MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Technical Report No. AITR-1566 January, 1995

Segmentation of Brain Tissue from Magnetic Resonance Images

Tina Kapur

This publication can be retrieved by anonymous ftp to publications.ai.mit.edu.

Abstract

Segmentation of medical imagery is a challenging problem due to the complexity of the images, as well as to the absence of models of the anatomy that fully capture the possible deformations in each structure. Brain tissue is a particularly complex structure, and its segmentation is an important step for studies in temporal change detection of morphology, as well as for 3D visualization in surgical planning. In this paper, we present a method for segmentation of brain tissue from magnetic resonance images that is a combination of three existing techniques from the Computer Vision literature: EM segmentation, binary morphology, and active contour models. Each of these techniques has been customized for the problem of brain tissue segmentation in a way that the resultant method is more robust than its components. Finally, we present the results of a parallel implementation of this method on IBM's supercomputer Power Visualization System for a database of 20 brain scans each with 256x256x124 voxels and validate those against segmentations generated by neuroanatomy experts.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for this research was provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contracts N00014-91-J-4038 and N00014-94-01-0994, and in part by a grant from IBM.

Segmentation of Brain Tissue from Magnetic Resonance Images

by

Tina Kapur

Submitted to the Department of Electrical Engineering and Computer Science
on January 6, 1995, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

Segmentation of medical imagery is a challenging problem due to the complexity of the images, as well as to the absence of models of the anatomy that fully capture the possible deformations in each structure. Brain tissue is a particularly complex structure, and its segmentation is an important step for studies in temporal change detection of morphology, as well as for 3D visualization in surgical planning. In this paper, we present a method for segmentation of brain tissue from magnetic resonance images that is a combination of three existing techniques from the Computer Vision literature: EM segmentation, binary morphology, and active contour models. Each of these techniques has been customized for the problem of brain tissue segmentation in a way that the resultant method is more robust than its components. Finally, we present the results of a parallel implementation of this method on IBM's supercomputer Power Visualization System for a database of 20 brain scans each with 256x256x124 voxels and validate those against segmentations generated by neuroanatomy experts.

Thesis Supervisor: W. Eric L. Grimson
Title: Associate Professor of Electrical Engineering and Computer Science

Acknowledgments

I would like to thank my advisor, Prof. Eric Grimson, for his support and guidance. This work, and almost every other aspect of my life at MIT, has been enhanced because of his enthusiasm and constant encouragement to explore and to learn. I would also like to thank Dr. Ron Kikinis at Harvard Medical School for suggesting this particular project to me, the countless lessons in neuroanatomy, and for being a unique liaison between the fields of computer vision and medicine. I would like to thank IBM for providing a generous grant that partially supported this work, and Dr. Russ Taylor of IBM for his enthusiastic participation in this project.

I would especially like to thank Sandy Wells for explaining his EM segmenter to me. I would like to thank Gil Ettinger for sharing his expertise in medical imaging and parallel programming with me. Thanks to Dr. Martha Shenton, Hiroto Hokama, and Cindy Wible of Brigham and Women's Hospital for providing manual segmentations for validation work done in this thesis.

I would like to thank Jose Robles, David Wetherall, Gil Ettinger, Sandy Wells, Greg Klanderman, Tao Alter, Oded Maron, Holly Yanco, Mike Wessler, Raquel Romano, Lily Lee, Carlo Maley, Dave Baggett, Anita Flynn, Pisti Morocz, Sharon Peled, Shalini Bhargava, Gianna Gifford, and Caprice Sikorsky for being my friends. I would also like to thank Marianna Jakab, Dr. Langham Gleason, Dr. Petra Huppi, Simon Warfield, John Martin, Mike Halle, Simon Warfield, Adam Shostack, and Mark Anderson for making the Surgical Planning Lab an enjoyable place to work.

Special thanks to my husband Atin Malaviya, and to my family: Rekha Malaviya, Anand Malaviya, Seema Winsor, and David Winsor for their love through it all.

Contents

1 Introduction **6**

1.1 The Segmentation Problem 7

1.2 Modalities for Medical Image Acquisition 8

1.3 Image Formation in MRI 8

1.4 Characteristics of Medical Imagery 9

1.5 The Brain Segmentation Problem 11

2 Background **16**

2.1 EM Segmentation 16

2.2 Mathematical Morphology 25

2.3 Deformable Models 27

3 Our Method **33**

3.1 Description of Input Data 33

3.2 Constructing a Model for the Brain 35

3.3 Segmenting the Brain Tissue 36

3.3.1 Gain Correction 37

3.3.2 Removal of Thin Connectors Using Morphological Operations 37

3.3.3 Extraction of Brain Surface using Deformable Contour Models 41

3.4 Implementation on IBM Power Visualization System 46

3.4.1 Architecture of Power Visualization System 51

3.4.2 Sample Running Times 51

4 The Validation Problem And Results **52**

4.1 The Validation Problem 52

4.2 Validation Methods 54

4.3 Our Validation Scheme and Results 54

4.3.1 Quantitative Results from Schizophrenia Study 55

4.3.2 Qualitative Results 56

4.3.3 Surgical Planning 56

4.3.4 NewBorn Study 62

5 Discussion and Conclusion **72**

5.1 Generalization to Other Anatomical Structures 74

5.2 Conclusion 75

6 Related Work **76**

6.1 Deformable Model Based Methods 77

6.1.1 Example Application Systems 78

6.2 Morphology Based Methods 81

6.2.1 Example Application Systems 81

6.3 Statistical Methods 82

6.3.1 Example Application Systems 82

Chapter 1

Introduction

As recently as a few years ago, most radiologists only had available to them pictures of several cross sections of a brain or an abdomen on a light board, plus their own 3d reconstruction abilities, to make a clinical diagnosis or to evaluate the results of a therapy on a patient. Surgeons had the same resources and the additional challenge of planning a path to the lesions or tumor. In recent years, the interdisciplinary field of medical image processing has produced several automatic and semi-automatic tools to assist medical practitioners and researchers in these and other tasks. For instance, tools for 3D visualization of anatomy (i.e. reconstruction and rendering), for surgical planning as well as pedagogical purposes, are available in several hospitals and research laboratories. In addition, tools are available for tasks such as the registration of data for the same patient from different modalities so that the total information available to the clinician is increased, and for measuring the change in volume of a particular structure in a series of images taken over time, to evaluate a therapy. In addition to producing new tools for medical applications, the goals of medical image processing include increased automation of the existing tools that have proven useful to the medical community yet still require assistance from experts.

1.1 The Segmentation Problem

We are interested in the problem of segmentation of medical data into anatomical structures, which is an interesting and important step for many medical imaging applications that are currently under development. For instance, the Enhanced Reality Visualization project for providing surgeons with intra-operative "x-ray vision" [1] assumes that a segmentation of structures such as brain tissue, tumor, and ventricles of the patient are available for fusion with live video of the patient. At the Surgical Planning Lab (SPL) of Brigham and Women's Hospital (BWH), Shenton et al. use segmented temporal lobes from magnetic resonance brain scans for studying the relationship of schizophrenia to volumetric changes in the region [2]. At University of Southern California, Sandor et. al require a segmentation of the intra-cranial cavity in order to produce a labeling for the cortex [3]. The Multiple Sclerosis project, also at SPL, BWH, [4] for studying the growth of lesions over time uses a segmentation of the intra-cranial cavity to register various temporal MR scans of a patient to a baseline scan.

In order to understand the issues in medical image segmentation, in contrast with segmentation of, say, images of indoor environments, or snapshots of an assembly line, which are the kinds of images with which general purpose visual segmentation systems deal, we need an understanding of the salient characteristics of medical imagery. To that end, we provide a brief introduction to some popular medical imaging modalities, a high level description of the image formation process in magnetic resonance imaging since that is the modality used to acquire the images in this work, and a discussion of the salient characteristics of medical data that make the segmentation task challenging. We conclude this chapter with the statement of the specific segmentation task we address in this work, and an outline of the rest of this paper.

1.2 Modalities for Medical Image Acquisition

Various imaging modalities are available for acquiring complementary information for different aspects of anatomy. Examples are MRI (Magnetic Resonance Imaging), PET (Positron Emission Tomography), SPECT (Single Photon Emission Computed Tomography), MRS (Magnetic Resonance Spectroscopy), Ultrasound, and X-ray imaging including CT (Computed Tomography). The modalities have different strengths and thus are used under different circumstances. For instance, MRI produces high contrast between soft tissues, and is therefore useful for detecting lesions in the brain. CT captures bone with accuracy, and is used for dosage planning in radiation therapy. PET, SPECT, and MRS are typically used to provide functional information with MRI also gaining usage in that domain. A typical image acquisition consists of exposing patients to the imaging equipment, sometimes with contrast enhancing agents or markers, and generating an image of their anatomy. This image can be a 2D projection of a 3D scene as is produced with X-ray or ultrasound, or it can be a full 3D image, as generated by CT or MRI.

1.3 Image Formation in MRI

In this section we give a high level view of the process by which an MR image is acquired.

MRI exploits the inherent magnetic moment and angular momentum of certain atomic nuclei. The nucleus of the hydrogen atom (proton) is used in biologic tissue imaging due to its abundance in the human body and its large magnetic moment. When the patient is positioned in the bore of the imaging magnet, protons in the tissues experience a strong static magnetic field and precess at a characteristic frequency that is a function solely of the magnetic field strength, and does not depend, for instance, on the tissue to which the protons belong. An excitation magnetic field is applied at this characteristic frequency to alter the orientation of precession of the protons. The protons relax to their steady state after the excitation field is stopped.

The reason MRI is useful is because protons in different tissues relax to their steady state at different rates. MRI essentially measures the components of the magnitude vector of the precession orientation at different times and thus differentiates tissues. These measures are encoded in 3D using methods for slice selection, frequency encoding, and phase encoding. Slice selection is performed by exciting thin cross-sections of tissue one at a time. Frequency encoding is achieved by varying the returned frequency of the measured signal, and phase encoding is done by spatially varying the returned phase of the measured signal.

In the interest of clinical efficiency, patient comfort, and minimized artifacts due to physiologic patient motion, rapid scanning techniques for MRI are popular. In this work we use gradient-echo images, that are obtained using a rapid scanning technique. Sample MRI data used in this work is shown in Figure 1-1. Readers interested in details of MRI and of this particular acquisition method are referred to [5], [6].

1.4 Characteristics of Medical Imagery

While the nature of medical imagery allows a segmentation system to ignore issues like occlusion, illumination, and pose determination that would be of significance to a more general purpose visual image segmentation system, there are other issues, as discussed below, that such a system needs to address. The objects to be segmented from medical imagery are actual anatomical structures, which are often non rigid and complex in shape, and exhibit considerable variability from person to person. This, combined with an absence of explicit shape models that capture the deformations in anatomy, makes the segmentation task challenging. Magnetic resonance images are further complicated due to the limitations in the imaging equipment: inhomogeneities in the receiver or transmit coils leads to a non linear gain artifact in the images, and large differences in magnetic susceptibilities of adjacent tissue leads to distortion of the gradient magnetic field, and hence a spatial susceptibility artifact in the images. In addition, the signal is degraded by motion artifacts that may appear in the images due to movement of the patient during the scan.

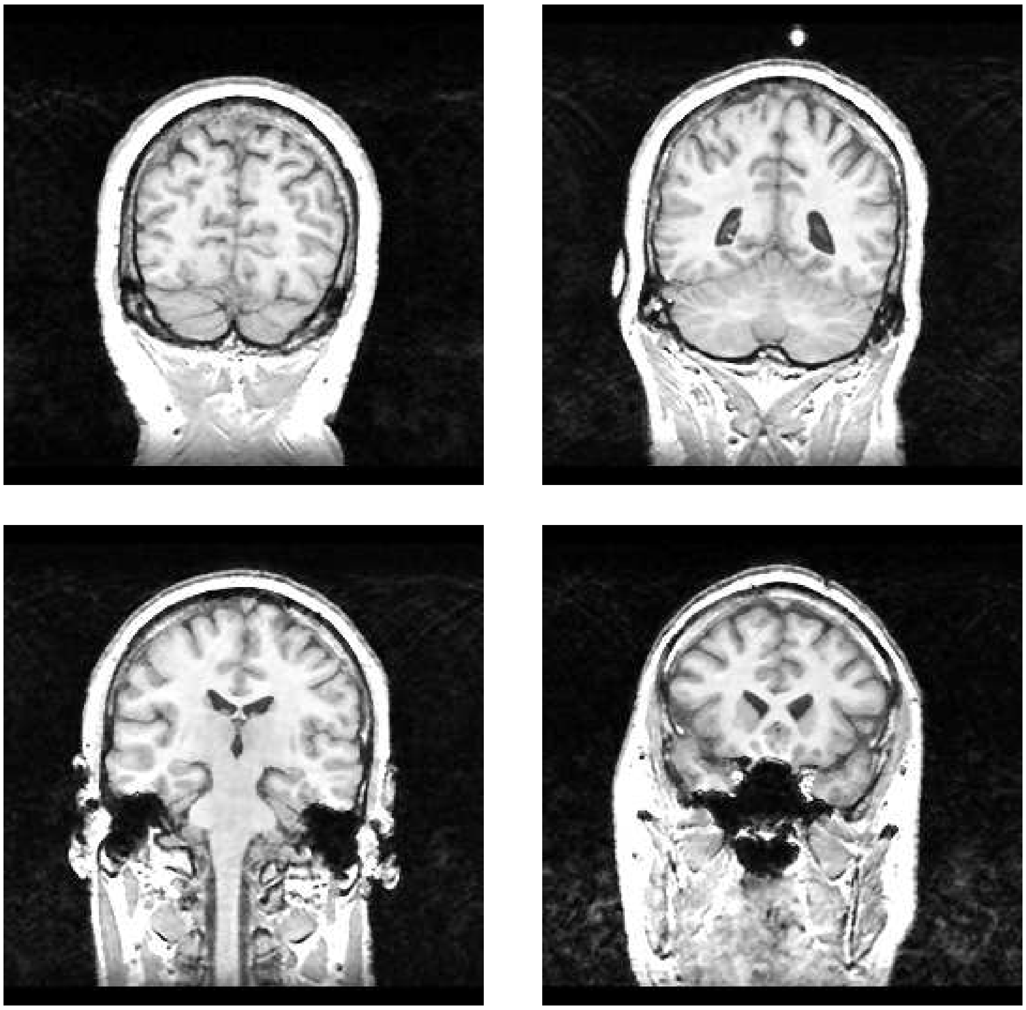


Figure 1-1: Sample Coronal Cross Sections from A Gradient Echo MR Scan.

1.5 The Brain Segmentation Problem

In this work, we address the specific segmentation task of extracting the brain tissue from magnetic resonance images, which serves the enhanced reality visualization, brain volume measurement, cortical labeling, and change detection applications mentioned earlier. Figure 1-2 shows a photograph from an anatomy atlas of a cut through a cadaver head, and Figures 1-3 and 1-4 respectively show a photograph of the whole brain isolated by actual dissection, and a photograph of a cut through the isolated brain. We present a method that accomplishes a similar isolation of the brain tissue (see Figure 1-5) from digitized magnetic resonance images such as the ones shown in Figure 1-1.

In Chapter 2 we present background on the techniques used in this work. In Chapter 3 we present a method for the segmentation that combines the strengths of some of the existing technology with some new ideas and is more robust than its individual components. We also discuss our parallel implementation of the method on IBM's supercomputer, the Power Visualization System. In Chapter 4 we discuss issues regarding validation of medical image segmentation, and present the results of our method on a database of 20 brains, each of size 256x256x124 voxels. In Chapter 5 we describe a chronology of our development, and discuss potential generalizations of our methodology to segment other structures from medical imagery. In Chapter 6 we survey other related work in medical image segmentation.

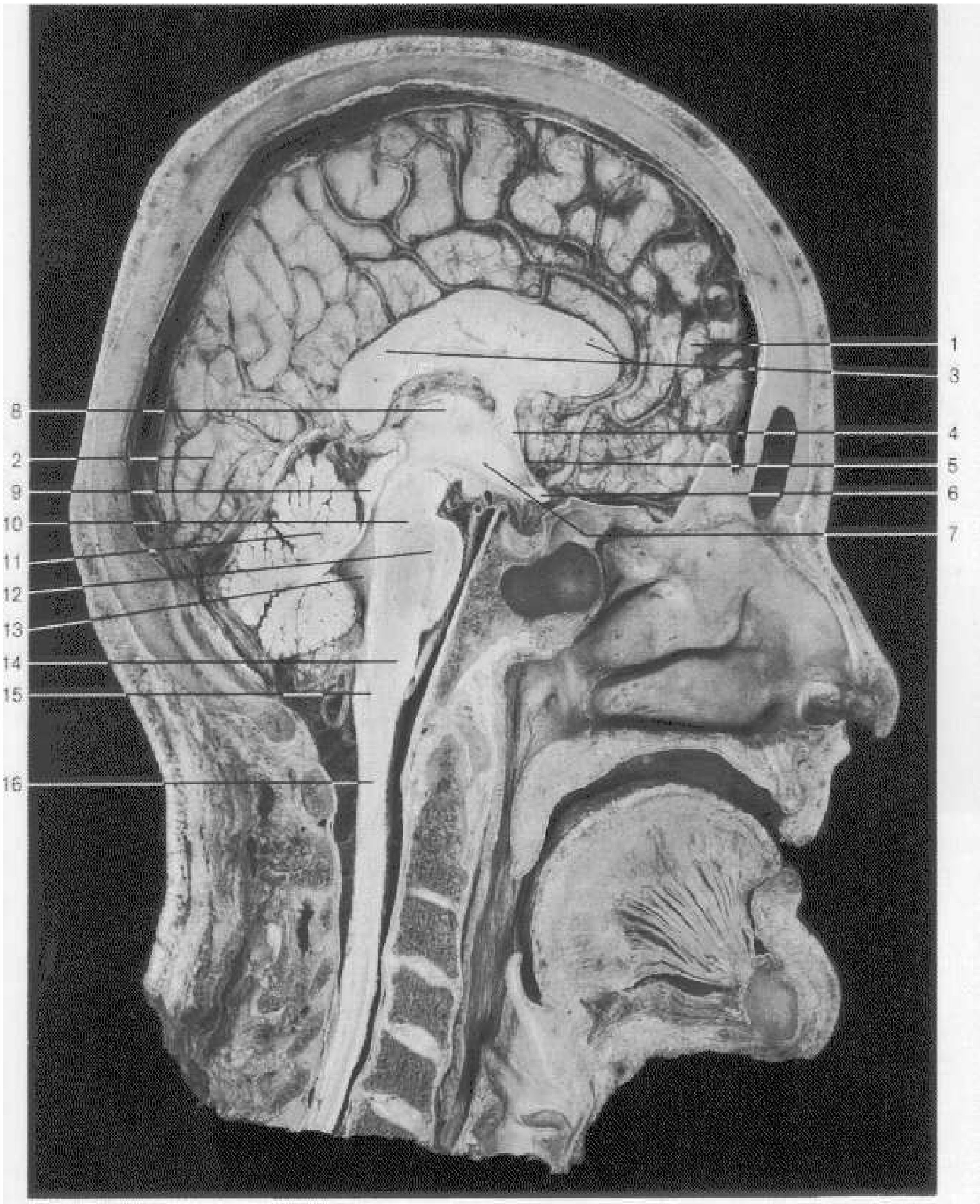


Figure 1-2: A Photograph of a Sagittal Cross Section of a Cadaver Head. Labels 1-16 Point to Various Brain Structures.

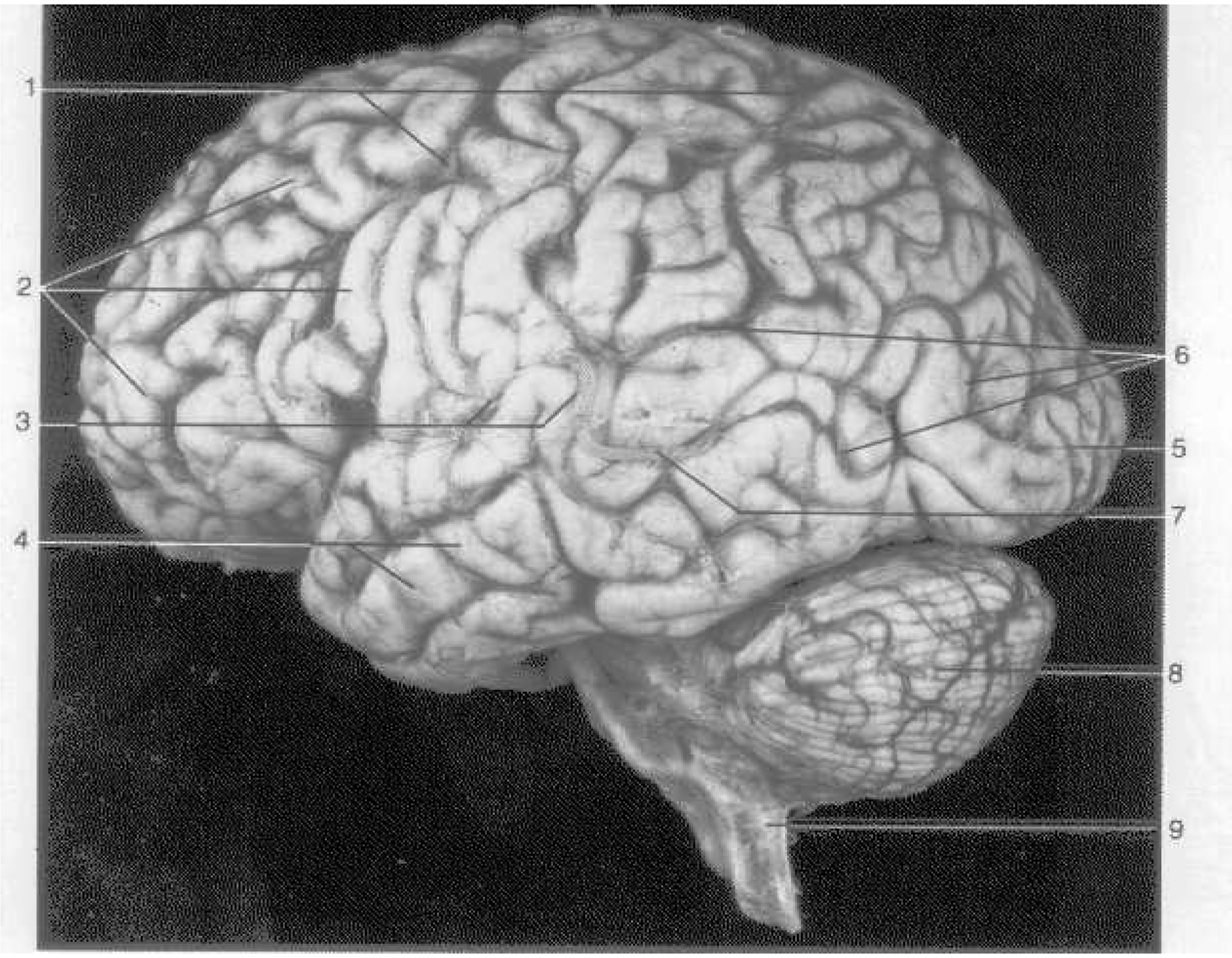


Figure 1-3: A Photograph of an Isolated Cadaver Brain. Labels 1-9 Point to Various Brain Structures.

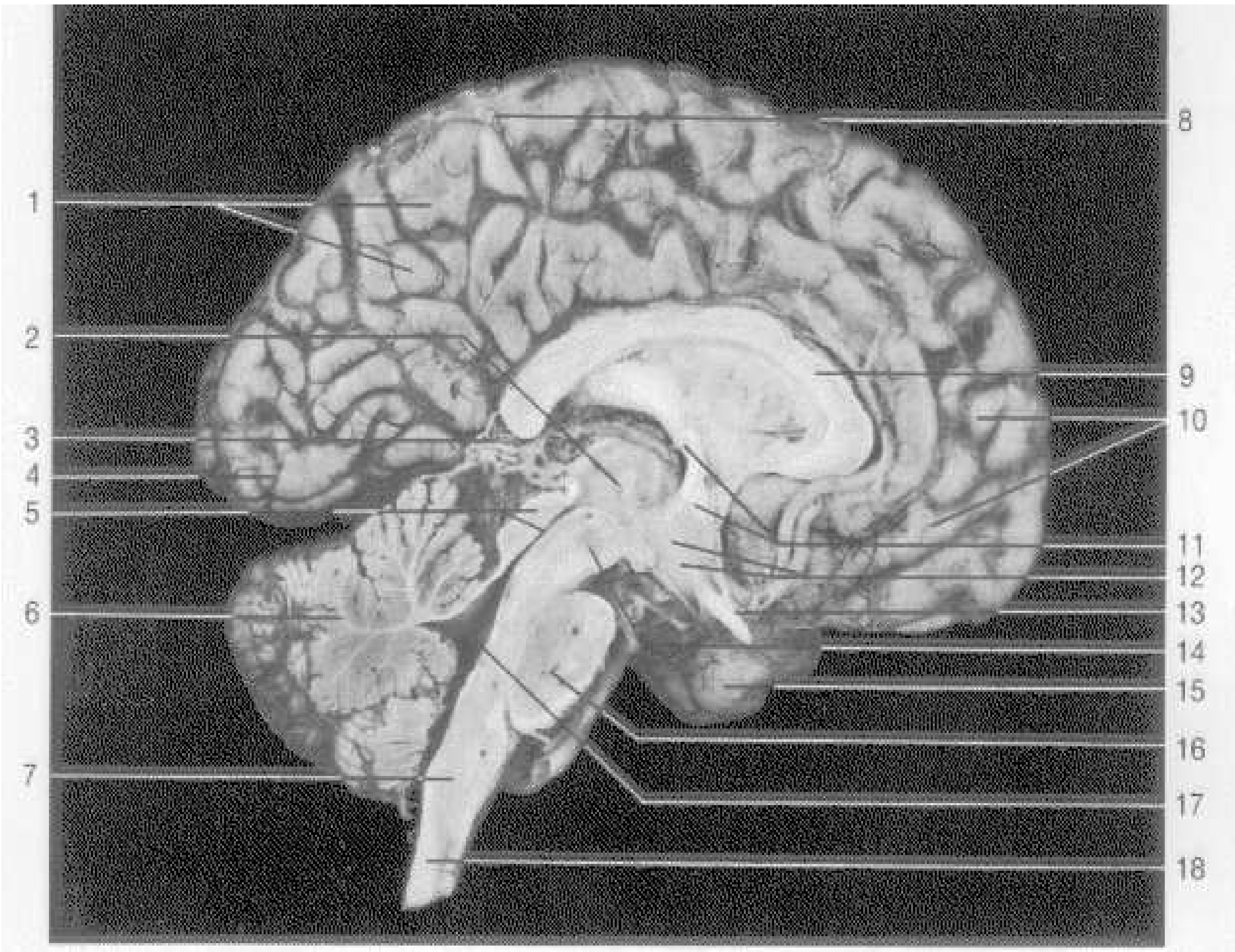


Figure 1-4: A Photograph of a Section through an Isolated Cadaver Brain. Labels 1-18 Point to Various Brain Structures.

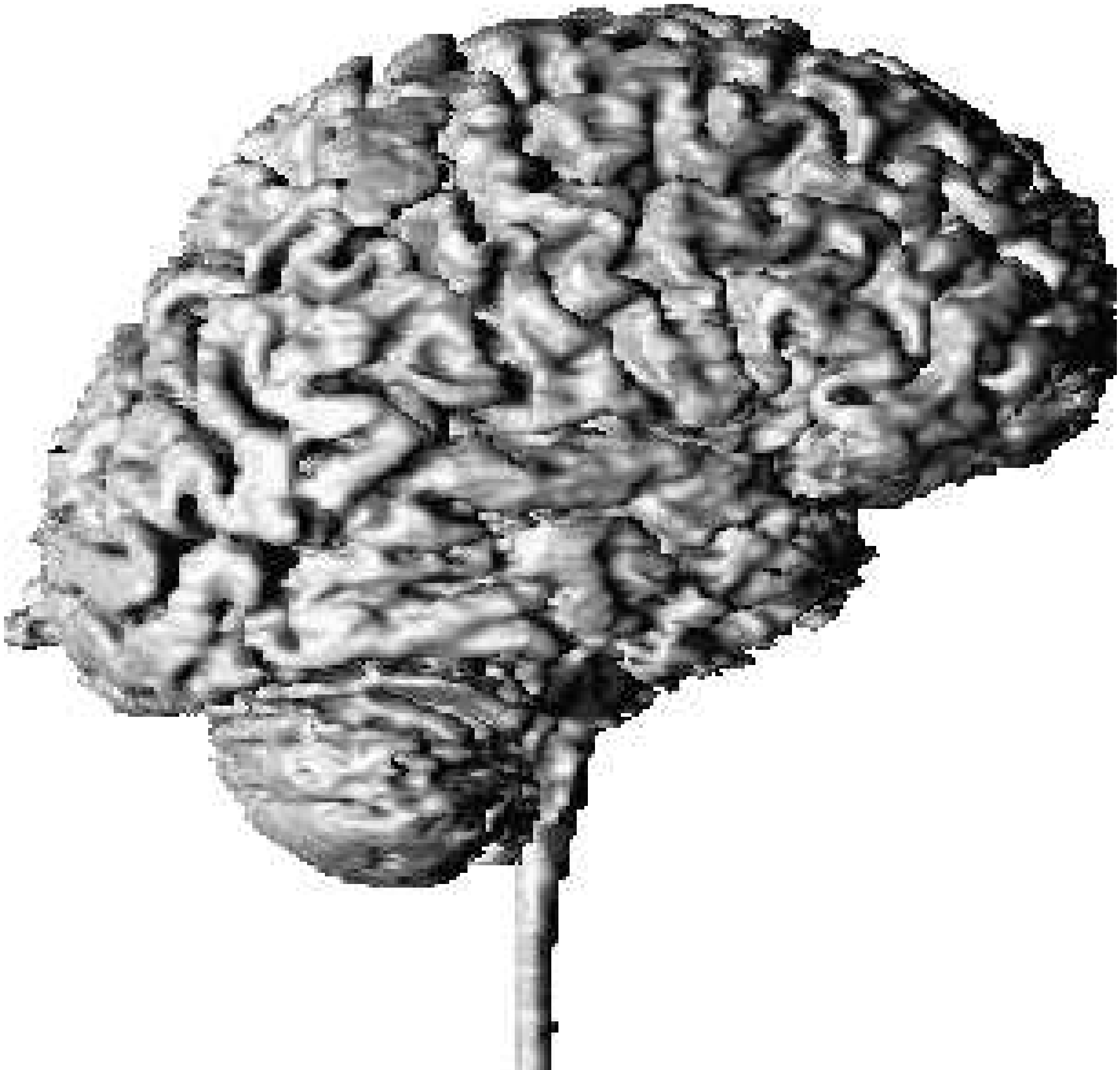


Figure 1-5: Segmented Brain Tissue from Magnetic Resonance Images Using Our System

Chapter 2

Background

Anatomical structures are frequently non-rigid and exhibit substantial morphological variation from subject to subject. Hence the task of segmenting these structures from medical imagery is one of identification of a region in an image with only approximate knowledge of its shape, size, gray level appearance, and spatial location. Different segmentation applications have available to them some knowledge in each of these categories, and the challenge is to combine it all in a manner such that the lack of information in one category is offset by the information in the others. In our method for segmentation of brain tissue from magnetic resonance images, we combine methods that exploit gray level, topological and spatial information in the images. The specific techniques we use are: EM segmentation for an intensity based correction and classification of the data, binary morphology and connectivity for incorporation of relative topological information, and balloon based deformable contours for addition of spatial information to the segmentation process.

2.1 EM Segmentation

Traditional intensity based segmentation relies on elements of the same tissue type having MR intensities that are clustered around a mean characteristic value, and relies on each cluster being well separated from other tissue clusters. Most MR scanners, however, have inhomogeneities in the imaging equipment, which give rise to a smoothly varying, non-linear gain field. While the human visual system easily compensates for this field, the gain can perturb the intensity distributions, causing them to overlap significantly and thus lead to substantial misclassification in traditional intensity based classification methods (e.g. [7]). (See Figure 2-1 for histograms of the intensity distribution in the head in an MR scan, and of the intensity distribution in segmented brain from the same scan. A visual comparison of the brain histogram with the complete head histogram shows that the the intensities for the brain neither form distinct peaks in the histogram, nor are clearly separated from the rest of the structures in the head in any other way, which are the two criteria that traditional intensity based segmentation relies on.)

Wells et al. [8] introduced a powerful statistical method that uses knowledge of tissue properties and RF coil inhomogeneity to correct for the spatial distortion in the signal. The method, EM segmentation, produces simultaneous classification of tissue classes and estimation of the gain field due to inhomogeneities in the RF coil in MRI. The following is a description of the details of the method presented in the same words as [9] and [8].

Intra- and inter-scan MRI intensity inhomogeneities are modeled with a spatially-varying factor called the *gain field* that multiplies the intensity data. The application of a logarithmic transformation to the intensities allows the artifact to be modeled as an additive *bias* field.

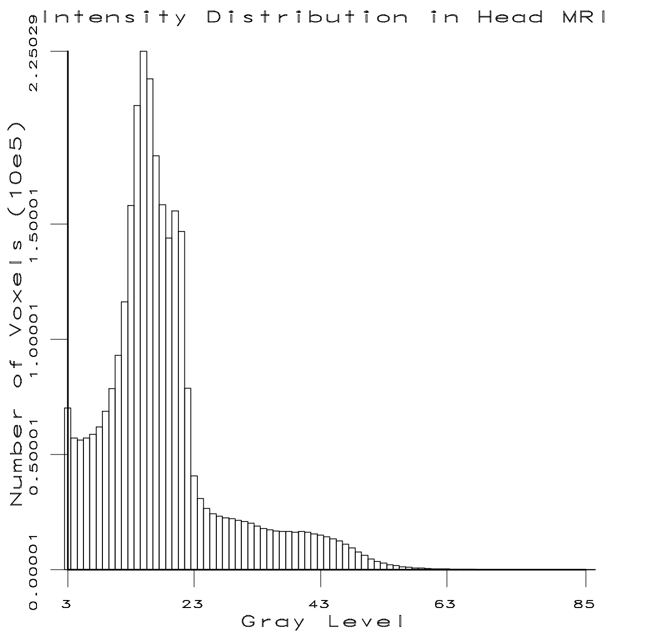
A Bayesian approach is used for estimating the bias field that represents the gain artifact in log-transformed MR intensity data. A logarithmic transform of the intensity data is computed as follows,

$$Y_i = g(X_i) = (\ln([X_i]_1), \ln([X_i]_2), \ldots, \ln([X_i]_m))^T \quad , \tag{2.1}$$

where X_i is the observed MRI signal intensity at the i-th voxel, and m is the dimension of the MRI signal. [1]

The distribution for observed values is modeled as a normal distribution with the

[1]For example, m will be two for a double-echo MRI acquisition.



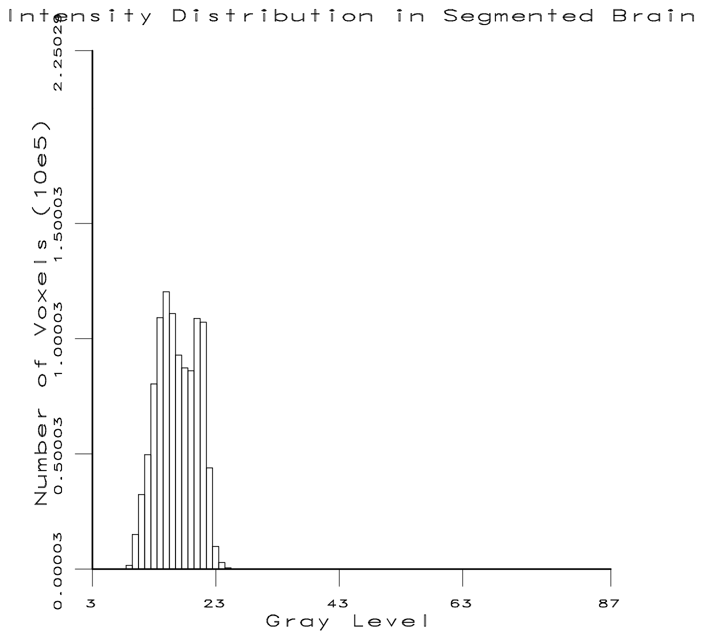


Figure 2-1: Intensity Distributions for the Head and Brain in MRI

incorporation of an explicit bias field:

$$p(Y_i \mid \Gamma_i, \beta_i) = \mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i) \ \ , \tag{2.2}$$

where

$$\mathrm{G}_{\psi_{\Gamma_i}}(x) \equiv (2\pi)^{-\frac{m}{2}} |\psi_{\Gamma_i}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} x^T \psi_{\Gamma_i}^{-1} x\right)$$

is the m-dimensional Gaussian distribution with variance ψ_{Γ_i}, and where

Y_i	is the observed log-transformed intensities at the i^{th} voxel
Γ_i	is the tissue class at the i^{th} voxel[2]
$\mu(x)$	is the mean intensity for tissue class x
ψ_x	is the covariance matrix for tissue class x
β_i	is the bias field at the i^{th} voxel.

Here Y_i, $\mu(x)$, and β_i are represented by m-dimensional column vectors, while ψ_x is represented by an m-by-m matrix. Note that the bias field has a separate value for each component of the log-intensity signal at each voxel. In words, Equation 2.2 states that the probability of observing a particular image intensity, given knowledge of the tissue class and the bias field, is given by a gaussian distribution centered at the biased mean intensity for the class.

A stationary prior (before the image data is seen) probability distribution on tissue class is used, it is denoted

$$p(\Gamma_i) \ \ . \tag{2.3}$$

If this probability is uniform over tissue classes, this method devolves to a maximum-likelihood approach to the tissue classification component.

The entire bias field is denoted by $\beta = (\beta_0, \beta_1, \ldots, \beta_{n-1})^T$, where n is the number of voxels of data. The bias field is modeled by a n-dimensional zero mean Gaussian prior probability density. This model captures the smoothness that is apparent in

[2] For example, when segmenting brain tissue $\Gamma_i \in \{$*white-matter, gray-matter, cerebro-spinal fluid*$\}$.

these inhomogeneities:

$$p(\beta) = \mathrm{G}_{\psi_\beta}(\beta) \quad , \tag{2.4}$$

where

$$\mathrm{G}_{\psi_\beta}(x) \equiv (2\pi)^{-\frac{n}{2}} |\psi_\beta|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} x^T \psi_\beta^{-1} x\right)$$

is the n-dimensional Gaussian distribution. The $n \times n$ covariance matrix for the entire bias field is denoted ψ_β. Although ψ_β is too large to manipulate directly in practice, it is shown in [9] that tractable estimators result when ψ_β is chosen so that its inverse is banded. Having a banded inverse implies that only the relationship between nearby bias components contributes to the probability, while more distant bias values are independent of one another.

It is assumed that the bias field and the tissue classes are statistically independent, which follows if the intensity inhomogeneities originate in the equipment. Using the definition of conditional probability the joint probability on intensity and tissue class conditioned on the bias field is obtained as follows,

$$p(Y_i, \Gamma_i \mid \beta_i) = p(Y_i \mid \Gamma_i, \beta_i) p(\Gamma_i) \quad , \tag{2.5}$$

and the conditional probability of intensity alone is obtained by computing a marginal over tissue class:

$$p(Y_i \mid \beta_i) = \sum_{\Gamma_i} p(Y_i, \Gamma_i \mid \beta_i) = \sum_{\Gamma_i} p(Y_i \mid \Gamma_i, \beta_i) p(\Gamma_i) \quad . \tag{2.6}$$

Thus, this modeling has led to a class-independent intensity model that is a mixture of Gaussian populations (one population for each tissue class). Since this model is a Gaussian mixture, rather than a pure Gaussian, the estimators derived below will be non-linear.

Statistical independence of the voxel intensities is assumed (in other words, the noise in the MR signal is spatially white). The probability density for the entire image is then written as

$$p(Y \mid \beta) = \prod_i p(Y_i \mid \beta_i) \quad . \tag{2.7}$$

Bayes' rule is then be used to obtain the posterior probability of the bias field, given observed intensity data as follows,

$$p(\beta \mid Y) = \frac{p(Y \mid \beta)p(\beta)}{p(Y)} \quad , \tag{2.8}$$

where $p(Y)$ is an unimportant normalizing constant.

Having obtained the posterior probability on the bias field, the maximum-a-posteriori (MAP) principle is used to formulate an estimate of the bias field as the value of β having the largest posterior probability,

$$\hat{\beta} = \arg\max_{\beta} p(\beta \mid Y) \quad . \tag{2.9}$$

A necessary condition for a maximum of the posterior probability of β is that its gradient with respect to β be zero. An equivalent zero-gradient condition on the logarithm of the posterior probability is used,

$$\left[\frac{\partial}{\partial[\beta_i]_k} \ln p(\beta \mid Y)\right]_{\beta=\hat{\beta}} = 0 \qquad \forall_{i,k} \quad , \tag{2.10}$$

where $[\beta_i]_k$ is the k-th component of the bias field at voxel i. Installing the statistical modeling of Equations 2.2-2.9 yields the following expression for the zero gradient condition:

$$\left[\frac{\partial}{\partial[\beta_i]_k}\left(\sum_j \ln p(Y_j \mid \beta_j) + \ln p(\beta)\right)\right]_{\beta=\hat{\beta}} = 0 \qquad \forall_{i,k} \quad .$$

Since only the i-th term of the sum depends on β_i, differentiating the logarithms gives

$$\left[\frac{\frac{\partial}{\partial[\beta_i]_k} p(Y_i \mid \beta_i)}{p(Y_i \mid \beta_i)} + \frac{\frac{\partial}{\partial[\beta_i]_k} p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}} = 0 \qquad \forall_{i,k} \quad .$$

Using Equations 2.2 and 2.6, this may be re-written as

$$\left[\frac{\sum_{\Gamma_i} p(\Gamma_i)\frac{\partial}{\partial[\beta_i]_k}\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)}{\sum_{\Gamma_i} p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)}+\frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}}=0 \qquad \forall_{i,k} \ .$$

Differentiating the Gaussian expression in the first term yields

$$\left[\frac{\sum_{\Gamma_i} p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)\left[\psi_{\Gamma_i}^{-1}(Y_i-\mu(\Gamma_i)-\beta_i)\right]_k}{\sum_{\Gamma_i} p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)}+\frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}}=0 \qquad \forall_{i,k} \ .$$

This expression may be written more compactly as

$$\left[\sum_j W_{ij}\left[\psi_j^{-1}(Y_i-\mu_j-\beta_i)\right]_k+\frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}}=0 \qquad \forall_{i,k} \ , \tag{2.11}$$

with the following definition of W_{ij}, (which are called the *weights*),

$$W_{ij}\equiv\frac{\left[p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)\right]_{\Gamma_i=tissue\text{-}class\text{-}j}}{\sum_{\Gamma_i} p(\Gamma_i)\mathrm{G}_{\psi_{\Gamma_i}}(Y_i-\mu(\Gamma_i)-\beta_i)} \ , \tag{2.12}$$

where subscripts i and j refer to voxel index and tissue class respectively, and defining

$$\mu_j\equiv\mu(tissue\text{-}class\text{-}j)$$

as the mean intensity of tissue class j. Equation 2.11 may be re-expressed as follows,

$$\left[\sum_j W_{ij}\left[\psi_j^{-1}(Y_i-\mu_j)\right]_k-\sum_j W_{ij}\left[\psi_j^{-1}\beta_i\right]_k+\frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}}=0 \qquad \forall_{i,k} \ ,$$

or as

$$\left[\left[\overline{R}_i\right]_k-\left[\overline{\psi^{-1}}_{ii}\beta_i\right]_k+\frac{\frac{\partial}{\partial[\beta_i]_k}p(\beta)}{p(\beta)}\right]_{\beta=\hat{\beta}}=0 \qquad \forall_{i,k} \tag{2.13}$$

with the following definitions for the mean residual,

$$\overline{R}_i\equiv\sum_j W_{ij}\psi_j^{-1}(Y_i-\mu_j) \ , \tag{2.14}$$

and the mean inverse covariance,

$$\overline{\psi^{-1}}_{ik} \equiv \begin{cases} \sum_j W_{ij}\psi_j^{-1} & \text{if } i = k \\ 0 & \text{otherwise} \end{cases} \quad . \tag{2.15}$$

The mean residuals and mean inverse covariances defined above are averages taken over the tissue classes, weighted according to W_{ij}.

Equation 2.13 may be re-expressed in matrix notation as

$$\left[\overline{R} - \overline{\psi^{-1}}\beta + \frac{\nabla_\beta p(\beta)}{p(\beta)} \right]_{\beta=\hat{\beta}} = \mathbf{0} \quad .$$

Differentiating the last term yields the following,

$$\overline{R} - \overline{\psi^{-1}}\hat{\beta} - \psi_\beta^{-1}\hat{\beta} = \mathbf{0} \quad .$$

Finally, the zero-gradient condition for the bias field estimator may be written concisely as

$$\hat{\beta} = H\overline{R} \quad , \tag{2.16}$$

where the linear operator H is defined by

$$H \equiv \left[\overline{\psi^{-1}} + \psi_\beta^{-1} \right]^{-1} \quad , \tag{2.17}$$

that is, the bias field estimate is derived by applying the linear operator H to the mean residual field, and H is determined by the mean covariance of the tissue class intensities and the covariance of the bias field.

The bias field estimator of Equation 2.16 has some resemblance to being a linear estimator in Y of the bias field β. It is not a linear estimator, however, owing to the fact that the W_{ij} (the "weights") that appear in the expression for $\overline{R}$ and H are themselves non-linear functions of Y (Equation 2.12).

The statistical modeling presented above results in a formulation of the problem of estimating the bias field as a non-linear optimization problem embodied in Equation

2.16. This optimization depends on the mean residual of observed intensities and the mean intensity of each tissue class, and on the mean covariance of the tissue class intensities and the covariance of the bias field. Next is described the approach used to obtaining solutions (estimates).

The Expectation-Maximization (EM) algorithm is used to obtain bias field estimates from the non-linear estimator of Equation 2.16. It is often used in estimation problems where some of the data are "missing." In this application, the missing data is knowledge of the tissue classes. (If they were known, then estimating the bias field would be straightforward.)

In this application, the EM algorithm iteratively alternates evaluations of the expressions appearing in Equations 2.16 and 2.12,

$$W_{ij} \leftarrow \frac{\left[p(\Gamma_i) \mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)\right]_{\Gamma_i = \textit{tissue-class-j}}}{\sum_{\Gamma_i} p(\Gamma_i) \mathrm{G}_{\psi_{\Gamma_i}}(Y_i - \mu(\Gamma_i) - \beta_i)} \quad , \tag{2.18}$$

$$\hat{\beta} \leftarrow H\overline{R} \quad . \tag{2.19}$$

In other words, Equation 2.12 is used to estimate the weights given an estimated bias field, then Equation 2.19 is used to estimate the bias, given estimates of the weights.

As frequently occurs in application of the EM algorithm, the two components of the iteration have simple interpretations. Equation 2.18 (the *E-Step*) is equivalent to calculating the posterior tissue class probabilities (a good indicator of tissue class) when the bias field is known. Equation 2.19 (the *M-Step*) is equivalent to a MAP estimator of the bias field when the tissue probabilities W are known.

The iteration may be started on either expression. Initial values for the weights will be needed to start with Equation 2.19, and initial values for the bias field will be needed to start with Equation 2.18.

In principle, given $\mu(\Gamma_i)$, ψ_β, and ψ_j, we could use the EM algorithm to obtain the needed estimates. In practice, we cannot directly measure ψ_β, and iterated moving-average low-pass filters have been used for the operator H in Equation 2.19. These filters have a low computational cost that is independent of the width of their support (amount of smoothing). That the optimal H is a linear low-pass filter, when tissue class is constant is argued in [9].

This method has been used to successfully segment 1000 scans in a multiple sclerosis study without manual intervention, which makes it the most aggressively tested intensity based classifier around. Figure 2-2 shows the results of the EM segmenter on a sagittal section obtained using a surface coil. In the top row, on the left is the input image and on the right is the initial gray matter probability, which is equivalent to the use of linear methods of intensity correction. In the second row, on the left is the final probability for gray matter obtained using EM segmentation, and on the right is the bias estimate. In the third row, on the left is the resultant classification of the input image, and on the right is the final intensity corrected image obtained from EM segmentation.

2.2 Mathematical Morphology

Image morphology provides a way to incorporate neighborhood and distance information into algorithms (see [10] [11] for detailed treatment of morphological operators). The basic idea in morphology is to convolve an image with a given mask (known as the structuring element), and to binarize the result of the convolution using a given function. Choice of convolution mask and binarization function depend on the particular morphological operator being used.

Binary morphology has been used in several segmentation systems, and we provide here functional descriptions of morphological as applicable in this work.

- **Erosion:** An erosion operation on an image I containing labels 0 and 1, with a structuring element S, changes the value of pixel i in I from 1 to 0, if the result of convolving S with I, centered at i, is less than some predetermined value. We have set this value to be the area of S, which is basically the number of pixels that are 1 in the structuring element itself. The structuring element (also known as the *erosion kernel*) determines the details of how a particular

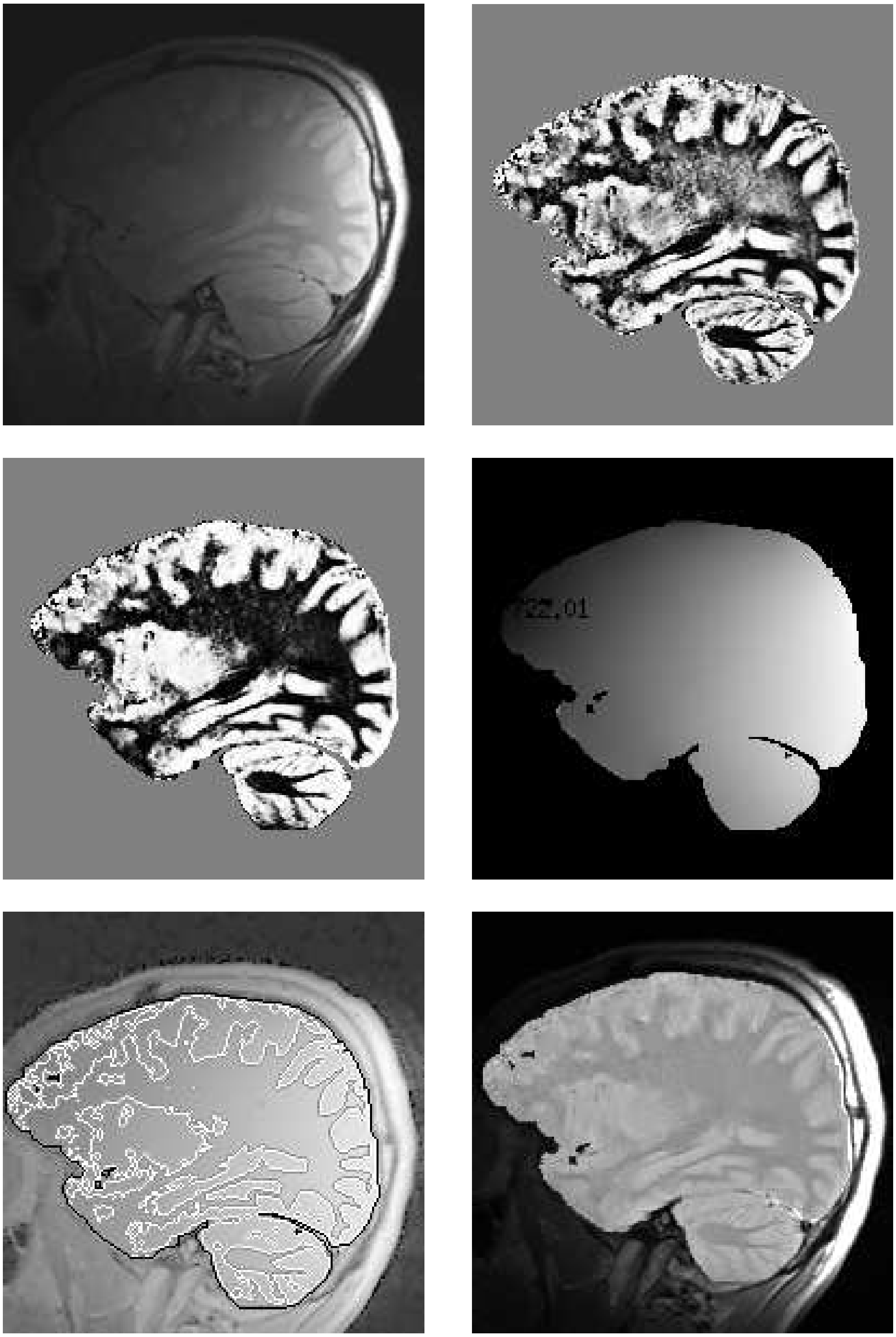


Figure 2-2: Results of EM Segmentation (Borrowed from [8])

erosion thins boundaries.

- **Dilation** Dual to erosion, a dilation operation on an image I containing labels 0 and 1, with a structuring element S, changes the value of pixel i in I from 0 to 1, if the result of convolving S with I, centered at i, is more than some predetermined value. We have set this value to be zero. The structuring element (also known as the *dilation kernel*) determines the details of how a particular dilation grows boundaries in an image.

- **Conditional Dilation** A conditional dilation is a dilation operation with the added condition that only pixels that are 1 in a second binary image, I_c, (the image on which the dilation is conditioned), will be turned changed to 1 by the dilation process. It is equivalent to masking the results of the dilation by the image I_c.

- **Opening** An opening operation consists of an erosion followed by a dilation with the same structuring element.

- **Closing** A closing operation consists of a dilation followed by an erosion with the same structuring element.

Figure 2-3 shows (from top to bottom) a binarized MR cross section, erosion of the MR image with a circular structuring element of radius 3, dilation of the largest connected component in the eroded image with a circular structuring element of radius 4.

2.3 Deformable Models

Deformable models are a popular component of medical image segmentation systems due to their ability to encode approximate shape constraints. For segmentation purposes, anatomical structures can be modeled using stacks of deformable contours in 2D or using actual 3D deformable surfaces. Also known as active contour models, these provide a method for minimizing an objective function to obtain a contour of

Figure 2-3: Examples of Erosion and Dilation on a Binary Brain Image

interest, especially if an approximate location of the contour is available. In this section, we present some background on two of the earlier deformable contour models: snakes [12] and balloons [13] in order to familiarize the reader with the essential components of such models.

Snakes

A deformable contour is a planar curve which has an initial position and an objective function associated with it. A special class of deformable contours called *snakes* was introduced by Witkin, Kass and Terzopoulos [12] in which the initial position is specified interactively by the user and the objective function is referred to as the *energy* of the snake. By analogy to physical systems, the snake slithers to minimize its energy over time. This energy of the snake (E_{snake}) is expressed as a sum of two components: the *internal energy* of the snake ($E_{internal}$) and the *external energy* of the snake ($E_{external}$).

$$E_{snake} = E_{internal} + E_{external} \quad . \tag{2.20}$$

The internal energy term imposes a piecewise smoothness constraint on the snake by preferring low first and second derivatives along the contour as shown in Equation 2.21:

$$E_{internal} = \int_s (w_1(s)||v'(s)||^2 \ + \ w_2(s)||v''(s)||^2)ds \quad , \tag{2.21}$$

where s is arclength, derivatives are with respect to s, and $v(s)$ stands for the ordered pair $(x(s), y(s))$, which denotes a point along the contour. The choice of w_1 and w_2 reflects the penalty associated with first and second derivatives along the contour respectively. w_1 is also known as the coefficient of elasticity, and w_2 as the coefficient of rigidity for the snake.

The external energy term in Equation 2.20 is responsible for attracting the snake to interesting features in the image. The exact expression for $E_{external}$ would depend on the characteristics of the features of interest. For example, if points of high brightness gradient magnitude in the image are interesting for a particular image, then the expression for the external energy can be as follows:

$$E_{external} = -|| \bigtriangledown I(v(s))||^2 \quad .$$

Finding a local minima for E_{snake} from equation 2.20 corresponds to solving the following Euler-Lagrange equation for v:

$$-(w_1 v')' + (w_2 v'')'' + \bigtriangledown E_{external}(v) = 0 \quad .$$

with boundary conditions specifying if the snake is a closed contour, or the derivatives are discontinuous at the end points.

This equation is then written in matrix form as

$$Av = F \quad ,$$

where

$$F(v) = - \bigtriangledown E_{external} \quad . \tag{2.22}$$

Here A is a pentadiagonal banded matrix that represents the “stiffness” properties of the snake, v is the position vector of the snake, and F is gradient of the external energy of the snake, or the external force acting on it.

The evolution equation:

$$\frac{dv}{dt} - Av = F \quad ,$$

is solved to obtain the v that is closest to the initial position. As $\frac{dv}{dt}$ tends to zero, we get a solution to the system $Av = F$.

Formulating this evolution problem using finite differences with time step τ, we obtain a system of the form [13]:

$$(I + \tau A)v^t = v^{t-1} + \tau F(v^{t-1}) \quad ,$$

where v^t denotes the position vector of the snake at time t, and I is the identity matrix. The system is considered to have reached equilibrium when the difference between v^t and v^{t-1} is below some threshold.

Balloons

The Balloon model for deformable contours is a substantial extension of the snake model. It modifies the snake energy to include a "balloon" force, which can either be an inflation force, or a deflation force. The external force F in Equation 2.22 is changed to

$$F = k_1 \mathbf{n}(s) + k \frac{\nabla E_{external}}{||\nabla E_{external}||} \quad , \tag{2.23}$$

where $\mathbf{n}(s)$ is a unit vector normal to the contour at point $v(s)$, and $|k_1|$ is the amplitude of this normal force. Changing the sign of k_1 makes this normal force inflating or deflating. Comparing Equation 2.23 to Equation 2.22 shows that the balloon model uses only the direction of the gradient of the external force, while snakes employ the magnitude of this force. This normalization of the external force, along with careful selection of the magnitude of k, constrains the motion of each point on the contour to at most one pixel per iteration. This avoids instabilities due to time discretization, and is discussed in detail in [13].

Contrasting the Balloon and Snake models, we note that incorporation of the normal force in the balloon model allows the initial position of the contour to be further from the intended final position, while still enabling convergence. As well, in the balloon model, the initial position can lie either inside or outside the intended contour, while the snake model requires the initial position to surround the intended contour if regions of zero image force lie between the initial and intended contours, since snakes can only collapse to a point in the absence of image forces.

The deformable contour model we use in our segmentation method is a refinement on the balloon model and is presented in the next chapter.

Snakes and balloons have been typically used with creative external forces to segment various anatomical structures. Using snakes to track 2D contours in 3D, coupling of the snakes to region information in the image, using simulated annealing to find the global minimum of the snake energy, and imposing shape priors are notable extensions to the original snakes and are discussed in Chapter 6.

Chapter 3

Our Method

In this chapter we present the details of our method for segmentation of brain tissue from magnetic resonance imagery. As mentioned earlier, our method combines three distinct techniques from computer vision in a manner that leverages the strengths of each one. Specifically, the techniques we use are EM segmentation for correction of gain due to RF coil inhomogeneities in the data, binary morphology and connectivity to incorporate topological information, and active contours to add spatial knowledge to the segmentation process. We begin the discussion with a description of the input data to our algorithm, followed by a description of the model of the brain tissue that is implicit in our algorithm, and then present details of each of the three steps in our algorithm. We conclude the chapter with a brief section on a parallel implementation of the algorithm that we are currently using at the Surgical Planning Laboratory of Brigham and Women's hospital to routinely segment brains for clinical and research purposes.

3.1 Description of Input Data

We use magnetic resonance scans of the head as input to our algorithm (see Figure 3-1 for sample images). We are currently working with gradient echo images acquired using a General Electric Signa 1.5 Tesla clinical MR imager. The voxel size is 1 mm x 1 mm x 1.5 mm, and there are 256 x 256 x 124 voxels per data set (see [2] for details).

The signal at each voxel represents the aggregate response of the corresponding 1.5mm^3 volume in the patient's head. The 1.5mm^3 volume in the patient contains millions of cells, and in a normal (non-pathological) case these cells belong to various anatomical structures in the head such as brain tissue, cerebrospinal fluid (csf), meninges (the protective membranes surrounding the brain), skull, muscle, fat, skin, or air (ee Figure 3-2). Pathology introduces the additional classes of edema, tumor, hemorrhage, or other abnormality.

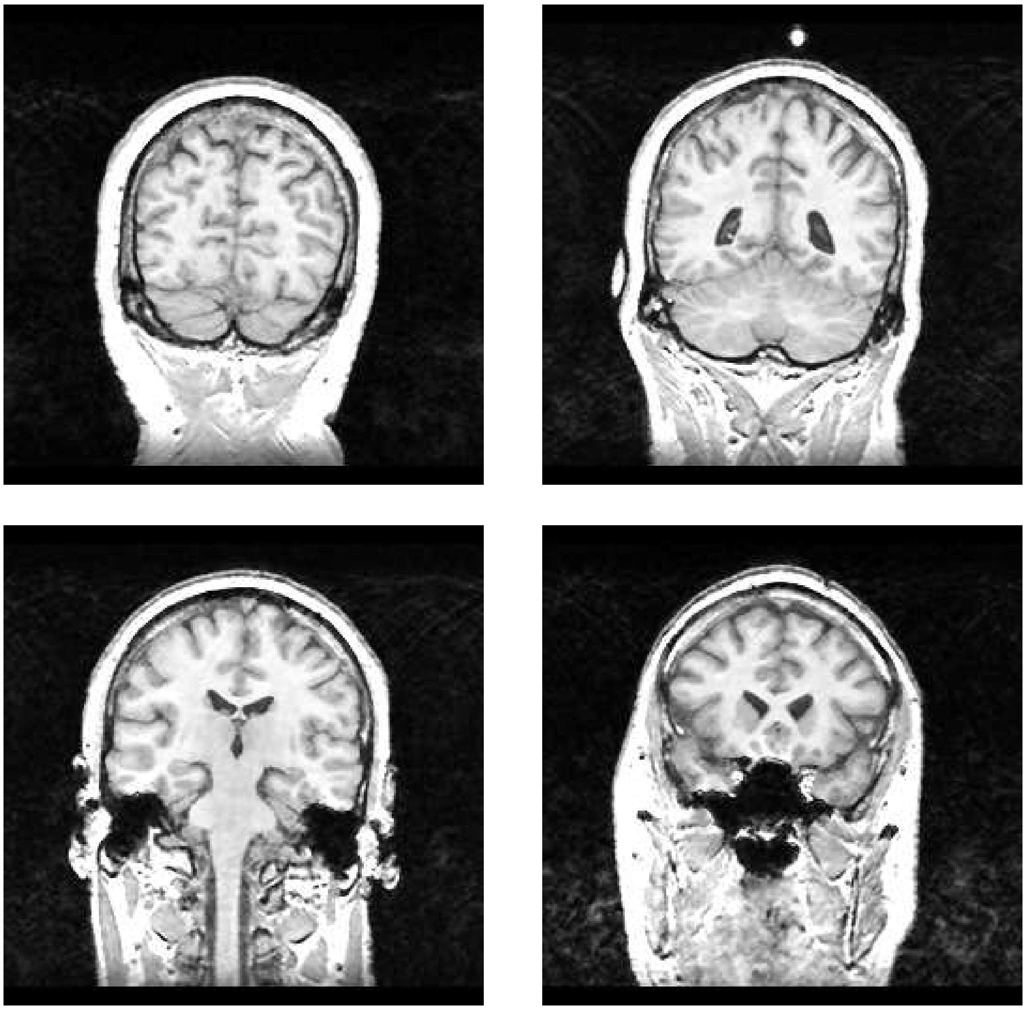


Figure 3-1: Sample Slices from a Gradient Echo MR Scan

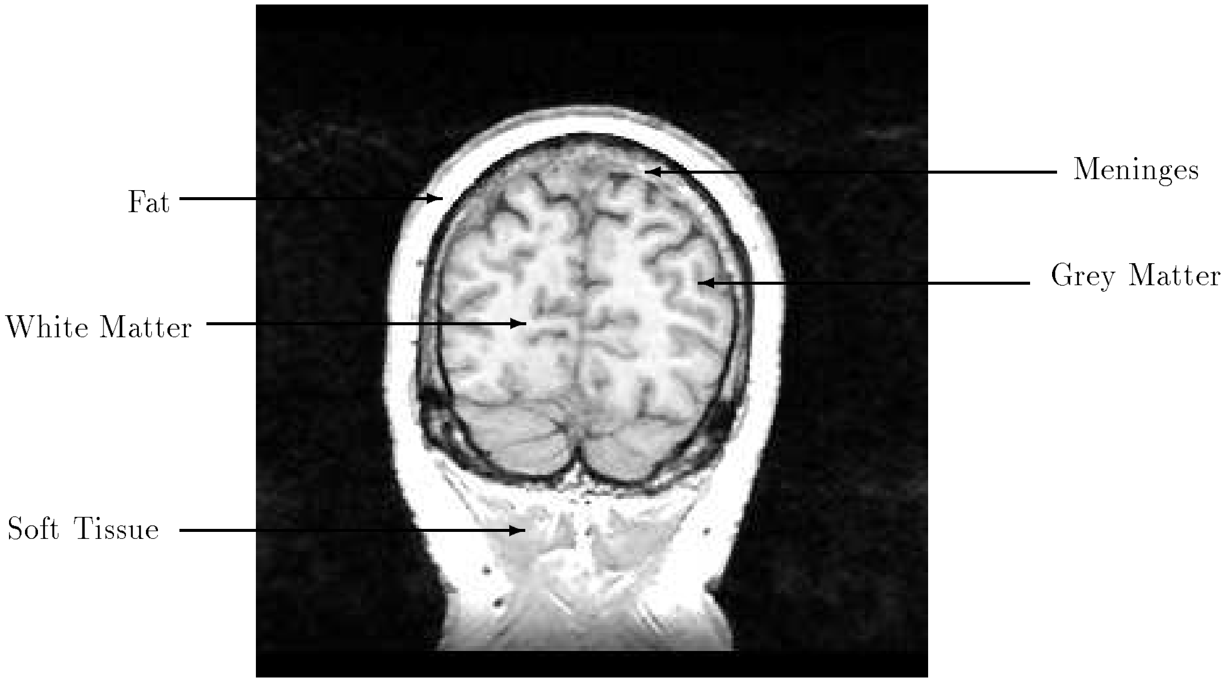


Figure 3-2: An Annotated Gradient Echo MR Slice (Air, Csf and Cranium are Dark in These Images)

3.2 Constructing a Model for the Brain

Our first task was to construct a model for the brain that would guide our segmentation process. This model is implicit in our algorithms and was constructed based on expert opinion from the Surgical Planning Lab, Brigham and Women's Hospital, and Harvard Medical School, in the form of conversations with experts or examination of segmentations done manually by them. We model the brain as the largest connected region (obtained by the morphological processing of intensity classified voxels) approximately in the center of the head surrounded by csf and meninges. The meninges are surrounded by the cranium. Blood vessels and nerves connect the cranium to the brain tissue. In particular, the connectors are, in order of increasing thickness: bridging veins from the cerebral cortex to dura, and from dura to the skull, the second cranial nerve, or optic nerve, the vertebral arteries around foramun magnum, and the external carotid artery in the region of the temporalis muscle. There is some natural overlap in the intensity distributions of brain versus non-brain structures. Additional overlap in the intensity distributions may be introduced due to two limitations of the imaging process: (1) presence of spatial inhomogeneities in the sensitivity of the scanning equipment introduces a smoothly varying non-linear gain artifact in the images, and (2) large differences in magnetic susceptibilities of adjacent tissue leads to distortion of the gradient magnetic field, and hence a susceptibility artifact in the images. In addition, spurious connections between the brain and the rest of the head may be introduced due to movement of the patient during the acquisition of the scan.

This model of the brain strongly suggests the use of intensity distributions as well as absolute and relative spatial arrangement of the various structures in the head to aid the segmentation process. In fact, as noted in the Chapter 6, several groups have approached this task with techniques that indicate a similar underlying model for the brain tissue [14] [3].

3.3 Segmenting the Brain Tissue

Using the model described above, we divide the segmentation of the brain into three steps.

The first step is to remove the gain artifact from the data and to classify the voxels into four classes: white matter, grey matter, csf, and skin (or fat), purely on the basis of their signal intensities. Due to natural overlap between intensity distributions of the various structures, misclassifications are likely at this stage. In particular, muscle is likely to be classified as gray matter, fat classified as white matter, and nerve fibers classified as white or gray matter.

The next step aims to reduce some of this misclassification by using neighborhood and connectivity information. It uses morphological operators to "shave off" the nerve fibers and muscles connecting the brain tissue to the cranium, and then uses connectivity to find the largest connected component of white and grey matter in the image. The strategy is that misclassified fat, muscle, and nerve fibers will get cut off from the central largest component, which is the brain tissue. Due to the variation in the size of the connectors from the brain tissue to the cranium, often the brain tissue is not isolated at the end of this step, which is when we use the third step.

The third step uses expert input to annihilate connections between the brain and spurious structures in a few carefully chosen slices of the data, and then employs region based deformable contours to propagate the manually drawn contours to the rest of the volume.

Sections 3.3.1, 3.3.2 and 3.3.3 provide details on each of the three steps of our segmenter.

3.3.1 Gain Correction

Initially, we use EM segmentation to correct for the gain introduced in the data by the imaging equipment. We use a single channel, non-parametric, multi-class implementation of the segmenter that is described in [8]. Training points are used for white matter, grey matter, csf and skin, and therefore the resultant tissue classifications correspond to these four classes. The output of this stage is a set of classified voxels.

Figures 3-3 and 3-4 show two separate runs of EM segmenter on two different MR scans. The top left image in each figure is the input given to the segmenter, and the rest of the images (top to bottom, left to right) show the tissue classifications generated by the EM segmenter in successive iterations. Typically convergence is achieved in 5-10 iterations.

3.3.2 Removal of Thin Connectors Using Morphological Operations

We have tried two methods for incorporating neighborhood information using morphological operations into the tissue-labeled image obtained from EM Segmentation. These methods are similar to the work in [15] [16].

First Method

- Perform an erosion operation on the input with a spherical structuring element with radius corresponding to the thickness of the connectors between brain and the cranium, so that it eliminates connections from the brain to any misclassified

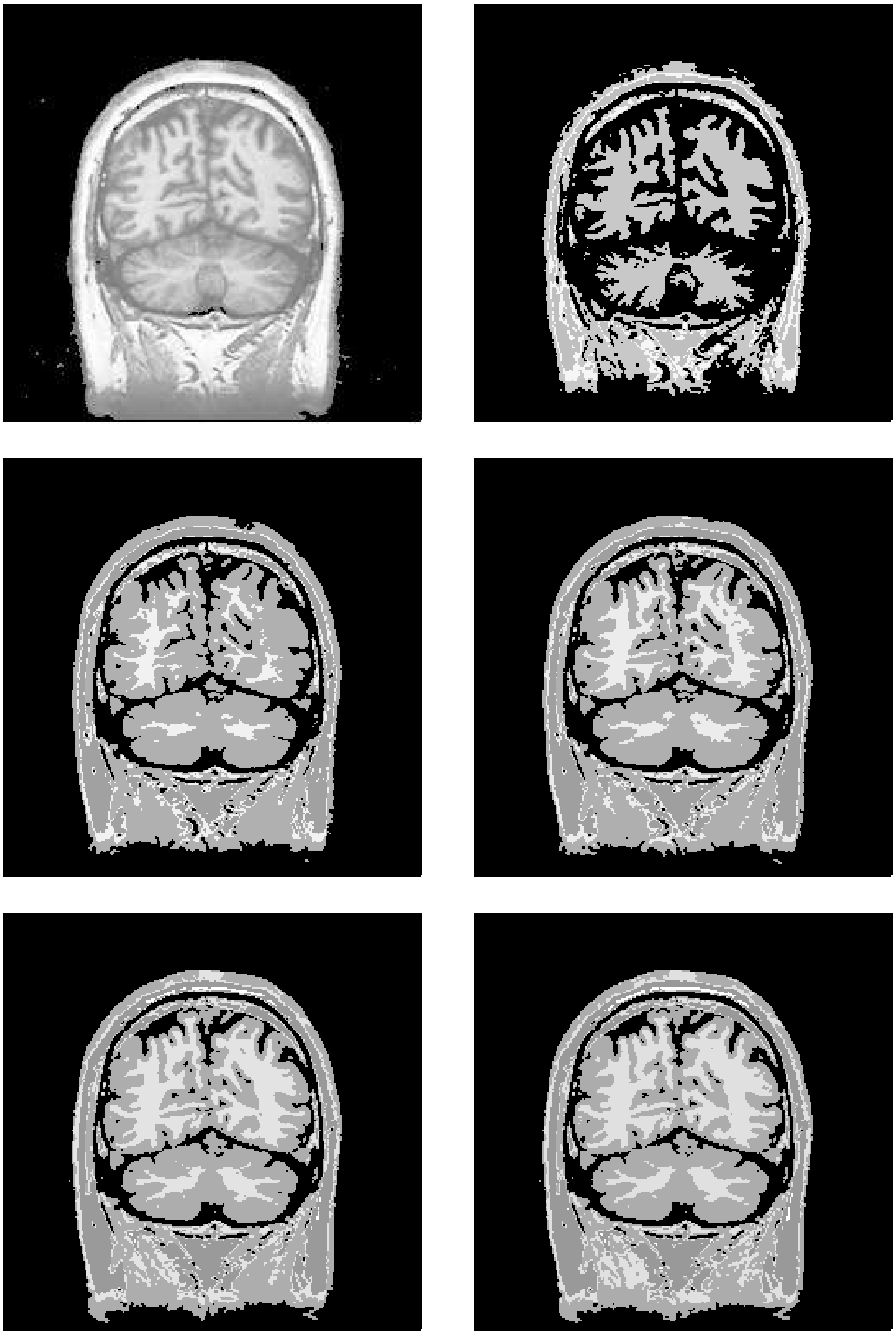


Figure 3-3: Top to Bottom, Left to Right: Input Image and Tissue Classification Generated by Successive Iterations of the EM Segmenter(white matter is brightest, grey matter is medium grey, and csf and air are black)

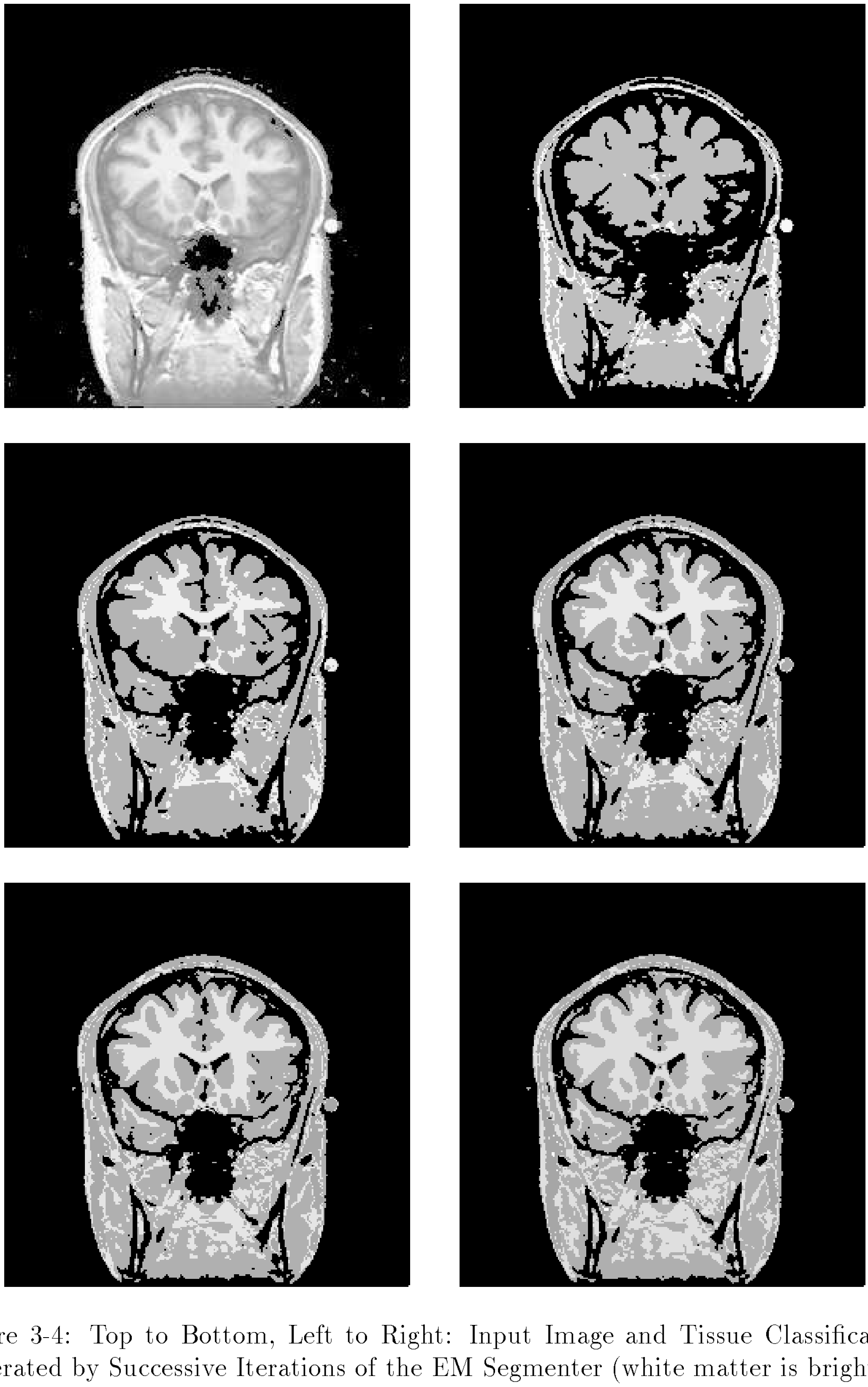


Figure 3-4: Top to Bottom, Left to Right: Input Image and Tissue Classification Generated by Successive Iterations of the EM Segmenter (white matter is brightest, grey matter is medium grey, and csf and air are black)

non-brain structure.

- Find the largest connected component with tissue labels corresponding to the brain.

- Dilate the brain component obtained in the previous step by a structuring element comparable in size to the one used in the erosion, conditioned on the brain labels in the input image. This corresponds approximately to restoring the boundary of the brain component that were distorted in the erosion step.

Second Method

- Perform an opening operation on the input with a spherical structuring element with radius corresponding to the thickness of the connectors between brain and the cranium, so that it eliminates connections from the brain to any misclassified non-brain structure.

- Find the largest connected component with tissue labels corresponding to white matter,

- Dilate the white matter component obtained in the previous step by a structuring element that reflects the nominal thickness of grey matter in the neo-cortex, conditioned on the grey matter labeling of the input image.

We usually use the first method in our routine segmentations because eroding the white matter in the second method results in loss of substantial information due to its highly convoluted surface characteristics.

The results of this stage is an improved segmentation of the tissue types, which incorporates rough shape information, and which has removed some of the artifacts due to pure intensity classification.

Figures 3-5 and 3-6 show the results of using the first method on the same two slices that were used to illustrate the EM segmenter in the previous section. Note that final result of the example shown in Figure 3-5 (the bottom right image) is indeed a segmentation of the brain tissue, while the end result of the example shown in Figure 3-6 has not succeeded in eliminating from consideration all the non-brain pixels that were classified as white matter or grey matter by the EM segmenter. In the next section we will present the third step of our algorithm which corrects the segmentation for the example shown in Figure 3-6. It should be pointed out here that the slices shown here are cross-sections of a full 3D data set, and all the algorithms used so far have been operational in 3D. This means that even though connectors between the brain and non-brain structures in Figure 3-6 are not visible in this slice, the two are connected in 3D.

3.3.3 Extraction of Brain Surface using Deformable Contour Models

The third step in our segmentation algorithm is the use of deformable contour models to refine the result of the brain tissue estimate obtained using morphological operations, since morphology is often unable to remove all connectors from the brain to the cranium (as shown in the example in Figure 3-6). The intuition behind this step is to incorporate substantial spatial information into the segmentation process via the manual initialization of the brain boundary in a few slices and use that to refine the results obtained thus far. We are exploring the possibility of constructing a spatial prior on the brain that would supply this information and obviate manual intervention, but meanwhile we do require expert navigation at this stage. For this step, the input volume is treated as a stack of parallel 2D slices where each slice shows a cross-section of the patient's head. We model the brain surface as a set of adjacent boundary contours for each slice of the scanned volume and use deformable contours to find each of these contours. We think of this step as operational in "2.5D" since the estimate of the brain is obtained by processing the data as a 3-dimensional volume, and smoothness of the brain boundary between adjacent slices is imposed by the propagation of contours between adjacent slice.

In this section, we first describe our experience with implementing the snake and balloon deformable contour models, and then present a variant on the balloon model,

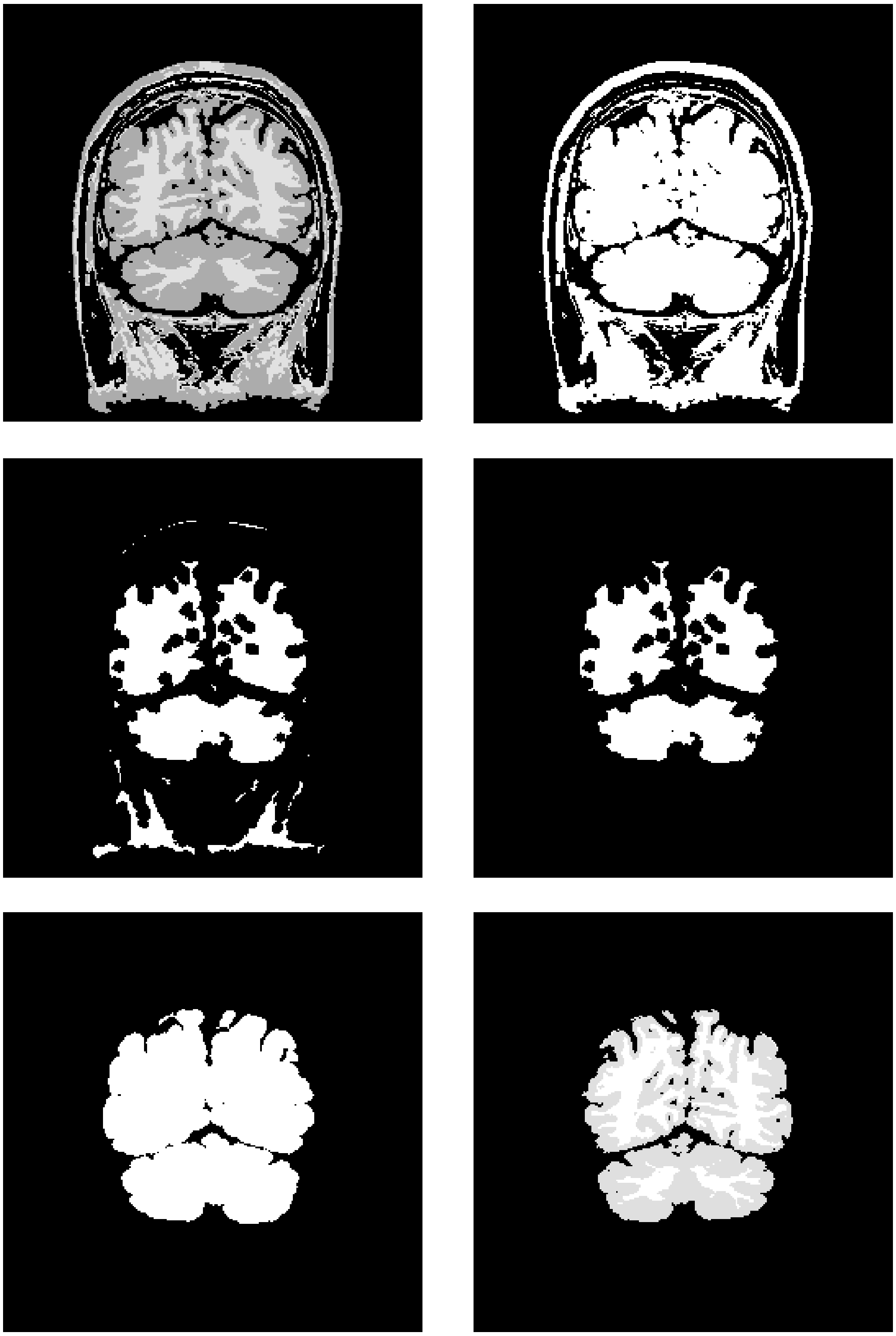


Figure 3-5: Top to Bottom, Left to Right: EM Segmentation from Figure 3-4, Binarized Image, Eroded Image, Largest Connected Component in Eroded Image, Dilated Connected Component, Conditionally Dilated Connected Component

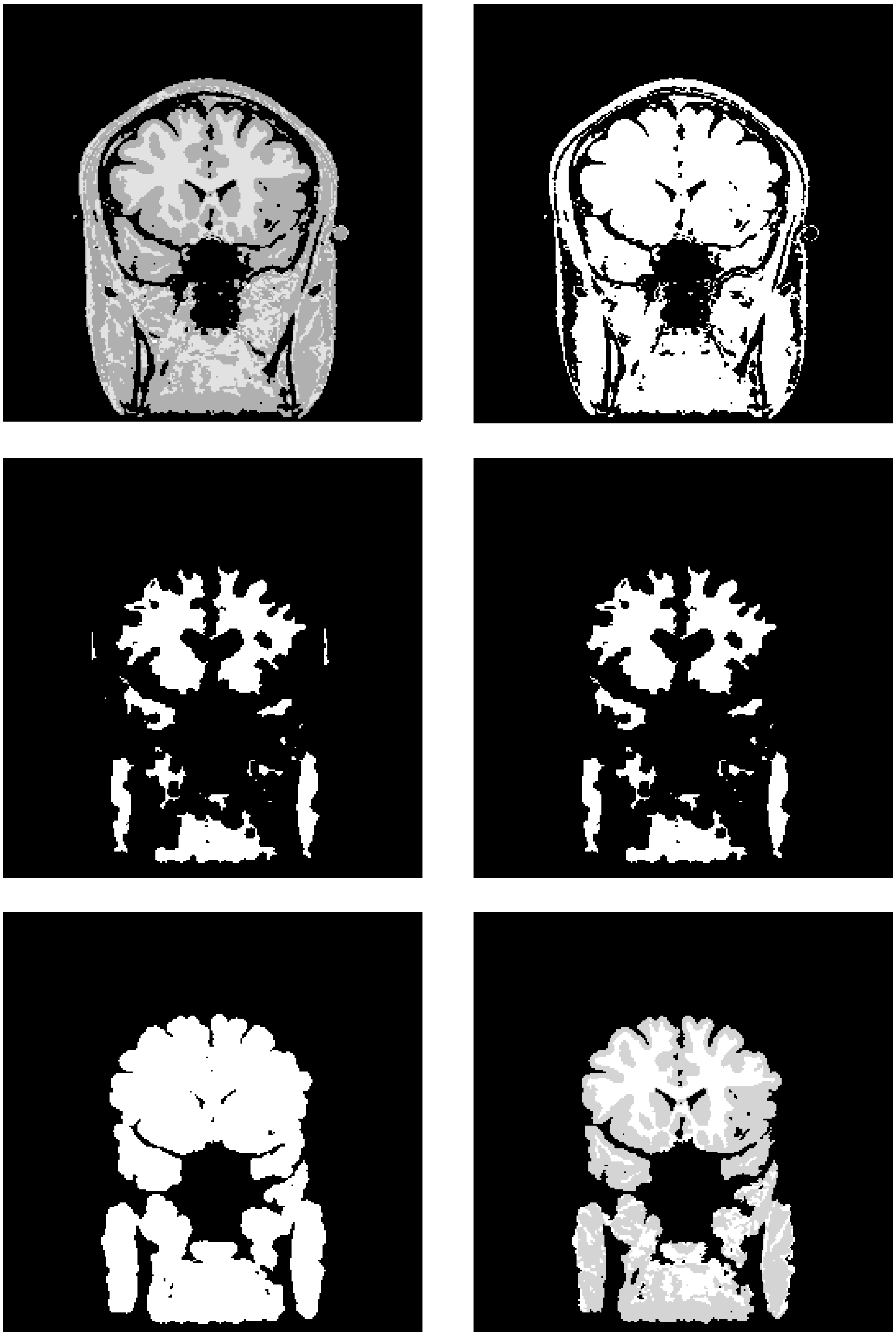


Figure 3-6: Top to Bottom, Left to Right: EM Segmentation from Figure 3-3, Binarized Image, Eroded Image, Largest Connected Component in Eroded Image, Dilated Connected Component, Conditionally Dilated Connected Component

customized for the extraction of the brain surface.

On Snakes

We found that snakes provide a useful framework for extraction of specific features from an image. The internal and external components of the snake energy helped us separate the specifications of the expected shape of the features of interest from specifications of the expected brightness patterns along the features of interest. However, extraction of the brain does not seem to be an application that best exploits the strengths of the snake model.

Snakes work well if the internal energy closely models shape characteristics of the contour, so that corruptions in the image forces can be correctly compensated for by the smoothness criterion. Translating this to the snake equations, this means that w_1 and w_2 in Equation 2.21 can greatly influence the behavior of the snake, and thus need to be determined carefully for each application. If an application requires the detection of smooth contours, large values of w_1 and w_2 adequately model the scenario, and snakes serve the task well. It is especially complicated to use the snake model when these coefficients vary along the contour, and the contour is allowed to change its length over time.

A contour as complex as the boundary of the brain would be inadequately modeled by constant values of w_1 and w_2 in the snake model. A more accurate model along these lines would probably contain more detailed information about how these coefficients should vary along the contour. However, it is not obvious to us how to build such a detailed model.

Another reason why we need a more specific model than snakes for our application is that the boundary of the brain is surrounded by points of high brightness gradient (the surrounding matter being soft tissue of the neck or the meninges), which makes insufficient the use of any of the functions suggested in [12] as the external energy terms of the evolution equation.

On Balloons

The balloon model, with its normal force which allows the initial position to be inside or outside the intended contour, came closer to our requirement, but needed further customization, since it treats the direction of the normal force as a parameter of the contour itself, and therefore, the contour can either expand or contract as a whole. This does not model our situation correctly, since we propagate manual initialization from slice to slice, and cannot guarantee that the initial position of the contour will always be completely inside or outside the expected equilibrium position. A simple addition to the balloon model, as described next, fixed this problem.

Customized Balloon Model

We customized the balloon model to exploit the estimate of the brain volume created using the first two steps of our segmentation process. Instead of using a fixed predetermined direction for the balloon force by selecting a sign for the factor k_1 in Equation 2.23 which is independent of the image characteristics, we define a signed balloon force direction vector, B, with one entry per voxel of the input data. The sign at the ith position of vector B indicates whether the voxel i exerts a force along the inward or outward normal to the evolving contour. This vector B is determined using the brain estimate obtained after applying the morphological operations as discussed in the previous section. If voxel i is classified as brain tissue at the end of the morphological routines, then B[i] gets a positive sign, so that the voxel pushes the contour in the direction of the normal towards the boundary of the brain estimate, otherwise B[i] gets a negative sign. The external force now becomes:

$$F = kB(s)\mathbf{n}(\mathbf{s}) \ ,$$

where $B(s)$ is the direction of the balloon force exerted by the image at the point $v(s)$ of the contour, and $\mathbf{n}(\mathbf{s})$ is the normal to the local tangent vector to the contour at s. The unsigned constant k is used to determine the magnitude of the balloon force. If $B(s)$ is positive, motion is along $\mathbf{n}(\mathbf{s})$, which corresponds to an inflating force applied to the contour at s. Negative $B(s)$ results in a deflating force applied to the contour at $v(s)$.

This force is used with the same evolution equation as before:

$$(I_d + \tau A)v^t = v^{t-1} + \tau F(v_{t-1}) \ , \tag{3.1}$$

Below we present some illustrations of iterating snakes using our customized balloon model. In Figure 3-7, the top left shows the result of the morphological operators as obtained in Figure 3-6 in the previous section. The top right image shows the initial location of the snake, which was provided manually. The last four images in Figure 3-7 (top to bottom, left to right), and the first five images of Figure 3-8 show positions of the snake as it iterates to convergence. The bottom right image in Figure 3-8 shows the cross-section of the brain enclosed by the final snake configuration. It should be noted here that even though Figures 3-7 and 3-8 illustrate the ability of the customized balloons to converge to an acceptable result from a very rough initialization, this does not imply that initialization can be arbitrary. Figures 3-9 and 3-10 show an example of a run with a different initial position that causes the snake to be stuck in a local minimum which causes the temporal lobes to be excluded from region enclosed by the final snake configuration.

We have presented a method for segmentation of brain tissue from MR images, that combines intensity, topological, and spatial information into the process. Next, we briefly describe the implementation of this method on a supercomputer, IBM Power Visualization System.

3.4 Implementation on IBM Power Visualization System

Implementation of the EM segmentation, morphological operations, and connectivity has been done on IBM POWER Visualization Server (PVS). In this section we present

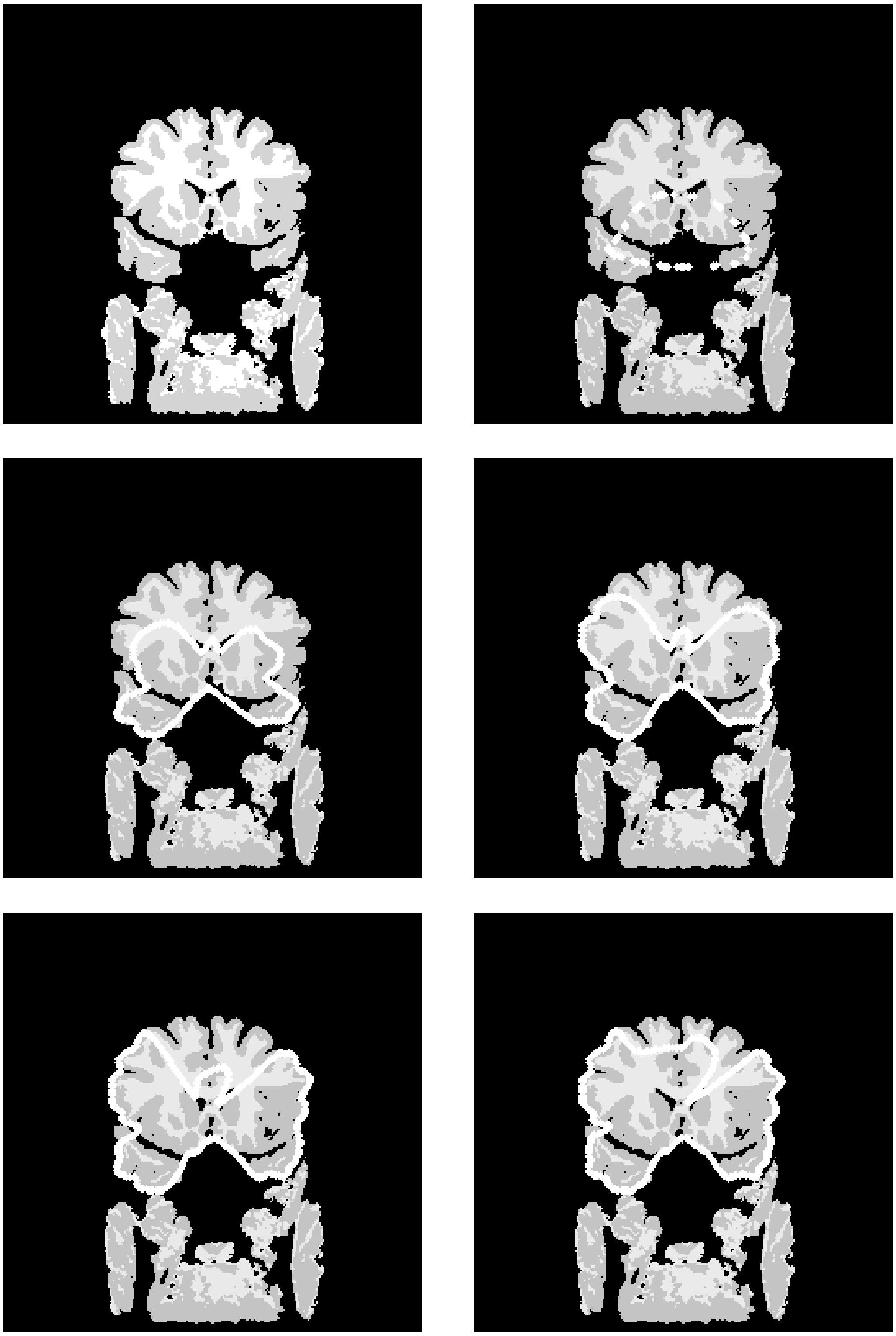


Figure 3-7: Top to Bottom, Left to Right: Result at the end of the Morphology Step, Interactively Specified Initial Position of Snake(in white), First Few Iterations of the Customized Balloon Model

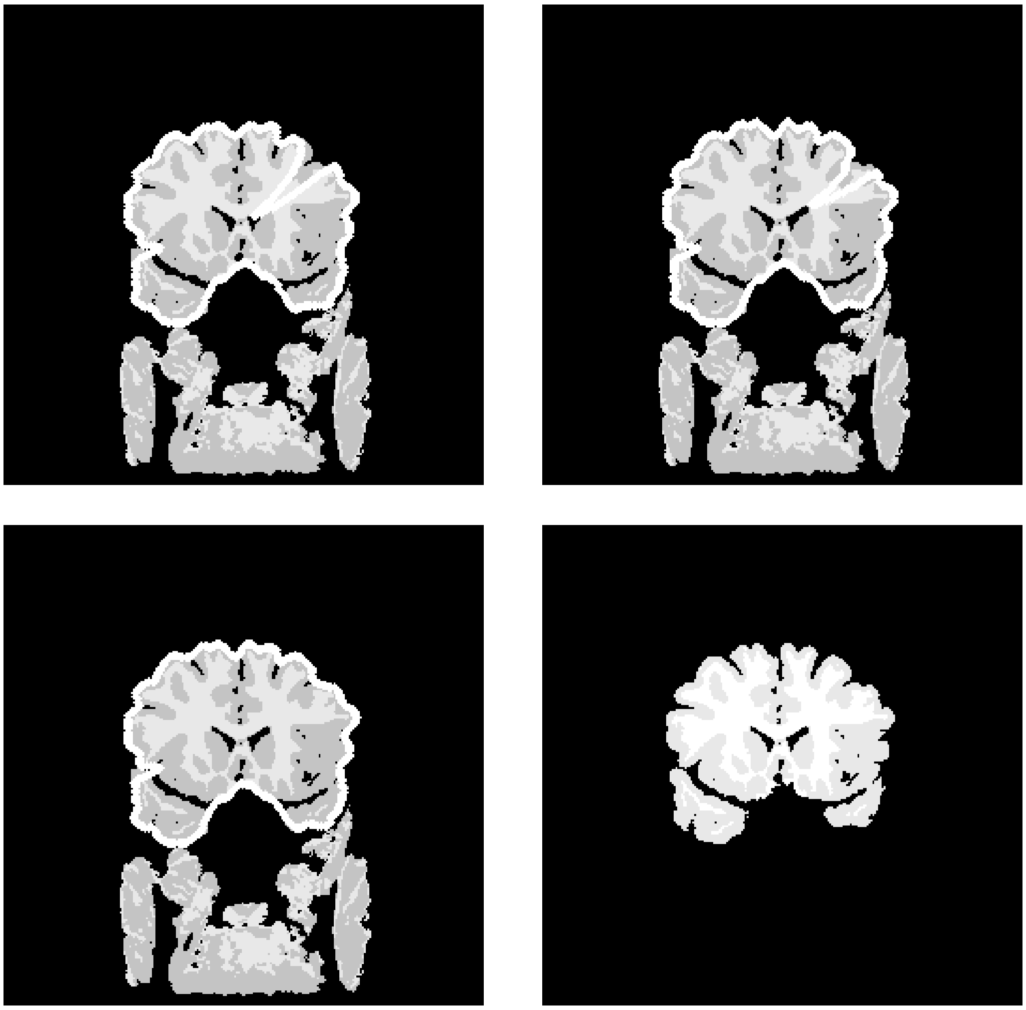


Figure 3-8: Top to Bottom, Left to Right: First 3 Images show Further Iterations of the Customized Balloon Model (continued from Figure 3-7) and the Bottom Right Image Shows the Brain Tissue Enclosed by the Final Position of the Snake.

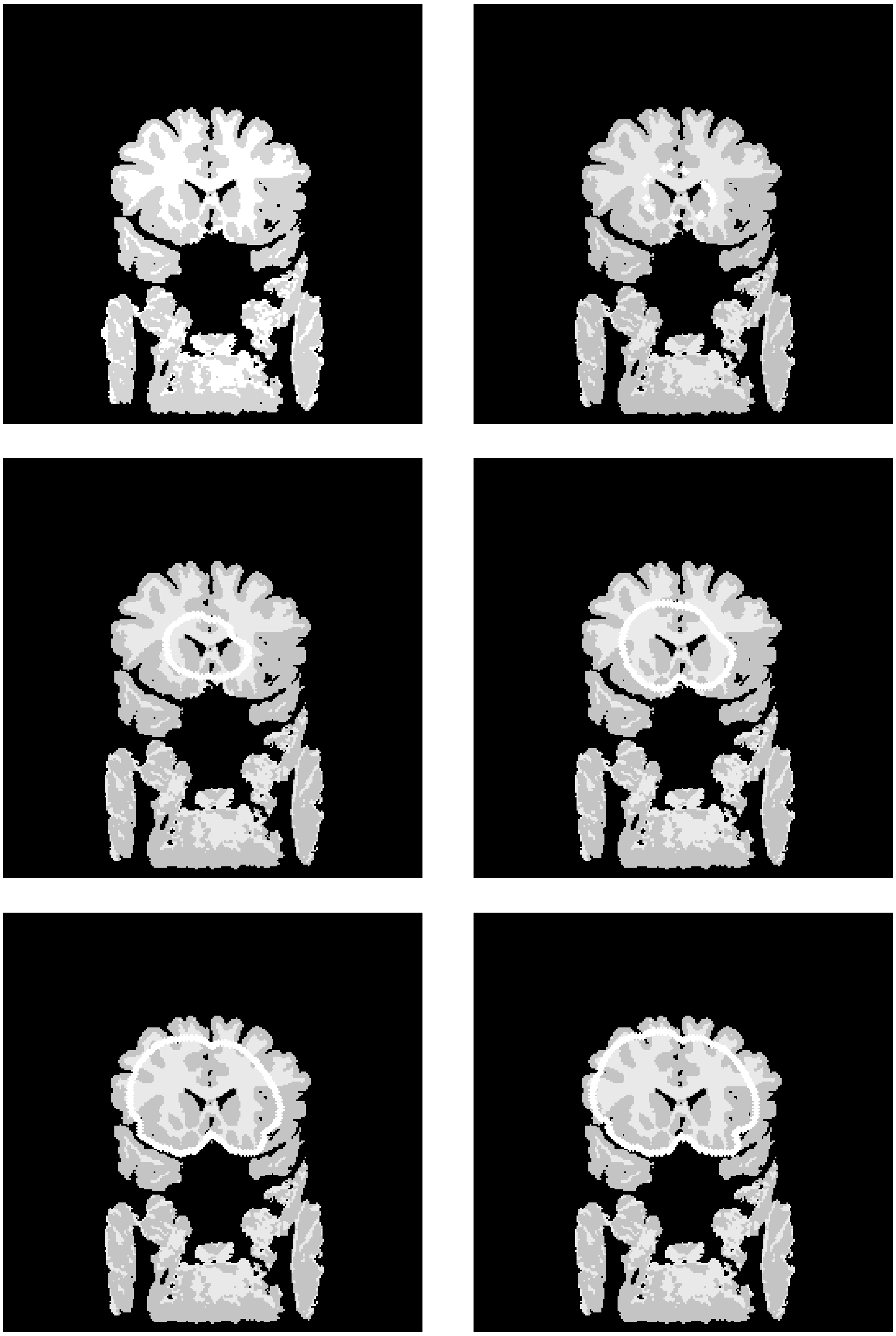


Figure 3-9: Top to Bottom, Left to Right: Result at the end of the Morphology Step, A Different Interactively Specified Initial Position of Snake(in white), A few Iterations of the Customized Balloon Model.

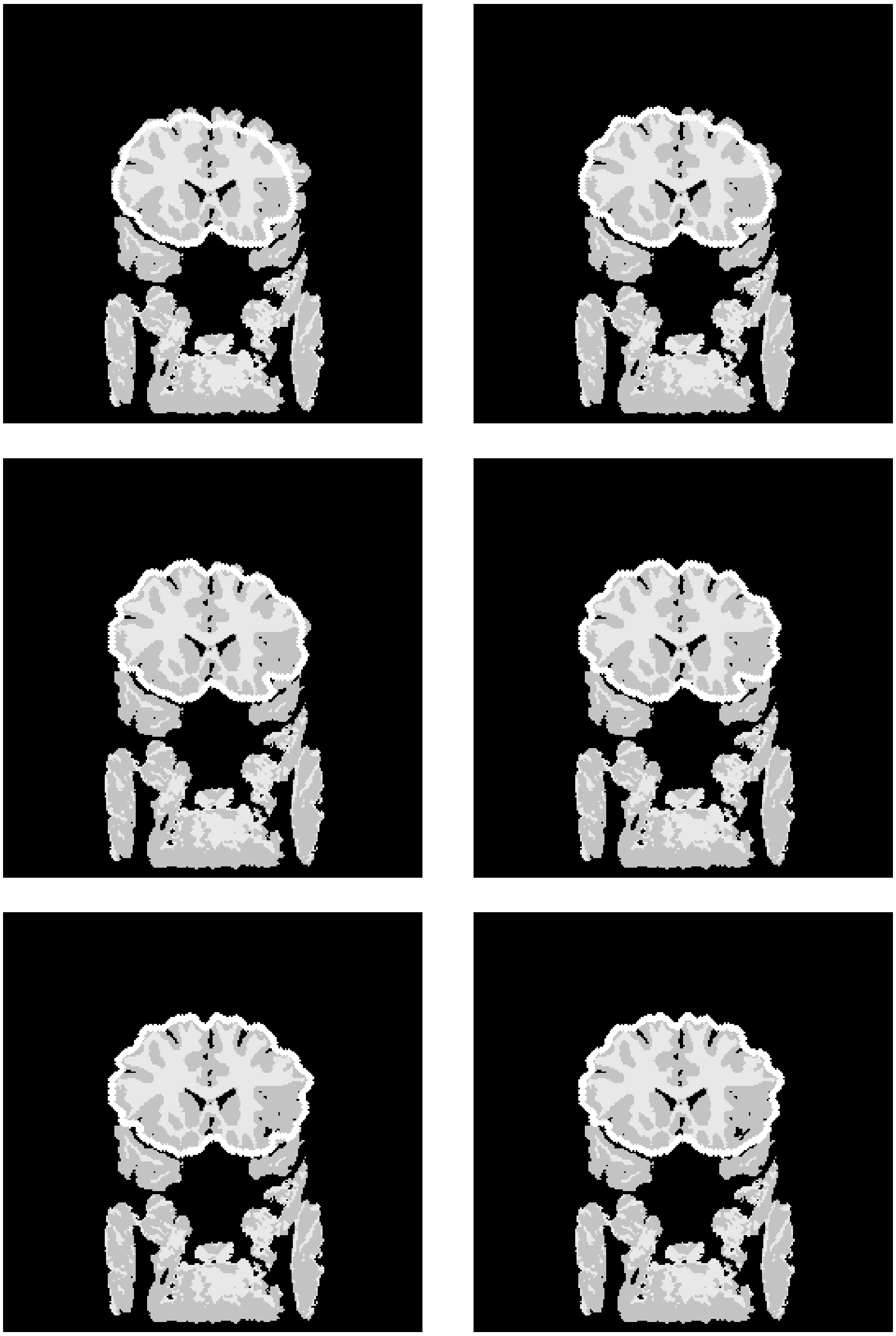


Figure 3-10: Top to Bottom, Left to Right: Further Iterations of the Customized Balloon Model (continued from Figure 3-9) and the Bottom Right Image Shows Convergence of the Snake to a Unacceptable Local Minimum that Excludes the Temporal Lobes.

Algorithm	Time Per Iteration	Number of Iterations in Average Case
EM Segmentation	30s	5-10
Erosion/ Dilation	10s (kernel radius 3)	1
8-connectivity	30s	n/a

Table 3.1: Sample Running Times for Segmentation Steps

a brief description of the architecture of the PVS [17], along with sample running times for each algorithm.

3.4.1 Architecture of Power Visualization System

The PVS is a platform for scientific visualization, imaging, and graphics applications. It is composed of a parallel processor server, a video controller, and a disk array. The server at the Surgical Planning Lab, Brigham and Women's Hospital contains 32 processors, 500 MB of shared memory, HiPPI channels, all interconnected by a 1280 MB/s backplane. Each processor consists of a 40 MHz Intel i860 microprocessor, 16 MB of private memory, and a buffered interface to the high-speed backplane. The server attaches to an IBM RS6000, which serves as the support processor, and provides access to standard networks. The video controller takes images over a 100MB/s HiPPI channel from the server and displays them at a resolution of 1280x1024. The disk array subsystem provides data storage for 20GB, and communicates with the server at sustained data transfer rate of 55MB/s using a HiPPI channel.

3.4.2 Sample Running Times

Table 3.1 shows sample running times for different steps of the segmentation process on a 256x256x124 data set. The intent of this table is simply to give the reader a rough idea of how long each of these steps takes on this parallel machine, and to show that these running times make the process realistic for use in actual surgical planning tasks.

Chapter 4

The Validation Problem And Results

While several systems for segmentation of medical data are currently in use in various research laboratories and hospitals, we have yet to come across one that employs a satisfactory validation method to its results. The import of this statement is not that these systems perform unsatisfactorily, rather it is an illustration of the fact that the medical image processing community is reaching the consensus that validating segmentations of medical data is a hard, and maybe even a poorly defined, problem. The importance of this issue is underlined by recent publications dedicated to it [18] [19].

4.1 The Validation Problem

A validation method can be thought of as a combination of two components. One component is the notion of a "ground truth" or a "golden mean" against which the results of an algorithm are to be judged. The second component is a measure for establishing the deviation of the results from this ground truth. For their second component, most validation schemes use standard statistical methods of finding means, modes, variances, standard deviations, or root mean squared errors. The first component requires developing the notion of a ground truth for a segmentation algorithm and that is where we all tend to run into difficulty. The problem is not that there is no ground truth for medical data, for certainly there is; the problem is that the ground truth is not typically available to the segmentation systems in any form that they can use readily. For instance, in the case of segmentation of brain tissue from magnetic resonance images, there is indeed a true boundary of the brain tissue for each patient, but we don't know what it is. Approximations to the true boundary can be obtained in the form of manual segmentation by experts of neuroanatomy, but experiments have shown that there can be up to 15% variations in classifications generated by different experts [15] [8]. We could test our method on synthetic images, but it is difficult to generate synthetic images that capture the complexity and deformability of the human head. We could test our method on a phantom brain, but we are not aware of the existence of one. We could place fiducials or markers randomly in the patient in known areas of the brain, and use that as a subsampled ground truth to measure the classifications against, but that it far too invasive to be practical. A reasonable alternative would be to use cadaver brains for marker based validation, but we are not aware of any systems that actually do this. One way of working around the ground truth issue would be to note that segmentation is usually a pre-processing step for other applications, and instead to measure the suitability of various segmentation algorithms to particular applications based on the performance of the applications that build upon them. Of course, this "lazy evaluation" would only be useful in applications where the validation problem is easier, for example in the case of registration of data from different modalities where external markers can be used [20] to facilitate validation.

In the rest of this chapter we summarize the methods that have been typically used for validation of segmentation algorithms, their strengths and weaknesses, and then present our results based on the validation scheme that we have chosen.

#	Segmentation Task	Strengths	Weaknesses
1	Brain Tissue[14] & Cardiac Vasculature [21]	Simple	Subjective
2	White and Grey Matter [8], [16], [15]	Less Subjective	15% variation in manual segmentations & Labor Intensive
3	White and Grey Matter [8], Cardiac Walls & Corpus Callosum [22]	Scientific Approach	Not Sufficient in all applications
4	Cardiac Motion [23]	Reliable	Invasive
5	Brain Tissue [24]	Reliable	Labor Intensive

Table 4.1: Strengths and Weaknesses of Validation Methods

4.2 Validation Methods

We have considered five types of schemes that have been proposed by different groups for validation of segmentation algorithms. These are as follows:

- Method 1: Visual Inspection
- Method 2: Comparison with Manual Segmentation
- Method 3: Testing on Synthetic Data
- Method 4: Use of Fiducials on Patients
- Method 5: Use of Fiducials and/or Cadavers

Table 4.1 presents for each of these methods a summary of the particular segmentation applications for which it been used, as well as its strengths and weaknesses.

4.3 Our Validation Scheme and Results

We use a combination of Methods 1 & 2, i.e. visual inspection and comparison with manual segmentation for validating our results. In spite of the weaknesses of these methods, they are acceptable to us because our results are mostly used for visualization purposes in surgical planning. So in some sense, we are measuring the performance of our segmenter by the performance of the application that build upon its results. Only in this case, the application is visualization of the segmented data, and the performance of that can only be measured by the clinicians and surgeons who use the segmented data for surgical planning or other purposes. We show 3D renderings of our results that have been used in this application in Section 4.3.3.

Recently our segmenter has gained usage in two volumetric studies underway at Harvard Medical School and Brigham and Women's Hospital: one is a Schizophrenia study that examines the white matter and grey matter content in the cerebellum, [2], and the other is a study of the change in brain composition in newborns [25]. For the Schizophrenia study, we had available to us manual segmentations for 20 scans, against which we compare our results and present some quantifications of our performance in Section 4.3.1. For the newborn study, our results were validated by experts visually and considered acceptable, and illustrations of the segmentations are shown in Section 4.3.4.

4.3.1 Quantitative Results from Schizophrenia Study

In the last three months, we have used our system to segment cases for various purposes. Of these, we have manual segmentations for 20 cases that are being used in a Schizophrenia study [2] (the rest were only performed using our system sisince manual segmentation for each case can take upto 6 expert-hours) against which we compared our results. Table 4.2 shows a simple comparison of the number and percentage of pixels that are classified differently by our segmenter as compared to the manual segmentation. Tables 4.3 and 4.4 show a comparison of the brain tissue edges generated by our segmentations with edges generated by the manual segmentations. The comparison is performed in two steps: first by finding the mean distance (city-block) from each edge in our segmentation to edges in the manual segmentation (which basically penalizes false positives) as shown in Table 4.3, and then by finding the mean distance from an edge in the manual segmentation to edges in our segmentation (which penalizes false negatives) as shown in Table 4.4. We also include some other statistics on the edges, such as the percent of edges in our segmentation that coincide with (or are one and two pixels away from) edges in the manual segmentation. We repeat this test to compute the number of edges in the manual segmentation that coincide with, or are within one or two pixels of the edges of our segmentation. The idea behind these tests is to convey the "goodness of fit" between our segmentation and ones performed manually. As can be seen by the numbers in these tables, we are usually within 2-3% of manual segmentations, which is considered acceptable performance in the applications we are dealing with [15].

We illustrate these results in Figures 4-1– 4-5 using five randomly chosen slices from different data sets. In each of these figures, the top left image shows the edges of brain in our segmentation overlaid on the input image (in white), and the top right image shows the edges of the brain in the manual segmentation overlaid on the input image (also in white). The bottom left image shows our segmentation overlaid with the edges of the brain from the manual segmentation, and the bottom right shows the manual segmentation overlaid with the edges of the brain in our segmentation. The purpose of the top two images in each of the figures is to give an idea of how well the boundaries of our segmentation and the manual segmentation agree with the original grayscale images, and the purpose of the bottom two images in each figure is to illustrate how well these two sets of edges agree with each other.

4.3.2 Qualitative Results

In this section, we present some visual illustrations of the performance of our segmenter in two other applications have been involved in, but don't have available to us manual segmentations to compare our results against. In Section 4.3.3 we present some of the segmentations that we have generated for visualization in surgical planning, and in Section 4.3.4 we present some segmentations generated for the volumetric study of brains of preterm infants.

4.3.3 Surgical Planning

We are now generating routine segmentations for surgical planning at the Surgical Planning Lab of Brigham and Women's Hospital. It takes us about 15-20 minutes to

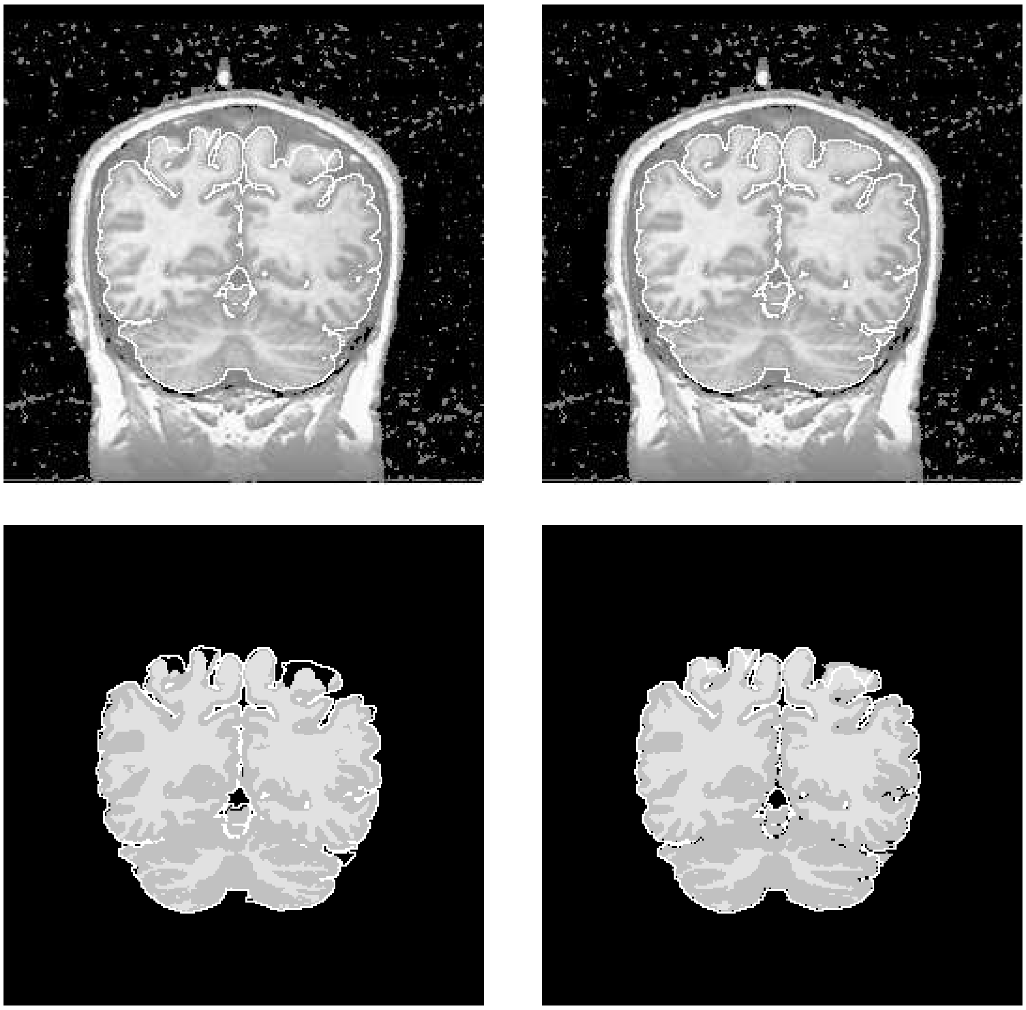


Figure 4-1: Top Left: Input Image Overlaid With Brain Edges from our Segmentation, Top Right: Input Image Overlaid With Brain Edges from Manual Segmentation, Bottom Left: our Segmentation Overlaid with Brain Edges from Manual Segmentation, Bottom Right: Manual Segmentation Overlaid with Brain Edges from our Segmentation

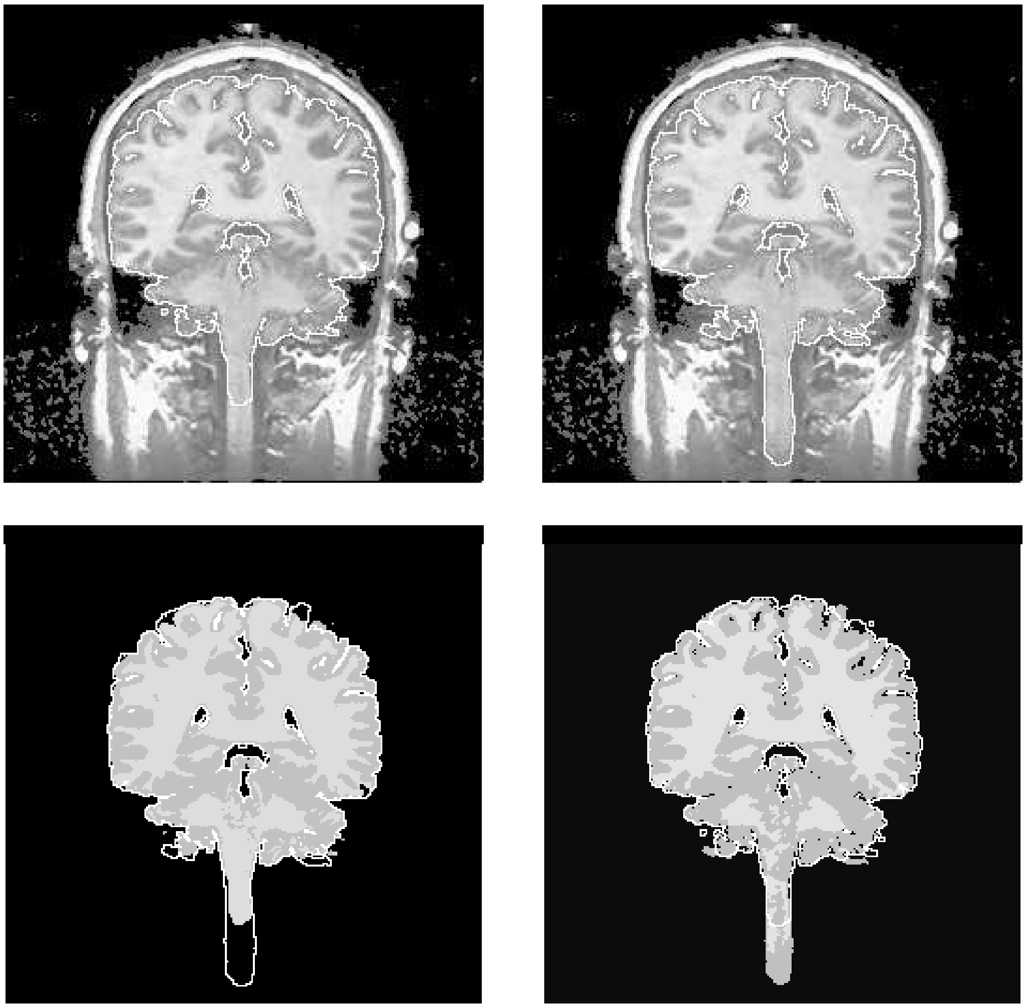


Figure 4-2: Top Left: Input Image Overlaid With Brain Edges from our Segmentation, Top Right: Input Image Overlaid With Brain Edges from Manual Segmentation, Bottom Left: our Segmentation Overlaid with Brain Edges from Manual Segmentation, Bottom Right: Manual Segmentation Overlaid with Brain Edges from our Segmentation

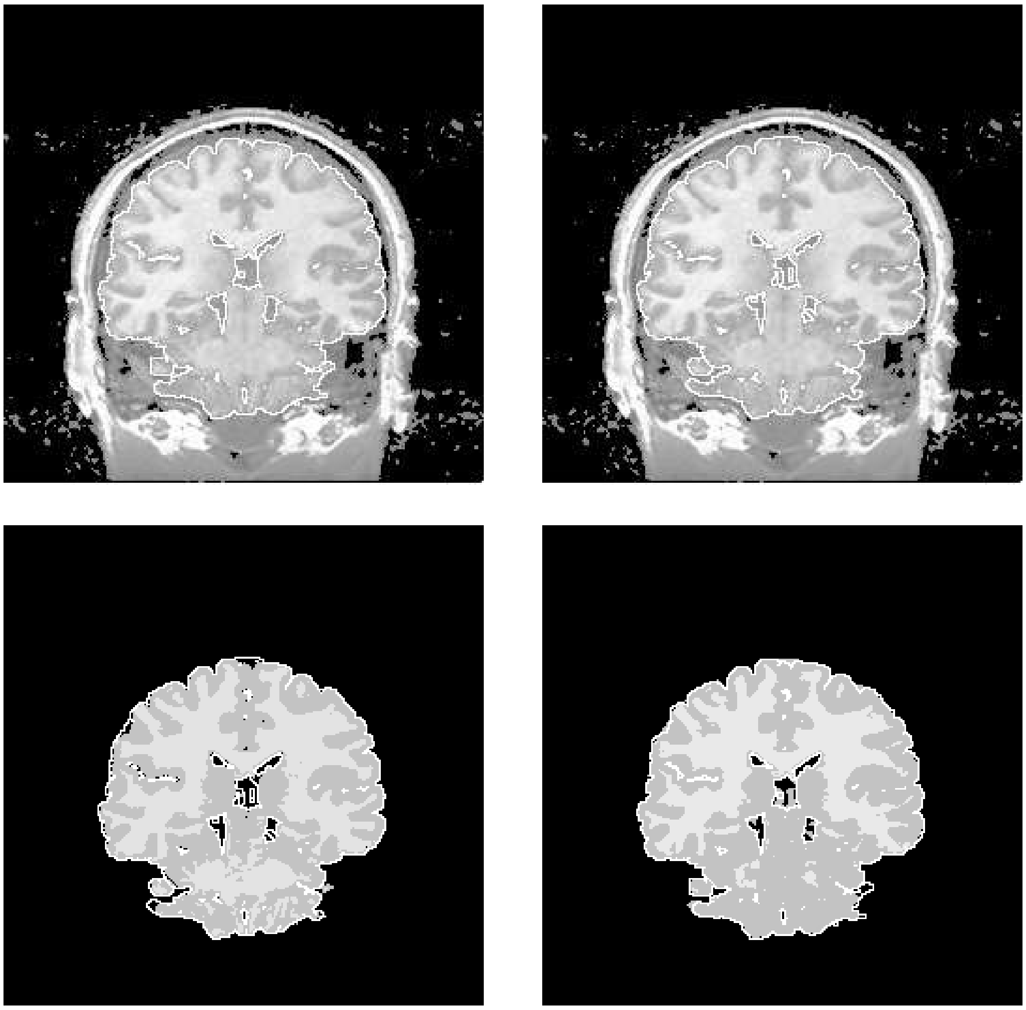


Figure 4-3: Top Left: Input Image Overlaid With Brain Edges from our Segmentation, Top Right: Input Image Overlaid With Brain Edges from Manual Segmentation, Bottom Left: our Segmentation Overlaid with Brain Edges from Manual Segmentation, Bottom Right: Manual Segmentation Overlaid with Brain Edges from our Segmentation

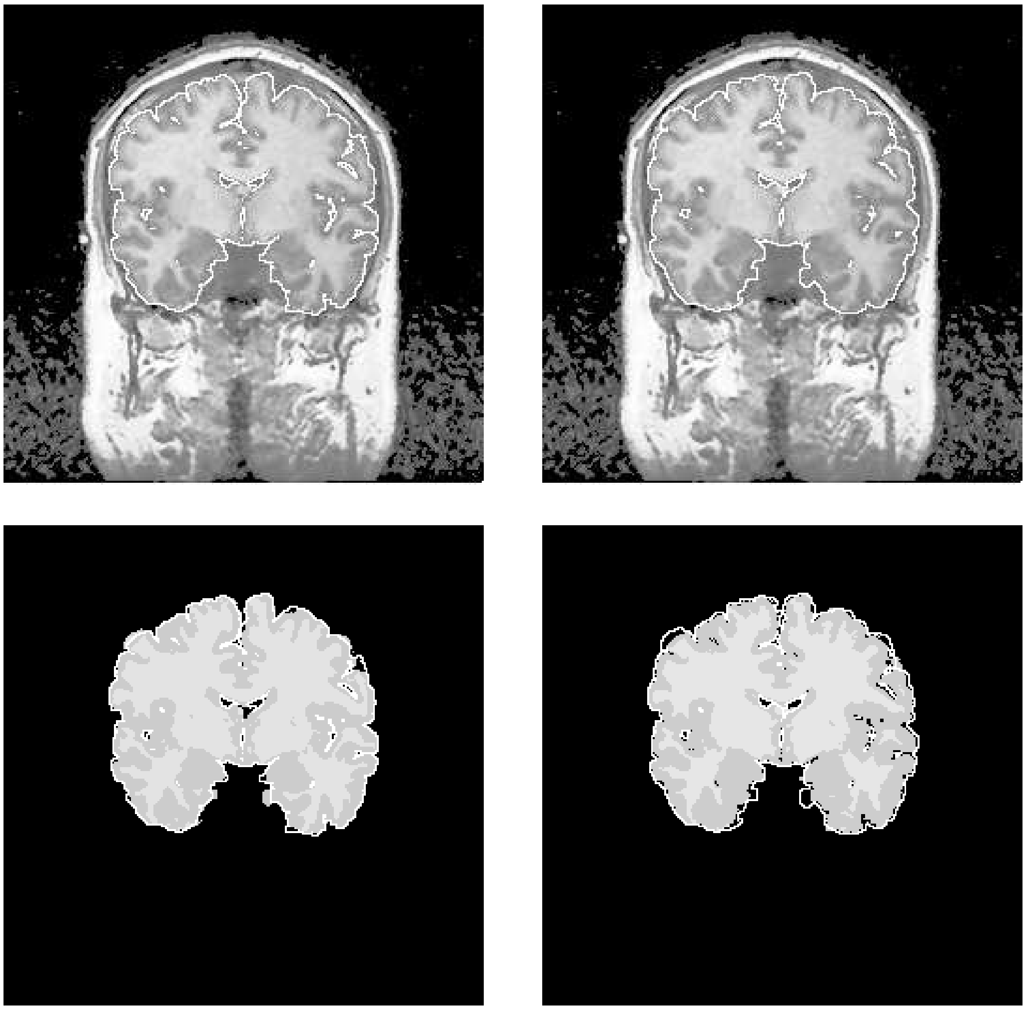


Figure 4-4: Top Left: Input Image Overlaid With Brain Edges from our Segmentation, Top Right: Input Image Overlaid With Brain Edges from Manual Segmentation, Bottom Left: our Segmentation Overlaid with Brain Edges from Manual Segmentation, Bottom Right: Manual Segmentation Overlaid with Brain Edges from our Segmentation

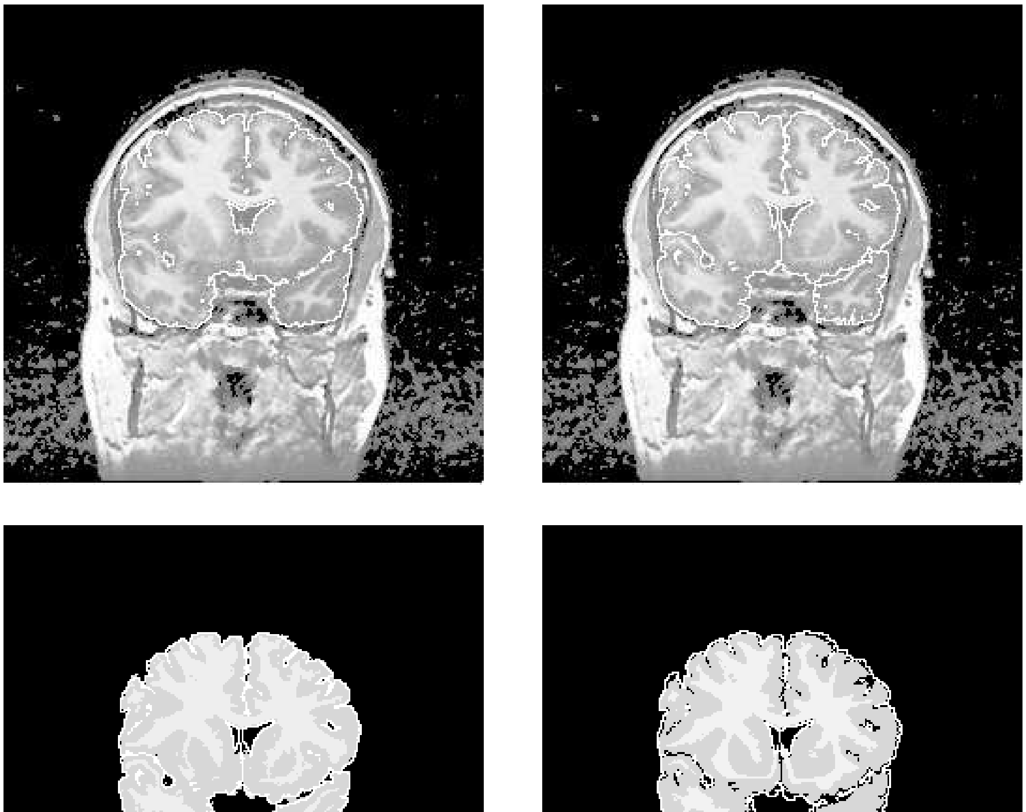


Figure 4-5: Top Left: Input Image Overlaid With Brain Edges from our Segmentation, Top Right: Input Image Overlaid With Brain Edges from Manual Segmentation, Bottom Left: our Segmentation Overlaid with Brain Edges from Manual Segmentation, Bottom Right: Manual Segmentation Overlaid with Brain Edges from our Segmentation

Case No.	Number of Different Voxels	Percentage Difference
1	120463	1.72
2	129648	1.85
3	153631	2.19
4	124335	1.77
5	150297	2.14
6	133743	1.91
7	192502	2.75
8	221291	3.16
9	188528	2.69
10	171776	2.45
11	176724	2.52
12	144990	2.07
13	136753	1.95
14	184274	2.63
15	142237	2.03
16	258752	3.69
17	194908	2.78
18	197774	2.82
19	101356	1.45

Table 4.2: Difference in Classification between Our Method and Manual Segmentation

process a single case, which makes the method attractive to clinicians who could otherwise spend 4-6 hours creating the segmentations with less automatic tools. Typically clinicians overlay these segmentations with manually segmented pathological areas of the brain to create the final visualizations. Figures 4-6 and 4-7 are two views each of the brain surface and the white matter surface for two different healthy patients. Figures 4-8 show two views of a patient with subdural haematoma as well as a tumor.

4.3.4 NewBorn Study

We have segmented brains for a quantitative study of brain development in preterm infants by Huppi et al. [25]. In order to examine brain composition, MR scans of 14 preterm infants of different gestational age were used, and our segmenter was used to define the total brain volume, as well as percentages of csf, unmyelinated and myelinated white matter, and total grey matter. In this case, we defined two separate classes for the myelinated and unmyelinated white matter, as compared to

Case No.	Mean d	% at d=0	% at d=1	% at d=2	% at d>2
1	1.10	19.84	69.05	7.65	3.47
2	1.23	8.19	73.49	12.67	5.65
3	0.84	31.49	61.27	5.10	2.13
4	1.07	20.91	66.45	7.29	5.35
5	1.11	11.00	75.19	9.27	4.54
6	0.85	28.20	65.20	4.82	1.78
7	0.76	38.58	55.20	4.42	1.80
8	0.81	34.74	57.45	5.04	2.78
9	0.52	57.18	39.12	2.29	1.41
10	1.21	22.34	63.32	6.78	7.56
11	0.77	42.97	47.00	5.14	4.89
12	2.22	15.27	61.86	8.13	14.74
13	1.67	17.96	58.14	8.43	15.47
14	1.36	13.12	68.58	8.84	9.47
15	1.59	13.63	63.66	10.33	12.38
16	6.44	37.79	31.49	4.58	26.14
17	1.29	42.25	40.81	5.32	11.63
18	1.26	41.16	40.10	5.94	12.80
19	1.15	16.53	71.84	8.29	3.35

Table 4.3: Measuring False Positives (d is the distance in pixels between an edge in the manual segmentation and the closest edge in our segmentation).

Case No.	Mean d	% at d=0	% at d=1	% at d=2	% at d>2
1	1.20	10.37	72.57	12.39	4.67
2	1.34	21.41	69.02	4.46	5.11
3	1.51	7.44	66.06	15.97	10.54
4	1.69	10.97	72.20	11.81	5.03
5	1.08	12.49	78.51	6.56	2.43
6	1.46	6.46	69.14	16.03	8.37
7	1.44	6.67	65.49	18.50	9.34
8	1.32	8.34	67.57	19.73	4.36
9	2.09	5.33	48.32	28.47	17.88
10	1.10	14.91	69.71	11.58	3.80
11	1.34	8.76	63.61	22.36	5.27
12	1.25	14.82	71.71	9.71	3.76
13	1.17	14.21	70.97	10.74	4.07
14	1.17	13.08	75.40	8.00	3.52
15	1.00	17.91	72.09	7.48	2.53
16	1.48	9.00	55.81	26.81	8.39
17	2.26	8.87	56.99	24.13	10.02
18	1.43	11.28	56.98	25.07	6.68
19	1.42	15.12	68.81	10.04	6.03

Table 4.4: Measuring False Negatives (d is the distance in pixels between an edge in our segmentation and the closest edge in the manual segmentation).

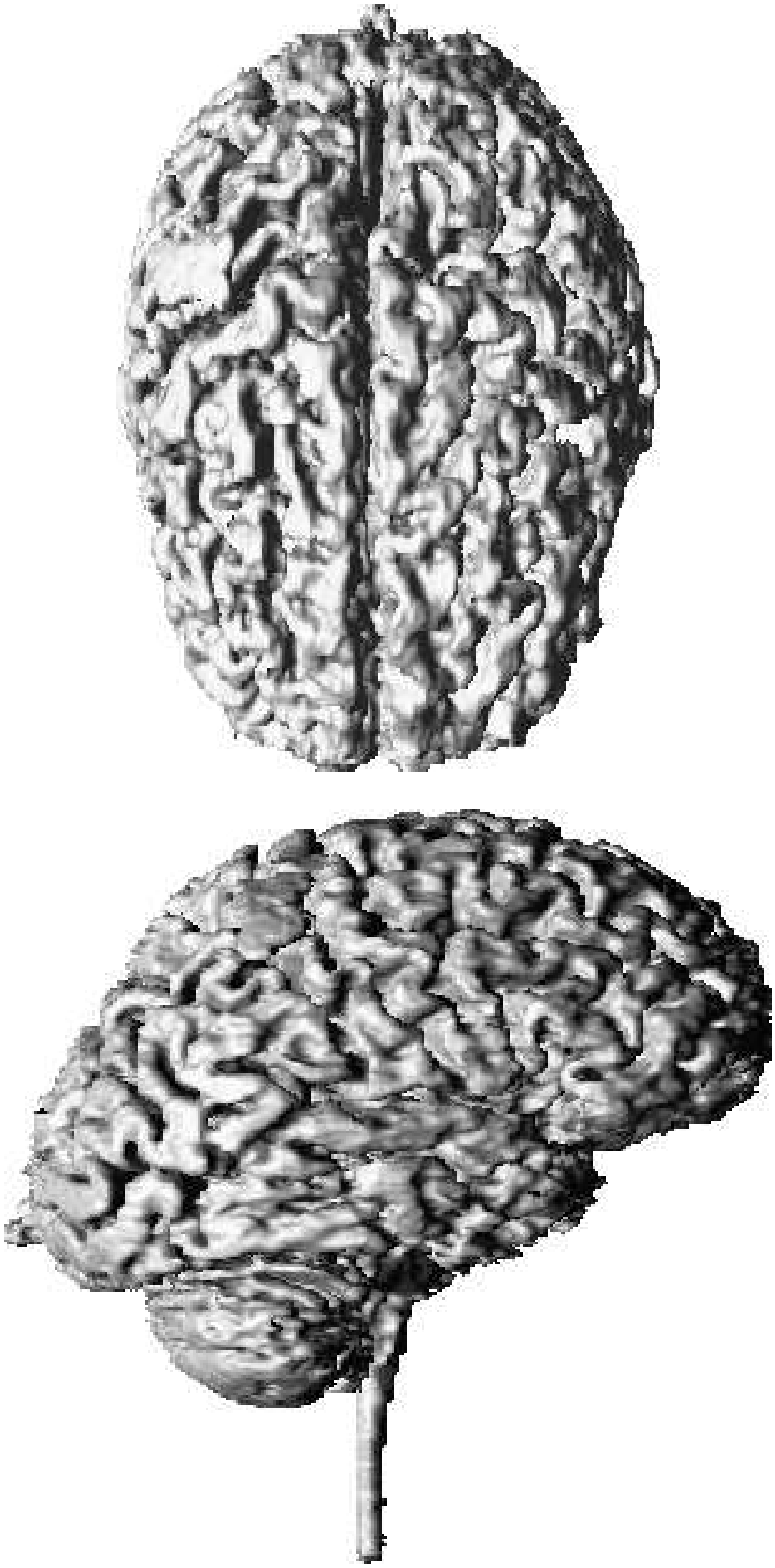


Figure 4-6: Top: Top View of a Reconstructed Brain Surface, Bottom: Left View of the same Reconstructed Brain Surface

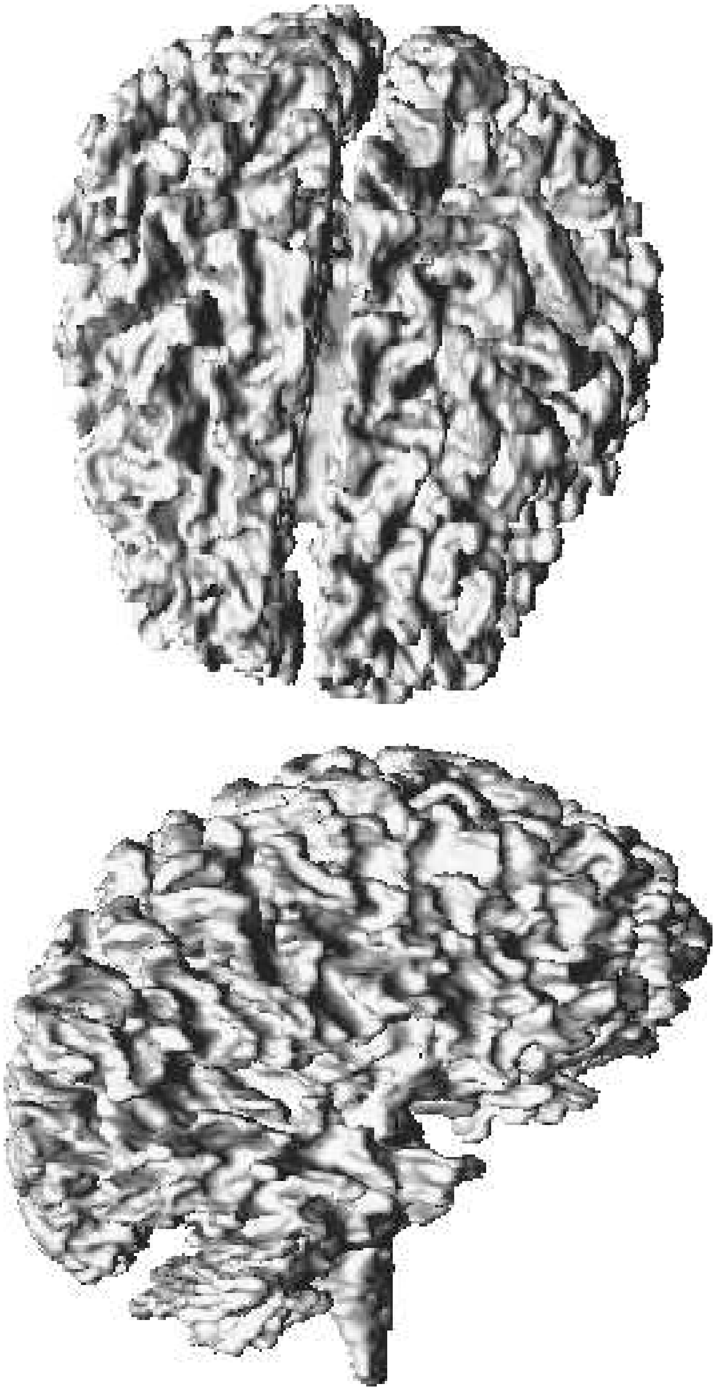


Figure 4-7: Top: Top View of a Reconstructed White Matter Surface, Bottom: Left View of the same Reconstructed White Matter Surface

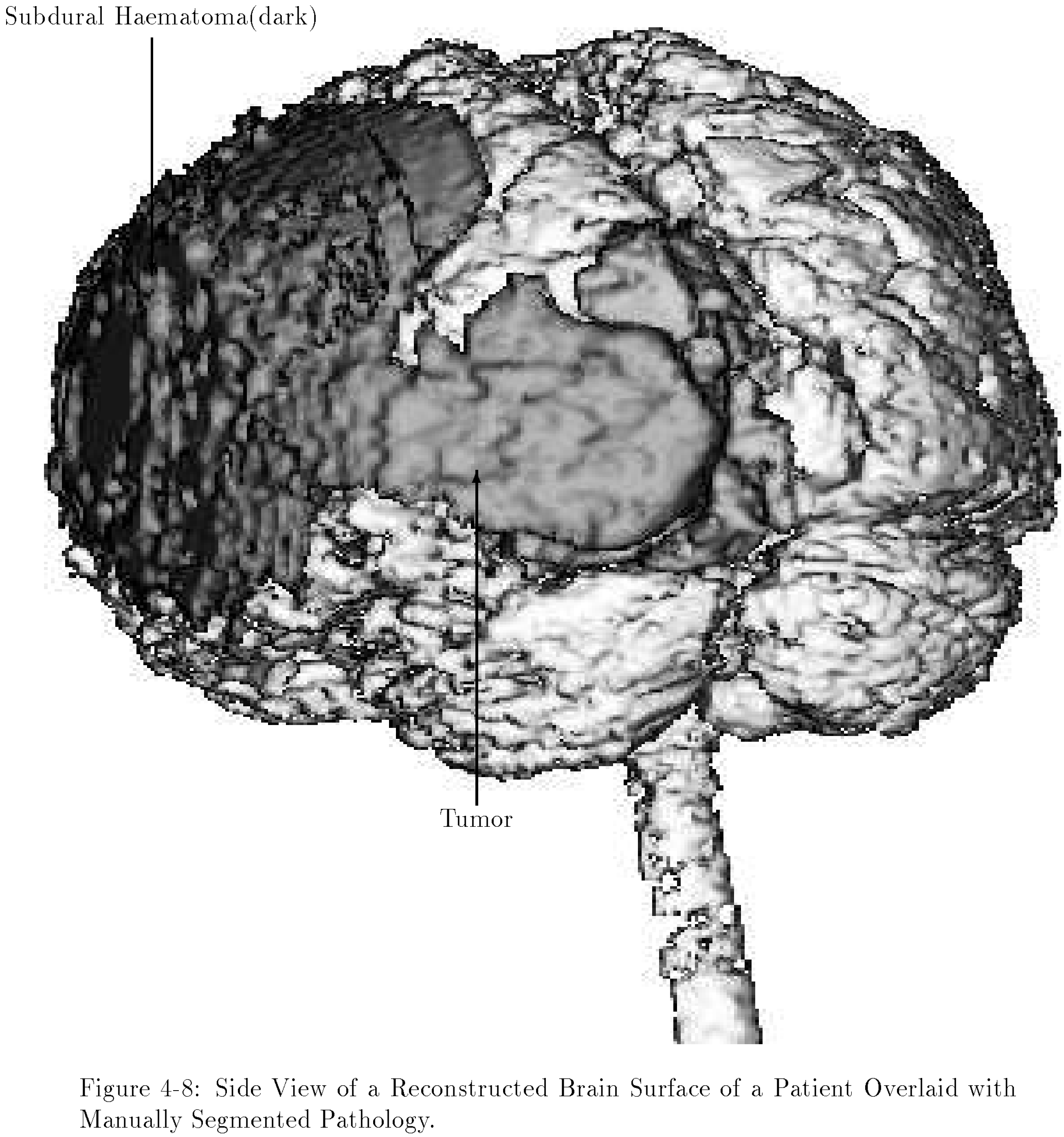


Figure 4-8: Side View of a Reconstructed Brain Surface of a Patient Overlaid with Manually Segmented Pathology.

the case of adult brains where we only use one tissue class for white matter. Results of the study indicate that in healthy preterm infants, myelinated white matter volume increases with gestational age, unmyelinated white matter decreases, and grey matter increases. Some results of the segmentation are show in Figures 4-9 – 4-11. The top two images in each of these figures are the two echoes of the data that were used to generate the tissue segmentations, which are shown at the bottom of each figure.

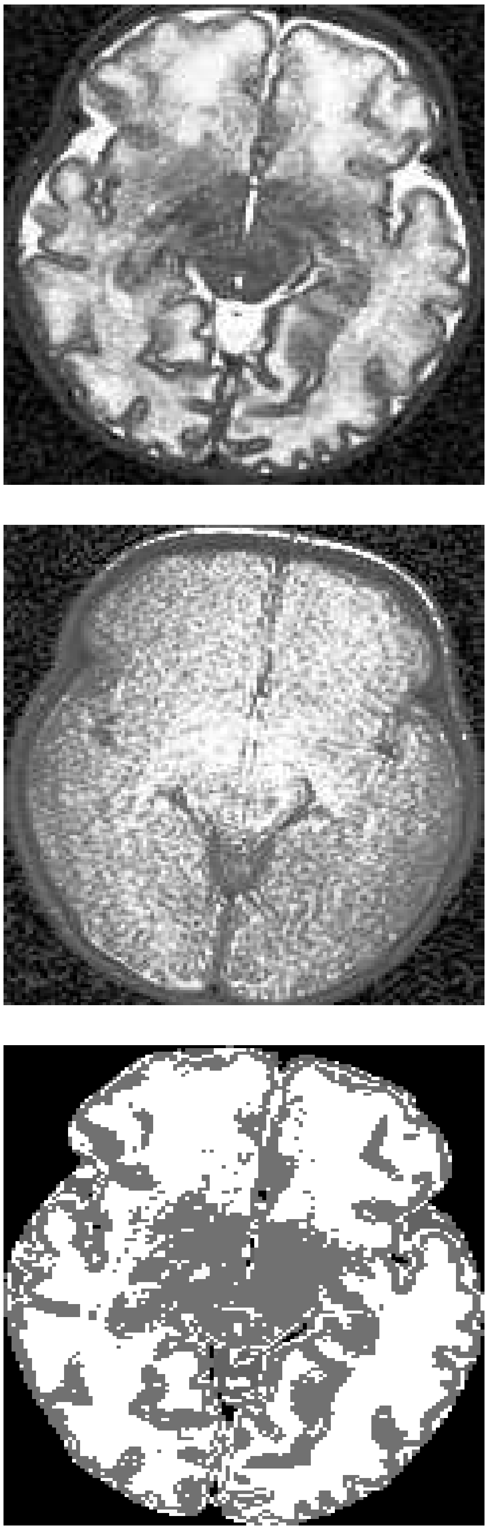


Figure 4-9: Top: First Input Echo, Middle: Second Input Echo, Bottom: Segmentation into Gray (dark) and White (bright) Matter for a Neo-Natal Brain

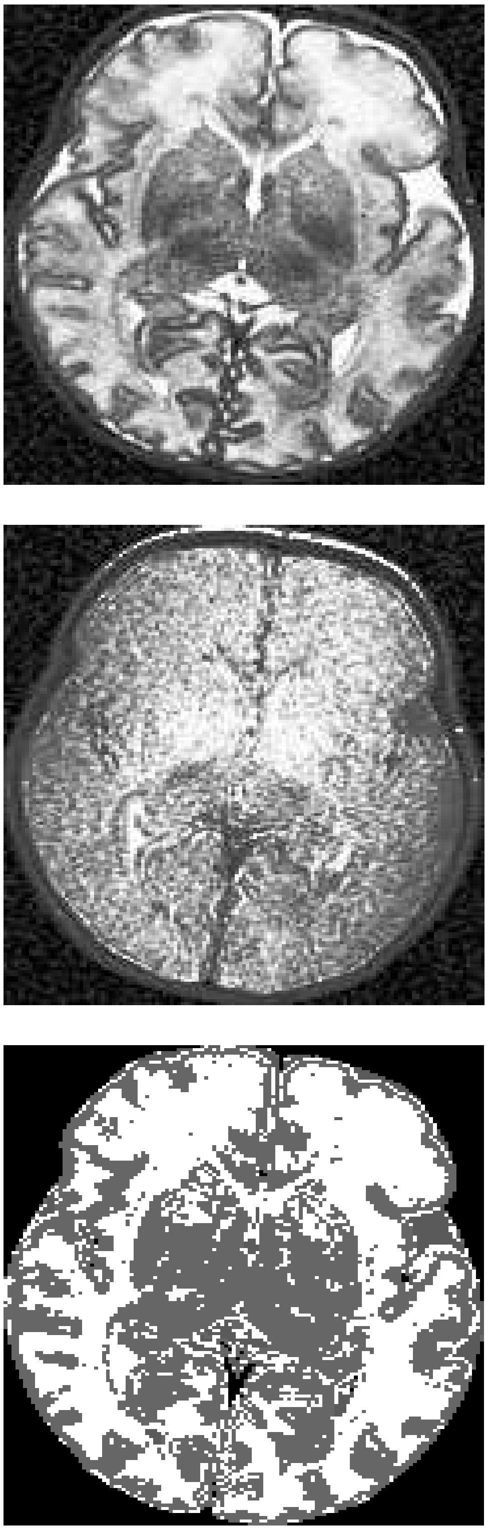


Figure 4-10: Top: First Input Echo, Middle: Second Input Echo, Bottom: Segmentation into Gray (dark) and White (bright) Matter for a Neo-Natal Brain

Figure 4-11: Top: First Input Echo, Middle: Second Input Echo, Bottom: Segmentation into Gray (dark) and White (bright) Matter for a Neo-Natal Brain

Chapter 5

Discussion and Conclusion

This work began as an exploration of the suitability of snakes for the segmentation of brain tissue, and gradually, mostly due to the application oriented nature of our collaboration with a hospital, turned into a search for a practical solution to the segmentation task. We started with the classical snakes, and coupled those to various types of external forces that seemed, at the time at least, to capture the salient features of our problem. Our first attempt was to use just a gradient magnitude based external force, which wasn't very successful because the boundary of the brain tissue is usually surrounded by points of high gradient magnitude due to other structures, and "correct" initialization became arduous.

The second attempt was to build a smarter objective function and to obviate the initialization step. To this end, we extracted the surface of the skin by thresholding and taking the extremities of each scan line, and used that as the initial position for the snake in each cross section of the head. We created two separate objective functions, and iterated the snakes to minimize each of those in order. The first objective function was intended to take the snakes from the surface of the skin to the surface of the skull, and the second one was supposed to start from the surface of the skull, and move in to the surface of the brain. Since skin is usually a lot brighter than skull, and surrounds the skull, the first objective function essentially preferred low brightness values. The supposition was that in areas of high brightness such as skin and subcutaneous fat, the smoothness constraints will dominate and cause the snake to shrink until it gets to the areas of low intensity which is when the preference for low intensities would cause the snake to reach equilibrium. Then we used the second objective function which employed a gradient based force to allow the snake to lock onto the surface of the brain tissue. This solution was also quickly rendered unusable due to the presence of spurious local minima between the skin surface and the brain surface due to structures such as the dura mater and soft tissue around the brain stem.

The third attempt was to incorporate balloon forces in the snake model, and initialize the snake inside (but not necessarily close to) the brain surface in order to avoid the local minima that were posed by the structures between the skin and the brain in the previous attempt. The problem here was that we could not use results of iterations on one slice as the starting location for an adjacent slice because associating a direction with the balloon force is tantamount to assuming that brain boundaries either only expand or only contract as we move along the volume in a give direction. The fourth attempt incorporated region based forces into the balloon model by using a topological mask for the brain which defined a direction for the balloon force at each point in the contour, so that we did not have to decide a priori whether the balloon was going to expand or contract from its initial position. This made it possible to "track" the brain boundary between adjacent slices because we could use the result of iterating the snake on one slice as the starting position for the snake on adjacent slices, without worrying about the relative size of the brain in the two slices. This is basically what we use as the final step of our segmentation algorithm in its current state. This step still requires manual initialization of the brain boundary in a few slices, and is an undesirable way of adding spatial information to the segmenter for that very reason. An alternative method would be a spatial prior that we have been working to extract from a reference data set based on ratio of the distances from the centroid of the head to the skin surface and to the brain for each possible pair of angles in a standard spherical coordinate system, but the results are preliminary at this point.

Our method for computing a topological mask for the brain uses grayscale morphology and could also be improved. Currently, we do not account for brain size in the kernels used for morphology, which, in hindsight, is an obvious limitation.

5.1 Generalization to Other Anatomical Structures

Though in this paper we have developed the details for the segmentation of a particular anatomical structure, the brain, there are other structures that share the characteristics of the brain that have been exploited in this process, and hence, this method is extensible to those structures also. The kidney and the heart are two examples of such structures. In order to extend our process to other structures, one could apply the following methodology:

- Identify the various relevant intensity distributions in the data using Adaptive Segmentation
- Incorporate topological information using customized morphological operators or other techniques
- Incorporate spatial information using deformable contours, or other customized spatial priors.

The key word in the above method is "customize", since most of the effort in applying this segmenter to any other application will be hidden in doing exactly this. In some sense, this customization step corresponds to the construction of a model for each anatomical structure, which can never really be dispensed with. However, explicit shape modeling promises a more elegant approach to creating general purpose segmentation methods, and is an attractive direction to pursue.

5.2 Conclusion

We have presented a method for segmentation of brain tissue in magnetic resonance images. This method combines three existing techniques into an effective tool for segmentation. Adaptive segmentation is used for intensity correction in the data, morphology and connectivity are used to incorporate topological information in the algorithm, and active contours are used to supply spatial information.

Our implementation is on IBM Power Visualization System, a parallel supercomputer, and has been in use at the Surgical Planning Lab at Brigham and Women's hospital for extracting brain tissue for visualization for surgical planning and other purposes, as well as for volumetric studies for research purposes.

Chapter 6

Related Work

Anatomical structures are frequently non-rigid and exhibit substantial morphological variation from subject to subject. Hence the task of segmenting these from medical imagery is one of identification of a region in an image with only approximate knowledge of its shape, size, gray level appearance, and spatial location. Different segmentation applications have available to them some knowledge in each of these categories, and the challenge is to combine it all in a manner such that the lack of information in one category is offset by the information in the others. For instance, Gerig et al. combine spatial and gray level information to track lesion growth in white matter by characterizing the lesions as concentric blobs of increasing or decreasing brightness in the white matter [16]. Note that in this particular task there is little prior information available about the size or shape of the lesions. In a different task of quantifying the motion of the heart, Duncan et al. use shape constraints to segment the beating heart in a time sequence of images using Fourier snakes [23].

In this chapter, we overview several existing segmentation systems and classify them into three main categories based on the techniques they use to exploit a priori knowledge of anatomy (see Figure 6-1). The three categories are: deformable models based methods, statistical methods, and morphology based methods. Within each of these categories, we will specify which systems encode knowledge of anatomy in explicit atlases, and which ones use anatomical models that are implicit in the heuristics employed by them. Several of the systems we discuss here are in fact eligible for membership in more than one categories, but the classification is based on the primary features of the system as perceived by us. Figure 6-1 shows a diagramatic representation of our classification.

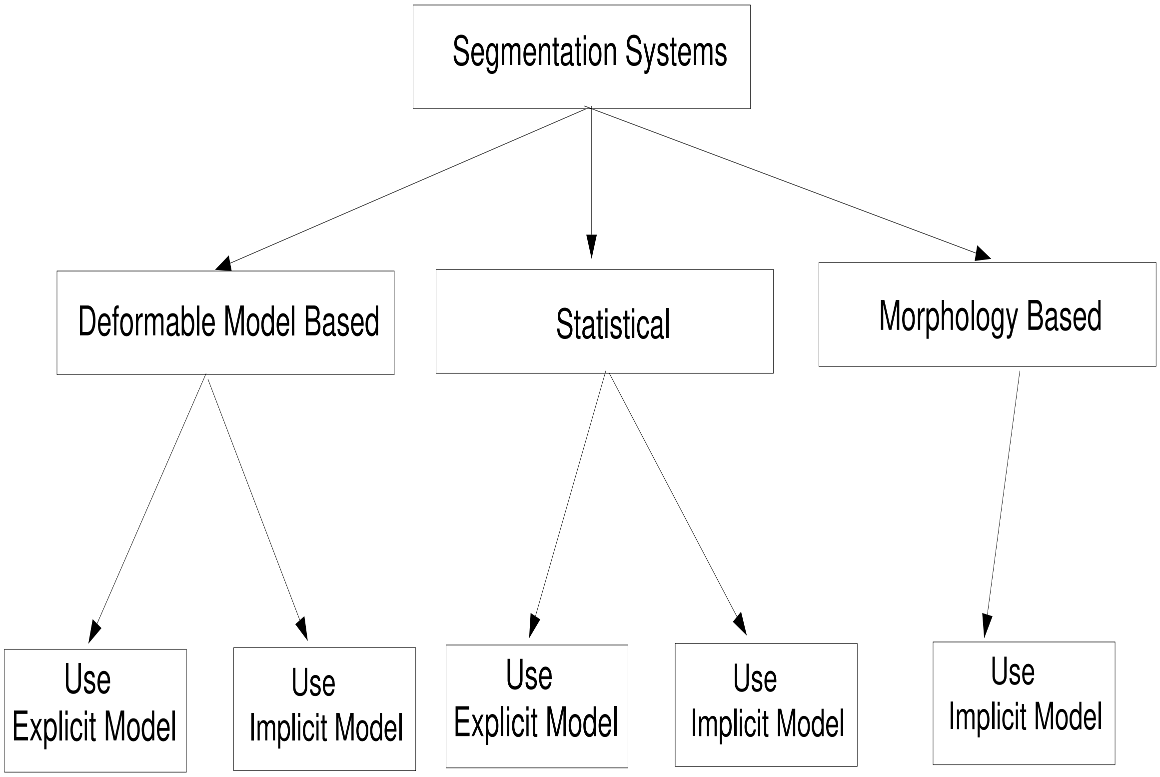


Figure 6-1: Classification of Segmentation Systems

6.1 Deformable Model Based Methods

Deformable models are a fairly popular component of medical image segmentation systems due to their ability to encode approximate shape constraints. In such systems, anatomical structures are typically modeled using stacks of deformable contours in 2D or using actual 3D deformable surfaces. In this section we describe some systems that are employ snakes, balloons and other deformable models for segmentation purposes.

6.1.1 Example Application Systems

Snakes and balloons have been typically used with creative external forces to segment various anatomical structures. Using snakes to track 2D contours in 3D, coupling of the snakes to region information in the image, using simulated annealing to find the global minimum of the snake energy, and imposing shape priors are notable extensions to the original snakes.

Hyche et al. at Georgia Institute of Technology use classical snakes to detect motion of coronary arteries in cardiac angiography (X-ray) motion sequences from a pig [21]. A radio-opaque contrast agent is injected to improve identification of the arteries. They accurately mark the position of the artery in one frame of the cardiac cycle and then use the snake model to track the position of the artery in subsequent frames. They report a need for data acquisition at 30fps or more for the algorithm to perform satisfactorily.

Poon et al. at the University of Technology, Sydney, Australia detect cervical carcinoma and ventricles from MR images using a snake approach that incorporates region information as well [26]. The number of regions in the image and a topologically correct initialization of contours is provided interactively. A region based image energy term is created that minimizes gray level variance within regions and maximizes it between regions. An external constraint energy term is added to penalize the difference between the number of regions in the image, and the number of regions created by the contours. The snake energy is minimized using simulated annealing to avoid being trapped in local minima. A cooling parameter, T, analogous to temperature is defined. Starting at a high temperature, each contour is modified by repositioning one of its points. The new location is generated randomly, but the maximum displacement is adapted to the cooling schedule to allow large scale modifications at high temperatures and small scale modifications at low temperatures. As is characteristic of most simulated annealing implementations, this one is computationally intensive. Syn et al. [27] use a more efficient simulated annealing by reformulating the snake forces as Gibbs functions. This defines a Markov dependency over adjacent control point triplets, allowing the use of simulated annealing to optimize over a search range defined on both sides of the initial contour. Grzeszczuk and Levin present a stochastic deformation method for segmentation [28] in which image forces are derived from the statistical properties of previously segmented images, and simulated annealing is used to find the globally optimal contour of interest.

Staib and Duncan use a Fourier parameterization for the active contour model (Fourier snakes)[29] which expresses a curve in terms of an orthonormal basis represented as follows:

$$\begin{aligned} v(t) &= \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} \\ &= \begin{bmatrix} a_0 \\ b_0 \end{bmatrix} + \sum_{k=1}^{\infty} \begin{bmatrix} a_k b_k \\ c_k d_k \end{bmatrix} \begin{bmatrix} cos(kt) \\ sin(kt) \end{bmatrix} \end{aligned}$$

where $v(t)$ is the contour vector consisting of the x and y coordinates and a_k, b_k, c_k,and d_k are the corresponding fourier coefficients, The curve is thus represented by

$$p = (a_0, b_0, a_1, b_1, c_1, d_1, \cdots).$$

To segment a region from an image, an explicit fourier description of a boundary contour is created using sample segmentations, and is used to first search for a rough match in the image which is then refined using gradient based image forces. This method has been used in the segmentation of the corpus callosum boundary from MR data and to detect the epicardial and endocardial borders for the left ventricle from cardiac MRI sequences. The data is processed slice by slice in 2D, and in the latter application the contours are from each frame are stacked and run through a minimum-span surface tessellation algorithm to form the endocardial and epicardial surfaces. Also noteworthy is the addition of region based information into this framework by Chakraborty et al. [22].

Gindi et al. use "atlas based snakes" to segment the putamen, a low contrast mid-brain structure from T1 weighted MR images [30]. The explicit "atlas" is a normal data set in which locations of some landmarks as well as the boundary of the putamen have been marked manually. For a given data set, the segmentation consists of two steps. The first step is a coarse alignment of the data with the atlas using a 3D affine transform that brings manually identified landmarks for the data into correspondence with those for the atlas. The data set is then resliced to obtain cross-sections that correspond to the slices in the atlas data. The transform is applied to the atlas contours for the putamen to obtain an initial estimate of the putamen in each slice of the transformed data. The second step is to refine this estimate using the snake model described in 2.3 with the following additions. Each step of the snake evolution has two phases. In the first phase, the snake evolves to minimize external forces, internal forces, and squared distance from the atlas snake for this data set, which is initially the reference contour supplied by the atlas. And in the second phase, the atlas snake for this particular data is computed from the reference atlas snake and the position of the snake in the first phase by using an affine transform that minimizes the squared distance between the two.

Goble et al. create an explicit deformable model for the brain which is a composite of four different 3D snakes or "active surfaces" – one for each of the temporal lobes, one for the brain stem, and a combined one for the frontal, parietal and occipital lobes [24]. They interactively initialize this model in the data and iterate it in conjunction with morphology based image forces to segment the brain from MRPAGE data.

In our experience, snake-like models serve best in an interactive setting because of the need for manual initialization of the initial positions and the adjusting of parameters for acceptable performance. Fourier snakes obviate the need for manual initialization by constructing shape models from sample data and are attractive for smooth surfaces like that of the corpus callosum, but would require a prohibitive amount of computation if applied to surfaces as complex as that of the brain. Atlas based snakes also remain to be demonstrated on a surface other than that of the smooth putamen. It is also worthwhile to note that the efficacy of snakes as an interactive tool depends greatly on the presence of a powerful graphical user interface that allows the user to easily, naturally, and efficiently control the different forces acting on different parts of the snake.

6.2 Morphology Based Methods

Image morphology provides a way to incorporate neighborhood and distance information into algorithms (see [10] [11] for detailed treatment of morphological operators). Binary morphology has been used in several segmentation systems, some of which are discussed below.

6.2.1 Example Application Systems

Brummer et al. use a series of 2D morphological and connectivity operations to extract the brain tissue from MR images [14]. They start with a Rayleigh distribution based method for generating thresholds for the brain. Then they employ the following four steps used to extract a first approximation to a brain mask: erosion with circular structuring element, 8-connectivity on eroded image, dilation of the labels generated in previous step with structuring element slightly larger than used in the erosion, masking label dilated image with result of the first step. An image overlap test to determine which dilated connected components should be used in each slice is done by selecting the slice which contains the largest cross section of the brain, and then selecting in neighbouring slices only the connected components that satisfy some measure of overlap with the largest cross-section. The resultant approximation to the brain is refined using a distance from the skin surface metric.

Though several heuristics have been incorporated to produce reasonable results, unnecessary complexity has been introduced in the algorithm by ignoring the 3D nature of the problem, and trying to solve it in 2D.

Work by Gerig et al. [16] does use a similar semi-automatic approach in 3D to segment the brain from T1 and T2 weighted MR images. Sandor and Leahy use 3D morphology and connectivity to mark the various sulci in the cortex in isolated brains from MR images [3]. In a similar vein, O'brien et al. at Georgia Tech detect coronary vessels in Angiographic image sequences (X-ray) using operations such as region growing and skeletonization [31].

Though morphology provides a simple and efficient way for incorporating distance and neighborhood information in segmentation, it carries the potential risk of distorting structure boundaries if used inappropriately. If the structure being eroded is convex, then dilation is indeed dual to erosion in the sense that a dilation following an erosion will restore the object boundaries eroded by the erosion. However, in case of non-convex objects like the brain tissue, if the kernel size is comparable to the folds of the brain, then dilation can not restore the effect of the erosion, and therefore the resultant object boundaries will be distorted.

6.3 Statistical Methods

Maximum likelihood, Bayesian decision theory, and principal component analysis have been used for intensity based as well as location based segmentation of medical imagery as illustrated below.

6.3.1 Example Application Systems

Wells et al. [8] present a powerful method for simultaneous classification of tissue classes and estimation of the gain field due to inhomogeneities in the RF coil in MRI. Various tissue classes are represented as a Parzen distribution, and the Expectation-Maximization algorithm is used to iteratively estimate the tissue classes and the gain field. The "E" step of the algorithm estimates the expected tissue class at each voxel using maximum likelihood, and the "M" step computes the gain artifact based on the tissue class estimates and prior knowledge of smoothness of the gain field. The two components are iterated to convergence in this manner. This method has been used to successfully segment 1000 scans in a multiple sclerosis study without manual intervention, which makes it the most aggressively tested intensity based classifier around.

A different statistical approach is illustrated in the atlas based segmentation method of Evans et al. at McGill University, Montreal [32]. They have constructed an average brain volume by registering over 300 normal brains in Tailarach space, and then averaging the voxels independently. Their approach to segmentation consists of registering a given data set to their average brain model, and then transferring the probabilities of different structure locations from the model to the data to produce a labeling of the data. In particular, they have used this method in conjunction with an intensity based classifier to differentiate between white matter lesions (which have gray level appearance similar to that of grey matter) and grey matter in the brain by using the atlas to generate a spatial prior for the actual grey matter.

While the idea of completely eliminating expert intervention makes explicit shape modeling an extremely appealing approach, the contribution of such averaged models is limited in segmentation of detailed structures that exhibit considerable variation from person to person. The neocortex, for instance, with its intricate sulco-gyral patterns, when averaged over several data sets will result in a smooth surface for the model. In this case, such a model would be of limited help in segmentation, since the critical structures have been smoothed.

Cootes et al. describe a different representation for the shape and grey level appearance of anatomical objects called "active shape model"[33]. The two components of this work are a "point distribution model" that describes the shape and gray level appearance of an object, and a local optimization method that improves an initial alignment of the model with a given data set. In the point distribution model, principal component analysis is used to separately capture the modes of variation of the shape as well as the gray level appearance of the region of interest from sample segmentations. An approximate location of the model in the data is achieved and refined by the use of genetic algorithms. Demonstrated segmentations include the left ventricle in echocardiograms, and brain ventricles in MRI.

Bibliography

[1] W.E.L. Grimson, T. Lozano-Pérez, W. Wells, et al. An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced realigy visualization. In *Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA., june 1994. IEEE.

[2] M. Shenton, R. Kikinis, F. Jolesz, et al. Abnormalities of the Left Temporal Lobe and Thought Disorder in Schizophrenia. *N. Engl. J. Med.*, (327):604–612, 1992.

[3] S. Sandor and R. Leahy. A 3d morphological algorithm for automated labelling of the cortex in magnetic resonance brain images. In *AAAI Spring Symposium Applications of Computer Vision in Medical Image Processing*, Paulo Alto, CA., March 1994. AAAI.

[4] G. Ettinger, W. Grimson, T. Lozano-Pérez, W. Wells, S. White, and R. Kikinis. Automatic Registration for Multiple Sclerosis Change Detection. In *Proceedings of the IEEE Workshop on Biomedical Image Analysis*, Seattle, WA., 1994. IEEE.

[5] C. Westbook and C. Kaut. *MRI in Practice*. Blackwell Scientific, 1993.

[6] Wehrli. Principles of nuclear magnetic resonance.

[7] H.E. Cline, W.E. Lorensen, R. Kikinis, and F. Jolesz. Three-Dimensional Segmentation of MR Images of the Head Using Probability and Connectivity. *JCAT*, 14(6):1037–1045, 1990.

[8] W. Wells, R. Kikinis, W. Grimson, and F. Jolesz. Statistical Intensity Correction and Segmentation of Magnetic Resonance Image Data. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[9] W. Wells, R. Kikinis, W. Grimson, and F. Jolesz. Statistical Intensity Correction and Segmentation of Magnetic Resonance Image Data. In *Proceedings of the First Conference on Computer Vision, Virtual Reality and Robotics in Medicine*. Springer, 1995.

[10] J. Serra. *Image Analysis and Mathematical Morphology*. London Academic, 1982.

[11] R.M. Haralick, S.R. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *IEEE Transactions PAMI*, 9:532–550, 1987.

[12] A. Witkin, M. Kass, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, june 1988.

[13] L. Cohen. On active contour models and balloons. *Computer Vision, Graphics and Image Processing: Image Understanding*, 53(2):211–218, march 1991.

[14] M.E. Brummer, R.M. Mersereau, R.L. Eisner, and R.R.J. Lewine. Automatic detection of brain contours in MRI data sets. *IEEE Transactions on Medical Imaging*, 12(2):153–166, june 1993.

[15] R. Kikinis, M. Shenton, F.A. Jolesz, G. Gerig, J. Martin, M. Anderson, D. Metcalf, C. Guttmann, R.W. McCarley, W. Lorensen, and H. Cline. Quantitative Analysis of Brain and Cerebrospinal Fluid Spaces with MR Imaging. *JMRI*, 2:619–629, 1992.

[16] G. Gerig, W. Kuoni, R. Kikinis, and O. Kübler. Medical Imaging and Computer Vision: an Integrated Approach for Diagnosis and Planning. In *Proc. 11'th DAGM Symposium*, pages 425–443. Springer, 1989.

[17] *IBM POWER Visualization System: Programmer's Technical Reference*. Number SC38-0487-00. First edition edition, 1993.

[18] Y.J. Zhang. Segmentation evaluation and comparison: a study of various algorithms. In *SPIE Proceedings of Visual Communication Image Processing*. SPIE, November 1993.

[19] C.N. de Graaf, Koster A.S.E, K.L. Vincken, and M.A. Viergever. A methodology for teh validation of image segmentation methods. In *Proceedings Fifth Annual IEEE Symposium on Computer-Based Medical Systems*, pages 17–24. SPIE, June 1992.

[20] J.P. Mellor. *Real Time Camera Calibration.* Master's Thesis, Massachusetts Institute of Technology, Cambridge, MA, 1995.

[21] Hyche. In *Proceedings of the First Conference on Visualization in Biomedical Computing*. SPIE, 1992.

[22] A. Chakraborty, L.H. Staib, and J.S. Duncan. An integrated approach to boundary finding in medical images. In *Proceedings of the IEEE Workshop on Biomedical Image Analysis*, Seattle, WA., 1994. IEEE.

[23] P. Shi, A. Amini, G. Robinson, A. Sinusas, C.T. Constable, and J.S. Duncan. Shape-based 4d left ventricular myocardial function analysis. In *Proceedings of the IEEE Workshop on Biomedical Image Analysis*, Seattle, WA., 1994. IEEE.

[24] J. Gobel and J. Snell and K. Hinckley and N. Kassell. A Real-Time System for 3D Neurosurgical Planning. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[25] P. S. Huppi, M. Tsuji, T. Kapur, P. Barnes, R. Kikinis, and J. Volpe. Quantitative brain development in preterm infants assessed by 3d mri. *In Preparation.*

[26] C. Poon, M. Braun, R. Fahrig, A. Ginige, and A. Dorrell. Segmentation of medical images using an active contour model incorporating region-based image features. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[27] M. Syn, J. Gosling, R. Prager, L. Berman, and J. Crowley. Tracking the inter-frame deformations of structures in 3d ultrasound imaging. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[28] R. Grzeszczuk and D. Levin. Brownian strings: segmenting images with stochastically defformable contours. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[29] L.H. Staib and J.S. Duncan. Boundary finding with parametrically deformable contour models. *PAMI*, November 1992.

[30] C.H. Chang, W. Wang, A. Rangarajan, R.T. Constable, I.G. Zubal, and G. Gindi. Segmentation of mr brain scans using atlas – guided deformable contours. In *AAAI Spring Symposium: Applications of Computer Vision in Medical Image Processing*, Paulo Alto, CA., March 1994. AAAI.

[31] "J. O'Brien and N. Ezquerra". Automated Segmentation of Coronary Vessels in Angiographic Image Sequences Utilizing Temporal, Spatial and Structural Constraints. In *Proceedings of the Third Conference on Visualization in Biomedical Computing*. SPIE, 1994.

[32] D. L. Collins, T.M. Peters, W. Dai, and A.C. Evans. Model based segmentation of individual brain structures from mri data. In *Proceedings of the First Conference on Visualization in Biomedical Computing*. SPIE, 1992.

[33] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. Use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6), 1994.