Brain Segmentation and the Generation of Cortical Surfaces

Mukta Joshi,* Jing Cui,† Keith Doolittle,* Sarang Joshi,* David Van Essen,† Lei Wang,* and Michael I. Miller‡

*Department of Electrical Engineering, Washington University, St. Louis, Missouri 63130; †Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri; and ‡Center for Imaging Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, Maryland 21218

Received August 5, 1998

This paper describes methods for white matter segmentation in brain images and the generation of cortical surfaces from the segmentations. We have developed a system that allows a user to start with a brain volume, obtained by modalities such as MRI or cryosection, and constructs a complete digital representation of the cortical surface. The methodology consists of three basic components: local parametric modeling and Bayesian segmentation; surface generation and local quadratic coordinate fitting; and surface editing. Segmentations are computed by parametrically fitting known density functions to the histogram of the image using the expectation maximization algorithm [DLR77]. The parametric fits are obtained locally rather than globally over the whole volume to overcome local variations in gray levels. To represent the boundary of the gray and white matter we use triangulated meshes generated using isosurface generation algorithms [GH95]. A complete system of local parametric quadratic charts [JWM95] is superimposed on the triangulated graph to facilitate smoothing and geodesic curve tracking. Algorithms for surface editing include extraction of the largest closed surface. Results for several macaque brains are presented comparing automated and hand surface generation.

1. INTRODUCTION

As we move into the 21st century, medical imaging is becoming more common place. It is emerging as a very powerful tool in diagnostics and in the understanding of anatomy. We and others have been studying cortical brain anatomy in trying to understand the folding and geometry of the cortex [Ess97, EDJM97, OKA90, RGK92, MFB94, Hol95, ATP96, DB96, TSL96, MDBD98]. It is our goal to have a complete methodology which allows for the generation of a topologically correct cortical surface from image data. A mathematical representation for the surface consisting of a family

of local quadratic charts [JWM95] makes it possible to map one surface geometry to another in the context of deformable templates [JGM97, MBC97]. With the cortical surface digitally and mathematically represented, it is possible to automatically trace sulci and gyri on the surface [KGM98, STA96, CT98, MBD98]. These features are of importance in the understanding of the functionality of brain structures [EDSS, DEA96].

As an illustration, shown in Fig. 1 is the 3D rendering of a typical macaque brain (Fig. 1a). A section through this brain reveals the gray and white matter regions (Fig. 1b). The triangulated mesh representation of the cortical surface shows the folds of the cortex (Fig. 1c). Fig. 1d shows the intersection of this surface with a section, illustrating that the surface constructed sits at the interface of the gray and white matter.

The principal focus of this paper is on the description of a semiautomated methodology for the generation of cortical surfaces from image data. There has been a great deal of work done on segmentation and boundary finding. Early image segmentation based on MAP methods on Markov random fields are described in [HE82, GD84, DE87]. Deformable contour methods and more global shape models have been used as well [MT95, XP97, DS93]. In [ZY96] Zhu and Yuille describe an algorithm which combine the more global approaches with local segmentation.

Our work is most akin to the beautiful approach described in [CSW97], essentially extending their methodology to incorporate automated renormalization of local inhomogeneities in the image field. This coupled with connectedness notions [CSW97] and local quadratic charts provides an extended framework for generation of surfaces.

There are three basic aspects to this method.

1. Local parametric Bayesian segmentation: The first step in the procedure is the segmentation of the brain image data into multiple compartments including white matter and gray matter. This is done by parametrically fitting known density functions to the histogram of the image using the expectation maximization (EM) algorithm [DLR77, KGWIK96, WGKJ96]
defined locally throughout the volume estimating the means and variances of the local histogram of the data.

2. Isosurface generation: We use triangulated meshes to represent the outer surface of the white matter generated using isosurface generation algorithms [GH95]. Superimposed on the graph is a complete system of local parametric quadratic charts representing the surface as a twice continuously differentiable manifold supporting curvature calculations such as used in [J WM'95, SZ92, Gri94, TG95].

3. Segmentation and surface editing: Manual interactions with segmentations and surfaces are enabled.
via segmentation and surface editing tools. Algorithms are implemented for extracting the largest closed surface and removing shells, along with triangle reduction algorithms, and surface smoothing algorithms.

Given such representations including the local smooth coordinate systems, sulcal, gyral, and geodesic curve generation methods based on dynamic programming algorithms as described in [KGM98] may be incorporated. The subsequent sections describe the segmentation algorithm via probabilistic modeling and Bayesian hypothesis testing, along with surface editing algorithms. Section 2.3 presents results of our methods on six macaque brain hemispheres and the Visible Human brain hemisphere. Methods used to quantify the accuracy of the segmentations and the surfaces are also discussed.

2. SEGMENTATION

2.1. Data Acquisition

Macaque cryosection. Much of the results shown in this paper are generated using cryosection images of macaque brains. These images were acquired at Dr. David Van Essen’s Laboratories at the Washington University School of Medicine. The image of the frozen brain is acquired with a charged coupled device camera. A section of 0.1 mm thickness is removed and the next image is acquired. This procedure is called cryosection imaging. For all cryosection macaque brains shown $n_1 \times n_2 = 640 \times 480$, $n_3 \approx 80$ representing on the average 0.1 mm resolution in plane and 0.1 mm resolution cross plane.

Visible Man. The Visible Man data set is a high resolution data set obtained by the cryosection imaging procedure, as described earlier. The resolution for this set is 0.33 mm in plane and 1 mm cross plane.

2.2 Bayesian Segmentation

Mixture modeling. Regions of the brain not containing cerebrospinal fluid are largely composed of gray or white matter. Shown in Fig. 2a is a section from the cryosection image of a macaque brain together with the histogram for that section. Figure 2b depicts the section with the hand segmentation of the gray matter overlaid on it. Using this hand segmentation we could separate the histogram for the gray matter which is shown below it. Figure 2c shows a similar histogram obtained from the hand segmentation of the white matter as shown above it. The histogram shows the distribution of the pixels over the (0–255) intensity range. The histogram is modeled as a mixture, a superposition of distributions representing the gray and white matter. The goal is to construct first and second order parametric representations for the mixture.

Define the entire image volume $X = [x(i), i \in \mathcal{I}]$, where $\mathcal{I}$ is the three-dimensional voxel lattice, $\mathcal{I} = n_1 \cdot n_2 \cdot n_3$, and $x(i) \in [0, 1, \ldots, 255]$. Define local neighborhoods of each of the voxels $i \in \mathcal{I}$ to be $N(i) \subset \mathcal{I}$, with the histogram of the number of pixels of value $v$ in neighborhood $i$ defined according to $h_i(v) = \sum_{j \in N(i)} \delta[X(j) - v]$, where $\delta[X(j) - v] = 1$ if $X(j) = v$ and 0 otherwise. Suppressing subscript $i$, model the histogram data as a superposition of $M$ weighted functions,

$$\sum_{m=0}^{M-1} \alpha_m f_m(v - \mu_m) / \sigma_m,$$ (1)

where $\alpha_m$ is the area parameter of the density function $f_m$, and $\mu_m$ and $\sigma_m$ are the mean and standard deviation, respectively. The maximum-likelihood estimates of $(\alpha_m, \mu_m, \sigma_m)$ are computed locally throughout the volume using the iterative EM algorithm (see Appendix for the EM algorithm).

Arbitrary histogram shapes. We have used both predefined (Gaussians) parametric forms for the fitting of the histogram data as well as arbitrary histograms with amplitude scale and shift as parameters. For functions which cannot be represented in closed form, we estimate the amplitudes, scales, and shifts using the EM algorithm by numerically solving the maximization in the M-step. In the context of the brain segmentation problem the following procedure is followed for fitting the histogram using arbitrary discrete functions. Take a sample set of points, for example, a section through the brain. Segment this section manually into various regions, $m = 1, 2, \ldots, M$, with the segmentations thus obtained generating histograms for the $M$ regions. This produces empirical curves that represent the distribution of points in each region. Take these as the basis functions $f_m$, $m = 1, 2, \ldots, M$, $M$ the number of regions. Estimate the parameters for these functions to fit the histograms of all the sections in the brain. As an illustration, brain 87A was segmented.

**FIG. 3.** (a) A section of brain 87A shown with the two types of histogram fits and the respective segmentations. (b) Histogram with the arbitrary functions fit: segmentation using Bayes test on this fit. (c) The section overlaid with the hand segmentation. (d) Histogram with the Gaussian function fit: Segmentation using Bayes test on this fit.

**FIG. 5.** (a) Two consecutive sections and the side view of the brain 95BL; (b) the same sections with the segmentation, via global threshold overlaid on the data; (c) these sections with the segmentations based on the local parametric fits overlaid on the data.
using both the Gaussian fits and the arbitrary function fits without any kind of human interaction. Figure 3 shows two sections through brain 87A. Figure 3a shows the corresponding histograms superimposed with the estimated histogram using arbitrary curves. Figure 3b shows the histogram superimposed with the Gaussian fits and Fig. 3c shows the segmentation using these fits. Notice how the fit using arbitrary curves is more accurate than that using the Gaussian predefined form.

The measure of similarity between two distributions can be quantified using the Kullback-Leibler distance or the relative entropy between two distributions. The relative entropy between two distributions \( p \) and \( q \) is defined as \([\text{CT91}]\)

\[
D(p|q) = E[\log(p/q)] = \sum_x p(x) \log \frac{p(x)}{q(x)}. \tag{2}
\]

This is not a distance in the true sense as it is not symmetric and it does not satisfy the triangle inequality.

To quantify closeness of the segmentation methods, the improvement in the fits due to the use of arbitrary, empirical distributions over the use of Gaussians was calculated. For the sections studied, the cross-entropy for the arbitrary histogram model was 0.19; the cross-entropy for the true data was calculated. For the sections studied, the cross-entropy between the Gaussian model and the true data was 0.09.

To apply Bayesian hypothesis testing, we use the log likelihood ratio. With \( \mu_g \) and \( \mu_w \) representing the means and \( \sigma_g \) and \( \sigma_w \) representing the standard deviations of the gray and white matter, respectively, the optimum Bayes threshold becomes

\[
\frac{v - \mu_g}{\sigma_g} \bigg/ \frac{v - \mu_w}{\sigma_w} \Leftrightarrow \ln \eta \tag{3}
\]

where \( \eta \) is the system threshold, which is determined by \( \alpha_g \) and \( \alpha_w \), probability of hypothesis gray and white, respectively, and \( C_{kl} \) is the cost for choosing hypothesis \( H = k \) given \( H = l \). The costs determine the threshold and significantly influence the segmentation. The results shown hereafter assume a Gaussian fitting model. We have taken the costs to be equal between white and gray matter, and we choose the prior probabilities \( \alpha_g \) and \( \alpha_w \) to be the relative proportion of white and gray matter within a neighborhood of the parametric histogram fitting.

2.3. Quantifying the Accuracy of Segmentation

The cryosection data we have been examining exhibit variations in the contrast from one section to another, due to the changes in the parameters of the imaging system. The EM algorithm has been implemented over arbitrarily defined neighborhoods in the image. For all of the data being studied, the model was modified to estimate the parameters over planar neighborhoods \( N(i) \) only since the data reflected variations of this kind. More generally, \( N(i) \) can be defined as any subregion of the image. Figure 4 shows the change of the means of the gray and white matter with sections in each of the six macaque hemispheres.

To illustrate the necessity of using the local planar neighborhood, we have compared the results of the global model which assumes the whole volume as the neighborhood and the local model that assumes the planar neighborhood.

Figure 5a illustrates an example of a brain that has poor contrast coherence. The side view (right panel) shows the variation. Shown in Fig. 5b is the result of applying Bayes segmentation assuming \( N(i) \) to be the entire volume, giving a global threshold. A global threshold causes the model to miss entire sections, as the intensities are lower than the global threshold chosen. Notice that section 51 is not segmented (middle panel of Fig. 5b).

Shown in Fig. 5c is the result of using the local parametric fitting. Notice in Fig. 5c, middle panel, that the section 51 is segmented into white matter and gray matter. Notice from Fig. 5c, right panel, that the model segments the entire volume.

2.4. Quantifying the Accuracy of Segmentations

To quantify the accuracy of the automatic segmentations we have compared them to reference segmentations generated by trained individuals. For measuring the accuracy we use the normalized variational distance, between positive summable functions. The expression for this is given in Section 6.1 in the appendix.
The left two columns of Table 1 show the accuracy of the segmentations obtained using the global and local fitting of the parametric mixtures with global and planar neighborhoods, respectively. Notice the improvement of the local parametric fitting over the global parameterization. For the brain 87A, which has relatively small intensity inhomogeneity, the local parametric fitting improved only marginally from 0.185 to 0.17. For brains 95BL and 93I, where there was a marked variation in illumination (as shown in Fig. 4), the accuracy improved significantly (95BL, 0.24 to 0.12; 93I, 0.19 to 0.14).

The manual segmentations are most uncertain at the boundary of the gray and white matter [BCH '95]. The confidence level of marking white matter away from the boundary is higher. For this reason, we have also calculated the normalized variational distance neglecting a single layer of pixels on the boundary. The two columns on the right of Table 1 show the accuracies for

![Fig. 4](image_url) Plots of the means of white and gray matter against the section number for each of the eight data sets. Each mean is the empirical estimate found by the EM algorithm for each section. The means are fairly constant for data sets of 87A and 95CL; all others demonstrate large variations in their parameters.

![Fig. 11](image_url) (a) The surface of brain 95CL with Euler number 2; (b) a section showing the surface embedded in the volume; (c) the surface with the curvature map and the major sulci traced; (d) the surface with the major gyri traced.

![Fig. 12](image_url) (a) The surface of the superior temporal gyrus (STG) of the Visible Man using parametric models with the STG gyrus crown traced by the dynamic programming algorithm; (b) the Van Essen hand-generated surface of the STG in the Visible Man; (c) the Van Essen-generated surface (red) together with the Bayesian parametric surface (blue) embedded in the region of the STG.
the segmentations with this boundary removed. To illustrate the relative quality of the automated segmentations, four experts manually segmented the white matter in a collection of sections of a monkey cortex data set. Table 2 shows the distance (without boundary) of the four manual segmentations with respect to a single separate expert segmentation. The right column shows the distance for the automated segmentation obtained by local parametric representation of the data and Bayes thresholding. Notice that the accuracy obtained by the automated segmentation is comparable to that achieved by human experts.

2.5. Comparison to Renormalization

In order to segment images we have used the Bayes method. There are many alternative approaches. Histogram equalization and contrast enhancing algorithms have been historically used in an attempt to facilitate image segmentation. We term such approaches normalization methods. To illustrate an alternative approach we can perform normalization of the imagery using standard linear re-mapping of between the means. For this, we have taken the estimated means of each brain gray and white matter $\mu_g$ and $\mu_w$ and mapped to the fixed values of 100 and 200, respectively. All pixels with values that lie between $0 - \mu_g$ are linearly mapped to $0$–100. Similarly, the range of pixel values from $\mu_g$ to $\mu_w$ are mapped to 100–200, and the values $>\mu_w$ to values $>200$. Figure 6 shows the results of this method for the brain 95BL that has marked illumination variation. Shown in Fig. 6b are the axial and sagittal views of 95BL. These images show low contrast and large variations in the illumination across sections. In Fig. 6c are shown the same views, after renormalization. This image renormalization causes the volume to have uniform gray values. A global threshold can be applied to obtain a segmentation of the white matter, the global threshold being the average of the two means (100 and 200), i.e., 150.

Shown in Table 3 are the results of segmentations using the renormalization. The left two columns show the accuracies of the segmentations by the global parametric method and the remapped method, respectively. The right two columns show the distances when the boundary of the manual segmentations is not considered. In general, the remapping improves the accuracy over the global parametric method; however, the renormalization cannot perform better than the Bayes approach from an information point of view. Compare the numbers in the two tables (Tables 1 and 3).

3. SURFACE GENERATION

Once the white matter segmentation is obtained, the next step is to generate a surface that represents the outer boundary of this white matter volume thus defining the interface between the gray matter and white matter.

Isosurface generation. For generating triangulated surfaces we have used an isosurface generation algorithm by Gueziec and Hummel [GH95] that uses tetrahedral decomposition. This algorithm is similar to the “Marching Cubes” algorithm [LC87]. The Marching Cubes algorithm has some inherent ambiguities that can result in surfaces containing holes in them [VGW94, TG96]. The tetrahedral decomposition algorithm generates a tessellation of triangles for an isosurface of a given intensity. It decomposes each voxel (eight points) into five tetrahedra. The resulting surface is a closed and valid triangulation made up of the triangles that bound these tetrahedra. We have implemented this algorithm and have used it for generating all the surfaces presented in this paper. This algorithm produces a very large number of triangles that makes rendering of the surfaces computationally intense.
Quadratic chart generation. We define surfaces to be manifolds supported by local coordinate systems with at least two derivatives allowing for the generation of curvature. We follow the automated procedure of [Ham93] described in [JWM95] for calculating these quadratic patches. Consider a surface $M$. We consider all the points on the graph that are directly connected to a vertex $i \in M$. These points form the neighborhood $N(i)$.

For each point $i \in M$, the surface is expressed as the graph of the function $z(x, y) = f(x, y)$, such that it is locally quadratically approximated by $f(p) = f_x(p) = f_y(p) = 0$, with

$$f(x, y) = \frac{1}{2} \left( f_{xx} x^2 + 2f_{xy} xy + f_{yy} y^2 \right),$$

with the $2 \times 2$ shape operator defined via the curva-

TABLE 3

<table>
<thead>
<tr>
<th>Brain</th>
<th>With boundary</th>
<th>Without boundary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Global parametric</td>
<td>Normalized imagery</td>
</tr>
<tr>
<td>87A</td>
<td>0.185</td>
<td>0.174</td>
</tr>
<tr>
<td>95BL</td>
<td>0.246</td>
<td>0.130</td>
</tr>
<tr>
<td>95CL</td>
<td>0.210</td>
<td>0.209</td>
</tr>
<tr>
<td>93G</td>
<td>0.150</td>
<td>0.156</td>
</tr>
<tr>
<td>90C</td>
<td>0.105</td>
<td>0.105</td>
</tr>
<tr>
<td>93I</td>
<td>0.196</td>
<td>0.138</td>
</tr>
</tbody>
</table>

Note. Table shows the distances of segmentation obtained by global thresholding (left) and thresholding after remapping (right), for the eight macaque data sets.

Triangle reduction and redistribution. Marching Cubes and tetrahedrons generate triangulations that tend to be over parameterized. Work on mesh decimation have been done [SZL92, HDD93]. In this paper, we have developed a process for the reduction of the number of triangles needed to represent a surface while preserving the original shape and geometry. This involves a recursive procedure that repeatedly removes triangles at every stage giving different levels of detail at each stage. The vertices and edges that can be removed are the ones which will result in the least amount of shape change. There are two criteria that are used to eliminate vertices and edges to achieve this reduction. First, the vertices that lie on a low curvature region of the surface can be removed. Thus, if the curvature at a vertex is below a specified minimum then that vertex is removed. The hole formed due to its removal is then retriangulated. Second, edges that are relatively small in length contribute to long slender triangles mostly of very small areas. Removing these edges does not adversely change the geometry. In eliminating an edge, essentially two vertices are merged. This eliminates all triangles that shared the edge reducing the number of triangles considerably.

To regularize the surface, the triangles are “redistributed.” This is an algorithm used to make the triangles
on the surface as equilateral as possible while preserving topology [JGM97] based on minimizing a cost function along the gradients in the tangent space of the manifold.

The above three passes over the surface constitute an iteration. This process is repeated until no more vertices or edges can be removed, resulting in a surface with the minimum possible triangles under the specified minimum curvature and edge lengths required to be maintained. For the surfaces generated in this paper, there was a 60% reduction in the total number of triangles from the original triangulation following such a procedure.

Smoothing. The surfaces generated from voxel data are not smooth. In order to smooth the surface we use the local coordinate charts and locally approximate the surface as a quadratic patch. The vertices on the original surface are then moved to the nearest point on the parametric patch, thus smoothing the surface locally. The neighborhood chosen to fit the quadratic patch can be varied as well.

Extract. In order to remove the numerous external disjoint surfaces and cavities or shells that are formed we have implemented an algorithm called “Extract.” Vertices on the isosurface are “colored” in a simple recursive region growing algorithm. Neighboring vertices are given the same color and the number of vertices belonging to a certain color are kept track of. When all vertices have been colored, the program outputs only those vertices and corresponding triangles that belong to the largest color group, i.e., the largest connected set.

Local Euler number calculation. The cortex of the brain is considered to be a single sheet without any holes or toroidal structure in it. Since we generate the cortical surface as the isosurface enclosing the white matter, we require the white matter segmentation to be devoid of holes and toroidal structures. This implies that the white matter is topologically equivalent to a sphere and falls in the set of simply connected structures. A simply connected volume or surface enclosing the simply connected volume is identified by the Euler number, whose value is 2 for any simply connected surface. The Euler number is calculated on a triangulated surface according to

$$\chi = f - e + v,$$

where v is the number of vertices, f is the number of faces (polygons/triangles), and e is the number of edges. Since the segmentation process is not constrained, the white matter component may not be simply connected. Thus, the surface may have toroidal structures. In order to identify the location of these structures we use an algorithm that locally checks the Euler number. The whole surface can be traversed in patches, a patch corresponding to a portion of the surface that is located at one of the vertices and spans n neighborhoods. These patches will not be closed surfaces. A valid patch that is a part of a simply connected closed surface should be equivalent to a plane and thus its Euler number must be 1. If the Euler number is not 1 then the patch is highlighted indicating a topological problem. The region corresponding to the patch is then indicated in the segmentation so that the segmentation can be modified to improve the topology.

The surface of the brain 95CL was modified until the Euler number was 2. This surface is shown in Fig. 11a.

Once the local quadratic charts are generated, on the triangulated surface a curvature map may obtained by computing the principal curvature at each node. Based on the curvature map dynamic programming algorithms finding the minimum cost route from the start point to the end point along points of extremal curvature (for sulcal and gyral curves) and along points that minimize the total length for geodesic curves can be implemented [KGM98, Kha96].

4. RESULTS OF SURFACE GENERATION

Figure 7 shows the surfaces that have been generated for the six macaque data sets. The procedure begins with local parametric fitting to the data histogram, from which the Bayes threshold for each planar neighborhood is calculated. Manual modifications using the editing tools were applied as well. The isosurface generation algorithm produces the cortical surfaces from the segmentations with smoothing and retriangulation applied to improve the quality of the surface.

To examine surface accuracy with respect to the volumes from which they were generated, the surfaces are embedded in the respective volumes. Shown in Fig. 8 are sections through each brain, showing the gray level cryosection data in gray and the points of intersection of the surface with the volume in red. Notice the near exact intersection of the surface with the gray-white matter boundary.

4.1. Segmentation and Surface Editing

Obtaining a single surface. We are interested in generating a single surface at the interface of gray and white matter. There may be many disconnected surfaces that are generated due to the disconnected nature of the segmentation. After the isosurface has been generated it is desirable to remove any smaller surfaces that are not attached to the original surface. These smaller surfaces can be produced, for example, by single one-valued voxels in the segmentation that are not connected to any other one-valued voxels. To remove these smaller disconnected surfaces we use the
The implementation of the algorithm "Extract" described in Section 3. To illustrate the difference caused by applying this algorithm, shown in Fig. 9 is a surface of the macaque brain 95BL before and after applying Extract to it.

Manual modification. Tables 1 and 2 illustrate that the errors in segmentations are comparable to the variation in accuracy of different manual segmentations. However, the quality of the surfaces obtained from these segmentations cannot be directly judged from these numbers. Surface accuracy is determined by the topology. We have found that slight changes in the segmentations can cause large changes in the topology of the resulting surface.

To illustrate, the segmentation of the 87A monkey brain was modified slightly in a small portion of the

FIG. 7. Surfaces generated by applying the isosurface generation algorithm with triangle reduction and the extract algorithm.

FIG. 9. (a) Surface of 95BL; (b) surface of 95BL after applying "Extract."
brain. Two sections from this portion are shown in Fig. 10a with the region where the changes were made encircled. Figure 10b shows the surface generated from this segmentation. Figure 10c shows the modified segmentation for the sections in Fig. 10a. Figure 10d shows the surface generated from this modified segmentation.

The change in normalized variational distance between segmentations due to this modification was from 0.1553 to 0.1549—an insignificantly small amount of improvement. However, the change in surface topology was very visible. This difference is shown in the right column.

We have spent varying amounts of time on each of the brains to manually modify parts of the segmentations to make the resulting surfaces more topologically correct. For brains 87A, 90C, 93G, 93I, and 95BL, hand editing was less than 1 h total for each of the brains. Brain 95CL was modified until the Euler number of the surface was 2; this required several days. The brain is shown in Fig. 11.

Figure 11 shows different results on the brain 95CL. The segmentation of 95CL was hand modified to improve the topology of the surface. Approximately 35 h were spent on this brain and the resultant surface with Euler number 2 (correct topology) is shown in Fig. 11a. Figure 11b shows a section through the volume with the surface embedded. Figure 11c shows the surface with the curvature coding. The major sulci are traced in white. Figure 11d shows the curvature map and shows the major gyri colored in white.

5. VISIBL HUMAN BRAIN

The method was used to generate the surface of the region around the superior temporal gyrus (STG) in the Visible Man data set. A part of the whole brain that corresponds to the STG was cut out and segmented. The resolution of this data set was 0.33 mm in plane and 1 mm cross-plane. The size of the region cut was 55 × 90 × 60 pixels. This data set produced a 289,280 triangle surface and is shown in Fig. 12a.

A surface was generated by marking the region of the white matter by visual inspection of the data by human experts. This work was done at Dr. Van Essen’s laboratory at the Washington University School of Medicine [DE96]. For the sake of comparison, the expert hand-generated surface is shown in Fig. 12b compared to the semiautomated one in Fig. 12a. For comparison in the volume we have embedded both surfaces in the Visible Man volume data. This is shown in Fig. 12c. The dark blue line represents the expert generated surface, and the red line represents the automated parametric surface which was generated. Notice how accurately the algorithm has found the boundary between the gray and white matter in this human brain data set. The principal difference is that the hand-generated surface is at the middle of the cortical mantle, called layer 4. Having generated the surface with its quadratic charts for the STG region, we applied the dynamic programming algorithms of Khaneja to generate the gyral crowns of Khaneja. This is shown superimposed in white.

6. APPENDIX

6.1. Normalized Variational Distance

Let $\mathcal{P}^1$ and $\mathcal{P}^2$ be two measures over the lattice $\mathcal{L}$, where $P_i$ is the value of the measure at $i \in \mathcal{L}$. The variational distance, $d_{12}$, between $\mathcal{P}^1$ and $\mathcal{P}^2$ is defined as [CT91]

$$d_{12} = \frac{1}{2} \left| \mathcal{P}^1 - \mathcal{P}^2 \right|_1 = \frac{1}{2} \sum_{i \in \mathcal{L}} \left| P_i^1 - P_i^2 \right|_1,$$

where $\left| \mathcal{P}^1 - \mathcal{P}^2 \right|_1$ is the $L_1$ distance between the two measures. Let $\mathcal{M}$ be a manual segmentation and $\mathcal{A}$ an automated segmentation. Define the segmentation masks as $P_i^j \in \{\mathcal{M}, \mathcal{A}\}$, $i \in \mathcal{L}$, $\mathcal{L}$ being the lattice of voxels representing the volume.

$$P_i^{\mathcal{M}} = 1 \text{ if manual (automatic) segmentation is white,} = 0 \text{ if manual (automatic) segmentation is gray.}$$

The normalized variational distance, $d_{\mathcal{M}, \mathcal{A}}$, becomes

$$d_{\mathcal{M}, \mathcal{A}} = \frac{1}{2} \left| P^{\mathcal{A}} - P^{\mathcal{M}} \right|_1 = \frac{1}{2} \sum_{i \in \mathcal{L}} \left| P_i^{\mathcal{A}} - P_i^{\mathcal{M}} \right|_1.$$

6.2. The EM Algorithm for Estimating the Parameters for a Mixture Density

The intensity of each pixel is drawn from one of $M$ parametric densities of the form $\alpha_m f_m(x|\mu, \sigma)$, where $\alpha$, $\mu$, and $\sigma$ are the parameters and $\sum_{m=1}^{M} \alpha_m = 1$; define the vector of parameters $\theta = ([\alpha_1, \mu_1, \sigma_1], \ldots, [\alpha_M, \mu_M, \sigma_M])$. The complete data (CD) becomes the set of pairs $(x(1), i_1), \ldots, (x(j), i_j), \ldots , (x(N), i_N)$, where $i_j \in \{1, 2, \ldots, M\}$ is the tag associated with the $j$th sample that identifies which of the parametric density functions the sample $x(j)$ was drawn, with incomplete data (ID) are the collected samples $x = (x(1), x(2), \ldots, x(j), \ldots, x(N))$, where $x(j)$ is the data value at pixel $j$ and $N$ is the number of pixels in the image.
complete to incomplete data is the many to one mapping discarding the tag in the pair that defines the complete data. Consider the indicator function \( \delta[m - i_j] \) defined as \( \delta[m - i_j] = 1 \) if \( m = i_j \) and 0 otherwise, where \( i_j \) is the tag attached to the sample \( x(j) \). The complete data log-likelihood is

\[
L_{CD}(\theta) = \sum_{j=1}^{N} \sum_{m=1}^{M} \delta[m - i_j] \ln \alpha_m p_m(x(j) | \theta).
\] (6)

The EM algorithm requires two steps, the E-step and the M-step.

The E-step. Assume parameter \( \theta^k = (\alpha_k, \mu_k, \sigma_k) \) for iteration \( k \); then with \( E[\delta[m - i_j] x(j), \theta^k] = \alpha_k^0 p_m(x(j) | \theta^k) \) this gives

\[
Q(\theta | \theta^k) = E[L_{CD}(\theta) | x, \theta^k] = \sum_{j=1}^{N} \sum_{m=1}^{M} E[\delta[m - i_j] \ln \alpha_m p_m(x(j) | x, \theta^k)]
\]

\[= \sum_{j=1}^{N} \sum_{m=1}^{M} \alpha_m^k p_m(x(j) | \theta^k) \ln \alpha_m p_m(x(j) | \theta^k).\]

When the \( N \) data samples are grouped into \( V \) bins of size \( \Delta v \), \( h(v) \) the normalized histogram of the image with sampled probability density \( f_m(v | \theta^k) = \int_{v-\Delta v}^{v+\Delta v} p_m(u | \theta^k) du \), then

\[
Q(\theta | \theta^k) = \sum_{v=0}^{V-1} \sum_{m=1}^{M} h(v) \frac{\alpha_m^k f_m(v | \theta^k)}{\sum_{i=1}^{M} \alpha_m^k f_i(v | \theta^k)} \ln \alpha_m f_m(v | \theta^k).\] (7)

The M-step. The maximization step is carried out with respect to the parameter to be estimated:

\[
\frac{\partial Q}{\partial \alpha_m} = \sum_{v=0}^{V-1} h(v) \frac{\alpha_m^k f_m(v | \theta^k)}{\sum_{i=1}^{M} \alpha_i^k f_i(v | \theta^k)} \frac{1}{\alpha_m} = 0 \quad \text{subject to} \quad \sum_{m=1}^{M} \alpha_m = 1,
\] (8)

\[
\frac{\partial Q}{\partial \mu_m} = \sum_{v=0}^{V-1} h(v) \frac{\alpha_m^k f_m(v | \theta^k)}{\sum_{i=1}^{M} \alpha_i^k f_i(v | \theta^k)} \frac{1}{\mu_m} = 0,
\] (9)

Then for the Gaussian case the iteration becomes

\[
\alpha_{m+1}^k = \alpha_m^k \sum_{v=0}^{V-1} h(v) \frac{G_m^k(v)}{s^k(v)},
\] (11)

\[
\mu_{m+1}^k = \sum_{v=0}^{V-1} h(v) \frac{G_m^k(v)}{s^k(v)} / \sum_{v=0}^{V-1} h(v) \frac{G_m^k(v)}{s^k(v)},
\]

and

\[
(\sigma_{m+1}^k)^2 = \sum_{v=0}^{V-1} (v - \mu)^2 h(v) \frac{G_m^k(v)}{s^k(v)} / \sum_{v=0}^{V-1} h(v) \frac{G_m^k(v)}{s^k(v)},
\] (12)

where

\[
G_m^k(v) = \int_{v-\Delta v}^{v+\Delta v} \frac{1}{\sqrt{2\pi\sigma_m}} e^{-\frac{(x-\mu_m)^2}{2\sigma_m^2}} dx.
\] (13)

Theorem 6.1. The stable points of the above iterative EM algorithm satisfy the maximizer conditions for the Maximum Likelihood estimators of the parameters \( \alpha_m \), \( \mu_m \) and \( \sigma_m \).

Examine the case \( M = 2 \) without loss of generality. From Eq. (11), stable points satisfy the condition that as \( k \rightarrow \infty \) then stable points satisfy \( \lim_{k \rightarrow \infty} (\alpha, \mu, \sigma)^{k+1} = \lim_{k \rightarrow \infty} (\alpha, \mu, \sigma)^k = (\alpha_0, \mu_0, \sigma_0) \) implying \( \sum_{v=0}^{V-1} h(v) (G_m(v | \theta^0)) / \sum_{m=1}^{M} \alpha_m G_m(v | \theta^0) = 1 \) for all \( m \). To see this is satisfied, use the incomplete data log-likelihood \( L_{ID} = \sum_{v=0}^{V-1} h(v) \ln (\alpha_0 G_1(v | \theta^0) + \alpha_2 G_2(v | \theta^0)) \), to examine the maximizer conditions at \( \alpha_1 \):
\[ \frac{\partial L_{ID}}{\partial \mu_m} = \sum_{v=0}^{V-1} h(v) \frac{G_m(\mu^*)}{\sum_{m=1}^2 \alpha_m G_m(\mu^*)} = 1. \] (15)

\[ \frac{\partial L_{ID}}{\partial \sigma_m} = \sum_{v=0}^{V-1} \frac{h(v)}{\alpha_1 G_1(\mu^*) + \alpha_2 G_2(\mu^*)} \frac{1}{\sqrt{2\pi\sigma_m}} e^{-\frac{(v-\mu_m)^2}{2\sigma_m^2}}, \] (16)

Substituting \( \mu_m \) as the estimate given by Eq. (11), the above equation becomes zero implying that the estimate of \( \mu \) satisfies the maximizer condition for the ML estimate:

\[ \frac{\partial L_{ID}}{\partial \sigma_m} = \sum_{v=0}^{V-1} \frac{h(v)}{\alpha_1 G_1(\mu^*) + \alpha_2 G_2(\mu^*)} \cdot G_1(\mu^*) \frac{(v - \mu_m)^2}{\sigma_m^2} - 1. \] (17)

Substituting Eq. (12) in this expression gives zero. A similar proof applies for the \( \sigma_2 \). Thus the estimates given by Eq. (12) satisfy the maximizer condition for the ML estimate and therefore they are the ML estimates of the variances.

**ACKNOWLEDGMENTS**

We wish to thank Heather Drury for the helpful discussions on the Euler patch test algorithm and Shashi Kumar for histological sectioning of the cryosection macaque brains. Also we wish to thank Dr. Michael Vannier for his constructive criticism. Parts of the results shown in this paper for segmentation were done with the help of Analyze AVW 1.0 from the Biomedical Imaging Resources of the Mayo Foundation.

**REFERENCES**


