patterns across subjects given the rapid changes during early brain development. With the combined white matter prior, the discrimination between the two different white matter classes is primarily driven by the image intensities. The atlas was created by averaging three semi-automatic segmentations registered using affine transformation. Each segmentation was done by a human rater that selects samples for each tissue types for k-nearest neighbor segmentation. The outputs of the k-nearest neighbor classification are then edited by manual outlining. The number of subjects is insufficient to create prior probabilities that reflect the variability in the population. At this point, I am limited by the size of the datasets and the amount of time for manual processing, but the development of an improved probabilistic atlas is part of the future work in UNC Psychiatry. To compensate for the lack of available data, an additional blurring is applied to the average segmentations. The blurred spatial probabilities simulate an atlas with a higher level of population variability. The segmentation performance is improved by using the blurred version of the limited atlas as it accounts a more flexible prior. Details of the automatic segmentation method are presented in the next section.
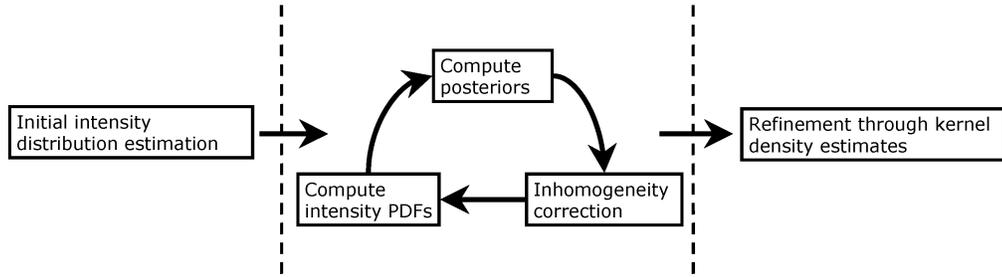
FIGURE 4.4. The segmentation framework for newborn brain MRI.

4.2. Method

My newborn brain MRI segmentation framework is composed of three major steps, as shown in Figure 4.4. First, it obtains rough estimates of the class intensity clusters. It then iteratively performs inhomogeneity correction and parametric classification. Finally, it refines the segmentation using non-parametric kernel density estimates.

4.2.1. Estimation of Intensity Distributions. The segmentation of newborn brain MRI involves classifying each voxel into different categories $\mathcal{C}$, where $\mathcal{C}$ is defined to be {myelinated white matter, non-myelinated white matter, gray matter, and cerebrospinal fluid}. The first step in the segmentation process is to determine rough estimates of the class intensity distributions. The method obtains samples for class $c \in \mathcal{C}$ at location k with high atlas prior probability values, for example $Pr(S_k = c) > 0.9$ as presented by Cocosco *et al.* [**17**] and described previously.

The white matter samples are constrained to have low image gradient magnitude values to avoid choosing samples near the transition regions between myelinated and non-myelinated white matter and at white/gray matter boundaries. The value used for the gradient magnitude of the collection of the 3-D images is the 2-norm of the vector of the individual gradient magnitudes of the different modalities,

$$G_k = \sqrt{|\nabla I_{k,1}|^2 + \ldots + |\nabla I_{k,n}|^2} \tag{4.2.1}$$

51

where $I_{k,d}$ is the intensity of the dth image at location k. I only retain samples for the white matter class with multimodal gradient value G_k lower than the average of the G_k values over the white matter prior,

$$G_{threshold} = \frac{\sum_k Pr(S_k = white\, matter)\, G_k}{\sum_k Pr(S_k = white\, matter)}. \qquad (4.2.2)$$

The 2-norm gradient magnitude metric is more sensitive to noise compared to the vector field gradient magnitude metric described in [51]. This is a desired property since we want to avoid sampling noisy regions.

The obtained intensity samples is then processed to remove outliers and false positives. I use the Minimum Covariance Determinant (MCD) estimator [80] (see Section 2.3.1) to generate the robust mean and covariance estimates of the unimodal distributions (gray matter and csf). The MCD estimator computes the robust mean and covariance that have the smallest determinant of covariance and covers at least half of the data. For the bi-modal white matter distribution, I use a robust graph based clustering method, similar to the one described in [17]. The clustering method creates the minimum spanning tree (MST) graph [21] from the sample points and breaks long edges to form the clusters [27]. The minimum spanning tree is the graph where all the points are connected such that the total edge lengths are minimized. The MST graph does not have any closed loops (cycles). The removal of samples with relatively high image gradient helps in the MST clustering process, as shown in Figure 4.5.

The algorithm searches for myelinated white matter and non-myelinated white matter intensity clusters by iteratively breaking long edges of the MST. At each iteration, we break an undirected edge $e(v, w)$ that connects vertices v and w if it is longer than $A(v) \times T$ or $A(w) \times T$. $A(v)$ is the average length of edges incident on vertex v, $A(v) = \frac{1}{|N(v)|} \sum_q |e(v, q)|$, $q \in N(v)$ where $N(v)$ is the set of vertices that are neighbors of v. T is a scalar distance multiplier that determines the sensitivity to the differences in intensity features. The edge breaking results in subtrees where each subtree forms an intensity cluster. For each detected cluster, an intensity location estimate is computed. The cluster
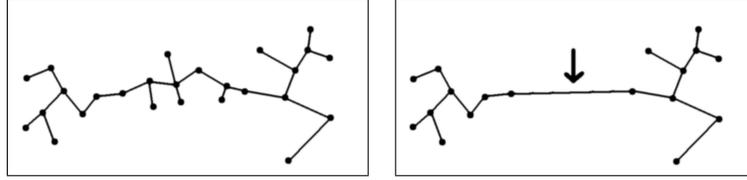
FIGURE 4.5. Illustrations of the Minimum Spanning Trees for white matter obtained using different sampling strategies. Left: Samples with high probability values. Right: Samples with high probability values and low gradient magnitude. Choosing only samples with low gradient magnitude helps to remove samples from the transition regions between myelinated white matter and non-myelinated white matter and gray/white boundary voxels. This is crucial for clustering based on edge breaking. As seen on the right picture, breaking the longest edge marked by the arrow would give two well separated clusters.

intensity location estimate provides an approximation of where most points in the cluster are distributed in the intensity feature space. The iterative algorithm terminates when two clusters are found with intensity location estimates that are in the proper order. For example, the order of intensities for the classes in T2w from darkest to brightest is myelinated white matter, gray matter, non-myelinated white matter, followed by csf. Here again, domain knowledge is used to help assign proper clusters to tissue types.

The intensity location estimates for the two white matter classes are computed using the MCD estimator. I use the robust MCD mean values, as opposed to the standard location estimates such as the mean or median, to make sure that the algorithm obtain reasonable sample clusters. The standard location estimates such as mean or median are generally not robust enough for the noisy newborn MRI data. The mean value could be skewed by a single outlier sample, while the median value only uses one sample point and ignores contributions of other samples. The initial intensity distributions for non-myelinated white matter and myelinated white matter are computed as the MCD mean and covariance estimates of the largest detected clusters. The MCD estimator therefore serves to estimate the initial intensity distributions. The initial gray matter and csf distributions are the MCD estimates of the atlas sampled data. The initial white matter distributions are the MCD estimates of the atlas samples that are clustered and

pruned using MST. The steps involved in the intensity distribution estimation are listed in Algorithm 1.

Algorithm 1 Initial intensity distribution estimation

 1: Obtain samples by thresholding atlas prior probabilities
 2: Remove white matter samples with gradient magnitude higher than $G_{threshold}$
 3: Compute robust mean intensity values for gray matter and csf (μ_{gm} and μ_{csf}) using the MCD estimator
 4: Construct Minimum Spanning Tree from white matter samples
 5: $T \leftarrow 2$
 6: **repeat**
 7:     Break edges longer than $T \times A$, where A is the average length of connected neighbor edges
 8:     Find largest myelinated white matter cluster, where $\mu_{myelinated} < \mu_{gm}$ in T2w
 9:     Find largest non-myelinated white matter cluster, where $\mu_{gm} < \mu_{non-myelinated} < \mu_{csf}$ in T2w
10:     $T \leftarrow T - 0.01$
11: **until** both white matter clusters are found or $T \leq 1$
12: **if** $T < 1$ **then**
13:     Algorithm fails
14: **end if**
15: Compute white matter Gaussian distribution parameters from detected clusters

---

4.2.2. Intensity Inhomogeneity Correction. Newborn brain MR images exhibit higher intensity variability for each tissue and lower intensity contrast compared to adult brain MRI. These two factors severely hamper the estimation of intensity inhomogeneity. Histogram based intensity inhomogeneity estimation methods, such as the ones proposed by Sled *et al.* [**85**] and Styner *et al.* [**87**], are likely to have difficulties in obtaining the optimal solution. The histogram of a newborn brain MR image is generally smooth with weak or non-existent maximas.

In the case of inhomogeneity correction of newborn brain MRI, the spatial information is useful to deal with the low intensity contrast. I have chosen to use the method developed by van Leemput *et al.* [**92**]. The scheme uses the spatial class posterior probabilities to estimate the intensity inhomogeneity, which helps to overcome problems with low contrast and high variability. The inhomogeneity estimation method is part of the iterative

generalized Expectation-Maximization algorithm. It interleaves classification with inhomogeneity estimation at each iteration. The Gaussian distributions obtained from the initial segmentation step are used as initial parameters for the iterative inhomogeneity estimation algorithm.

The segmentation method uses the spatial priors defined by the newborn brain atlas for the class posteriors, following the Bayes rule

$$p(S_k|I_k, \theta) = \frac{p(I_k|S_k, \theta)Pr(S_k)}{p(I_k)}. \qquad (4.2.3)$$

The myelinated and non-myelinated white matter shares the same atlas prior $Pr(S_k = \{white matter\})$ that is divided into the prior values for myelinated and myelinated white matter using global weights w_i, $Pr(S_k = \{myelinated\ white\ matter\}) = w_1\ Pr(S_k = \{white matter\})$ and $Pr(S_k = \{unmyelinated white matter\}) = w_2 Pr(S_k = \{white matter\})$ The global class prior weights w_i can be tuned based on the age of the newborns to be segmented. For the results presented here, I set the global class priors such that white matter is more likely to be not myelinated: $w_1 = 0.2$, $w_2 = 0.8$. The use of the atlas spatial prior probabilities $Pr(S_k)$ helps resolve ambiguities that are caused by the low image contrast following the formulation used in [**93**].

4.2.3. Segmentation Refinement. The class intensity likelihoods are modeled as Gaussian probability density functions in the segmentation and inhomogeneity correction to obtain an optimal parametric solution. The use of the parametric Gaussian distribution eases the computation of the maximum likelihood estimate. However, Gaussian distributions can have significant overlap and therefore result in degenerate decision boundaries. In order to capture the complex and ambiguous intensity characteristics of newborn brain MRI, the method switches from the parametric Gaussian distribution to a non-parametric distribution estimate. I refine the classification by sampling the inhomogeneity corrected images, pruning the outliers and false positives from the intensity samples, and then estimating the intensity distribution using kernel density functions [**27, 39**].

The non-parametric intensity probability density function for each class c is estimated as follows:

$$\hat{p}(I_k|S_k = c, \theta) = \sum_{j=1}^{N_i} w_{c,j} \, K_h(I_k - Y_{c,j}) \tag{4.2.4}$$

where K_h is the Gaussian kernel with standard deviation h, N_i is the number of training samples for class c, and $Y_{c,j}$ is the jth training sample for the class c. Each training sample has an associated weight $w_{c,j}$, where for each class c, $\sum_{j=1}^{N_i} w_{c,j} = 1$. The kernel density estimates are used to produce the final classification results, which are the class posterior probabilities:

$$\hat{p}(S_k|I_k, \theta) = \frac{\sum_c \hat{p}(I_k|S_k = c, \theta) Pr(S_k = c)}{\sum_{c'} \hat{p}(I_k|S_k = c', \theta) Pr(S_k = c')} \tag{4.2.5}$$

The atlas spatial prior probabilities are also used at this stage. The spatial priors are combined with the non-parametric kernel densities to provide class posterior probabilities that are capable of capturing more complex intensity characteristics.

The set of training samples Y for the kernel density estimates are obtained by sampling the MR images using the previously obtained posterior probabilities. Each sample $Y_{c,j}$ for class c is obtained by selecting features at location k where the following condition is satisfied:

$$\arg\max_{S_k} p(S_k|I_k, \theta) = c. \tag{4.2.6}$$

The samples are pruned and clustered using the robust MST-based method proposed by Cocosco *et al.* [17]. This step removes the false positives and outliers in the intensity data resulting from using Gaussian distribution estimates in the previous step.

The method proposed by Girolami *et al.* [36] is applied to efficiently estimate the kernel density function. This method speeds up the density estimation process by reducing the size of the training set. The weights $w_{c,j}$ are chosen to minimize the integrated squared error between the true density function and the estimated kernel density function. Redundant training features are assigned lower weight values compared to characteristic training features. This minimization process for the sample weight assignment is similar

to the quadratic optimization process for Support Vector Machines, for which an efficient solution exists [82]. The samples with zero weights are removed from the training set, which effectively eliminates the redundant features in the training set. Compared to other fast density estimation techniques such as pre-binning [83] and multi-scale selection using hyperdiscs [68], this method has the advantage of having only one user specified parameter: the kernel width or the standard deviation of the Gaussian kernels.

4.3. Results and Validation

Validation of the automatic segmentation results of the newborn brain MR images is difficult because a gold standard does not exist. The common standard, manual segmentations, are difficult to obtain since highly convoluted structures in low-contrast, noisy data are very hard to trace. In addition to that, the myelinated white matter and the non-myelinated white matter have ambiguous boundaries, which would make manual segmentation results highly variable and difficult to reproduce. A limited validation of the segmentation results has been performed by restricting the validation only to the 2D coronal slices of five datasets. Two human raters assign discrete labels to each voxel in the slice. White matter is divided into two distinct classes (myelinated and non-myelinated) and the degree of myelination is not specified because it would be extremely difficult for the raters to consistently assign a continuous weight for myelination.

Figure 4.6 shows the coronal view of the MR images along with two sets of manual segmentation slices done by different raters. The four cases are drawn from a large neonatal study at UNC Chapel Hill to assess early brain development in normal and high risk children [105, 34]. Figure 4.7 shows surface renderings of an example segmentation result. Figure 4.8 shows the coronal view of the automatic segmentation results for four subjects. The 3D volumes for the automatically segmented structures are listed in Table 4.1. Visual inspection of the results show that the myelinated white matter regions are mostly distributed near the spine (central posterior) and internal capsule. The presence of small regions of myelinated white matter around the regions associated with the sensory and motor cortex can also be observed.

The validation measures described in Appendix A are used to compare the 2D segmentations. Since the comparison is performed on images with a finite collection of segmentation labels, Cohen's kappa measure is used to measure the segmentation variability. The kappa values comparing the two manual segmentations and the manual segmentation against the automatic segmentation is shown in Table 4.2. Dice similarity coefficient is also used to evaluate the inter-rater variability and the relative performance

of the automatic segmentation method. The DSC values reflecting inter-rater variability is shown in Table 4.3. The overlap comparison between the manual raters and the automatic segmentation method is shown in Table 4.4 and Table 4.5. Since the validation is only done on 2D slices, the number of samples is low and together with the high complexity of the folding structures this consequently leads to relatively low overlap values.

The kappa values show that there is insufficient level of reliability for the two manual segmentations. The non-myelinated white matter and gray matter classes have higher number of observations compared to the other classes and therefore dominate the kappa measurements. The kappa values are low because the segmentations of the brain tissue classes tend to be ambiguous. The overlap measures show that the automatic segmentation method has similar level of variability as the two manual segmentations. The overlap values of the cerebrosinal fluid regions for the automatic method are generally lower due to misclassifications in the ambiguous partial volume regions.

The kappa and DSC values related to the proposed automatic segmentation method show that the method have similar level of variability to the inter-rater variability for the human raters. This demonstrates that the proposed method has comparable performance to a manual rater, given the limitations due to the restricted number of validation samples and the noise and ambiguity in the image data.

(a)	(b)	(c)	(d)
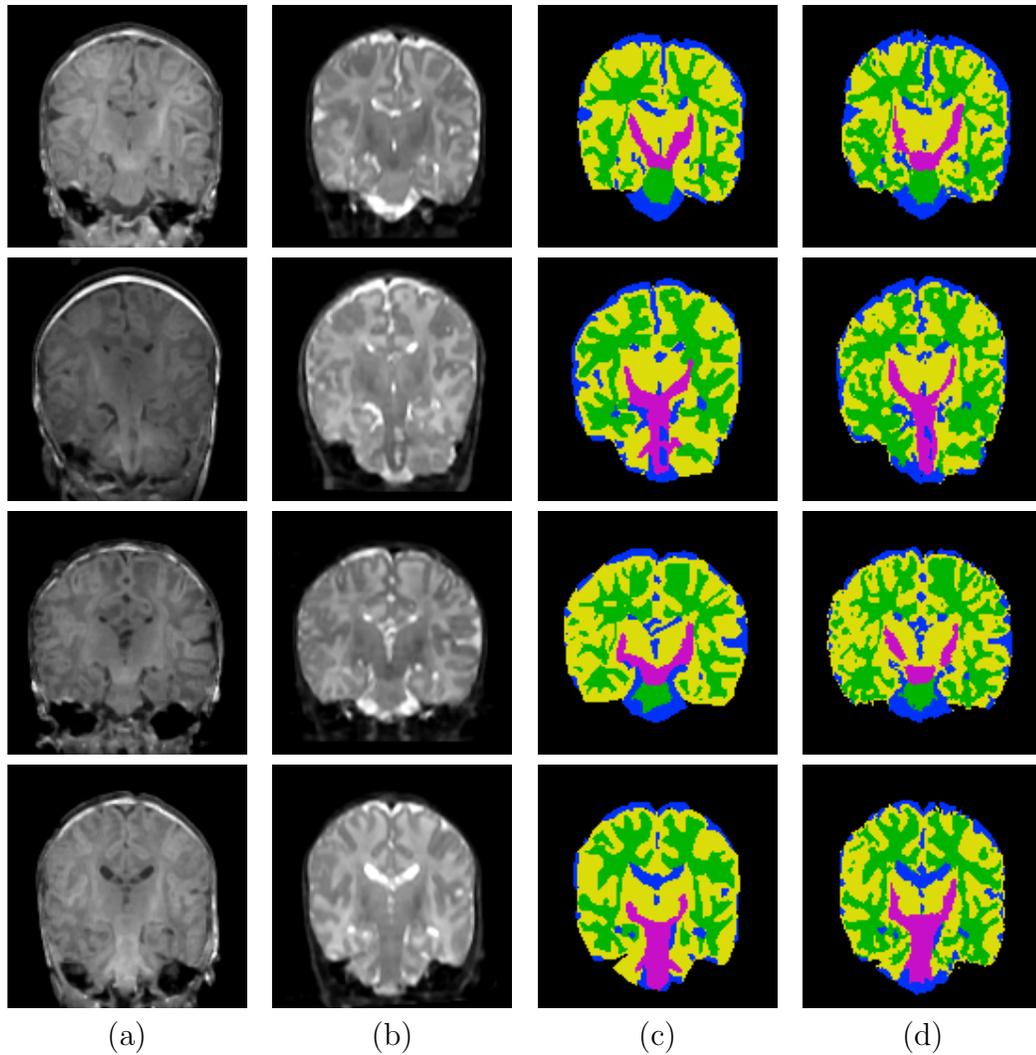
FIGURE 4.6. The newborn MR images along with the manually segmented labels. From left to right: (a) T1w image, (b) T2w image, (c) color image showing the segmentation obtained by the first human rater, and (d) color image showing the segmentation obtained by the second human rater. Purple is myelinated white matter, green is non-myelinated white matter, yellow is gray matter, and blue is csf. From top to bottom: subject 0096, 0117, 0118, and 0123.

(a)   (b)   (c)   (d)
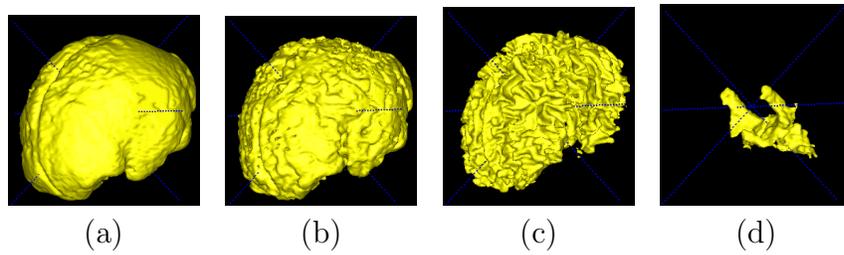
FIGURE 4.7. Surface renderings of the segmented structures of newborn subject 0123. From left to right: (a) intra cranial volume, (b) gray matter, (c) non-myelinated white matter, and (d) myelinated white matter.

(a)   (b)   (c)   (d)   (e)
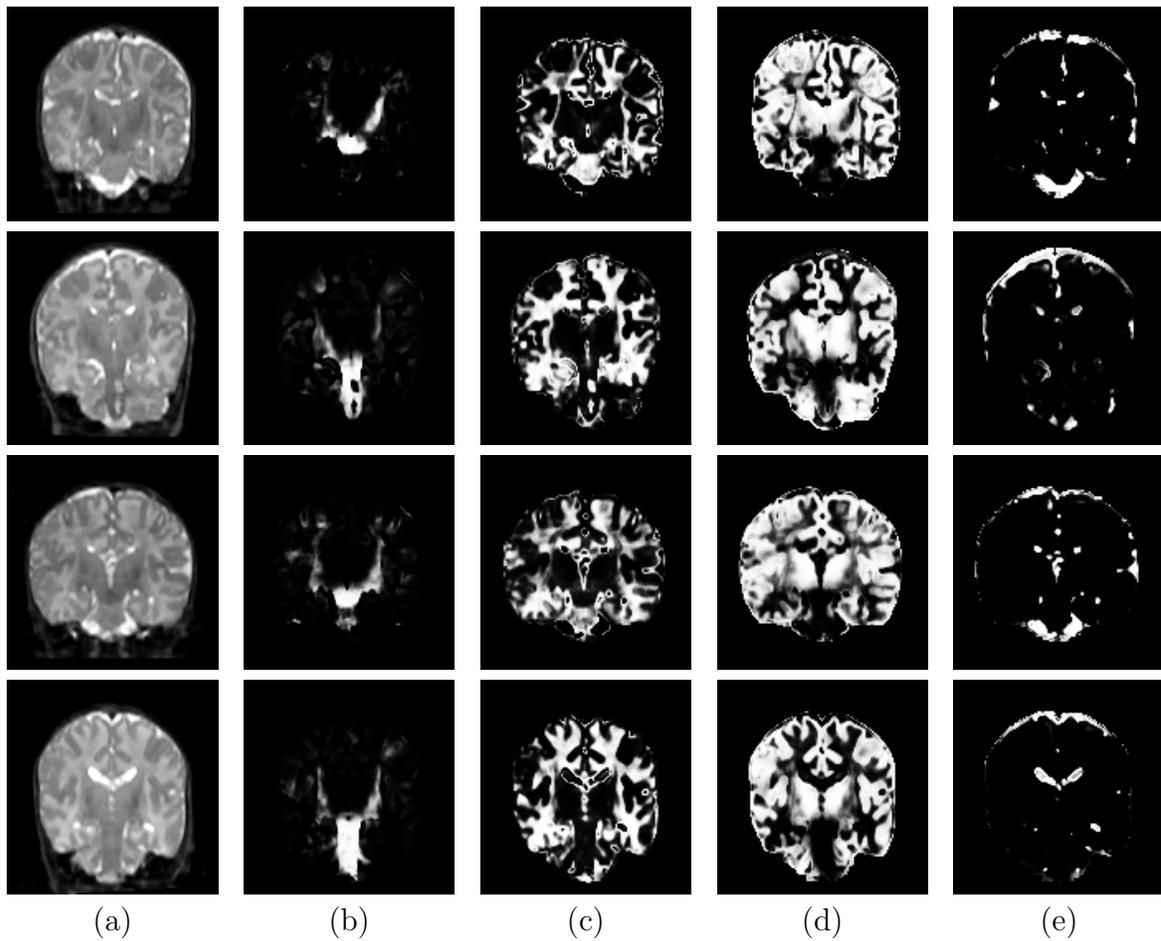
FIGURE 4.8. Coronal view of the 3D automatic newborn brain MRI segmentation results. From left to right: (a) the T2w image and the class posterior probabilities for (b) myelinated white matter, (c) non-myelinated white matter, (d) gray matter, and (e) cerebrospinal fluid. From top to bottom: subject 0096, 0117, 0118, and 0123.

61

Subject	ICV	Myelinated WM	Non-myelinated WM	Gray Matter	CSF
0096	504724	15353	157160	289133	43078
0117	527885	12678	234706	250161	30340
0118	514760	11480	193307	255849	54124
0123	499775	28487	170227	252056	49005

TABLE 4.1. The volumes of the segmented structures for four newborn subjects. These include the intra cranial volume (ICV) and the volumes of the individual structures (myelinated white matter, non-myelinated white matter, gray matter, and cerebrospinal fluid). All volumes are measured in cubic millimeters.

Subject	Rater 1 vs Rater 2	Rater 1 vs Automatic	Rater 2 vs Automatic
0096	0.658	0.604	0.558
0117	0.627	0.577	0.587
0118	0.603	0.561	0.500
0123	0.625	0.626	0.542

TABLE 4.2. The kappa coefficients that measure the level of agreement between manual raters, first manual rater against the automatic method, and second manual rater against the automatic method.

Subject	Myelinated White Matter	Non-myelinated White Matter	Gray Matter	CSF
0096	0.715	0.767	0.777	0.738
0117	0.760	0.771	0.741	0.662
0118	0.683	0.738	0.752	0.696
0123	0.787	0.757	0.750	0.639

TABLE 4.3. The Dice similarity values that measure the overlap between the two manual segmentations of newborn brain MRI.

Subject	Myelinated White Matter	Non-myelinated White Matter	Gray Matter	CSF
0096	0.634	0.676	0.809	0.681
0117	0.637	0.725	0.776	0.491
0118	0.681	0.661	0.782	0.598
0123	0.777	0.724	0.790	0.569

TABLE 4.4. The Dice similarity values that measure the overlap between the segmentation results of the first human rater and the automatic method for newborn brains.

Subject	Myelinated White Matter	Non-myelinated White Matter	Gray Matter	CSF
0096	0.774	0.649	0.777	0.598
0117	0.651	0.739	0.760	0.577
0118	0.606	0.630	0.742	0.601
0123	0.719	0.694	0.732	0.478

TABLE 4.5. The Dice similarity values that measure the overlap between the segmentation results of the second human rater and the automatic method for newborn brains.

4.4. Newborn Brain Population Study

The newborn segmentation framework described in this chapter has been applied to 74 healthy newborn brain MRI datasets (40 males and 34 females). From the segmentation of the growing newborn brains, Gilmore *et al.* [**35**] observed that there are significant differences in regional growth rates for the cortical gray matter structure (Figure 4.9) that are consistent with results from previous studies. The regional growth of the unmyelinated white matter structure, as seen in Figure 4.10, is less pronounced. For both gray matter and unmyelinated white matter, the posterior regions tend to grow faster than the anterior regions. This provides an indication that sensory-motor development occurs earlier than cognitive development. The regional differences in gray matter growth is consistent with previous studies in synapse development. This pattern (posterior region growing faster than anterior region) is also observed in unmyelinated white matter, though less pronounced.

Cerebral asymmetry was present at birth, with the left hemisphere being consistently larger than the right. This is the opposite asymmetry pattern observed in older children and adults. When accounting for gender, males are observed to have larger total brain volumes compared to females at birth. The larger total brain volume is mainly due to increased amount of white and gray matter in the male subjects.

The study has found that early development is characterized by robust growth of cortical gray matter compared to white matter and of the posterior regions compare to the anterior regions. The growth patterns observed in the newborn subjects display some characteristics that are the opposite of those observed in adults and older children. The newborn brain volumes are found to be roughly 35% of the adult brain volumes. This indicates that there is enormous growth occuring between birth and adulthood. Further studies still need to be done to learn more about early brain development. In particular, studies comparing healthy subjects and subjects-at-risk for neurological diseases and psychiatric disorders would be of great value in determining abnormal growth patterns.
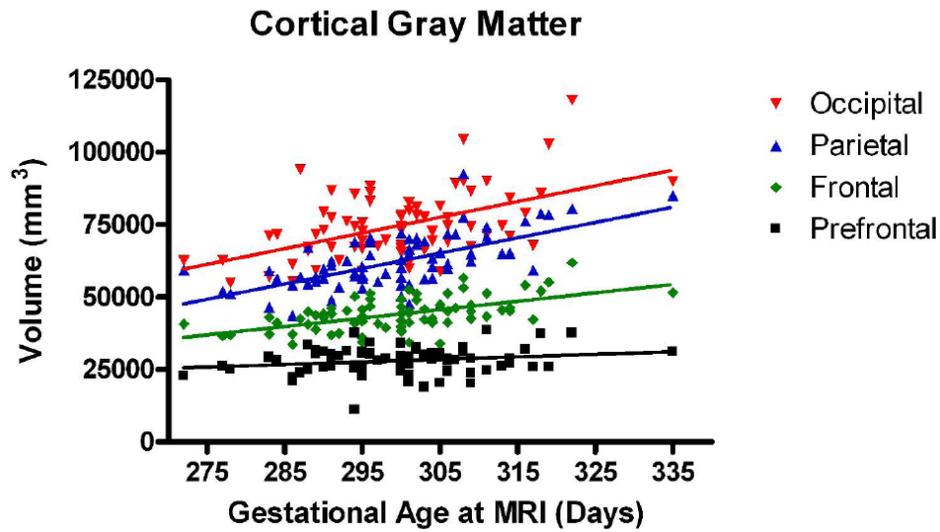
FIGURE 4.9. Regional growth of cortical gray matter in newborns. The growth rates (slopes) of the posterior regions tend to be significantly higher than the anterior regions. Image provided by John Gilmore [**35**].
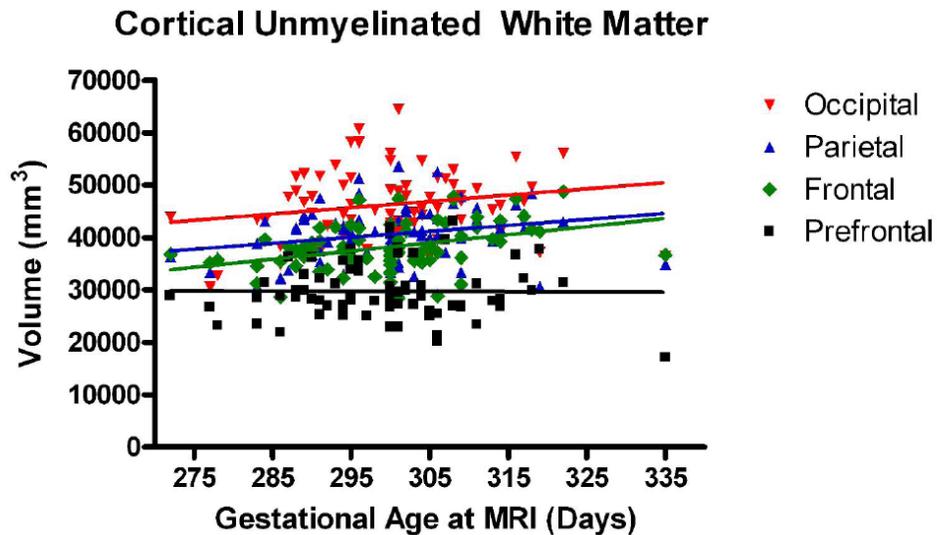
FIGURE 4.10. Regional growth of cortical unmyelinated white matter in newborns. The growth rates (slopes) of the posterior regions tend to be slightly higher than the anterior regions. Image provided by John Gilmore [**35**].

4.5. Conclusions

In this chapter, I have presented an atlas-based automatic segmentation framework for segmenting newborn brains from multimodal MR images. The problem related to the low signal-to-noise ratio in newborn brain MRI is handled by combining a spatial prior (brain atlas) with robust parameter estimation techniques. The method uses two robust parameter estimation methods: graph-based clustering and MCD estimator. They are used to obtain initial parameter estimates for the EM image segmentation and to refine the final segmentation results.

Visual inspection of the results shows that the major structures are segmented consistently. Segmentation of partial voluming regions is still insufficient and is an inherent problem with voxel-based classification. The segmentation results of the four cases presented in the previous section show that the new method can cope with variable brain shapes. The locations of the early myelination accross the subjects appear to be similar. The method has been applied in a study involving more than 74 newborn subjects, with optimal reproducibility since the method is fully automatic. This study is made possible using the automatic newborn brain tissue segmentation method described in this chapter. It is by far the largest neonate sample set analyzed to date and revealed very interesting results as described in [**35**].

Due to the lack of a gold standard, I have only performed a limited validation of the results. Validation is done only on selected 2D slices and not on the whole 3D brain volume. The κ coefficient values and volume overlap measures show that our segmentation results have similar level of variability to the inter-rater variability for manual segmentations. The lack of a complete 3D ground truth data is an issue that can be solved using simulated datasets. A discussion on possible future work related to newborn brain segmentation can be found in Chapter 6.

CHAPTER 5

Simulation Data for Objective Validation

This chapter presents a novel framework for generating synthetic brain tumor MRI data with known underlying ground truth. Section 5.1 describes the motivation and challenges for objective validation of image segmentations. The new method for generation of pathological ground truth with tumor and edema from a healthy ground truth is covered in Section 5.2. A method for generating synthetic MR images, which act as test data for evaluation of different segmentation methods, is presented in Section 5.3. Example results and validation of the simulation process are described in Section 5.4

5.1. Background

Integral to any image segmentation framework is the validation of the segmentation results, where the results are compared against a standard (ground truth) using various measures. A segmentation algorithm must undergo systematic validation to ensure that it performs reliably in a clinical setting or a medical study. As discussed in Section 1.1, the usual standard used for validating the segmentation results of the automatic methods is the manual segmentation results done by human experts. However, different investigators are likely to employ different image acquisition parameters and different manual segmentation methods. A compounding issue is that any manual segmentation method suffers from lack of reliability and reproducibility. Even if a rich set of manual segmentations are available, they may not reflect the ground truth, and the true gold standard may need to be estimated [97]. For the practical problems addressed in this dissertation (segmentation of brain tumor and newborn brain MRI), validation is particularly difficult

since the tumor boundary definition can be ambiguous and no gold standard exists for newborn brain MRI.

In this chapter, emphasis will be placed on the validation of segmentations of brain MRI with tumor and edema. In most cases, priority is given to pathological structures such as tumor and edema, so comparisons of the segmentation of structures other than the brain tumor are not done. There is a high degree of variability for the segmentations of edema done by human raters due to the inherent ambiguity, so if manual segmentations are available they are generally less than ideal.

Brain MRI exhibiting tumor is difficult to segment due to a combination of the following factors:

- The deformation of brain tissue due to tumor mass effect or volume expansion.
- The infiltration of brain tissue by tumor and edema (swelling). Edema appears around tumor mainly in the white matter regions.
- The gradual transition between tumor, edema, and surrounding brain tissue. This results in the ambiguity of the structural boundaries.
- The T1w MRI with contrast enhancement, typically using a gadolinium agent, is the standard modality for identifying tumors. This modality results in active tumor tissue appearing with bright intensity. Unfortunately, blood vessels also appear bright while parts of tumor that are necrotic do not have higher levels of intensity. Therefore, the information provided by the intensities in this modality is not always consistent, and it is generally impossible to segment the tumor by thresholding the intensities in this image modality.

In order to provide objective assessments of segmentation performance, there is a need for an objective 3D ground truth with associated MR images that exhibit the same major segmentation challenges as that of common, realistic scans of a tumor patient. For this purpose, I develop a method that generates realistic looking MR images with the associated ground truth by simulating brain tumor growth.
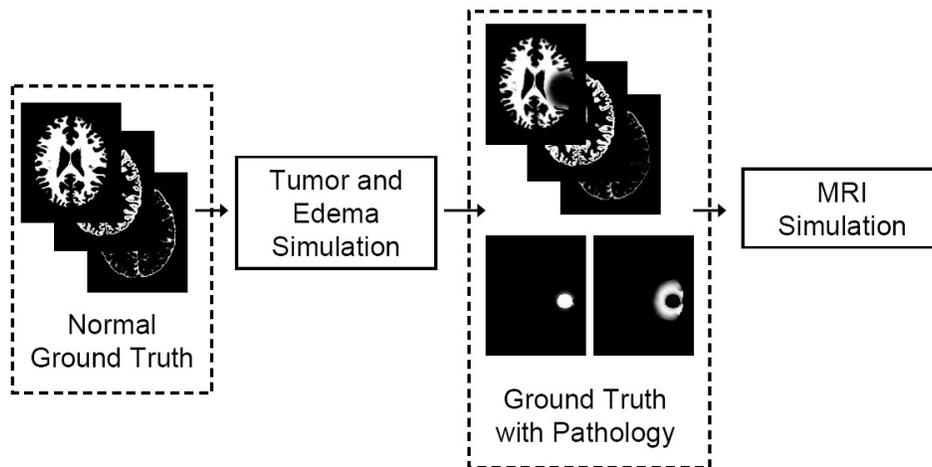
FIGURE 5.1. Overview of the generation of validation data. A well defined ground truth for normal brains is modified following a tumor and edema growth model. The normal brain ground truth contains the probabilities for white matter, gray matter, and csf drawn from the BrainWeb data. The healthy tissue probabilities are modified to take into account mass effect and infiltration and new pathological probabilities are added (tumor and edema). The modified ground truth is then used to create the synthetic multi-modal MR images. The pathological simulation is described in Figure 5.2 and the MRI simulation is described in Figure 5.5.

Rexilius *et al.* [79] proposed a framework for generating digital brain phantoms with tumor. They used a biomechanical linear elastic finite element model to simulate the tumor mass effect. In their method, he MRI of a healthy subject is deformed and a tumor structure from a real subject is inserted into the MRI. Their model for edema is computed from the distances to the tumor boundary and the white matter mask. This is insufficient to simulate real infiltration properties since infiltration can occur in regions away from tumor. Such regions are typically connected through white matter fibers. Moreover, the framework of Rexilius *et al.* only considers contrast enhancement inside tumors, without contrast enhancement of blood vessels.

Models for brain tumor expansion and edema has been proposed by Nagashima *et al.* [73], Clatz *et al.* [14, 15], and Mohamed *et al.* [69, 70]. More recently, Clatz *et al.* developed a realistic tumor growth model that simulates the main effects of tumor growth (mass effect and infiltration) using simple computational models. Clatz *et al.* use

a biomechanical finite element model to simulate mass effect, and they use a reaction-diffusion process that is modulated by the diffusion tensor field to simulate the infiltration by tumor cells and edema. In this chapter, I propose a new method for generating pathological ground truth by applying their mass effect and infiltration model to a well defined ground truth for healthy brains. Additionally, I propose to extend the Clatz *et al.* model by using random pressure directions and by simulating the effect of volume expansion on the white matter fibers by warping the diffusion tensors and making them more isotropic depending on the magnitude of local deformations.

I develop a method for generating realistic-appearing contrast enhanced T1 weighted MR images (a standard modality for diagnosis) by simulating the accumulation of contrast agents in the brain. The corresponding multi-modal MR images are generated from the simulated ground truth and textures synthesized from samples of a real tumor MRI data. Figure 5.1 shows an overview of the proposed method. The simulation method is capable of generating 3D whole brain ground truth that exhibits the effects of a real tumor on normal brains, along with simulated multi-modal MR images that are challenging to segment. The method for generating synthetic brain tumor ground truth does not attempt to simulate the complete process of real tumor growth. Instead, I aim to generate sets of realistic looking images, with the associated ground truth, that are almost as challenging to segment as real brain tumor MRI. The simulation data, along with the probabilistic ground truth, is designed to validate segmentations such as the ones generated by the tumor segmentation method described in Chapter 3.
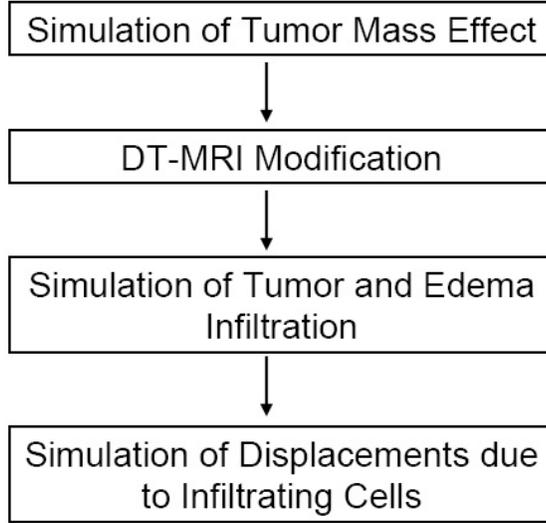
FIGURE 5.2. Overview of the simplified tumor and edema growth model. The model is composed of four sequential processes, where it simulates the deformation due to tumor expansion, the modification of DT-MRI due to the deformation, the infiltration of brain tissue by tumor cells and edema, and the displacements of tissue due to the infiltrating cells.

5.2. Generation of Pathological Ground Truth

Tumor and edema growth involves many concurrently occurring processes. As proposed by Wasserman *et al.* [98], the growth model may involve biomechanics, nutrient distribution, and metabolic processes. Since my goal is not to model tumor growth per se, I have chosen to simplify the model and use three separate sequential processes for efficiency, as shown in Figure 5.2. First, I simulate the deformation that is due to tumor mass effect using a biomechanical model. It is then followed by the simulation of the infiltration process using reaction-diffusion guided by diffusion tensor information. Finally, I compute the deformation that is due to tumor infiltration of brain tissue and the mass effect of edema. The BrainWeb dataset [16], which contains multimodal MR images along with spatial probabilities of normal brain structures, is used as the healthy brain ground truth that is transformed into a pathological ground truth. Figure 5.3 shows subject 04 from a collection of 20 normals from the BrainWeb datasets [4], which is used to generate the results described in this chapter. The dataset of subject 04 includes

FIGURE 5.3. Axial views of subject 04 from the BrainWeb dataset of twenty normals, which provides a standard for validation of normal brain MRI segmentation. From left to right: the T1w image and the spatial probabilities for white matter, gray matter, csf, and blood vessels.

spatial probabilities for each voxel being white matter, gray matter, cerebrospinal fluid, and blood vessel.

5.2.1. Mass Effect. The effect of tumor volume expansion on surrounding tissues is modeled using continuum mechanics [**38**]. The initial tumor seed region is defined manually and then undergoes simulated deformation that represents tumor mass effect. This initialization can also be done automatically given some prior knowledge of the spatial distribution, configuration, and image intensity characteristics of various brain tumor types. Meningiomas, for example, tend to be uniformly enhancing, to possess smooth borders and to originate from meningothelial cells associated with the arachnoid and dura matter. Glioblastomas, on the other hand, tend to be ring enhancing with irregular borders and almost always arise within the white matter. Metastatic lesions tend to have uniform or ring-like contrast enhancement, are often relatively spherical, and can appear in any location. The work described in this chapter provides examples of tumors that would likely represent metastatic lesions or glioblastomas, but the approach is generalizable to any tumor type.

In the initial tumor region, the tumor probabilities are set to one, $p_{tumor}(x) = 1$, and tissue or fluid probabilities are set to zero. The set of spatial probabilities for healthy tissue, along with the new tumor probabilities, are deformed according to a biomechanical model of brain tissue as described below. Brain deformation is modeled using the classic

72

linear elasticity model. The constitutive equation that relates stress and strain is

$$\sigma = E\epsilon \tag{5.2.1}$$

and the corresponding linear strain-displacement equation is

$$\epsilon = \frac{1}{2}(\nabla + \nabla^T)u \tag{5.2.2}$$

where σ denotes the stress tensor, E denotes the elasticity tensor, ϵ denotes strain, and u denotes the displacement. The elasticity tensor E is a function of the Young modulus and the Poisson ratio [41]. Following Clatz *et al.* [15], I use the linearized homogeneous version of the constitutive equation proposed by Miller [66], where brain tissue (white and gray matter) is assigned the value of 694 Pa for the Young modulus and 0.4 for the Poisson ratio. The falx cerebri, the fold of dura matter that divides the left and right brain hemispheres, is considered to be a stiff material with the value of $200,000$ Pa for the Young modulus and 0.4 for the Poisson ratio. The skull is considered fixed and brain tissue slides along contact with it.

The volume expansion due to tumor mass effect is modeled using a homogeneous pressure that is applied to tissues surrounding tumor [50, 69, 98]. The displacement field solution satisfies the static equilibrium equation

$$\text{div}(\sigma) + f_{ext} = 0 \tag{5.2.3}$$

with f_{ext} being the external forces applied to the model. The external forces that act on the tumor surface are formulated as follows

$$f_{ext} = P\,A\,\mathcal{VMF}(n, \kappa) \tag{5.2.4}$$

where P is the constant pressure (in Pa), A is the surface area, and $\mathcal{VMF}(n, \kappa)$ is a direction drawn randomly from the von Mises-Fisher distribution with mean direction n and concentration parameter κ [60]. The von Mises-Fisher distribution can be considered as the directional analogue to the multivariate normal distribution, where we use the

surface normal n as the mean direction and we use κ as the parameter that is inversely proportional to the spread or variability of the directions. The use of randomly generated directions increases the variability of the generated tumor shape and thus presents more challenge in segmenting the brain tumor. With regard to the displacements u, the following boundary conditions are applied:

(1) The sliding boundary condition in the regions where brain tissue contacts the skull, represented by

$$u \cdot n = 0$$

where n is the normal direction for the element boundary [65].

(2) The pressure inside the ventricular system is considered negligible relative to the pressure induced by tumor on the brain tissue, so the ventricular nodes are allowed to move freely.

The biomechanical problem is discretized using the finite element method, similar to the approaches used by Ferrant *et al.* [29] for inter-operative registration and Kyriacou *et al.* [50] for tumor mass effect simulation. I use the method proposed by Persson and Strang [76] to generate the tetrahedral mesh. The method has the advantage of having a relatively simple implementation and being generalizable to any number of dimensions (provided a corresponding Delaunay tessellation implementation). The method is composed of three steps: selection of points, tessellation of the points, and adjustment of the point locations. After the tessellation process, the points are adjusted so that edge lengths are optimal and so that edge lengths do not cross the external boundary or the internal structural boundaries. The edge lengths in the 3D tetrahedral mesh are optimal when they match a given distribution function. For an edge that connects two points x and y, the ideal edge length is proportional to the distribution function evaluated at the edge midpoint $z = \frac{x+y}{2}$. I have chosen to use the following distribution function,

$$f_{edge}(z) = \begin{cases} 1 & \text{if } \psi(z) > 1 \\ \psi(z) & \text{otherwise} \end{cases} \tag{5.2.5}$$

74

where $\psi(z)$ is the distance from z to the closest structural boundary. Assuming that the distance functions ψ_c for each class c is defined to be positive inside the relevant structure and zero otherwise, $\psi(z) = \min_c \psi_c(z)$. Using this function results in smaller tetrahedra near structural boundaries and larger tetrahedra in the internal regions. This behavior is desirable since the brain structures typically have complex and detailed shapes at the boundaries, which can be better interpolated using small-sized tetrahedra.

The displacement solution for the linear elastic model is computed by minimizing the potential energy,

$$E = \int_\Omega \frac{1}{2} \mathrm{Tr}[\epsilon^T \sigma] \, dx + \int_\Omega f_{ext}^T u \, dx. \tag{5.2.6}$$

Details are available in [**15, 41**]. The displacement solutions are computed by taking into account the boundary conditions [**5**]. The linearized growth process tends to result in slow deformations, so the model solution is computed iteratively and integrated until the volume of the expanded tumor exceeds a particular threshold.

5.2.2. Modification of Diffusion Tensors. Tumor infiltration and edema generally occur along white matter fibers, where diffusion is more likely. The properties of the white matter fiber within the brain is reflected in diffusion tensor MR images (DT-MRI). Since the BrainWeb datasets [**16, 4**] do not contain average diffusion tensor images, I generate average tensors from 5 normal subjects. The subjects are drawn at random from a dataset that contains 100 subjects, designed to study differences across age groups [**72**] (age range is approximately 18 - 74 years). I registered the 5 DT-MR images to the T2w image provided by BrainWeb by matching the associated mean diffusivity (MD) images to the T2w image using affine transformation and mutual information [**58**]. The tensors are mapped and reoriented following the finite strain reorientation strategy proposed by Alexander *et al.* [**1**].

In order to generate a realistic tumor mass effect, I simulate the expansion process. However, the inverse of the expansive deformation is required to resample the DT images. A true inverse may not exist since the expansive deformation may not be smooth and invertible, so I estimate the inverse of the deformation field using an iterative process.

Given a displacement field u that maps x to y, $y = x + u(x)$, I compute u^{-1} where $x = y + u^{-1}(y)$. This is done by estimating the inverse mapping for y, denoted by $\hat{x}$. The vector $\hat{x}$ must minimize the mapping residual

$$d(\hat{x}, y) = (\hat{x} + u(\hat{x})) - y \tag{5.2.7}$$

which expresses the distance between the forward map of the inverse estimate and the current location. The ideal value of $\hat{x}$ is computed by doing iterative minimizations (e.g., by Newton's method) at each location y in the target image. For an ideal inverse mapping, $d(\hat{x}, y)$ should be close to zero at any given y. The inverse displacement field u^{-1} is obtained directly from the $\hat{x}$ estimate, $u^{-1}(y) = \hat{x} - y$.

The average tensors at each voxel are computed using the efficient log-Euclidean tensor framework proposed by Arsigny *et al.* [2, 3]. Given image coordinate x, the average diffusion tensor is

$$D_\mu(x) = \text{Exp}\left(\sum_{i=1}^{N} \frac{1}{N}\text{Log}(D_i(x))\right) \tag{5.2.8}$$

where Exp is the matrix exponential function and Log is the matrix logarithm function. The Log function linearizes the space, allowing us to use efficient linear operations to manipulate the tensors. The Exp function maps the result of the linear operations back to the original space of diffusion tensors.

The fast tensor calculus proposed by Arsigny *et al.* [2, 3] is an approximation of the tensor calculus framework based on affine-invariant Riemannian metrics, such as the one proposed by Fletcher *et al.* [30]. The approach taken by Fletcher *et al.* treats the positive definite matrices representing the diffusion tensors as points on a manifold. The computation of distances and averages are done through linear approximations of the manifold (the tangent planes). Arsigny *et al.* use a single approximation for all the tangent planes, while Fletcher *et al.* explicitly compute the tangent planes for the relevant points. This turns the averaging process proposed by Fletcher *et al.* into an optimization process, while reducing the averaging process proposed by Arsigny *et al.*

into a closed form equation. The two approaches yield very similar results with regard to interpolation. Comparing the two methods, the determinant of the tensors are monotonically interpolated with no swelling in the tensors as commonly observed with trilinear interpolation. However, the approximation of Arsigny *et al.* yields tensor averages with trace that are always larger or equal to the affine-invariant counterpart, and the result is generally more anisotropic [3].

I have found that registration and reorientation of DT-MRI may not be sufficient to generate edema that appears realistic. White matter fibers around tumor tend to be displaced, and as observed by Lu *et al.* [57] in regions near the tumor, the mean diffusivity (MD) tends to be decreased while the fractional anisotropy (FA) tends to be decreased. These observations can be attributed to the destruction of white matter fibers due to tumor growth, which makes tensors more isotropic. Therefore, it was desirable to reflect this destruction in the simulator. To the best of my knowledge, the interactions between tumor growth and diffusion tensors are not fully understood, so I make the following assumptions:

(1) Local volume expansion reduces tensor coherence and results in more isotropic tensors. Tumor tends to destroy white matter fibers, so water is no longer restricted to flow in specific directions.

(2) Local volume compression or shrinking does not modify tensor information. I have observed that in real tumor DT-MRI some fibers can appear condensed without being destroyed.

The influence of pure tumor mass effect on DT-MRI is modeled using a combination of image warping and nonlinear interpolation. The displacement of white matter fibers is simulated by warping the DT-MRI following the strategy described in [1], where a rigid rotation is applied to each individual tensor. The rigid rotation is computed based on the local warping property. Given the displacement field u, I compute the local affine transform $F = I_{3 \times 3} + \nabla u$. This transform is decomposed into a rigid rotation component R and a linear deformation component W, $F = RW$. The reoriented tensor D is obtained

using the following equation

$$D = RD_0R^T \tag{5.2.9}$$

where D_0 is the resampled original tensor. The destruction of tensor information is modeled as a nonlinear interpolation between the original tensor and the isotropic version of the tensor. The isotropic version of a given tensor D is formulated as the identity matrix multiplied by the scaled determinant value of the original tensor

$$D_{iso} = (2|D|)^{\frac{1}{3}}I_{3\times3}. \tag{5.2.10}$$

A scale factor of 2 is used for the tensor determinant. This value is found by repeated experiments to find sufficiently realistic looking MD images. The transformed diffusion tensor is computed as follows (see [2] for details):

$$D'(x) = \text{Exp}\left(\alpha\text{Log}(D(x)) + (1-\alpha)\text{Log}(D_{iso}(x))\right). \tag{5.2.11}$$

The interpolation weight α is inversely proportional to the amount of volume expansion

$$\alpha(x) = exp\left(-\frac{[max(1,|J(x)|)-1]^2}{2s_J^2}\right) \tag{5.2.12}$$

where J is the Jacobian matrix of coordinate mapping function and s_J reflects the amount of expansion that results in significant destruction of fibers. In regions with a high amount of volume expansion (low values of α), the tensors become homogenized and no longer have preferred directions for diffusion. In regions with local volume compression, the determinant of the Jacobian is less than one and the original tensor is maintained since $\alpha = 1$. This behavior is chosen to simulate destruction of white matter fibers due to expansive mass effect while ignoring compression effects. An example application of this model to a registered DT-MRI is shown in Figure 5.4.

5.2.3. Tumor Infiltration and Edema. In order to simulate the growth and spreading of tumor cells following the preferred diffusion directions in the brain, the spatial probability that a particular location is infiltrated by pathological cells or fluid (edema)

78
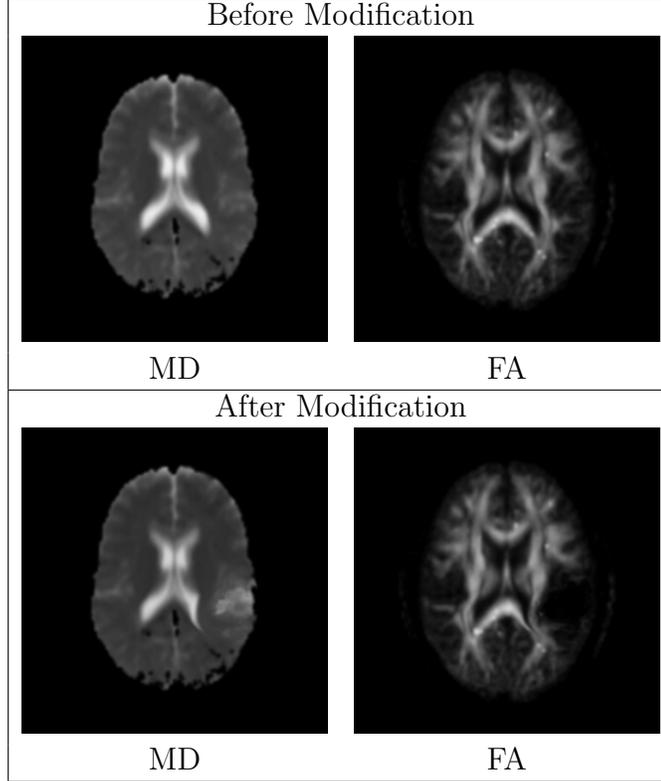
FIGURE 5.4. Visualization of diffusion tensor MRI by axial views of 3D Mean Diffusivity (MD) and Fractional Anisotropy (FA) scalar images. The modified DT-MRI has higher MD and lower FA in the regions surrounding tumor, which models the destruction of the fibers. The MD image shows that the ventricle near the tumor is slightly deformed. The FA image shows that the white matter fibers near the tumor region are pushed away.

is evolved using a reaction-diffusion model guided by the modified DT-MRI [15]. More precisely, the change for $p_{infiltrated} = \phi$ in time is governed by

$$\frac{\partial \phi}{\partial t} = \text{div}(c_d D' \nabla \phi) + c_r \phi \tag{5.2.13}$$

where c_d is the diffusion rate, D' is the diffusion tensor that has been modified using the method described in section 5.2.2, and c_r is the reaction rate. The diffusion rate c_d depends on the local tissue type. White matter is more likely to be infiltrated than gray matter, while csf is not likely to be infiltrated at all. The reaction rate or the growth term c_r is a constant. The diffusion tensors D' are normalized so that the trace of each tensor

79

is within the range of $[0, 1]$. The evolution is stopped when the volume of infiltrated brain regions exceed a predefined fraction of the brain tissue volume.

The infiltrating tumor cells and edema also tend to displace nearby tissue. The effect of the infiltrating cells is modeled using the following equilibrium equation [15]:

$$\text{div}(\sigma - \lambda I_{3\times 3}\phi) + f_{ext} = 0 \tag{5.2.14}$$

where λ is the coupling factor that describes the contribution of an infiltrating tumor to the internal forces. The equilibrium equation can be interpreted as the application of body forces $-\lambda\nabla\phi$ to the classic linear elastic model, which models the outward forces proportional to the concentration of tumor cells.

Brain tissue can be infiltrated by edema (swelling) and/or tumor cells. Since edema regions can also contain tumor cells, it is difficult to classify or separate the infiltrating component into distinct tumor and edema regions. The separation is approximated by assigning the regions formed early in the infiltration process as tumor and assigning the regions formed later in the infiltration process as edema. This approximation yields the following spatial probability functions for tumor and edema:

$$
\begin{aligned}
p_{tumor}(x) &= p_{mass-effect}(x) \\
&\quad + [\phi(x, t_{early}) \times p_{tissue}(x)]
\end{aligned}
\tag{5.2.15}
$$

$$
p_{edema}(x) = [\phi(x, t_{final}) - \phi(x, t_{early})] \times p_{tissue}(x)
\tag{5.2.16}
$$

where $p_{mass-effect}$ is the deformed initial tumor probability according to the mass effect model, p_{tissue} is the probability of brain tissue (white matter or gray matter), t_{final} is the time where the infiltration process is stopped, and t_{early} is a fraction of the total time that indicates when edema begins to occur. The choice for the value of t_{early} depends on the type of tumor being modeled. For example, an appropriate model for gliomas would typically have a large value for t_{early} since active tumor cells in gliomas tend to infiltrate large regions. Alternatively, an appropriate model for meningiomas with large

surrounding edema would involve a small value for t_{early} since most of the infiltrating regions should be attributed as edema.

In summary, the simulation of the pathological effects of brain tumor and edema is performed through the following steps:

(1) Manual definition of an initial tumor seed region in the space of a healthy ground truth data (e.g., the BrainWeb data).

(2) Simulation of deformation of brain tissue due to tumor mass effect, given the anatomical description (the BrainWeb classification), the initial seed region, and the constant pressure value P at the tumor surface. The deformation is modeled as a linear biomechanical equation and computed iteratively to mimic the possible non-linear deformations.

(3) Warping the average diffusion tensor MR images using the tumor mass effect displacements. Destruction of white matter fibers due to tumor is simulated by making tensors more isotropic depending on the magnitude of deformation.

(4) Simulation of tissue infiltration using the DT-MRI guided reaction-diffusion equation (Equation 5.2.13) to account for infiltration of tissue by tumor cells and edema.

(5) Simulation of deformation of brain tissue due to edema mass effect, given the infiltration probabilities and the value of λ for Equation 5.2.14.

5.3. Generation of MR Images

For the purpose of validating segmentation methods, there is a need for a set of synthetic MR images that corresponds to the pathological ground truth. These images function as test data for the segmentation methods to be validated. The generation of synthetic tumor MRI involves the simulation of two processes: contrast enhancement in T1w MRI due to the use of contrast agents (the standard modality for tumor diagnosis), and generation of intensity patterns similar to those observed in real MRI. These are detailed in the following two subsections. Contrast enhancement is simulated using a model of the contrast agent accumulation process, while the generation of MRI intensity patterns is accomplished using texture synthesis. A conceptual view of the combination of the two processes for generating a synthetic contrast enhanced T1w image is shown in Figure 5.5.
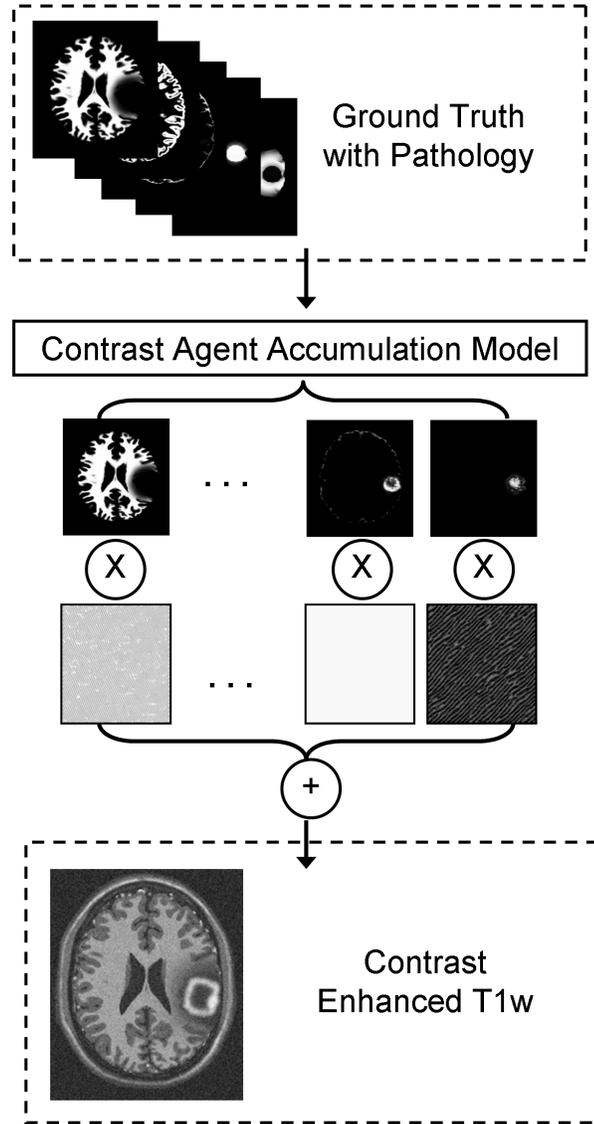
FIGURE 5.5. Generation of a synthetic contrast enhanced T1w image. The modified ground truth (the probabilities for white matter, gray matter, csf, tumor, and edema) is first used to determine where contrast agent is likely to accumulate. This is then followed by a combination of synthesized textures modulated by the spatial probabilities. In this figure I only show the probabilities and textures for white matter, non-highlighted tumor, and highlighted tumor or csf regions. For generating the T1w and T2w modalities without contrast enhancement, the contrast agent accumulation is not simulated and the texture combination is done directly using the modified ground truth.

5.3.1. Contrast Agent Accumulation. One of the particular challenges in segmenting brain tumor MRI are inconsistencies in the contrast enhanced T1w image, which can be attributed to biological processes such as tumor formation, blood flow, and cell death. The contrast agent is generally accumulated in regions other than the active tumor regions. Particularly, the blood vessels within the brain are almost always enhanced. Brain tissue may also appear enhanced if there is leakage of contrast agent due to the breakdown of the blood-brain barrier. Conversely, the contrast agent does not accumulate in the necrotic parts of the tumor at all. The necrotic regions are generally found in the core tumor regions.

I explicitly model the accumulation of the contrast agent in active tumor tissue and blood vessels in order to generate non-homogeneous contrast enhanced T1w images that are more challenging to segment. The spatial probability for the accumulation of contrast agent, $p_{accum} = \gamma$, is evolved using a reaction-diffusion equation that models the spread of contrast agent within blood vessel and tumor regions while excluding necrotic regions:

$$
\begin{aligned}
\frac{\partial \gamma}{\partial t} \;=\;& \operatorname{div}(a_d \, \nabla \gamma) \\
& + a_{source} \, I\{x \in X_{source}\} \, \gamma \\
& - a_{sink} \, I\{x \in X_{sink}\} \, \gamma.
\end{aligned}
\tag{5.3.1}
$$

Here, each I is an indicator function, a_d is the diffusion rate for the contrast agent, a_{source} is the source coefficient, and a_{sink} is the sink coefficient. The value of a_d depends on the structure type at location x. I assign high values of a_d in blood vessel regions, moderate values of a_d in tumor tissue, and low values of a_d in healthy tissue. The selection of the values of a_d for each class models the fact that contrast agent is more likely to spread in blood vessel regions than in tumor tissue and is not likely spread to healthy tissue at all. This corresponds to the actual biological process, where contrast agent is injected intravenously and then transported to the active tumor regions through the brain arteries. Healthy brain tissue generally does not accumulate contrast agent due to the blood-brain

barrier. Conversely, there tends to be a higher uptake of the contrast agent within and around tumor structures due to increased tumor metabolism and possible leakage of the blood-brain barrier.

X_{source} and X_{sink} in Equation 5.3.1 are sets of points that act as sources or sinks, respectively. The source points X_{source} are chosen at random from a probability function that indicates likely blood vessel regions or likely tumor regions that are close to the tumor boundary. These source regions correspond to regions that likely accumulate contrast agent and thus appear enhanced. The sink points X_{sink} are chosen at random from a probability function that indicates likely tumor regions that are close to the tumor core. The internal tumor regions are typically necrotic and thus do not accumulate contrast agent. The probability that a location is at the boundary or the core regions is computed using the distance maps and expressed as half-normal distributions. For example, when drawing points that are at the tumor border the following probability function is used:

$$p_{tumor-border}(x) = p_{tumor}(x) \times \mathcal{HN}(\psi_{tumor}(x), \omega) \tag{5.3.2}$$

where $\psi_{tumor}(x)$ indicates the distance from a location x to the nearest tumor boundary point. $\mathcal{HN}(z, \omega)$ denotes the half-normal distribution with parameter ω, which is defined as follows:

$$\mathcal{HN}(z, \omega) = \begin{cases} 0 & \text{if } z < 0 \\ \frac{2\omega}{\pi} exp\left(-\frac{z^2 \omega^2}{\pi}\right) & \text{if } z \geq 0 \,. \end{cases} \tag{5.3.3}$$

For a parameter value ω, $\mathcal{HN}(z, \omega)$ is a distribution with mean $\frac{1}{\omega}$ and variance $\frac{\pi-2}{2\omega^2}$. The border extent of the active tumor region or the standard deviation for $\mathcal{HN}(\psi_{tumor}(x), \omega)$ is a user-specified parameter value that is inversely proportional to ω. Tumors with nearly uniform enhancement can be simulated by drawing source points from a uniform distribution (within tumor), as opposed to a half-normal distribution, while non-enhancing tumors can be simulated by replacing the tumor source points with an empty set.

I initialize γ so that tumor and blood vessel regions have random probability of accumulating contrast agent: $\gamma(x, t = 0) = \mathcal{U}(0, 1)$. The initialization using the random
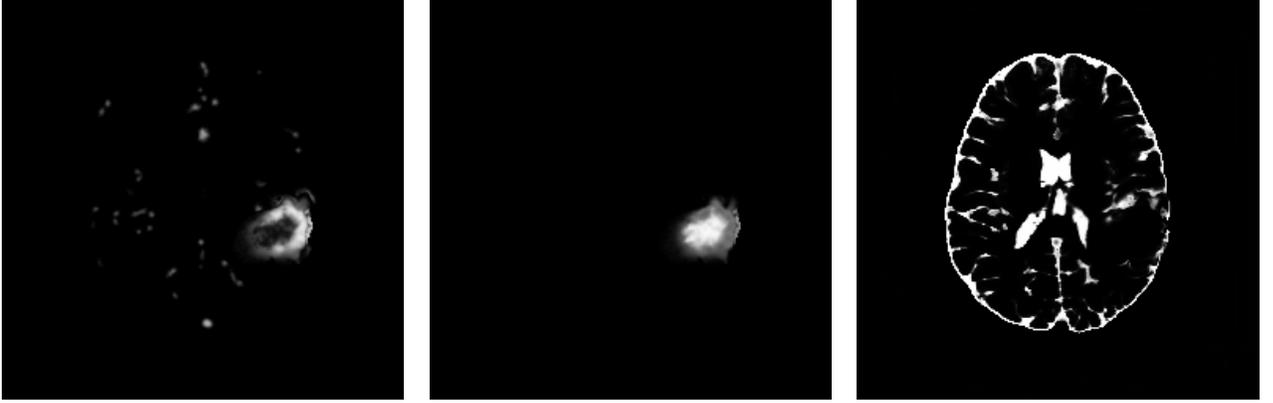
FIGURE 5.6. Axial view of the generated probabilities related to contrast enhancement for the SimTumor001 dataset. From left to right: probability for highlighted csf or tumor, probability for non-highlighted tumor, and probability for non-highlighted csf. Tumor periphery and blood vessel regions are the regions most likely to appear highlighted in the contrast enhanced T1w modality.

variables $\mathcal{U}(0,1)$ drawn from the uniform probability in $[0,1]$ ensures that the reaction-diffusion process is capable of generating complex patterns of enhancement. The probability that a location x would appear highlighted in the contrast enhanced T1w image is the probability that the structure in that location is either tumor or blood vessel and that it has accumulated contrast agent,

$$p_{enhanced}(x) = p_{accum}(x) \times [p_{vessel}(x) + p_{tumor}(x)]. \qquad (5.3.4)$$

Figure 5.6 shows an example of the generated contrast enhancement probabilities, while Figure 5.7 provides a comparison between real and synthetic contrast enhanced T1w MRI. The enhancement probabilities are generated using the method discussed in this subsection, and the synthetic MRI is generated using the method covered in the next subsection. The proposed contrast agent model accounts for the fact that blood vessel and active tumor regions are highlighted and that the necrosis regions are not highlighted. However, the model only accounts for the deformation of healthy blood vessels and does not account for the fact that new blood vessels can be formed due to the presence of tumor.
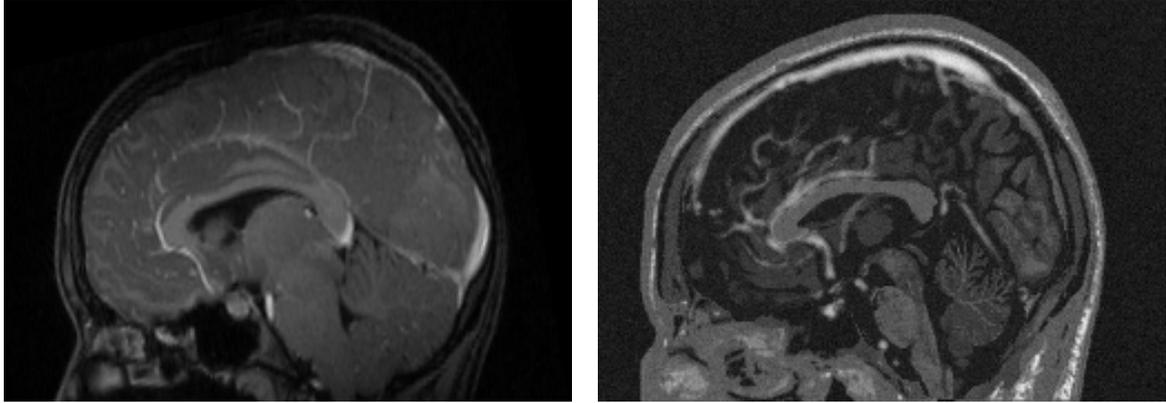
FIGURE 5.7. Sagittal view of the contrast enhanced T1w MRI for a real tumor (left) and a synthetic tumor (right) generated using the new method described in this chapter. Both images show contrast enhancement in the superior sagittal sinus and the anterior cerebral artery.
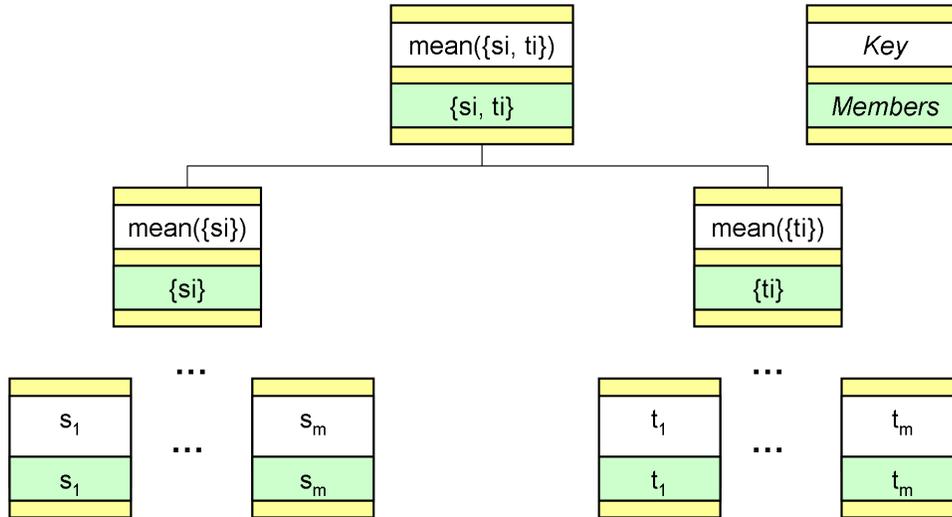
FIGURE 5.8. Tree Structured Vector Quantization using a binary tree data structure. Each node in the tree represents a cluster, which contains a set of data samples and is represented by a key (the data mean). The child nodes of the parent node are formed by splitting the clusters represented by the parent, typically done using the k-means algorithm.

5.3.2. Texture Synthesis. I generate tumors that contain intensity patterns found in real tumor MRI using the texture synthesis algorithm proposed by Wei and Levoy [**99**]. This approach only relies on actual samples from actual tumor MRI scans and does not make restrictive assumptions on the intensity distributions. The algorithm starts with an image that contains random noise [**61**] and then proceeds to modify the image by finding neighborhood matches in the input texture. The neighborhood search is done deterministically across scales and is made efficient by clustering the texture neighborhood features. Rather than performing the search by comparing a random neighborhood with all the data samples, their method uses the Tree Structured Vector Quantization (TSVQ) technique to limit the search to the relevant clusters. An example application of TSVQ using the binary tree data structure is shown in Figure 5.8.

The synthetic MR images are generated by linearly combining the texture synthesis results for each structure. To simulate partial voluming and the ambiguity in the boundary, the textures are weighted by the fuzzy class probabilities. For a modality k, the

synthetic MR intensity for each location x is

$$I_k^{synth}(x) = \sum_{c=1}^{N_{C_k}} m_{c,k}\, p_c(x)\, T_{c,k}(x) + \mathcal{N}_{0,\sigma_k} \tag{5.3.5}$$

where c indexes the N_{C_k} different classes for the modality k. For the T1w and T2w modality, the set of brain structure classes is composed of white matter, gray matter, csf, tumor, and edema. For the contrast enhanced T1w modality, the set of brain structure classes is composed of white matter, gray matter, non-enhancing csf, non-enhancing tumor, edema, and the class for all contrast enhanced structures. The contrast between different classes is adjusted via the user-specified coefficients $m_{c,k}$, which are chosen to generate realistic-appearing MRIs. As an example, a higher value of $m_{1,k}$ is chosen for white matter (e.g., 1.2) and a lower value of $m_{2,k}$ is chosen for gray matter (e.g., 0.7) when generating T1w images with good white-gray matter difference. The probabilities from the pathological ground truth are represented by p_c. The images $T_{c,k}$ are generated using texture synthesis from actual tumor MRI samples. Noise in the image data is simulated using $\mathcal{N}_{0,\sigma_k}$, which is randomly generated from a normal distribution with zero mean and standard deviation σ_k that is voxelwise independent.

5.4. Results and Evaluation

I generated five examples of the synthetic MR images using the simulation method described in this chapter (labeled SimTumor001 - SimTumor005). Figure 5.9 shows an example of a peripheral tumor with significant mass effect and surrounding edema, as seen in real MRI and in the synthetic MRI (SimTumor001). In both the real and synthetic MRI there are significant deformations of the surrounding healthy tissue due to tumor and there are ambiguities in the definition of tumor-edema boundaries. The five synthetic brain tumor MRI datasets are shown in Figures 5.10 and 5.11. In all cases, tumor deforms other structures and edema infiltrates brain tissue. The contrast enhanced T1w images also show complex patterns of highlighted intensities, which include active regions at the tumor periphery and the blood vessel regions. SimTumor001 shows a tumor with significant mass effect and a large surrounding edema. SimTumor002 shows a tumor that displaces the right ventricle from below and a moderate extent of edema. SimTumor003 shows a tumor that displaces the falx cerebri. SimTumor004 shows a tumor that displaces the left ventricle from the internal regions. SimTumor005 shows a small tumor in the anterior region with nearly uniform enhancement. The associated ground truth for all cases are shown in Figures 5.12 and 5.13. The ground truth is represented as a set of spatial probability maps for tissue and pathology. This provides advantage over binary data or class membership data since this allows a validation scheme to use probabilistic statistical analysis rather than simple volume comparison.

In order to verify that the synthetic MRI and the ground truth matches human perception and high level knowledge, I performed a limited comparison of the tumor volumes. The tumor structures obtained from the ground truth were compared to the segmentations drawn by a human expert and the results of a user guided semi-automatic segmentation method using level set evolution [**40, 104**]. Following standard practice, the segmentations were primarily driven by the contrast enhanced T1w images. The measures used for comparison are the Dice similarity coefficient and the average surface distances, which are discussed in Appendix A. The volumetric values for tumor and
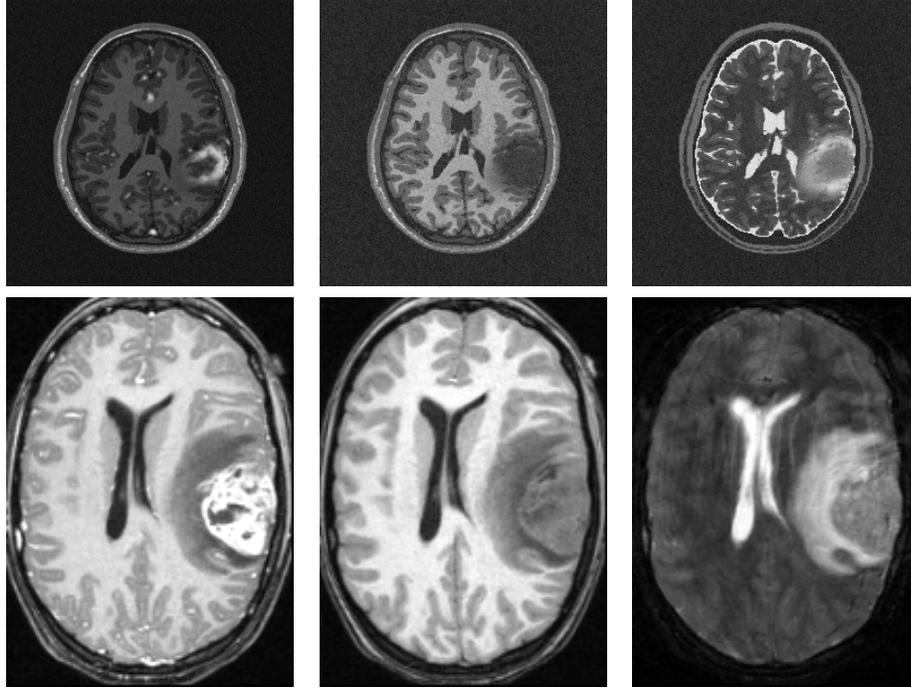
FIGURE 5.9. The MR images of the synthetic dataset SimTumor001 compared to real MR images of a subject with malignant tumor and surrounding edema. Top: axial view of the synthetic 3D MR images generated using the simulation method described in this chapter. Bottom: axial view of real 3D MR images. From left to right: contrast enhanced T1w, T1w, and T2w images.

edema are shown in Table 5.1. Volumes are measured as the integral of the spatial probabilities of the relevant structure. Table 5.2 shows the quantitative comparison results between the synthetic ground truth and the manually drawn segmentations, while Table 5.3 shows the quantitative comparison results between the synthetic ground truth and the semi automated segmentations. The surface distances differ less than $1.5\,mm$ in average, for both the comparison against the manual drawings and against the semi-automatic method. The difference between the user-guided segmentation results and the synthetic ground truth is mainly in the definition of the extent of tumor boundaries. The definition of tumor extent is generally ambiguous due to surrounding edema. Compared to the manual segmentation results, the semi-automatic segmentation results more closely resemble the simulated ground truth as the level set evolution generates more detailed tumor contours.
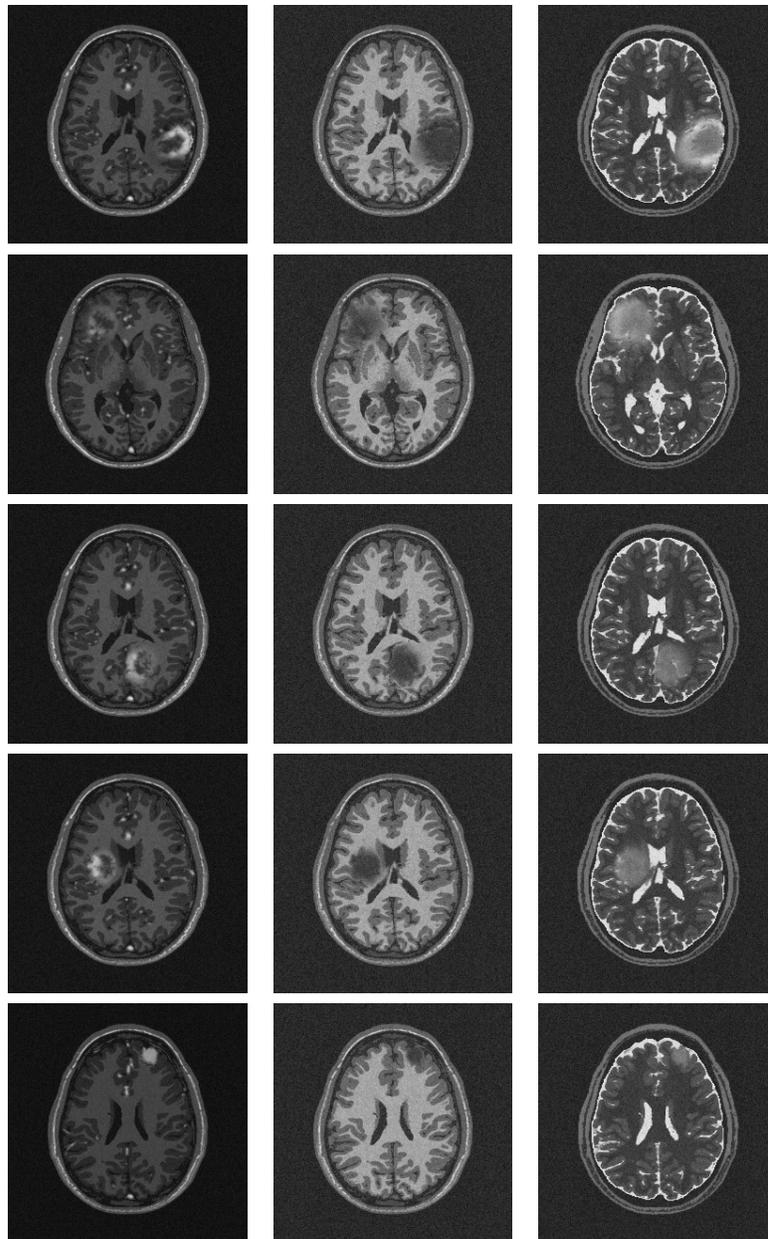
FIGURE 5.10. Axial views of the MR images of the synthetic datasets. From top to bottom: the SimTumor001, SimTumor002, SimTumor003, SimTumor004, and SimTumor005 MRI datasets. From left to right: contrast enhanced T1w, T1w, and T2w images.
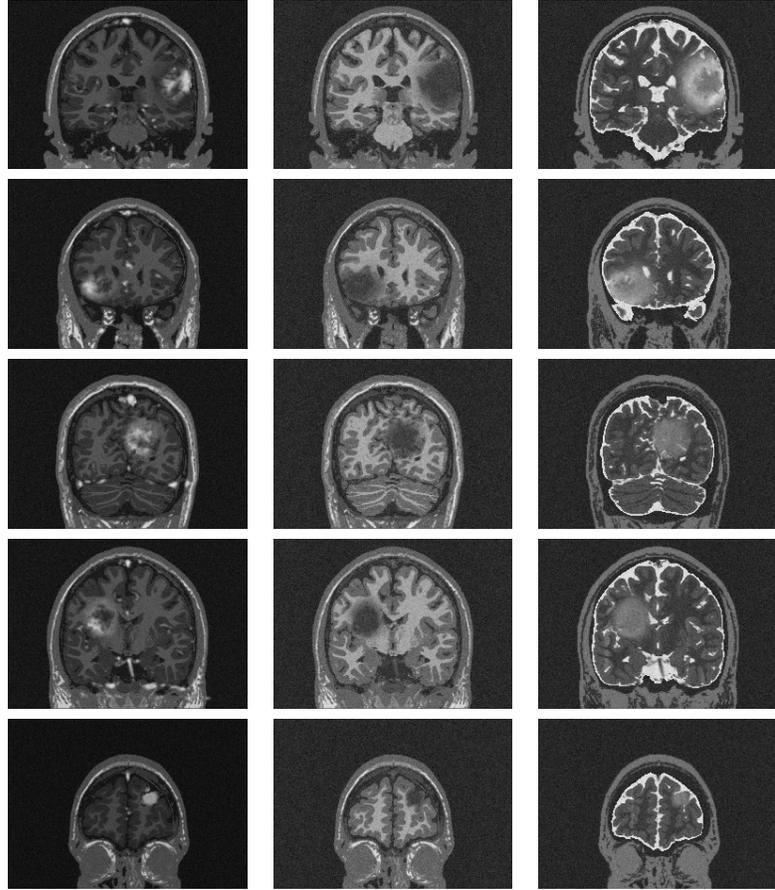
FIGURE 5.11. Coronal views of the MR images of the synthetic datasets. From top to bottom: the SimTumor001, SimTumor002, SimTumor003, SimTumor004, and SimTumor005 MRI datasets. From left to right: contrast enhanced T1w, T1w, and T2w images.
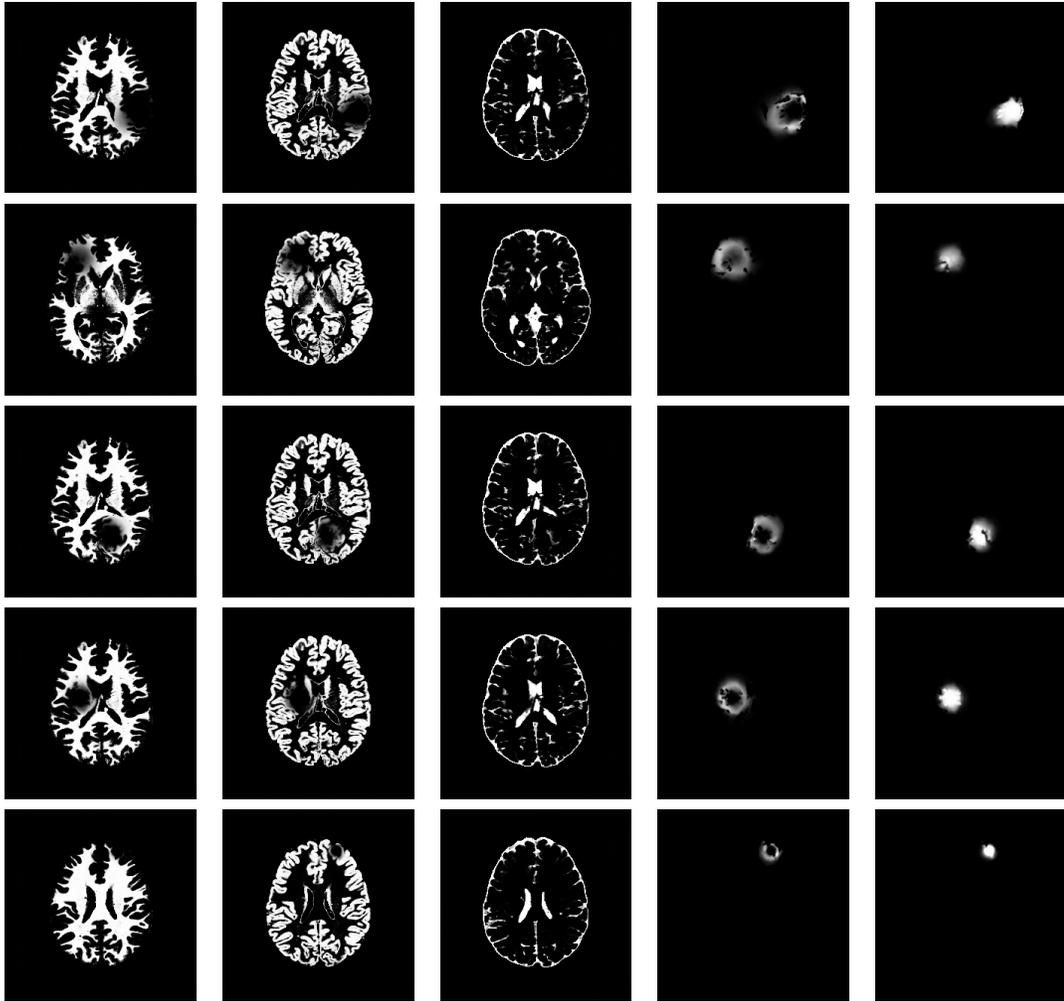
FIGURE 5.12. Axial views of the ground truth for the 3D synthetic brain tumor MRI data sets. From top to bottom: spatial probabilities for the SimTumor001, SimTumor002, SimTumor003, SimTumor004, and SimTumor005 datasets. From left to right: the class probabilities for white matter, gray matter, csf, tumor, and edema.
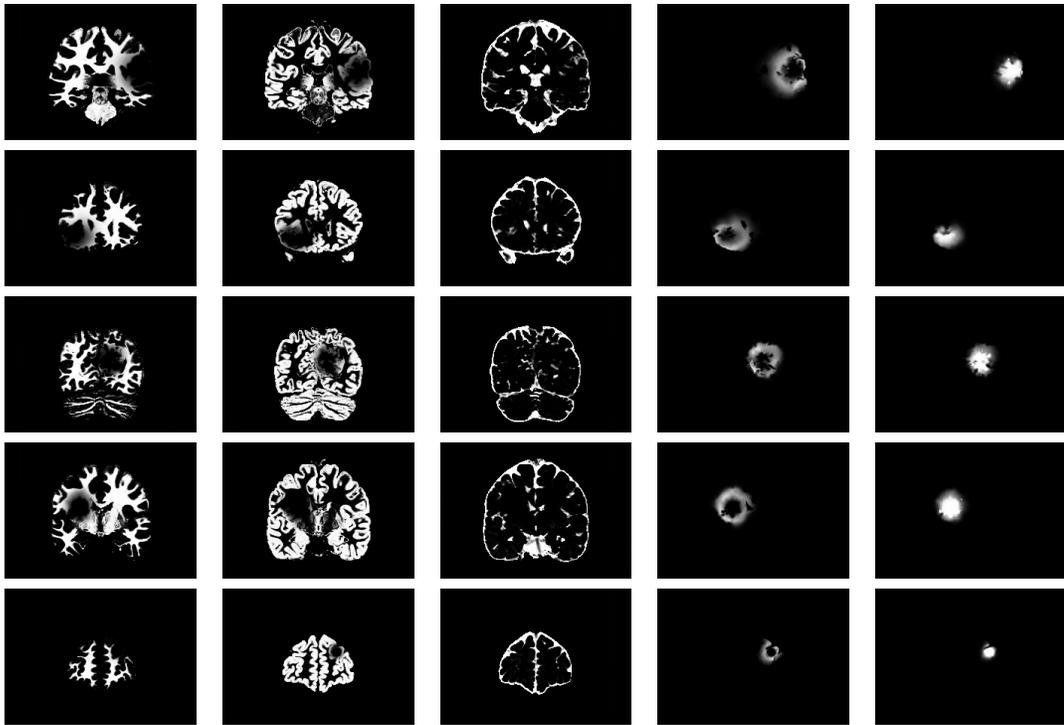
FIGURE 5.13. Coronal views of the ground truth for the 3D synthetic brain tumor MRI data sets. From top to bottom: spatial probabilities for the SimTumor001, SimTumor002, SimTumor003, SimTumor004, and SimTumor005 datasets. From left to right: the class probabilities for white matter, gray matter, csf, tumor, and edema.

Dataset	Tumor Volume (mm^3)	Edema Volume (mm^3)
SimTumor001	25659.3	36602.6
SimTumor002	21893.9	19208.2
SimTumor003	23379.6	3977.7
SimTumor004	19940.0	5925.7
SimTumor005	3113.1	195.8

TABLE 5.1. Volumes of the tumor and edema structures in the synthetic datasets.

Dataset	DSC (%)	Average surface distance (mm)
SimTumor001	84.3	1.261
SimTumor002	82.1	1.555
SimTumor003	77.4	1.752
SimTumor004	75.2	2.048
SimTumor005	72.4	1.137

TABLE 5.2. Comparison of the synthetic ground truth to the segmentations drawn by a human expert for the simulated brain tumor MRI datasets.

Dataset	DSC (%)	Average surface distance (mm)
SimTumor001	88.03	1.19
SimTumor002	84.66	1.566
SimTumor003	79.66	2.032
SimTumor004	81.79	1.681
SimTumor005	90.14	0.366

TABLE 5.3. Comparison of the synthetic ground truth to semi-automatic segmentation for the simulated brain tumor MRI datasets.

5.5. Conclusions

In this chapter, I have presented a new method for generating synthetic MR images with tumor and edema, together with the associated ground truth, from a healthy brain ground truth. The process for generating a synthetic brain tumor dataset is summarized in Figure 5.14. I have performed a limited validation by comparing the synthetic ground truth with the tumor segmentations done by human raters. The results verify that there is a satisfactory level of agreement between the tumors perceived within the synthetic MRI and the synthetic ground truth. The synthetic brain tumor MRI along with the associated ground truth provide the means for performing objective validation of different brain tumor MRI segmentation frameworks. Given a segmentation framework for brain tumor MRI, it can be tested using the synthetic multimodal brain tumor MRI as input images. We can then measure its performance by comparing the segmentation results and the synthetic ground truth. Compared to validation against manual segmentations, this approach has the advantage of having consistent, known ground truth for the whole brain. This capability is novel as most validations done so far were focused on tumor only and not performed on the infiltrated and deformed healthy tissue.

Brain tumor growth is a very complex process, and it is extremely challenging to account for all the variables that govern the process. One possible extension to the method proposed in this chapter is the simulation of the formation of new blood vessels (angiogenesis). Tumor cells are known to generate biological signals that induce formation of blood vessels to supply additional energy for the increased metabolism. Bullitt *et al.* [10] also observed that vessels in and around the tumor tend to have larger variability in the curve angles and become more tortuous. Simulating blood vessel formation and shape changes will allow for the generation of more realistic mass effect and infiltration models and improve the appearance of the synthetic contrast enhanced T1w image. The simulation of the deformation due to tumor mass effect could be improved by using more complex computational model such as the biphasic models proposed by Miga *et al.* [64, 65] and Nagashima *et al.* [74]. Another possible extension is a more detailed

modeling of the changes in csf volume and flow. If the intracranial pressure is high, there tends to be a loss of csf volume (which may not be restricted to the ventricles). If brain tumor blocks ventricular outflow, the csf volume can increase. Other possible extensions to the simulation framework and potential applications will be discussed in the future work section of Chapter 6.
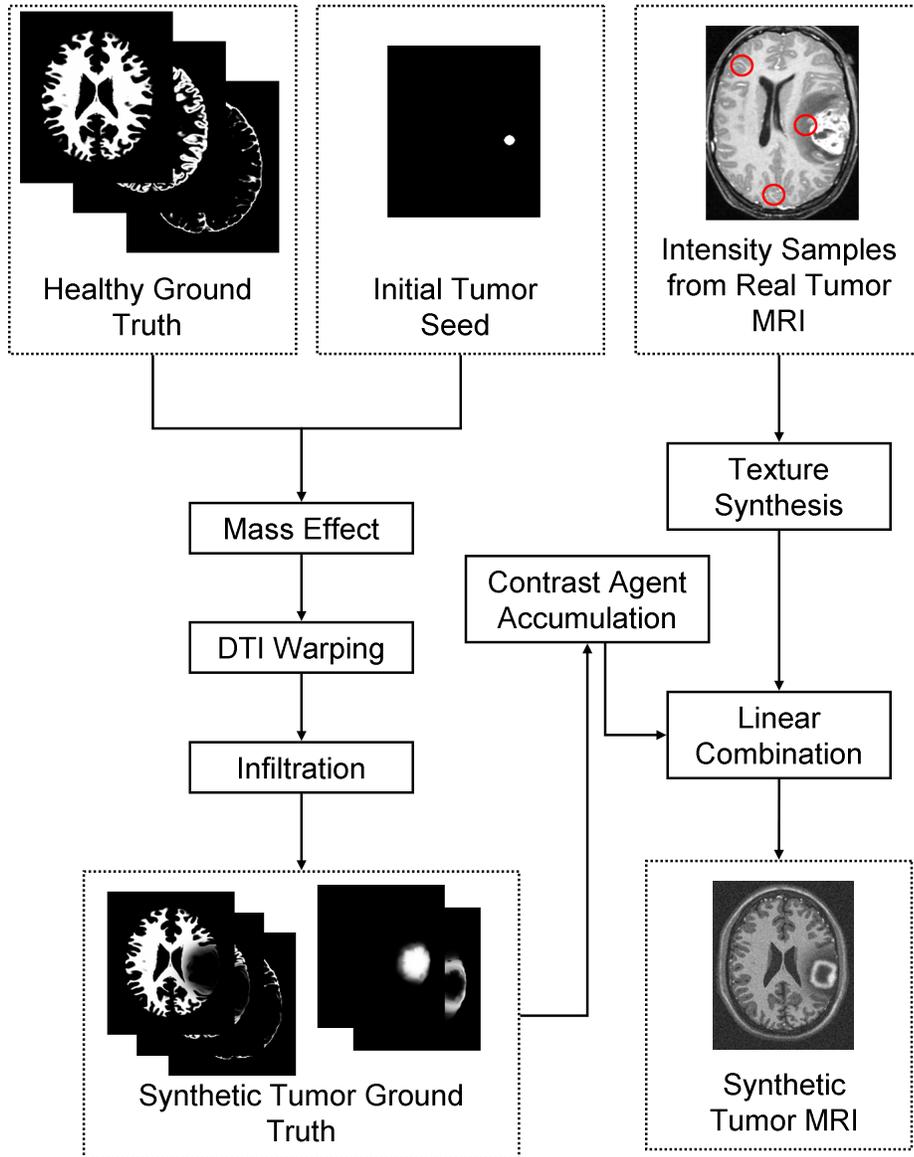
FIGURE 5.14. Summary of the generation of synthetic brain tumor ground truth together with the associated brain tumor MRI (here only the contrast enhanced T1 image is shown).

CHAPTER 6

Discussion and Future Work

This chapter reviews and discusses the contributions of this dissertation and presents some possibilities for future work. Section 6.1 reviews and discusses the list of contributions presented in the first chapter. Section 6.2 discusses future research possibilities and potential application areas. Section 6.3 concludes with a summary.

6.1. Review of Contributions

This section provides a summary of the contributions of this dissertation. Each contribution is presented along with a discussion on how it was accomplished in this dissertation.

(1) *Image segmentation using a modified Expectation-Maximization (EM) algorithm: the novelties of this approach are its use of robust parameter estimation techniques and its automatic detection of the feature space clusters for the mixture model.*

The method used for image segmentation is based on the Expectation-Maximization algorithm with Gaussian finite mixture model. Chapter 2 investigated the limitations of the standard EM algorithm and presented extensions that adapt the algorithm for cases with significant deviations from the reference brain atlas. These deviations typically cause the data samples to be contaminated with a large proportion of outliers. This necessitated and motivated the use of robust combinatorial parameter estimation methods such as the Minimum Covariance Determinant (MCD) estimator for determining the proper parameters for the image likelihood $p(I|S, \theta)$.

Chapters 3 and 4 presented the adaptations of the robust EM algorithm for the segmentation of MRI of adult brains with tumor and of newborn brains. The highly variable nature of pathology can include the appearance of new structures such as edema. Edema, which is a swelling of the tissue, often appear around brain tumor. However, certain types of brain tumor (e.g., gliomas) are not associated with significant edema growth. To account for the possibility of the existence of additional structures (hence additional components in the mixture model), the EM algorithm is modified to perform a search for the existence of additional clusters in the intensity feature space based on prior knowledge on the intensity properties. To my knowledge, the brain tumor segmentation method described in Chapter 3 is the first fully automatic segmentation system that provides 3D segmentations of the whole brain tissue, tumor, and edema.

(2) *Generation of an augmented feature space for image segmentation through the use of spatial constraints such as location, curvature, and adjacency.*

The segmentation method uses samples obtain from the brain atlas. Since the subject images have significant deviations from the atlas, there is a need to restrict the sampling regions. Chapters 3 and 4 presented the strategies for obtaining samples based on known properties related to the structures and image geometry for specific anatomical regions. For brain tumor segmentation, tumor samples are restricted to non-spurious regions and edema samples are restricted to regions that are adjacent to tumor. For newborn brain segmentation, white matter samples are restricted to regions with low image gradient magnitude so that the ambiguous transitional white matter regions are excluded.

(3) *Application of the proposed segmentation framework for healthy brains as well as images that exhibit deviations due to pathology (brain tumor) and growth (newborn brains).*

In Chapters 3 and 4, the segmentation method was applied to real brain tumor MRI and newborn brain MRI. Validation of the segmented structures was
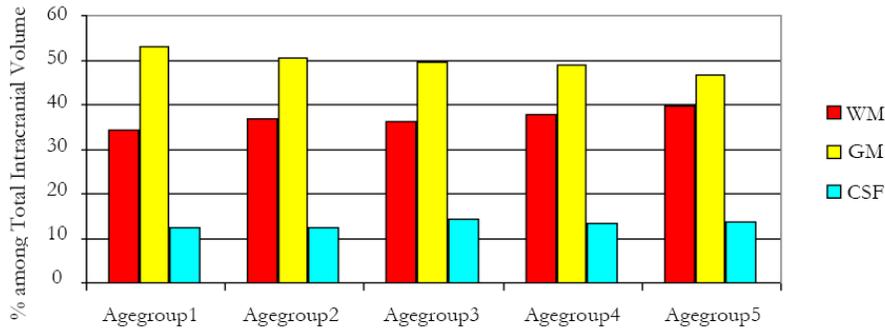
FIGURE 6.1. Volumes of brain structures for healthy subjects across age groups (divided by decades from young to old). The gray matter volume tend to decrease with age, which is in agreement with previously published results. Image provided by Bénédicte Mortamet [**72**].

performed by comparing the automatic results against manually obtained results. The automatic segmentation framework performed with satisfactory level of agreement to manual raters, with the advantage of being fully reproducible. The proposed segmentation framework has been used in several clinical studies [**72, 35**]. These studies reported facts that agree with the knowledge of the biological brain trends.

Segmentation tools based on the techniques described in Chapter 2 is currently in active use at multiple image analysis research laboratories. The segmentation tool developed using the atlas-based robust EM methodology has been applied to over 2000 adult subjects in UNC and Duke. It has been applied in a validation study with a traveling phantom on 10 scanners across the United States, and a separate validation study with 60 controls scanned twice. The tool is a critical component of a study on structural volumetric changes within the brain related to aging, and it has been applied succesfully to approximately 100 subjects with ages ranging from 18-74 years. The structural volumes for different brain tissue types for the different age groups are shown in Figure 6.1. In addition to UNC and Duke, the tool is in active use at Helsinki (Finland), Lausanne (Switzerland), and Liége (Belgium). It has also been extended to segment brains of children at one and two years of age.

The newborn brain MRI segmentation method (Chapter 4) has been applied to healthy subjects in UNC Department of Psychiatry and prematurely born subjects in the University Hospital of Geneva. In Section 4.4, I described a study that made use of the newborn brain MRI segmentation tool [35]. This study analyzed 74 newborns and found interesting clinical observations on the early growth patterns for the brain. The results of that study show that there is a characteristic pattern of regional brain structures that are distinct from those of adults and older children. That study presents the analysis of the largest collection of newborn datasets to date, which was made possible through the use of the new automated segmentation method.

(4) *A method for generating pathological ground truth (tumor and edema) from image data with known healthy ground truth by combining a linear elastic biomechanical model with random surface tractions and a reaction-diffusion process guided by diffusion tensor imaging (DTI). The simulation of a new pathological ground truth is guided by the underlying biological processes.*

Chapter 5 described a method for simulating brain tumor and edema given a healthy ground truth. Brain tissue is modeled using the classic linear elastic biomechanical model which are deformed using tumor surface forces. The force directions are drawn randomly from the von Mises-Fisher [60] distribution so that the final tumor shapes are more variable and more challenging to segment. Infiltration due to tumor and edema is simulated through a reaction-diffusion process which is guided by a diffusion tensor field. The diffusion tensor field that guides the reaction-diffusion is obtained from an average diffusion tensor of multiple subjects that is modified to account for tumor effects. The tensors are warped to simulate displacement due to tumor mass effect and made isotropic to simulate the destruction of white matter fibers.

(5) *Simulation of the accumulation of contrast agent for a brain tumor subject to generate contrast enhanced T1w MRI, which is the standard diagnostic imaging*

103

modality. The accumulation model is guided by the underlying biological processes.

The contrast enhanced T1w MRI modality is the current standard channel for tumor diagnosis. The simulation of the contrast enhancement due to the contrast agent (typically gadolinium) that is injected into the blood stream is proposed in Chapter 5. The infusion and spread of the contrast agent is simulated using a reaction-diffusion process which has a source and a sink term for the reaction. The source term indicates the regions where the contrast agent would be accumulated. This term is distributed in specific regions within the brain, particularly the blood vessels and the peripheral tumor regions. The choice of the regions are drawn at random from a spatial probability distribution that provides high probability values for the targeted regions. The sink term is similarly drawn at random from a spatial probability distribution that assigns high probability values in the internal tumor regions (the likely necrotic regions), where the concentration of contrast agent is likely to be low.

(6) *Simulation of MR images with brain tumor and edema using textures synthesized from real tumor MRI samples. The synthetic MR images and the associated ground truth provides the means for objective evaluation of different segmentation schemes.*

Chapter 5 discussed the process of generating test data for comparing brain tumor segmentation methods. Rather than simulating the complex MRI signal acquisition process, I proposed to use the linear combination of textures synthesized using an approximate Markovian model [**99**]. The textures are synthesized from samples that are obtained from real brain tumor MRI, so that the appearance of individual structures has the same relative property as a real scan. Each texture's contribution to a voxel is weighted using the class posterior probability, which mimics the partial voluming within real MR images.

The generated synthetic MR images along with the ground truth provides an objective means for evaluating different segmentation methods. As shown in Chapter 5, the synthetic images appear reasonably realistic and the synthetic ground truth is comparable to manual segmentations of the simulated images. Using the simulated images with known underlying anatomy, results from multiple methods can be compared to the ground truth without a need to declare a single manual segmentation as a gold standard. The synthetic data can also be used to evaluate the accuracy of various deformable registration schemes which are often designed to map tumor data to a healthy template. The mapping can be tested objectively and accurately since both the healthy and pathological anatomy are known.

6.2. Future Work

The tasks of segmenting images and evaluating segmentation results are challenging and involve different factors that might be domain specific. Therefore, this dissertation has only explored some of the issues specific to brain tumor and newborn brain segmentations. Possible future directions of research are listed in the following two subsections. The first subsection covers possible extensions for the automatic segmentation framework, while the second subsection covers possible extensions for the brain tumor simulation framework.

6.2.1. Segmentation of Brain MRI. Brain MR images present significant challenges for segmentation yet there is a wealth of prior knowledge that could be used for analysis. There are also many possibilities for using more advanced models of image appearance and anatomical brain structures. The combination of these factors might lead to various future directions for research in image segmentation.

(1) *Maximum a Posteriori image segmentation.* In Chapter 2, the image segmentation framework was developed from the ML segmentation framework combined with robust combinatorial methods for finding the maximum likelihood estimate.

The framework could be further improved by making use of the more flexible Maximum a Posteriori (MAP) segmentation framework that is summarized in Appendix B. This will make it possible to use the more sophisticated image models used in the framework proposed by Tu and Zhu [**91**]. Their framework unifies many different standard image segmentation techniques such as level set evolution, maximum likelihood estimation, and region merging and splitting.

The MAP segmentation framework estimates the segmentation parameters using the Markov Chain Monte Carlo (MCMC) algorithm, as opposed to the EM algorithm for ML segmentation. The combinatorial robust parameter estimation techniques described in Chapter 2 can be adapted to the MCMC technique to create proposal probabilities from the data that explicitly excludes some outliers. Combining the robust parameter estimation with the richer set of segmentation operations is likely to generate an improved method that can be used to segment a wider variety of brain MRI subjects.

(2) *Joint estimation of deformation and label assignments.* A potential issue that was not handled by the proposed method is large deformation of brain structures [**54**]. When there is large deformation (e.g., when tumor is present), using the brain atlas may lead to incorrect sampling. In this case, the atlas-based samples would be severely contaminated and the model estimation may yield incorrect results. The improper spatial priors would also limit the segmentation quality, as the segmentation output cannot differ greatly from the atlas. The proposed robust EM method can still handle some level of deformation due to the use of robust estimators that deal with outliers, but having an explicit model of the image deformation would lead to better segmentations.

In such cases, the atlas priors $Pr(S_k = c)$ need to be deformed to match the subject MRI adding an extra parameter to the priors. The segmentation method will need to estimate the best value for the label assignment S, image parameter θ, and the deformation for $Pr(S) = Z$. The deformed spatial priors

Z become a new parameter that needs to be estimated. Estimating the best estimate for the three values using a joint probability would be computationally difficult. A possible approach is to follow the approach using partial maximizations similar to the Iterative Conditional Modes (ICM) algorithm proposed by Besag [8]. With this approach, we iteratively compute the best estimate of an individual parameter set while holding the rest fixed. For example, we find the best label assignments S while holding θ and Z fixed in one iteration, and we estimate the deformation for Z while holding S and θ fixed in the next iteration. In different iterations, $Pr(S) = Z$ is treated either as a spatial prior or as a parameter. In fact, Neal and Hinton[75] proposed that the EM algorithm can be viewed as iterative partial maximizations of a function $F(p, \theta)$. During the nth iteration, the algorithm computes $p^{(n)}$ that maximizes $F(p, \theta^{(n-1)})$ in the E-step, and computes $\theta^{(n)}$ that maximizes $F(p^{(n)}, \theta)$ in the M-step. The computation of the deformation for the priors Z can be seen as an extension where we optimize $F(p, \theta, Z)$.

The theoretical impact of deforming $Pr(S) = Z$ in an EM-based approach is unclear. Deforming Z would have impact in both the E-step and the M-step, where Z acts as the spatial prior during the E-step and a parameter to be optimized during the M-step. In the E-step, the optimized lower bound for the image likelihood is formed by using a distribution that has the smallest Kullback-Leibler (KL) divergence to the class posterior $p(S|I, \theta)$, which is the class posterior itself (Section 2.2). Modifying the prior Z will also modify the class posterior, so the lower bound will be inaccurate if the KL divergence to the true class posterior becomes too large. In the M-step, the deformation of Z is treated as a parameter that needs to be optimized so that the priors match the observed image data. The ideal matching criteria still needs to be investigated. In particular, the estimation of deformation is not straightforward for images with pathology. This is due to the fact that we need to map priors that do not

107

contain a pathological structure to images that contain a pathological structure. The creation of new structures is a challenge in designing reliable matching criteria for the deformation.

(3) *Image segmentation using higher level features.* The segmentation method presented in this dissertation detects abnormal regions in the brain based on the atlas and image intensities. Other properties can also be used for this process; these can include geometric properties such as curvature or brain asymmetry [45], or blood vessel information from MR angiography. For example, brain tumor typically generates gross brain asymmetry, and knowledge of blood vessel information could help to isolate active tumor regions as tumors are typically associated with the formation of new arterial branches. More sophisticated robust parameter estimation schemes compared to the ones described in this dissertation may be necessary for such features. The schemes proposed in this dissertation assume that the data samples can be approximated with a Gaussian or that they exist in a Euclidean metric space.

(4) *Statistical models for pathological structures.* Although a structure like brain tumor can appear in many different sizes and shapes, an explicit statistical model could be useful for improving the accuracy of the segmentation. A model that describes the likely shape and the deviations from it could function as a mechanism to constrain the search for the optimal solution. An example of such a statistical shape model is the Principal Geodesic Analysis of the m-rep medial shape model proposed by Fletcher *et al.* [30].

Another approach, proposed by Mohamed *et al.* [70], uses a statistical model of deformations outside of the tumor regions to isolate the likely biomechanical model for the tumor and edema. This approach could be combined with the segmentation approach proposed in this dissertation to further isolate likely tumor and edema regions.

(5) *Diffusion tensor imaging (DTI) for image segmentation.* Diffusion tensor imaging is an interesting new modality that provides higher level structural information, namely the likely directions of water diffusion. Lenglet *et al.* [**52**] proposed a DTI segmentation framework that has been applied for brain white matter by evolving a surface based on the tensor data. A similar approach, or one based on probabilistic pattern recognition techniques, could be applied to obtain additional feature information that would help in isolating the relevant structures. Since DTI allows the derivation of the local white matter fiber structures, it may be particularly useful in segmenting anatomical deviations. In the case of brain tumors, white matter fibers tend to be displaced or destroyed around a brain tumor which indicates that the deformation of white matter fibers might be a vital cue for the presence of tumor. In the case of newborn brains, the knowledge of white matter fiber locations could help in isolating myelination regions since the process of myelination occurs along the white matter fibers and alters diffusion properties.

(6) *Image indexing and retrieval.* The segmentation framework presented in this dissertation is fully automatic and thus can be consistently applied to a large set of MR images. This could be of significant value in the indexing of images based on some clinical criteria (e.g., pathology, growth). For example, the framework can be combined with the brain indexing system proposed by Liu *et al.* [**53**] to retrieve images with clinically relevant features such as existence of brain tumors, location of tumors, tumor volumes, etc.

(7) *Extension to other anatomical regions.* The segmentation of other structures than the brain using MRI or other modalities such as CT (Computed Tomography) is another possible extension of the robust maximum likelihood segmentation framework. However, the methods described in this dissertation are focused on the brain only with its well defined structural properties. Additionally, the brain structure is encased in the skull, which is generally fixed or immobile. The

skull provides a strong cue for the registration process between brains. These two factors make it challenging to adapt the method to other anatomical regions. Extending this method to other anatomical regions with fewer motion constraints as compared to the brain would likely require the use of a highly robust and accurate registration scheme.

6.2.2. Brain Tumor MRI Simulator. The aim of the new brain tumor MRI simulator is to generate sufficiently realistic MR images that are challenging to segment. Since the brain is a complex structure, it is extremely challenging to account for all the processes and variables involved. Many possible avenues are available for extending the brain tumor MRI simulator.

(1) *Extended modeling of brain tumor growth.* In this dissertation, the proposed synthetic brain tumor MRI simulator is mainly driven towards the generation of test images that empirically exhibit the major pathological effects seen in real images. The simulator can be extended to include the complex interactions between the deformation process, the infiltration process, the nutrient and chemical interactions, and formation of new blood vessels (angiogenesis). Zheng *et al.* [**107**] proposed a more detailed model for 2D data, however a full 3D implementation for the brain would be a significant challenge.

In particular, simulating the formation of new blood vessels and the shape changes of blood vessels within tumor regions would allow for the generation of more realistic mass effect and infiltration models and improve the appearance of the synthetic contrast enhanced T1w image. The modeling of the physical changes in csf volume and structure would improve the degree of realism in the image as the csf can be heavily influenced by tumor growth. If the intracranial pressure is increased due to tumor, there tends to be a loss of csf volume (which may not be restricted to the ventricles). If brain tumor growth blocks ventricular outflow, the csf volume can get larger and the ventricle shape would change.

(2) *Simulation of other pathologies in brain MRI.* The simulator described in this dissertation can also be extended to other cases that deviate from healthy adult MRI. A rather direct extension into the simulation of lesions in cases of vascular strokes or multiple sclerosis, for example. These cases typically present multiple regions with tissue infiltration and small scale deformation that could be generated using similar methods.

(3) *Simulation of aging in brain MRI.* Another possible extension to the simulator is the development of new growth models for the validation of segmentations in age-related studies. There is a lack of consensus as to the correct segmentation in the very young (newborn infants) and the very old age groups (older than 70 years). In both age groups, there is very low differentiation between white matter and gray matter. In the case of newborn infants, the white matter undergoes a growth process called myelination, which is mainly an infiltrative process. The lack of reliable ground truth for this age group makes validation difficult: a typical solution is to restrict the validation to only a part of the 3D volume [**78**]. In elderly subjects, the ventricles are typically enlarged. However, the increase of ventricular volume may be governed by the loss of tissue integrity (i.e., a change in tissue elasticity) and not by an increased ventricular pressure, so a biomechanical model with expansion due to ventricular pressure similar to the one described in this dissertation may not be appropriate.

(4) *Public web interface for simulated brain tumor MRI.* The methodology for generating synthesized tumor MRI could be further developed into web-accessible system where a user could interactively select locality, size, shape, and type of tumors by setting some variables (similar to the BrainWeb interface [**16**]). Image datasets generated by such a tool might find widespread use in validation of segmentation methods, comparison of different segmentation and registration strategies, and training or teaching. For example, the effect of voxel size and slice thickness on tumor volume estimates might be studied systematically.

Moreover, a series of images with embedded tumors of various size and shape might be used to evaluate well-established standards for tumor size measurements like the one-dimensional RECIST (Response Evaluation Criteria in Solid Tumours) criterion, which uses the maximum diameter of the structure measured only in axial cross-sections [90]. Systematic studies and evaluations would eventually lead to improved methodologies.

6.3. Summary

This dissertation presented a strategy for segmenting brain MR images based on a reference population model and measures of deviations from the model. I proposed the use of robust parameter estimation techniques in combination with the standard EM algorithm to compute the optimal image segmentation. This approach has been shown to be practically effective in isolating the relevant image parameters for images with significant deviations from the brain atlas. The combination of population models and robust methods for estimating the image likelihood functions created a novel segmentation strategy that has been applied to two particularly challenging cases: brain tumor MRI and newborn brain MRI.

A novel automatic segmentation scheme for brain tumors with adjoining edema was presented in Chapter 3. This scheme made use of standard non-contrast enhanced multimodal MRI, with an explicit requirement that one of the modalities is the T2-weighted modality. The method uses the robust MCD estimator described in Section 2.3.1 to find outliers in the image data sampled using the brain atlas, allowing the detection of tumor and edema as abnormal regions. Most methods so far have been applicable only to enhancing, homogeneous tumors. Furthermore, they require user-guidance in training a supervised classifier or obtaining a rough outline of the region of interest. The proposed brain tumor MRI segmentation technique also automatically identifies the presence or absence of nearby edema, which is novel and has not been presented before. This feature

is highly relevant for clinicians, as the edema region often may require secondary analysis and treatment after the primary focus to the tumor region.

A similar strategy has also been applied for the segmentation of newborn brain MR images with low contrast-to-noise ratio (Chapter 4). The newborn brain MRI segmentation method makes use of spatial priors to deal with the low image contrast and the robust parameter estimation techniques described in Section 2.3 to deal with the noisy, ambiguous image intensity values. So far, segmentation of newborn brain MR images has been done using manual or semi-automatic methods [96]. The proposed method provides an alternative that is fully automatic.

The validation results for newborn brains shown in Section 4.3 demonstrate that the performance of the segmentation framework is promising. Since the segmentation scheme is objective and fully reproducible, it has been used effectively in clinical studies that require analysis of a large population. The growth patterns of early brain development discussed in Section 4.4 was observed using segmentation tools that made use of the concepts and strategies presented in this dissertation. The resulting segmentations of newborn brain MR images have shown some interesting results on early brain development with respect to cerebral asymmetry, gender differences, and anterior-posterior growth patterns [35].

In Chapter 5, I presented a method for generating new ground truth with tumor and edema from a normal brain ground truth. I also described a method for generating synthetic multi-modal MR images that exhibit segmentation challenges similar to real tumor MRI. Objective evaluation of different segmentation methods can be done using a set of synthetic images with variations of tumor size, location, extent of surrounding edema, and contrast enhancement properties. The synthetic 3D MRI with the associated ground truth also allows for the validation of the segmentation of the whole brain, which includes white matter, gray matter, csf, and edema. The capability to obtain synthetic datasets is promising as most validations done so far were focused on only the tumor structures,

and no validations are typically performed on the segmentations of the infiltrated and deformed tissue.

In closing, this dissertation presented novel work in automatic MRI segmentation and the validation of such segmentations. The methods proposed in this dissertation involve almost no user interaction, so they provide objective results that are fully reproducible. The methods have the potential to be extended and applied to other pathological and growth processes in the brain or other parts of the body. The ultimate goal of the work described in this dissertation is the creation of a system that performs the following tasks:

- Automatic segmentation of anatomical structures with clinically interesting deviations from an expected model.

- Validation of segmentation results by comparing the automatic segmentation results to a simulated ground truth from the expected model.

The combination of the segmentation and validation frameworks might have significant potential value for clinical studies involving large populations. It provides a way to overcome the challenging routine of manual segmentation and to generate objective quantitative measures of the segmentation performance for quality control. The software tools developed for this dissertation form the initial steps toward the creation of the combined framework. They have been applied to a large number of real clinical cases with promising results.

APPENDIX A

Validation Measures

In this appendix, I review the measures used for comparing different segmentation results. These validation measures are used in Chapters 3 and 4 to provide a quantitative evaluation of the performance of the proposed automatic segmentation methods. They are also used in Chapter 5 to test the synthetic ground truth for the brain tumor MRI against the ground truth perceived by human raters. Comparisons are performed on the discretized version of the segmentation results. Section A.1 presents the measures used for comparing two binary segmentations. Section A.2 discusses the method of comparing two segmentations that have multiple (non-binary) label assignments.

A.1. Comparison of Binary Labels

In order to provide direct comparisons for structures of interest, the segmentation results (class posterior probabilities) from the methods proposed in Chapters 3 and 4 are converted to a set of binary images or masks. Each mask indicate the voxel locations that belong to a particular anatomy. The VALMET validation tool [32] is used to generate quantitative validation measures from the binary images associated with the structures of interest.

For a given binary segmentation volume A and a ground truth volume B, the relevant measures are:

(1) Dice similarity coefficient [26], which measures the ratio of volume overlap and the average volume:

$$DSC(A, B) = 2\frac{|A \cap B|}{(|A| + |B|)}.$$
(A.1.1)

(2) Jaccard similarity coefficient [**44**], which measures the ratio of volume overlap and the volume of the union of the segmentations:

$$JSC(A, B) = \frac{|A \cap B|}{|A \cup B|}. \tag{A.1.2}$$

The Jaccard and Dice similarity coefficients are related through the following equation:

$$
\begin{aligned}
JSC(A, B) &= \frac{2\,|A \cap B|}{2\,(|A| + |B| - |A \cap B|)} \\
&= \frac{2\,|A \cap B|}{|A| + |B|} \times \frac{1}{2 - \frac{2\,|A \cap B|}{|A| + |B|}} \\
&= \frac{DSC(A, B)}{2 - DSC(A, B)}.
\end{aligned}
\tag{A.1.3}
$$

In general, $JSC(A, B)$ tends to be lower than $DSC(A, B)$.

(3) Average of closest surface distances. At each point within the surface, the closest distance from the segmentation to the ground truth is averaged. This can be implemented efficiently using distance transforms of the binary objects [**23**]. More specifically,

$$Ave(A, B) = \frac{1}{|\delta A|} \sum_{x\,\in \delta A} \min_{y \in \delta B} distance(x, y) \tag{A.1.4}$$

where δA is set of the surface points of A and δB is the set of surface points of B. The distances can be restricted so that only distances to points going inside or outside the object are considered.

(4) The symmetric Hausdorff distance between the surface points ins δA and δB,

$$H(\delta A, \delta B) = max\{dist_H(\delta A, \delta B),\ dist_H(\delta B, \delta A)\} \tag{A.1.5}$$

where for a pair of collections of points P and Q, $dist_H(P, Q)$ is the (non-symmetric) Hausdorff distance. It is defined as the maximum of the closest pairwise point distances, $dist_H(P, Q) = \max_{p \in P} \min_{q \in Q} distance(p, q)$.

116

In ideal conditions, the measures provide a quantitative summary of the performance of a segmentation scheme. The volume overlap measures are normalized, where 0 indicates complete dissimilarity and 1 indicates complete agreement. A good segmentation scheme yields large values for the DSC and JSC measures (above 80%) and low distance values (less than $1\ mm$). However, the measures involving volume overlap (DSC and JSC) are strongly influenced by the volumes of the segmented objects. They tend to report low performance for segmentations of small structures. The average surface distance can also report low performance when there is an outlier in the computed distances.

A.2. Comparison of Non-binary Labels

The measures listed in the previous section only work on binary segmentations, so they can only be applied to each distinct brain structure separately. For segmentations of the whole brain with multiple labels, such as the case for newborn brain MRI, there is a need for a summary measure that indicate the differences between two segmentations. The summary measure chosen is Cohen's kappa [18], which measures the level of agreement of two raters. Suppose that the number of voxels (observations) within the image is N and that each segmentation maps the image to $M = |\mathcal{C}|$ number of classes / labels, then the level of agreement κ is defined as follows

$$
\begin{aligned}
\kappa &= \frac{p(agreement) - p(agreement\ by\ chance)}{1 - p(agreement\ by\ chance)} \\
&= \frac{\sum_{i=1}^{M} agreements(C_i) - \sum_{i=1}^{M} ef(C_i)}{N - \sum_{i=1}^{M} ef(C_i)} \quad \text{(A.2.1)}
\end{aligned}
$$

where $agreements(C_i)$ is the number of agreements between two segmentations for the class C_i. $ef(C_i)$ is the expected frequency of agreement by chance for class C_i, assuming that the two segmentations are statistically independent,

$$
ef(C_i) = \frac{1}{N}[\#\ of\ times\ rater\ 1\ assigns\ C_i] \times [\#\ of\ times\ rater\ 2\ assigns\ C_i] \quad \text{(A.2.2)}
$$

117

The κ values are normalized, 0 indicates independence and 1 indicates complete agreement. Large inter-rater variability results in small κ values. κ values greater than 0.7 is generally interpreted to reflect a satisfactory level of reliability. Cohen's kappa places equal weight on the samples of each class, so classes with larger number of observations will have more influence on the final result.

APPENDIX B

Maximum a Posteriori Image Segmentation

This appendix provides a summary of the Maximum a Posteriori (MAP) image segmentation approach, which provides possible extensions to the Maximum Likelihood approach described in Chapter 2. The following two sections describe the basic ideas for MAP segmentation and a method for estimating the segmentation parameters.

B.1. Introduction

The maximum a posteriori (MAP) estimator is a widely used Bayes estimator which gives the mode of the posterior distribution $p(W|I)$. More specifically, a MAP estimator computes the tuple of label image and model parameters $\hat{W}$, where

$$\hat{W} = \arg\max_{W} p(W|I) = \arg\max_{W} p(I|W)Pr(W) \tag{B.1.1}$$

As opposed to the ML estimator, the MAP estimator allows the modeling of the prior knowledge of the world state $W = (S, \theta)$. As indicated in Equation B.1.1, one has a choice of computing the MAP estimate using $p(W|I)$ or the product $p(I|W)p(W)$. The posterior probability $p(W|I)$ is referred to as the discriminative model, which can be used directly without computing $p(I|W)$ if an explicit definition is available. The image data likelihood term $p(I|W)$ is referred to as the generative model, which describes the image that will be likely observed given the segmentation labels and the model parameters.

It is generally extremely challenging to use an explicit discriminative model $p(W|I)$ for segmentation, since it involves training over all possible image observations. In most cases, the generative model $p(I|W)$ is easier to describe and can be approximated reliably using predefined statistical models. This motivates the use of a Bayesian segmentation framework that combines the generative model $p(I|W)$ and the prior $Pr(W)$. The

Bayesian viewpoint is particularly useful when there is a limited number of images available and when there exists some prior knowledge on the image content ($Pr(S)$) and the characteristics of the image appearance ($Pr(\theta)$).

B.2. Parameter Estimation using MCMC

In the ML segmentation formulation, the EM algorithm provides a framework for estimating the segmentation labels and the model parameters separately. In the MAP formulation the segmentation labels S and the model parameters θ need to be updated simultaneously since the dependence on θ is not used explicitly. The simultaneous estimation of the segmentation labels and the model parameters complicates the segmentation problem. Finding the best MAP estimate can be done using the Markov Chain Monte Carlo (MCMC) method, where we traverse the space of the world state W in a random fashion yet still guided by the relevant probabilities. MCMC methods can be relatively slow to converge, but compared to the EM algorithm it is more flexible. It can handle changes in the size or dimension of the model parameters θ and does not require explicit maximization of the objective function $p(W|I)$. More details on the MCMC method and its application to image analysis can be found in the book by Winkler [101].

MCMC combines two concepts: the solution of a numerical problem through the use of random samples (Monte Carlo) and the stochastic transitions of a changing variable that depends only on the current state (Markov chain). Consider the problem of computing the expectation of a random variable X, this can be solved by a Monte Carlo scheme as follows:

$$\mathcal{E}[X] = \int p(x)dx = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{B.2.1}$$

where the set of x_i values are N randomly generated samples of the random variable X from the distribution $p(X)$. Markov Chain Monte Carlo method generates the x_i using Markov chain (i.e. x_{i+1} is generated based only on x_i) where the chain has $p(X)$ as its stationary distribution. Related to the solution of the MAP segmentation problem is the problem of finding the mode of the distribution $p(X)$. Within the MCMC framework,

this can be implemented by isolating the samples that are traversed most often. A common approach for optimization with MCMC is by using simulated annealing, where we iteratively draw samples from $(p(X))^{\frac{1}{T}}$ where T signifies the temperature. Simulated annealing begins with a high temperature value which makes all solutions equally likely, and at each iteration the temperature is lowered until it converges to the most likely samples drawn from $p(X)$.

The fact that we can use $Pr(S, \theta)$ combined with the flexibility of the MCMC algorithm makes it possible to use a more sophisticated model (θ). Tu and Zhu [91] described a Bayesian image segmentation framework for 2D images that unifies the common image segmentation strategies. Their unified image segmentation framework combines the following operations:

(1) Multi-scale edge detection and partitioning.

(2) Clustering of observed image intensity data to form image likelihood.

(3) Deformable segmentation through region competition.

(4) Region merging and splitting.

The data-driven Markov Chain Monte Carlo (DDMCMC) segmentation framework [91] divides a 2D image into M disjoint regions. That is, the segmentation label image S is defined as $S = \cup_{i=1}^{M} R_i$, where $R_i \cap R_j = 0 \; \forall i \neq j$. Each region R_i have its own appearance model θ_i and is assigned an index l_i for the type of distributions representing the intensities within the region. The extended world state W for the segmentation becomes:

$$W = (M, \{(R_i, l_i, \theta_i) | i = 1, 2, \ldots, M\}). \tag{B.2.2}$$

As opposed to the world state $W = (S, \{(\mu_c, \Sigma_c) | c \in \mathcal{C}\})$ used within the example ML segmentation framework, this extended model allows the definition of the prior probability $Pr(W)$ as a product of the following prior probabilities:

(1) A prior on the number of regions, for example if it is not be preferable to have too many regions we can use $Pr(M) \propto \frac{1}{M}$

(2) A prior on the smoothness of the region boundaries defined by each R_i.

(3) A prior on the size of each region R_i, when the size of the underlying structures are known this can be used to constrain the segmentation to meaningful results.

The index l_i allows the algorithm to switch between different family of image likelihood models for each region R_i. Tu and Zhu's method uses four models for the likely appearance within the region: voxelwise independent Gaussian distribution, nonparametric intensity histogram or kernel density estimates, texture model, and a Bezier-spline model for regions that have smooth variations.

The solution space for the extended W is extremely large and likely contains many local maximas. Due to the combinatorial nature of the possible world state W, explicit optimization is practically impossible. The DDMCMC segmentation framework uses Markov Chain Monte Carlo to traverse the solution space. The method moves to a new state W' based on the current W estimate using the following Markov chain dynamics:

(1) Boundary diffusion through region competition [108]. The boundary Γ_{ij} between regions R_i and R_j undergoes motion based on the difference of the log likelihood given the region model parameters:

$$\frac{\partial \Gamma_{ij}(s)}{\partial t} = \left[log\frac{p(I_k|\theta_i, l_i)}{p(I_k|\theta_j, l_j)} + Brownian\ motion \right] \vec{n}(s) \tag{B.2.3}$$

where s is the parametrization of the boundary, k is the image location that corresponds to $\Gamma_i j(s)$ and $\vec{n}(s)$ is the normal of the boundary curve.

(2) Image appearance model adaptation. This dynamic involves finding the values θ_i for each region R_i that maximizes the regional log likelihood $log\ p(I_{R_i}|\theta_i, l_i)$.

(3) Region merging and splitting. A region R_i can be split into two regions or two regions R_i and R_j can be merged depending on how likely the Markov chain proposes to move to the new state. This can be implemented using the Metropolis-Hastings algorithm [63]. Barbu *et al.* [6] described a more efficient merging and splitting operations, where multiple regions are split and merged using Swendsen-Wang graph cuts.

(4) Switching of image appearance models. This dynamic involves changing the family of distributions for $p(I_{R_i}|\theta_i)$ by choosing a new index l_i for a given region R_i.

Each move within the solution space corresponds to a common image segmentation operations.

BIBLIOGRAPHY

[1] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee. Spatial transformations of diffusion tensor magnetic resonance images. *IEEE Trans. Med. Imaging*, 20(11):1131–1139, 2001.

[2] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Fast and simple calculus on tensors in the log-euclidean framework. In *Proc. MICCAI 2005*, LNCS 3749, pages 115–122, 2005.

[3] Vincent Arsigny, Pierre Fillard, Xavier Pennec, and Nicholas Ayache. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine*, 56(2):411–421, August 2006.

[4] B. Aubert-Broche, M. Griffin, G. B. Pike, A. C. Evans, and D. L. Collins. Twenty new digital brain phantoms for creation of validation of image data bases. *IEEE Trans Med Imaging*, 25(11):1410–1416, November 2006.

[5] O. Axelsson and V. A. Barker. *Finite element solution of boundary value problems.* Academic Press, 1984.

[6] Adrian Barbu and Song-Chun Zhu. Generalizing Swendsen-Wang to sampling arbitrary posterior probabilities. *PAMI*, 27, 2005.

[7] Thomas Bayes. Studies in the history of probability and statistics: IX. Thomas Bayes's essay toward solving a problem in the doctrine of chances. *Biometrika*, 45:296–315, 1958.

[8] J. Besag. On the statistical analysis of dirty pictures (with discussions). *J Royal Statistical Society, Series B*, 48:259–302, 1986.

[9] J. P. Boardman, K. Bhatia, S. Counsell, J. Allsop, O. Kapellou, M. A. Rutherford, A. D. Edwards, J. V. Hajnal, and D. Rueckert. An evaluation of deformation-based morphometry in the developing human brain and detection of volume changes associated with preterm birth. In *Proc. MICCAI 2003*, 2003.

[10] E. Bullitt, G. Gerig, S. M. Pizer, and S. R. Aylward. Measuring tortuosity of the intracerebral vasculature from MRA images. *IEEE Transactions on Medical Imaging*, 22:1163–1171, 2003.

[11] Wray Buntine. Computation with the exponential family and graphical models, 1996. Tutorial given at NATO Workshop on Learning in Graphical Models, Erice, Italy, September 1996.

[12] A. Christmann and S. Van Aelst. Robust estimation of Cronbach's alpha. *J of Multivariate Analysis*, 97:1660–1674, 2006.

[13] M. C. Clark, L. O. Hall, D. B. Goldgof, R. Velthuizen, F. R. Murtagh, and M.S . Silbiger. Automatic tumor-segmentation using knowledge-based techniques. *IEEE Transactions on Medical Imaging*, 117:187–201, 1998.

[14] Oliver Clatz, Pierre-Yves Bondiau, Simon K. Warfield Hervé Delingette Grégoire Malandalin, Maxme Sermesant, and Nicholas Ayache. In silico tumor growth: Application to glioblastomas. In *Proc. MICCAI 2004*, volume LNCS 3217, pages 337–345, 2004.

[15] Olivier Clatz, Maxime Sermesant, Pierre-Yves Bondiau, Hervé Delingette, Simon K. Warfield, Grégoire Malandain, and Nicholas Ayache. Realistic simulation of the 3D growth of brain tumors in MR images including diffusion and mass effect. *IEEE Trans. Med. Imaging*, 24(10):1344–1346, October 2005.

[16] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans. BrainWeb: Online interface to a 3D MRI simulated brain database. *NeuroImage*, 5(4), 1997.

[17] Chris A. Cocosco, Alex P. Zijdenbos, and Alan C. Evans. A fully automatic and robust brain MRI tissue classification method. *Medical Image Analysis*, 7:513–527, December 2003.

[18] J. Cohen. A coefficient of agreement for nominal scales. *Educational and psychological measurements*, 20:37–46, 1960.

[19] D. L. Collins, A. P. Zijdenbos, W. F. C. Baaré, and A. C. Evans. ANIMAL+INSECT: Improved cortical structure segmentation. In A. Kuba, M. Sámal, and A. Todd-Pokropek, editors, *Information Processing in Medical Imaging*, volume 1613 of *Lecture Notes in Computer Science*, pages 210–223. Springer, Heidelberg, Germany, 1999.

[20] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans. Medical Imaging*, 17(3):463–468, June 1998.

[21] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. MIT Press, Cambridge, MA, 2nd edition edition, 2001.

[22] M. Bach Cuadra, J. Gomez, P. Hagmann, C. Pollo, J..-G. Villemure, B. M. Dawant, and J. Ph. Thiran. Atlas-based segmentation of pathological brains using a model of tumor growth. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*, volume 2488 of *LNCS*, pages 380–387. Springer Verlag, September 2002.

[23] Per-Erik Danielsson. Euclidean distance mapping. *Computer Graphics and Image Processing*, 14:227–248, 1980.

[24] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(2):224–227, April 1979.

[25] A.P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc.*, 39:1–38, 1977.

[26] Lee R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, July 1945.

[27] R. O. Duda, P. E. Hart, and D.G. Stork. *Pattern Classification*. Wiley, second edition, 2001.

[28] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters. 3D statistical neuroanatomical models from 305 MRI volumes. In *Proc. IEEE Nuclear Science Symposium and Medical Imaging Conference*, pages 1813–1817, 1993.

[29] M. Ferrant, A. Nabavi, B. Macq, Ferenc A. Jolesz, Ron Kikinis, and S. K. Warfield. Registration of 3D interoperative MR images of the brain using finite element biomechanical model. *IEEE Trans. Med. Imaging*, 20(12):1384–1397, 2001.

[30] P. T. Fletcher and S. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *CVAMIA 2004*, LNCS 3117, pages 87–98, 2004.

[31] L. M. Fletcher-Heath, L. O. Hall, D. B. Goldgof, and F. R. Murtagh. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artifical Intelligence in Medicine*, 21:43–63, 2001.

[32] G. Gerig, M. Jomier, and M. Chakos. VALMET: a new validation tool for assessing and improving 3D object segmentation. In W Niessen and M Viergever, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2001*, volume 2208, pages 516–523, New York, October 2001. Springer Verlag.

[33] David T. Gering, W. Eric L. Grimson, and R. Kikinis. Recognizing deviations from normalcy for brain tumor segmentation. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*, volume 2488 of *LNCS*, pages 388–395. Springer Verlag, September 2002.

[34] J.H. Gilmore, G. Zhai, K. Wilber, J.K. Smith, W. Lin, and G. Gerig. 3T magnetic resonance imaging of the brain in newborns. *Psychiatry Research Neuroimaging*, 132:81–85, 2004.

[35] John H. Gilmore, Weili Lin, Marcel W. Prastawa, Christopher B. Looney, Y. Sampath K. Vetsa, Rebecca C. Knickmeyer amd Dianne D. Evans, J. Keith Smith, Robert M. Hamer, Jeffrey A. Lieberman, and Guido Gerig. Regional gray matter

growth, cerebral asymmetry, and sexual dimorphism in the neonatal brain. *Journal of Neuroscience*, 27(6):1255–1260, February 2007.

[36] Mark Girolami and Chao He. Probability density estimation from optimally condensed data samples. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(10):1253–1264, October 2003.

[37] Robert I. Grossman and David M. Yousem. *Neuroradiology: The Requisites*. Mosby, 2003.

[38] Morton E. Gurtin. *An introduction to continuum mechanics*. Academic Press, 1981.

[39] T. Hastie, R. Tibshirani, and J. H. Friedman. *The Elements of Statistical Learning*. Springer-Verlag, 2001.

[40] S. Ho, E. Bullitt, and G. Gerig. Level set evolution with region competition: Automatic 3-D segmentation of brain tumors. In R. Katsuri, D. Laurendeau, and C. Suen, editors, *Proc. 16th International Conference on Pattern Recognition*, pages 532–535. IEEE Computer Society, August 2002.

[41] Thomas J. R. Hughes. *The finite element method: linear static and dynamic finite element analysis*. Dover, 2000.

[42] P.S. Hüppi, S. Warfield, R. Kikinis, P.D. Barnes, G.P. Zientara, F.A. Jolesz, M.K. Tsuji, and J.J. Volpe. Quantitative magnetic resonance imaging of brain development in premature and normal newborns. *Ann Neurol*, 43:224–235, 1998.

[43] Terrie E. Inder, Simon K. Warfield, Hong Wang, Petra S. Hüppi, and Joseph J. Volpe. Abnormal cerebral structure is present at term in premature infants. *Pediatrics*, 115(2):286–294, 2005.

[44] P. Jaccard. The distribution of flora in the alpine zone. *New Phytologist*, 11:37–50, 1912.

[45] Sarang. Joshi, Peter. Lorenzen, Guido. Gerig, and Elizabeth. Bullitt. Structural and radiometric asymmetry in brain images. *Medical Image Analysis*, 7:155–170, 2003.

[46] E. R. Kandel, J. H. Schwartz, and T. M. Jessel. *Principles of Neural Science*. McGraw Hill, 4th edition edition, 2000.

[47] M. Kaus, S.K. Warfield, A. Nabavi, Peter M. Black, Ferenc A. Jolesz, and Ron Kikinis. Automated segmentation of MR images of brain tumors. *Radiology*, 218(2):586–591, 2001.

[48] M. R. Kaus, S. K. Warfield, A. Nabavi, E. Chatzidakis, P. M. Black, Jolesz F. A., and Kikinis R. Segmentation of meningiomas and low grade gliomas in MRI. In Chris Taylor and Alan Colchester, editors, *MICCAI*, volume 1679 of *LNCS*, pages 1–10. Springer, Sept 1999.

[49] Solomon Kullback and Richard A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):70–86, 1951.

[50] Stelios K. Kyriacou, Christos Davatzikos, S. James Zinreich, and R. Nick Bryan. Nonlinear elastic registration of brain images with tumor pathology using a biomechanical mode. *IEEE Trans. Med. Imaging*, 18(7):580–592, 1999.

[51] H. C. Lee and D. R. Cok. Detecting boundaries in a vector field. *IEEE Trans. Signal Processing*, 39(5):1181–1194, 1991.

[52] C. Lenglet, M. Rousson, and R. Deriche. DTI segmentation by statistical surface evolution. *IEEE Trans. Med Imaging*, 25(6):685–700, 2006.

[53] Y. Liu, F. Dellaert, W. E. Rothfus, A. Moore, J. Schneider, and T. Kanade. Classification-driven pathological neuroimage retrievl using statistical asymmetry measures. In *Proc. MICCAI 2001*, 2001.

[54] P. Lorenzen, S. Joshi, G. Gerig, and E. Bullitt. Tumor-induced structural radiometric asymmetry in brain images. In *Proceedings of Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)*, 2001.

[55] P. Lorenzen, M. Prastawa, B. Davis, G. Gerig, E. Bullitt, and S. Joshi. Multi-modal image set registration and atlas formation. *Medical Image Analysis*, 10:440–451, 2006.

[56] Peter Lorenzen. *Multi-modal Image Registration and Atlas Formation*. PhD thesis, University of North Carolina at Chapel Hill, 2006.

[57] Stanley Lu, Daniel Ahn, Glyn Johnson, and Soonmee Cha. Peritumoral diffusion tensor imaging of high-grade gliomas and metastatic brain tumors. *AJNR*, 24(5):937–941, May 2003.

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

[59] Mikhail B. Malyutov and M Lu. Robust modification of the EM-algorithm for parametric multitrajectory estimation in noise and clutter. In *NATO Science Series III: Computers and Systems Sciences*, volume 198, pages 575–586, 2005.

[60] K. V. M. Mardia and P. E. Jupp. *Directional Statistics*. John Wiley and Sons, 2000.

[61] M. Matsumoto and T. Nishimura. Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Trans. on Modeling and Computer Simulation*, 8(1):3–30, January 1998.

[62] J. Matsuzawa, M. Matsui, T. Konishi, K. Noguchi, R.C. Gur, W. Bilder, and T. Miyawaki. Age-related volumetric changes of brain gray and white matter in healthy infants and children. *Cerebral Cortex*, 11(4):335–342, April 2001.

[63] N. Metropolis, M.N Rosenbluth, A. W. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1092, 1953.

[64] M. Miga, K. Paulsen, F. E. Kennedy, A. Hartov, and D. Roberts. Model-updated image guided neurosurgery using the finite element method: incorporation of the falx cerebri. In *Proc. MICCAI 1999*, LNCS 1679. Springer-Verlag, 1999.

[65] M. I. Miga, K. D. Paulsen, J. M. Lemery, S. D. Eisner, A. Hartov, F. E. Kennedy, and Roberts D. W. Model-updated image guidance: initial clinical experience with gravity induced brain deformation. *IEEE Trans Med Imaging*, 18:866–874, 1999.

[66] Karol Miller. *Biomechanics of brain for computer integrated surgery*. PhD thesis, Warsaw Univ. Technol., Warsaw, Poland, 2002.

[67] T. Minka. Expectation-maximization as lower bound maximization, 1998. Tutorial published on the web at http://www-white.media.mit.edu/ tpminka/papers/em.html.

[68] P. Mitra, C. A. Murthy, and S. K. Pal. Density based multiscale data condensation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(6), 2002.

[69] Ashraf Mohamed and Christos Davatzikos. Finite element modeling of brain tumor mass-effect from 3D medical images. In *Proc. MICCAI 2005*, LNCS 3749, pages 400–408, 2005.

[70] Ashraf Mohamed, Evangelia I Zacharaki, Dinggang Shen, and Christos Davatzikos. Deformable registration of brain tumor images via a statistical model of tumor-induced deformation. *Medical Image Analysis*, 10(5):752–763, October 2006.

[71] N. Moon, E. Bullitt, K. Van Leemput, and G. Gerig. Automatic brain and tumor segmentation. In Takeyoshi Dohi and Ron Kikinis, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2002*, volume 2489 of *LNCS*, pages 372–379. Springer Verlag, September 2002.

[72] Bénédicte Mortamet, Donglin Zeng, Guido Gerig, Marcel Prastawa, and Elizabeth Bullitt. Effects of healthy aging measured by intracranical compartment volumes

using a designed MR brain database. In *Proc. MICCAI 2005*, LNCS 3749, pages 383–391, 2005.

[73] T. Nagashima, B. Horwitz, and S. I. Rapoport. A mathematical model for vasogenic brain edema. *Adv Neurol*, 52:317–326, 1990.

[74] T. Nagashima, Y. Tada, S. Hamano, M. Sakakura, K. Masaoka, N. Tamaki, and S. Matsumoto. The finite element analysis of brain oedema associated with intracranial meningiomas. *Acta Neurochir Suppl*, 51:155–157, 1990.

[75] R. Neal and G. Hinton. A view of the EM algorithm that justifies incremental, sparse, and other variants. In M. I. Jordan, editor, *Learning in Graphical Models*. Kluwer, 1998.

[76] P.-O. Persson and G. Strang. A simple mesh generator in MATLAB. *SIAM Review*, 46(2):329–345, 2004.

[77] Marcel Prastawa, Elizabeth Bullitt, Nathan Moon, Koen Van Leemput, and Guido Gerig. Automatic brain tumor segmentation by subject specific modification of atlas priors. *Academic Radiology*, 10:1341–1348, December 2003.

[78] Marcel Prastawa, John H. Gilmore, Weili Lin, and Guido Gerig. Automatic segmentation of MR images of the developing newborn brain. *Medical Image Analysis*, 9:457–466, 2005.

[79] Jan Rexilius, Horst K Hahn, Mathias Schlüter, Sven Kohle, Holger Bourquain, Joachim Böttcher, and Heinz-Otto Peitgen. A framework for the generation of realistic brain tumor phantoms and applications. In *Proc. MICCAI 2004*, LNCS 3217, pages 243–250, 2004.

[80] P. J. Rousseeuw and K. Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41(3):212–223, 1999.

[81] Mary Rutherford, editor. *MRI of the Neonatal Brain*. W.B. Saunders, 2002.

[82] B. Schölkopf, J. Platt, J. Shawe-Taylor, A. Smola, and R. Williamson. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13:1443–1471, 2001.

[83] D. W. Scott and S. J. Sheather. Kernel density estimation with binned data. *Comm Statistics- Theory and Methods*, 14:1353–1359, 1985.

[84] Thomas. B. Sebastian, Huseyin. Tek, Joseph. J. Crisco, and Benjamin. B. Kimia. Segmentation of carpal bones from ct images using skeletally coupled deformable models. *Medical Image Analysis*, 7(1):21–45, March 2003.

[85] John G. Sled, Alex P. Zijdenbos, and Alan C. Evans. A nonparametric method for automatic correction of tissue nonuniformity in MRI data. *IEEE Trans. Medical Imaging*, 17(1):87–97, 1998.

[86] Charles Stewart. Robust parameter estimation in computer vision. *SIAM Review*, 41(3):513–537, 1999.

[87] M. Styner, Ch. Brechbuhler, Gabor Szekely, and G. Gerig. Parametric estimate of intensity inhomogeneities applied to MRI. *IEEE TMI*, 19(3):153–165, 2000.

[88] H. Tek and B.B. Kimia. Image segmentation by reaction-diffusion bubbles. In *International Conference on Computer Vision (ICCV'95)*, pages 156–162, 1995.

[89] H. Tek and B.B. Kimia. Volumetric segmentation of medical images by three-dimensional bubbles. *Computer Vision and Image Understanding (CVIU)*, 65(2):246–258, 1997.

[90] P. Therasse, S. G. Arbuck, E. A. Eisenhauer, J. Wanders, R. S. Kaplan, and L. Rubinstein et al. New guidelines to evaluate the response to treatment in solid tumors. *J. Nat. Cancer Inst.*, 92:205–216, 2000.

[91] Zhuowen Tu and Song-Chun Zhu. Image segmentation by data driven Markov chain Monte Carlo. *IEEE PAMI*, 24:c57–673, 2002.

[92] K. van 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*, 18:885–896, October 1999.

[93] K. van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Medical Imaging*, 18:897–908, October 1999.

[94] Koen van Leemput. Probabilistic brain atlas encoding using bayesian inference. In *Proc. MICCAI 2006*, LNCS 4190, pages 704–711, 2006.

[95] Koen van Leemput, Frederik Maes, Dirk Vandermeulen, Alan Colchester, and Paul Suetens. Automated segmentation of multiple sclerosis lesions by model outlier detection. *IEEE Trans Med Imaging*, 20(8):677–688, 2001.

[96] Simon K. Warfield, Michael Kaus, Ferenc A. Jolesz, and Ron Kikinis. Adaptive, template moderated, spatially varying statistical classification. *Medical Image Analysis*, 4(1):43–55, Mar 2000.

[97] Simon K. Warfield, Kelly H. Zou, and William M. Wells. Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. *IEEE Trans. Med. Imaging*, 23(7):903–921, 2004.

[98] Richard Wasserman, Raj Acharya, Claudio Sibata, and K. H. Shin. A patient-specific in vivo tumor model. *Mathematical Biosciences*, 136:111–140, 1996.

[99] L.-Y. Wei and M. Levoy. Fast texture synthesis using tree-structured vector quantization. In *Proc. SIGGRAPH 2000*, pages 479–488, 2000.

[100] W. M. Wells, R. Kikinis, W. E. L. Grimson, and F. Jolesz. Adaptive segmentation of MRI data. *IEEE Transactions on Medical Imaging*, 15:429–442, August 1996.

[101] Gerhard Winkler. *Image analysis, random fields and Markov chain Monte Carlo methods: a mathematical introduction*. Springer-Verlag, second edition, 2003.

[102] Anthony Brinston Wolbarst. *Looking within: how X-Ray, CT, MRI ultrasound, and other medical images are created and how they help physicians save lives*. University of California Press, 1999.

[103] C. F. J. Wu. On the convergence properties of the EM algorithm. *Ann. Stat*, 11:95–103, 1983.

[104] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. J. Gee, and G. Gerig. User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *NeuroImage*, 31:1116–1128, 2006.

[105] Guihuao Zhai, Weili Lin, Kathy Wilber, Guido Gerig, and John H. Gilmore. Comparison of regional white matter diffusion in healthy neonates and adults using a 3T head-only MR scanner. *Radiology*, 229:673–681, DEC 2003.

[106] 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 Trans Med Ima*, 20(1):44–57, January 2001.

[107] X. Zheng, S. M. Wise, and V. Cristini. Nonlinear simulation of tumor necrosis, neo-vascularization and tissue invasion via and adaptive finite-element/level-set method. *Bulletin of Mathematical Biology*, 67:211–259, 2005.

[108] S. Zhu, T. S. Lee, and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. In *ICCV'95*, pages 416–423, 1995.