

Three-dimensional Model-based Segmentation of Brain MRI

András Kelemen, Gábor Székely, Guido Gerig
Swiss Federal Institute of Technology
Communication Technology Laboratory
ETH-Zentrum, CH-8092 Zurich, Switzerland
akelemen@vision.ee.ethz.ch

Abstract

This paper presents a new technique for the automatic model-based segmentation of 3-D objects from volumetric image data. The development closely follows the seminal work of Cootes et al. [5] but presents various new solutions to come up with a true 3-D technique rather than a slice-by-slice 2-D processing. 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 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 a parametric surface net which is normalized to get an invariant object-centered coordinate system. Surface descriptions are expanded into series of spherical harmonics which provide parametric representations of object shapes. Gray-level information is represented by 1-D profiles normal to the surface. The alignment is based on the well-accepted stereotactic coordinate system since the driving application is the segmentation of brain objects. Shape statistics are calculated from the parametric shape representations rather than from the spatial coordinates of sets of points. After initializing the mean shape in a new data set on the basis of the alignment coordinates, the model elastically deforms in accordance to displacement forces across the surface but is restricted only by shape deformation constraints. The technique has been applied to segment left and right hippocampal structures from a large series of 3-D magnetic resonance scans taken from a schizophrenia study.

1 Introduction

Segmentation of anatomical objects from large 3D medical data sets, obtained from routine Magnetic Resonance Imaging (MRI) examinations, for example, represents a necessary yet difficult issue in medical image 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 [7].

In many clinical applications such as computer assisted neurosurgery or radiotherapy planning, a large number of organs have to 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 has to be analyzed. In order to get statistically significant results, a large number of data set has to 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.

Elastically deformable contour and surface models, so-called snakes [8], have been proposed as tools for supporting manual object delineation. While such procedures can be extended to 3-D [24, 3], their initialization is a critical issue. Most often, the initial guess must be very close to the sought contour to guarantee a satisfying result [14]. An excellent overview of elastically deformable models can be found in [11]. The primary reason for the need of a precise snake initialization is the presence of disturbing attractors in the image. These do not belong to the desired object contour but force the snake into local energy minima. If the deformation of a snake could be limited to shapes within the normal variation of a class of object boundaries, the procedure would be more robust.

Elastically deformable parametric models offer a straightforward way for the inclusion of prior knowledge in the image segmentation process. This is done by incor-

porating prior statistics to constrain the variation of the parameters of the elastic model. Such procedures have been developed by Vemuri and Radisavljevic [26] using a hybrid primitive called deformable superquadrics constructed in a multi-resolution wavelet base, or by Staib and Duncan [18] 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 [15] 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 the smallest 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 [17] and Nastar and Ayache [13], 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) [4, 6], 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.

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 2-D outlines [18] and 3-D object surfaces [19, 20] using Fourier-descriptors. In our previous work [21] 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 [23] 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 [2] 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.

2 Segmentation concept

The 3-D segmentation discussed here is based on a statistical model, 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 section. The purely geometrical statistical model proposed in our earlier paper [21] has been extended by incorporating gray-valued profiles across the organ surface, implementing the concept proposed by Cootes and Taylor [5, 6] to the third dimension.

The matching process is initialized using the average geometrical model resulting from this training phase. A two-stage algorithm, described in section 4, 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 along the surface, which becomes possible only by using a dual representation of the object both as a collection of sample points and as a parametrized surface.

3 Generation of 3-D statistical model

The concept proposed in this paper results in an *automatic selection of a large set of labeled surface points*. This is done using a uniform parametrization of closed surfaces and by calculating an invariant object centered description (Brechtbühler et al. [1, 2]). The alignment of parametrized object surfaces approximates a correspondence of surface points. The training set consists of a series of segmented volumetric objects obtained by experts using interactive segmentation.

3.1 Interactive expert segmentation

Today's routine practice for 3-D segmentation involves slice-by-slice processing of high-resolution volume data.

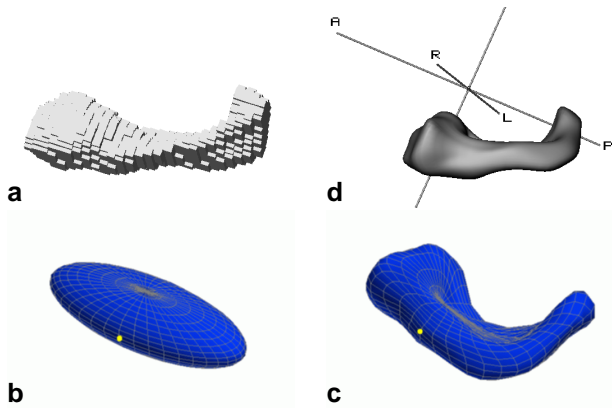


Figure 1. Model building: Interactive segmentation (a), reconstruction from shape descriptors up to degree one (b), reconstruction up to degree ten (c) and normalization in object space (d).

Working on large series of similar scans, human observers knowledgeable in anatomy become experts and produce highly 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 an interactive tool called "slice editor". The series of binary regions segmented from consecutive slices form volumetric voxel objects. Figure 1a illustrates the result of an expert segmentation of the left hippocampus.

3.2 Surface parametrization

The surface of a closed voxel object is most often stored as a mesh based on vertices having three spatial coordinates, although presenting two degrees of freedom. Brechbühler et al. [2] developed a surface parametrization of arbitrary simply connected objects based on those two parameters. The embedding of a convoluted object surface in the surface of the unit sphere of the parameter space was formulated as a constrained non-linear optimization problem. The method achieves a homogeneous density and minimal distortion of the parameter net projected onto the object surface. Further, objects of similar shape produce very similar parameter nets (up to a rotation in parameter space; see below) due to the uniqueness of the solution and robustness of the iterative optimization [2].

3.3 Spherical harmonic shape descriptors

The parametrization gives us three explicit functions defining the object surface as $\mathbf{x}(\theta, \phi) = (x(\theta, \phi), y(\theta, \phi), z(\theta, \phi))$. This surface description is used to expand a 3-D shape into a complete set of spherical harmonics. The series takes the form

$$\mathbf{x}(\theta, \phi) = \sum_{l=0}^K \sum_{m=-l}^l \mathbf{c}_l^m Y_l^m(\theta, \phi). \quad (1)$$

The coefficients \mathbf{c}_l^m are three-dimensional 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. [2]. All the \mathbf{c}_l^m with components (x, y, z) define the vector

$$\mathbf{c} = (c_{x_0}^0, c_{y_0}^0, c_{z_0}^0, c_{x_1}^{-1}, c_{x_1}^0, c_{x_1}^1, c_{y_1}^{-1}, c_{y_1}^0, c_{y_1}^1, c_{z_1}^{-1}, c_{z_1}^0, c_{z_1}^1, \dots, c_{x_K}^{-K}, \dots, c_{z_K}^K)^T.$$

3.4 Surface correspondence and object alignment

The surface parametrization, i.e., the 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. However, a 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 position and orientation of objects in original coordinate space has to be normalized before starting a comparison. For example, parameters for aligning objects can be obtained by calculating an object-centered coordinate system. 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 landmarks.

Object-centered invariant surface parametrization

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,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 Figure 2). 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 assessment of the quality of point correspondences, with a view to potential improvements, are research questions of current interest ([9, 22, 16]).

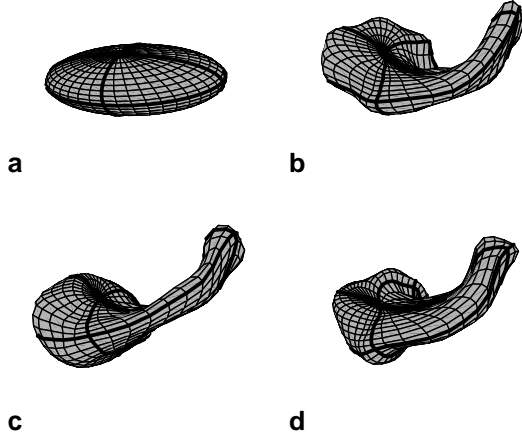


Figure 2. Corresponding parameter values for $\theta = \pi/2$, $\phi = 0, \pi$, and $\phi = \pi/2, 3\pi/2$ (thick lines) illustrated on an ellipsoid (a) and on three individual left-hippocampi.

Alignment in object space Our driving application is the automatic segmentation of brain objects. 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 inter-hemispheric fissure by a midsagittal plane and the definition of the anterior and posterior commissure (AC-PC) (see Fig. 3). Each data set is transformed into canonical coordinates by 3-D rotation and scaling.

3.5 Shape statistics

After transformation to canonical coordinates, the object descriptors are related to the same reference system and can be directly compared. An established procedure for describing a class of objects follows [5], where the calculation are

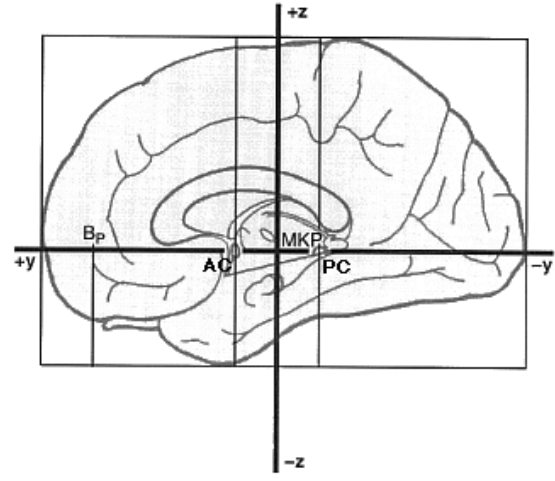


Figure 3. Stereotactic coordinate system used for object space normalization

carried out in the domain of shape descriptors rather than the Cartesian coordinates of points in object space. The mean model is determined by averaging the descriptors \mathbf{c}_j of the N individual shapes (see Fig. 4).

$$\bar{\mathbf{c}} = \frac{1}{N} \sum_{j=1}^N \mathbf{c}_j \quad (2)$$

Eigenanalysis of the covariance matrix Σ results in eigenvalues and eigenvectors representing the significant modes of shape variation.

$$\Sigma = \frac{1}{N-1} \sum_j (\mathbf{c}_j - \bar{\mathbf{c}}) \cdot (\mathbf{c}_j - \bar{\mathbf{c}})^T \quad (3)$$

$$\Sigma = \mathbf{P}_c \Lambda \mathbf{P}_c^T = (\mathbf{P}_c \Lambda^{1/2}) (\mathbf{P}_c \Lambda^{1/2})^T, \quad (4)$$

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

$$\mathbf{c}_j = \bar{\mathbf{c}} + \mathbf{P}_c \mathbf{b}_j \quad (5)$$

Figure 5 illustrates the largest two eigenmodes of the hippocampus training set. Truncating the number of eigenmodes, corresponding to the eigenvalues sorted by size, restricts deformations to the major modes of variation. Figure 7 illustrates the square root of eigenvalues sorted by size (dotted line) and the components of one typical vector \mathbf{b}_j .

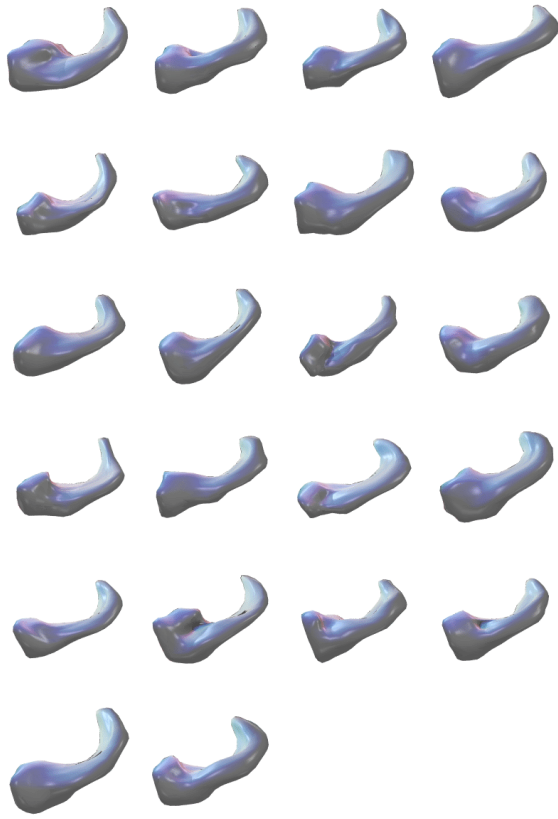


Figure 4. Illustration of all 22 left hippocampal structures of the training sets, normalized and reconstructed from their descriptors.

3.6 Modeling the gray level environment of surface models

Following the work of Cootes et al. [5, 6] we augment the geometric shape models by incorporating information about the gray level environment of the model surface. We examine the statistics of the image intensity along 1-D profiles orthogonal to the object surface at a discrete set of sampling points. Equal processing of each part of the model surface is ensured by choosing a homogeneous distribution of sampling points and profiles over the 3-D surface. Because the objects are parametrized by the two spherical coordinates (θ, ϕ) , the straightforward method would be to use a regular mesh of these parameters. This, however, would result in a highly irregular mesh on a spherical surface giving a dense sampling at the poles and a sparse sampling along the equator (see Figure 8(a) and Figure 1(b,c)). A perfectly regular sampling of a spherical surface does not exist, but we can find a good approximation by icosahedron subdivi-

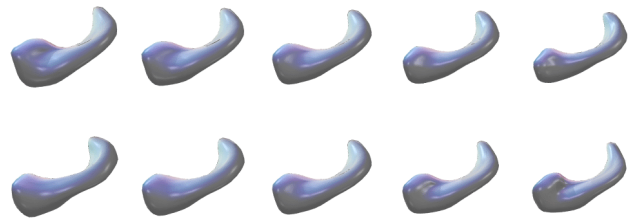


Figure 5. Largest two modes of variation for $b_j = -2\sqrt{\lambda_j} \dots 2\sqrt{\lambda_j}$. In the middle column, $b_j = 0$ represents the mean model.

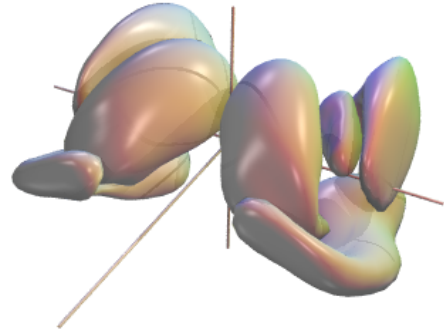


Figure 6. Mean models of left and right thalamus, globus pallidus, putamen and hippocampus.

sion, 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 Figure 8(b). 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.

We have chosen a subdivision of $k = 10$ which gives us $n = 12 + 30(k - 1) + 20 \frac{(k-1)(k-2)}{2} = 1002$ vertices. Computing the θ_i and ϕ_i values at each vertex coordinate i of the subdivided icosahedron and substituting them into

$$\mathbf{x}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \sum_{l=0}^K \sum_{m=-l}^l \mathbf{c}_l^m Y_l^m(\theta_i, \phi_i), \quad (6)$$

$$i = 1 \dots 1002 \quad (7)$$

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

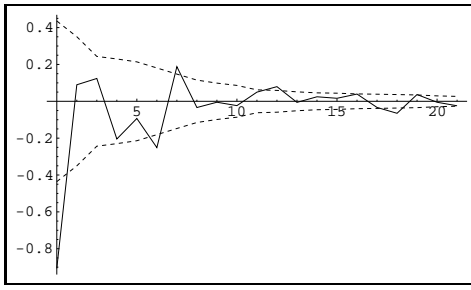


Figure 7. Statistics of shape deformation. The dotted line represents the square root of eigenvalues $\sqrt{\lambda_j}$ sorted by decreasing size. The continuous line illustrates the components of an individual vector b_j , which describes the deviation of the shape c_j from the mean shape \bar{c} .

$$\mathbf{x} = \mathbf{A}\mathbf{c}, \quad (8)$$

where \mathbf{x} represents the coordinates in object space and \mathbf{c} the spherical harmonics descriptors. \mathbf{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 \mathbf{w}_{ij} of n_p sample points. The distance between sample points is the length of one voxel edge. The profiles are oriented normal to the object surface and centered at the surface points \mathbf{x}_{ij} , as illustrated in Figure 9. For each sample point i we can obtain a mean profile by averaging over the sample objects N :

$$\bar{\mathbf{w}}_i = \frac{1}{N} \sum_{j=1}^N \mathbf{w}_{ij}. \quad (9)$$

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

Cootes et al. [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.

4 Segmentation by model fitting

Until now we have only described the creation of a flexible 3D model including geometric shape, gray level envi-

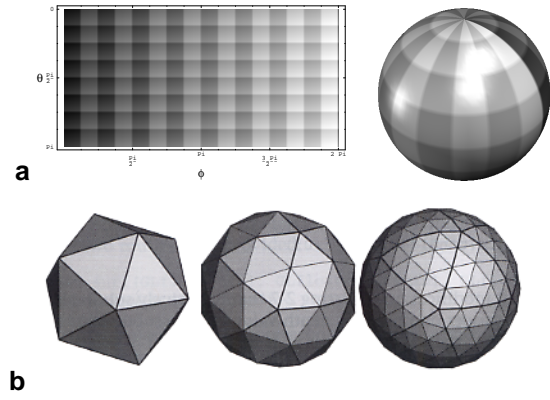


Figure 8. Sampling methods of spherical surfaces: regular mesh in spherical coordinates (a), icosahedron subdivision (b).

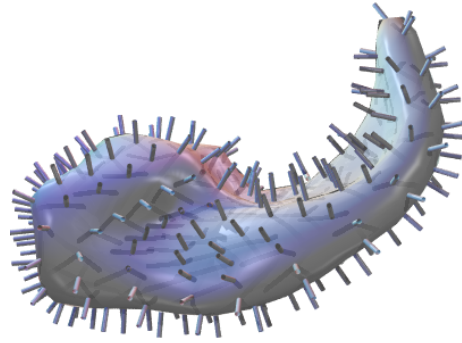


Figure 9. Illustration of an individual left hippocampal shape with its profile vectors shown from the left side of the brain.

ronment and statistics about normal shape variability. We now perform the segmentation step by elastically fitting this model to new 3D 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 surface will be elastically deformed until it best matches the new gray value environment.

4.1 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 we are planning to replace the determination of the symmetry plane and the AC/PC line by automatic methods [10, 25, 12]. 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.

4.2 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

$$\mathbf{n}_i = \sum_{l=0}^K \sum_{m=-l}^l \mathbf{c}_l^m \frac{\partial Y_l^m}{\partial \theta} \times \sum_{l=0}^K \sum_{m=-l}^l \mathbf{c}_l^m \frac{\partial Y_l^m}{\partial \phi}. \quad (10)$$

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 \mathbf{A} .

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 \mathbf{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:

$$d_{Maha}^2(s) = (\mathbf{w}(s) - \bar{\mathbf{w}}) \Sigma_{\mathbf{w}}^{-1} (\mathbf{w}(s) - \bar{\mathbf{w}}) \quad (11)$$

where $\mathbf{w}(s)$ represents the sub-interval 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 $d_{Maha}^2(s)$. Suppose s_{best} is the shift between the two profiles providing the best fit. We choose a displacement vector $d\mathbf{x}$ for each model point which is parallel to the profile, in the direction of the best fit and has magnitude s_{best} .

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

$$d\mathbf{x} = (dx_1, dy_1, dz_1, \dots, dz_n), \quad (12)$$

we then adjust the shape parameters to move the model surface towards 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 calculated displacements $d\mathbf{x}$, 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.

The shape statistics, as described in section 3.5, can be expressed by

$$\mathbf{c} = \bar{\mathbf{c}} + \mathbf{P}_c \mathbf{b}. \quad (13)$$

Multiplying both sides of the above equation by \mathbf{A} we get the dual surface description by a set of surface points:

$$\mathbf{A} \mathbf{c} = \mathbf{A} \bar{\mathbf{c}} + \mathbf{A} \mathbf{P}_c \mathbf{b} \quad (14)$$

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_x \mathbf{b}, \quad (15)$$

where \mathbf{P}_x denotes the product $\mathbf{A} \mathbf{P}_c$ which represents the matrix of modes of shape variation expressed in object coordinate space. Recall that \mathbf{P}_c is the matrix of eigenvectors in the shape descriptor space defined by the components of the elliptic harmonic descriptors \mathbf{c} . Interestingly, weight vectors \mathbf{b}_i of individual shapes, which express the deviation from the mean model, stay the same in both shape representation schemes.

We seek an adjustment $d\mathbf{b}$ to \mathbf{b} which causes a deformation in eigenspace which matches the optimal deformation \mathbf{x} as closely as possible.

$$(\mathbf{x} + d\mathbf{x}) = \bar{\mathbf{x}} + \mathbf{P}_x (\mathbf{b} + d\mathbf{b}). \quad (16)$$

Subtracting eq.(15) from eq.(16) we get

$$d\mathbf{x} = \mathbf{P}_x d\mathbf{b}. \quad (17)$$

This is an overdetermined 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 5 to 10). Therefore a least squares approximation to the solution can be obtained using standard linear algebra.

The entire procedure is repeated iteratively and starts with the average model such that $\mathbf{b}_{i=0} = \mathbf{0}$. At each iteration step, we compute a new set of displacements from the match of profiles and update the shape deviation vector \mathbf{b} until the shape stops to vary.

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. And on the other hand, after the weights have been updated by

$$\mathbf{b}_{t+1} = \mathbf{b}_t + d\mathbf{b}_t, \quad (18)$$

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

5 Results

Figure 10(a) shows the initial placement of the left hippocampus model (white line) together with the hand-segmented contour (gray line) on a sagittal 2D slice (top) and as a 3D scene (bottom) viewed from the right side of the head. Images (b), (c), and (d) show the iterative progress of the fit. After 100 iterations the model gives a sufficiently close fit to the data. The model used in this example had 5 degrees of freedom, and the model profiles had total length of 11 sample points while the extracted profiles had a total length of 19 sample points. The whole segmentation process takes about 2 minutes 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 22 data sets where the hippocampus had been manually segmented, and additionally to 8 data sets where it had not. The performance of the automatic segmentation has been tested by comparisons with manually segmented object shapes, which are well recognized as a gold standard given the lack of ground truth. **A** represents the model shape obtained by human experts, **B** the result of the new model-based segmentation.

The overlap measure $(\mathbf{A} \cup \mathbf{B}) / (\mathbf{A} \cap \mathbf{B})$ shown in Figure 11 is calculated on binary voxel maps, created by intersection of the object surfaces with the voxel grid. The resulting measure is very sensitive to even small differences in overlap, both inside and outside of the object model, and 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 1 voxel edge.

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 expression. We therefore avoid the discretization errors by projecting the surfaces back to a voxel grid.

$$\oint \|\mathbf{x}(\mathbf{u})\|^2 d\mathbf{u} = \sum_{l=0}^{\infty} \sum_{m=-l}^l |\mathbf{c}|^2 = 4\pi \cdot \text{MSD}, \quad (19)$$

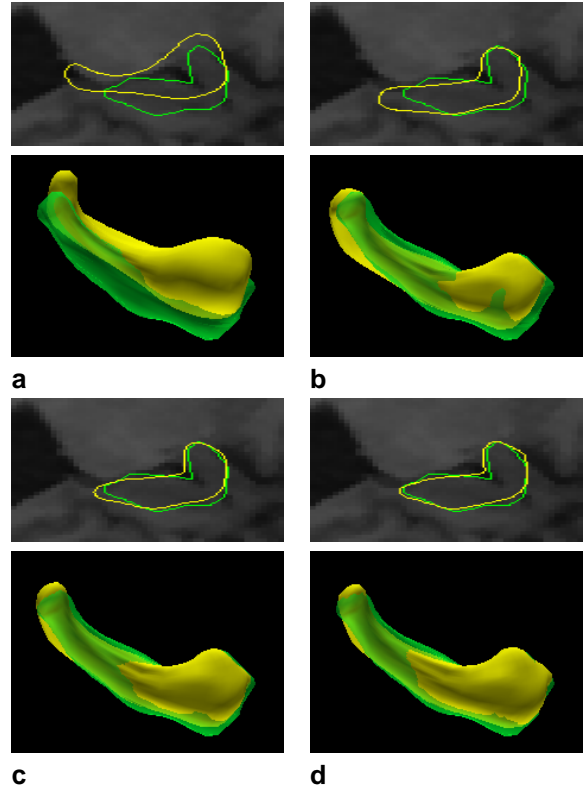


Figure 10. Segmentation results of a left hippocampus on sagittal 2D slices and 3D views from the left hand side. The images a-d were taken after 0, 30, 60 and 100 iterations.

where MSD stands for *mean squared distance* measured from the origin of the coordinate system.

Figure 12 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 grey 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.

6 Conclusions

We present a new 3-D segmentation technique that provides fully automatic segmentation of objects from volumetric image data. Tests with a large series of image data demonstrated that the method was robust and provides re-

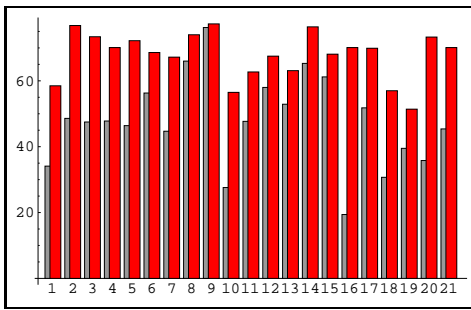


Figure 11. Overlap measure $(A \cup B)/(A \cap B)$ in percentage calculated between manually (A) and automatic (B) segmented left hippocampi of 21 individuals. Bars in light grey illustrate the measure at initialization and in dark grey after deformation.

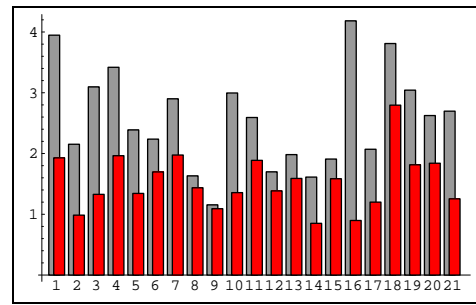


Figure 12. Average distances in mm calculated between manually and automatic segmented left hippocampi of 21 individuals. The bars in light grey illustrate the mean distance of the initialization of the model in a new data set and the dark bars the final mean distance of surfaces to the model surface. The length of the hippocampus is about 40 mm.

producible 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 model has been derived from a series of training data sets. Thereby, the model represents a realistic shape rather than a simple geometric 3-D figure. Furthermore, information about the statistics of a normal shape deformation helps to constrain a new solution. 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.

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

Statistical shape models: To overcome the problem of getting a reproducible interactive definition of a set of key 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.

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].

Dual shape representations: Our approach makes use of

two shape representations which are used in a vice-versa fashion, taking advantages 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 modelling of gray level information near the object boundaries provides valuable additional information for a model placement and improves the robustness and stability of the optimization scheme. An early version of our segmentation [21] 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.

We noticed that the convergence is faster if only a small number of modes (usually 5) are involved, while a larger number of modes (usually 10) 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.

Validation so far has been only done by visually comparing the automatic segmentation with the results of interactive outlining by experts, (see Fig. 10), which is a common “gold standard” for comparisons. We are currently working on a quantitative validation study.

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 a large number of data sets as collected in clinical studies, for example in schizophrenia studies. This will provide new statistical models with increased number of samples for normal controls and for different patient categories.

Figure 6 displays average models of four different brain objects. 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.

Acknowledgements

Ron Kikinis and Martha Shenton, Brigham and Women's Hospital, Harvard Medical School, Boston kindly provided the original MR and segmented data sets. We are thankful to Christian Brechbühler and Jonathan Oakley from our lab for providing the software for surface parametrization and shape description, and for carefully proof-reading the manuscript, respectively. Jens-Peer Kuska is acknowledged for making us Mathview3d available which converts *Mathematica*TM 3D objects to *PovRay* ray-tracer files.

References

- [1] C. Brechbühler, G. Gerig, and O. Kübler. Surface parametrization and shape description. In *Visualization in Biomedical Computing*, pages 80–89, 1992.
- [2] C. Brechbühler, G. Gerig, and O. Kübler. Parametrization of closed surfaces for 3-D shape description. *CVGIP: Image Understanding*, 61:154–170, 1995.
- [3] I. Cohen, L. Cohen, and N. Ayache. Using Deformable Surfaces to Segment 3D Images and Infer Differential Structures. *CVGIP: Image Understanding*, 56(2):242–263, 1992.
- [4] T. Cootes and C. Taylor. Active Shape Models – ‘Smart Snakes’. In *British Mach. Vision Conf.*, pages 266–275. Springer-Verlag, 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 and Vision Computing*, 12(6):355–366, July 1994. Electronic version: <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. *Computer Vision and Image Understanding*, 61(1):38–59, Jan. 1995.
- [7] G. Gerig, J. Martin, R. Kikinis, O. Kübler, M. Shenton, and F. Jolesz. Automatic Segmentation of Dual-Echo MR Head Data. In *IPMI'91*, pages 175–187. Wye, GB, 1991.
- [8] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Comp. Vision*, 1(4):321–331, 1988.
- [9] A. Kotchegg and C. Taylor. Automatic Construction of Eigenshape Models by Genetic Algorithm. In *Information Processing in Medical Imaging*, pages 1–14. Springer, 1997.
- [10] P. Marais, R. Guillemaud, M. Sakuma, A. Zisserman, and M. Brady. Visual Cerebral Asymmetry. In *Visualization in Biomedical Computing*, pages 411–416, 1996.
- [11] T. McInerney and D. Terzopoulos. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*, 1(2):91–108, 1996.
- [12] 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. *The Journal of Nuclear Medicine*, 34(2):322–329, Feb. 1993.
- [13] C. Nastar and N. Ayache. Frequency-based Nonrigid Motion Analysis: Application to Four Dimensional Medical Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1067–1079, Nov. 1996.
- [14] N. Neuenchwander, P. Fua, G. Székely, and O. Kübler. Initializing Snakes. In *CVPR'94*, pages 658–663, 1994.
- [15] A. Pentland and A. Sclaroff. Closed-Form Solutions for Physically Based Shape Modelling and Recognition. *IEEE PAMI*, 13(7):715–729, 1991.
- [16] A. Rangarajan, H. Chui, and F. Bookstein. The Softassign Procrustes Matching Algorithm. In *Information Processing in Medical Imaging*, pages 29–42. Springer, 1997.
- [17] S. Sclaroff and A. Pentland. On Modal Modelling for Medical Images: Underconstrained Shape Description and Data Compression. In *IEEE Workshop on Biomed. Image Anal.*, pages 70–79, Seattle, Washington, USA, 1994.
- [18] L. Staib and J. Duncan. Boundary Finding with Parametrically Deformable Models. *IEEE PAMI*, 14(11):1061–1075, 1992.
- [19] L. Staib and J. Duncan. Deformable Fourier models for surface finding in 3D images. In *VBC'92*, pages 90–194, 1992.
- [20] L. Staib and J. Duncan. Model-based Deformable Surface Finding for Medical Images. *IEEE Trans. Med. Imaging*, 15(5):1–12, 1996.
- [21] G. Székely, A. Kelemen, C. Brechbühler, and G. Gerig. Segmentation of 2-D and 3-D objects from MRI volume data using constrained elastic deformations of flexible Fourier contour and surface models. *Medical Image Analysis*, 1(1):19–34, 1996.
- [22] H. Tagare. Non-rigid Curve Correspondence for Estimating Heart Motion. In *Information Processing in Medical Imaging*, pages 489–494. Springer, 1997.
- [23] J. Talairach and P. Tournoux. *Co-planar stereotaxic atlas of the human brain*. Thieme, Stuttgart, 1988.
- [24] D. Terzopoulos, A. Witkin, and M. Kass. Symmetry-Seeking Models and 3D Object Reconstruction. *Int. J. Comp. Vision*, 1(3):211–221, 1988.
- [25] J. P. Thirion, S. Prima, and G. Subsol. Statistical Analysis of Dissymmetry in Volumetric Medical Images. Technical Report 3178, INRIA, June 1997. Electronic version: <http://www.inria.fr/RRRT/RR-3178.html>.
- [26] B. Vemuri and A. Radisavljevic. Multiresolution Stochastic Hybrid Shape Models with Fractal Priors. *ACM Trans. Graphics*, 13(2):177–200, 1994.