Elastic Model-Based Segmentation of 3-D Neuroradiological Data Sets

András Kelemen, Gábor Székely,* and Guido Gerig

Abstract— **This paper presents a new technique for the automatic model-based segmentation of three-dimensional (3-D) objects from volumetric image data. The development closely follows the seminal work of Taylor and Cootes on active shape models, but is based on a hierarchical parametric object description rather than a point distribution model. The segmentation system includes both the building of statistical models and the automatic segmentation of new image data sets via a restricted elastic deformation of shape models. Geometric models are derived from a sample set of image data which have been segmented by experts. The surfaces of these binary objects are converted into parametric surface representations, which are normalized to get an invariant object-centered coordinate system. Surface representations are expanded into series of spherical harmonics which provide parametric descriptions of object shapes. It is shown that invariant object surface parametrization provides a good approximation to automatically determine object homology in terms of sets of corresponding sets of surface points. Gray-level information near object boundaries is represented by 1-D intensity profiles normal to the surface. Considering automatic segmentation of brain structures as our driving application, our choice of coordinates for object alignment was the well-accepted stereotactic coordinate system. Major variation of object shapes around the mean shape, also referred to as shape eigenmodes, are calculated in shape parameter space rather than the feature space of point coordinates. Segmentation makes use of the object shape statistics by restricting possible elastic deformations into the range of the training shapes. The mean shapes are initialized in a new data set by specifying the landmarks of the stereotactic coordinate system. The model elastically deforms, driven by the displacement forces across the object's surface, which are generated by matching local intensity profiles. Elastical deformations are limited by setting bounds for the maximum variations in eigenmode space. The technique has been applied to automatically segment left and right hippocampus, thalamus, putamen, and globus pallidus from volumetric magnetic resonance scans taken from schizophrenia studies. The results have been validated by comparison of automatic segmentation with the results obtained by interactive expert segmentation.**

Index Terms— **Automatic 3-D segmentation, elastically deformable surface models, statistical shape models.**

Manuscript received March 1, 1999; revised June 18, 1999. The Associate Editors responsible for coordinating the review of this paper and recommending its publication were M. Viergever and W. Niessen. *Asterisk indicates corresponding author*.

A. Kelemen and *G. Székely are with the Computer Vision Laboratory, Swiss Federal Institute of Technology, ETH-Zentrum, CH-8092 Zurich, Switzerland (e-mail: szekely@vision.ee.ethz.ch).

G. Gerig is with the Department of Computer Science and Psychiatry, University of North Carolina, Chapel Hill, NC 27599-3175 USA and with the Department of Psychiatry, University of North Carolina, Chapel Hill, NC 27599-3175 USA.

Publisher Item Identifier S 0278-0062(99)09109-0.

I. Introduction

SEGMENTATION of anatomical objects from large three-dimensional (3-D) medical data sets, obtained from routine magnetic resonance imaging (MRI) examinations, for example, represents a necessary yet difficult issue in medical image analysis. With the steady increase of routine use of 3-D imaging methods such as MRI, computer tomography (CT), and 3-D ultrasound in radiological diagnosis, monitoring, radiotherapy, and surgical planning, for example, there is a clear need for improved and efficient methods for the extraction of anatomical structures and for a description by morphometric analysis. In some limited applications, segmentation can be achieved with minimal user interaction by applying simple and efficient image processing methods, which can be applied routinely [8].

In many clinical applications, such as computer assisted neurosurgery or radiotherapy planning, a large number of organs must be identified in the radiological data sets. While a careful and time-consuming analysis may be acceptable for outlining complex pathological objects, no real justification for such a procedure can be found for the delineation of normal healthy organs at risk.

Delineation of organ boundaries is also necessary in various types of clinical studies, where the correlation between morphological changes and therapeutical actions or clinical diagnosis must be analyzed. In order to get statistically significant results, a large number of data sets must be segmented. For such applications manual segmentation becomes questionable, not only because of the amount of work, but also with regard to the poor reproducibility of the results. The necessity of obtaining high reproducibility and the need to increase efficiency motivates the development of computer-assisted automated procedures.

II. Model-Based Segmentation of 3-D Radiological Data

Elastically deformable contour and surface models, so-called snakes [9], have been proposed as tools for supporting manual object delineation. While such procedures can be extended to 3-D [4], [29], their initialization is a critical issue. Most often, the initial guess must be very close to the sought contour to guarantee a satisfying result [16]. An excellent overview of elastically deformable models can be found in [13]. The primary reason for the need of a precise snake initialization is the presence of disturbing attractors in the image. These attractors do not belong to the object contour,

but force the snake into undesired local energy minima. The procedure could become more robust if the deformation of a snake would be limited to shapes within the normal variation of a class of object boundaries.

Elastically deformable parametric models offer a straightforward way for the inclusion of prior knowledge in the image segmentation process. This is done by incorporating prior statistics to constrain the variation of the parameters of the elastic model. Such procedures have been developed by Vemuri and Radisavljevic [31], using a hybrid primitive called deformable superquadrics constructed in a multiresolution wavelet base, or by Staib and Duncan [23] for deformable Fourier models.

For complex shapes described by a large number of often highly correlated parameters, the use of such priors may become tedious. The modal analysis as proposed by Pentland and Sclaroff [17] offers a promising alternative by changing the basis from the original modeling functions to the eigenmodes of the deformation matrix. The dominant part of the deformations can thus be characterized by only a few eigenmodes, substantially reducing the dimensionality of the object descriptor space. Methods using modal analysis have been successfully applied to medical image analysis by Sclaroff and Pentland [20] and Nastar and Ayache [15], for example.

Cootes *et al.* [5] combined the power of parametric deformable shape descriptors with statistical modal analysis. They use active shape models, which restrict the possible deformations using the statistics of training samples. Object shapes are described by the point distribution model (PDM) [6], [7], which represents the object outline by a subset of boundary points. There must be a one-to-one correspondence between these points in the different outlines of the training set. After normalization to size, orientation, and position, they provide the basis for the statistical analysis of the object shape deformations. The mean point positions and their modes of variation (i.e., the eigenvectors corresponding to the largest eigenvalues of their covariance matrix) are used for limiting the object deformations to a reasonable linear subspace of the complete parameter space. Principal component analysis has also been used for the characterization of anatomical shape variability, using other shape parametrization schemes such as invariant moments [18], [19], for example.

For a large training set containing several anatomical structures, the generation of the PDM parametrization becomes very tedious and, because of the lack of a reasonable automatization, can be a source of errors, suggesting alternative approaches for form parametrization. Staib and Duncan have already demonstrated segmentation by parametrically deformable elastic models for two-dimensional (2-D) outlines [23] and 3-D object surfaces [22], [24] using Fourier descriptors. In our previous work [26] we combined the statistical modal analysis with parametrization based on 2-D Fourier-descriptors. Using spatial normalization based on the generally accepted Talairach coordinate system [28], we demonstrated that fully automatic segmentation of organ contours on 2-D image slices can be achieved. In this previous paper, the feasibility of a 3-D extension of this method has already been investigated. We have demonstrated that, based on a general surface parametrization scheme [3], the concept can be generalized for 3-D organ surfaces with spherical topology, using spherical harmonics as shape descriptors. This paper summarizes the basic concepts of the newly developed 3-D segmentation system and also presents evaluation results, using a collection of 22 volumetric MR brain data sets.

The 3-D segmentation discussed here is based on statistical shape models generated from a collection of manually segmented MR image data sets of different subjects. The process can be divided into two major phases: a model-building stage and the automatic segmentation of large series of data sets.

- In the training phase, the results of interactive segmentation of sample data sets are used to create a statistical shape model which describes the average as well as the major linear variation modes.
- The model is placed into new unknown data sets and is elastically deformed to optimally fit the measured data.

The generation of the statistical model will be discussed in detail in the following sections. The purely geometrical statistical model proposed in our earlier paper [26] has been extended by incorporating gray-valued profiles across the organ surface, implementing the concept proposed by Cootes and Taylor [5], [6] for 3-D models.

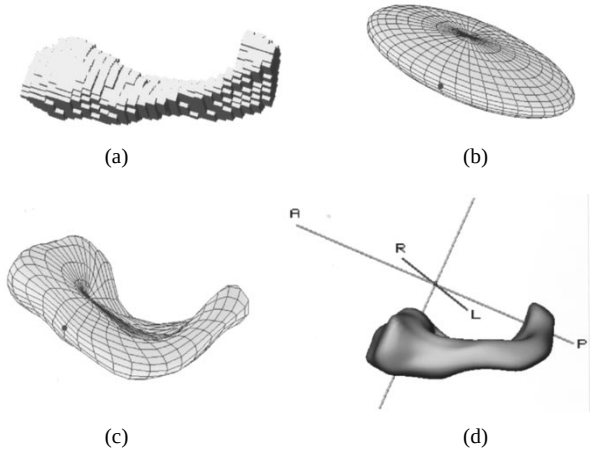
The matching process is initialized using the average geometrical model resulting from this training phase. A two-stage algorithm, described in Section VII, is used to deform this model to optimally fit the features of a new data set, while still restricting the deformations to the variability allowed by the statistical model. This algorithm makes full use of the gray-value profiles normal to the surface, which is efficiently calculated by using a dual representation of the object both as a collection of sample points and as a parametrized surface.

III. 3-D Object Model

A. Training Set

Today's routine practice for 3-D segmentation involves slice-by-slice manual processing of high-resolution volume data. Working on a large series of similar scans, human observers knowlegeable in anatomy become experts and produce reliable segmentation results, although at the cost of a considerable amount of time per data set. Realistic figures are several hours to one day per volume data set, for only a small set of structures. Regions in 2-D image slices corresponding to cross sections of 3-D objects are outlined and painted by different interactive tools, performing delineation purely manually [25] or partially supported by image data using energy minimization schemes discussed above. The series of binary regions segmented from consecutive slices form volumetric voxel objects. Fig. 1(a) illustrates the result of an expert segmentation of the left hippocampus from a magnetic resonance head dataset.

Our initial training set, consisting of 30 male brain MR volumes, a courtesy of the Harvard Medical School in Boston, has been processed this way. These datasets have been acquired and deep gray-matter structures have been processed in the



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

Fig. 1. Model building. (a) Interactive segmentation of a left hippocampus. (b) Reconstruction from surface shape descriptor up to degree one. (c) Reconstruction up to degree ten. (d) Normalization of the shape pose in object space.

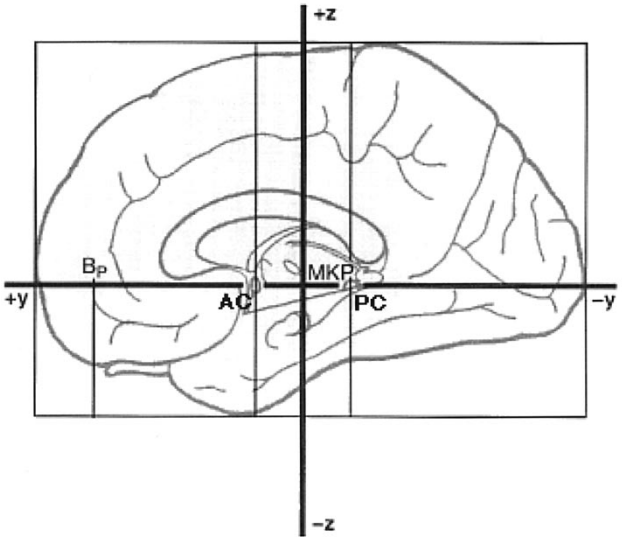


Fig. 2. Stereotactic coordinate system used for object space normalization.

frame of a comparative psychiatric study [21]. In each volume, six brain regions have been manually labeled in both hemispheres. These regions include the amygdala/hippocampus complex, the parahippocampal gyrus, thalamus, caudate nucleus, putamen, and globus pallidus. Fig. 6 compares the segmented four different brain objects for an individual case with the average models.

In order to demonstrate some model building aspects which require a larger training set, we will also refer to a set of 71 corpus callosum outlines, a courtesy of the European BIOMORPH project.

B. Parametric Shape Representation

In the following, the term parametrization will be used in two different ways. On the one hand, the process which maps two parameter values (s, t) to each point on a surface is called surface parametrization. These surface mappings parametrize an object shape with respect to surface coordinates $\boldsymbol{v}(s, t)$:

$$\boldsymbol{v}(s,t) = \begin{pmatrix} v_1(s,t) \\ v_2(s,t) \\ v_3(s,t) \end{pmatrix} = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix}. \tag{1}$$

To make a clear distinction, surface parametrization will refer to this mapping procedure. In the following, we will consider only surfaces with spherical topology, which is true for a broad class of anatomically interesting organ boundaries, in particular for all studied structures of the basal ganglia. Such surfaces can be parametrized by two polar variables (θ and ϕ) and therefore defined by three explicit functions over them

$$\boldsymbol{v}(\theta,\phi) = \begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix}. \tag{2}$$

Second, shape parametrization denotes the computation of object shape descriptors parametrizing these coordinate functions. One possibility is an expansion over a complete set of basis functions. With respect to computing elastic shape deformations, the choice of basis functions is not critical. B-splines or wavelets could be used, as well as other local representations. As discussed later, shape correspondence among multiple individuals is obtained by rotating a (θ, ϕ) parameter net over the object surface to a canonical position based on global surface parametrization. We therefore make use of the hierarchical shape representation offered by spherical harmonics, resulting in the following (truncated) series expansion:

$$\boldsymbol{v}(\theta,\phi,\boldsymbol{p}) = \sum_{k=0}^{K} \sum_{m=-k}^{k} \boldsymbol{c}_k^m Y_k^m(\theta,\phi), \tag{3}$$

where

$$\boldsymbol{c}_k^m = \begin{pmatrix} c_{x_k}^m \\ c_{y_k}^m \\ c_{z_k}^m \end{pmatrix}. \tag{4}$$

The coefficients $\boldsymbol{c}_l^m$ are 3-D vectors with components $c_{x_l}^m$, $c_{y_l}^m$, and $c_{z_l}^m$ with degree l and order m. A detailed description can be found in Brechbühler *et al.* [3]. All the $\boldsymbol{c}_l^m$ with components (x, y, z) define the shape description vector

$$\boldsymbol{p} = (c_{x_0}^0, c_{y_0}^0, c_{z_0}^0, c_{x_1}^{-1}, c_{x_1}^0, c_{x_1}^1, c_{y_1}^{-1}, c_{y_1}^0, c_{y_1}^1 \; c_{z_1}^{-1}, c_{z_1}^0, c_{z_1}^1, \cdots, c_{x_K}^{-K}, \cdots, c_{z_K}^K)^\top.$$

Fig. 1(b) and (c) illustrates the hierarchical property of spherical harmonics: reconstructing the shape from coefficients up to degree one results in an ellipsoid. Incorporating more descriptors [up to degree ten in Fig. 1(c)] increases the level of details and more closely approximates the original shape.

C. Anatomic Reference Coordinate System

Our driving application is the automatic segmentation of deep gray-matter structures of the human brain. We begin by choosing the standard stereotactic coordinate system proposed by Talairach for global alignment of the head image data sets. Basic features used for alignment are the approximation of the interhemispheric fissure by a midsagittal plane and the definition of the anterior and posterior commissure (AC-PC) (see Fig. 2). The identification of the symmetry plane of

the brain and the position of the AC-PC line is performed manually by selecting reference points on 2-D slices of the volumetric images. Each data set is transformed into canonical coordinates by 3-D rotation and scaling as illustrated in Fig. 1(d).

In comparison to a fully object-centered spatial normalization the segmentation method described in this paper can incorporate small deviations of translation and orientation into the shape statistics. This allows us to reproducibly define a global coordinate system, based on a small set of significant external landmarks for initializing shape models at their most likely positions.

In our earlier work [26] we found that in medical images there is no real justification for separating similarity transformation from shape deformation because of the strong correlation between the position and shape of organs. Accordingly, for images representing anatomy the relative position, rotation, and size of healthy organs is restricted in a similar and correlated way to their elastic deformation. We therefore introduced models incorporating this full biological variability with respect to a natural reference system, which resulted in a much more robust segmentation process in 3-D. The Talairach reference used here is a straightforward extension to our earlier AC-PC based 2-D coordinate system.

IV. Correspondence by Parametrization

During the present study we established surface correspondence between the items of the training set by an area-preserving parametrization followed by the object-oriented normalization of its starting point. This fully automatic procedure, which will be described in more detail in this section, has certain strong limitations and provides only a first step in establishing correspondence. However, preliminary 3-D studies have revealed that corrections resulting from feature-based correspondence search based on curvature [27], e.g., are minor and arc-length based parametrization provides a good first approximation.

The shape representation proposed in this paper results in a continuous mapping function between similar objects. This is done using a uniform parametrization of closed surfaces and by calculating an invariant object-centered description (Brechbühler *et al.* [2], [3]). By sampling of the spherical reference surface, this method can also be used to generate corresponding pairs of surface points.

A key step in the shape description of a surface is its mapping to the parameter space, the sphere. Any point on the surface must map to exactly one point on the sphere and vice versa. The location on the sphere corresponding to a surface point defines the surface parameters of the point. It can be represented as two polar or three Cartesian coordinates, related through the bijection

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}.$$

Mapping a surface to the sphere assigns parameters to every surface point. The mapping must be continuous, i.e., neighboring points in one space must map to neighbors in the other space. Our approach is to achieve a correspondence between different objects by constructing a mapping that preserves areas. Based on the voxel representation, such a mapping assigns the square facets on the object surface to a portion of the surface of the unit sphere. It is not possible, in general, to map every surface facet to a spherical square. Distortions cannot be avoided, but they should be minimal.

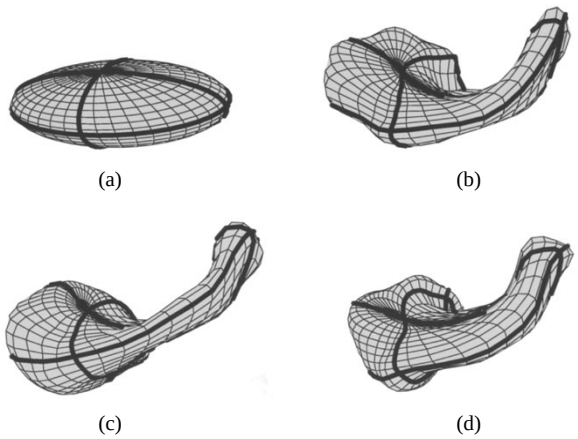


Fig. 3. (a) Corresponding parameter values for $\theta = \pi/2$, $\phi = 0, \pi$, and $\phi = \pi/2, 3\pi/2$ (thick lines) illustrated on an ellipsoid. (b)–(c) On three individual left hippocampal structures.

The surface parametrization, i.e., the embedding of the object surface graph into the surface of the unit sphere, is solved as a constrained optimization problem, looking for the optimal coordinates of all vertices [3]. However, the resulting representation of the surface by a parameter net with homogeneous cells is so far only determined up to a 3-D rotation in parameter space. Point-to-point correspondence of surfaces of different objects would require parameters which do not depend on the relative position of the parameter net. The object can be rotated to a canonical position in parameter space by making use of the hierarchical shape description provided by spherical harmonic descriptors. The coefficients of the spherical harmonic function of different degrees provide a measure of the spatial frequency constituents that compose the structure. As higher frequency components are included, more detailed features of the object appear. To define a standard position, we only consider the contribution of the spherical harmonics of degree one, which is an ellipsoid representing the coarse elongation of the object in 3-D space. We rotate the parameter space so that the north pole ($\theta = 0$) will be at one end of the shortest main axis, and the point where the zero meridian ($\phi = 0$) crosses the equator ($\theta = \pi/2$) is at one end of the longest main axis. Fig. 1(b) and (c) illustrates the location of the middle main axis on the reconstruction up to degree one and ten, respectively.

Objects of similar shape will get a standard parametrization which becomes comparable, i.e., parameter coordinates (θ, ϕ) are located in similar regions of the object shape across the set of objects (see Fig. 3). Corresponding points on different object surfaces are therefore found by calculating a canonical parametrization, rather than by interactive selection of labeled sets of 3-D points.

The normalization techniques described here require the precondition that coefficients of degree one represent a real ellipsoid. If, however, the ellipsoid degenerates to an ellipsoid of revolution or a sphere, the technique will fail to derive stable main axes. Objects of higher symmetries, such as regular polygons and polyhedra, are a good example for the limitations of the normalization technique.

V. Capturing Statistical Information of Shape

After transformation to canonical coordinates, the object descriptors are related to the same reference system and can be directly compared. An existing procedure for describing a class of objects follows our 2-D method as described in Székely *et al.* [26], where the calculations are carried out in the domain of shape descriptors, rather than the Cartesian coordinates of points in object space.

A. Principal Component Analysis of a Set of Shapes

The mean model is determined by averaging the descriptors $\boldsymbol{c}_j$ of the N individual shapes (see Fig. 4)

$$\overline{\boldsymbol{c}} = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{c}_j. \tag{5}$$

Eigenanalysis of the covariance matrix $\boldsymbol{\Sigma}$ results in eigenvalues and eigenvectors representing the significant modes of shape variation

$$\boldsymbol{\Sigma} = \frac{1}{N-1}\sum_{j} (\boldsymbol{c}_j - \overline{\boldsymbol{c}}) \cdot (\boldsymbol{c}_j - \overline{\boldsymbol{c}})^T \tag{6}$$

$$\boldsymbol{\Sigma P}_c = \boldsymbol{\Lambda P}_c \tag{7}$$

where the columns of $\boldsymbol{P}_c$ hold the eigenvectors and the diagonal matrix $\boldsymbol{\Lambda}$ the eigenvalues λ_j of Σ. Vectors $\boldsymbol{b}_j$ describe the deviation of individual shapes $\boldsymbol{c}_j$ from the mean shape, using weights in eigenvector space, and are given below

$$\boldsymbol{c}_j = \overline{\boldsymbol{c}} + \boldsymbol{P}_c \boldsymbol{b}_j. \tag{8}$$

Fig. 5 illustrates the largest two eigenmodes of the hippocampus training set, while Fig. 7 depicts the square root of eigenvalues sorted by size (dotted line) together with components of one individual vector $\boldsymbol{b}_j$. As after the first few eigenvectors the variance becomes very small, the first t largest eigenmodes have been taken for building a flexible model that explains the biological variability of the hippocampal shape. Any shape in this linear subspace is approximated by combining the mean shape and a weighted sum of the deviations obtained from the first few modes

$$\boldsymbol{c} = \overline{\boldsymbol{c}} + \boldsymbol{P}_c \boldsymbol{b} \tag{9}$$

where $\boldsymbol{b}$ is a vector of weights, one for each eigenvector and, since eigenvectors are orthogonal, $\boldsymbol{P}_c^T \boldsymbol{P}_c = \boldsymbol{I}$, $\boldsymbol{b}$ of a given shape $\boldsymbol{c}$ can be computed using

$$\boldsymbol{b} = \boldsymbol{P}_c^T (\boldsymbol{c} - \overline{\boldsymbol{c}}). \tag{10}$$

The vector $\boldsymbol{b}$ can also be thought of as a new and more compact representation of the shape in the new basis of the deformation

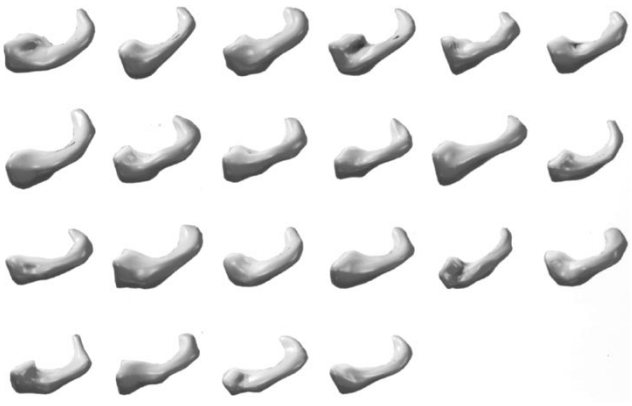


Fig. 4. Illustration of all 22 left hippocampal structures of the training sets, normalized and reconstructed from their descriptors.

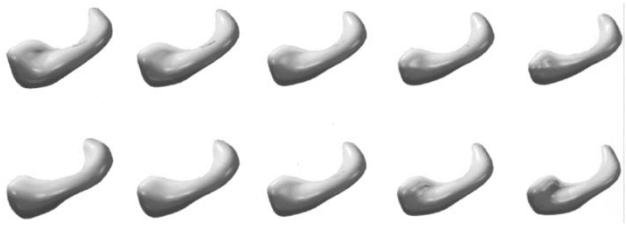


Fig. 5. Largest two modes of variation for $b_j = -2\sqrt{\lambda_j} \cdots 2\sqrt{\lambda_j}$. In the middle column, $b_j = 0$ represents the mean model.

modes, instead of the spherical harmonics. Equation (10) describes how to generate new examples of the shapes by varying the parameters $\boldsymbol{b}$ within suitable limits, so that the new shapes remain similar to those in the training set. The limits for each element of $\boldsymbol{b}$ are derived by examining the distributions of the parameter values required to generate the training set. If Gaussian distributions are assumed, the variances of the elements of $\boldsymbol{b}$ are given by the corresponding eigenvalues.

To choose the appropriate number of eigenmodes for the shape representation, the following has to be taken into consideration. Supposing Fourier harmonics up to degree n_l has been used, there will be $3n_l^2$ free parameters describing the shape. This results in a covariance matrix of the size $(3n_l^2) \times (3n_l^2)$ and theoretically in $3n_l^2$ different eigenmodes. However, if the training set only consists of $N < (3n_l^2)$ samples, there will be only $N-1$ linearly independent columns or rows in Σ and also that many eigenmodes in $\boldsymbol{P}_c$. It follows that the number of modes t should be smaller than both N and $3n_l^2$

$$t < \min(N, 3n_l^2). \tag{11}$$

One method for calculating t is to choose the smallest number of modes such that the sum of their variances explain a sufficiently large proportion of λ_T, the total variance of all the independent variables, where

$$\lambda_T = \sum_{k=1}^{\min(N, 3n_l^2)} \lambda_k \tag{12}$$

and λ_k is the kth diagonal element of $\boldsymbol{\Lambda}$.

Neglecting eigenmodes corresponding to small eigenvalues is only reasonable if the shape variation is globally distributed

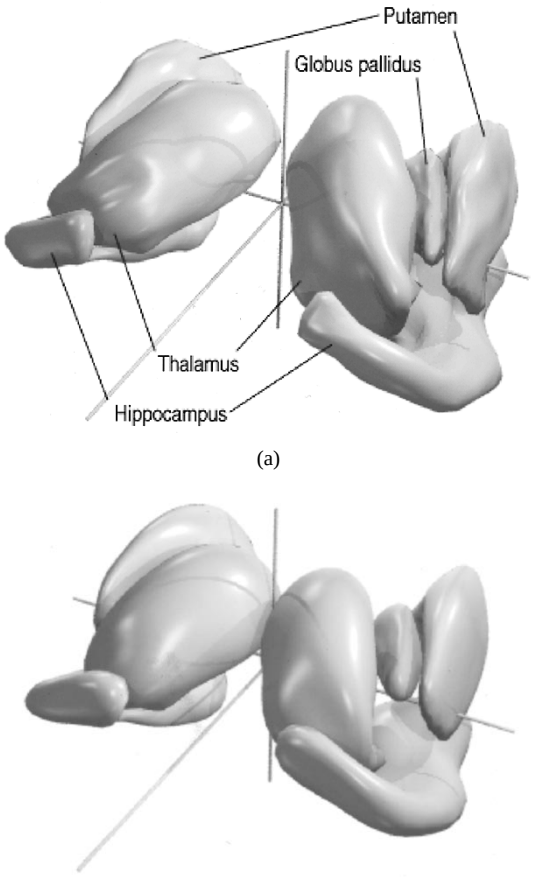


Fig. 6. (a) Left and right thalamus, globus pallidus, putamen, and hippocampus in one individual case. (b) Their average models computed from all cases.

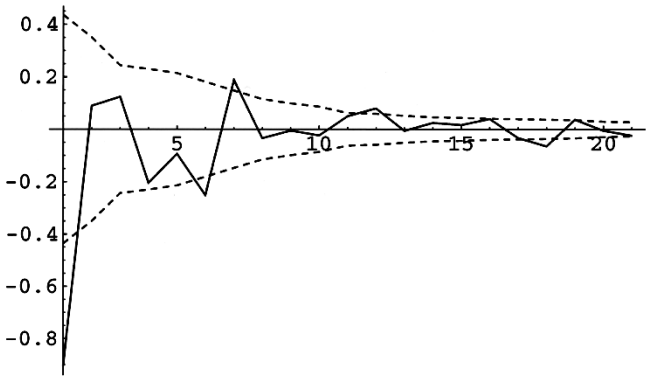


Fig. 7. Statistics of shape deformation. The dotted line represents the square root of eigenvalues $\sqrt{\lambda_j}$ sorted by decreasing size. The continuous line illustrates the components of an individual vector b_j, which describes the deviation of the shape $\boldsymbol{c}_j$ from the mean shape $\bar{\boldsymbol{c}}$.

along the whole object surface. Some diseases are thought to be correlated with very small localized anatomical differences. Eigenmodes describing highly localized (but still significant) variations should not therefore be discarded, even if the corresponding eigenvalue is small. Thus, choosing eigenmodes related to maximal local deformations results in a better representation of the training set. To investigate the influence of local deformations on the selected eigenmodes, we have resampled the surface of our models and characterized each deformation mode by the surface point with the largest deviation d_k. For our models, both sorting criteria λ_k and d_k produced identical ordering of deformation modes.

B. Validation of Statistical Models

In the example above, 22 samples of hippocampi have been used to derive a statistical model. Applying (12) we find that the ten largest eigenvalues express 99% of the variation represented in the training set. It is important to note that this does not correspond to the true anatomical variability of the organ shape, only that of the limited representation provided by the selected parametric shape descriptors. Accordingly, the parametrized 22 shapes of the training set can be described with minor error using the model, but no information is provided for shapes not included in the initial population.

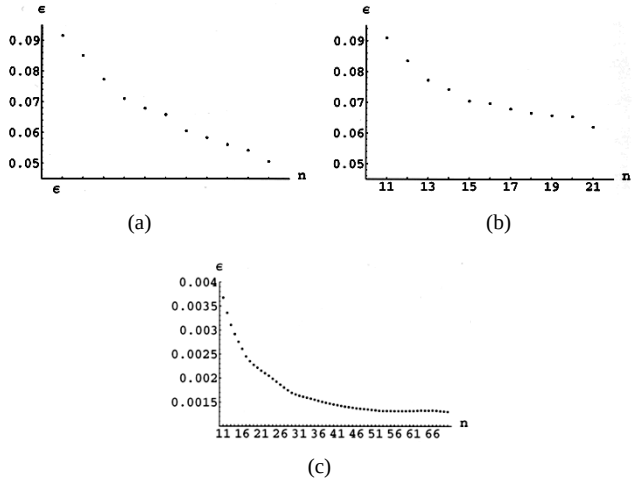
The description error of a shape not included in the training set can be computed by first projecting its descriptors $\boldsymbol{c}$ into the subspace of the major eigenvectors (10), then approximating coefficients $\tilde{\boldsymbol{c}}$ from the projection (9), and finally comparing $\boldsymbol{c}$ and $\tilde{\boldsymbol{c}}$. The difference between $\boldsymbol{c}$ and $\tilde{\boldsymbol{c}}$ is given by the Euclidean distance of the two vectors

$$\epsilon = D_{\text{Eucl}}(\boldsymbol{c}, \tilde{\boldsymbol{c}}). \tag{13}$$

It must be mentioned that the value of ϵ is an absolute measure of shape similarity in an abstract, object-dependent parameter space and, consequently, only allows comparisons between representations of the same organs.

To demonstrate the predictive ability of the statistical model, we investigated how the quality of the model increases while incorporating more individuals in the training set. We first have built a statistical model using 11 arbitrarily chosen shapes out of the entire set, as needed, to be able to compute ten deformation modes and determine the above measure of segmentation error for the remaining shapes not included in the training set. Repeating the computations for a statistically significant subset of the [22 choose 11] combinations of the entire set and finally averaging the errors, we obtain a measure for the 11-shape model, as shown in Fig. 8(a) and (b), by the first data points. The size of the training set is then increased one by one and the average error is computed, based on segmentation results using the largest ten eigenmodes to obtain the rest of the data points in the same figure. One can observe that the average error decreases as the model grows.

While this decrease of the error is significant in practically all cases, for our 3-D training sets even the insertion of the last item leads to a significant improvement of the model. As a comparison, a similar experiment for the statistical model generated from 71 2-D outlines of the corpus callosum on midsaggital brain slices has been performed. The result is shown in Fig. 8(c). One can clearly see that, in this case, the model becomes saturated and the addition of a certain number of new shapes to the training set does not increase the model's information content. On the other hand, adding new individuals to our 3-D models could further increase their ability to describe unseen shapes of the same kind. This apparent deficiency of our 3-D models also explains the



(a) (b) (c)

Fig. 8. Predictive ability of the models derived from the training sets of 22 hippocampi in the (a) left and (b) right hemisphere. (c) The same performance curve computed for the training set containing 71 corpora callosa. It displays the average representation error ϵ for shapes not included in the training set when using the ten largest eigenmodes, while increasing the number of shapes used for the generation of the statistical model from $n = 11$ to $n = N - 1$, where N is the number of available shapes for the organ in question.

difference in segmentation quality we only experience in 3-D, between shapes included and not included in the training set, as discussed later in Section VIII.

VI. Modeling Gray-Level Environment

Organ geometry represents only a part of the full information provided by the original volumetric image data sets. In addition to organ shape, radiological interpretation heavily relies on local brightness and contrast information. Previous work clearly demonstrated that augmenting geometric models with information about the gray-level environment of the model surface significantly improves the robustness of the segmentation [6]. Therefore, we examine the statistics of the image intensity along one-dimensional (1-D) profiles, orthogonal to the object surface at a discrete set of sampling points.

A. Sampling of Model Surface

Equal processing of each part of the model surface is ensured by choosing a homogeneous distribution of sampling points and profiles over the surface parameter space. A perfectly regular sampling of a spherical surface does not exist, but we can find a good approximation by icosahedron subdivision, a technique often used in computer graphics to triangulate and display spheres at different scales. The algorithm takes an icosahedron inscribed in a sphere and subdivides its faces, as shown in Fig. 9. The newly introduced vertices lie slightly inside the sphere, so we push them to the surface by properly normalizing their distance to the center to unity.

The necessary level of subdivision depends, of course, on the size of the object. For the hippocampus structures analyzed in this paper, a subdivision of $k = 10$ resulted in a sampling distance of about one voxel on the surface. This subdivision results in $n = 12 + 30(k-1) + 20((k-1)(k-2)/2) = 1002$

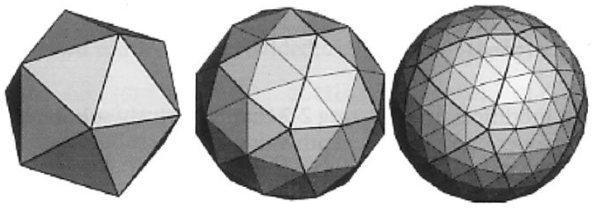


Fig. 9. Nearly regular sampling of spherical surfaces by icosahedron subdivision.

vertices. Computing the θ_i and ϕ_i values at each vertex coordinate i of the subdivided icosahedron and substituting them into

$$\boldsymbol{x}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \sum_{l=0}^{K} \sum_{m=-l}^{l} \boldsymbol{c}_l^m Y_l^m(\theta_i, \phi_i), \qquad i = 1 \cdots 1002 \tag{14}$$

we obtain a dual description of the object surface by the coordinates of a set of surface points $\boldsymbol{x}_i$. The equation above can be written in a more compact matrix form as

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{c} \tag{15}$$

where $\boldsymbol{x}$ represents the coordinates in object space and $\boldsymbol{c}$ the spherical harmonics descriptors. $\boldsymbol{A}$ consists of the function values of $Y_l^m(\theta_i, \phi_i)$, one for each dimension, and describes the mapping between shape description space and object space coordinates.

For every surface point i in each data set j, we can extract a profile $\boldsymbol{w}_{ij}$ of n_p sample points. The distance between sample points is the length of one voxel. The profiles are oriented normally to the object surface and centered at the surface points $\boldsymbol{x}_{ij}$, as illustrated in Fig. 10. For each sample point i we can obtain a mean profile by averaging over the sample objects N

$$\overline{\boldsymbol{w}}_i = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{w}_{ij}. \tag{16}$$

We calculate a $n_p \times n_p$ covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{w}_i}$, which gives us a statistical description of the expected profiles at each sample point

$$\boldsymbol{\Sigma}_{\boldsymbol{w}_i} = \frac{1}{N-1} \sum_{j=1}^{N} (\boldsymbol{w}_{ij} - \overline{\boldsymbol{w}}_i)(\boldsymbol{w}_{ij} - \overline{\boldsymbol{w}}_i)^T. \tag{17}$$

Cootes *et al.* in [6] propose normalized derivative profiles giving invariance to uniform scaling of gray levels and constant shift. For our applications, however, we achieved best results using unnormalized original gray-level profiles, as all our data sets have been acquired under the same imaging conditions. This allows us to avoid the information loss caused by any normalization procedure.

B. Dual Surface Representation

The points of the sampled surface can be considered as a new representation of the same object, which can be obtained

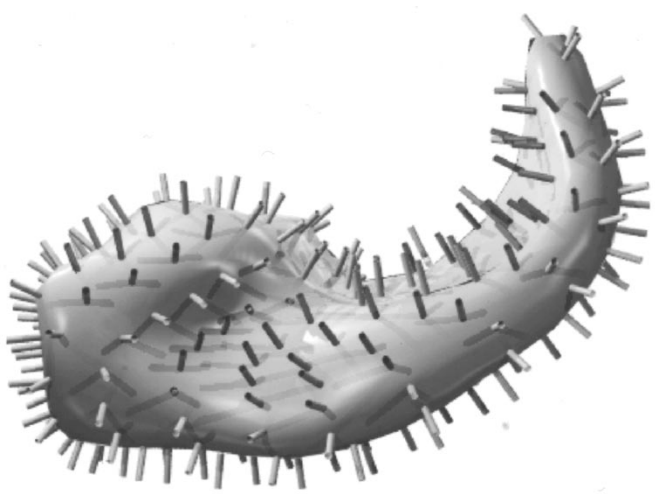


Fig. 10. Illustration of an individual left hippocampal shape with its profile vectors shown from the left side of the brain.

from the spherical harmonic descriptors by the linear transformation described by (15). The deformation modes have been previously derived, based on the spherical harmonic coefficients (referred as parameter space). To examine how the modes can be converted to the description of the surface based on sample points (referred as object space), we investigate the covariance matrix. In parameter space it is defined by

$$\boldsymbol{\Sigma}_c = \mathrm{Cov}[d\boldsymbol{c}] = \mathrm{E}[d\boldsymbol{c}\, d\boldsymbol{c}^T]. \tag{18}$$

Writing the same equation in object space and substituting the transformation matrix $\boldsymbol{A}$, we obtain

$$\boldsymbol{\Sigma}_x = \mathrm{Cov}[d\boldsymbol{x}] = \mathrm{E}[d\boldsymbol{x}\, d\boldsymbol{x}^T] = \mathrm{E}[\boldsymbol{A} d\boldsymbol{c}\, d\boldsymbol{c}^T \boldsymbol{A}^T] = \boldsymbol{A}\boldsymbol{\Sigma}_c \boldsymbol{A}^T \tag{19}$$

where $d\boldsymbol{x}$ denotes the deviation of an individual $\boldsymbol{x}$ from the average $\overline{\boldsymbol{x}}$ over the whole population. Performing principal component analysis on $\boldsymbol{\Sigma}_x$

$$\boldsymbol{\Sigma}_x \boldsymbol{P}_x = \boldsymbol{\Lambda}_x \boldsymbol{P}_x \tag{20}$$

and substituting (15) and (19) into (20) and multiplying both sides by $\boldsymbol{A}^T$, we obtain

$$\boldsymbol{A}^T \boldsymbol{A}\boldsymbol{\Sigma}_c \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{P}_c = \boldsymbol{\Lambda}_x \boldsymbol{A}^T \boldsymbol{A}\boldsymbol{P}_c. \tag{21}$$

Comparing (21) with (7), it can be seen that $\boldsymbol{P}_c$ and $\boldsymbol{P}_x$ describe the same deformation modes if and only if $\boldsymbol{A}^T\boldsymbol{A} = \alpha\boldsymbol{I}$, where α is a scalar and $\boldsymbol{I}$ the identity matrix. This requirement is fulfilled if $\boldsymbol{A}$ is an orthogonal matrix. Furthermore, if $\boldsymbol{A}$ is orthonormal, then $\alpha = 1$.

In the 2-D case, columns of $\boldsymbol{A}$ are regularly sampled versions of $\cos(kt)$ and $\sin(kt)$ functions (with $k = 0 \cdots K$, and $0 \leq t < 2\pi$) which are known to be orthogonal since they also build the orthogonal basis vectors of the discrete Fourier transformation. It follows that eigenvectors $\boldsymbol{P}_x$ and eigenvalues $\boldsymbol{\Lambda}_x$ in object space can be easily computed from those defined in parameter space using the following equations:

$$\boldsymbol{P}_x = \boldsymbol{A}\boldsymbol{P}_c \tag{22}$$

$$\boldsymbol{\Lambda}_x = \alpha\boldsymbol{\Lambda}_c. \tag{23}$$

$\boldsymbol{P}_x$ in this case describes real eigenmodes. Thus, the obtained statistical shape representation is identical to point distribution models introduced by Cootes *et al.* [6] with the exception that in our case point-to-point correspondence has been automatically approximated, rather than manually determined.

Contrary to the special 2-D case, in 3-D matrix $\boldsymbol{A}$ is nonorthogonal because of the not perfectly regular sampling of the spherical surface. In our application, it is more beneficial that corresponding deformation modes describe the same alteration in object and parameter space than their exact orthogonality, hence, we introduce quasi-eigenvectors $\boldsymbol{P}'_x$.

The shape statistics, as described in Section V, can be expressed by (9). Multiplying both sides of this equation by $\boldsymbol{A}$ we get the dual surface description by a set of surface points

$$\boldsymbol{A}\boldsymbol{c} = \boldsymbol{A}\overline{\boldsymbol{c}} + \boldsymbol{A}\boldsymbol{P}_c\boldsymbol{b} \tag{24}$$

$$\boldsymbol{x} = \overline{\boldsymbol{x}} + \boldsymbol{P}'_x\boldsymbol{b} \tag{25}$$

where $\boldsymbol{P}'_x$ denotes the product $\boldsymbol{A}\boldsymbol{P}_c$ which represents the matrix of modes of shape variation expressed in object coordinate space. Recall that $\boldsymbol{P}_c$ is the matrix of eigenvectors in the shape descriptor space, defined by the components of the elliptic harmonic descriptors $\boldsymbol{c}$. Thus, $\boldsymbol{P}'_x$ is not a real matrix of eigenvectors since its column vectors are also nonorthogonal, although they still represent the same shape deformations as eigenvectors in $\boldsymbol{P}_c$. This deviation from orthogonality can be characterized by the ratio between the average of the nondiagonal and diagonal elements of $\boldsymbol{P}'_x\boldsymbol{P}'^T_x$, which was 15.2626 in our case. Therefore, weight vectors $\boldsymbol{b}_i$ of individual shapes, which express the deviation from the mean model, remain the same in both shape representation schemes.

VII. Iterative Segmentation Scheme

Until now we have only described the creation of a flexible 3-D model, including geometric shape, gray-level environment and statistics about normal shape variability. We now perform the segmentation step by elastically fitting this model to new 3-D data sets. This is achieved with the following two steps,

- Initialization is done by transforming the model's coordinate system into that of the new data set.
- The elastic deformation of the surface until it best matches the new gray-value environment.

A. Initialization of Segmentation

Since the model has been built based on a normalization to the Talairach coordinate system, the determination of the symmetry plane of the brain, and the position of the AC/PC line becomes an integral part of the initialization. Currently this is done manually, but the determination of the symmetry plane and the AC/PC line by can be replaced in the future by an automatic method [11], [12], [14], [30]. To derive the position of the midsagittal plane, the user specifies the position of three or more points lying on the interhemispheric plane. The program computes parameters of a plane which have the best least squares fit to the given points. The more points are specified, the more robust is the fit. Having determined the

plane, the user finally marks the location of the AC and PC. A translation vector and a rotation matrix are computed to transform the model's coordinate system into the image space of the new data set.

B. Elastic Deformation of Model Shape

We introduced two different representations of a surface, one based on the spherical harmonic descriptors and a second one based on the subdivided icosahedron. We attempt to use the advantages of both representations in our procedure. Spherical harmonic descriptors were necessary to find a correspondence between similar surfaces and they also allow the exact analytical computation of surface normals by

$$\boldsymbol{n}_i = \sum_{l=0}^{K} \sum_{m=-l}^{l} \boldsymbol{c}_l^m \frac{\partial Y_l^m}{\partial \theta} \times \sum_{l=0}^{K} \sum_{m=-l}^{l} \boldsymbol{c}_l^m \frac{\partial Y_l^m}{\partial \phi}. \tag{26}$$

However, they only represent a global description of an object shape. The surface points, on the other hand, give a local representation, which is essential to carry out an iterative refinement of the model, as will be described in the next section. Thus, we decided to keep both representations during the matching process, the relation between the two being tractable via the matrix $\boldsymbol{A}$.

C. Calculating Displacements for Surface Points

After initialization of the surface model, we calculate the displacement vector at each surface sample point which would move that point to a new position closer to the sought object. Since there is a model of a gray-level profile for each point, the search tries to find an adjacent region which better matches this profile. A profile $\boldsymbol{w}$ of length l $(>n_p)$ normal to the surface is extracted at each model point. This new profile is shifted along the model profile in discrete steps s to find the position of the best match. This is given as the square of the Mahalanobis distance

$$d^2_{\text{Maha}}(s) = (\boldsymbol{w}(s) - \overline{\boldsymbol{w}})\boldsymbol{\Sigma}_w^{-1}(\boldsymbol{w}(s) - \overline{\boldsymbol{w}})T \tag{27}$$

where $\boldsymbol{w}(s)$ represents the subinterval of the extracted profile at step s having a length of n_p. The location of the best fit is thus the one with minimal $d^2_{\text{Maha}}(s)$. Suppose s_{best} is the shift between the two profiles providing the best fit. We choose a displacement vector $d\boldsymbol{x}$ for each model point which is parallel to the profile in the direction of the best fit and has magnitude s_{best}. Fig. 11 illustrates this process.

D. Adjusting Shape Parameters

Having generated 3-D displacement vectors for each of the n model points

$$d\boldsymbol{x} = (dx_1, dy_1, dz_1, \cdots, dz_n) \tag{28}$$

we then adjust the shape parameters to move the model surface toward a new position. Since rotation, translation, and scale are already incorporated in the model statistics, we do not have to deal with them separately. Of more concern are the calculated displacements $d\boldsymbol{x}$, as these could freely deform the shape of the object. In order to keep their resulting shape consistent with the statistical model, we restrict possible deformations by considering only the first few modes of variation. This will be solved by minimizing a sum of squares of differences between actual model point locations and the suggested new positions.

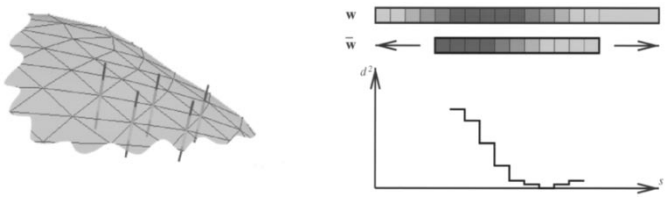


Fig. 11. Illustration of the surface matching process. (a) Part of the model's triangulated surface with longer extracted (in black) and shorter model profiles (in gray). (b) The computation of a suggested movement for a single surface point.

The shape statistics, as described by (25), represent the matrix of modes of shape variation expressed in object coordinate space. We seek an adjustment $d\boldsymbol{b}$ to $\boldsymbol{b}$, which causes a deformation in eigenspace which matches the optimal deformation $\boldsymbol{x}$ as closely as possible

$$(\boldsymbol{x} + d\boldsymbol{x}) = \overline{\boldsymbol{x}} + \boldsymbol{P}'_x(\boldsymbol{b} + d\boldsymbol{b}). \tag{29}$$

Subtracting (25) from (29) we get

$$d\boldsymbol{x} = \boldsymbol{P}'_x\, d\boldsymbol{b}. \tag{30}$$

This is an over-determined set of linear equations where the number of equations $(3n)$ is much larger than the number of variables (the number of modes is usually restricted from around five to ten). Therefore, a least squares approximation to the solution can be obtained using standard methods of linear algebra. Because of the orthogonality of $\boldsymbol{P}_x$ in 2-D, the least squares solution could be obtained by $d\boldsymbol{b} = \boldsymbol{P}_x^T d\boldsymbol{x}$. In 3-D, the nonorthogonality of $\boldsymbol{P}'_x$ does not allow solving (30) this simple way. With this object, the general purpose least squares routine F04JAF from the NAG Fortran library has been applied to obtain $d\boldsymbol{b}$.

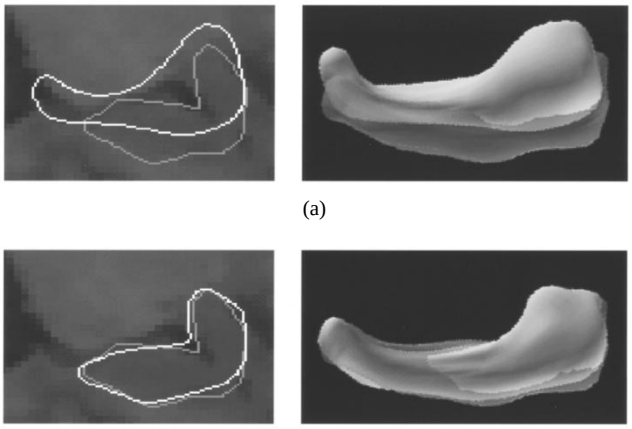
The entire procedure is repeated iteratively and starts with the average model such that $\boldsymbol{b}_{t=0} = \boldsymbol{0}$. At each iteration step, we compute a new set of displacements from the match of profiles and update the shape deviation vector $\boldsymbol{b}$ until the variation of the shape remains below a threshold value for a certain number of iterations.

E. Shape Constraints

There are two different kind of constraints we apply to keep the resulting shape consistent with the shape model. On the one hand, there is a limited number of eigenmodes due to the small number of individuals and the restriction of the number of modes. On the other hand, after the weights have been updated by

$$\boldsymbol{b}_{t+1} = \boldsymbol{b}_t + d\boldsymbol{b}_t \tag{31}$$

we constrain the components b_i of $\boldsymbol{b}$, using the standard deviation defined by the statistical model which is given by the eigenvalues $\sqrt{\lambda_i}$ (see Fig. 7). Thus, each component of



(a) (b)

Fig. 12. Segmentation result of a left hippocampus on sagittal 3-D slices and 3-D views from the left hand side. (a) Image has been taken after initialization. (b) Image illustrates the final result after about 100 iterations.

$\boldsymbol{b}_{i,\,t+1}$ lying outside of the interval $\pm a\sqrt{\lambda_i}$ will be truncated where the constant a is set to 2.

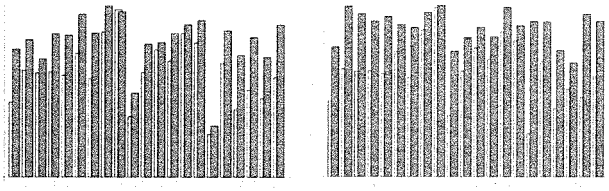
VIII. Results and Validation of Segmentation

Fig. 12(a) shows the initial placement of the left hippocampus model (white line) together with the hand-segmented contour (gray line) on a sagittal 3-D slice and as a 3-D scene viewed from the right side of the head. After about 100 iterations, the model gives a sufficiently close fit to the data. The model used in this example had five degrees of freedom, and model profiles had a total length of 11 sample points, while the extracted profiles a length of 19 sample points. The whole segmentation process takes about 2 min on a Sun Ultra 1 workstation and runs fully automatically after initializing the model with a new data set.

The above procedure has been applied to all 21 data sets where the hippocampus had been manually segmented. To make optimal use of the relatively small data set, 21 models have been built, leaving out one shape each time and applying the technique to this specific shape. The performance of the automatic segmentation has been tested by comparisons with manually segmented object shapes which were used as a widely accepted standard, given the lack of ground truth. $\boldsymbol{A}$ represents the model shape obtained by human experts, $\boldsymbol{B}$ the result of the new model-based segmentation.

The overlap measure $(\boldsymbol{A}\cap\boldsymbol{B})/(\boldsymbol{A}\cup\boldsymbol{B})$ shown in Fig. 13(a) is calculated on binary voxel maps created by intersection of the object surfaces with the voxel grid. We therefore avoid the discretization errors by projecting the surfaces back to a voxel grid. The resulting measure is very sensitive to even small differences in overlap, both inside and outside of the object model, and is therefore a strong test for segmentation accuracy. For example, two voxel cubes of a volume of $10 \times 10 \times 10$ shifted by one voxel along the space diagonal direction results in only a 57% overlap $(729/1271)$, although the mean distance of surfaces is roughly one voxel.

The calculation of the mean distance of surfaces can be determined in an elegant way directly from the coefficients of the spherical harmonic expansion using Parseval's theorem



(a) (b)

Fig. 13. Overlap measure $(\boldsymbol{A}\cap\boldsymbol{B})/(A\cup B)$ in percentage calculated between manually and automatic segmented left hippocampi of 21 individuals. Bars in light gray illustrate the measure at initialization and in dark gray after deformation. In image (a) segmentations have been carried out with the leave-one-out method while in (b), all shapes have been included to build the model.

relating the energy of the 1-D continuous signal $f(t)$ to its Fourier coefficients (a_n, b_n):

$$\frac{1}{T}\int_{-T}^{T}[f(t)]^2\,dt = \frac{a_0^2}{2} + \sum_{1}^{\infty}(a_n^2 + b_n^2). \tag{32}$$

The equation also applies to other orthogonal basis functions, such as spherical harmonics, and to higher dimensions as well. This way, the average distance of a closed surface $\boldsymbol{x}(\boldsymbol{u})$ from the coordinate origin can be described as

$$\oint ||\boldsymbol{x}(\boldsymbol{u})||^2\,d\boldsymbol{u} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}|\boldsymbol{c}_l^m|^2 = 4\pi\cdot\text{MSD} \tag{33}$$

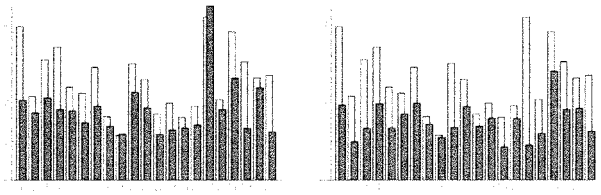
where MSD stands for mean squared distance measured from the origin of the coordinate system. Similarly, the average distance between two surfaces given by $\boldsymbol{x}_1(\boldsymbol{u})$ and $\boldsymbol{x}_2(\boldsymbol{u})$ or $\boldsymbol{c}_{l1}^m$ and $\boldsymbol{c}_{l2}^m$ can be written as

$$\oint ||\boldsymbol{x}_1(\boldsymbol{u}) - \boldsymbol{x}_2(\boldsymbol{u})||^2\,d\boldsymbol{u} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l}|\boldsymbol{c}_{l1}^m - \boldsymbol{c}_{l2}^m|^2 \tag{34}$$

providing an elegant way to calculate an error measure based on average surface distance from the spherical harmonic coefficients of the model and the segmentation result.

Fig. 14(a) nicely illustrates how the mean distance of surfaces is reduced by the iterative elastic deformation of the model. Again, we take the human expert's segmentation as ground truth and compare its surface with the result of the automatic segmentation. The bars in light gray illustrate the mean distance of the initialization of the model in a new data set and the dark bars, the final mean distance of surfaces to the model surface. The horizontal axis lists the series of 21 normal controls and schizophrenics that were used in this study.

To illustrate the performance drop caused by using statistically suboptimal models obtained from a training set too small to represent the entire class of shapes, we also show segmentation results in Fig. 13(b) and Fig. 14(b) which have been computed with a model including all 21 shapes. Experiments with the statistically optimal 2-D model of the corpus callosum showed that including a shape into a saturated model does not significantly influence the quality of the segmentation. However, the restricted size of our training set did not allow us to generate such saturated shape models, even after including all available samples, leading to a slight



(a) (b)

Fig. 14. Average distances in millimeters, calculated between manually and automatically segmented left hippocampi of 21 individuals. The bars in light gray illustrate the mean distance of the initialization of the model in a new data set and the dark bars the final mean distance of surfaces to the model surface. The length of the hippocampus is about 40 mm. (a) Segmentations have been carried out with the leave-one-out method. (b) All shapes have been included to build the model.

degradation of the segmentation results. In other words, our statistical 3-D models does not contain enough information to represent all possible shapes of a certain organ calling for the compilation of a larger training set.

IX. Conclusions

We present a new model-based 3-D segmentation technique that provides automatic segmentation of objects from volumetric image data. Tests with a large series of volumetric image data taken from different patient studies demonstrated that the method was robust and provides reproducible results.

The new technique uses elastic deformation of surface models, which carry statistical information of normal geometric shape variation and statistics about gray levels near the object surface. Our models has been derived from a series of interactively segmented training data set. Thereby, the model represents a real anatomical shape rather than a simple geometric 3-D figure, as obtained by CAD modeling, for example. Furthermore, information about the statistics of a normal shape deformation helps to constrain the elastic deformations. This is an important advantage since 3-D snake and balloon techniques are known to be prone to elastically deform to any smooth object shape and to be trapped by disturbing attractors not part of the sought shape.

Our approach has been significantly influenced by the research work of Cootes, Taylor *et al.* [5], [6]. However, the extension of their original 2-D method to a true 3-D volumetric segmentation technique required various new solutions to single steps of the procedure.

1) Parametric Shape Representation: The most prominent distinction from [6] is that we use a parametric 3-D object shape representation, rather than a point distribution model, and that shape statistics are calculated in the space of these shape parameters, rather than point coordinates.

2) Statistical Shape Models: To overcome the problem of getting a reproducible interactive definition of a set of homologous points in 3-D space, the approach presented herein proposes an automatic definition of surface meshes with homogeneous distribution of nodes defined in a standard canonical position.

3) Object Alignment: We define the position and orientation of objects in a global coordinate system which is defined by the type of application. Small translations and rotations of objects with respect to this coordinate system are part of the statistical model. Therefore, we do not separate a similarity transform for alignment and an elastic transform for remaining shape deformations as in [5].

4) Dual Shape Representations: Our approach makes use of two shape representations which are used in a vice versa fashion, taking advantage of shape descriptors holding a compact global object characterization and of a set of surface points giving access to local shape properties.

Similar to the experience of Cootes *et al.* [5], we too found that the modeling of gray-level information near the object boundaries provides valuable additional information for a model placement and improves the robustness and stability of the iterative optimization scheme. An early version of our segmentation [26] used an energy minimization concept similar to standard snake techniques. This method was very sensitive to the quality of the initialization and prone to be trapped by local energy minima. The additional use of gray-level profile information represents a strong restriction to the number of possible solutions and was demonstrated to be robust, even in the presence of considerable mismatch between initialization and a new object.

Validation has been done by defining shape distance metrics and comparing the results of interactive outlining by experts, which is a common gold standard for comparisons, with the shapes obtained by model-based segmentation (see Fig. 14).

We noticed that the convergence is faster if only a small number of modes (usually five) are involved, while a larger number of modes (usually ten) is required to find the exact contour. Thus, we plan to apply a relaxation method which gradually increases the number of modes. The convergence criteria is set by the size of the deformation of a surface.

The fundamental difficulty of the application of parametric statistical models for 3-D organ segmentation remains the efficient establishment of correspondence between the single objects of the training set. This is a major research area at the moment and different approaches are under investigation [1], [10], [27]. The method proposed by Kotcheff [10] is of particular interest, as it addresses not only the problem of correspondence but, at the same time, the question of the underlying distribution model for the shape parameters under investigation. This is another basic matter of concern if principle component analysis applied. Using PCA implies a Gaussian noise model on the parametrization, which cannot be expected in general. The idea of using reparametrization for correspondence establishment in order to change the underlying parameter error distribution would offer an fundamentally new way for estimating correspondence in a whole shape population and is currently investigated. As an alternative, independent component analysis, looking for higher order correlations in the data, can also be applied to cope with the problem of non-Gaussian distribution of the analyzed shape parameters. The application of more sophisticated methods will, however, be limited by the relatively small number of individual samples, as compared to the degrees of freedom of the model.

The set of statistical models and the automatic and efficient segmentation technique (only a few minutes per data set) open

new possibilities for the processing of a large number of data sets as they are collected in clinical studies, for example, in schizophrenia research. This will provide new statistical models with increased number of samples for normal controls and for different patient categories. These statistical models represent the first step in building an anatomical atlas based on a set of surfaces of anatomical shapes. Whereas the current segmentation technique would segment a series of objects independently, a future development could provide a combined modeling of several anatomical structures. The representation of anatomical objects by normalized shape descriptors further exploits its access to morphometric parameters. After segmenting a new set of image data, morphological properties of objects are available for comparative studies.

ACKNOWLEDGMENT

The authors thank R. Kikinis and M. Shenton, Brigham and Women's Hospital, Harvard Medical School, Boston, for kindly provided the original MR and segmented data sets. The authors are thankful to C. Brechbühler of the Image Science Laboratory for providing the software for surface parametrization and shape description. They further acknowledge MR datasets provided by the European BIOMORPH consortium (BIOMED 2 95-0845). Finally, the authors would like to thank the anonymous reviewers for their comments and suggestions.

REFERENCES

[1] F. L. Bookstein, "Shape and the information in medical images: A decade of the morphometric synthesis," *Comp. Vision Image Understanding,* vol. 66, no. 2, pp. 97–118, May 1997.

[2] C. Brechbühler, G. Gerig, and O. Kübler, "Surface parametrization and shape description," *Visualization Biomed. Comput.,* pp. 80–89, 1992.

[3] ——, "Parametrization of closed surfaces for 3D shape description," *CVGIP: Image Understanding,* vol. 61, pp. 154–170, 1995.

[4] I. Cohen, L. D. Cohen, and N. Ayache, "Using deformable surfaces to segment 3-D images and infer differential structures," *CVGIP: Image Understanding,* vol. 56, no. 2, pp. 242–263, 1992.

[5] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," *Image Vision Comput.,* vol. 12, no. 6, pp. 355–366, 1994. Available: http://s10d.smb.man.ac.uk/publications/index.htm

[6] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models—Their training and application," *Comput. Vision Image Understanding,* vol. 61, no. 1, pp. 38–59, 1995.

[7] T. F. Cootes and C. J. Taylor, "Active shape models—Smart snakes," in *Proc. British Machine Vision Conf.* Berlin, Germany: Springer-Verlag, 1992, pp. 266–275.

[8] G. Gerig, J. Martin, R. Kikinis, O. Kübler, M. Shenton, and F. Jolesz, "Automatic segmentation of dual-echo MR head data," in *Proc. IPMI'91,* Wye, U.K., 1991, pp. 175–187.

[9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comp. Vision,* vol. 1, no. 4, pp. 321–331, 1988.

[10] A. C. W. Kotcheff and C. J. Taylor, "Automatic construction of eigenshape models by genetic algorithm," in *Information Processing in Medical Imaging.* Berlin, Germany: Springer-Verlag, 1997, pp. 1–14.

[11] F. Kruggel and G. Lohmann, "Automatical adaption of the steretactical coordinate system in brain MRI datasets," in *Information Processing in Medical Imaging.* Berlin, Germany: Springer-Verlag, 1997, pp. 471–476.

[12] P. Marais, R. Guillemaud, M. Sakuma, A. Zisserman, and M. Brady, "Visual cerebral asymmetry," *Visualization Biomed. Comput.,* pp. 411–416, 1996.

[13] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Med. Image Anal.,* vol. 1, no. 2, pp. 91–108, 1996.

[14] S. Minoshima, R. A. Koeppe, M. A. Mintun, K. L. Berger, S. F. Taylor, K. A. Frey, and D. E. Kuhl, "Automated detection of the intercommissural line for stereotactic localization of functional brain images," *J. Nucl. Med.,* vol. 34, no. 2, pp. 322–329, 1993.

[15] Ch. Nastar and N. Ayache, "Frequency-based nonrigid motion analysis: Application to four dimensional medical images," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 18, pp. 1067–1079, Nov. 1996.

[16] W. Neuenschwander, P. Fua, G. Székely, and O. Kübler, "Initializing snakes," in *Proc. CVPR'94,* 1994, pp. 658–663.

[17] A. Pentland and A. P. Sclaroff, "Closed-form solutions for physically based shape modeling and recognition," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 13, pp. 715–729, July, 1991.

[18] F. Poupon, J.-F. Mangin, V. Frouin, and I. Mangin, "3D multi-object deformable templates based on moment invariants," in *Proc. Xth Scandinavian Conf. Image Analysis,* Lappeenranta, Finland, June 1997, pp. 149–156.

[19] F. Poupon, J.-F. Mangin, D. Hasboun, C. Poupon, V. Frouin, and I. Mangin, "Multi-object deformable templates dedicated to the segmentation of brain deep structures," in *Proc. 1st Int. Conf. Medical Image Computing Computer Assisted Intervention* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 1998, vol. 1496, pp. 1134–1143.

[20] S. Sclaroff and A. P. Pentland, "On modal modeling for medical images: Underconstrained shape description and data compression," in *Proc. IEEE Workshop Biomedical Image Analysis,* Seattle, WA, 1994, pp. 70–79.

[21] M. Shenton, "Abnormalities of the temporal lobe and thought disorders in schizofrenia. A quantitative magnetic resonance imaging study," *New England J. Med.,* vol. 327, pp. 604–612, 1992.

[22] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 14, pp. 1061–1075, Nov. 1992.

[23] L. H. Staib and J. S. Duncan, "Deformable Fourier models for surface finding in 3D images," in *Proc. VBC'92,* 1992, pp. 90–194.

[24] ——, "Model-based deformable surface finding for medical images," *IEEE Trans. Med. Imag.,* vol. 15, pp. 1–12, May, 1996.

[25] G. Székely, M. Bajka, J. Hug, M. Manestar, P. Groscurth, and U. Haller, "Anatomical model generation for laparoscopic surgery simulation," in *2nd Visible Human Project Conf. Proc.* (Abstracts of the National Library of Medicine), pp. 45–46, Oct. 1998.

[26] G. Székely, A. Kelemen, Ch. Brechbühler, and G. Gerig, "Segmentation of 2D and 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models," *Med. Image Anal.,* vol. 1, no. 1, pp. 19–34, 1996.

[27] H. D. Tagare, "Non-rigid curve correspondence for estimating heart motion," in *Information Processing in Medical Imaging.* Berlin, Germany: Springer-Verlag, 1997, pp. 489–494.

[28] J. Talairach and P. Tournoux, *Co-Planar Stereotaxic Atlas of the Human Brain.* Stuttgart, Germany: Thieme, 1988.

[29] D. Terzopoulos, A. Witkin, and M. Kass, "Symmetry-deeking models and 3-D object reconstruction," *Int. J. Comp. Vision,* vol. 1, no. 3, pp. 211–221, 1988.

[30] J. Ph. Thirion, S. Prima, and G. Subsol, "Statistical analysis of dissymmetry in volumetric medical images," INRIA, Tech. Rep. 3178, 1997. Available: http://www.inria.fr/RRRT/RR-3178.html

[31] B. C. Vemuri and A. Radisavljevic, "Multiresolution stochastic hybrid shape models with fractal priors," *ACM Trans. Graph.,* vol. 13, no. 2, pp. 177–200, 1994.