Diss. ETH No.

Elastic Model-Based Segmentation of 2-D and 3-D Neuroradiological Data Sets

A dissertation submitted to the

SWISS FEDERAL INSTITUTE OF TECHNOLOGY

for the degree of

Doctor of Technical Science

presented by

ANDRÁS KELEMEN

Dipl. El. Ing TU Budapest
born December 22, 1968
citizen of Hungary

accepted on the recommendation of
Prof. Dr. G. Gerig, examiner
Prof. Dr. C. Taylor, coexaminer
Dr. G. Székely, coexaminer

Zurich 1998

Abstract

The general problem addressed in this thesis is the automatic model-based segmentation of two- and three-dimensional objects from medical images and volumetric image data. The development of the 2-D segmentation procedure closely follows the seminal work of [Cootes *et al.* 1994a]. The work done being necessary to allow the extension of this technique to 3-D, presenting solutions to the difficulties arising in application to true volume data as opposed to slice-by-slice 2-D processing.

The segmentation system includes both the building of statistical models and the automatic segmentation of unseen image data sets via a constrained elastic deformation of models. Geometric models are derived from a sample set of image data which has been segmented by experts; a training set. The surfaces of these binary objects are converted into a parametric surface net which is normalized to yield 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 stereotactic coordinate system, primarily chosen for its general acceptibility in applications of neuroimaging. 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, restricted only by shape deformation constraints.

The technique has been applied to the segmentation of the corpus callosum and the left hippocampus from a large series of 2-D and 3-D magnetic resonance scans. Validation was done by comparing the results of our automated segmentation approach to those performed manually by experts.

Kurzfassung

Diese Dissertation behandelt das allgemeine Problem der automatischen modellbasierten Segmentierung zwei- und dreidimensionaler Objekte aus medizinischen Bild- und Volumendaten. Die Entwicklung der Segmentierungsmethode in 2-D folgt der grundlegenden Arbeit von Cootes, Taylor, Cooper und Graham aus dem Jahre 1995. Hier werden neue Konzepte vorgestellt, die es ermöglichen, diese Technik auf 3-D zu erweitern und auf echte Volumendaten anzuwenden, im Gegensatz zur 2-D-Verarbeitung einzelner Schichten.

Das Segmentierungssystem umfasst den Aufbau statistischer Modelle und die automatische Segmentierung unbekannter Bilddatensätze mittels eingeschränkter elastischer Deformation der Modelle. Der Trainingssatz ist einer Stichprobe von durch Experten segmentierten Bilddaten; aus ihm leiten sich die geometrischen Modelle ab. Wir konstruieren eine bijektive Abbildung der Oberflächen der binären Objekte auf die Einheitskugel; eine Normierung dieser Parametrisierung ergibt ein invariantes objektzentriertes Koordinatensystem. Die Reihenentwicklung der Koordinatenfunktionen der Oberfläche in Kugelfunktionen liefert eine parametrische Darstellung der Form der Objekte. Die zugehörige Grauwertinformation wird durch 1-D Profile dargestellt, die das Volumen orthogonal zur Oberfläche abtasten. Die Ausgangslage der Modelle basiert auf dem stereotaktischen Koordinatensystem, das sich vor allem wegen seiner allgemeinen Verwendung in der Neuroanatomie anbietet. Formstatistiken werden aus den parametrischen Formdarstellungen errechnet und nicht aus räumlichen Koordinaten von Punktmengen. Nachdem das Durchschnittsmodell in einem neuen Datensatz auf Grund der stereotaktischen Koordinaten platziert ist, deformiert und verschiebt es sich elastisch unter der Wirkung der Verschiebungskräfte, die sich aus der Übereinstimmung der Grauwerprofile auf der Oberfläche mit dem Datenvolumen ergeben. Das statistische Modell definiert die zulässigen Deformationen.

Das Verfahren wurde zur Segmentierung des Corpus Callosum und des linken Hippocampus aus einer grossen Serie von 2-D- und 3-D-Magnetresonanzbildern eingesetzt. Zur Validierung wurden die Resultate unserer automatischen Segmentierung mit den manuell von Experten durchgeführten verglichen.

1

Acknowledgements

In the first place I wish to thank Prof. Dr. Guido Gerig for his continuous, valuable and motivating support throughout my work, for involving me into the BIOMORPH project which greatly helped me to learn research activities most closely related to the fascinating field of medical image processing and shape morphometry, and finally for providing a pleasant working environment at the ETH in Zurich.

My supervisor and co-examiner, Dr. Gábor Székely greatly helped me

I am greatly indepted to Prof. Dr. Christopher Taylor from whom I received many helpful suggestions. I am very grateful for his spontaneous willingness to co-examine the thesis.

Special thanks to Dr. Christian Brechbühler

Last but not least I wish to acknowledge all the members of the Image Science Laboratory

Contents

List of Figures

List of Tables

2

Introduction

2.1 Medical image segmentation

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. Today's routine practice for 3-D segmentation involves slice-by-slice processing of high-resolution volume data. 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 2.1 illustrates the result of an expert segmentation of the left hippocampus.

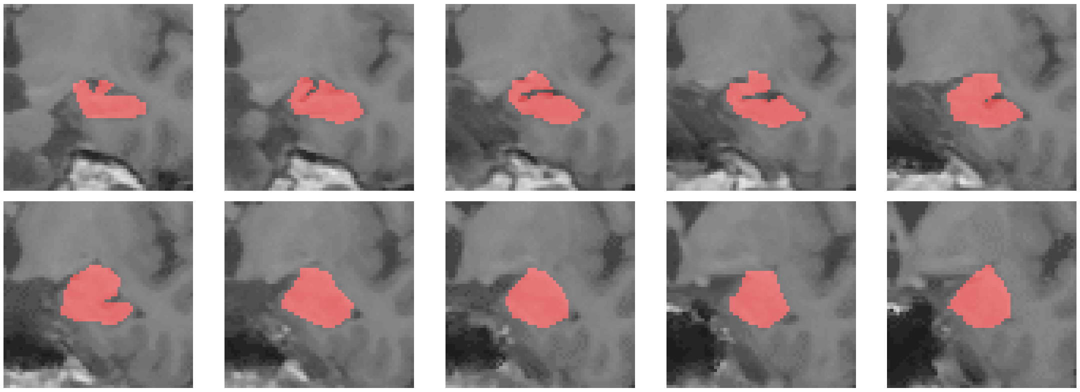


Figure 2.1: Ten consecutive slices of manual segmentation of the left hippocampus overlaid with the original MRI data set.

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. Many attempts has been done to achieve semi- or fully automatic segmentation of medical datasets by incorporating human expertise.

2.2 Segmentation techniques

Intensity based segmentation techniques on the low-level end, such as multi-spectral classification, provide a fast, reproducible and fully automatic way to segment medical images. The only knowledge, however, these methods have about the sought objects is their intensity distribution. Other important high-level features like location and shape of organs, which a human expert would help to find the objects, are completely ignored and also difficult to include because of the low-level nature of the procedure.

Elastically deformable contour and surface models, so-called snakes, have been proposed as tools for supporting manual object delineation. Besides intensity, these techniques use additional structural knowledge about the underlying image in form of edges. On the other hand, through interactive initialization automation and reproducibility is lost. Furthermore, while such procedures can be extended to 3-D, their initialization is a critical issue. Most often, the initial guess must be very close to the sought contour to guarantee a satisfying result.

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.

For complex shapes described by a large number of often highly correlated parameters, the use of such priors may become tedious. The modal analysis 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

The power of parametric deformable shape descriptors with statistical modal analysis has been combined by [Cootes *et al.* 1994a]. 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), 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.

However, 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.

2.3 Thesis Overview

The goal of the presented thesis is to develop an automatic atlas-based technique for segmentation of MR volumes of human neuroanatomy. The proposed segmentation technique involves generation of an anatomical atlas in the first stage, and then elastic matching of the atlas to individual anatomy in the second stage.

The atlas should be derived by the statistical evaluation of a set of manually segmented 3-D brain regions. Computation of statistical descriptors like average and covariance out of a given set of shapes assumes correspondence between the objects. In case of smooth surfaces, establishing homology by hand is infeasible because identification of stable landmarks cannot be carried out.

An alternative way to approximate correspondence is provided by curve- and surface-parametrization and subsequent normalization of the parameters. Mapping positions along the boundary of similar objects onto a common parameter space automatically identifies homology among them. This procedure can be also applied to 3-D surfaces and is not limited to 2-D contours. A beneficial side-effect of correspondence by parametrization is the more compact object description than the set-of-points representation behind point distribution models.

Statistical description by average shape and covariance matrix obtained from normalized coefficients of a training set can be computed. Applying modal analysis to the covariance matrix reduces significantly the dimensionality of the shape description.

Having generated statistical models of organs, they has to be matched to individual medical data sets. Because of the biological variability of human anatomy, elastic deformations of the atlas is needed to achieve a sufficient fit. By using statistical models, the matching algorithm can be constrained to only allow shapes which are consistent with the natural variability of the training set.

New methods have been implemented first for 2-D objects, since testing and validation can be carried out less costly in two dimensions. Only the most mature 2-D technique have been implemented for 3-D surfaces. This concept is also reflected in the organization of the thesis, which basically follows the historical development of the new 3-D segmentation framework.

2.4 Organization of the Thesis

Chapter 2 (this chapter) describes the research field of medical image segmentation. It states the goal of our work, which is the development of an automatic segmentation tool for 2-D and 3-D organs of the human brain based on elastic matching of a statistical atlas to individual anatomy.

Chapter 3 reviews previous work on related fields. It contains a description of deformable models, since they represent the basis of our segmentation process. Curve and surface parametrization techniques are then presented, followed by a discussion on representing anatomical knowledge using deformable atlases of neurology, and capturing biological variability of organs by statistical methods.

Chapter 4 gives detailed overview on Fourier curve parametrization and its 3-D counterpart, the parametrization of surfaces by spherical harmonic basis functions. These methods have been applied to approximate correspondence between individual anatomy in order to compute average models of shapes.

Chapter 5 describes our attempts towards building statistical models of 2-D organs with the help of arc-length parametrization to solve the correspondence problem. An example presents the limitations of this method and proposes solution by reparametrization. We also see the gradual development of an elastic matching algorithm based on statistical models.

Chapter 6 presents the generalization of the techniques discussed in chapter 5 to the third dimension. While most of the methods could be directly converted to 3-D, there were some issues which needed different treatment, these will receive more attention. Additionally, we address the topics of multi-shape models and robustness of multi-variate statistics derived from small number of samples.

Chapter 7 draws conclusions from the presented thesis and gives suggestions for further work.

3

Review of related work

3.1 Image segmentation

Segmentation of anatomical objects from large three-dimensional (3-D) data sets has become increasingly important during the last decade. New techniques like magnetic resonance imagery (MRI), computed tomography imagery (CTI), nuclear medicine imagery (NMI) or ultrasound imagery (USI) in the medical field or confocal scanning laser microscopy (CLSM) in biology produce thousands of volumetric images each year. Although these new devices offer unique views into the internal human anatomy or microscopic structures, computer based methods still lack in efficiency and in accuracy to extract and quantify embedded structures. The spectrum of activities relying on accurate data ranges from radiotherapy planning, to surgery, to diagnosis, to morphometric studies.

Segmenting structures from 3-D images is difficult because of the complexity and variability of the shapes of interest, and the artifacts typical to sampled data, such as noise, bias field, and spatial aliasing. This prevents low-level image processing techniques, based only on local information, from generating reliable and accurate segmentation results. These model-free techniques often need expert interaction to achieve acceptable results. On the other side, user interaction considerably slows down the segmentation process, and it introduces additional errors and uncertainty into it.

Deformable models has been thoroughly researched and proven to be a promising model-based approach to computer-assisted image processing. Their ability to incorporate both *a priori* knowledge about the shape and environment of the object, and constraints from the image data itself, makes them suitable in particular for model-based image segmentation. Depending on the mathematical model behind them, deformable models can represent a vast variety of biological structures and their variability across different individuals. One must mention that deformable models are also extensively applied in other fields of medical image analysis, like matching and motion tracking [McInerney and Terzopoulos 1996].

There are several surveys discussing 3-D segmentation methods in general (e.g. [Ayache 1995]), this section, however, will only focus on to the present work more closely related fields of *deformable models* and *model based segmentation*.

3.1.1 Deformable models

The mathematical background of deformable models consists of three major parts: *geometry* representing object shape, *behavior* constraining the way object shape may vary, and *fitting algorithm* providing a method to fit the model to the measured data. Geometric representations can be divided into *nonparametric* and *parametric* classes and the latter into a large variety of groups ranging from fully *local* to fully *global* representations.

Nonparametric representations simply describe the curve or surface using geometric primitives such as line segments or polygons, while parametric representations usually provide a more compact way to describe an object and also allow of resampling object boundaries. Global descriptions enable shape comparisons while losing the connection to the contour itself, meaning that local changes of the boundary effect all of the descriptors. *Hybrid* models has been introduced to merge the advantages of *local* and *global* representations.

The *behavior* assigned to a certain geometric model constrains its deformation capacity in an intuitively meaningful way. Traditionally deformable models are placed into a physical framework defined by Lagrangian dynamics. This physical interpretation views deformable models as elastic bodies which respond naturally to applied forces and constrains. Typically, deformation energy functions defined in terms of the geometric descriptors are associated with the deformable model. The value of the models energy increases as the model deforms away from a natural rest position and usually include terms that constrain the smoothness of the curve. In the Lagrangian setting, the deformation energy gives rise to elastic *internal forces.* On the other hand, the external potential energy defined in terms of the data of interest, to which the model is to be fitted, gives rise to *external forces* which deform the model such that it fits the data. Some authors distinguish a separate category of *global forces* which effect the whole model and are usually introduced artificially to enforce specific behavior tasks such as shrinking or expanding of the model. The Lagrangian formalism allows to include interactive manipulations performed by the user into this framework. *Springs* allow the user to drag one point of the model and acts as a global force, while a volcano ads a Gaussian heap to the gradient image to push it away and can be characterized as external force. Although, this physics-based terminology of *forces* only applies fully to one certain category of deformable models, popularly known as 'snakes' [Kass *et al.* 1988], other types of deformable models also inherit the very specific properties of *internal* and *external forces.*

The *fitting algorithm* is the actual mathematical method to compute the elastic deformation of the *geometry* based on its defined *behavior.* Physics based models usually use Euler-Lagrange differential equations to evolve the deformable model while parametric models are most often apply multi-dimensional energy minimization techniques to compute the equilibrium between *internal* and *external forces.*

3.2 Traditional *Active Contour Models*

Basically all deformable models can be originated from the seminal technique first proposed by [Kass *et al.* 1988] in 1988. We briefly outline the mathematical background of their "snakes" model.

Traditional snakes used for image segmentation are polygonal curves to which an objective function is associated that combines an image term (*external force*) that measures either edge strength or edge proximity, and a regularization term (*internal force*) that accounts for tension and curvature. The curve is deformed so as to optimize the score and, as a result, to match the image edges.

The original snakes introduced by Kass *et al.* [Kass *et al.* 1988] are modeled as time-dependent 2-D curves defined parametrically as

$$\boldsymbol{v}(s,t) = (x(s,t), y(s,t))_{0\leq s\leq 1} \ , \tag{3.1}$$

where s is proportional to the arc length, t the current time, and x and y the curve's image coordinates. The snake deforms itself as time progresses so as to minimize an image potential $E_i(\boldsymbol{v})$, with

$$E_I(\boldsymbol{v}) = -\int_0^1 P(\boldsymbol{v(s,t)})ds \tag{3.2}$$

where $P(\boldsymbol{v(s,t)})$ is a function of the image. One typical choice is to take $P(\boldsymbol{v(s,t)})$ to be equal to the magnitude of the image gradient, that is

$$P(\boldsymbol{v(s,t)}) = |\nabla I(\boldsymbol{v(s,t)})| \ , \tag{3.3}$$

where I is either the image itself or the image convolved by a Gaussian kernel.

Alternatively, one can take $P(\boldsymbol{v(s,t)})$ to be the Euclidean distance to the closest edge-point in an edge-map computed using an operator such as the edge-detector proposed by [Canny 1986]. The results of all these approaches are very similar.

Unfortunately, whatever the choice of P, $E_I(\boldsymbol{v})$ is typically not a convex functional which can affect the minimization in an undesirable way. To overcome this problem, a convex regularization term $E_D(\boldsymbol{v})$ is introduced. As a consequence, a total energy term $E(\boldsymbol{v})$ that is the sum of $E_I(\boldsymbol{v})$ and $E_D(\boldsymbol{v})$ is then minimized. Using the elastic rod model, $E_D(\boldsymbol{v})$ is taken to be

$$E_D(\boldsymbol{v}) = \frac{1}{2}\int_0^1 \alpha(s)\left|\frac{\partial\boldsymbol{v}(s,t)}{\partial s}\right|^2 + \beta(s)\left|\frac{\partial^2\boldsymbol{v}(s,t)}{\partial s^2}\right|^2 ds \ , \tag{3.4}$$

where $\alpha(s)$ and $\beta(s)$ are arbitrary functions that regulate the curve's tension and rigidity.

The original approach has been extended by [Terzopoulos 1987] introducing dynamic terms such as kinetic energy and velocity dependent friction. The kinetic energy term $E_K(\boldsymbol{v})$ allows the snake to move in a physically reasonable manner and is set to

$$E_K(\boldsymbol{v}) = \frac{1}{2}\int_0^1 \mu(s)\left|\frac{\partial\boldsymbol{v}(s,t)}{\partial t}\right|^2 ds \ , \tag{3.5}$$

where $\mu(s)$ represents the mass of the snake. The objective function associated to the snake is then an energy term of the form

$$E(\boldsymbol{v}) = E_K(\boldsymbol{v}) + E_I(\boldsymbol{v}) + E_D(\boldsymbol{v}) \tag{3.6}$$

Snakes in expression (3.6) are regarded as an energy conserving physical system. I order to achieve a more realistic physical behavior, energy dissipation should be allowed also. For this purpose, the Rayleigh dissipation functional

$$D(\boldsymbol{v}) = \frac{1}{2}\int_0^1 \gamma(s)|\boldsymbol{v}_t|^2 ds \tag{3.7}$$

is incorporated where $\gamma(s)$ is the damping coefficient. The segmentation process performed with snakes is governed by the minimization of the term $\int E(\boldsymbol{v})+D(\boldsymbol{v})dt$. This amounts to using Hamilton's principle in variational calculus to derive the *Euler-Lagrange* equations of motion.

$$\mu\frac{\partial^2\boldsymbol{v}(s,t)}{\partial t^2}+\gamma\frac{\partial\boldsymbol{v}(s,t)}{\partial t}-\frac{\partial}{\partial s}\left(\alpha(s)\frac{\partial\boldsymbol{v}(s,t)}{\partial s}\right)+\frac{\partial^2}{\partial s^2}\left(\beta(s)\frac{\partial^2\boldsymbol{v}(s,t)}{\partial s^2}\right) = -\nabla P(\boldsymbol{v}(s,t)) \tag{3.8}$$

Presuming α and β are taken to be constant and supplied by the user, the equations of motion for the snake can be written as

$$\begin{aligned} \mu\frac{\partial^2 x(s,t)}{\partial t^2}+\gamma\frac{\partial x(s,t)}{\partial t}-\alpha\frac{\partial^2 x(s,t)}{\partial s^2}+\beta\frac{\partial^4 x(s,t)}{\partial s^4} &= -\frac{\partial P}{\partial x} \\ \mu\frac{\partial^2 y(s,t)}{\partial t^2}+\gamma\frac{\partial y(s,t)}{\partial t}-\alpha\frac{\partial^2 y(s,t)}{\partial s^2}+\beta\frac{\partial^4 y(s,t)}{\partial s^4} &= -\frac{\partial P}{\partial y} \end{aligned} \tag{3.9}$$

Solving these equations in a computer system means that the evolution of the snake can be performed only in discrete time steps. The model $\boldsymbol{v}$ is taken to be a polygonal curve defined by a set of vertices. It is customary to use central and backward finite differences, respectively to approximate time and spatial derivatives.

3.3 Parametric models

Another way of representing the model is to use a parametric vector $\boldsymbol{p}$ rather than a set of points. Since parametric models are of particular interest with respect to the presented work, they will be outlined in more detail. We first give an overview of curve and surface parametrization techniques, then summarize methods applied to these to achieve elastic deformation and finally present optimization techniques applied to solve the energy minimization problem arising in the model fitting process.

3.3.1 Curve and surface parametrizations

Parametric models use a parameter vector $\boldsymbol{p}$ rather than a set of points as their geometric representation. The general 2-D and 3-D models $\boldsymbol{v}(\boldsymbol{p})$ can be defined by the mappings:

$$\begin{array}{rrcl} \boldsymbol{v}(s;\boldsymbol{p}): & \Omega\subset\mathcal{R} & \rightarrow & \mathcal{R}^2 \\ & s & \mapsto & \boldsymbol{v}(s;\boldsymbol{p}) = (v_1(s;\boldsymbol{p}), v_2(s;\boldsymbol{p}))^T \end{array}$$

$$\begin{array}{rrcl} \boldsymbol{v}(s,t;\boldsymbol{p}): & \Omega\subset\mathcal{R}^2 & \rightarrow & \mathcal{R}^3 \\ & s,t & \mapsto & \boldsymbol{v}(s,t;\boldsymbol{p}) = \begin{pmatrix} v_1(s,t;\boldsymbol{p}) \\ v_2(s,t;\boldsymbol{p}) \\ v_3(s,t;\boldsymbol{p}) \end{pmatrix} \end{array}$$

Coordinate functions in $\boldsymbol{v}(s;\boldsymbol{p})$ and $\boldsymbol{v}(s,t;\boldsymbol{p})$ can be either represented in a geometric way by assigning unique functions to each of them or they can be taken as a series of basis functions. Both classes house parametrizations ranging from *local* to *global.* In the category of geometric parametrizations belong objects represented by a *set of points* at the local extreme and *superquadrics* at the global end. Similarly, in the category of parametrizations by basis functions *B-splines* provide a local and *Fourier* and *spherical harmonics* bases a global representation, while *wavelets* compromise between the two inheriting advantages of both. *Hybrid parametrizations* offer an alternative way to extend local descriptions by global ones and achieve a more flexible model. In the following these techniques will be introduced one by one in more detail.

Point sets

The most simple and direct way to represent objects is by taking a set of points which lie on its boundary or surface. This can be expressed in 2-D as

$$\boldsymbol{v}(s,\boldsymbol{p}) = (x_0, y_0, x_1, y_1, \ldots, x_s, y_s, \ldots, x_{n-1}, y_{n-1})^T \ , \tag{3.10}$$

and in 3-D as

$$\boldsymbol{v}(s,\boldsymbol{p}) = (x_0, y_0, z_0, x_1, y_1, z_1, \ldots, x_s, y_s, z_s, \ldots, x_{n-1}, y_{n-1}, z_{n-1})^T \ , \tag{3.11}$$

where $s = 0 \ldots n-1$ parameter indexes the points and the coordinates (x_s, y_s), respectively (x_s, y_s, z_s) are the actual shape descriptors. Since point sets are not mapped onto a surface, there is no need for a second parameter t. One can notice that point sets does not necessarily have any topology, hence they can represent objects of arbitrary shape.

Point set representations have been used as a basis for *Point Distribution Models* (PDMs) by [Cootes *et al.* 1994a] to represent 2-D objects such as hands, faces, and organs in medical images. They have been applied by [Caunce and Taylor 1998] to compute a PDM of the 3-dimensional cortical sulci of the human brain. [Bookstein 1997] employed them in his morphometric synthesis of the midsagittal plane of the brain.

Superquadrics

A superquadrics ellipsoid, is defined as

$$\boldsymbol{v}(a, a_1, a_2, a_3, \varepsilon_1, \varepsilon_2) = a \begin{pmatrix} a_1 C_u^{\varepsilon_1} C_v^{\varepsilon_2} \\ a_2 C_u^{\varepsilon_1} S_v^{\varepsilon_2} \\ a_3 S_u^{\varepsilon_1} \end{pmatrix} \quad \begin{array}{c} -\pi/2 \le u \le \pi/2 \\ -\pi \le v < \pi \end{array} \tag{3.12}$$

where $S_w^\varepsilon = \operatorname{sgn}(\sin w)|\sin w|^\varepsilon$ and $C_w^\varepsilon = \operatorname{sgn}(\cos w)|\cos w|^\varepsilon$. Here, $a \ge 0$ is a scale parameter, $0 \le a_1, a_2, a_3 \le 1$ ere aspect ratios and $\varepsilon_1, \varepsilon_2 \ge 0$ are “squareness” parameters.

This six degrees of freedom model in its plain form is rarely employed in elastic fitting techniques because of its limited capability in describing detailed surfaces. It customary to extend it by “tapering” and “bending” terms to achieve higher flexibility. Combining superquadrics with surface representations by series of basis functions results in *hybrid models* as discussed later in this section.

B-splines

B-splines are a computationally convenient representations for functions. In the B-spline form, the coordinate functions $v_1(s)$ and $v_2(s)$ are constructed as a weighted sum of N_B basis functions $B_n(s)$, $n = 0, \ldots, N_B - 1$. In the regular case, each basis function consists of d polynomials each defined over a span of the s-axis. The spans have unit length and are joined at evenly spaced knots. The constructed spline function is

$$x(s) = \sum_{n=0}^{N_B-1} x_n B_n(s) \tag{3.13}$$

where x_n are the weights applied to the respective basis functions $B_n(s)$. B-splines provide a local parametrization since one particular weight only affect a small well defined part of the whole curve. By convention, B-spline basis functions are constructed in such a way that they sum to 1 at all points:

$$\sum_{n=0}^{N_B-1} B_n(s) = 1 \text{ for all } s. \tag{3.14}$$

In the simple case of a quadratic B-spline with knots spaced regularly at unit intervals, the first B-spline basis function has the form

$$B_0(s) = \begin{cases} s^2/2 & \text{if } 0 \leq s < 1 \\ \frac{3}{4} - (s - \frac{3}{2})^2 & \text{if } 1 \leq s < 2 \\ (s-3)^2/2 & \text{if } 2 \leq s < 3 \\ 0 & \text{otherwise} \end{cases} \tag{3.15}$$

and the others are simply translated copies:

$$B_n(s) = B_0(s - n). \tag{3.16}$$

A more detailed introduction to B-splines and deformable models based on them can be find in the book *Active Contours* by [Bake and Isard 1998].

Fourier

The contour of a simply connected object is represented by a closed curve with coordinates $(v_1(s), v_2(s))$, with s ranging from 0 to $2\,\pi$. The coordinate functions can be expanded in a Fourier-series. Restricting the series to degree K results in the parametric description

$$\boldsymbol{v}(s, \boldsymbol{p}) = \begin{pmatrix} v_1(s, \boldsymbol{p}) \\ v_2(s, \boldsymbol{p}) \end{pmatrix} = \begin{pmatrix} a_0 \\ c_0 \end{pmatrix} + \sum_{k=1}^{K} \begin{pmatrix} a_k & b_k \\ c_k & d_k \end{pmatrix} \cdot \begin{pmatrix} cos(ks) \\ sin(ks) \end{pmatrix}. \tag{3.17}$$

The outline is now completely described by the parameter-vector

$$\boldsymbol{p} = (a_0 \ldots\ a_K, b_1 \ldots b_K, c_0 \ldots c_K, d_1 \ldots d_K)^T. \tag{3.18}$$

The parameters can be easily calculated from the sampling points of the outline $\boldsymbol{q}_0, \boldsymbol{q}_1, \ldots, \boldsymbol{q}_P$ with $\boldsymbol{q}_0 = \boldsymbol{q}_P$. The resulting parametric shape description can be made invariant under similarity transformation by shifting, rotating and scaling according to the actual displacement, orientation and size of the ellipse determined by the first degree terms of the Fourier series. Similarly, the starting point is moved to a canonical position. The theorem of Fourier curve parametrization as developed by [Kuhl and Giardina 1982] will be discussed in the next chapter in more detail.

Fourier parametrization has been extended by [Staib and Duncan 1992b] to 3-D. In this representation each of the three coordinate functions, $v_1(s,t)$, $v_2(s,t)$, and $v_3(s,t)$, are represented by the sum:

$$f(u,v) = \sum_{m=0}^{2K} \sum_{l=0}^{2K} \lambda_{m,l} \begin{pmatrix} a_{m,l} \cos 2\pi mu \cos 2\pi lv + b_{m,l} \sin 2\pi mu \cos 2\pi lv + \\ c_{m,l} \cos 2\pi mu \sin 2\pi lv + d_{m,l} \sin 2\pi mu \sin 2\pi lv \end{pmatrix} \tag{3.19}$$

where:

$$\lambda_{m,l} = \begin{cases} 1 \text{ for } m = 0,\ l = 0 \\ 2 \text{ for } m > 0,\ l = 0 \quad \text{or} \quad m = 0,\ l > 0 \\ 4 \text{ for } m > 0,\ l > 0 \end{cases}$$

Using this parametrization, four 3-D topology classes can be described: tori (closed tubes), open surfaces (with one edge), tubes (open surfaces with two edges) and closed surfaces (no edges, sphere topology).

Spherical Harmonics

offer a more elegant and compact surface representation than 3-D Fourier descriptors. Selecting the spherical harmonic functions (Y_l^m denotes the function of degree l and order m (see [Greiner and Diehl 1986]) as a basis, the coordinate functions $v_1(s,t)$, $v_2(s,t)$, and $v_3(s,t)$ can be written as

$$\boldsymbol{v}(\theta, \phi, \boldsymbol{p}) = \sum_{k=0}^{K} \sum_{m=-k}^{k} \boldsymbol{c}_k^m \, Y_k^m(\theta, \phi) \quad , \tag{3.20}$$

where

$$\boldsymbol{c}_k^m = \begin{pmatrix} c1_k^m \\ c2_k^m \\ c3_k^m \end{pmatrix} \quad . \tag{3.21}$$

The expansion is restricted to the first $K + 1$ terms.

Y_l^m denotes the spherical harmonic function of degree l and order m. 3-D surface models based on this representation has been applied in the present thesis as described in chapter 6, therefore it will be discussed in more soundly in the next chapter.

Wavelets

Wavelets are a mathematical tool for hierarchically decomposing functions. The wavelet analysis procedure is to adopt a wavelet prototype function or *mother wavelet.* Temporal (or spatial) analysis is performed with a contracted, high-frequency version of the mother wavelet, while frequency analysis is performed with a dilated, low-frequency version of the same wavelet. Thus wavelets ere able to combine the local and global properties of the same function.

Comparing wavelet and Fourier transforms, the most interesting dissimilarity between these two kind of transforms is that individual wavelet functions are localized in space, while Fourier sine and cosine are not. Another difference is that wavelet transforms do not have a single set of basis functions like the Fourier transform, which utilizes just the sine and cosine functions. Instead, wavelet transforms have an infinite set of possible basis functions such that depending on the application an optimal mother wavelet can be chosen.

$$\begin{tikzcd}
\boldsymbol{c}^{n} \arrow[r, "\boldsymbol{A}^{n}"] \arrow[dr, "\boldsymbol{B}^{n}"'] & \boldsymbol{c}^{n-1} \arrow[r, "\boldsymbol{A}^{n-1}"] \arrow[dr, "\boldsymbol{B}^{n-1}"'] & \boldsymbol{c}^{n-2} \ \ldots & {} \arrow[r, "\boldsymbol{A}^{1}"] \arrow[dr, "\boldsymbol{B}^{1}"'] & \boldsymbol{c}^{0} \\
 & \boldsymbol{d}^{n-1} & \boldsymbol{d}^{n-2} \ \ldots & & \boldsymbol{d}^{0}
\end{tikzcd}$$

Figure 3.1: The filter bank.

Following [Finkelstein and Salesin 1994], Figure 3.1 explains the composition of wavelet coefficients starting with a discrete signal $\boldsymbol{c}^n$, expressed as a column vector of samples $[c_1^n, \ldots, c_m^n]$. In the first step, multiplying $\boldsymbol{c}^n$ with the matrix $\boldsymbol{A}^n$ a filtered and downscaled version of it, $\boldsymbol{c}^{n-1}$, is computed

$$\boldsymbol{c}^{n-1} = \boldsymbol{A}^n \boldsymbol{c}^n \ , \tag{3.22}$$

where $\boldsymbol{A}^n$ is an $m' \times m$ matrix. If $\boldsymbol{A}^n$ is appropriately chosen, it is possible to capture the lost details as another signal $\boldsymbol{d}^{n-1}$, computed by

$$\boldsymbol{d}^{n-1} = \boldsymbol{B}^n \boldsymbol{c}^n \ , \tag{3.23}$$

where $\boldsymbol{B}^n$ is an $(m - m') \times m$ matrix, which is related to matrix $\boldsymbol{A}^n$. The procedure for splitting $\boldsymbol{c}^n$, into a low-resolution part $\boldsymbol{c}^{n-1}$ and a detail part $\boldsymbol{d}^{n-1}$ can be applied recursively to the new signal $\boldsymbol{c}^{n-1}$. This recursive process is known as a *filter bank*. Since the original signal $\boldsymbol{c}^n$ can be recovered from the sequence $\boldsymbol{c}^0, \boldsymbol{d}^0, \boldsymbol{d}^1, \ldots, \boldsymbol{d}^{n-1}$, this sequence can be thought of as a transform of the original signal, known as a *wavelet transform*. The transform can also be expressed using basis functions, called *scaling functions*. In many applications the piecewise constant Haar basis is used because of its simplicity. However, for applications such as curve editing and deformable models B-splines of degree 2 or 3 are a more appropriate choice due to their continuity (see [Stollnitz *et al.* 1995a, Stollnitz *et al.* 1995b, Stollnitz *et al.* 1996]).

Hybrid models

Vemuri *et al.* in [Vemuri and Radisavljevic 1994] introduce an extension to the superquadrics model by a displacement vector function in a wavelet basis. This *hybrid* *model* have the property of being able smoothly to scale up or down the range of possible deformations and the number of parameter required to characterize them. Thus, the model comprises global and local shape parameters.

Another parametrically deformable model has been introduced by [Terzopoulos and Metaxas 1991] which uses *Deformable Superquadrics* to fit the model to range data. They propose to use superquadrics with 6 parameters as models with global and local deformation properties inherited from ellipsoids and membrane splines.

3.3.2 Elastic deformation of parametric models

To achieve elastic deformation, parametric models are usually handled differently from traditional physics-based deformable models. In the parametric approach curves and surfaces are represented by their descriptors rather than by coordinates of contour points. External and internal energies are expressed in terms of contour parameters. Multi-dimensional nonlinear optimization techniques can be applied to simultaneously adjust shape parameters and minimize the total energy.

Energy minimization using direct optimization

The snake technique as proposed by [Kass *et al.* 1988] tries to find the position of a curve $\boldsymbol{v}(t,\boldsymbol{p})$, which minimizes the energy

$$E(\boldsymbol{p}) = E(\boldsymbol{v}(t,\boldsymbol{p})) = E_I(\boldsymbol{v}(t,\boldsymbol{p})) + E_D(\boldsymbol{v}(t,\boldsymbol{p})) \; . \tag{3.24}$$

By varying $\boldsymbol{p}$, the curve deforms itself to minimize the image energy

$$E_I(\boldsymbol{v}(t,\boldsymbol{p})) = -\int_0^{2\pi} P(\boldsymbol{v}(t,\boldsymbol{p}))\, dt \; , \tag{3.25}$$

searching for an optimal position in the image, described by the potential $P(\boldsymbol{v}(t,\boldsymbol{p}))$. Comparing the equation above with eq. 3.2 one can notice that the integration interval changed to $[0, 2\pi]$. This is due to the fact that parametric curves are usually closed and mapped to the unit sphere, while original snakes were open curves mapped to the linear interval $[0, 1]$. A typical choice takes $P(\boldsymbol{v}(t,\boldsymbol{p})))$ equal to the negative magnitude of the image gradient as given in equation 3.3.

Contrary to traditional physics-based snakes, parametric models based on a limited number of basis functions cannot deform in an entirely free manner. Thus, energy terms such as *tension*, *rigidity*, *mass*, and *energy dissipation* become redundant in this framework. The deformation term $E_D(\boldsymbol{v})$ is called the internal energy of the parametric snake and serves as a regularization force. It restricts the elongation and bending of the snakes, and normally depends on the first and second derivatives of the curve $\boldsymbol{v}(t,\boldsymbol{p})$. To balance between external and internal energy the parameter α is introduced

$$E(\boldsymbol{v}(t,\boldsymbol{p})) = \alpha E_I(\boldsymbol{v}(t,\boldsymbol{p})) + (1-\alpha) E_D(\boldsymbol{v}(t,\boldsymbol{p})) \; . \tag{3.26}$$

A different energy model has been proposed by [Staib and Duncan 1992a] which better suits their Fourier curve and surface parametrizations. The new model makes use of the normal direction of the parametrized curve and of the image gradient to achieve a higher selectivity. Normalizing the image potential by the contour length allows contraction and dilation of curves without affecting the energy function.

$$E_I(\boldsymbol{v}(t,\boldsymbol{p})) = \pm \frac{\displaystyle\int_0^{2\pi} \nabla I(\boldsymbol{v}(t,\boldsymbol{p})) \cdot \dot{\boldsymbol{v}}_\perp(t,\boldsymbol{p})\, dt}{\displaystyle\int_0^{2\pi} \|\dot{\boldsymbol{v}}(t,\boldsymbol{p})\|\, dt} \tag{3.27}$$

The sign of the energy $E_I(\boldsymbol{v})$ will decide about segmenting bright objects on a dark background or vice versa. The polarity of the boundary can be neglected by using the absolute value of the dot product in the integration term.

One has to realize that cutting the Fourier expansion at a finite degree serves already as a regularization, leaving out high frequency variations of the coordinate functions. However, the internal energy cannot be neglected, basically because of two reasons

- Even Fourier models of low degree can generate sharp cusps, and
- the parametrized contour can cross itself.

Whereas self-crossings of the outline are expensive to detect, discontinuities of the tangent can be evaluated from the curve parametrization. At such location the curvature $\kappa(t_0,\boldsymbol{p})$ and its derivative $\dot{\kappa}(t_0,\boldsymbol{p})$ become both infinite. While high curvature of the boundary should not be excluded a priori, the curvature derivative is chosen to indicate discontinuities (see also [Staib and Duncan 1992a]), providing the following expression for the internal energy of the Fourier-snake.

$$E_D(\boldsymbol{v}(t,\boldsymbol{p})) = \int_0^{2\pi} \dot{\kappa}^2(t,\boldsymbol{p}) \cdot \|\dot{\boldsymbol{v}}(t,\boldsymbol{p})\|^2\; dt$$

3.3.3 Optimization techniques

Detailed contours and especially surfaces often require a large number of coefficients to capture and describe them by a given parametrization technique. The dimensionality of the problem is usually of the order of several tens to a few hundreds, with the latter to be more realistic. Not every optimization method is able to handle that many free parameters and it is often unavoidable to invent a new method to a specific problem. Possible choices reach from deterministic techniques like conjugate gradient optimization to stochastic methods such as genetic algorithms.

The objective function to solve is not in general convex, but depends ultimately on the gray level surface shape of the image. Smoothing of the image can be used to avoid getting trapped in a local maximum.

3.4 Representing anatomical knowledge

Many attempts has been done to achieve semi- or fully automatic segmentation of medical datasets by incorporating human expertise. This section gives a short overview of the different model-based techniques.

3.4.1 Deformable atlases

The general goal of *deformable atlases* is to remove structural variation between individuals by matching an atlas image to each individual using non-rigid registration, in order to have a common coordinate system for comparison. Shape differences between the atlas and individual's anatomy are contained in the non-rigid transformation.

The customary classification of registration methods using the terms *rigid* and *non-rigid* is less intuitive, since the rigid group only contains registration by rotation and translation, while *similarity* and *affine transforms* are characterized together with *elastic transforms* as non-rigid. Therefore we propose to use the classification introduced by [Evans *et al.* 1994] which suggests the classes *parametric* and *non-parametric*. Thus, in the parametric group will belong, besides the previous rigid class, the similarity and affine transforms, while the rest of the non-rigid group such as *fluid-registration* and *thin-plate splines* is characterized as non-parametric. In the following we briefly discuss the most relevant approaches to atlas based registration.

Bookstein in [Bookstein 1997] proposed a thin-plate spline interpolation approach to the description of deformations specified by a finite set of corresponding point pairs. This method requires the identification of large number of homologous pairs of anatomic point landmarks throughout the the image because local changes to the spline have global effects. These landmarks must be very reliable, because the thin-plate spline forces these points exactly into correspondence.

Bajcsy and Kovačič in [Bajcsy and Kovacic 1989] present a multiresolution elastic matching process to fit a voxel based anatomical atlas to individual tomography scans. In their approach first the model and the input data are aligned globally. After global alignment the the model is deformed like a piece of rubber simulating the mechanical properties of an isotropic homogeneous body by partial differential equations. Deformation proceeds step-by-step in a coarse-to-fine strategy, increasing the local similarity and global coherence. The driving force of the elastic matching is derived from a local similarity measure of image intensities between the region at some particular position in the object and corresponding regions in the other object. A local search is performed to find the best possible local match for the displacement vector at a given position. The process is carried out at different scales of the images proceeding from large to small gradually improving the overall fit.

Evans *et al.* in [Evans *et al.* 1994] developed an MRI-based probabilistic atlas of neuroanatomy by registering and averaging the brain of about 300 normal young males. The Talairach coordinate system served as the basis of the registration. It is defined by the anterior-posterior commissural (AC-PC) line and a set of perpendiculars to this crossing at the midpoint of AC-PC. All brain volumes has been linearly transformed to this coordinate system and then averaged. The resulting composite MRI-intensity atlas has proven to be useful as a qualitative indicator of local anatomical variability. For the purpose of quantitative analysis the authors constructed a 3-D volume-of-interest atlas by manual outlining of 64 individual brain regions. In this approach, the labeled atlas has been deformed to match an individual image volume by manual identification of corresponding 3-D landmarks, using a thin-plate spline algorithm for non-linear deformation of atlas coordinates.

The non-linear matching method has also been applied to compute an average of 16 brain MRI volume. By this study Evans *et al.* demonstrated the nonlinear variability in brain anatomy, since the non-linear average brain had a sharper appearance than the linear one.

Christensen *et al.* in [Christensen *et al.* 1996] present two physical models for non-rigid registration of the brain. The transformations are constrained to be consistent with the physical properties of deformable elastic solids in the first method and those of viscous fluids in the second. Viscous fluid models are less constraining than elastic models and allow long-distance, nonlinear deformations of small subregions. In these formulations, however, no matter what model is used, the deformed configuration of the atlas is always determined by driving the deformation using only pixel-by-pixel intensity difference between images.

Davatzikos and Prince in [Davatzikos and Prince 1994] propose a method where they first identify the boundary of homologous brain regions in two images to be registered, and establish a one-to-one mapping between them. Based on this mapping, they deform the boundaries in one image into those in the other image. The rest of the image is deformed by solving the equations describing the deformation of an elastic body using the boundary deformations as input. In this approach, although the mapping may be accurate on the boundary, the farther away the structure is from the boundary, the more error there is, because only information from object boundaries is used for registration.

Wang and Staib in [Wang and Staib 1998] incorporate statistical shape information into physical model based registration. The algorithm they employ is based on a physical model, a gray level similarity measure and a consistency measure between corresponding boundary points. The statistical shape information is embedded in the boundary finding with correspondence process applied to the study. This method uses statistical point models with shape, and shape variation, generated from sets of examples by principal component analysis of the covariance matrix.

3.4.2 Mechanical modes of variation

Martin *et al.* in [Martin *et al.* 1994] proposed another framework for analyzing the shape of structures within the human brain. They developed a mathematical model to describe the deformation of any brain structure whose shape is affected by both gross and detailed physical processes. Their model decomposes the total shape deformation into physical modes of variation obtained from finite element analysis, and experimental modes of variation obtained from sample data using principal component analysis.

The authors use this mathematical model to classify diseases that affect the shape of the ventricular system of the brain. They find that because ventricular shape is affected not only by pathology but also by overall brain shape, disease discrimination can be difficult. Martin *et al.* claim that by modeling the brain's elastic properties, they are able to compensate for some of the nonpathological modes of ventricular shape variation. This allows them to experimentally characterize and separate modes of variation that are indicative of disease processes.

3.4.3 Rule-based expert systems

Although they are not closely related to model-based segmentation techniques, rule-based expert systems represent an alternative way to capture human expertise. [Raya 1990] describes a two stage rule-based system to segment 3-D MRIs of the brain. The proposed technique applies in the first stage domain-independent knowledge of segmentation, grouping and shape analysis, while the second stage embodies domain-dependent knowledge of anatomy and physiology. A similar procedure has been introduced by [Stansfield 1986] to segment coronary vessels from digital angiograms of the chest.

3.4.4 Statistical shape models

Shapes of anatomical structures show off a certain variability over different individuals. The most plausible way to capture this variability is to use distributions. Statistically deformable models has been introduced by several researchers during the last few years, and they are proved to be powerful in capturing expert knowledge about object shape in model based segmentation techniques. Statistical modeling requires knowledge of the a priori probability $p(\boldsymbol{p})$ of the model's parameter vector $\boldsymbol{p}$ and sometimes also the probability of $\boldsymbol{p}$ conditioned on the sensed data, i.e., $p(\boldsymbol{p}|\boldsymbol{D})$. Unfortunately, in surface-modeling applications we rarely if ever have this kind of complete knowledge about the probabilistic structure of the problem. Typically, we have a number of representative samples of what we want to model. These samples are used to estimate the unknown probability which in turn are used as if they were the true values. In the following three different approaches will be presented, among which the last had larger influence to the present work, therefore we devote more attention to it.

Solving the correspondence problem

Estimating the distribution $p(\boldsymbol{p})$ of given individual shapes $\boldsymbol{p}_i$ requires that the descriptors of the individuals can be directly compared, in other words, that there is correspondence between them. This means, e.g., that the expression $(\boldsymbol{p}_1 + \boldsymbol{p}_2)/2$ yields the descriptors of the average shape. That is, correspondence or homology of $\boldsymbol{p}_i$-s is essential to derive statistics on them.

In order to determine correspondence we have two distinct classes of methods. These being the discrete and continuous cases. Considering the first class, when the contour has to be parametrized by a set of discrete points stored in $\boldsymbol{p}$, correspondence can be established manually by annotating landmarks along the contour of the objects. Given objects of the same class and the homologies between them, the point sets can be aligned using Kendall's Procrustes algorithm (see [Bookstein 1997]). This involves centering and scaling each point set such that the sums of squared distances of each point set is unity and then evaluating the Euclidean distance between the point sets after superimposing the two forms.

However, when the point sets are automatically derived from images, the difficult problem of establishing correspondence and rejecting non-homologies and outliers remains. Rangarajan *et al.* [Rangarajan *et al.* 1997] developed their *Softassign Procrustes Matching Algorithm* to handle these cases. They formulate the matching algorithm as two interlocking optimimization problems – one on the rotation between the two point sets and the other on the point-to-point feature correspondence. The algorithm iteratively alternates between solving for the rotation and the correspondence until a minimum is achieved.

We discuss now the second case, when the contours are given as continuous functions. [Tagare 1997] defines a correspondence between two closed curves $c(\boldsymbol{p}_1)$ and $c(\boldsymbol{p}_2)$ as a subset of the product $c(\boldsymbol{p}_1) \times c(\boldsymbol{p}_2)$. It is shown that the product space $c(\boldsymbol{p}_1) \times c(\boldsymbol{p}_2)$ is the surface of a torus and that corresponding point pairs on $c(\boldsymbol{p}_1)$ and $c(\boldsymbol{p}_2)$ define a closed curve on the surface of it which curve is nondecreasing with respect to arc length. Using this property a dissimilarity function between the shapes $c(\boldsymbol{p}_1)$ and $c(\boldsymbol{p}_2)$ is defined and the correspondence is found by minimizing this function.

Correspondence can be approximated between curves or surfaces using length- or area-preserving parametrizations, respectively. Taking a continuous surface as example, global parametrizations map each point of the surface onto parameter values. Thus, surfaces with simple sphere-topology can be mapped onto the spherical parameters (θ, ϕ), which establishes one-to-one correspondences between surface points. Assuming a uniform parametrization, meaning that equal areas on the object surface are assigned to equal areas on the corresponding sphere, and additionally assuming a standardized position of the starting point of the parametrization achieved by normalization, surface descriptors can be directly compared. This method, however, is only able to establish correspondence between highly similar contours and surfaces and it clearly fails if one part of the object is missing or there is an additional part attached to it.

Kotcheff and Taylor in [Kotcheff and Taylor 1997] describe a technique to enhance correspondence among a larger set of shape representations obtained, e.g. from manual annotation of object contours or from arc-length parametrizations. Based on a training set of similar objects deformation modes of the shape can be determined using principal component analysis of the covariance matrix, as shown in the next section. The authors construct an objective function that prefers concentrating the variance into a few deformation modes with large variances, this being a measure for higher correspondence. The objective function is minimized using a genetic algorithm which iteratively reparametrizes the contours.

Not belonging in any of the categories above, but important to mention is the method proposed by [Thirion and A. 1996] which uses crest lines to find correspondence between volumetric images. Crest lines are defined as lines of maximum or minimum of the first or second principal curvatures of an isointensity surface in the 3-D image, i.e. it can be either a ridge or a valley. Their method is based on the implicit representation of the surface, and the differential values of the isointensity surfaces are directly computed from the voxel image, without extracting any surface first. Possible applications are registration of volumetric images or, if combined with surface models, the computation of average shapes.

Applications

Vemuri and Radisavljevic in [Vemuri and Radisavljevic 1994] present a real 3-D model-fitting process in a probabilistic framework which applies prior distributions on the vector of geometric parameters that are being estimated. Their solution comprises a training phase and an operational phase. In the training phase human supervision via interaction and manipulation is needed to provide adequate training samples, while in the operation phase the model-fitting process becomes fully automatic and relies solely on statistics accumulated in the prior model.

Their hybrid 3-D shape model combines a global *superquadrics* parametrization with a local *wavelet* representation, already surveyed in the previous section. The training process collects statistics on the global parameters of the model. For the local parameters the energy, which governs the deformation of the model away from its natural form, is converted into a probability distribution over expected shapes. This is achieved via a technique of statistical mechanics, which converses energies into probabilities using Gibbs distribution:

$$p(\boldsymbol{p}) = \frac{1}{Z_p}\exp(-E_p(\boldsymbol{p})), \tag{3.28}$$

where E_p is the internal smoothness energy restoring a model toward a rest state and Z_p is a normalizing constant. The internal energy E_p of the model given by a quadratic form,

$$E_p = 1/2[(\boldsymbol{p}-\bar{\boldsymbol{p}})^T\boldsymbol{K}_p(\boldsymbol{p}-\bar{\boldsymbol{p}})]\,, \tag{3.29}$$

where $\bar{\boldsymbol{p}}$ the rest state of their model, and the matrix $\boldsymbol{K}_p$ corresponds to the stiffness of deformation. The equation turns the internal energy into a probability distribution over expected shapes, with lower-energy shapes being more likely. Similarly, the external energy is converted into the probability distribution:

$$p(\boldsymbol{D}|\boldsymbol{p}) = \frac{1}{Z_D}\exp(-E_D(\boldsymbol{D},\boldsymbol{p})), \tag{3.30}$$

where $E_D(\boldsymbol{D},\boldsymbol{p})$ is the edge-based potential energy synthesized from image data. Combining the two distributions (from equations 3.28 and 3.30, respectively) using Bayes' rule, we obtain the posterior distribution:

$$p(\boldsymbol{p}|\boldsymbol{D}) = \frac{p(\boldsymbol{D}|\boldsymbol{p})\;p(\boldsymbol{p})}{p(\boldsymbol{D})} = \frac{1}{Z}\exp(-E_p(\boldsymbol{p}) - E_D(\boldsymbol{D},\boldsymbol{p})). \tag{3.31}$$

Computing the maximum *a posteriori* estimate

$$\boldsymbol{p}_{map} = \arg\max_{\boldsymbol{p}} p(\boldsymbol{p}|\boldsymbol{D}) = \arg\max_{\boldsymbol{p}} \frac{p(\boldsymbol{D}|\boldsymbol{p})\;p(\boldsymbol{p})}{p(\boldsymbol{D})}, \tag{3.32}$$

i.e., the value of $\boldsymbol{p}$ that maximizes the conditional probability $p(\boldsymbol{p}|\boldsymbol{D})$, provides the same result as finding the minimum energy configuration of the physically based model. However, the advantage of the probabilistic framework is that it allows for incorporating a prior model. The authors trained their prior models of a gyrus and a hippocampus during a *supervised learning procedure* on several brain MR images and present segmentation result for these two objects.

Staib and Duncan in [Staib and Duncan 1996] apply a similar probabilistic formulation to fit their Fourier-based globally parametrized model (see Section 3.3.1) to image data. However, the way of transforming energies into probabilities is different. The expression 3.32 can be simplified by taking the logarithm and eliminating $p(\boldsymbol{D})$, the prior probability for image data, which is equal for all $\boldsymbol{p}$:

$$\boldsymbol{p}_{map} = \arg\max_{\boldsymbol{p}} [\ln p(\boldsymbol{p}) + \ln p(\boldsymbol{D}|\boldsymbol{p})] = \arg\max_{\boldsymbol{p}} M(\boldsymbol{p}, \boldsymbol{D}) \tag{3.33}$$

Modeling the image $\boldsymbol{D}$ as a noise-corrupted version of an image template, $t_{\boldsymbol{p}}$, and assuming a Gaussian additive noise model with zero mean and standard deviation σ_n, $M(\boldsymbol{p}, \boldsymbol{D})$ can be written as

$$M(\boldsymbol{p}, \boldsymbol{D}) = \ln p(\boldsymbol{p}) + \frac{k}{2\sigma_n^2} \int\int_{\mathcal{A}} [\boldsymbol{D}(\boldsymbol{v}(\boldsymbol{p}))] dA, \tag{3.34}$$

where dA is a surface element on the surface $\mathcal{A}$. The authors use a conjugate gradient method to maximize $M(\boldsymbol{p}, \boldsymbol{D})$.

This technique has been applied to a variety of synthetic and radiologic images. In the synthetic example the model's probability distribution was simulated and the initial surface was positioned roughly at the target object. Another example illustrates the segmentation of a dog's heart from 3-D MRI data. A prior model, $p(\boldsymbol{p})$, was determined from a sample of four images which were delineated manually with a closed surface. The prior probability distribution, characterized by the mean and standard deviation, was determined from the parameters of the sample. The mean of the distribution was used as the initial position for boundary finding in another image. Compared with contours hand-traced by an expert, the reported average error was 0.65 pixels in both cases.

Although the approaches above give a statistical framework to the model fitting process, the dimensionality of the problem remains the same as before. Optimization algorithms used for energy minimization still have to handle as many parameters as the shape description uses to represent the object of interest. Especially in case of complex 3-D shapes high number of parameters are needed to describe the form with sufficient amount of details. Energy minimization becomes a highly nonlinear multi-parametric optimization problem requiring computationally very intensive techniques. The question arises, if and how the dimensionality of the problem can be reduced. [Cootes *et al.* 1995a] describe an elegant solution. Their approach makes the assumption that $p(\boldsymbol{p})$ can be approximated by a joint Gaussian distribution and reduces the number of parameters using decorrelation.

Point Distribution Models has been introduced by the authors to learn patterns of variability from a training set of correctly annotated images. Image annotation is performed manually by placing manually a set of points representing the boundary of an object, internal features, or external ones. Points are placed in the same way on each of a training set of examples of the object marking significant features such as (1) curvature extrema along the boundary, (2) application-dependent landmarks, and (3) other points which can be interpolated from points of type 1 or 2; for instance, points placed at equal distances between two type 1 landmarks.

In order to be able to compare equivalent points from different shapes, they must be aligned. The required alignment is achieved by scaling, rotating, and translating the training shapes so that they correspond as closely as possible. The authors use a modification of the Procrustes method (see [Bookstein 1997]) to minimize a weighted sum of squares of distances between equivalent points of different shapes.

The ith shape in the set is represented by a vector $\boldsymbol{x}_i$ of n 2-D points:

$$\boldsymbol{x}_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \ldots, x_{ik}, y_{ik}, \ldots, x_{in-1}, y_{in-1})^T \tag{3.35}$$

A set of N example shapes can be represented by a cloud of N points in $2n$ dimensional space. The authors assume that these point build an isolated region within the space, which they call *allowable shape domain*. Making the assumption that the cloud is approximately ellipsoidal, the principal axes of a $2n$-D ellipsoid fitted to the data can be calculated by applying principal component analysis. Given a set of N aligned shapes, the mean shape $\bar{\boldsymbol{x}}$ is calculated using

$$\bar{\boldsymbol{x}} = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{x}_i \ , \tag{3.36}$$

and for each shape in the training set its deviation from the mean, $d\boldsymbol{x}_i$, is determined

$$d\boldsymbol{x}_i = \boldsymbol{x}_i - \bar{\boldsymbol{x}} \ . \tag{3.37}$$

Using the deviations a $2n \times 2n$ covariance matrix can be calculated:

$$\boldsymbol{S} = \frac{1}{N-1}\sum_{i=1}^{N} d\boldsymbol{x}_i d\boldsymbol{x}_i^T \tag{3.38}$$

The principal axes of the ellipsoid, defining the modes of variation of the shape, are described by $\boldsymbol{p}_k$, $(k = 1, \ldots, 2n)$, the unit eigenvectors of $\boldsymbol{S}$ such that

$$\boldsymbol{S}\boldsymbol{p}_k = \lambda_k \boldsymbol{p}_k \tag{3.39}$$

The eigenvectors of the covariance matrix corresponding to the largest eigenvalues describe the longest axes of the ellipsoid, and so the most significant modes of variation in the variables used to derive the covariance matrix. Most of the variation can usually be explained by a small number of modes, t. A shape in the training set can be approximated using the mean shape and a weighted sum of the largest t eigenmodes:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{P}\boldsymbol{b}, \tag{3.40}$$

where $\boldsymbol{P} = (\boldsymbol{p}_1 \boldsymbol{p}_2 \ldots \boldsymbol{p}_t)$ is the matrix of the first t eigenvectors, and $\boldsymbol{b} = (b_1 b_2 \ldots b_t)$ is a vector of weights. The above equation describe the way of generating new examples of the shapes by varying the weights within suitable limits, typically of the order of

$$-3\sqrt{\lambda_k} \leq b_k \leq 3\sqrt{\lambda_k}, \tag{3.41}$$

since most of the population lies within three standard deviation of the mean.

Assuming a nearly Gaussian distribution of the single contour points, this method reduces the information required to represent a class of similar shapes to a minimum. The average shape and the eigenmodes can be thought of as basis vectors of a global parametrization and the weights $\boldsymbol{b}$ as the actual shape descriptors.

Cootes *et al.* applied the technique of *point distribution models* to determine the modes of variation in several man-made and biological objects, such as a set of resistor shape, a set of heart ventricle shapes, and a set of hand shapes. These models has been embedded into a *snake*-like segmentation tool which iteratively enhances the initial model's pose and shape parameters while taking shape constraints into consideration. The authors call their technique *Active Shape Models* to make a clear distinction to *Active Contour Models* or *Snakes* without shape information behind them.

Although global optimization could also be applied to *point distribution models*, [Cootes *et al.* 1994a] propose a technique based on local search near the contour. They extend their contour model by information on its gray level environment using sampled profiles normal to the boundary. For each of these profiles a local search is performed to find a nearby position in the image which better matches with the sampled gray levels. Their technique computes in each iteration step a set of adjustments to model points along a normal to the boundary toward the strongest image edge within a certain distance. These adjustments are not in general consistent with the shape model, thus the shape is updated with respect to the model constraints, hereby preventing the model's shape to leave the *allowable shape domain*. The model is updated so that only deformations, being consistent with the deformation modes represented in the original training set, are allowed. The procedure is repeated iteratively until no significant changes result.

The transition, however, to 3-D *Point Distribution Models* of the same kind is hampered by the way contour point has to be manually identified. While the manual annotation of a 2-D contour can be carried out in a straightforward way, this procedure is rather hard to perform for 3-D surfaces in order to establish correspondences between them. Thus, to be able to build *Active Shape Models* in 3-D annotation has to be done automatically. [Caunce and Taylor 1998] propose and compare methods, based on the Iterative Closest Point Algorithm, to find corresponding points in sparse 3-D point sets. The algorithm is similar to the Softassign Procrustes Matching Algorithm previously discussed in this section. They generated point sets by determining the location of the sulci in 16 MRI volumes. Taken the sparsest one as a reference these point sets containing points between 4000 and 6000 were corresponded. The authors performed the statistical analysis described above to compute a *Point Distribution Model* of the cortical sulci and found the resulting modes of variation represent meaningful shape distortions. This method has the advantage that it makes no assumption about the topology of the shape the points belong to.

A less general but, regarding segmentation applications, more beneficial solution to the problem of establishing correspondence between 3-D shapes is provided by global parametrization and will be outlined in chapter 6 of the present thesis.

3.5 Discussion

Summarizing the three techniques for statistical model-based segmentation introduced above, Cootes and Taylor, provide a complete solution for 2-D with their *Point Distribution Models* (PDM). However, the manual annotation of 2-D contours, they apply in order to establish correspondence, prevents a straightforward generalization of their technique to 3-D. A recently published technique from [Caunce and Taylor 1998] is able to replace the manual procedure by an automated one, which identifies correspondences between arbitrary point sets in 3-D. Using this method, 3-D PDMs can be built. In absence of topology, however, these models are difficult used in segmentation processes. Furthermore, to reduce variance caused by sampling and only capture variance in shape, the method requires high number of points, uniformly distributed over the surface of objects of interest. This requirement is hard to fulfill, especially in case of manually segmented voxel surfaces.

The techniques proposed by [Vemuri and Radisavljevic 1994] and [Staib and Duncan 1996] make use of the properties of globally parametrized surfaces in order to approximate correspondence between them. The way, however, the two cited works obtain surface parametrization by user interaction, suggests that the requirement of perfect alignment cannot be fulfilled. In consequence of this, the statistics derived from them can represent an average shape and deformation modes only in a limited way. Furthermore, the *hybrid model* applied by Vemuri and Radisavljevic, combining *global* and *local* parametrizations, is suboptimal for statistical modeling, since only its global part can be used for the computations, which being a superquadrics does not allow of a detailed shape description.

Statistical modeling requires an adequate number of samples, to derive statistically robust estimations of distributions of model parameters. Therefore, a training set containing as few as four samples, e.g., as used by Staib and Duncan to compute their model of a dog's heart, is certainly too small for this purpose. On the other hand, numerous medical studies have been carried out lately, resulting in many high quality expert segmentations of 3-dimensional organs. The results of these are usually stored as binary volume data, and if they could be directly converted into parametric surface models, the problem of too small training sets would be easily solved. That is, instead of developing a new and proprietary interactive segmentation tools for each type of surface parametrization, existing techniques could be used.

We aim at building a model based automatic segmentation tool to process 3-dimensional medical data sets of healthy human neuroanatomy. In the first stage a statistical atlas of neurology has to be set up. To model two- and three-dimensional organs we have chosen global parametrizations by Fourier and spherical harmonics basis functions in place of parametrizations by point sets. Since we do not intend to handle pathology and healthy organs have high similarity over individuals, global parametrizations provide sufficiently good homology among them. The spherical harmonics parametrization has been preferred as they, in contrast to superquadrics, imposes no predetermined symmetries upon the object surface and has no preferred directions in space. It is complete; no additional formalisms are needed to rotate, translate, taper, bend, twist or locally deform the model, as the surface of any simply connected object can be represented to any degree of detail in the simple yet comprehensive formalism. In the next two chapters we show that manually segmented binary shapes of simple topology can be easily converted into Fourier or spherical harmonics descriptors. To derive a statistical model, we estimate the distribution of shape descriptors and in order to reduce reduce the dimensionality of the models, we borrow the technique of eigendeformations proposed by [Cootes *et al.* 1995a]. In a second stage the statistical atlas is matched to individual data sets, where the underlying parametrization provides not only a binary voxel segmentation but also shape characterization of organs.

4

Parametrization of closed curves and surfaces

Parametrically deformable models give rise to the question of obtaining parametrical descriptions of given pixel or voxel based object contours or surfaces, as they usually result from manual segmentation.

In the following the term "*parametrization*" will be used in two different senses. On the one hand, the process which maps one parameter value (s) to each point of a curve or two parameter values (s,t) to each point of a surface is called *curve* or *surface parametrization.*

$$\boldsymbol{v}(s) = \begin{pmatrix} v_1(s) \\ v_2(s) \end{pmatrix} = \begin{pmatrix} x(s) \\ y(s) \end{pmatrix} \tag{4.1}$$

in 2-D, or

$$\boldsymbol{v}(s,t) = \begin{pmatrix} v_1(s,t) \\ v_2(s,t) \\ v_3(s,t) \end{pmatrix} = \begin{pmatrix} x(s,t) \\ y(s,t) \\ z(s,t) \end{pmatrix} \tag{4.2}$$

in 3-D. On the other hand, these mappings can be used to give a mathematical representation of the contour by the coordinate functions in $\boldsymbol{v}(s)$ and $\boldsymbol{v}(s,t)$. As it has been shown in the previous chapter, the coordinate functions also depend on parameters (descriptors), e.g., the weights of some basis functions. Confusingly, the procedure to compute these descriptors is also called *parametrization.* To make a clear distinction, in the following *curve* or *surface parametrization* will refer to the mapping procedure and simply *parametrization* to the computation of shape descriptors.

The computation of the *parametrization* of a given binary object usually obtained from manual segmentation is a three-stage process. In the first stage, the contour or surface of the object is converted into an abstract data structure called *chain* or *crack code* in 2-D and *cuberille notion* in 3-D describing the shape as a set of line segments or a set of rectangular patches. Considering only the case of simply connected objects, in the second stage, based on these notations, curves are mapped on the unit circle Ω_2, while surfaces on the unit sphere Ω_3. Arc length parametrization guarantees constant speed for $\boldsymbol{v}(s)$ along the curve, while the equivalent criteria of area preservation for surfaces can be at most approximately fulfilled. In the third stage a mathematical representation of the object is computed by deriving the parameters (descriptors) of the chosen functions, e.g., Fourier basis or superquadrics.

4.1 Fourier coefficients of a closed curve

Kuhl and Giardina in [Kuhl and Giardina 1982] presented an algorithm to efficiently compute the Fourier coefficients of a 2-D contour. Their elegant approach does not require integration or the use of fast Fourier transform techniques. The algorithm takes the chain code of the contour as its input, which is easily obtained from a pixel-based image. The resulting Fourier descriptors can be made invariant with rotation, dilatation and translation of the contour, and also with the starting point of the parametrization on the contour, without losing information about the shape.

4.1.1 Obtaining the code of a contour

The *chain* code first described by Freeman approximates a contour by a sequence of piecewise linear fits that consist of eight standardized line segments. The code of a contour is then the chain V of length K, $V = a_1 a_2 a_3 \dots a_K$, where each link a_i is an integer between 0 and 7 oriented in the direction $(\pi/4)a_i$ and of length 1 or $\sqrt{2}$ depending, respectively, on whether a_i is even or odd. Figure 4.1**b** illustrates the chain code of the simple object shown in Figure 4.1**a**. One can notice that per definition the chain coded contour lies by half voxel inside the object introducing a systematic error of 0.5 *voxel edge* into the description. To generate it the definition of "object" and "background" in the image is required. *Crack* code has been introduced to overcome these limitations. It is defined to lye exactly on the border of background and object (Figure 4.1, image **c**) and can be simplified by replacing two segments building a corner by a diagonal one (image **d**). While the simplification considerably shortens V (see Table 4.1), it also introduces error into the representation. The choice of a coding technique depend on the application, chain codes are most appropriate in case of area-quantized images. In our application to describe hand segmented organs, we have used unsimplified crack codes since these are closest to the original contour.

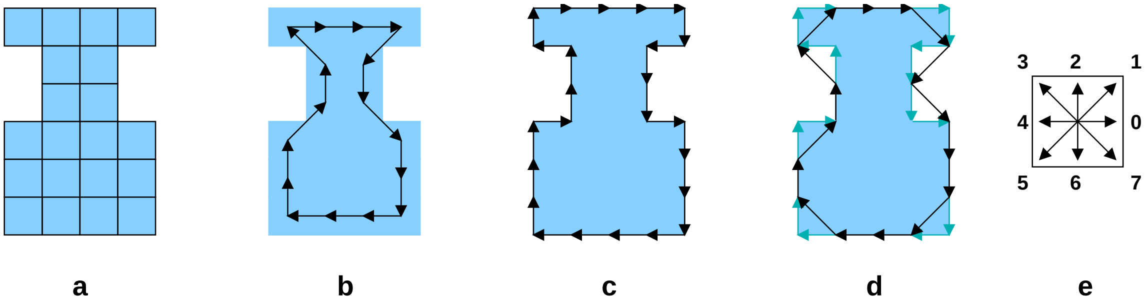


Figure 4.1: Different coding techniques of the simple object shown in **a**. The Freeman code (**b**) is defined to connect the middle points of the border pixels of the object. Images **c** and **d** illustrate the crack code and its simplification, respectively, while **e** shows the assignment of code numbers to directions.

$$\begin{aligned} V_{\text{chain}} &= 0005676644422123 \\ V_{\text{crack}} &= 0000646606664444222202242 \\ V_{\text{simplecrack}} &= 0075766544321231 \end{aligned}$$

Table 4.1: Code sequences corresponding to Figures 4.1(**b**,**c**, and **d**,)

4.1.2 Computing Fourier descriptors of a closed polygonal curve

The contour of a simply connected 2-D region is a closed curve which can be mapped onto the parameter space defined by the unit circle Ω_2. The contour is then parametrized by the polar coordinate ϕ, such that the position $\boldsymbol{v}(u)$ on the contour becomes a periodic function of ϕ. This suggests the harmonic Fourier basis being the preferred set of basis functions for the object description. Curves are usually parametrized with their arclength s measured from an arbitrarily chosen starting point on the contour. Thus, when the whole curve has the length L, the point of arclength s from the starting point is assigned to the parameter $\phi = \frac{2\pi s}{L}$. The mathematical treatment is simpler in the complex plane, i.e. when the parameter space is the circle $U = \{u \in \mathbb{C} \mid u^*u = 1\}$ and the mapping $z : U \mapsto \mathbb{C}$ represents the curve. (A full deduction of the complex formulas presented here can be found in Christian Brechbühler's Phd. thesis [Brechbühler-Miškuv 1995].)

The relation between s, ϕ, and u is

$$u = u_0 + ju_1 = e^{j\phi} = e^{\frac{2\pi js}{L}} \tag{4.3}$$

Fourier expansion

In the complex notation the vector function given in equation 4.1 transforms to the complex valued function $z(u)$. $z(u)$ is represented as a series of complex exponentials.

$$z(u) = \sum_{n=-\inf}^{\inf} z_n u^n \tag{4.4}$$

where the complex coefficient z_n can be expressed in polar notation, i.e.

$$z_n = r_n e^{j\psi_n} \ , \tag{4.5}$$

with $r_n \in \mathbb{R}, r_n \geq 0$, and $\psi \in \mathbb{R}$.

Determining the coefficients

The calculation of z_n for a given contour $z(u)$ is of practical interest. This is given by the formula:

$$z_n = \frac{1}{2\pi} \oint z(u)^{-n} |du| \tag{4.6}$$

In most applications, $z(u)$ describes a polygon and often we are not interested in the center of gravity of the contour. In this case, it is simplest to start from the derivative $\frac{d}{|du|}$ of 4.4 and derive another formula for z_n:

$$z_n = \frac{1}{jn} \oint z'(u)^{-n} |du| \tag{4.7}$$

Computing the Fourier descriptor for a closed polygon

When the curve is given as a polygon (e.g. by its Freeman-code), it can be expressed as a sum of straight line pieces. In the same way the integral 4.7 brakes up into a sum of partial integrals. The M sample points $z(u_k), k = 0 \ldots M$ of the Freeman curve define the transitions between the partial integrals. The arclength parametrization of the curve implies that the point $z(u)$ traverses the curve with constant speed; on each individual straight line piece this leads to

$$z'(u) = \frac{\Delta z}{|\Delta u|} = \frac{L \Delta z}{2\pi |\Delta u|} = \text{const} \tag{4.8}$$

This relation significantly simplifies the expression for z_n since the partial integrals of straight line segments can be substituted for their solution.

$$z_n = -\frac{1}{n^2} \sum_{k=0}^{M-1} z'[k] u^{-n} \,|_{u_k}^{u_{k+1}} \tag{4.9}$$

To evaluate this sum, it is enough to calculate one complex exponential for each term. As the lower bound of each term is equal to the upper bound of the previous term, the value of u^{-n} can be reused.

In the case of Freeman-code, Δz can only take eight different values, namely 1, $1+j$, j, $-1+j$, -1, $-1-j$, $-j$, and $1-j$. The same holds for z': these as well are completely determined by the code $0 \ldots 7$.

4.1.3 Normalization in object and parameter space

The normalization proposed by [Kuhl and Giardina 1982] is based on the ellipse defined by the 1st order Fourier descriptors and is carried out both in object and in parameter space. Normalization in object space effects the curve's position, orientation and size, while that in parameter space applies to the curve parametrization behind it. After normalization in object space the center of the 1st order ellipse of a normalized contour concurs with the coordinate origin, its main axis overlaps with the x-axis of the coordinate system has the length of 1. In parameter space the starting point of the parametrization is moved to a standard position defined by the crossing of the 1st order ellipse and its main axis.

The authors derive an error bound on the Fourier approximation which can be used to determine the number of harmonics required by a desired accuracy. Furthermore, they describe a classification and recognition procedure that is applicable to classes of objects which may occur in different orientation, sizes and translation. The following section briefly summarizes the mathematics of Fourier descriptors.

Dependence on starting point

To make the descriptors independent on the starting point of the parametrization, this can be shifted to a standard position, e.g. to the tip of the ellipse defined by the 1^{st} order Fourier descriptors. This can be thought of as a rotation in parameter space U given by the unit circle. The transformation is defined by

$$z_n \mid^V = z_n e^{jn\theta} \ , \tag{4.10}$$

where the notation $\mid^V$ marks the coefficients resulting from shifting by angle θ.

Dependence on rotational position

In the complex notation, rotation in object space by angle ψ is simply a multiplication by $e^-j\psi$. Applying it to 4.4 immediately reveals the coefficients of the rotated object.

$$z_n \mid^R = z_n e^{j\psi} \tag{4.11}$$

To achieve a standardized position of the curve, it is rotated in a way that its first ellipse's main axis matches the horizontal (real) coordinate axis.

Scale dependence

Scaling the objects by factor α leads to multiplying its coefficients by the same factor:

$$z_n \mid^S = \alpha z_n \tag{4.12}$$

The scaling factor α is usually set to normalize the half major axis to unity, meaning

$$\alpha = \frac{1}{|z_1| + |z_{-1}|} = \frac{1}{r_1 + r_{-1}} \tag{4.13}$$

Invariant Fourier descriptors

Ignoring z_0, that is setting $z_0 \mid^T = 0$, achieves translation invariance. Summing up all standardizations; the invariant coefficients are denoted $\tilde{z}_n$:

$$z_n \mid^{V,R,S,T} = \tilde{z}_n = z_n \frac{e^{j(n\theta - \psi)}}{r_1 + r_{-1}} \tag{4.14}$$

$$\tilde{z}_0 = 0 \tag{4.15}$$

4.1.4 Relations to real valued notation

In the complex notation of Fourier coefficients real and imaginary parts of z_n correspond to the x and y coordinates

$$\begin{pmatrix} x \\ y \end{pmatrix}_n = \begin{pmatrix} a_n & b_n \\ c_n & d_n \end{pmatrix} \begin{pmatrix} \sin \frac{2\pi ns}{L} \\ \cos \frac{2\pi ns}{L} \end{pmatrix} , \tag{4.16}$$

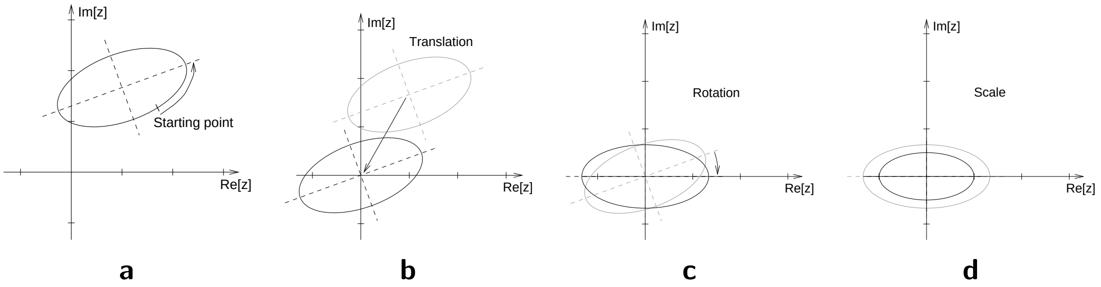


Figure 4.2: Normalization steps of Fourier coefficients; shifting of the starting point to the tip of the ellipse (**a**), moving the center of gravity to the coordinate origin (**b**), rotating the main axis of the ellipse to the real axis (**c**), and finally scaling the half major axis to unity (**d**).

where the real valued coefficients a_n, b_n, c_n, and d_n are defined as follows.

$$\begin{aligned} a_n &= \mathrm{Re} z_n + \mathrm{Re} z_{-n} && (4.17)\\ b_n &= -\mathrm{Im} z_n + \mathrm{Im} z_{-n} && (4.18)\\ c_n &= \mathrm{Im} z_n + \mathrm{Im} z_{-n} && (4.19)\\ d_n &= \mathrm{Re} z_n - \mathrm{Re} z_{-n} && (4.20)\\ & && (4.21) \end{aligned}$$

4.2 Description of surfaces by spherical harmonic functions

The problem of finding a similarly homogeneous parametrization of arbitrarily shaped surfaces proved to be more difficult. [Brechbühler *et al.* 1995] introduced only recently a new surface parametrization technique which can be considered as a complete generalization of Kuhl and Giardina's technique to three dimensions.

Earlier methods for mapping an object surface onto a sphere have been limited to represent only star-shaped or convex objects, as they start from an initial radial surface function $r(\theta, \phi)$. [Staib and Duncan 1992b] discuss the use of a parameter space with torus topology, which can be deformed into a tube by squeezing the torus cross-section to a thin ribbon. Closed surfaces are obtained by considering tubes whose ends close up to a point. This approach illustrates some principal difficulties which can also be found in other parametrization techniques.

- Warping a torus to a closed surface poses the problem that the *parameters have different rules.* One parameter defines a kind of spine along which cross-sections are stacked up.
- Squeezing a circle to line results in a *nonhomogeneous distribution of parameters* on the object surface.
- Warping a torus to a tube and finally to a closed surface causes the parametrization does *not* result in a *one-to-one mapping* of surface points to parameters.

The new method allows a uniform mapping of an object surface into a two-coordinate space with spherical topology. As a mapping of convoluted surface structures onto the surface of a sphere introduces distortions, optimization of the distribution of nodes in parameter space becomes necessary. This problem is solved by nonlinear optimization.

Parametrized surfaces can be expanded into spherical harmonics, hierarchically describing shape properties by spatial frequency constituents. Similarly to Fourier descriptors, spherical harmonic descriptors can also be made invariant to translation, rotation and scaling.

A summary of the surface parametrization procedure follows which is described in detail in [Brechbühler *et al.* 1992, Brechbühler *et al.* 1995].

4.2.1 The surface data structure

Medical CT or MRI images are examples of volumetric data. For each cuboidal cell (volume element or voxel) in a certain volume there are one or more measurements. When segmentation succeeds, one anatomical unit can be characterized by a binary data volume, in which every voxel contains either 1, which means it belongs to the unit, or 0, meaning it is in the background. The object is then the set of “1” voxels and can be pictured as a collection of small cubes, adopting the *cuberille notion* [Herman and Liu 1979]. The surface of a voxel object is a set of unit squares, all parallel to one of the three coordinate planes yz, xz, or xy. The edges and vertices that bound the faces are also parts of the surface, which is represented as a data structure that reflects geometry as well as neighborhood relations.

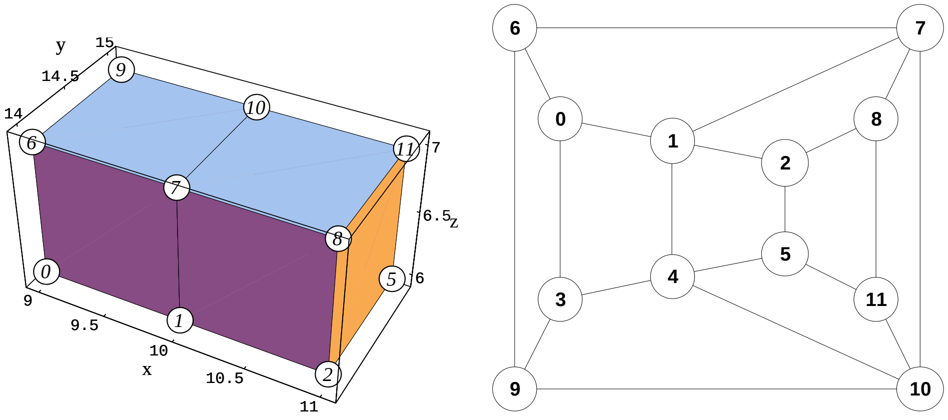


Figure 4.3: Left: A two-voxel example object illustrating the surface data structure, which focuses on the vertices. The numbers of the vertices are shown in circles (vertices 3 and 4 are hidden). The data structure entry for a vertex represents its Cartesian coordinates and a list of neighboring nodes. The entry for vertex 7 is $\{\{x_0 = 10,\ x_1 = 14,\ x_2 = 7\},\ \textit{neighbors}=\{6, 0, 1, 2, 8, 11, 10, 9\}\}$. The table below lists the complete surface data structure of this object. — **Right:** A flat diagram of the surface net for the same object. (courtesy of Ch. Brechbühler)

node nr.		x_0	x_1	x_2	neighbors
0	{{{	9,	14,	6},	{ 1, 7, 6, 9, 3, 4}},
1	{{	10,	14,	6},	{ 0, 3, 4, 5, 2, 8, 7, 6}},
2	{{	11,	14,	6},	{ 1, 4, 5, 11, 8, 7}},
3	{{	9,	15,	6},	{ 4, 1, 0, 6, 9, 10}},
4	{{	10,	15,	6},	{ 3, 9, 10, 11, 5, 2, 1, 0}},
5	{{	11,	15,	6},	{ 4, 10, 11, 8, 2, 1}},
6	{{	9,	14,	7},	{ 7, 10, 9, 3, 0, 1}},
7	{{	10,	14,	7},	{ 6, 0, 1, 2, 8, 11, 10, 9}},
8	{{	11,	14,	7},	{ 7, 1, 2, 5, 11, 10}},
9	{{	9,	15,	7},	{ 10, 4, 3, 0, 6, 7}},
10	{{	10,	15,	7},	{ 9, 6, 7, 8, 11, 5, 4, 3}},
11	{{	11,	15,	7},	{ 10, 7, 8, 2, 5, 4}}}

Table 4.2: The complete surface data structure of the two cube object.

4.2.2 Parametrization of closed surfaces

A key step in the shape description of a surface is its mapping to the parameter space, the sphere Ω_3. 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 *parameters* of the point. It can be represented as two polar or three Cartesian coordinates, related through the bijection

$$\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \sin\theta\,\cos\phi \\ \sin\theta\,\sin\phi \\ \cos\theta \end{pmatrix} .$$

Mapping a surface to the sphere assigns parameters to every surface point; therefore it is called *surface parametrization.* The mapping must be continuous, i.e. neighboring points in one space must map to neighbors in the other space. It is possible and desirable to construct a mapping that preserves areas (see Figure 4.8). Narrowing to the cuberille notion, Figure 4.4 symbolically illustrates this mapping of a selected facet from the object surface to a portion of Ω_3. It is not possible in general to map every surface facet to a spherical square: distortions cannot be avoided, but they should be minimal.

The parametrization, i.e., the embedding of the object surface graph into the surface of the unit sphere can be formulated as a constrained optimization problem. The following paragraphs define the meaning of *variables*, *objective* (goal function), *constraints* and *starting values* in this context.

Variables

The coordinates of all vertices vary in the optimization. Using two (e.g. spherical) coordinates per vertex would be the most economic representation with respect to storage space, but it would make the equal treatment of all spatial directions difficult and pose

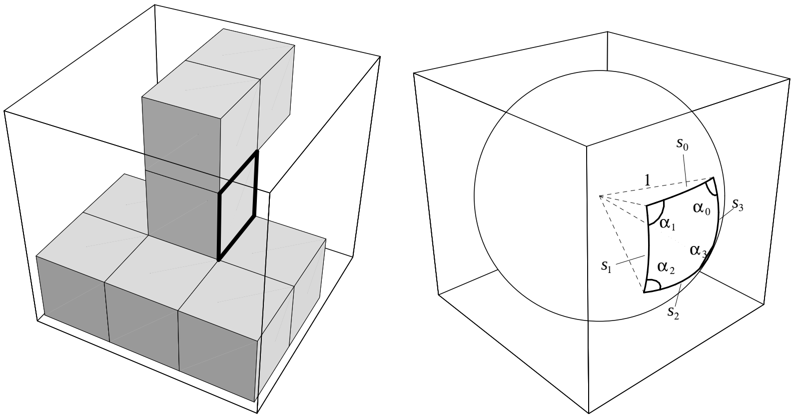


Figure 4.4: Every single face on the object's surface is mapped to a spherical quadrilateral. The sides of a spherical polygon are geodesic arcs on the sphere surface. As the sphere has unity radius, the length of a side s_i is equal to the corresponding center angle (in radian). The quadrilateral in this illustration is special in that its four sides $s_0 \cdots s_3$ are equal and its four angles $\alpha_0 \cdots \alpha_3$ are equal: it is the spherical analogue of a square. (courtesy of Ch. Brechbühler)

the problem of discontinuity and singularities in the parameter space. The authors prefer Cartesian coordinates (u, v, w) for representing a location on the sphere, introducing one virtual degree of freedom per vertex. The number of variables is three times the number of vertices.

Constraints

Two kinds of equalities and one kind of inequality constrain the values that the variables can take.

1. Every vertex must lie on the unit sphere in parameter space, i.e., $u_0^2 + u_1^2 + u_2^2 = 1$. This constraint compensates for the virtual degree of freedom and forces.
2. Ask for *area preservation*; any object surface region must map to a region of proportional area on the sphere. There is one constraint included for each elementary facet: the area of the spherical quadrilateral must be 4π divided by the total number faces.
3. All quadrilaterals on the sphere must remain convex; no angle α_k may become negative or exceed π.

Objective

The objective is to minimize the distortion of the surface net in the mapping. It is conceptually similar to angle preservation, and it must tend to make the shape of all the mapped faces as similar to their original square form as possible. To fulfill this goal perfectly, a facet should map to a "spherical square" (see Figure 4.4). This can never be reached exactly for all faces and a trade off has to be made between the distortions at different vertices. To achieve an optimum, the goal function maximizes $\sum_{i=0}^{3} \cos s_i$

summed over all spherical quadrilateral, which is the same as the sum of the cosines of the lengths of all edges.

Starting values

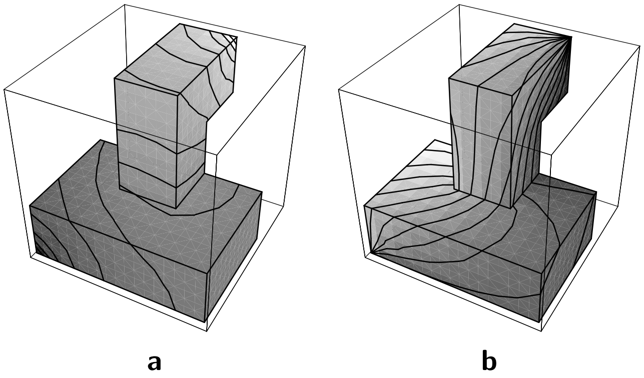


Figure 4.5: The simple object "duck" consisting of nine voxels is used for illustrating the initial parametrization. The north pole is at the lower left, the south pole at the upper right. Co-latitude is mapped on the object's surface as a grey value in *a*; iso-latitude lines are drawn every $\frac{\pi}{16}$. Longitude is shown in *b*; iso-longitude lines ("meridians") are $\frac{\pi}{8}$ apart. (courtesy of Ch. Brechbühler)

The variables of the optimization are the positions on the unit sphere to which the vertices are mapped. Therefore, starting values in this context means an initial mapping of the object's surface to the sphere. Two vertices are selected as the poles of the spherical parameter space and a two step diffusion process is applied to assign latitude and longitude coordinates to each vertex (see Figure 4.5).

Optimization method

The commonly available optimization routines can not be used for the resulting nonlinear constrained minimization problem of such a large scale, as they do not exploit its sparsity and information available about the constraints. A Newton-Lagrange algorithm has been developed by the author to find the constrained minimum of the goal function.

The solution of the nonlinear program defines the optimal parametrization of the object's surface. Figure 4.6 shows the starting point for the optimization; the right diagram is a combination of those in Figure 4.5. Figure 4.7 visualizes the result of the optimization in different ways. The same vertex as in Figure 4.6 is marked. Figure 4.8 demonstrates the importance of homogeneous surface parametrization.

4.2.3 Parametrization by spherical harmonics basis functions

The description of the surfaces of simply connected 3D objects in an arbitrary basis can be performed similarly to the 2D case. The surface will be parametrized by two variables, the θ and ϕ polar parameters, and will be defined by three explicit functions

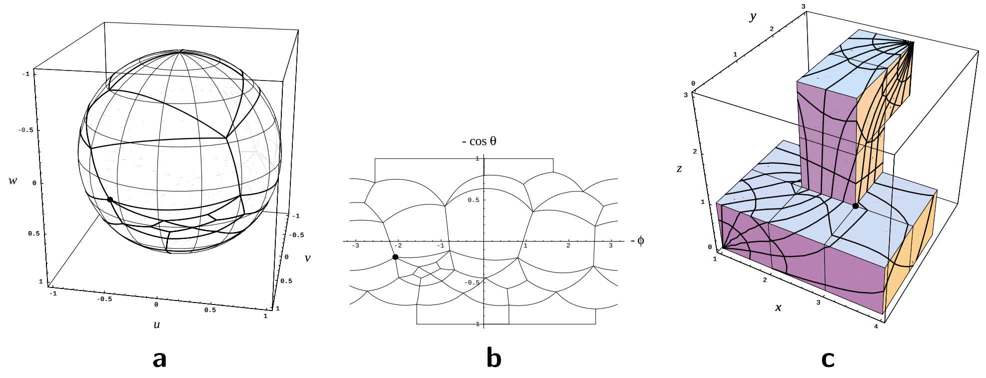


Figure 4.6: Diffusion yields the initial parametrization, which is plotted in the same three ways as the final result in Figure 4.7. **a** The surface net is plotted on the spherical parameter space. The thick lines depict the edges of the original square faces. The equidistance for both θ and ϕ is $\frac{\pi}{8}$. **b** ϕ and $\cos\theta$ are interpreted as Cartesian coordinates. The monotonic cosine function is applied to give a true-area cylindrical projection. The horizontal lines at ±1 are the poles. **c** Conversely, the globe coordinate grid is drawn over the object. (courtesy of Ch. Brechbühler)

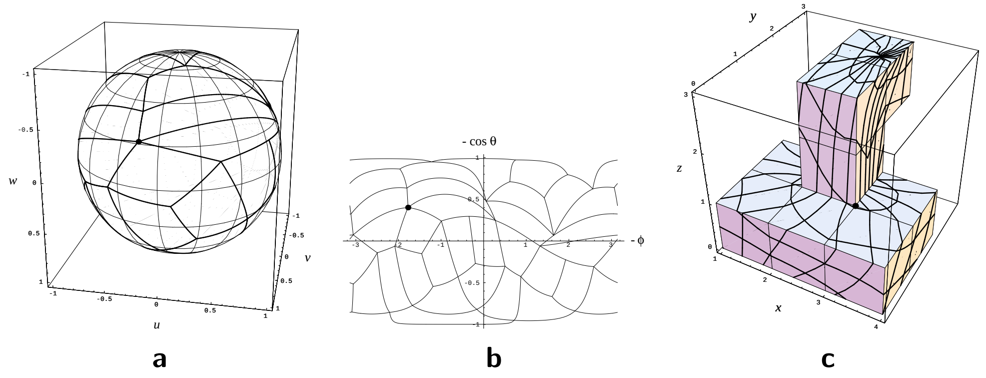


Figure 4.7: The result of the optimization, plotted in the same ways as in the previous figure. The areas of elementary facets in parameter space are now equal, and local distortions are minimized. The rotational position of the net on the sphere is arbitrary. **c** After optimization, the former poles have lost their prominent role. They have now the same importance as other point in parameter space and could lie anywhere on the surface. Only the use of polar coordinates for visualization gives them a conspicuous appearance. (courtesy of Ch. Brechbühler)

$$\boldsymbol{v}(\theta,\phi) = \begin{pmatrix} x(\theta,\phi) \\ y(\theta,\phi) \\ z(\theta,\phi) \end{pmatrix} \quad . \tag{4.22}$$

It has to be emphasized that this is not a radial function. Selecting the spherical harmonic functions (Y_l^m denotes the function of degree l and order m, see [Greiner and Diehl 1986]) as a basis, the coordinate functions can be written as

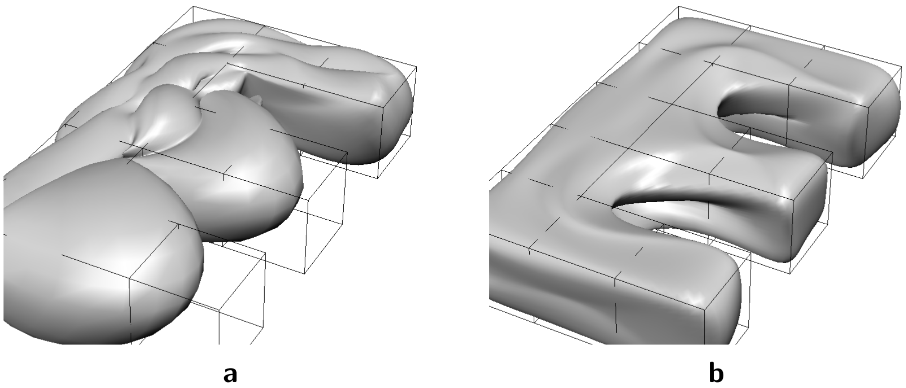


Figure 4.8: Homogeneous parameter distribution is important for shape description. The "E"-shaped object surface, indicated by a wireframe, is expanded into a series of spherical harmonics. The initial, non-uniform parametrization yields a poor shape representation (**a**); its optimization achieves a significant improvement (**b**). (courtesy of Ch. Brechbühler)

$$\boldsymbol{v}(\theta, \phi, \boldsymbol{p}) = \sum_{k=0}^{K} \sum_{m=-k}^{k} \boldsymbol{c}_k^m \; Y_k^m(\theta, \phi) \qquad , \tag{4.23}$$

where

$$\boldsymbol{c}_k^m = \begin{pmatrix} {c_x}_k^m \\ {c_y}_k^m \\ {c_z}_k^m \end{pmatrix} \qquad . \tag{4.24}$$

The expansion is restricted to the first $K+1$ terms. When the free variables θ and ϕ run over the whole sphere (e.g. $\theta = 0 \ldots \pi$, $\phi = 0 \ldots 2\pi$), the point $\boldsymbol{v}(\theta, \phi)$ runs over the whole surface of the object. The sphere Ω_3 is considered a perfectly symmetric surface without any singular points or preferred directions.

The surface is then described by the parameters

$$\begin{aligned} \boldsymbol{p} \;\; = \;\; & ({c_x}_0^0, {c_y}_0^0, {c_z}_0^0, {c_x}_1^{-1}, {c_x}_1^0, {c_x}_1^1, {c_y}_1^{-1}, {c_y}_1^0, {c_y}_1^1, \\ & {c_z}_1^{-1}, {c_z}_1^0, {c_z}_1^1, \ldots {c_x}_K^{-K} \ldots {c_z}_K^K)^\top \end{aligned}$$

Y_l^m denotes the spherical harmonic function of degree l and order m. The following definitions agree with [Press *et al.* 1988], page 252. The variable w is a scalar and will correspond to u_2 below.

Legendre polynomials

$$P_l(w) \;\; = \;\; \frac{1}{2^l l!} \, \frac{d^l}{dw^l} \, (w^2-1)^l \tag{4.25}$$

Associated Legendre polynomials

$$\begin{aligned} P_l^m(w) \;\; &= \;\; (-1)^m (1-w^2)^{\frac{m}{2}} \frac{d^m}{dw^m} P_l(w) \\ &= \;\; \frac{(-1)^m}{2^l l!} (1-w^2)^{\frac{m}{2}} \frac{d^{m+l}}{dw^{m+l}} (w^2-1)^l \end{aligned} \tag{4.26}$$

Y_l^m	$m=0$	$m=1$	$m=2$
$l=0$	$\frac{1}{2\sqrt{\pi}}$		
$l=1$	$\sqrt{\frac{3}{4\pi}}\cos(\theta)$	$-\sqrt{\frac{3}{8\pi}}\,e^{i\phi}\,\sin(\theta)$	
$l=2$	$\sqrt{\frac{5}{16\pi}}\left(-1+3\cos(\theta)^2\right)$	$-\sqrt{\frac{15}{8\pi}}\,e^{i\phi}\,\cos(\theta)\,\sin(\theta)$	$\sqrt{\frac{15}{2\pi}}\,e^{2i\phi}\sin(\theta)^2$

Table 4.3: Spherical harmonic basis functions up to degree 2.

Spherical harmonic functions

$$
\begin{aligned}
Y_l^m(\theta,\phi) &= \sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\;P_l^m(\cos\theta)\;\;e^{im\phi} && (4.27)\\
Y_l^{-m}(\theta,\phi) &= (-1)^m Y_l^{m*}(\theta,\phi) && (4.28)
\end{aligned}
$$

A list of the spherical harmonics up to degree 2 in table 4.3 exemplifies these definitions in polar coordinates (θ,ϕ). For convinience, the real and imaginary parts of the complex basis functions Y_l^m can be used as independent real valued basis functions. There is the same number of functions as in the complex basis, namely $2l+1$ for any non-negative l.

Figures 4.9 and 4.9 provide two different ways to visualize the basis functions. In figure 4.9 function values of $\mathrm{Re}\,Y_l^m$ up to degree 3 are mapped onto the surface of a unit sphere representing the parameter space in θ and ϕ, light areas corresponding to positive values and dark to negative values. In physics spherical harmonics are often shown as in figure4.9 where the radius of a unit sphere is modulated by the function values $|\mathrm{Re}\,Y_l^m|$.

One has to realize that the parametrization by spherical harmonics is just one possibility for the parametric description of contours. Alternative methods, as e.g. deformable superquadrics have also been proposed in the literature (see previous chapter). The author emphasizes that any reasonable parametric shape model can be used within the presented formalism.

4.2.4 Invariant descriptors

The coefficients obtained so far still depend on the relative position of the parameter net of the object surface, on the orientation of the object space, on the size of the object and finally on its position in space. By transforming the object to canonical positions in parameter space and object space, similarly as it has been done with Fourier parametrized 2-D contours, one can get rid of these dependencies.

Rotation independent descriptors

Rotation of the object to standard position in parameter space and object space needs three rotations in each spaces, when rotations are described by Euler angles. We wish to rotate the object in parameter space so that the north pole ($\theta=0$, on the u_2 axis) will be at one end of the shortest main axis of this first order ellipsoid and the point where the Greenwich meridian ($\phi=0$) crosses the equator ($\theta=\frac{\pi}{2}$, on the u_0 axis) is at one end of

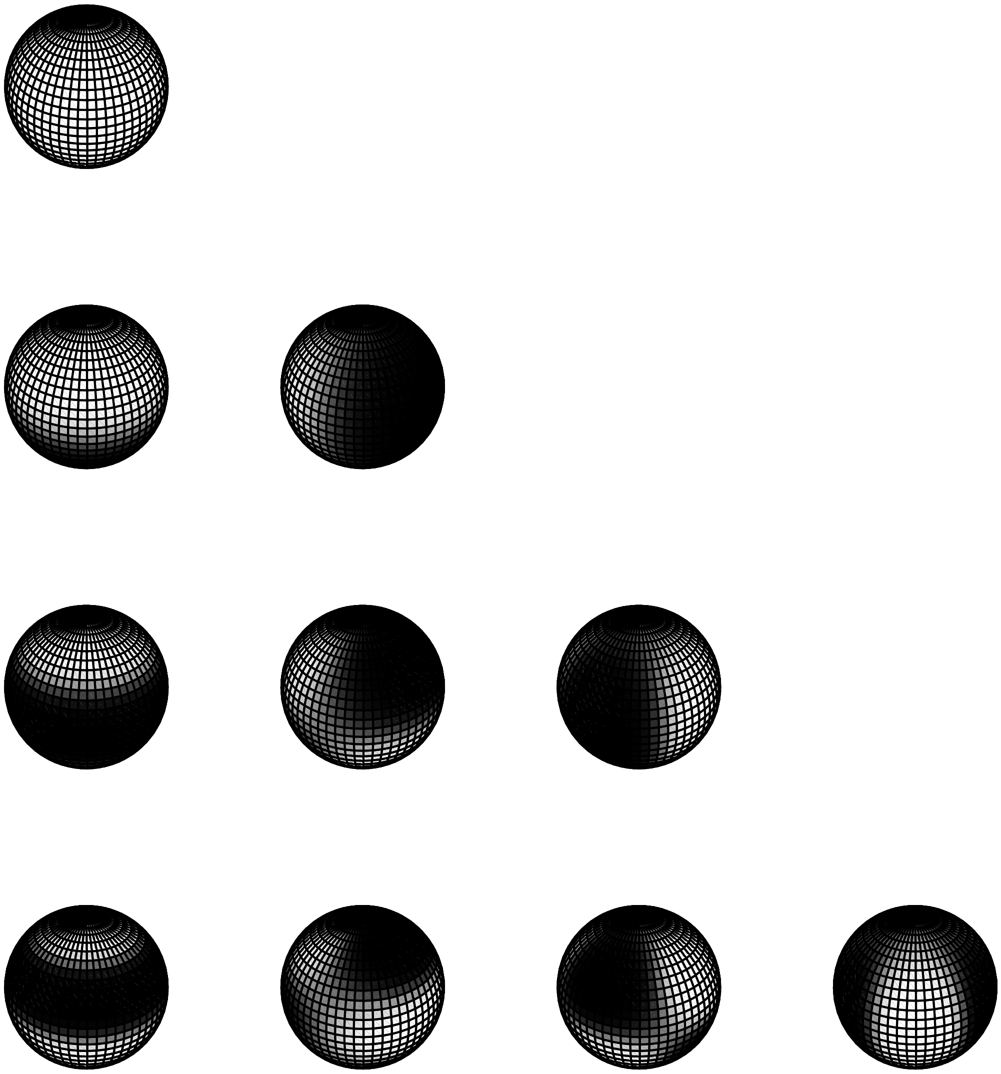


Figure 4.9: Real parts of spherical harmonic functions Y_l^m up to degree 3 are mapped onto the surface of a unit sphere, light areas corresponding to positive values and dark to negative values.

the longest main axis. To define a standard position only the contribution of the spherical harmonics of degree $l = 1$ in equation 4.23 are taken into consideration, these define an ellipsoid.

$$\boldsymbol{v}_1(\theta, \phi, \boldsymbol{p}) = \sum_{m=-1}^{1} \boldsymbol{c}_1^m \, Y_1^m(\theta, \phi) \tag{4.29}$$

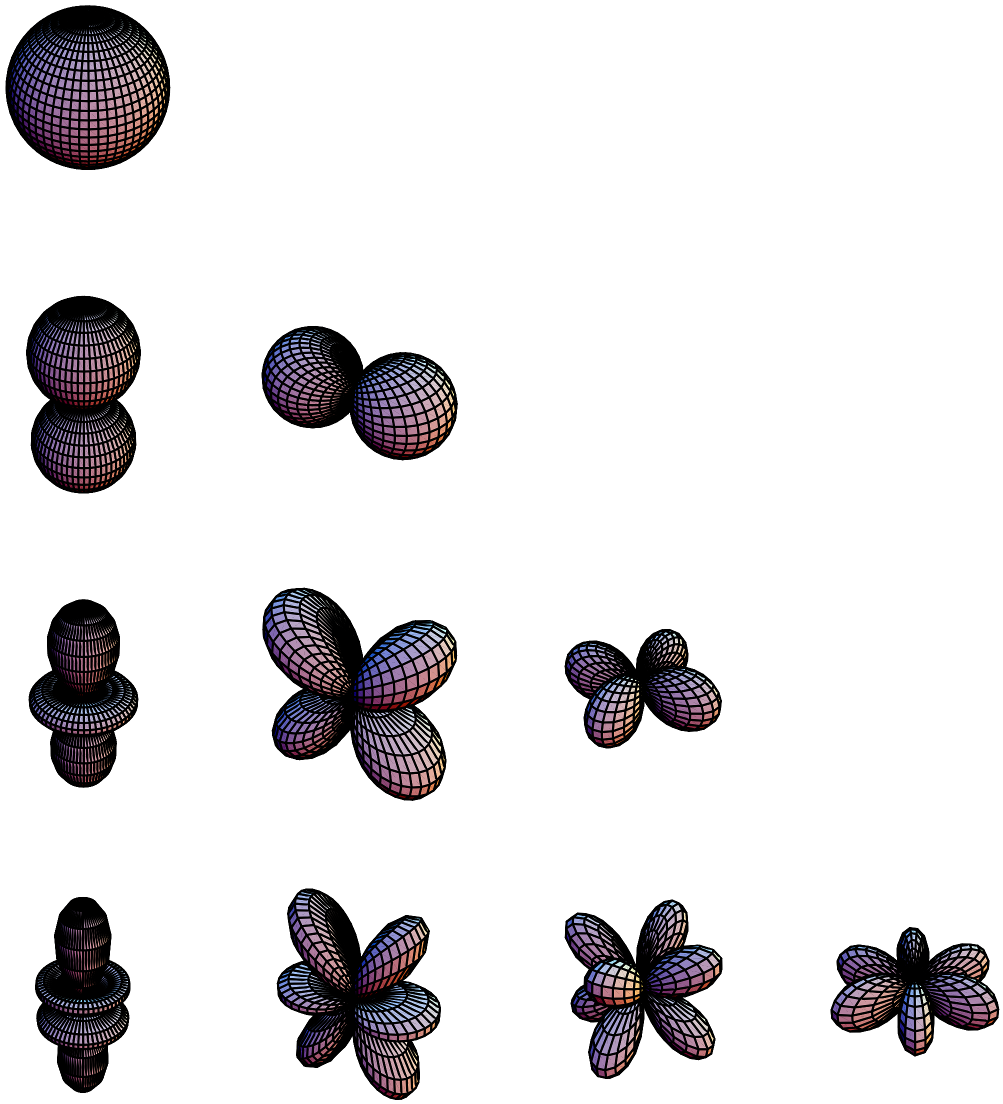


Figure 4.10: Real parts of spherical harmonic functions Y_l^m up to degree 3. The radius of a unit sphere is modulated by the function values $|\mathrm{Re}\, Y_l^m|$

Substituting the basis functions $Y_1^{-1} = \frac{\sqrt{3}}{2\sqrt{2\pi}}\,(u_0 - iu_1)$, $Y_1^0 = \frac{\sqrt{3}}{2\sqrt{\pi}} u_2$ and $Y_1^1 = -\frac{\sqrt{3}}{2\sqrt{2\pi}}(u_0 + iu_1)$ this sum can be written as

$$\boldsymbol{v}_1(\boldsymbol{u}) \;=\; \boldsymbol{A}\boldsymbol{u} \;=\; \boldsymbol{A}\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} \;=\; \boldsymbol{a}_1\, u_0 + \boldsymbol{a}_2\, u_1 + \boldsymbol{a}_3\, u_2 \quad , \tag{4.30}$$

where

$$\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3) \quad = \quad \frac{\sqrt{3}}{2\sqrt{2\pi}} \left(\boldsymbol{c}_1^{-1} - \boldsymbol{c}_1^{1}, i(\boldsymbol{c}_1^{-1} + \boldsymbol{c}_1^{1}), \sqrt{2}c_1^0 \right) \quad . \tag{4.31}$$

The rotation matrix $\boldsymbol{R}_u^T = (\hat{\boldsymbol{u}}_1, \hat{\boldsymbol{u}}_2, \hat{\boldsymbol{u}}_3)$, where $\hat{\boldsymbol{u}}_1$, $\hat{\boldsymbol{u}}_2$, and $\hat{\boldsymbol{u}}_3$ denote the unit eigenvectors of $\boldsymbol{A}$, can be then used to rotate the parameter net over the surface into the desired location. The roots of the eigenvalues $l_1^2 > l_2^2 > l_3^2$ represent half the length of the main axes of the ellipsoid. The rotation matrix is applied to the parameters $\boldsymbol{u}_i$ associated with each vertex i:

$$\boldsymbol{u}_i \mid^V = \boldsymbol{R}_u^T veu_i \ . \tag{4.32}$$

This new parametrization results in new coefficients $\boldsymbol{c}_l^m \mid^V$.

In the next step, the ellipsoid is rotated in the object space to make its main axes coincide with the coordinate axes, putting the longest ellipsoid axis along x and the shortest along z. This requires only the matrix multiplication

$$\boldsymbol{c}_l^m \mid^R = \boldsymbol{R}_x \ \boldsymbol{c}_l^m \mid^V \ , \tag{4.33}$$

where the rotation matrix is defined as $\boldsymbol{R}_x = \text{diagonal}(\frac{1}{l_1}, \frac{1}{l_2}, \frac{1}{l_2}) \boldsymbol{A}^T \boldsymbol{R}_u^T$.

Scale independence

Scaling invariance can be achieved by dividing all descriptors by l_1, the length of the longest main axis

$$\boldsymbol{c}_l^m \mid^S = \boldsymbol{R}_x \ \boldsymbol{c}_l^m \mid^R \ , \tag{4.34}$$

Invariant spherical harmonic descriptors

Ignoring the coefficients of degree $l = 0$, that is setting $\boldsymbol{c}_0^0 \mid^T = (0, 0, 0)^T$ achieves translation invariance.

4.3 Ambiguous cases of normalization

The normalization techniques described here for the 2-D and 3-D cases require the precondition that coefficients of degree 1 represent a real ellipse or a real ellipsoid, respectively. If, however, the ellipse degenerates to a circle and the ellipsoid to an ellipsoid of revolution or a sphere, the technique will fail to derive stabile main axes. Objects of higher symmetries, such as regular polygons and polyhedra are a good example for the limitations of the normalization technique (see Figure 4.11). In general, the method will fall through for shapes with properties which are not reflected in the first degree cofficients. In these cases other features have to be incorporated into the normalization. These features can involve, e.g., coefficients of higher degree or an external coordinate system. Wyskočil in his diploma work [Wyskocil 199798] proposes a method which uses higher degree Fourier coefficients to unambiguously normalize 2-D contours.

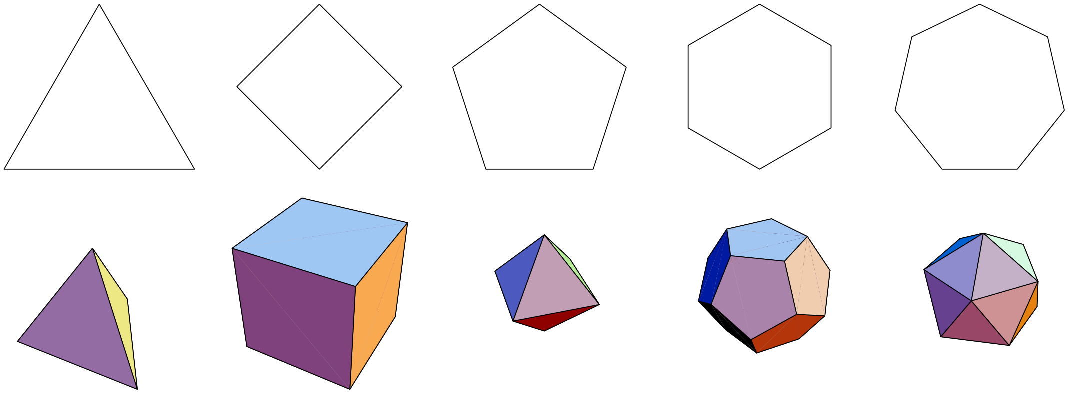


Figure 4.11: Because of the limitations of the normalization technique based solely on the first degree coefficients objects of higher symmetries, e.g. regular polygons and polyherda cannot be unambiguously aligned, since their fitst ellipse or ellipsoid degenerates to a circle or sphere.

The symmetries of the ellipse and ellipsoid have also to be taken into consideration while computing corresponding descriptors of similar objects. We will refer to this problem in the next chapters when we describe the generation of average models of anatomical organs.

5

Model-based segmentation of 2-D images

In this chapter we first summarize how to apply modal analysis to outlines extracted from 2D images. Two shape examples are presented. In the first example closed contours of the corpus callosum extracted from 2D mid-sagittal MR images are arc-length parametrized to establish correspondence between them. The other example, based on open contours of the femur obtained from X-ray images of the pelvis, illustrates the limitations of the arc-length parametrization and proposes automatic reparetrization to achieve homology.

The second part of the chapter presents the development of the elastic matching procedure for the corpus callosum in a historical order. Our first implementation separated shape deformation parameters from those of the similarity transform (translation, rotation and scaling). The proposed procedure is a combination of an initialization using an average model based on template matching technique followed by an elastic deformation restricted to the major eigenmodes. We demonstrate that the separation of the similarity transform from the elastic deformation often leads to unsatisfactory segmentation results, calling for a modified segmentation process that provides a unified framework for initialization and elastic model matching. Segmentation quality has been further improved by augmenting the model with information about its average image intensity environment. In the end of the chapter we present segmentation results and compare them to manual segmentations.

5.1 Model building

The corpus callosum outline on the midsagittal plane of the brain had to be segmented and parametrized for each image of the training set. To perform this task a reimplementation of Staib and Duncan's Fourier-snake program has been applied, with a manual initialization of the optimization. For the minimization of the total energy function of the Snakes the E04JBF routine of the NAG^{TM} library has been used, which applies a quasi-Newton algorithm. It is designed to find a minimum of a function of several variables which are either subject to fixed upper and lower bounds on the variables or unconstrained. No derivatives are needed, these are estimated with finite-differences. If the algorithm suspects a local minima or saddle point, it carries out a local search to move away.

The segmentation results can be seen in Figure 5.1. Fourier coefficients up to degree

25 have been used for the description. The resulting contours of the training set have been normalized to be invariant to translation, rotation, scaling in object space, and the starting point of the parametrization in parameter space. Moving the starting point to a canonical position establishes an approximate correspondence between individual contours. While this way of achieving correspondence cannot be as precise as manual annotation of contour points, it is obtained rapidly and without any human interaction. Additionally, one can think of very smooth contours which will not allow the natural placement of landmarks, in which case curve parametrization will give a better solution to the correspondence problem. The goodness of the correspondence is also illustrated in Figure 5.1, where along each contour eight dots are placed, such that they divide the length of the contour into eight equal intervals. Comparing the positions of these points among different individuals, it is obvious that the proposed parametrization provides a reasonable alternative to manual landmark placing.

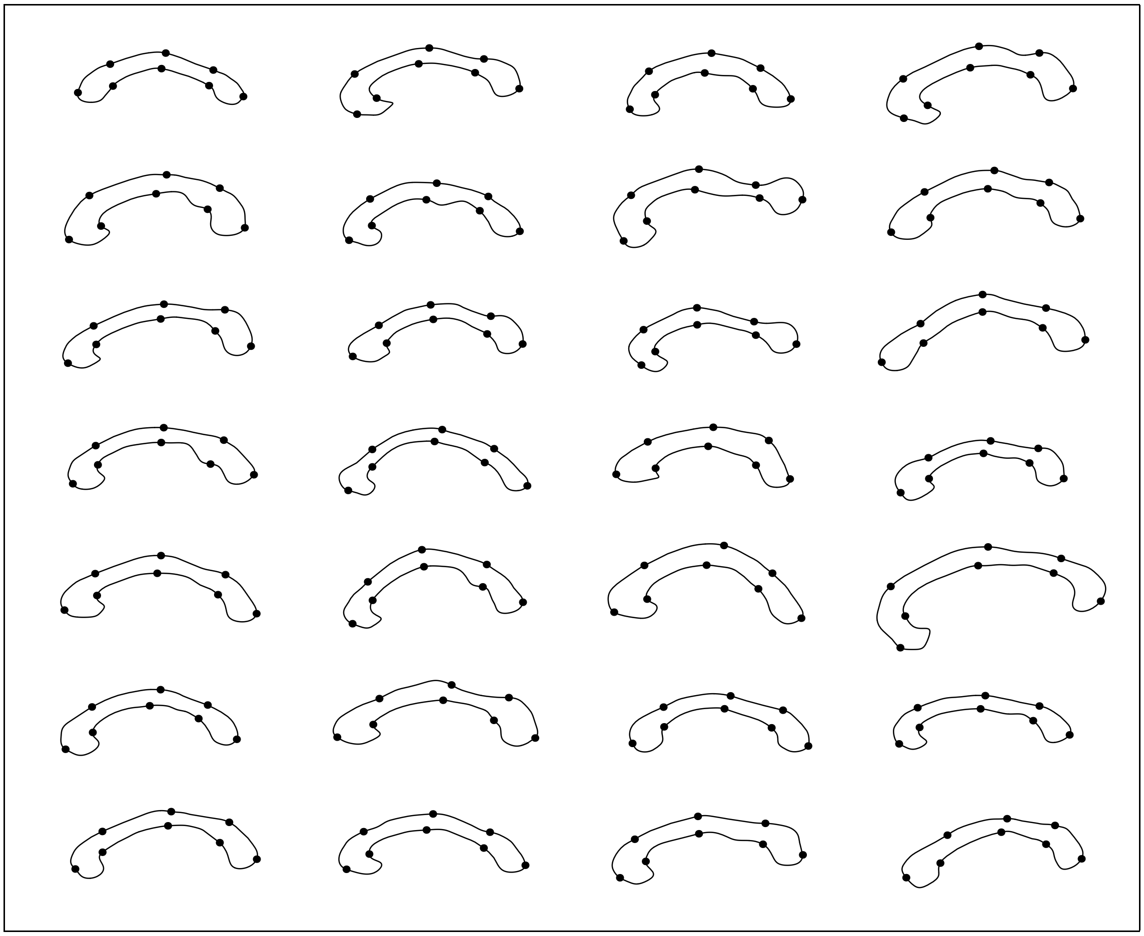


Figure 5.1: Complete ($N = 30$) set of shapes in the corpus callosum training set reconstructed from their Fourier descriptors. Along each contour eight points are marked which divide the length of the contour into eight equal intervaldivide the length of the contour into eight equal intervals.

The mean model has been determined from the normalized outlines by simply averaging their parameters $\boldsymbol{c}_j$.

$$\bar{\boldsymbol{c}} = \sum_{j=1}^{N} \boldsymbol{c}_j \tag{5.1}$$

In order to determine the major deformation modes defined by the above training set, we performed a principal component analysis of the covariance matrix of the normalized Fourier coefficients.

$$\text{deviation from mean:} \quad d\boldsymbol{c}_j = \boldsymbol{c}_j - \bar{\boldsymbol{c}} \tag{5.2}$$

$$\text{covariance matrix:} \quad \boldsymbol{\Sigma}_c = \frac{1}{N-1} \sum_{j=1}^{N} d\boldsymbol{c}_j d\boldsymbol{c}_j^T \tag{5.3}$$

$$\text{PCA:} \quad \boldsymbol{\Sigma}_c \boldsymbol{P}_c = \boldsymbol{\Lambda}_c \boldsymbol{P}_c \tag{5.4}$$

where the columns of $\boldsymbol{P}$ hold the eigenvectors and the diagonal matrix $\boldsymbol{\Lambda}$ the eigenvalues λ_j of $\boldsymbol{\Sigma}$. 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 shape of the corpus callosum outline. Any shape in the training set can be approximated using the mean shape and a weighted sum of the deviations obtained from the first few modes

$$\boldsymbol{c} = \bar{\boldsymbol{c}} + \boldsymbol{P}_c \boldsymbol{b} \ , \tag{5.5}$$

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

$$\boldsymbol{b} = \boldsymbol{P}_c^T (\boldsymbol{c} - \bar{\boldsymbol{c}}) \ . \tag{5.6}$$

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 modes instead of the Fourier harmonics. Eq. 5.6 describes how to generate new examples of the shapes by varying the parameters ($\boldsymbol{b}$) within suitable limits, so the new shapes will be 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 followings has to be taken into consideration. Supposing, Fourier harmonics up to degree K has been used, there will be $4K$ free parameters describing the shape, this results in a covariance matrix of the size $4K \times 4K$ and theoretically in $4K$ different eigenmodes. However, if the training set only consists of $N < 4K$ 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 $4K$:

$$t < \min(N, 4K) \ . \tag{5.7}$$

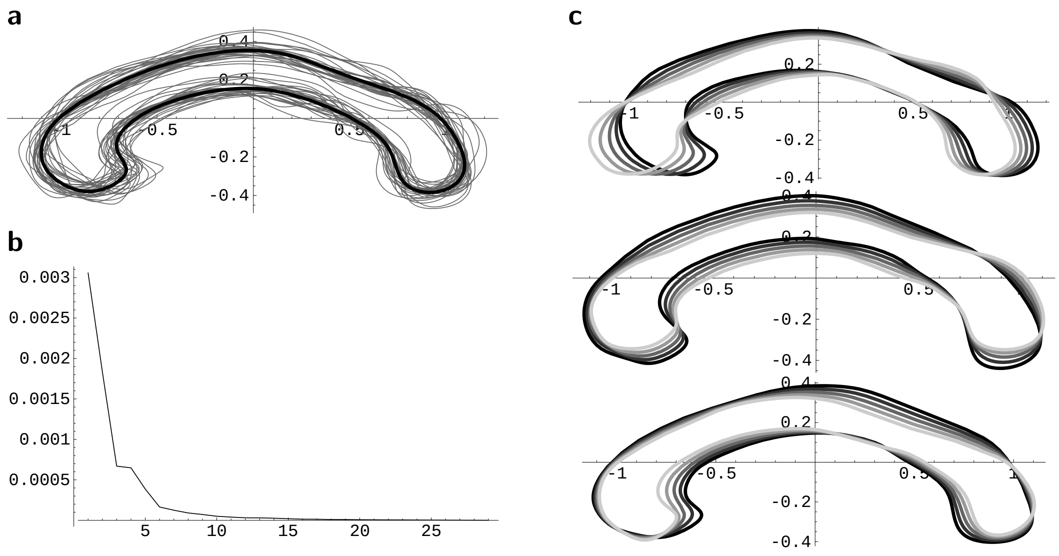


Figure 5.2: Principal component analysis of the covariance matrix computed from the normalized Fourier coefficients. The average contour is shown in **a** (thick line) together with the individual contours (gray lines). Resulting eigenvalues are illustrated in **b**, while the first three major deformation- or eigenmodes in **c**. The deformation range amounts from $-\sqrt{2}$ (light gray) to $\sqrt{2}$ (dark gray) eigenmodes.

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

$$\lambda_T = \sum_{k=1}^{\min(N,4K)} \lambda_k \tag{5.8}$$

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

5.1.1 Quality measure of the statistical model

In the example above 30 samples of *corpus callosum* have been used to derive a statistical model. Applying equation 5.8 we find that the 10 largest eigenvalues express 99% of the variation represented in the training set. This, however, only means that the 30 shapes of the training set can be described with minor error using the model, and does not say anything about 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 (eq. 5.6), then approximating coefficients $\tilde{\boldsymbol{c}}$ from the projection (eq. 5.5), 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:

$$\epsilon = D_{\text{Eucl}}(\boldsymbol{c}, \tilde{\boldsymbol{c}})\,. \tag{5.9}$$

To demonstrate how the quality of the model increases while incorporating more individuals, we first build a statistical model using a training set of 11 arbitrarily chosen shapes out of the entire set of 30, as needed to be able to compute 10 deformation modes, and determine the above measure for the remaining shapes not included in the training set. Repeating the computations for all possible combinations of 11 out of a set of 30 and averaging the errors, we obtain a measure of the 11-shape model as shown in Figure 5.3 by the first data point. The model is increased one-by-one and the average error is computed based on the largest 10 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, however, the decline becomes less and less significant. Later on in this chapter in Figure 5.14 we can recognize the asymptotic behavior of the curve which implies that the model becomes saturated and after a certain number of adding new shapes to the model does not increase its information content.

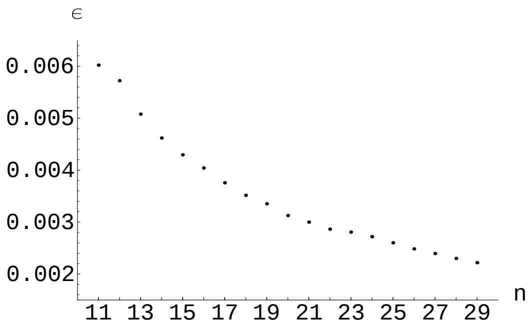


Figure 5.3: Quality measure curve computed for the 30 *corpura callosa* training set. It displays the average representation error ϵ for shapes not included in the training set when using the 10 largest eigenmodes while incresing the number of shapes involved in the computation from $n = 11$ to $n = 29$.

The model, however, shown in Figure 5.3 does not reach saturation after incorporating all shapes of the training set, which means that the amount of information currently included in the model could be further augmented by adding more individual shapes.

5.1.2 Modal analysis of the femur

In many cases, however, arc-length parametrization does not provide sufficient correspondence between contours. In a student project, Peter Wepf investigated shapes of the upper part of the femur. The training set has been obtained by manual segmentation of 41 X-ray images of the human pelvis and by subsequent B-spline parametrization of the open curves (see figure 5.5 **a**). The original X-ray images only show the upper head of the femur while its lower part has been cut away at an arbitrary position by the imaging system. This causes a high variation of the length of the femur's stem which prevents from establishing correspondence by arc length parametrization.

To solve the correspondence problem between the contours their tangential direction has been coded along the curve by Ψ-functions as illustrated in figure 5.4. Differential properties of the individual Ψ-functions have been used to reparametrize the contours by

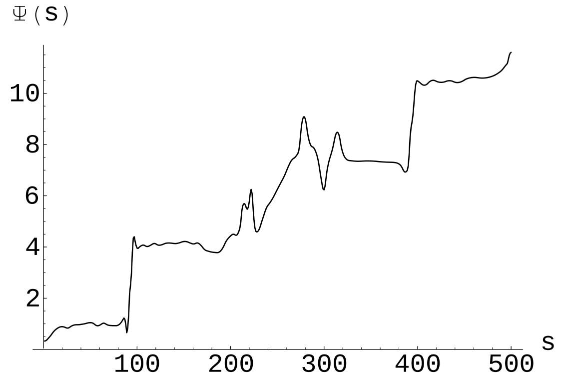


Figure 5.4: Ψ-function of one individual shape. It is used to code the contours tangential direction as a function of their arc-length s. The two most abrupt changes at $s \approx 100$ and $s \approx 400$ correspond to positions where the femur's lower part has been cut away.

assigning the same amount of parameter space to homologous intervals of the functions. This procedure is carried out fully automatically. After reparametrization deformation modes of the training set can be obtained. Figures 5.5 **b** and **c** show the resulting eigenvalues and eigenmodes. As expected, the largest deformation captured by the first eigenmode reflects the variation of the length of the femur's stem.

5.2 Initial placement of model contour by template matching

A segmentation of the corpus callosum from grey-valued images based on the deformation of the mean model requires a suitable initialization. Due to the normalization of the Fourier coefficients, the average model only expresses shape deformations up to a similarity transformation. Therefore, the initial placement has to provide a sufficiently good match between the model and the edges in the grey-valued image. A standard template matching procedure, which can be considered as a generalization of the Hough-transformation to arbitrary curves, was chosen to solve this first optimization problem by equidistantly sampling the possible parameters of the similarity transformation within a reasonable range. The goodness of fit was calculated for a Canny edge map on a relatively large scale, since the rigid transformation does not allow for elastic shape deformations. The objective function simply sums up the intensity values of the edge map along the contour, given by the averaged Fourier descriptors and the 4 parameters of the similarity transformation (x- and y-translation, rotation, and scale). To obtain the position of the best match the resulting 4-dimensional space has been searched at a reasonably fine grid of the parameters for a global maximum. Figure 5.6 illustrates the result of the initialization by template matching. Pictures **a, b**, and **c** show the original image, the Gaussian smoothed image ($\sigma = 5$ *pixel*), and the Canny gradient magnitude, respectively. Figure **d** presents an overlay of the optimal fit between model curve and gradient image.

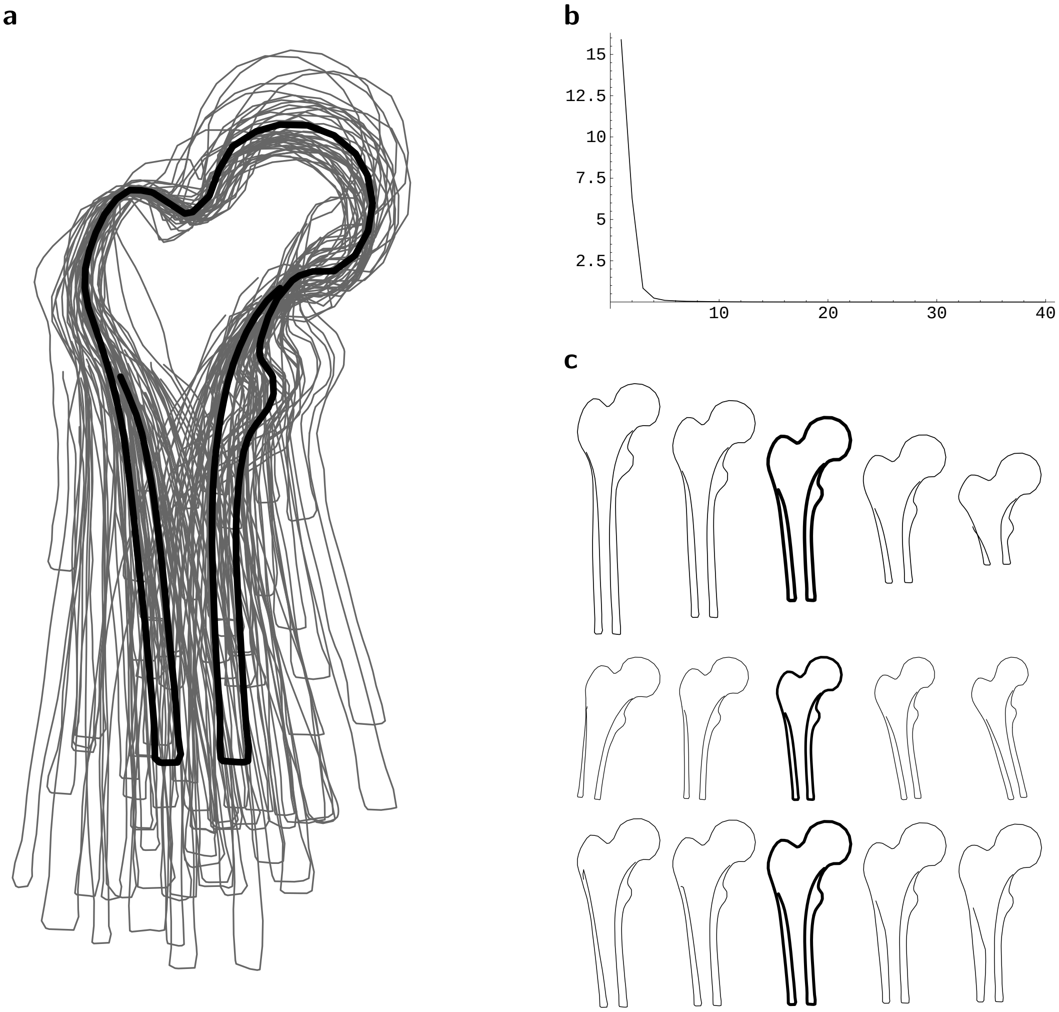


Figure 5.5: Principal component analysis of the femur shape. The average contour is shown in **a** (thick line) together with the individual contours (gray lines). Resulting eigenvalues are illustrated in **b**, while the first three major deformation- or eigenmodes in **c**. The deformation range amounts from $-\sqrt{2}$ (light gray) to $\sqrt{2}$ (dark gray) eigenmodes.

5.3 Segmentation by restricted elastic deformation

After initialization a modified version of the Fourier-snake program has been applied. Instead of optimizing in the complete space of the normalized Fourier coefficients, the optimization has been restricted in order to allow only deformations which are prominently represented in the training sample. In the first step the eigenvectors of the parameter covariance matrix are selected as a different set of basis functions in place of the harmonics. The restricted variation is achieved by choosing a subset of eigenmodes, usually the n largest ones, and calculating the optimization in this linear subspace. Starting with the result of the template matching, the model is elastically deformed until it fits (in a local optimum) the edges along object contour.

The procedure described above provided in many cases satisfactory segmentation re-

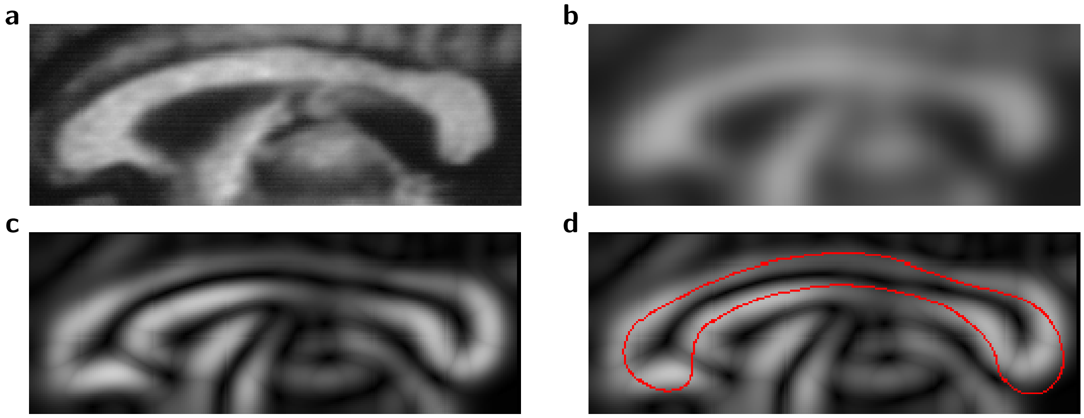


Figure 5.6: Template matching procedure. Pictures **a, b**, and **c** show the original image, the Gaussian smoothed image ($\sigma = 5\ pixel$), and the Canny gradient magnitude, respectively. Figure **d** presents an overlay of the optimal fit between model curve and gradient image.

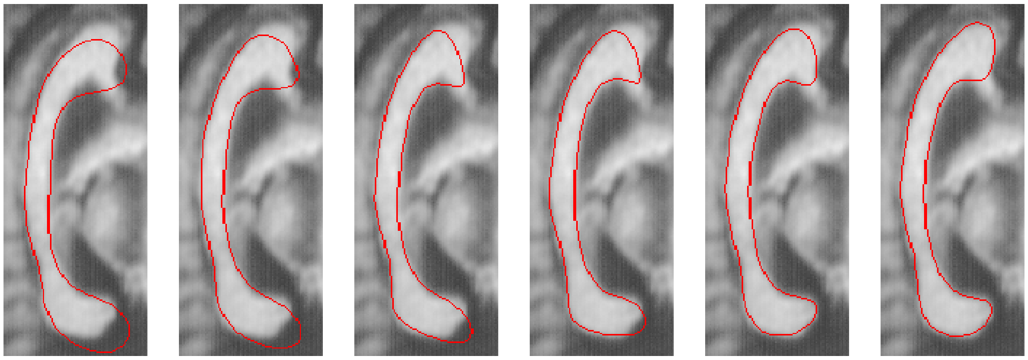


Figure 5.7: Iteration steps of the optimization procedure on an image not included in the training set. Starting with the best rigid match (left image), the model is elastically deformed until it matches the object contour (right image).

sults. In other cases, however, we met difficulties to find the correct contour. Figure 5.8 shows a few examples of unsuccessful segmentations. The analysis of the results has shown that the separation of the similarity transform (in the initialization step) from the elastic deformation is mainly responsible for this failure. As the models have been normalized before the statistical analysis, translation, rotation, and scaling has been actually excluded from the elastic deformation step, forbidding minor corrections of the similarity transform determined by the template matching in the initialization step. Restricting the deformations to the subspace of the dominant eigenmodes made this situation even worse by not allowing corrections through free elastic deformations. This has lead to the large segmentation errors demonstrated in Fig. 5.8.

To be able to refine the position of the contour during the elastic matching, the 4 parameters of the similarity transformation has also been added to the parameters which have to be optimized. Thus, the elastic model finally had $n + 4$ total degrees of freedom. This, however, only marginally improved the quality of the segmentation. We found two

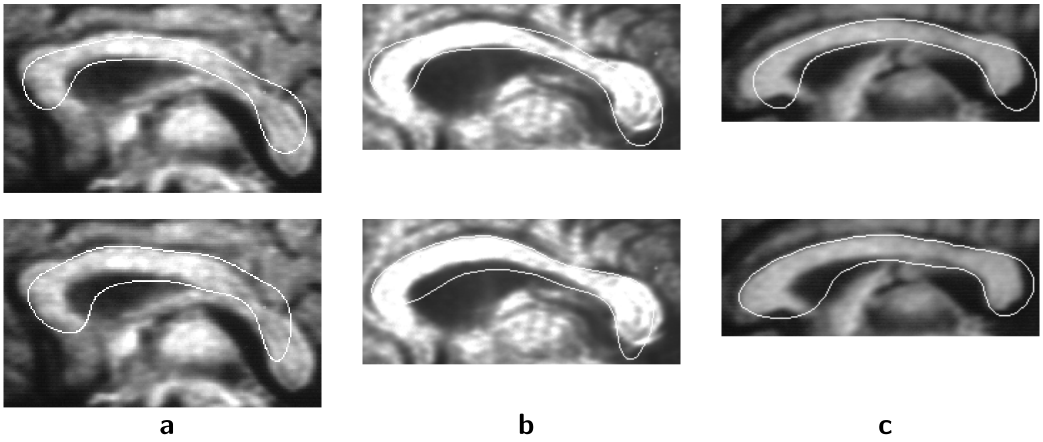


Figure 5.8: Segmentation examples illustrating segmentation failures. Images in the top row illustrate the initial placement of the model curves (using template matching), the bottom row shows the segmentation results. The *same* initial model curve is used for the different images.

major reasons for the failure:

- The separation of the parameters for the similarity transform (translation, rotation and scaling) and the Fourier coefficients in the model (small elastic deformation) is somewhat artificial, as the elastic deformation of normalized contours does not fully express the biological variations.
- The quasi-Newton algorithm used by the NAG library routine determines the local optimum nearest to the initialization. This *makes the optimization extremely sensitive to the initialization*, which is unsatisfactory since the Hough transform results in only a rough rigid match.

The following section describes how these problems can be solved by incorporating the similarity transform into the analysis of biological variability. This way all covariation of the pose and elastic deformation parameters can be handled in a consistent way, too. The selected solution also avoids the separation of the image analysis process into an initialization and elastic matching, leading to a two-step coarse-to-fine segmentation procedure.

5.4 Improving Fourier-snake segmentation

Alternative initialization techniques: One could try different ways to avoid the segmentation problems analyzed in the previous section.

- The desirable mixing of the effects of similarity transform and elastic deformation could be achieved by the incorporation of eigenvalues belonging to the largest eigenvectors into the initialization (template matching) procedure. However, the dimensionality of the sampled search space space would become by far too high and would create, especially in view of a generalization to 3D, an unacceptable computational burden.

- One could also try to extend the parameters of the elastic deformation with translation, rotation, and scaling. This would lead to a higher dimensional local optimization problem without allowing the handling of the optimization of the parameters in one coherent framework.

- The selected optimization procedure finds the nearest local optimum on a high-dimensional strongly non-convex goal function. This makes the procedure highly sensitive to non-desired local optima. However, one should realize that *the major problem does not lie in the initialization procedure.* The template matching usually provides a reasonable first approximation using the mean model, which still leads to unsatisfactory results due to the separation of the similarity transform and the elastic deformation described by the principal modes of the Fourier coefficients. This problem is addressed in the next paragraph.

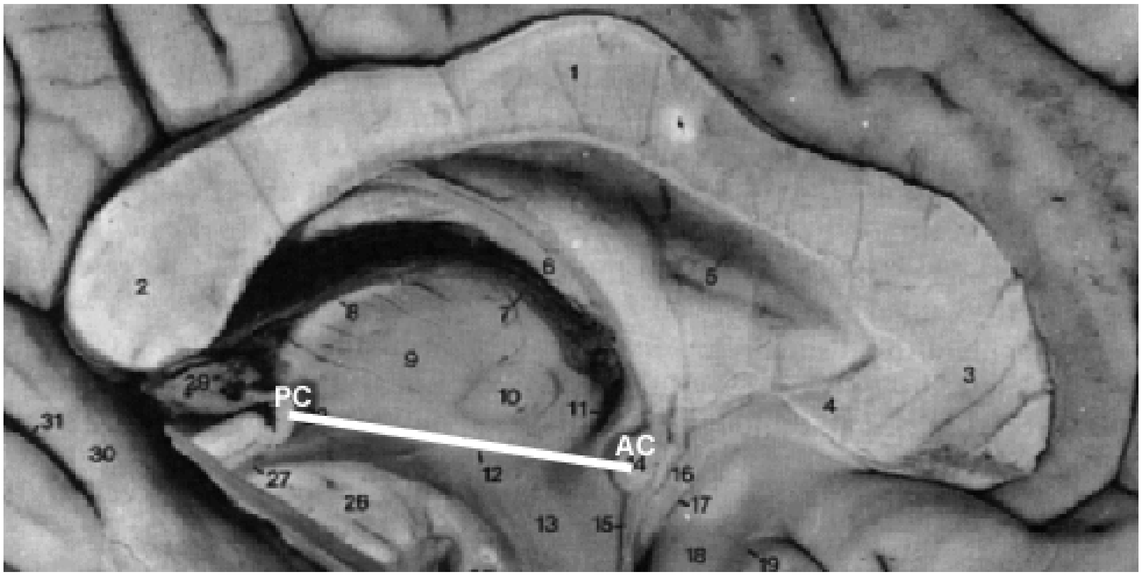


Figure 5.9: Position of the anterior commissure (AC) and the posterior commissure on a post mortem mid-sagittal brain slice. The connecting line represents the unit vector $\boldsymbol{e_x}$ of the external anatomical coordinate system.

5.4.1 Model incorporating full biological variability

We expect for images representing anatomy that the *relative position, rotation, and size* of healthy organs is restricted in a similar way as *their elastic deformation.* If we could define a coordinate system fixed to the anatomy, there would be no reason for a unrestricted similarity transform which precedes the elastic matching. The Fourier descriptors of the organ outlines *originally contain this information*, but we eliminate it by normalizing the coefficients. In the case of the corpus callosum, the AC/PC line is a generally accepted, well detectable geometric feature of the mid-sagittal images, which represents such a standard coordinate system. The line from the anterior to the posterior commissure (AC/PC line) as illustrated on Figure 5.9 has been manually extracted for each image of the training set. After determination of the Fourier coefficients, we apply normalization only for fixing the starting point of the curve parametrization. The standardization of the images, necessary for the determination of the deformation modes, is based on a normalization of the AC/PC line to a $\boldsymbol{e}_x$ unit vector. After that, the same statistical analysis of the test set can be performed as previously explained, providing a mean model (now including its relative position and size to the AC/PC line), and the deformation modes which incorporate the parameters of the similarity transform, too. The resulting eigenvalues are plotted in decreasing order in Figure 5.10 **b**. One can see that the remaining variation after the 10^{th} eigenvalue becomes negligible. Accordingly, all deformations have been restricted to the 10 largest eigenvectors.

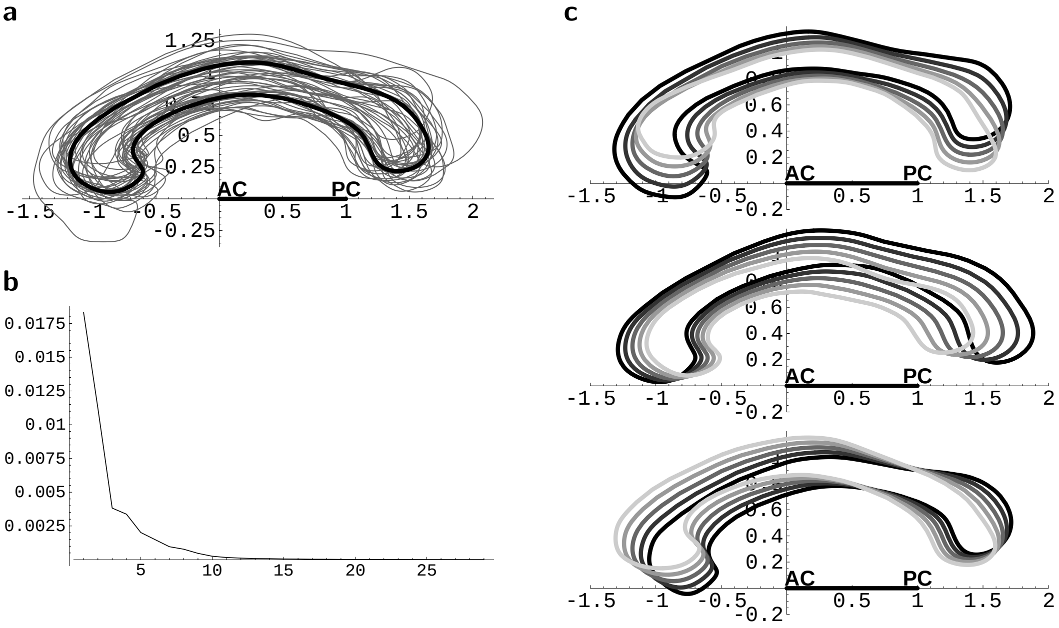


Figure 5.10: Principal component analysis of the covariance matrix computed from the AC/PC normalized Fourier coefficients. The average contour is shown in **a** (thick line) together with the individual contours (gray lines). Resulting eigenvalues are illustrated in **b**, while the first three major deformation- or eigenmodes in **c**. The deformation range amounts from $-\sqrt{2}$ (light gray) to $\sqrt{2}$ (dark gray) eigenmodes.

Figure 5.10 **a** illustrates the determination of the mean model. Figure 5.10 **c** presents the deformations according to the first three eigenmodes, with deformations in the range of $\pm\sqrt{2}$ eigenvalues. It shows how the pose and elastic deformation parameters are mixed together in the dominant eigendeformations.

Comparing the results using the normalized outlines as described in the previous section (Figure 5.2 **a**) and the average model resulting without normalization presented in the anatomical reference frame, it can be seen that the variance of the outlines became larger. This can be explained by the fact that the new normalization also keeps the contour's position relative to the anatomical coordinate system and not only its shape. The higher variation is also expressed by the more significant deviations represented by the eigenmodes (Figure 5.10 **c**) and the bigger eigenvalues (Figure 5.10 **b**) . At the same time, more characteristic shape features are retained in the average outline if no normalization is performed before the training.

5.4.2 Segmentation

The determination of the AC/PC line now becomes part of the segmentation, since the model is built based on a normalization to these landmarks. The AC/PC line is determined manually for the image under analysis. The flexible model, characterized by the mean contour and the eigenmodes, now incorporates *changes of the position and local deformations* of the generic model, which makes the initialization step obsolete. In order to make the optimization robust against local extrema we applied a *two-step coarse to fine* strategy.

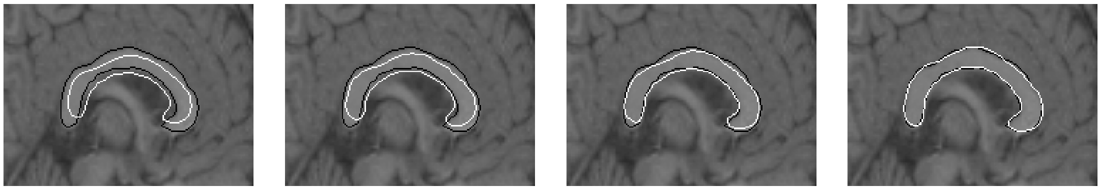


Figure 5.11: Segmentation of the corpus callosum using the manually determined AC/PC line as a reference coordinate system for model building and segmentation. The first image shows the initialization, resulting from a coarse match in the subspace of the largest four deformation modes. Images in the middle illustrate single steps of the optimization in the subspace of the largest 9 deformation modes (allowing fine adjustments). Finally the last image depicts the result of the segmentation

In the first step on the coarse level we use a procedure similar to the template matching, as described previously. However, we now use a small set of a few dominant modes and calculate the best match in this linear subspace of major deformations by equidistantly sampling the weights of the selected deformation modes in a reasonable range. This way the complete parameter space is explored, and the result of this coarse segmentation can be used as a reliable initialization for the following local optimization step. One should realize that the larger the number of eigenmodes participating in this coarse step is, the better the initial fit will be in the following step. On the other side, the computational burden is exponentially growing with the dimensionality of the sampled parameter space. The number of eigenmodes considered in this step can be selected by trading the fit quality and the computational requirements.

On the fine level the segmentation is performed in the homogeneous parameter space of the eigenvectors corresponding to the largest eigenvalues, now choosing a larger number of eigenvectors than on the coarse level. Due to the full incorporation of the parameters of the similarity transform into the eigendeformations, the fine tuning of the model pose can be performed within the variability determined by the training set. Figure 5.11 shows the results of the modified segmentation procedure, which achieved very good results in almost all cases.

For the sake of simplicity, our implementation uses a two-stage relaxation method as described above. However, the shape representation by deformation modes also allows for a gradient transition from course to fine levels, since course deformations are represented by eigenvectors corresponding to larger eigenvalues and smaller eigenvalues correspond to more subtle deformation modes. Thus, a gradual relaxation strategy can be achieved by starting the fit with the largest deformation mode and sequentially incorporating weaker modes as the fit converged on the previous level.

One should mention, that in some applications the identification of the small individual variations of the contour outline, which is not represented by even the whole set of the eigenmodes, is desirable. In this case our strategy can be extended with an additional finer level by applying an unrestricted deformation of the parametrized snake using the result of the previous two levels as initialization.

5.5 Modeling image intensity environment of contours

Shapes of boundaries are certainly not the only characteristic features of organs in medical images, their environment is usually also uniquely representative for a particular type of object. Objects' environment can be characterized by changes in image intensity along their boundary within a given zone inside and outside of the object. As it was shown by [Cootes *et al.* 1995a], incorporating this additional information into the statistical shape model and taking it into account during the segmentation process will further increase performance and provide better results. Following their suggestion, image intensity is sampled along profiles orthogonal to the boundary.

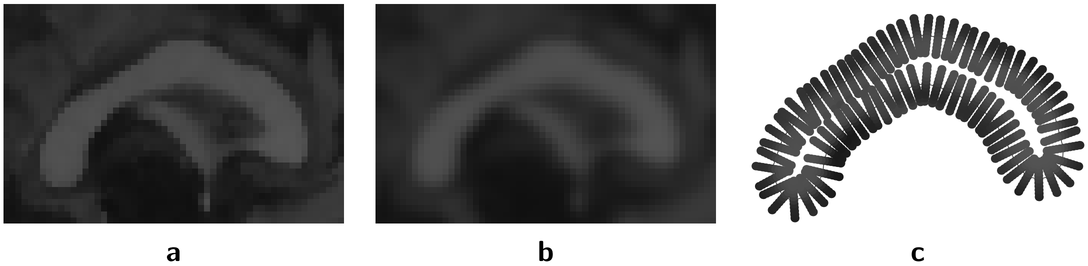


Figure 5.12: Extracting intensity profiles. The original image (**a**) is at first smoothed for regularization with a Gaussian of $\sigma = 1 pixel$ (**b**). Regular sampling of the contour is enabled by the parametric representation. Intensity profiles normal to the curve has been extracted at 80 equally distributed contour points. Along the profiles sampling points are placed at a distance of $1pixel - edge$ (**c**).

Figure 5.12 illustrates the process of extracting intensity profiles from one of the corpus callosum images (**a**). In a first step the image is smoothed with a Gaussian filter of $\sigma = 1pixel$ to make the latter optimization procedure less sensitive to noise and other disturbances in the image (**b**). Since the Fourier parametrization keeps proportions of distances in parameter and object space equal, thus, subdividing the parameter space into equal intervals profiles can be placed equidistantly along the contour. For every contour point i in each data set j we can extract a profile $\boldsymbol{w}_{ij}$ of n_p sample points. Sampling distances are designed to be smaller along the profiles than along the object contour to reflect the fact that abrupt changes in intensity are only expected orthogonal to the boundary. Profiles are sampled at one pixel distance in a stripe of ± 5 pixels while the distance between adjacent profiles is chosen to be about 2 pixels.

The profiles are oriented normal to the object boundary and centered at the contour points $\boldsymbol{x}_{ij}$, as illustrated in Figure 5.12 **c**. For each sample point i we can obtain a mean profile by averaging over the sample objects N:

$$\overline{\boldsymbol{w}}_i = \frac{1}{N}\sum_{j=1}^{N} \boldsymbol{w}_{ij}. \tag{5.10}$$

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.

5.5.1 Deformation modes of the contour points

The points of the sampled contour can be considered as a new representation of the same object which can be obtained from the Fourier descriptors by a linear transformation:

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{c} \ , \tag{5.11}$$

where $\boldsymbol{x}$ represents the coordinates in object space and $\boldsymbol{c}$ the Fourier descriptors. $\boldsymbol{A}$ consists of the function values $([\sin(kt), \cos(kt)],\ k = 0 \ldots K)$ of the Fourier parametrization, one for each dimension, and describes the mapping between shape description space and object space coordinates.

Deformation modes have been previously derived based on the Fourier coefficients. To examine how the modes can be converted to the contours' description in object space, we investigate the covariance matrix. In parameter space it is defined by

$$\boldsymbol{\Sigma}_c = \mathrm{Cov}[d\boldsymbol{c}] = \mathrm{E}[d\boldsymbol{c}\, d\boldsymbol{c}^T] \ . \tag{5.12}$$

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

$$\boldsymbol{\Sigma}_x = \mathrm{Cov}[d\boldsymbol{x}] = \mathrm{E}[d\boldsymbol{x}\, d\boldsymbol{x}^T] = \mathrm{E}[\boldsymbol{A}d\boldsymbol{c}\, d\boldsymbol{c}^T\boldsymbol{A}^T] = \boldsymbol{A}\boldsymbol{\Sigma}_c\boldsymbol{A}^T \ , \tag{5.13}$$

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 $\boldsymbol{\Sigma}_x$

$$\boldsymbol{\Sigma}_x\boldsymbol{P}_x = \boldsymbol{\Lambda}_x\boldsymbol{P}_x \ , \tag{5.14}$$

and substituting Eq. 5.11 and Eq. 5.13 into Eq. 5.14 and multiplying both sides by $\boldsymbol{A}^T$, we obtain

$$\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\Sigma}_c\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{P}_c = \boldsymbol{\Lambda}_x\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{P}_c \ . \tag{5.15}$$

Comparing Eq.5.15 with Eq. 5.2, 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 our case, the columns of $\boldsymbol{A}$ are regularly sampled versions of $\cos(kt)$ and $\sin(kt)$ functions (with $k = 0 \ldots K$, 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:

$$\begin{aligned} \boldsymbol{P}_x &= \boldsymbol{A}\boldsymbol{P}_c && (5.16) \\ \boldsymbol{\Lambda}_x &= \alpha\boldsymbol{\Lambda}_c \, . && (5.17) \end{aligned}$$

Figure 5.13 explains the relationships between contour parametrizations introduced above. In the following the vector $\boldsymbol{c}$ will be referred to as *Fourier descriptors* and vector $\boldsymbol{x}$ as *Contour points*. The latter can either be thought of as a local parametrization of the shape or as its reconstruction from the *Fourier descriptors*. On the right side in figure 5.13 the actual shape descriptors are the elements of vector $\boldsymbol{b}$, while the average shape and the eigenvectors of the covariance matrix build the basis of the parametrization and can be expressed either in $\boldsymbol{c}$ or $\boldsymbol{x}$. This will be referred to as *Statistical shape model* and vector $\boldsymbol{b}$ as *Shape parameters*.

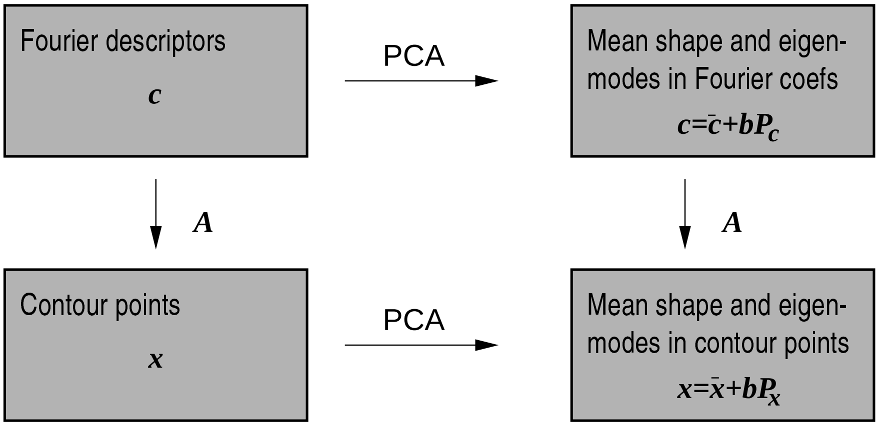


Figure 5.13: Overview of contour descriptions and how they are obtained from each other. Starting with a Fourier representation $\boldsymbol{c}$, contour points $\boldsymbol{x}$ can be computed at equally sampled parameter values. Having a set of similar contours, a more compact statistical description can be derived using principal component analysis.

5.5.2 New optimization strategy

The initialization step of the elastic fitting process remains the same as before while the optimization strategy has been changed to better accommodate with the augmented model. To give an initial estimate of the model contour in a new image the user has to specify the location of the AC/PC coordinate system by selecting the positions of the *anterior commissure* (AC) and *posterior commissure* (PC). Having the average location of the average shape, the position of the curve is iteratively refined until no significant changes result. Instead of using a multidimensional optimization which simultaneously modified all global shape parameters the new position of each contour point is determined individually based on its intensity profile before the global shape gets updated. Each iteration step can be divided into two stages:

- At each point of the model contour the image intensity environment is examined to find a displacement which moves the point to a better position.
- From these local displacements a global deformation of the shape is calculated which tries to satisfy all local adjustments without violating the shape constraints.

5.5.3 Calculating a set of displacements for the contour points

Starting with the average contour and given a profile model for each model point the search involves finding a nearby region which better matches the profile model. At a particular model point the algorithm extracts a profile, $\boldsymbol{w}$, from the processed image of some length, larger than that of the profile model, centered at the point and normal to the boundary. The profile model ($\bar{\boldsymbol{w}}$) is shifted along the extracted profile to find the point at which the model matches best. The fit of the model at a position s along the extracted profile can be calculated as follows;

$$f(s) = (\boldsymbol{w}(s) - \bar{\boldsymbol{w}})^T (\boldsymbol{w}(s) - \bar{\boldsymbol{w}}) \ , \tag{5.18}$$

or

$$f(s) = (\boldsymbol{w}(s) - \bar{\boldsymbol{w}})^T \boldsymbol{\Sigma}_w^{-1} (\boldsymbol{w}(s) - \bar{\boldsymbol{w}}) \ , \tag{5.19}$$

where $\boldsymbol{w}(s)$ is a sub-interval of $\boldsymbol{w}$ of length n_p centered at s, and $\boldsymbol{\Sigma}_w$ denotes the covariance matrix of the extracted profile. $f(s)$ in eq. 5.19 is the Mahalanobis distance of the sample from the mean intensity model and for normally distributed data it is proportional to the probability of obtaining $\boldsymbol{w}(s)$ from the measured distribution of image intensities. Using the latter equation to compute the fit has the advantage that values in the profile model having higher variances over the training set (being more unreliable) get weaker weights and hereby have less influence on the fit.

5.5.4 Calculating the adjustments to the shape parameters

Using the method described above a set of suggested displacements ($d\boldsymbol{x}$) can be calculated, one for each contour point. The word *"suggested"* implies that the statistical shape model cannot fulfill all displacements since there are only a limited number of deformation modes available and $d\boldsymbol{x}$ can move the contour points in as many different degrees of freedom as many contour points the shape model consists of. In other words, the dimensionality of $\boldsymbol{x}$ is larger than the dimensionality of $\boldsymbol{b}$, such that the suggested displacements are not in general consistent with the statistical shape model and only an approximation to the required deformation can be achieved.

The segmentation procedure starts with $\boldsymbol{b}_{n=0} = \boldsymbol{0}$ at iteration step $n = 0$, setting the shape to the average model:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{b}\boldsymbol{P}_{\boldsymbol{x}} = \bar{\boldsymbol{x}} \ . \tag{5.20}$$

In each iteration step an adjustment vector to the shape parameters, $d\boldsymbol{b}$, is computed which will best match the model to the suggested new positions. This is achieved by minimizing the differences between model points ($\boldsymbol{x}_{n-1} + d\boldsymbol{x}'_n$) and desired points ($\boldsymbol{x}_{n-1} + d\boldsymbol{x}_n$) in a least squares manner, such that

$$\boldsymbol{x}_{n-1} + d\boldsymbol{x}'_n = \bar{\boldsymbol{x}} + (\boldsymbol{b}_{n-1} + d\boldsymbol{b}_n)\boldsymbol{P}_{\boldsymbol{x}} \ . \tag{5.21}$$

In the special case when $\boldsymbol{P}_{\boldsymbol{x}}$ is orthogonal $d\boldsymbol{b}_n$ can be obtained from $d\boldsymbol{x}_n$ by

$$d\boldsymbol{b}_n = {\boldsymbol{P}_{\boldsymbol{x}}}^T d\boldsymbol{x}_n \ . \tag{5.22}$$

5.5.5 Applying constraints to the shape parameters

There are two ways to constrain the resulting shape. On the one hand, taking only the first few largest deformation modes and truncating weaker eigenvectors from the shape representation limits the range of possible deformations to those strongly present in the training set. On the other hand, a shape can be considered acceptable if its Mahalanobis distance from the average shape D_{m} is less than a suitable constant $D_{\max}$

$$D_{\mathrm{m}}^2 = \sum_{k=1}^{t} \frac{b_k^2}{\lambda_k} \leq D_{\max}^2 \ , \tag{5.23}$$

where t is the number of modes in the shape representation and λ_k are the corresponding eigenmodes. Equation 5.23 defines a hyper ellipsoid that the vector $\boldsymbol{b}$ should lie within. In case, updating $\boldsymbol{b}$ leads to an implausible shape, ie. $D_{\mathrm{m}} > D_{\max}$, $\boldsymbol{b}$ can be rescaled to lie on the closest point of the allowed hyper ellipsoid's surface using

$$\boldsymbol{b}_{\mathrm{new}} = \boldsymbol{b}_{\mathrm{old}} \frac{D_{\max}}{D_{\mathrm{m}}} \ . \tag{5.24}$$

However, projecting $\boldsymbol{b}$ back to the surface of the allowable shape domain can be considered as a hard constraint which often leads to oscillation instead of stable convergence. This can be avoided using soft constraints by adding the term αD_{m} to the goal function of the optimization. Thus, the difference of model and extracted profiles are minimized together with the shape's Mahalanobis distance.

5.5.6 Results and Validation

Within the confines of the BIOMORPH project, 71 new cases of *corpora callosa* have been manually segmented by a medical expert. Since these new segmentations are of better quality than our previous non-expert models, the original 30 cases have been replaced by the new ones, to build new shape and profile models.

Figure 5.14a illustrates the previously introduced (in section 5.1.1) quality measure curve for the new larger training set. It depicts the average description error ϵ for shapes not included in the training set when using the 10 largest eigenmodes while incresing

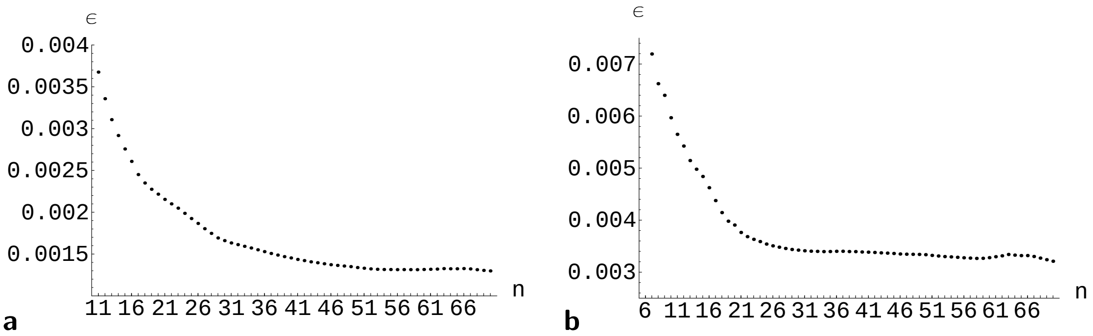


Figure 5.14: Quality measure curve computed for the 71 *corpura callosa* training set. It displays the average representation error ϵ for shapes not included in the training set when using the 10 largest eigenmodes while incresing the number of shapes involved in the computation from $n = 11$ to $n = 70$ on the **left**, and using the 5 largest eigenmodes from $n = 6$ to $n = 70$ on the **right**.

the number of shapes involved in the computation from $n = 11$ to $n = 70$. We can observe that incorporating more then about 50-55 shapes into the model does not lower considerably the average representation error. It follows that all linear variations expressed by 10 eigenmodes can be obtaind from a limited number of training shapes. Figure 5.14b shows the curve if we settle for only the largest 5 deformation modes. In this case the final representation error which the curve converges to is about twice as high as for the 10-eigenmodes model, on the other hand, only a smaller training set is needed to reach saturation. This points to the fact that larger deformation modes can be reliably predicted from smaller number of sample shapes, while more traning data is needed to obtail less significant deformation modes.

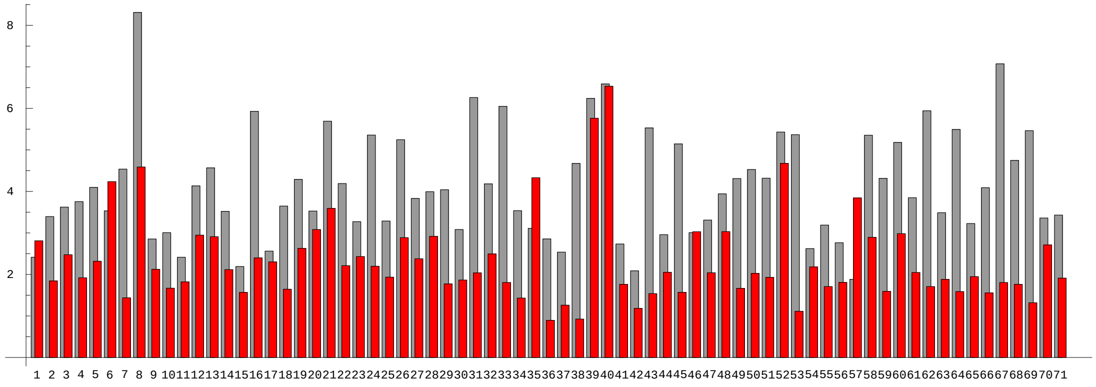


Figure 5.15: Average distances of corresponding contour points in pixels between manually and automatically segmented contours expressed in pixels. Bars in light grey illustrate the measure at initialization and in dark grey after deformation.

The new profile based segmentation technique has been applied to the 71 new mid-sagittal brain images. In the initialization step the location of the AC and PC landmarks has been determined by an expert manually. Having the AC/PC positions, all images

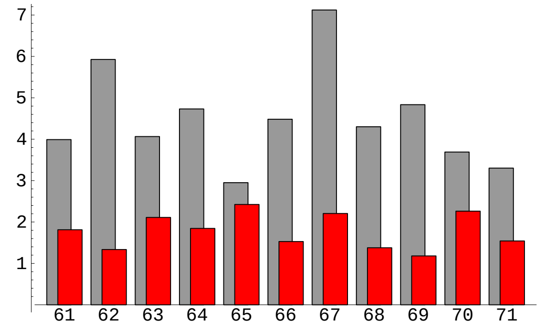


Figure 5.16: Average distances of corresponding contour points in pixels between manually and automatically segmented contours expressed in the length of a pixel edge for the 11 images not included in the training set. Bars in light grey illustrate the measure at initialization and in dark grey after deformation.

have been automatically segmented using 9 deformation modes and $\pm 2\sqrt{\lambda_i}$ bounds on the weights in $\boldsymbol{b}$. The method converged within 20 iterations steps for all of the cases taking less then one second on a Sun Ultra 1 workstation. To visualize the results the average distance of corresponding points has been determined between manual segmented contours and initial average, respectively fitted models. These distances are shown in Figure 6.17 for all 71 cases measured in pixels. Light gray bars correspond to the distance measure at initialization, while dark bars to the fitted model.

Consulting Figure 6.17 and visually comparing results achieved by manual and automatic segmentation, we found that the method succeeds in 64 of the 71 cases (90 %). Successful segmentation results are illustrated in Figure 5.17 and five segmentation failures in Figure 5.18. Images in the left column show initial average models (white lines) along with manually segmented contours (black lines). Images on the right depict the results of the fit (white lines) and also manually segmented contours (black lines).

Analyzing segmentation failures, we observed two reasons for the process going wrong. Firstly, if the contour of the object lies too far from the average (as in case 8), we can increase the number of deformation modes and relax the constraints to get closer to the object. Secondly, if the image is of low contrast (as in cases 35, 39, 40, and 52) one could include textural information about the object of interest and its background into the objective function to achieve better selectivity.

In order to test the robustness of the statistical model, we applied the technique to images not included in the training set. To do so, only the first 60 cases have been used to build the model. This model has then been used to segment the remaining 11 cases. Comparing the results shown in figure 5.16 we can see that excluding these cases from the model building stage does not seem to affect their segmentation.

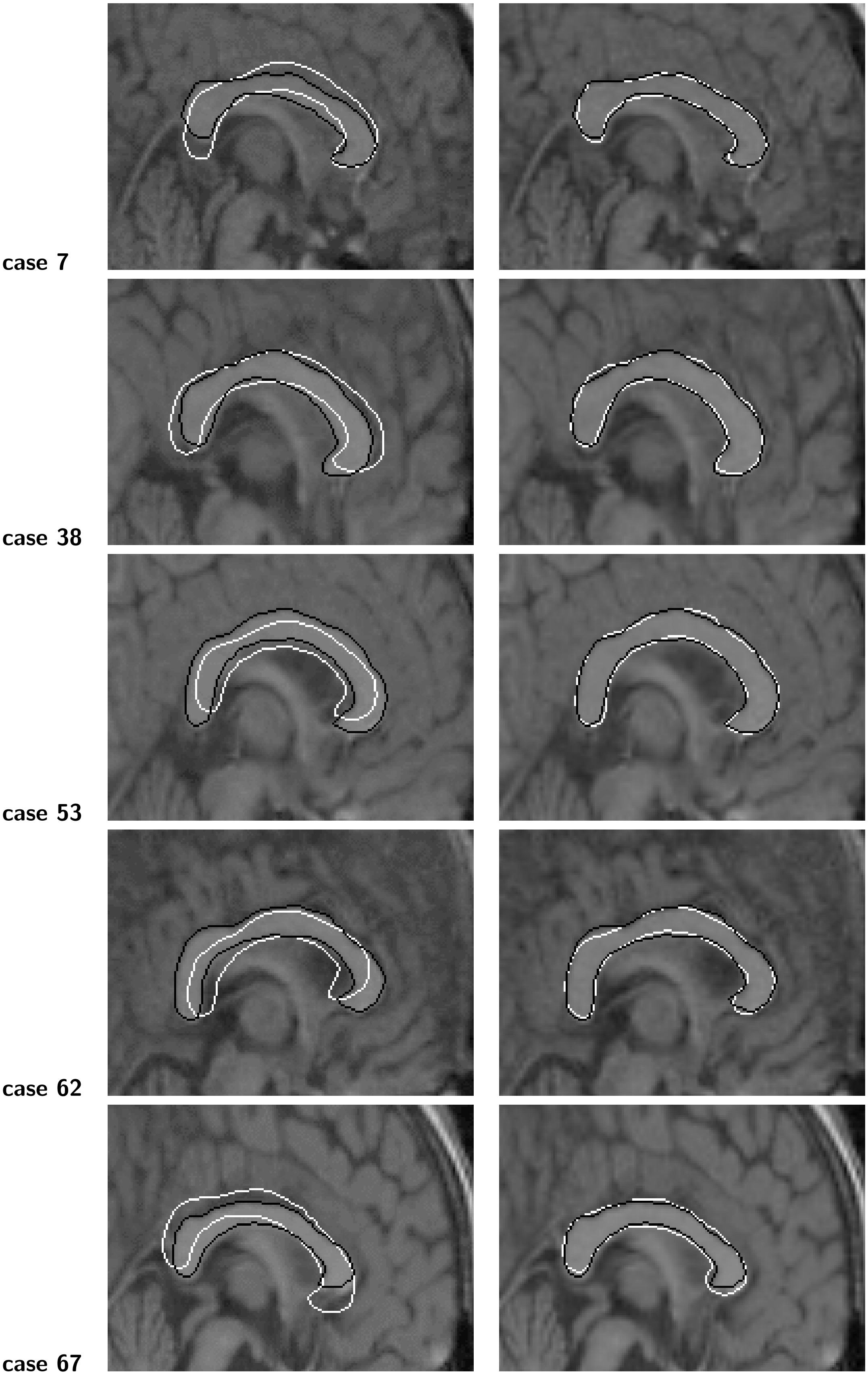


Figure 5.17: Successful segmentations. Images in the left column show initial average models (white lines) together with manually segmented contours (black lines). Images on the right depict the results of the fit (white lines) and also manually segmented contours (black lines).

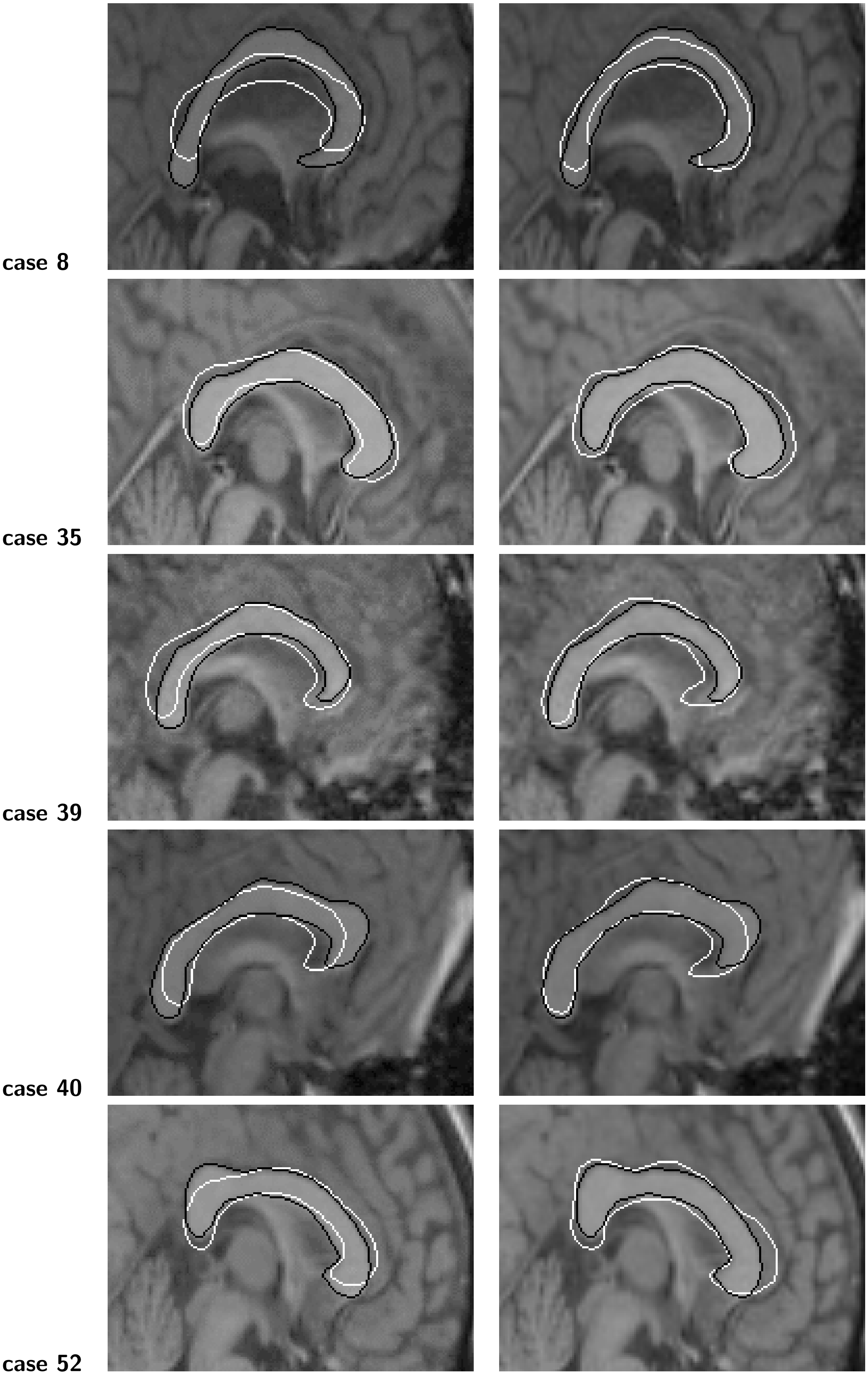


Figure 5.18: Segmentation failures. Images in the left column show initial average models (white lines) together with manually segmented contours (black lines). Images on the right depict the

6

Model-based segmentation of volumetric data

In the previous chapter we have seen that global parametrization can in certain cases successfully replace the tedious task of manual annotation of contour points in establishing correspondences between 2-D contours. It has also been shown that global parametrization techniques are not limited to 2-D contours but has recently been extended to surfaces of spherical topology as well. Since the only 2-D specific part of the proposed segmentation method is the Fourier parametrization, replacing it by an equivalent 3-D technique provides the straightforward generalization of the segmentation method to three dimensions. While the fact that the used surface parametrization technique can only be applied to simply connected objects is a true limitation, there is still a vast variety of applications, such as segmentation of anatomy, where objects of spherical topology are typical.

The 3-D segmentation discussed here is based on a statistical model, generated from a collection of manually segmented volumetric brain MR images 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.

By means of 3-D brain regions, the generation of the statistical model will be discussed in detail in the following section. Similarly to the 2-D method the purely geometrical statistical model has been extended by incorporating gray-valued profiles across the organ surface, implementing the concept proposed by [Cootes *et al.* 1994a, Cootes *et al.* 1995a] 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 6.2, 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 perpendicular to the surface, which can be calculated in an elegant way by using a dual representation of the object both as a collection of sample points and as a parametrized surface.

6.1 Generation of 3-D statistical model

The concept proposed here 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 ([Brechbühler *et al.* 1992, Brechbühler *et al.* 1995]) as reviewed in chapter 4. 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.

In the medical field brain regions usually fulfill the requirement for sphere topology. On the other hand, because of the poor contrast conditions deep gray brain regions show, they represent a real challenge to human experts.

6.1.1 Interactive expert segmentation

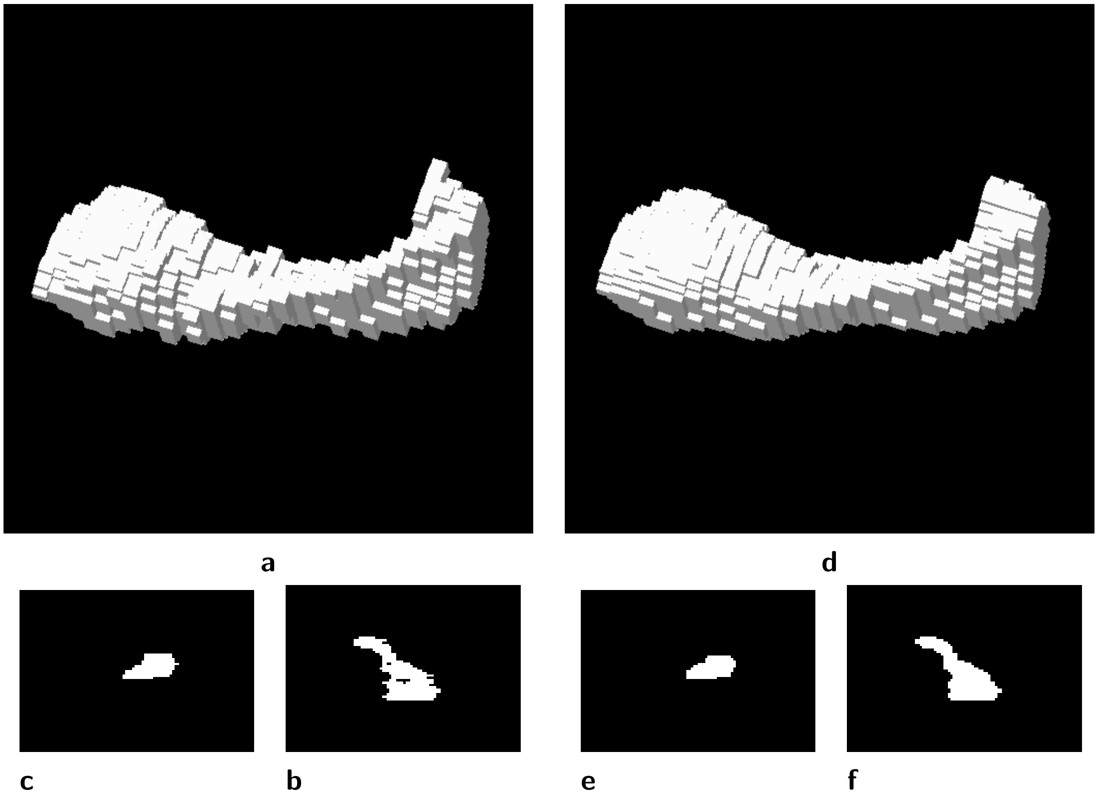


Figure 6.1: Preprocessing of the segmented volume in order to obtain smoother surfaces. Image **a** illustrates the raw volume reconstructed from its 2-D segmentation slices, while images **b** and **c** show two slices across the volume, one in the segmentation plane and the other one orthogonal to it. On the right the same representations can be seen, now taken from the preprocessed volume.

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

Our initial training set consisting of 30 age matched male brain MR volumes, a courtesy of Ron Kikinis and Martha Shenton at Harward Medical School in Boston, has also been processed this way. In each volume six brain regions has been manually labeled in both hemispheres, these regions include the *amigdala/hippocampus complex*, the *parahippocampual gyrus*, *thalamus*, *caudate nucleus*, *putamen*, and *globus pallidus*.

Figure 6.1 **a** illustrates the result of an expert segmentation of the left hippocampus. Images **b** and **c** display two orthogonal slices through the segmented volume and illustrate a typical effect usually seen in volumes created from 2-D slices. While **b** parallel to the segmentation plane shows a continuous contour, the slice parallel to it (**c**) illustrates the high variation between adjacent segmentation slices. Since the surface parametrization technique converges faster for smooth surfaces, all segmented volumes have been preprocessed to correct for the high inter-slice variability. To smoothen the surface a Gaussian filter of $\sigma = 0.5$voxel has been applied to the original binary image followed by thresholding to re-binarize the volume. The result is shown in Figure 6.1 **d**, **e**, and **f**.

6.1.2 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.* 1995] developed a surface parametrization of arbitrary simply connected objects based on those two parameters as it has been reviewed in chapter 4.

6.1.3 Surface correspondence and object alignment

The surface parametrization, i.e., the representation of the surface by a parameter net with homogeneous cells, is 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. 6.2(b,c) illustrates the location of the middle main axis on the reconstruction up to degree one and ten respectively.

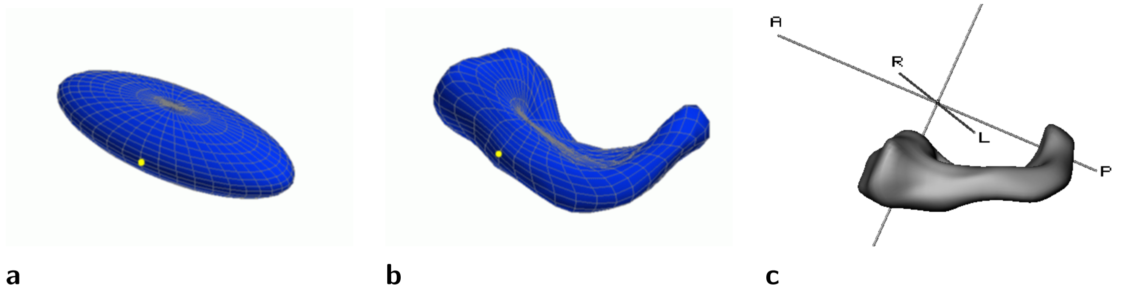


Figure 6.2: Model building: Reconstruction from shape descriptors up to degree one (a), reconstruction up to degree ten (b) and normalization in object space (c).

The symmetry of the first degree ellipsoid allows for eight different standard orientation of the object. Four of them can be explained by rotating the ellipsoid by π with respect to one of the three main axes. The other four arise from the rotated versions by mirroring them with respect to one of the three symmetry planes of the ellipsoid. The concatenation of two rotations or two mirrorings result in identity if they are done with respect to the same axis or symmetry plane. Further, two consecutive rotations in different axes give the same as one rotation in the third axis. Similarly, two mirrorings about different planes result in one rotation in the axis given by the intersection of the planes. This gives rise to eight different combinations. In the comparison of two descriptors, the minimal distance resulting from any of the eight flips is relevant. One reference parametrization has been chosen and the rest of the training set has been aligned to it. To achieve the alignment of one particular shape, the descriptors of all eight standard orientations has been compared with the reference descriptors by computing the Euclidean distance of the two description vectors and taking the one where this is minimum.

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 6.3). 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 assessment of the quality of point correspondences, with a view to potential improvements, are research questions of current interest

([Kotcheff and Taylor 1997, Tagare 1997, Rangarajan *et al.* 1997]).

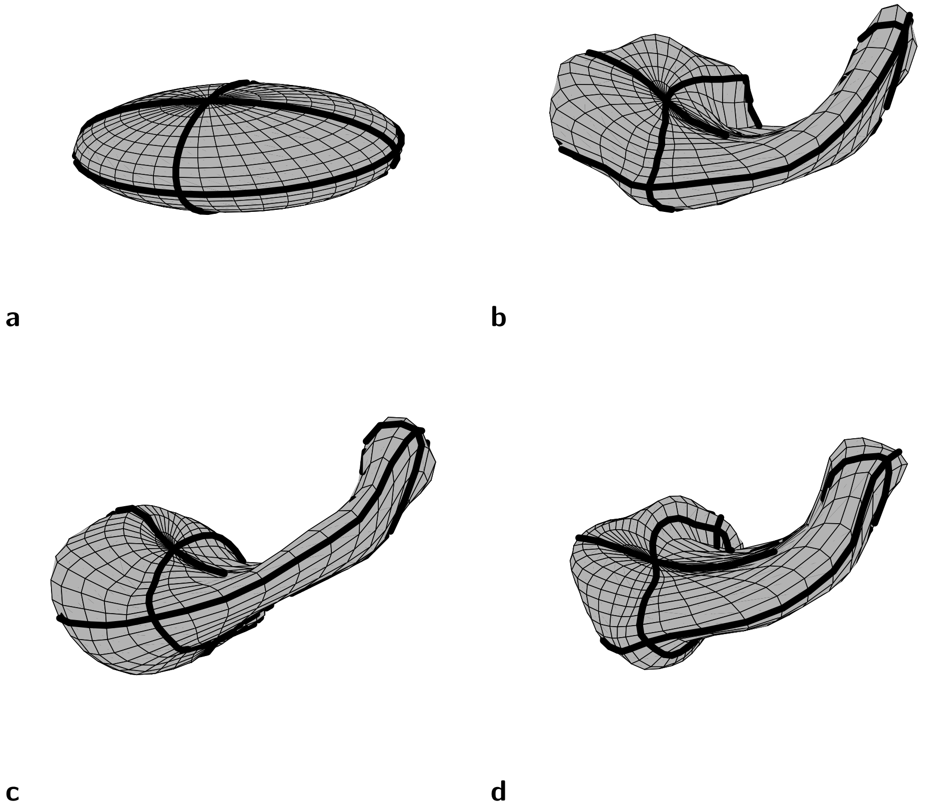


Figure 6.3: Corresponding parameter values for $\theta = \pi/2$, $\phi = 0, \pi$, and $\phi = \pi/2, 3\pi/2$ (thick lines) illustrated on an ellipsoid (a) and on three individual left-hippocampi.

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 and Tournoux 1988] for global alignment of the head image data sets. This can be seen as an extension of the coordinate system we have used for the *corpus callosum*. To be consistent with the literature the coordinate origin has been moved to the middle of the AC and PC points. Axis y traverses the brain from back (negative values) to front crossing the two feature points. Lying also in the midsagittal plane axis z is perpendicular to y and points from bottom to top. The third axis x, pointing from left to right, is defined to be orthogonal to the other two such that they build a right handed coordinate system. 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. 6.4). Each data set is transformed into canonical coordinates by 3-D rotation and scaling.

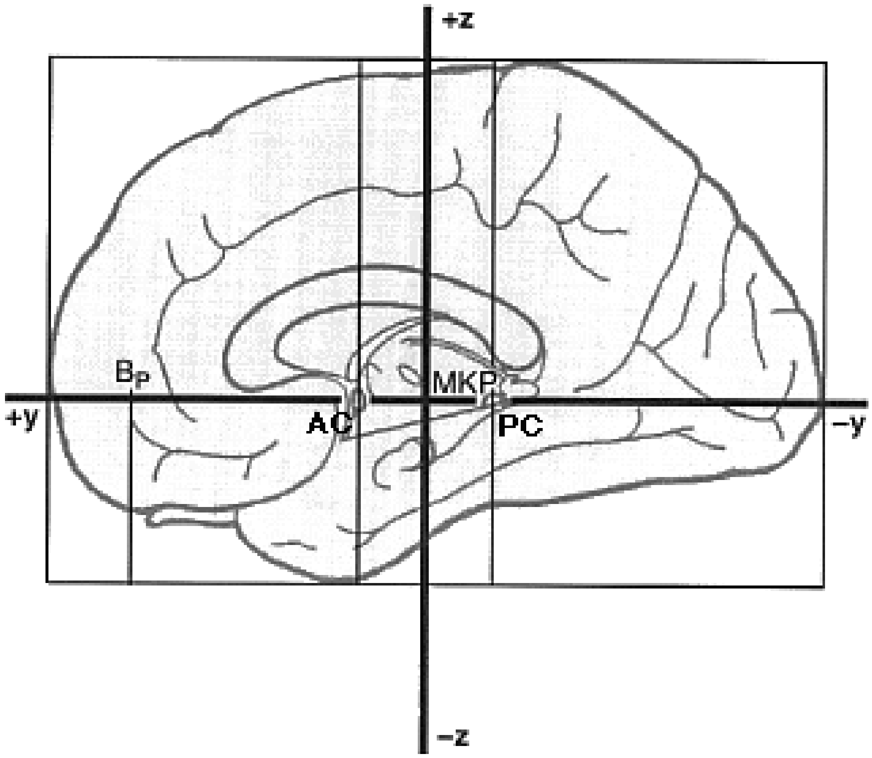


Figure 6.4: Stereotactic coordinate system used for object space normalization.

6.1.4 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 our 2-D method as described in the previous chapter, where the calculation are carried out in the domain of shape descriptors rather than the Cartesian coordinates of points in object space. This computation of deformation modes is basically identical to that of the 2-D case and is repeated here briefly for the sake of completeness. The mean model is determined by averaging the descriptors $\boldsymbol{c}_j$ of the N individual shapes (see Fig. 6.5).

$$\bar{\boldsymbol{c}} = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{c}_j \tag{6.1}$$

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

$$\boldsymbol{\Sigma} = \frac{1}{N-1} \sum_{j} (\boldsymbol{c}_j - \bar{\boldsymbol{c}}) \cdot (\boldsymbol{c}_j - \bar{\boldsymbol{c}})^T \tag{6.2}$$

$$\boldsymbol{\Sigma} = \boldsymbol{P_c} \boldsymbol{\Lambda} \boldsymbol{P_c}^T = (\boldsymbol{P_c} \boldsymbol{\Lambda}^{1/2})(\boldsymbol{P_c} \boldsymbol{\Lambda}^{1/2})^T, \tag{6.3}$$

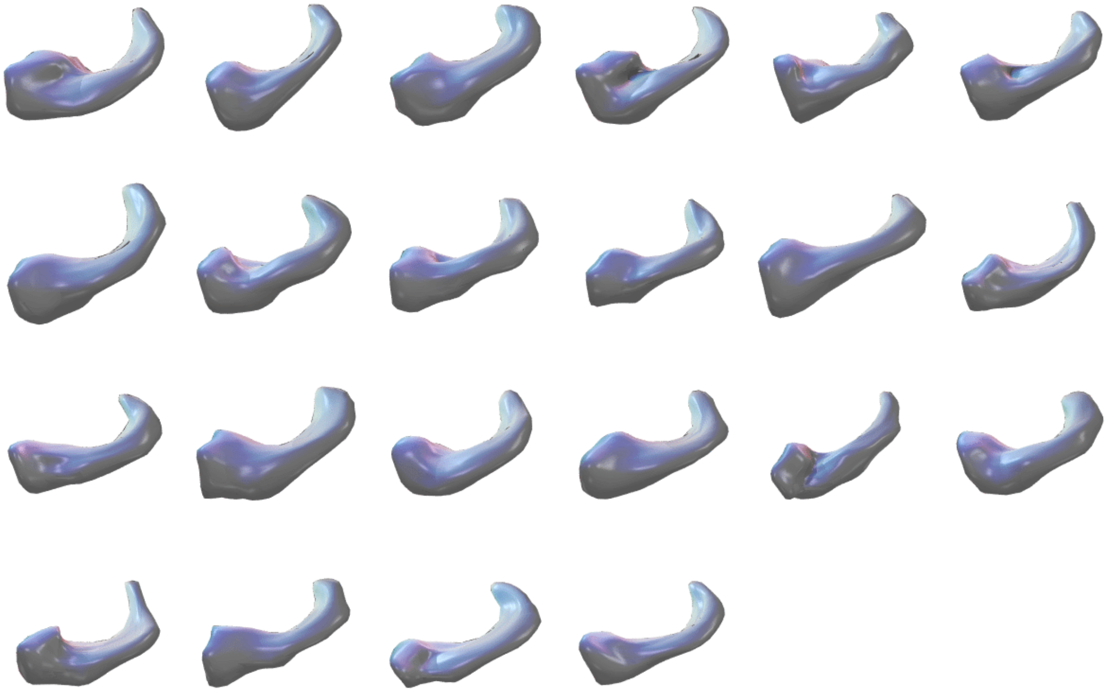


Figure 6.5: Illustration of all 22 left hippocampal structures of the training sets, normalized and reconstructed from their descriptors.

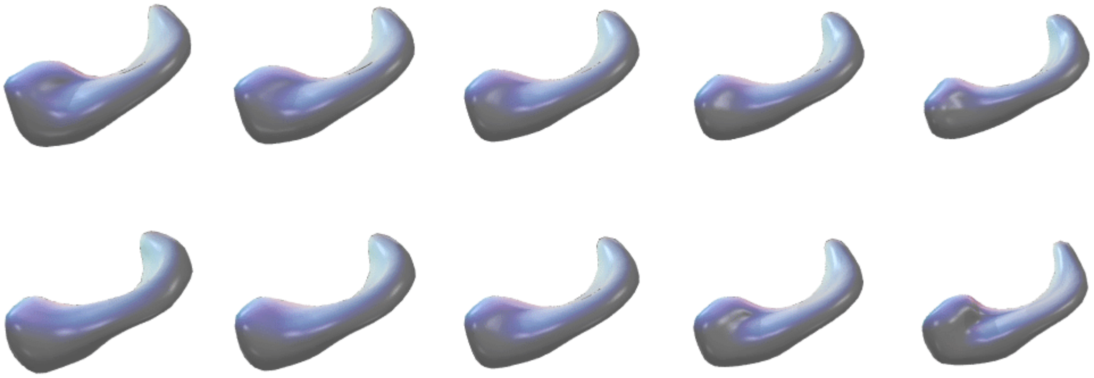


Figure 6.6: Largest two modes of variation for $b_j = -2\sqrt{\lambda_j} \ldots 2\sqrt{\lambda_j}$. In the middle column, $b_j = 0$ represents the mean model.

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

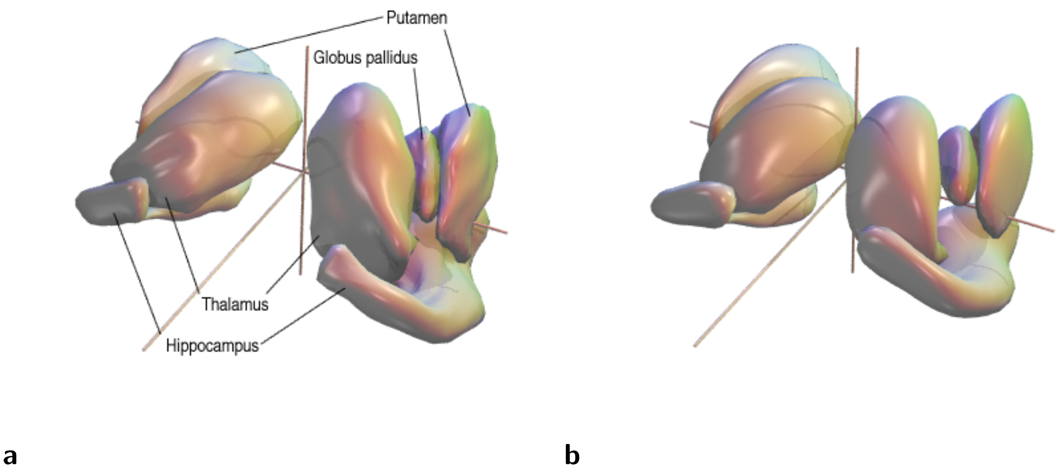


Figure 6.7: Left and right thalamus, globus pallidus, putamen and hippocampus in one individual case (**a**) and their average models computed from all 30 cases (**b**).

$$\boldsymbol{c}_j = \bar{\boldsymbol{c}} + \boldsymbol{P_c}\boldsymbol{b}_j \tag{6.4}$$

Figure 6.6 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 6.8 illustrates the square root of eigenvalues sorted by size (dotted line) and the components of one individual vector $\boldsymbol{b}_j$.

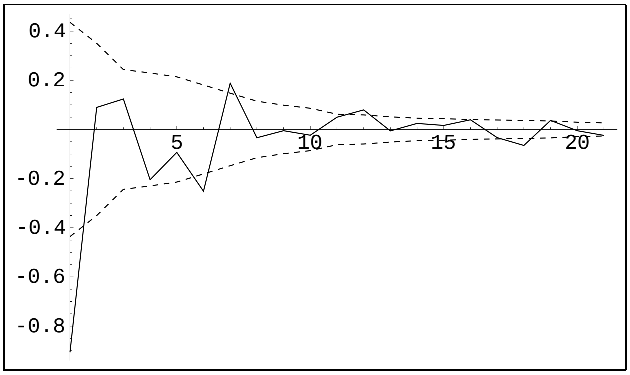


Figure 6.8: Statistics of shape deformation. The dotted line represents the square root of eigenvalues $\sqrt{\lambda_j}$ sorted by decreasing size. The continuous line illustrates the components of an individual vector b_j, which describes the deviation of the shape $\boldsymbol{c}_j$ from the mean shape $\bar{\boldsymbol{c}}$.

6.1.5 Quality measure of the 3-D statistical models

Figure 6.9a illustrates the previously introduced (in section 5.1.1) quality measure curves for the available 3-D training sets of left and right *hippocampus*, *putamen*, *globus pallidus*, and *thalamus*. As it can be seen, non of the eight models has reached the optimal error minimum. This means that adding new individuals could further increase the models' information content and thus, their ability to describe unseen shapes of the same kind. It also explains the difference in segmentation quality for shapes included and not included in the training set, which will be discussed in section 6.3.

6.1.6 Modeling the gray level environment of surface models

The geometric shape models are augmented 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 6.10(a) and Figure 6.2(b,c)). 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 Figure 6.10(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

$$\boldsymbol{x}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix} = \sum_{l=0}^{K} \sum_{m=-l}^{l} \boldsymbol{c}_l^m Y_l^m(\theta_i, \phi_i), \tag{6.5}$$

$$i = 1 \ldots 1002 \tag{6.6}$$

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

$$\boldsymbol{x} = \boldsymbol{A}\boldsymbol{c}, \tag{6.7}$$

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 normal to the object surface and centered at the surface points $\boldsymbol{x}_{ij}$, as illustrated in Figure 6.11. For each sample point i we can obtain a mean profile by averaging over the sample objects N:

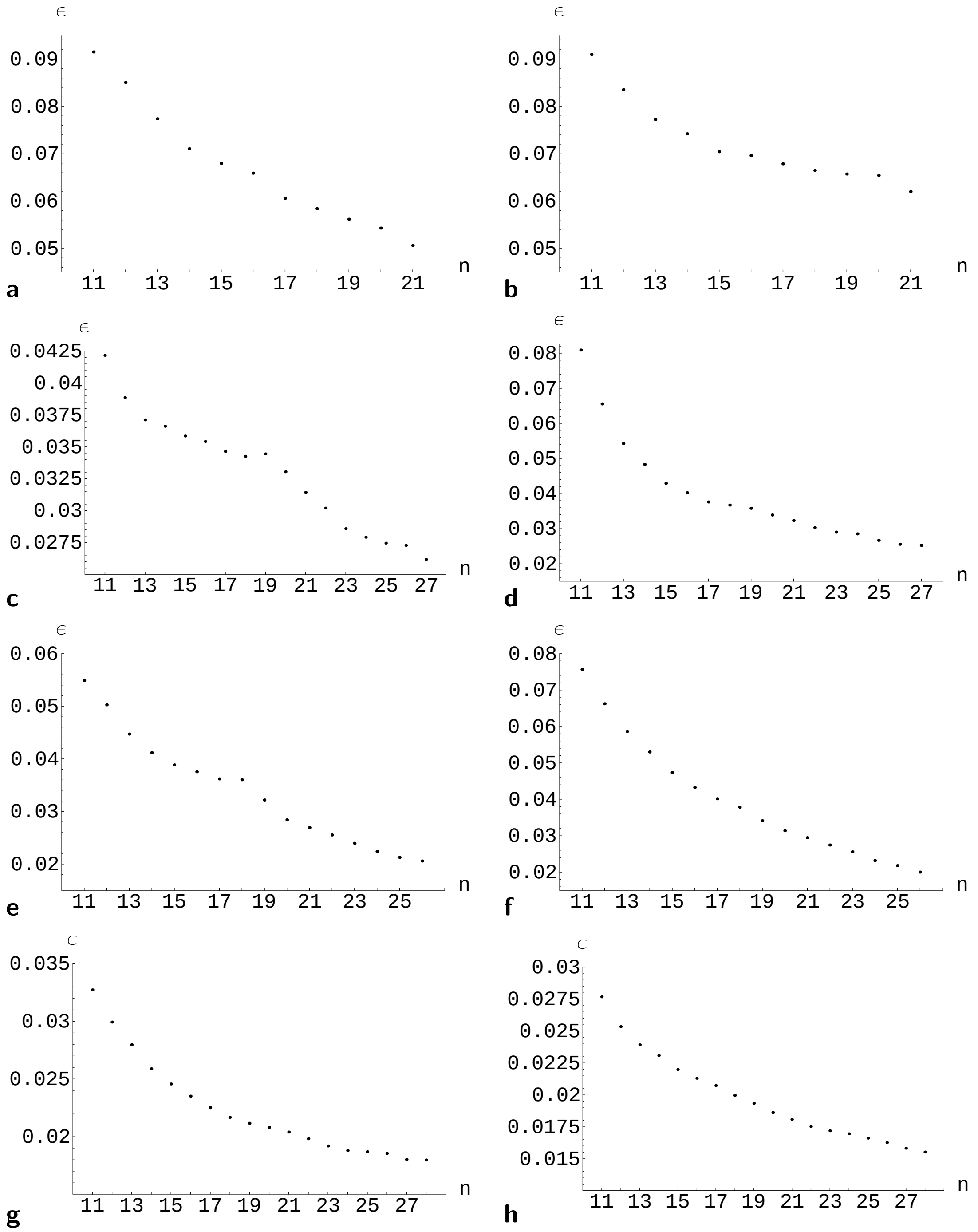


Figure 6.9: Quality measure curve computed for the training sets of left and right *hippocampus* (**a,b**), *putamen* (**c,d**), *globus pallidus* (**e,f**), and *thalamus* (**g,h**). It displays the average representation error ϵ for shapes not included in the training set when using the 10 largest eigenmodes while incresing the number of shapes involved in the computation from $n = 11$ to $n = N-1$, where N is the number of available shapes for the organ in question.

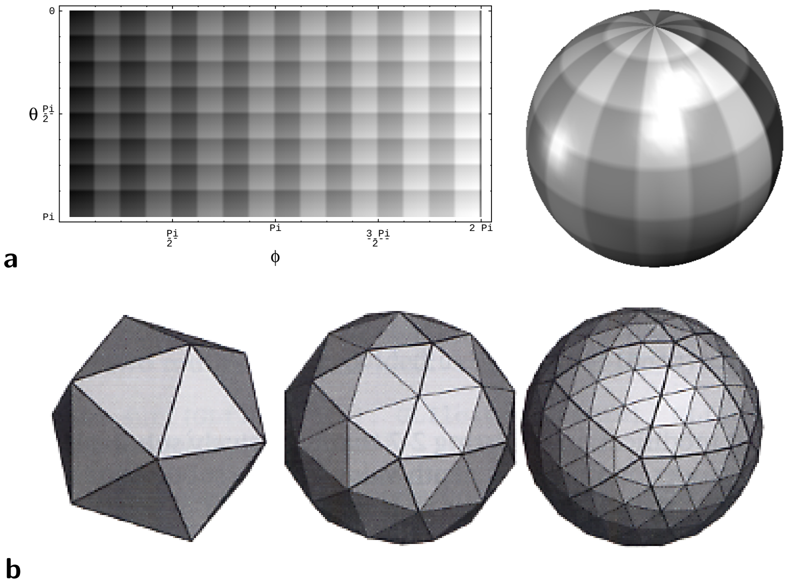


Figure 6.10: Sampling methods of spherical surfaces: regular mesh in spherical coordinates (a), icosahedron subdivision (b).

$$\bar{\boldsymbol{w}}_i = \frac{1}{N} \sum_{j=1}^{N} \boldsymbol{w}_{ij}. \tag{6.8}$$

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:

$$\boldsymbol{\Sigma}_{\boldsymbol{w}_i} = \frac{1}{N-1} \sum_{j=1}^{N} (\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i)(\boldsymbol{w}_{ij} - \bar{\boldsymbol{w}}_i)^T \ . \tag{6.9}$$

Cootes et al. in [Cootes *et al.* 1995a] 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.

6.2 Segmentation by model fitting

Until now we have only described the creation of a flexible 3D 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 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.

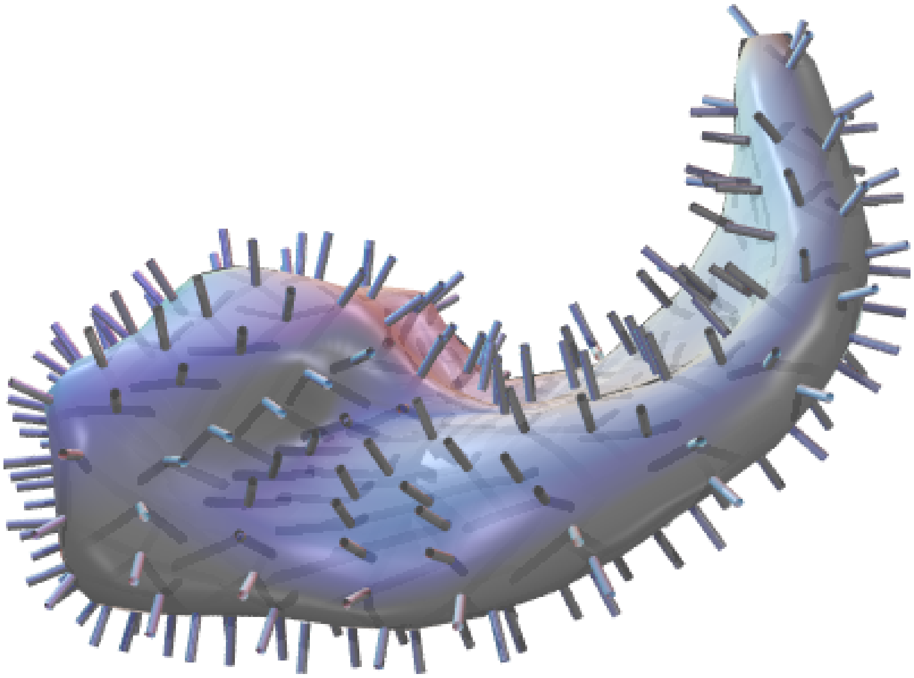


Figure 6.11: Illustration of an individual left hippocampal shape with its profile vectors shown from the left side of the brain.

- The elastic deformation of the surface until it best matches the new gray value environment.

6.2.1 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 the determination of the symmetry plane and the AC/PC line by can be replaced in future by an automatic method [Minoshima *et al.* 1993, Marais *et al.* 1996, Thirion *et al.* 1997, Kruggel and Lohmann 1997]. To derive the position of the midsagittal plane the user specifies the position of 3 or more points lying on the inter-hemispheric tissue. The program computes the 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 in the same way it happened for the segmentation of the the *corpus callosum*. 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.

6.2.2 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

$$\boldsymbol{n}_i = \sum_{l=0}^{K}\sum_{m=-l}^{l} \boldsymbol{c}_l^m \frac{\partial Y_l^m}{\partial\theta} \times \sum_{l=0}^{K}\sum_{m=-l}^{l} \boldsymbol{c}_l^m \frac{\partial Y_l^m}{\partial\phi}. \tag{6.10}$$

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 $\boldsymbol{A}$.

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 $\boldsymbol{w}$ 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:

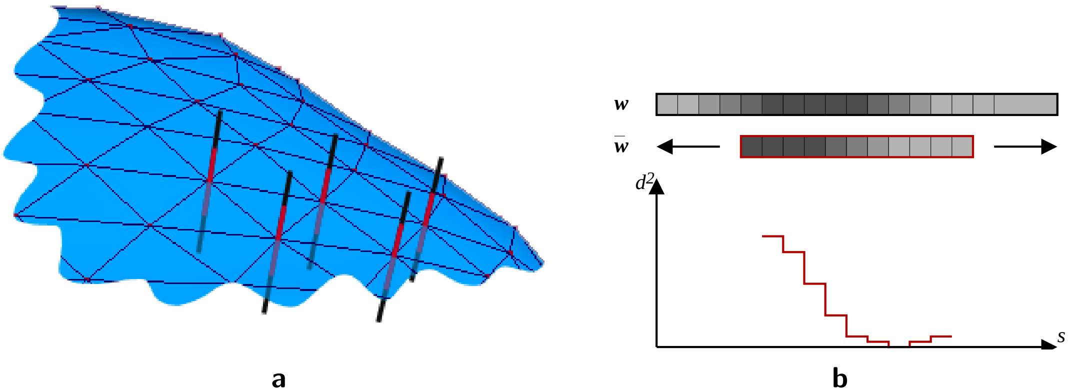


Figure 6.12: Illustration of the surface matching process. Image **a** shows part of the model's triangulated surface with longer extracted (in black) and shorter model profiles (in grey). Image **b** visualizes the computation of a suggested movement for a single surface point.

$$d^2_{Maha}(s) = (\boldsymbol{w}(s) - \bar{\boldsymbol{w}})\boldsymbol{\Sigma}_{\boldsymbol{w}}{}^{-1}(\boldsymbol{w}(s) - \bar{\boldsymbol{w}}) \tag{6.11}$$

where $\boldsymbol{w}(s)$ 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 $d^2_{Maha}(s)$. Suppose s_{best} is the shift between the two profiles providing the best fit. We choose a displacement vector $\boldsymbol{dx}$ for each model point which is parallel to the profile, in the direction of the best fit and has magnitude s_{best}. Figure 6.12 illustrates this process.

Adjusting the shape parameters

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

$$d\boldsymbol{x} = (dx_1, dy_1, dz_1, \ldots, dz_n), \tag{6.12}$$

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 $d\boldsymbol{x}$, 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.

The shape statistics, as described in section 6.1.4, can be expressed by

$$\boldsymbol{c} = \bar{\boldsymbol{c}} + \boldsymbol{P_c}\boldsymbol{b}\ . \tag{6.13}$$

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

$$\boldsymbol{A}\boldsymbol{c} = \boldsymbol{A}\bar{\boldsymbol{c}} + \boldsymbol{A}\boldsymbol{P_c}\boldsymbol{b} \tag{6.14}$$

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \boldsymbol{P_x}'\boldsymbol{b}\ , \tag{6.15}$$

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 $\boldsymbol{c}$. Contrary to 2-D, matrix $\boldsymbol{A}$ is non-orthogonal because of the unequal sampling of the spherical surface (compare Figures 5.13 and 6.13). Thus, $\boldsymbol{P_x}'$ is not a real matrix of eigenvectors since its column vectors are also non-orthogonal, although they still represent the same shape deformations as eigenvectors in $\boldsymbol{P_c}$. 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.

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.

$$(\boldsymbol{x} + d\boldsymbol{x}) = \bar{\boldsymbol{x}} + \boldsymbol{P_x}'(\mathrm{b} + d\boldsymbol{b})\ . \tag{6.16}$$

Subtracting eq.(6.15) from eq.(6.16) we get

$$d\boldsymbol{x} = \boldsymbol{P_x}'d\boldsymbol{b}\ . \tag{6.17}$$

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 5 to 10). Therefore a least squares approximation to the solution can be obtained using standard 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 non-orthogonality of $\boldsymbol{P_x}'$ does not allow of solving eq. 6.17 this simple way. With this object, the general purpose least squares routine F04JAF from the NAG^{TM} Fortran library has been applied to obtain $d\boldsymbol{b}$.

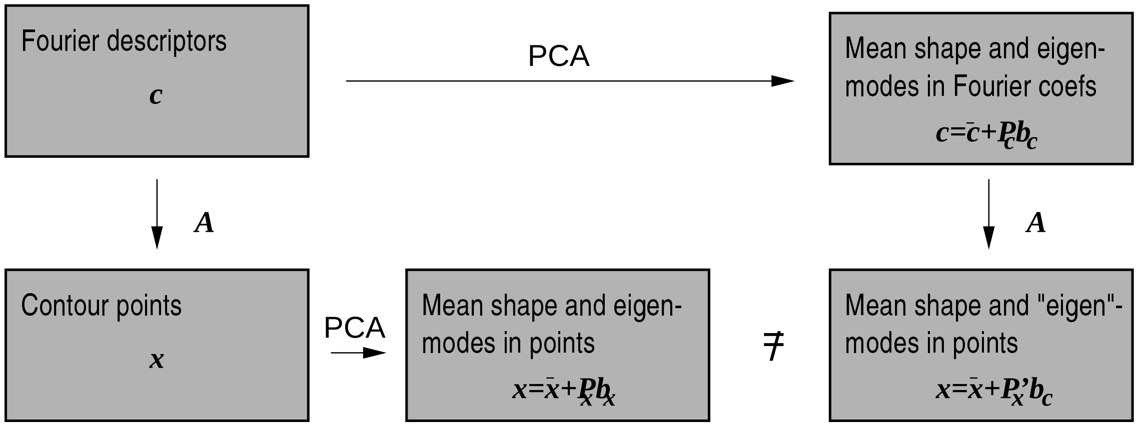


Figure 6.13: Overview of contour descriptions and how they are obtained from each other. Starting with a spherical harmonics representation $\boldsymbol{c}$, contour points $\boldsymbol{x}$ can be computed at sampled parameter values. Contrary to 2-D, reversing the order of multiplying by $\boldsymbol{A}$ and computing PCA yields different representations.

This more general algorithm can also be used in the special case of orthogonal $\boldsymbol{P_x}$ and would provide the same solution as equation 5.22.

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 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

$$\boldsymbol{b}_{t+1} = \boldsymbol{b}_t + d\boldsymbol{b}_t, \tag{6.18}$$

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. 6.8). 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.

6.2.3 Quality of gray level profiles

Cootes and Taylor propose to introduce weights into the calculation of the adjustments to shape parameters, $(d\boldsymbol{b})$. This results in modifying equation 6.17 to

$$\boldsymbol{P_x}^T \boldsymbol{Q} d\boldsymbol{x} = (\boldsymbol{P_x}^T \boldsymbol{Q} \boldsymbol{P_x}) d\boldsymbol{b}\ , \tag{6.19}$$

where $\boldsymbol{Q}$ is a diagonal matrix of weights, one for each surface point in $\boldsymbol{x}$. Weights can indicate the confidence we have in the suggested new position for the point. In their implementation weights in $\boldsymbol{Q}$ are inversely proportional to the length of $d\boldsymbol{x}_i$, the suggested displacement at surface point i:

$$q_i = \frac{1}{2 + |d\boldsymbol{x}_i|^2} \ . \tag{6.20}$$

The idea behind it is to give less weight for points which are found to be further away from the current model points than average, and may be outliers.

We developed a different weight model based on the quality of model profiles. For a successful segmentation model profiles have to fulfill the following two requirements:

- they have to be *selective*, meaning that they cross a strong edge or across several strong edges in the image, and
- *reliable*, meaning that $\boldsymbol{w}_i$ at a certain surface point i show low variability over individuals.

That is, model profiles $\bar{\boldsymbol{w}}_i$ of high selectivity and low variance are of better quality than ones with low selectivity and high variance. Selectivity is related to the derivative of the profile, $d\boldsymbol{w}_i$, while reliability can be expressed by the inverse of the covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{w}_i}$ (see equation 6.9) of the profile. The two can be combined into the quality measure

$$q_i = d\boldsymbol{w}_i^T \boldsymbol{\Sigma}_{\boldsymbol{w}_i}^{-1} d\boldsymbol{w}_i \ . \tag{6.21}$$

Visualizing this measure over the surface of the object, as shown in Figure 6.14 for the left hippocampus example, provides us information about the segmentation quality, we can expect at different locations.

6.3 Results

Figure 6.15(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. $\boldsymbol{A}$ represents the model shape obtained by human experts, $\boldsymbol{B}$ the result of the new model-based segmentation.

The overlap measure $(\boldsymbol{A} \cup \boldsymbol{B})/(\boldsymbol{A} \cap \boldsymbol{B})$ shown in Figure 6.16 is calculated on binary voxel maps, created by intersection of the object surfaces with the voxel grid. We therefore

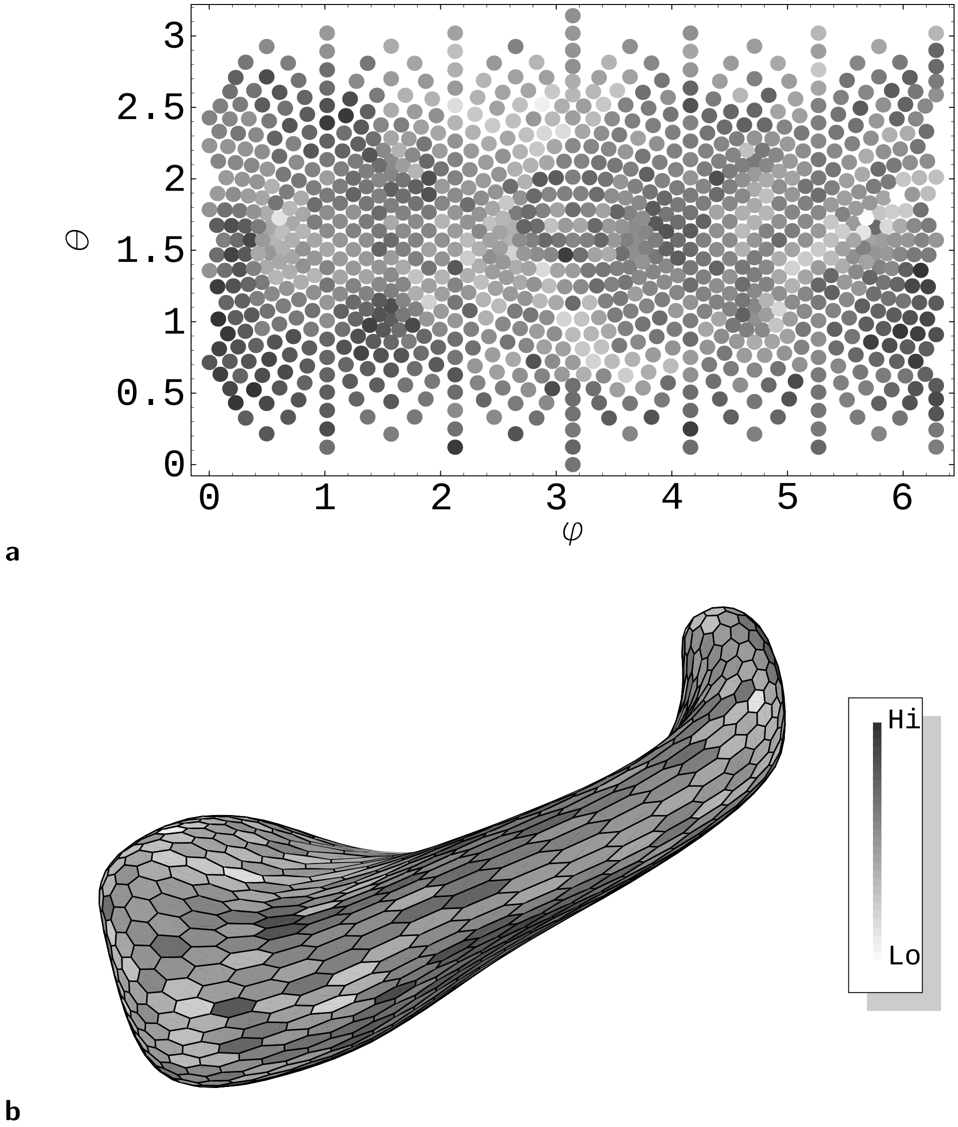


Figure 6.14: Quality measure of model profiles visualized in spherical coordinates ϕ and θ in image **a**, as well as over the surface of the left hippocampus model in image **b**. Darker dots correspond to higher profile quality.

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 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. Parseval's theorem says that the energy of the one-dimensional continuous signal $f(t)$ can

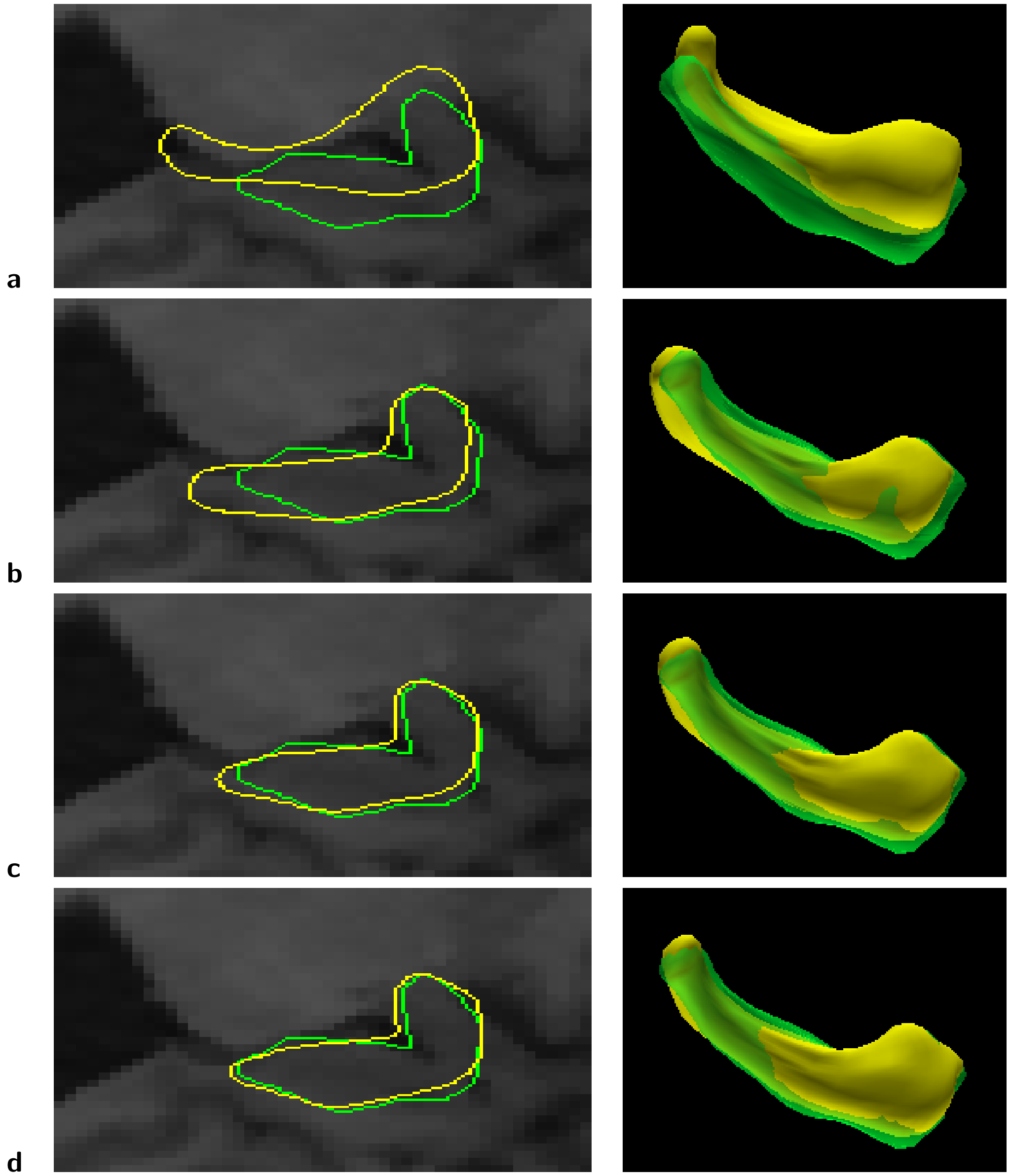


Figure 6.15: Segmentation results of a left hippocampus on sagittal 2D slices and 3D views from the left hand side. The images a-d were taken after 0, 30, 60 and 100 iterations.

be expressed by its Fourier coefficients (a_n, b_n) in the following way:

$$\frac{1}{T}\int_{-T}^{T}[f(t)]^2 dt = \frac{a_0^2}{2} + \sum_{1}^{\infty}(a_n^2 + b_n^2) \tag{6.22}$$

The equation also applies to other orthogonal basis functions, such as spherical harmonics, and to higher dimensions as well. Eq. 6.22 explains the computation of the average distance of a closed surface from the coordinate origin from the continuous coordinate functions $\boldsymbol{x}(\boldsymbol{u})$ and their descriptors $\boldsymbol{c}_l^m$.

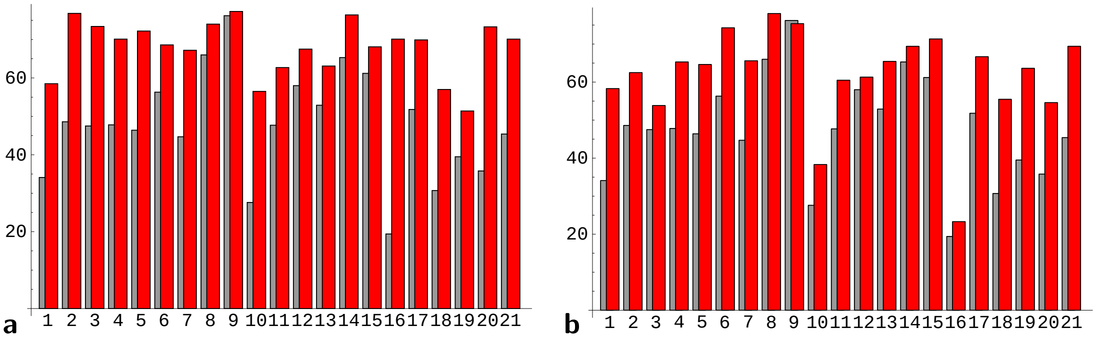


Figure 6.16: Overlap measure $(\boldsymbol{A} \cup \boldsymbol{B})/(\boldsymbol{A} \cap \boldsymbol{B})$ in percentage calculated between manually (A) and automatic (B) segmented left hippocampi of 21 individuals. Bars in light grey illustrate the measure at initialization and in dark grey after deformation.

$$\oint \parallel \boldsymbol{x}(\boldsymbol{u}) \parallel^2 d\boldsymbol{u} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mid \boldsymbol{c}_l^m \mid^2 = 4\pi \cdot \mathrm{MSD} \ , \tag{6.23}$$

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

$$\oint \parallel \boldsymbol{x}_1(\boldsymbol{u}) - \boldsymbol{x}_2(\boldsymbol{u}) \parallel^2 d\boldsymbol{u} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \mid \boldsymbol{c}_{l1}^m - \boldsymbol{c}_{l2}^m \mid^2 \quad . \tag{6.24}$$

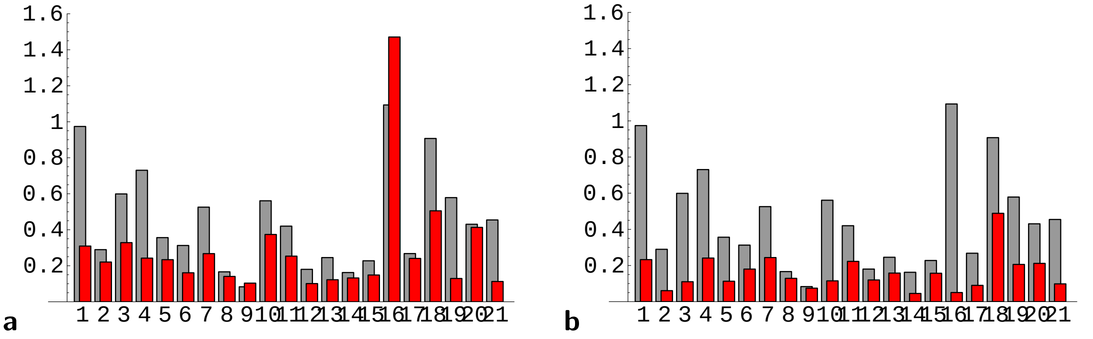


Figure 6.17: Average distances in mm calculated between manually and automatic segmented left hippocampi of 21 individuals. The bars in light grey illustrate the mean distance of the initialization of the model in a new data set and the dark bars the final mean distance of surfaces to the model surface. The length of the hippocampus is about 40 mm.

Figure 6.17 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.

6.4 Multi-shape models

We are particularly interested in building an atlas of neuroanatomy based on statistical surface models of brain regions. An atlas can be built by collecting single shapes into one model. To mathematically represent several organs as one model, their spherical harmonics shape descriptors can be concatenated into one single vector

$$\boldsymbol{c}_{j_{\text{multi}}} = \begin{pmatrix} \boldsymbol{c}_{j_1} \\ \vdots \\ \boldsymbol{c}_{j_k} \\ \vdots \\ \boldsymbol{c}_{j_K} \end{pmatrix} , \tag{6.25}$$

where $\boldsymbol{c}_{j_k}$ the parameter vector of shape k in individual j. Covariance matrix and deformation modes based on $\boldsymbol{c}_{i_{\text{multi}}}$ can be then similarly computed as for single shape models.

The above formulation requires that a complete segmentation exists for all individuals, in other words, it assumes that all organs have been segmented in all data sets. This, however, is not always the case, hence, we seek an alternative solution to compute the covariance matrix which can deal with incomplete data.

One can observe that the multi-shape covariance matrix $\boldsymbol{\Sigma}_{\text{multi}}$ is a block matrix and can be divided into single- or double-shape covariance matrices

$$\boldsymbol{\Sigma}_{\text{multi}} = \begin{pmatrix} \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1k} & \cdots & \boldsymbol{\Sigma}_{1K} \\ \vdots & \ddots & & & \\ \boldsymbol{\Sigma}_{k1} & & \boldsymbol{\Sigma}_{kk} & & \\ \vdots & & & \ddots & \\ \boldsymbol{\Sigma}_{K1} & & & & \boldsymbol{\Sigma}_{KK} \end{pmatrix} . \tag{6.26}$$

Single covariance matrices, $\boldsymbol{\Sigma}_{11}, \cdots, \boldsymbol{\Sigma}_{kk}, \cdots \boldsymbol{\Sigma}_{KK}$, build the diagonal of $\boldsymbol{\Sigma}_{\text{multi}}$ and are computed using equation 6.2. On the other hand, mixed covariance matrices of two objects and their transposed pairs can be found in the upper and lower triangle of the block matrix, respectively. These are given by the following formula:

$$\boldsymbol{\Sigma}_{k_1 k_2} = \frac{1}{N_{k_1 k_2} - 1} \sum_j (\boldsymbol{c}_{jk_1} - \bar{\boldsymbol{c}}_{k_1}) \cdot (\boldsymbol{c}_{jk_2} - \bar{\boldsymbol{c}}_{k_2})^T \text{ , with } k_1 \neq k_2 \text{ ,} \tag{6.27}$$

where $N_{k_1 k_2}$ is the number of cases where both objects, k_1 and k_2, exist. The advantage of dividing up $\boldsymbol{\Sigma}_{\text{multi}}$ is that for the computation of each partial $\boldsymbol{\Sigma}_{k_1 k_2}$ as many examples can be used as there are in the data set, without respect to other objects' existence.

To derive a statistical multi-shape model of the four organs *hippocampus* (H), *thalamus* (T), *putamen* (P), and *globus pallidus* (G) in the left hemisphere of the brain we had to deal with the incomplete data set of 30 patients shown in table 6.1.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
H	x	x	x	x	x	x	x		x	x	x	x	x	x	x
T	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
P	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
G	x	x	x	x	x	x	x	x	x	x	x	x	x		x

	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
H	x	x			x	x					x	x	x		x
T	x	x	x	x	x	x	x	x	x	x	x	x		x	x
P	x	x	x	x	x	x	x	x	x	x		x	x	x	
G	x	x	x	x	x	x	x	x	x	x		x	x	x	

Table 6.1: Data set of 30 manually segmented MR volumes. Existing segmentations of the organs *hippocampus* (H), *thalamus* (T), *putamen* (P), and *globus pallidus* (G) are marked by 'x'.

There are only $N_{\text{compl}} = 18$ complete segmentations, while the number of single- and dual-objects is given in table 6.2. This example illustrates the advantage of separately computing the elements of the multi-shape covariance matrix. Since the number of single- and dual-shape models is always higher than N_{compl}, the assembled covariance matrix

$$\mathbf{\Sigma}_{\text{left\_brain}} = \begin{pmatrix} \mathbf{\Sigma}_{HH} & \mathbf{\Sigma}_{HT} & \mathbf{\Sigma}_{HP} & \mathbf{\Sigma}_{HG} \\ \mathbf{\Sigma}_{TH} & \mathbf{\Sigma}_{TT} & \mathbf{\Sigma}_{TP} & \mathbf{\Sigma}_{TG} \\ \mathbf{\Sigma}_{PH} & \mathbf{\Sigma}_{PT} & \mathbf{\Sigma}_{PP} & \mathbf{\Sigma}_{PG} \\ \mathbf{\Sigma}_{GH} & \mathbf{\Sigma}_{GT} & \mathbf{\Sigma}_{GP} & \mathbf{\Sigma}_{GG} \end{pmatrix} \tag{6.28}$$

is likely to represent a more representative statistic than the one obtained from the 18 complete cases only.

$$\begin{array}{llll} N_H = 22 & N_{HT} = 21 & N_{TP} = 27 & N_{PG} = 27 \\ N_T = 29 & N_{HP} = 20 & N_{TG} = 26 & \\ N_P = 28 & N_{HG} = 19 & & \\ N_G = 27 & & & \end{array}$$

Table 6.2: Number of single- and dual-shape models in the data set shown above.

Figure 6.18 illustrates the resulting eigendeformation corresponding to the largest eigenvalue of $\mathbf{\Sigma}_{\text{left\_brain}}$ from three orthogonal views. From top to bottom the weight of the deformation varies from $-2\sqrt{\lambda_1}$ to $2\sqrt{\lambda_1}$, where λ_1 is the largest eigenmode. The scene is shown from left in the first, from back in the second and from above in the third column of the figure. For comparison, Fig. 6.19 shows first deformation modes separately computed for each of the four organs. It can be seen, how mixed mode covariance matrices couple the deformation of neighboring shapes, while separate models allow for uncorrelated deformations which can form anatomically meaningless constellations where adjacent organs get separated or come into intersection.

It is interesting to compare major eigenvalues of the coupled and the decoupled shapes.

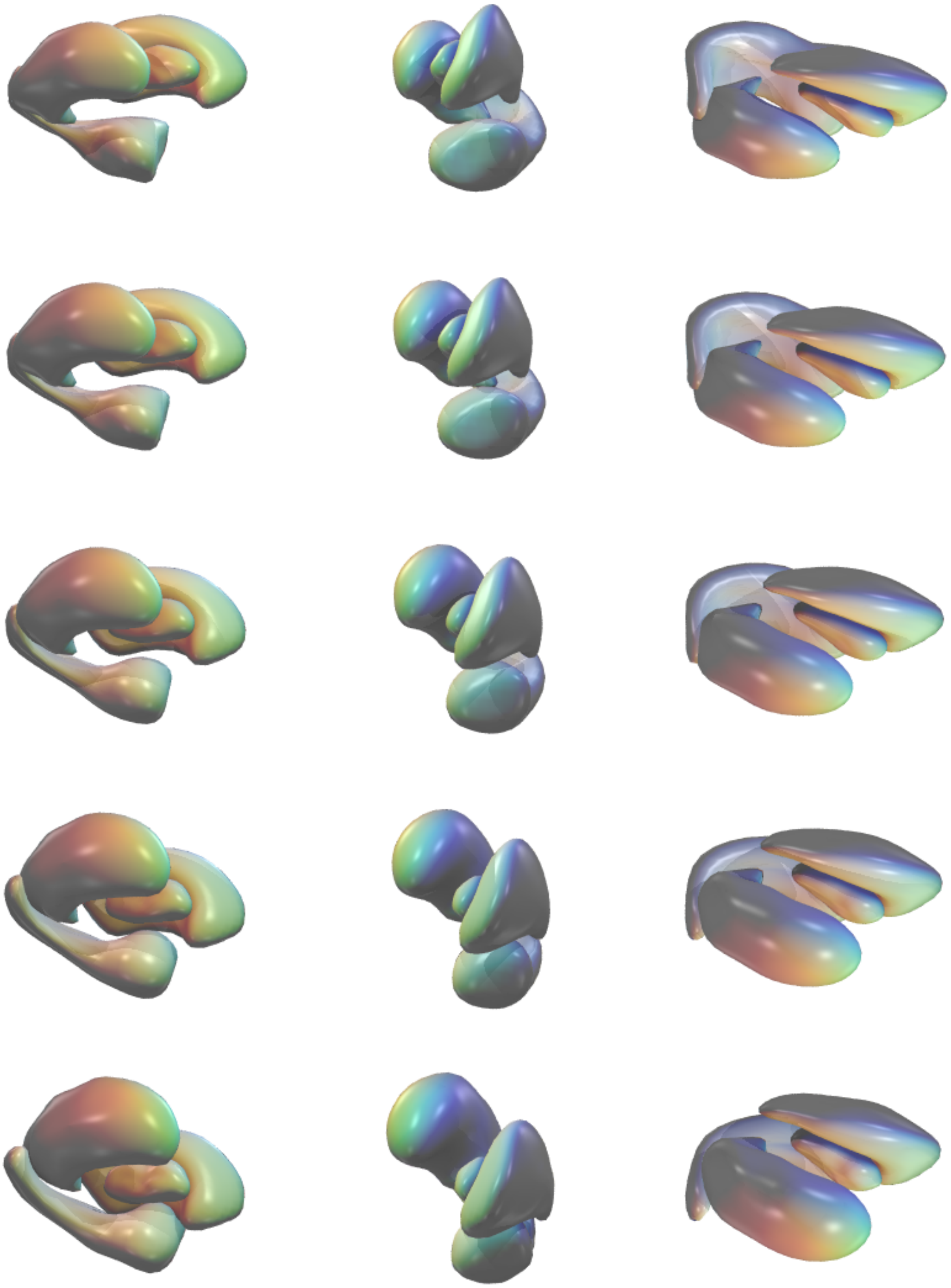


Figure 6.18: Resulting eigendeformation corresponding to the largest eigenvalue of $\boldsymbol{\Sigma}_{\mathrm{left\_brain}}$ from three orthogonal views. From top to bottom the weight of the deformation varies from $-2\sqrt{\lambda_1}$ to $2\sqrt{\lambda_1}$, where λ_1 is the largest eigenmode. The scene is shown from left in the first, from back in the second and from above in the third column of the figure.

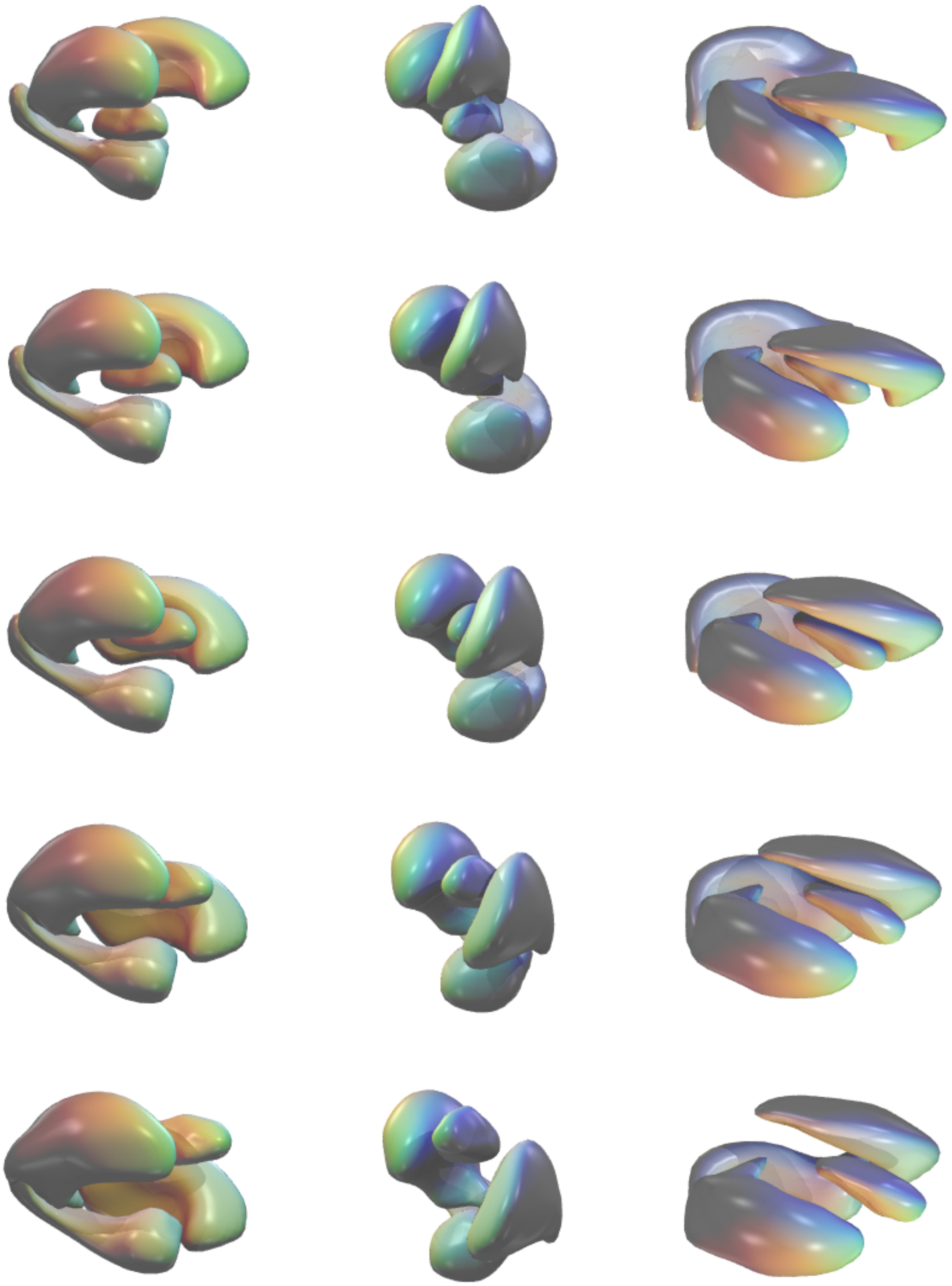
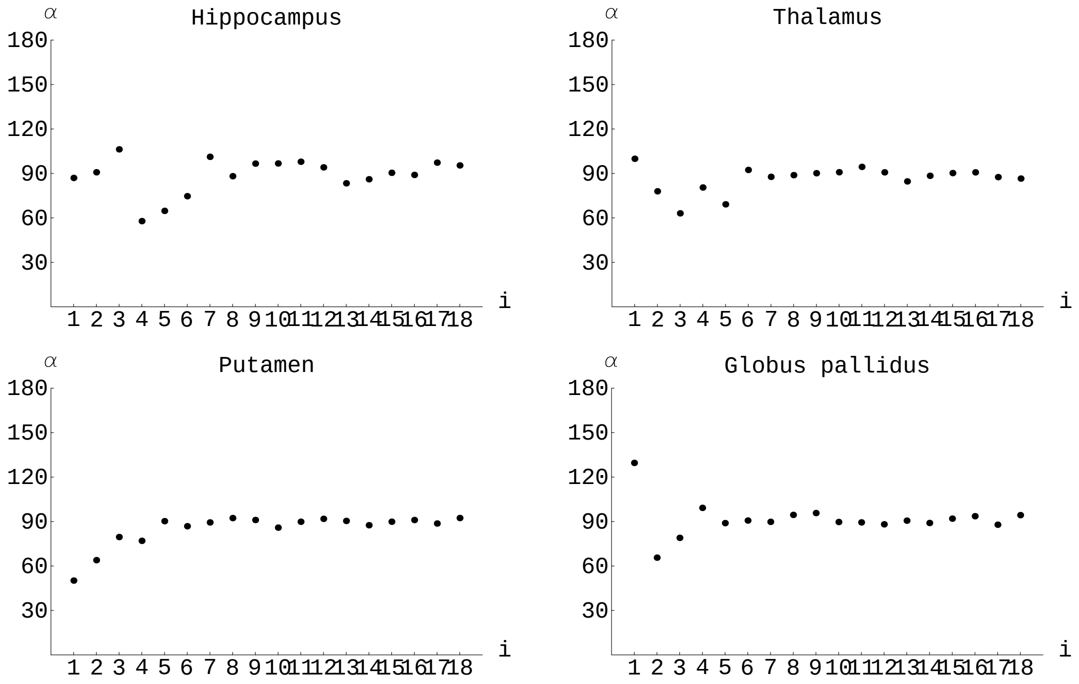


Figure 6.19: Eigendeformations corresponding to the largest eigenvalues of the separate covariance matrices $\mathbf{\Sigma}_{HH}$, $\mathbf{\Sigma}_{TT}$, $\mathbf{\Sigma}_{PP}$, and $\mathbf{\Sigma}_{GG}$, shown from three orthogonal views. From top to bottom the weight of the deformation varies from $-2\sqrt{\lambda_1}$ to $2\sqrt{\lambda_1}$, where λ_1 is the largest eigenmode. The scene is shown from the same directions as in Fig 6.18.

The angles

$$\alpha = \arccos(\boldsymbol{p}_{ci}^T \boldsymbol{p}_{di}) \tag{6.29}$$

of corresponding eigenvectors $\boldsymbol{p}_{ci}$ and $\boldsymbol{p}_{di}$ are illustrated in Fig. 6.20, where the indexes c and d denote coupled and decoupled cases, respectively. It can be shown (see Appendix A) that the expected angle of two independent random vectors is $\pi/2$ with approximate standard deviation $1/\sqrt{n}$, where n is the length of the vectors. In other words, they are orthogonal with high probbility. An angle different from $\pi/2$ indicates that the two vectors are correlated. In Fig. 6.20 we can observe that there is a larger correlation between the first eigenmodes in each object and smaller eigenmodes describe independent deformations.


Hippocampus

Thalamus

Putamen

Globus pallidus

Figure 6.20: Correlation of corresponding eigenvectors between coupled and separated multi-shape models. The correlation is given by the angle α of two vectors, where values different from $\alpha = 90^\circ$ indicate correlation.

6.5 Robustness of the multivariate statistics

The fact that our training set only consists of as few as n =22 to 30 shapes arises the question of how robust is the estimation of average shape and covariance matrix. The robustness of the estimation can be characterized by the expected variance of the estimated values, in our case the entries σ_{ij} in the covariance matrix $\boldsymbol{\Sigma}$. σ_{ij} can be thought of as the expectation of the random variable S_{ij}. We wish to compute the variance of S_{ij} in order to approximate the confidence we can have in the estimation of σ_{ij}. This is given as

$$\mathrm{Var}[S_{ij}] = \frac{\sigma_{ii}\sigma_{jj} + \sigma_{ij}}{n-1} \ . \tag{6.30}$$

A full deduction of this expression can be found in Appendix B. It says that the variance of the estimated covariance σ_{ij} from n observations is proportional to the values σ_{ii}, σ_{jj}, σ_{ij}, and inversely proportional to $(n-1)$.

6.5.1 Experiment

To visualize the result given in equation 6.30 we do an experiment with simulated data. We generate $n = 22$ samples of a 4-dimensional random vector with normal distribution of known mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$,

$$\boldsymbol{\mu} = \begin{pmatrix} 0 \\ 1 \\ 20 \\ 0 \end{pmatrix} , \ \boldsymbol{\Sigma} = \begin{pmatrix} 5 & 2 & -6 & 0 \\ 2 & 14 & -2 & 2 \\ -6 & -2 & 8 & 0 \\ 0 & 2 & 0 & 1 \end{pmatrix} \tag{6.31}$$

Applying equation 6.30 the expected standard deviation of each entries in $\boldsymbol{\Sigma}$ can be determined, assuming they are estimated from 22 samples (note: in a normal case, these values are not available, since $\boldsymbol{\Sigma}$ is unknown),

$$\begin{aligned} \sqrt{\mathrm{Var}[\boldsymbol{S}]} \ &= \begin{pmatrix} 5\sqrt{\frac{2}{21}} & \sqrt{\frac{74}{21}} & 2\sqrt{\frac{19}{21}} & \sqrt{\frac{5}{21}} \\ \sqrt{\frac{74}{21}} & 2\sqrt{\frac{14}{3}} & 2\sqrt{\frac{29}{21}} & \sqrt{\frac{6}{7}} \\ 2\sqrt{\frac{19}{21}} & 2\sqrt{\frac{29}{21}} & 8\sqrt{\frac{2}{21}} & 2\sqrt{\frac{2}{21}} \\ \sqrt{\frac{5}{21}} & \sqrt{\frac{6}{7}} & 2\sqrt{\frac{2}{21}} & \sqrt{\frac{2}{21}} \end{pmatrix} \\ &= \begin{pmatrix} 1.54303 & 1.87718 & 1.90238 & 0.48795 \\ 1.87718 & 4.32049 & 2.35028 & 0.92582 \\ 1.90238 & 2.35028 & 2.46885 & 0.61721 \\ 0.48795 & 0.92582 & 0.61721 & 0.30860 \end{pmatrix} \ . \end{aligned} \tag{6.32}$$

Figure 6.21 illustrates the covariance matrix $\boldsymbol{\Sigma}$ as a 3-D mesh together with the standard deviation of $\sqrt{\mathrm{Var}[\boldsymbol{S}]}$ of each entry in $\boldsymbol{\Sigma}$ as error bars.

Similarly to the procedure applied in computing shape variances, mean values and the covariance matrix of the simulated distribution has been estimated from the 22 cases using equations B.4 and B.5,

$$\boldsymbol{\Sigma}_{\mathrm{est}} = \begin{pmatrix} 4.85289 & 1.77946 & -5.90276 & -0.00761884 \\ 1.77946 & 14.273 & -1.72847 & 2.01807 \\ -5.90276 & -1.72847 & 7.93759 & 0.034534 \\ -0.00761884 & 2.01807 & 0.034534 & 0.925122 \end{pmatrix} \ . \tag{6.33}$$

In a real case, in which the covariance matrix is not known in advance, expected standard deviations $\sqrt{\mathrm{Var}[\boldsymbol{S}]}$ can only be approximated based on the estimated covariance matrix,

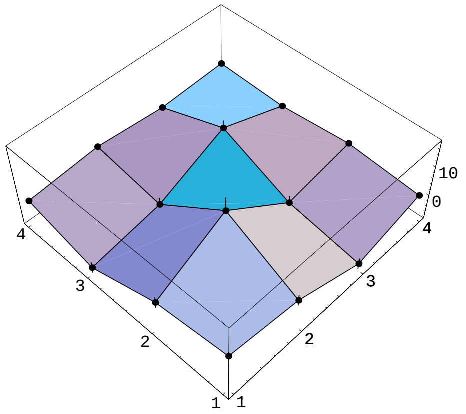


Figure 6.21: Covariance matrix $\mathbf{\Sigma}$ and its standard deviation if estimated from $n = 22$ samples. The covariance matrix is illustrated as a 3-D mesh while standard deviations are shown as error bars.

$$\sqrt{\mathrm{Var}[\boldsymbol{S}]}_{\mathrm{est}} = \begin{pmatrix} 1.49764 & 1.85718 & 1.86908 & 0.462374 \\ 1.85718 & 4.40473 & 2.35312 & 0.907031 \\ 1.86908 & 2.35312 & 2.44959 & 0.591384 \\ 0.462374 & 0.907031 & 0.591384 & 0.285499 \end{pmatrix} . \tag{6.34}$$

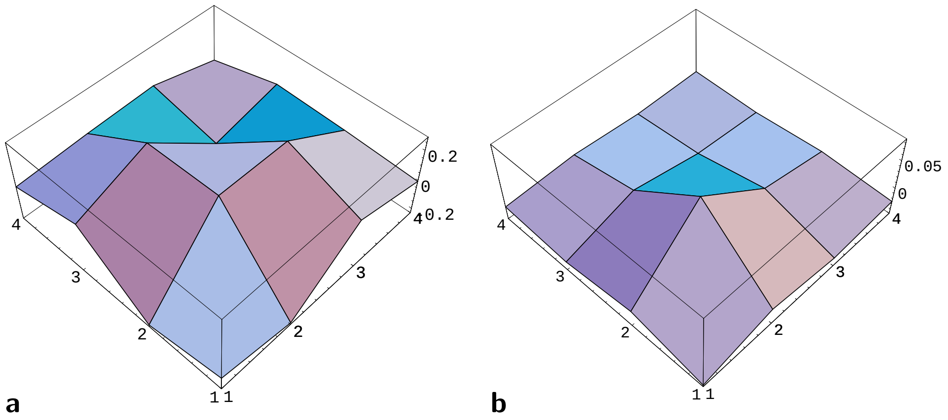


Figure 6.22: Absolute differences of computed and estimated covariance matrices (**a**) and their standard deviations (**b**).

In our experiment, empirical and estimated values can be compared. Absolute differences of computed and estimated covariance matrices and their standard deviations are shown in Figure 6.22 **a** and **b**, respectively. These images show that differences for the covariance matrix remain within the range of ± 0.2, while errors of standard deviations are within the range of ± 0.05.

The experiment indicates that estimations for $\mathbf{\Sigma}$ and $\sqrt{\mathrm{Var}[\boldsymbol{S}]}$ based on as few as 22 samples are reliable and come close to the the computed values. Since estimations does not depend on the dimensionality of the problem, these remain reliable and robust in multi-variate cases, such as 2- or 3-dimensional shape descriptors.

7

Conclusions

Automated, robust segmentation of medical images most often needs *a priori* anatomical knowledge. Typical cases are the segmentation of healthy organs, which present restricted anatomical variability, or the segmentation of organs if only incomplete evidence for boundaries is given by the grey-valued images, requiring an 'intelligent guess' about the position of the object boundary.

Geometric organ models and the statistics of their *normal (expected) variation* seem to offer a promising solution to this problem. We proposed the use of the Fourier parametrization of our models, followed by a statistical analysis of a training set, providing mean organ models and their eigendeformations. Elastic fit of the mean model in the subspace of eigenmodes restricts possible deformations and finds an optimal match between the model surface and boundary candidates.

In computing a statistical model for a set of similar shapes, correspondence between individuals is essential. We tried to solve the correspondence problem with respect to a possible extension to 3-D. Therefore we by design avoided to incorporate procedures which can be carried out only in 2-D such as manual annotation of contour points. We have shown that Fourier parametrization in 2-D provides a good approximation to the correspondence in many cases, and that its 3-D extension achieves similarly sufficient results. We also demonstrated the limitations of establishing correspondence by arc-length parametrization and presented a reparametrization technique based on Ψ-functions to overcome these.

The complete procedure has been demonstrated with the determination of the outline of the corpus callosum on a 2-D MRI slice of the human brain. We have shown that the selection of an anatomically defined reference coordinate system allows the inclusion of the variation of the spatial position and orientation into the description of the model variation. This procedure provides a homogeneous framework for the complete analysis, while adding only minimal, well defined user interaction to the procedure.

Segmentation quality and speed has been further improved by incorporating information about the objects image environment in form of one-dimensional intensity profiles. The procedure has been validated by comparing the results achieved by automatic segmentation to manual ones. Comparisons made for a statistically significant number of data sets have shown that our method performed well in 90% of the cases.

The initial placement of the average model is performed by a coarse elastic deformation defined by only the largest eigenmodes, helping to drive the following optimization into the "correct" local optimum. We noticed that the convergence is faster if only a small number of modes are involved, while a larger number of modes is required to find the exact contour. Thus, we 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.

Based on the 2-D procedure we presented 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 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 model has been derived from a series of training data sets. Thereby, the model represents a realistic shape rather than an artificial geometric model. 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.

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 optimization scheme. Early implementations were 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.

Our approach is based on the research work of [Cootes *et al.* 1994a]. 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 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.

Dual shape representations: Our approach makes use of two shape representations which are used in an alternating 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.

A new quality measure for intensity profiles has been proposed which expresses the combination of their selectivity and reliability. This can be included into the segmentation process by giving larger weight and hereby more confidence to suggested movements at more reliable boundary information.

Validation has been done by quantitatively comparing the automatic segmentation with the results of interactive outlining by experts, which is a common “gold standard” for comparisons. Two different measures has been computed, one based on the average distance between surfaces, and another one based on the intersection and union of two volumes. We have demonstrated that our results come close to the expert’s segmentations, although there were much less samples available for comparison than for the validation of the 2-D segmentations.

Application of multi-shape statistical models has also been investigated. By dividing the covariance matrix of multiple shapes into partial covariances only dependent on one or two objects, we are able to handle incomplete data sets without loss of information. Importance of multi-shape models has been demonstrated by an example using four nearby brain regions. If deformation modes are computed individually, the four shapes can deform into anatomically meaningless constellations where adjacent organs get separated or come into intersection. As presented, this can be avoided using coupled deformations derived from the multi-shape covariance matrix.

We also addressed the problem of small training sets by deriving a measure of reliability for the estimation of the covariance matrix. It has been shown that the variance of the estimated values is easy to compute and represents well the confidence we can have in the estimation. An example based on known distribution has been presented to show that even as small training sets as we had to deal with can provide a reliable estimation.

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.

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 model of several anatomical structures. The representation of anatomical objects by normalized shape descriptors makes access to explicite morphometric parameters. After segmenting a new set of image data, morphological properties of objects are available for comparative studies.

7.1 Outlook

There are still many issues especially in the 3-D part which could not be investigated further within the confines of the presented work. Our relatively small training sets did not allow an extensive testing of the proposed methods. Thus, the advantage of including the quality measure of profiles into the segmentation process has to be still validated on more cases. Similarly, to test the quality of the segmentation itself requires a larger set of expert segmentations.

The new segmentation framework gave impulse to new applications. New cases of *hippocampi* are currently under manual segmentation within the frames of the European Union’s BIOMORPH project. The data set available for the participants consists of 70 high resolution brain MRI-s, out of which 20 cases has already been processed. Expectedly, the statistical model derived from the 70 cases will considerably improve segmentation quality.

Another work, also carried out within the BIOMORPH Consortium, aims at building a brain mask and to fit it onto individual MR volumes using the technique described here. Automatic definition of a region of interest of the brain would help low level segmentation techniques such as statistical classification to concentrate only to that part of the volumetric data which holds the information.

The European Union's CONQUEST-project aims at providing quality assurance tools for cancer treatment resulting in better standardization eliminating the large variations in recurrence and survival rates between and within countries. The new 3-D segmentation technique is being applied within the project to speed up delineation of healthy tissues, in this specific case of the male abdomen, so that it can be done on a routine basis. The results of the segmentation can be used in treatment planning optimization to record dose distributions in healthy organs on a routine basis which is necessary get reliable statistical data on dose-response relationship of these subjects.

A

Angle between two random vectors

Given two normalized random vectors

$$\begin{aligned} \boldsymbol{X} &= \begin{pmatrix} X_0 \\ \vdots \\ X_{n-1} \end{pmatrix} \\ \boldsymbol{Y} &= \begin{pmatrix} Y_0 \\ \vdots \\ Y_{n-1} \end{pmatrix} , \end{aligned} \tag{A.1}$$

whith

$$\boldsymbol{X}, \boldsymbol{Y} \mapsto \Omega_n \ , \tag{A.2}$$

where all components have 0 mean and equal variance σ^2. For large n components can be considered as approximately independent. [1]. Per definition

$$\mathrm{E}[\boldsymbol{X}^T \boldsymbol{X}] = \mathrm{E}[\boldsymbol{Y}^T \boldsymbol{Y}] = 1 \ , \tag{A.3}$$

and from the approximate independence of components follows

$$\mathrm{E}[\boldsymbol{X}^T \boldsymbol{X}] = \mathrm{E}[\boldsymbol{Y}^T \boldsymbol{Y}] = n\sigma^2 \ . \tag{A.4}$$

Thus, variance of components can be approximated by

$$\sigma^2 \approx \frac{1}{n} \ . \tag{A.5}$$

The cosine of the angle between the two random vectors is then also a random variable. This is defined as

$$C = \cos\alpha = \boldsymbol{X}^T \boldsymbol{Y} \ . \tag{A.6}$$

---

[1] Since $\boldsymbol{X}$ and $\boldsymbol{Y}$ are normalized, their components cannot be completely independent. If $n-1$ components are given, the nth component can be determined by $X_n = \sqrt{1 - (X_0^2 + \ldots + X_{n-1}^2)}$

Expectation and variance of C can be computed by

$$\mathrm{E}[C] = \mathrm{E}[\boldsymbol{X}^T \boldsymbol{Y}] = n\mathrm{E}[X_i]\mathrm{E}[Y_i] = 0 \ , \tag{A.7}$$

and

$$\mathrm{Var}[C] = \mathrm{Var}[\boldsymbol{X}^T \boldsymbol{Y}] \approx n\mathrm{Var}[X_i]\mathrm{Var}[Y_i] = n\sigma^4 \approx n\frac{1}{n^2} = \frac{1}{n} \ . \tag{A.8}$$

Thus, two independent random vectors are expectedly orthogonal, since

$$\arccos \mathrm{E}[C] = \arccos 0 = \frac{\pi}{2} \ . \tag{A.9}$$

B

Variance of the estimated covariance matrix

In section 6.5 we address the problem of finding a measure for the robustness relating to the estimation of the covariance matrix computed from n observations. The robustness of the estimation can be characterized by the expected variance of the estimated values, in our case the elements σ_{ij} in the covariance matrix $\boldsymbol{\Sigma}$. σ_{ij} can be thought of as the expectation of the random variable S_{ij}. We wish to compute the variance of S_{ij} in order to approximate the confidence we can have in the estimation of σ_{ij}. For the sake of simplicity the computation will be done for a bi-variate covariance matrix and then generalized for the multi-variate case.

We define a 2-dimensional random vector of bivariate standard normal distribution

$$\boldsymbol{X} = \begin{pmatrix} X_0 \\ X_1 \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) . \tag{B.1}$$

Based on $\boldsymbol{X}$ we create the arbitrary 2-dimensional random vector $\boldsymbol{Y}$

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mu} = \begin{pmatrix} a_{00}X_0 + a_{01}X_1 \\ a_{10}X_0 + a_{11}X_1 \end{pmatrix} + \begin{pmatrix} \mu_0 \\ \mu_1 \end{pmatrix} , \tag{B.2}$$

with expectation $\boldsymbol{\mu}$ and covariance matrix

$$\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T = \begin{pmatrix} \sigma_{00} & \sigma_{01} \\ \sigma_{10} & \sigma_{11} \end{pmatrix} = \begin{pmatrix} a_{00}^2 + a_{01}^2 & a_{00}a_{10} + a_{01}a_{11} \\ a_{00}a_{10} + a_{01}a_{11} & a_{10}^2 + a_{11}^2 \end{pmatrix} . \tag{B.3}$$

Supposing, there are n observations available for $\boldsymbol{Y}$ means in $\boldsymbol{\mu}$ can be estimated by

$$\begin{aligned} \bar{Y_0} &= \frac{1}{n}\sum_{i=0}^{n-1} Y_{0i} \\ \bar{Y_1} &= \frac{1}{n}\sum_{i=0}^{n-1} Y_{1i} , \end{aligned} \tag{B.4}$$

and the covariance σ_{01} can be estimated using

$$S_{01} = \frac{1}{n-1}\sum_{i=0}^{n-1}(Y_0-\bar{Y_0})(Y_1-\bar{Y_1}) = \frac{1}{n-1}\left(\sum_{i=0}^{n-1}Y_0Y_1 - \sum_{i=0}^{n-1}Y_0\bar{Y_1} - \sum_{i=0}^{n-1}Y_0\bar{Y_1} + \sum_{i=0}^{n-1}\bar{Y_0}\bar{Y_1}\right) . \tag{B.5}$$

Multiplying both sides by $(n-1)$ and substituting first equations B.4 and then equation B.2 we obtain the following expressions ($\sum_{i=0}^{n-1}$ is replaced by $\sum$ for brevity)

$$\begin{aligned}
(n-1)S_{01} &= \sum Y_0Y_1 - \sum Y_0\bar{Y_1} - \sum Y_0\bar{Y_1} + \sum \bar{Y_0}\bar{Y_1} \\
&= \sum Y_0Y_1 - \frac{1}{n}\sum Y_0 \sum Y_1 - \frac{1}{n}\sum Y_0\bar{Y_1} + \frac{1}{n}\sum Y_0 \frac{1}{n}\sum Y_1 n \\
&= \sum Y_0Y_1 - \frac{1}{n}\sum Y_0 \sum Y_1 \\
&= \sum(a_{00}X_0 + a_{01}X_1 + \mu_0)(a_{10}X_0 + a_{11}X_1 + \mu_1) - \\
&\quad \frac{1}{n}\sum(a_{00}X_0 + a_{01}X_1 + \mu_0)\sum(a_{10}X_0 + a_{11}X_1 + \mu_1) ,
\end{aligned} \tag{B.6}$$

and after simplification we end up with

$$\begin{aligned}
(n-1)S_{01} &= a_{00}a_{10}\left(\sum X_0^2 - \frac{1}{n}\left(\sum X_0\right)^2\right) + \\
&\quad (a_{00}a_{11} + a_{01}a_{10})\left(\sum X_0X_1 - \frac{1}{n}\sum X_0 \cdot \sum X_1\right) + \\
&\quad a_{01}a_{11}\left(\sum X_1^2 - \frac{1}{n}\left(\sum X_1\right)^2\right) .
\end{aligned} \tag{B.7}$$

For the expression S_{01}, being itself a random variable, we can compute expectation $\mathrm{E}[S_{01}]$ and variance $\mathrm{Var}[S_{01}]$. Starting with $\mathrm{E}[S_{01}]$ we can write

$$\begin{aligned}
\mathrm{E}[(n-1)S_{01}] &= a_{00}a_{10}\left(n - \frac{1}{n}\cdot n\right) + \\
&\quad (a_{00}a_{11} + a_{01}a_{10})\left(0 - \frac{1}{n}\cdot 0 \cdot 0\right) + a_{01}a_{11}\left(n - \frac{1}{n}\cdot n\right) \\
&= (n-1)\cdot(a_{00}a_{10} + a_{01}a_{11}) ,
\end{aligned} \tag{B.8}$$

and finally dividing eq. B.8 by $(n-1)$ we obtain

$$\mathrm{E}[S_{01}] = a_{00}a_{10} + a_{01}a_{11} = \sigma_{01} \tag{B.9}$$

as expected. To compute $\mathrm{Var}[S_{01}]$ we first reformulate the expression $\sum X_0^2 - \frac{1}{n}\left(\sum X_0\right)^2$ from equation B.7 as follows:

$$\mathrm{exp}_1 = \sum_{i=0}^{n-1}X_{0i}^2 - \frac{1}{n}\sum_{i=0}^{n-1}X_{0i}\sum_{j=1}^{n}X_{0j}$$

$$\begin{aligned} &= \sum_{i=0}^{n-1} X_{0i}^2 - \frac{1}{n}\left(\sum_{i=0}^{n-1} X_{0i}^2 + \sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} 2X_{0i}X_{0j}\right) \\ &= \frac{n-1}{n}\sum_{i=0}^{n-1} X_{0i}^2 - \frac{1}{n}\sum_{i=0}^{n-2}\sum_{j=i+1}^{n-1} 2X_{0i}X_{0j}\ . \qquad \text{(B.10)} \end{aligned}$$

Similarly, the expression $\sum X_0X_1 - \frac{1}{n}\sum X_0 \cdot \sum X_1$ can be written as

$$\exp_2 = \frac{n-1}{n}\sum_{i=0}^{n-1} X_{0i}X_{1i} - \frac{1}{n}\sum_{i=0}^{n-1}\sum_{i=0,i\neq j}^{n-1} X_{0i}X_{1j} \qquad \text{(B.11)}$$

Applying the rule for the variance of the product of two random values A and B

$$\mathrm{Var}[AB] = \mathrm{Var}[A]\ \mathrm{Var}[B] + \mathrm{Var}[A]\ \mathrm{E}[B]^2 + \mathrm{Var}[B]\ \mathrm{E}[A]^2\ , \qquad \text{(B.12)}$$

and the rule to compute the variance of the sum $\alpha A + \beta B$

$$\mathrm{Var}[\alpha A + \beta B] = \alpha^2 \mathrm{Var}[A] + 2\alpha\beta \mathrm{Cov}[A, B] + \beta^2 \mathrm{Var}[B]\ , \qquad \text{(B.13)}$$

the variances of $\exp_1$ and $\exp_2$ result as

$$\begin{aligned} \mathrm{Var}[\exp_1] &= \frac{n-1}{n}2n + \frac{1}{n}2(n-1)n \\ &= 2(n-1) \qquad \text{(B.14)} \\ \mathrm{Var}[\exp_2] &= (n-1) \qquad \text{(B.15)} \end{aligned}$$

Putting it all together, the variance of expression B.7 can be written as

$$\begin{aligned} \mathrm{Var}[(n-1)S_{01}] &= (n-1)\left(2a_{00}^2a_{10}^2 + a_{01}^2a_{11}^2 + (a_{00}a_{11} + a_{01}a_{10})^2\right) \\ &= (n-1)\left(\sigma_{00}\sigma_{11} + \sigma_{01}\right)\ , \qquad \text{(B.16)} \end{aligned}$$

and multiplying eq. B.16 by $(n-1)$ we obtain the expression for the variance of the estimation of covariance σ_{01} based on n observations

$$\mathrm{Var}[S_{01}] = \frac{\sigma_{00}\sigma_{11} + \sigma_{01}}{n-1}\ . \qquad \text{(B.17)}$$

This can be generalized from the bi-variate to the multivariate case:

$$\mathrm{Var}[S_{ij}] = \frac{\sigma_{ii}\sigma_{jj} + \sigma_{ij}}{n-1}\ , \qquad \text{(B.18)}$$

saying that the variance of the estimated covariance σ_{ij} from n observations is proportinal to the values σ_{ii}, σ_{jj}, σ_{ij}, and inversely proportional to $(n-1)$.

Bibliography

[Andreasen *et al.* 1994] N.C. Andreasen, S. Arndt, V. Swayze Il, T. Cizadlo, M. Flaum, D. O'Leary, J.C. Ehrhardt, and V.T.C. Yuh. Thalamic Abnormalities in Schizophrenia Visualized Through Magnetic Resonance Image Averaging. *Science*, 266:294–298, October 1994.

[Ashton *et al.* 1997] E.A. Ashton, K.J. Parker, M.J. Berg, and Ch.W. Chen. A Novel Volumetric Feature Extraction Technique with Applications to MR Images. *IEEE Transactions on Medical Imaging*, 16(4):365–371, August 1997.

[Ayache 1995] N. Ayache. Medical computer vision, virtual reality and robotics. *Image and Vision Computing*, 13(4):295–313, May 1995.

[Bajcsy and Kovacic 1989] R. Bajcsy and S. Kovacic. Multiresolution elastic matching. *Computer Vision, Graphics and Image Processing*, 46:1, 1989.

[Bake and Isard 1998] A. Bake and M. Isard. *Active Contours*. Springer, 1998.

[Bookstein 1997] F.L. Bookstein. Shape and the Information in Medical Images: A Decade of the Morphometric Synthesis. *Computer Vision and Image Understanding*, 66(2):97–118, May 1997.

[Brechbühler *et al.* 1992] C. Brechbühler, G. Gerig, and O. Kübler. Surface parametrization and shape description. In *Visualization in Biomedical Computing*, pages 80–89, 1992.

[Brechbühler *et al.* 1995] C. Brechbühler, G. Gerig, and O. Kübler. Parametrization of closed surfaces for 3-D shape description. *CVGIP: Image Understanding*, 61:154–170, 1995.

[Brechbühler-Miškuv 1995] Ch. Brechbühler-Miškuv. *Description and Analysis of 3-D Shapes by Parametrization of Closed Surfaces*. PhD thesis, ETH Zurich, Switzerland, 1995.

[Canny 1986] J. Canny. A computational approach to edge detection. 8(6):679–698, November 1986.

[Caunce and Taylor 1998] A. Caunce and C.J. Taylor. 3D Point Distribution Models of the Cortical Sulci. In *Sixth International Proceedings on Computer Vision*, pages 402–407, Bombay, India, January 1998. IEEE, Narosa Publishing House.

[Chakraborty *et al.* 1996] A. Chakraborty, L.H. Staib, and J.S. Duncan. Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information. *IEEE Transactions on Medical Imaging*, 15(6):859–870, December 1996.

[Christensen *et al.* 1996] G.E. Christensen, M.I. Miller, and M.W. Vannier. Individualizing neuroanatomical alases using a massively parallel computer. *Computer*, pages 32–38, January 1996.

[Cohen *et al.* 1992] I. Cohen, L.D. Cohen, and N. Ayache. Using Deformable Surfaces to Segment 3D Images and Infer Differential Structures. *CVGIP: Image Understanding*, 56(2):242–263, 1992.

[Cootes and Taylor 1992] T.F. Cootes and C.J. Taylor. Active Shape Models – 'Smart Snakes'. In *British Mach. Vision Conf.*, pages 266–275. Springer-Verlag, 1992.

[Cootes and Taylor 1996] T.F. Cootes and C.J. Taylor. Data Driven Refinement of Active Shape Model Search. In R.B. Fisher and E. Trucco, editors, *British Machine Vision Conference*, pages 383–392. BMVA Press, 1996.

[Cootes *et al.* 1993] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam. The Use of Active Shape Models For Locating Structures in Medical Images. In *IPMI*, pages 33 – 47. Flagstaff USA, 1993.

[Cootes *et al.* 1994a] 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, 1994. Electronic version: http://s10d.smb.man.ac.uk/publications/index.htm.

[Cootes *et al.* 1994b] T.F. Cootes, C.J. Taylor, and A. Lantis. Active Shape Models: Evaluation of Multi-Resolution Method for Improving Image Search. In E. Hancock, editor, *Proc. British Machine Vision Conference*, volume 1, pages 327–226. BMVA Press, 1994.

[Cootes *et al.* 1995a] 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, 1995.

[Cootes *et al.* 1995b] T.F. Cootes, G.J. Page, C.B. Jackson, and C.J. Taylor. Statistical Grey-Level Models for Object Location and Identification. In D. Pycock, editor, *Proc. British Machine Vision Conference*, pages 533–542. BMVA Press, 1995.

[Davatzikos and Prince 1994] C. Davatzikos and J. Prince. Brain image registration based on curve mapping. In *IEEE Workshop Biomedical Image Anal.*, pages 38–59, 1994.

[Delibasis *et al.* 1997] K. Delibasis, P.E. Undrill, and G.G. Cameron. Designing Fourier Descriptor-Based Geometric Models for Object Interpretation in Medical Images Using Genetic Algorithms. *Computer Vision and Image Understanding*, 66(3):286–300, June 1997.

[Delingette 1997] H. Delingette. General Object Reconstruction based on Simplex Meshes. Technical Report 3111, INRIA, 1997.

[Di Mauro *et al.* 1996] E.C. Di Mauro, T.F. Cootes, C.J. Taylor, and A. Lantis. Active Shape Model Search using Pairwise Geometric Histograms. In R.B. Fisher and E. Trucco, editors, *British Machine Vision Conference*, pages 353–362. BMVA Press, 1996.

[Evans *et al.* 1994] A.C. Evans, M. Kamber, D.L. Collins, and D. MacDonald. An MRI-based Probabilistic Atlas of Neuroanatomy. In S.D. Shorvon, editor, *Magnetic Resonance Scanning and Epilepsy*, pages 263–274. Plenum Press, New York, 1994.

[Finkelstein and Salesin 1994] A. Finkelstein and D.H. Salesin. Multiresolution Curves. In *SIGGRAPH Computer Graphics Proceedings*, pages 261–268, July 1994.

[Fritsch *et al.* 1997] D. Fritsch, S. Pizer, L. Yu, V. Johnson, and E. Chaney. Segmentation of Medical Image Objects using Deformable Shape Loci. In *Information Processing in Medical Imaging, 15th International Conference*, pages 127–140, Berlin, Germany, 1997. Springer-Verlag.

[Gerig *et al.* 1991] 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.

[Graps A. 1995] Graps A. An Introduction to Wavelets. *IEEE Computational Science and Engineering*, 2(2), Summer 1995.

[Greiner and Diehl 1986] W. Greiner and H. Diehl. *Theoretische Physik - Ein Lehr- und Übungsbuch für Anfangssemester*, volume 3: Elektrodynamik. Verlag Harri Deutsch, Zürich und Frankfurt am Main, 1986.

[Herman and Liu 1979] G.T. Herman and H.K. Liu. Three-dimensional display of human organs from computer tomograms. *Computer Graphics and Image Processing*, 9(1):1–21, January 1979.

[Höhne *et al.* 1992] K.H. Höhne, M. Bomans, M. Riemer, R. Schubert, U. Tiede, and W. Lierse. A Volume-based Anatomical Atlas. *IEEE Computer Graphics and Application*, 12(4):72–77, July 1992.

[Kass *et al.* 1988] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. Comp. Vision*, 1(4):321–331, 1988.

[Kotcheff and Taylor 1997] A.C.W Kotcheff and C.J Taylor. Automatic Construction of Eigenshape Models by Genetic Algorithm. In *Information Processing in Medical Imaging*, pages 1–14. Springer, 1997.

[Kruggel and Lohmann 1997] F. Kruggel and G. Lohmann. Automatical Adaption of the Steretactical Coordinate System in Brain MRI Datasets. In *Information Processing in Medical Imaging*, pages 471–476. Sringer, June 1997.

[Kuhl and Giardina 1982] F.P. Kuhl and C.R. Giardina. Elliptic Fourier Features of Closed Contour. *Computer Graphics and Image Processing*, 18:236–258, 1982.

[Marais *et al.* 1996] P. Marais, R. Guillemaud, M. Sakuma, A. Zisserman, and M. Brady. Visual Cerebral Asymmetry. In *Visualization in Biomedical Computing*, pages 411–416, 1996.

[Martin *et al.* 1994] J. Martin, A. Pentland, and R. Kikinis. Shape Analysis of Brain Structures Using Physical and Experimental Modes. In *IEEE Proceedings on Computer Vision and Pattern Recognition*, pages 752–755. IEEE, IEEE Computer Society Press, June 1994.

[McInerney and Terzopoulos 1996] T. McInerney and D. Terzopoulos. Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[Minoshima *et al.* 1993] 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, 1993.

[Montagnat and Delingette 1997] J. Montagnat and H. Delingette. A Hybrid Framework for Surface Registration and Deformable Models. In *Computer Vision and Pattern Recognition, CVPR'97*, pages 1041–1046, San Juan, Puerto Rico, 1997.

[Nastar and Ayache 1994] C. Nastar and N. Ayache. Classification of Nonrigid Motion in 3D Images using Physics-Based Vibration Analysis. In *IEEE Workshop on Biomed. Image Anal.*, pages 61–69, Seattle, Washington, USA, 1994.

[Nastar and Ayache 1996] Ch. 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, 1996.

[Neuenschwander *et al.* 1994] W. Neuenschwander, P. Fua, G. Székely, and O. Kübler. Initializing Snakes. In *CVPR'94*, pages 658–663, 1994.

[Neuenschwander *et al.* 1997] W. Neuenschwander, P. Fua, G. Székely, and O. Kübler. Velcro Surfaces: Fast Initialization of Deformable Models. *Computer Vision and Image Understanding*, 65(2):237–245, February 1997.

[Neuenschwander 1996] W. Neuenschwander. *Elastic Deformable Contour and Surface Models for 2-D and 3-D Image Segmentation.* PhD thesis, ETH Zurich, 1996.

[Pentland and Sclaroff 1991] A. Pentland and A.P. Sclaroff. Closed-Form Solutions for Physically Based Shape Modelling and Recognition. *IEEE PAMI*, 13(7):715–729, 1991.

[Press *et al.* 1988] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. Vetterling. *Numerical recipes in C - The art of scientific computing.* Cambridge University Press, Cambridge, 1988.

[Rangarajan *et al.* 1997] A. Rangarajan, H. Chui, and F.L. Bookstein. The Softassign Procrustes Matching Algorithm. In *Information Processing in Medical Imaging*, pages 29–42. Springer, 1997.

[Raya 1990] S.P. Raya. Low-Level Segmentation of 3-D Magnetic Resonance Brain Images – A Rule-Based System. *IEEE Trans. on Medical Imaging*, 9(3):327–337, September 1990.

[Rusell *et al.* 1997] J.M Rusell, T.S. Early, J.C. Patterson, J.L. Martin, J. Villanueva-Meyer, and M.D. McGee. Temporal Lobe Perfusion Asymmetries in Schizophrenia. *The Journal of Nuclear Medicine*, 38(4):607–612, April 1997.

[Sclaroff and Pentland 1994] S. Sclaroff and A.P. 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.

[Staib and Duncan 1989] L.H. Staib and J.S. Duncan. Parametrically deformable contour models. In *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pages 98–103, June 1989.

[Staib and Duncan 1992a] L.H. Staib and J.S. Duncan. Boundary Finding with Parametrically Deformable Models. *IEEE Transactions on Pattern Analysis and Machine Intelligece*, 14(11):1061–1075, November 1992.

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

[Staib and Duncan 1996] L.H. Staib and J.S. Duncan. Model-based Deformable Surface Finding for Medical Images. *IEEE Trans. Med. Imaging*, 15(5):1–12, 1996.

[Staib *et al.* 1997] L.H. Staib, A. Chakraborty, and J.S. Duncan. An Integrated Approach for Locating Neuroanatomical Structure from MRI. *Int. J. Patt. Recog. Art. Intell., Special Issue on MR Brain Image Analysis*, 1997.

[Stansfield 1986] S.A. Stansfield. ANGY: A Rule-Based System for Automatic Segmentation of Coronary Vessels From Digital Subtracted Angiograms. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(2):188–199, 1986. March.

[Stollnitz *et al.* 1995a] E.J. Stollnitz, T.D. DeRose, and D.H. Salesin. Wavelets for Computer Graphics: A Primer, Part 1. *IEEE Computer Graphics and Application*, pages 76–84, May 1995.

[Stollnitz *et al.* 1995b] E.J. Stollnitz, T.D. DeRose, and D.H. Salesin. Wavelets for Computer Graphics: A Primer, Part 2. *IEEE Computer Graphics and Application*, pages 75–85, July 1995.

[Stollnitz *et al.* 1996] E.J. Stollnitz, T.D. DeRose, and D.H. Salesin. *Wavelets for Computer Graphics, Theory and Applications.* Morgan Kaufmann Publishers, Inc., 1996.

[Székely *et al.* 1996] G. Székely, A. Kelemen, Ch. 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.

[Tagare 1997] H.D. Tagare. Non-rigid Curve Correspondence for Estimating Heart Motion . In *Information Processing in Medical Imaging*, pages 489–494. Springer, 1997.

[Talairach and Tournoux 1988] J. Talairach and P. Tournoux. *Co-planar stereotaxic atlas of the human brain.* Thieme, Stuttgart, 1988.

[Terzopoulos and Metaxas 1991] D. Terzopoulos and D. Metaxas. Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics. *IEEE PAMI*, 13(7):703–714, 1991.

[Terzopoulos *et al.* 1988] D. Terzopoulos, A. Witkin, and M. Kass. Symmetry-Seeking Models and 3D Object Reconstruction. *Int. J. Comp. Vision*, 1(3):211–221, 1988.

[Terzopoulos 1987] D. Terzopoulos. On matching deformable models to images. *Topical Meeting on Machine Vision Tech. Digest Series*, 12:160–167, 1987.

[Thirion and A. 1996] J.-Ph. Thirion and Gourdon A. Computing the Differential Characteristics of Isointensity Surfaces. *Computer Vision and Image Understanding*, 61(2):190–202, March 1996.

[Thirion *et al.* 1997] J. Ph. Thirion, S. Prima, and G. Subsol. Statistical Analysis of Dissymmetry in Volumetric Medical Images. Technical Report 3178, INRIA, 1997. Electronic version: http://www.inria.fr/RRRT/RR-3178.html.

[Thompson and Toga 19967] P.M. Thompson and A.W. Toga. Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brai atlas based on random field transformations. *Medical Image Analysis*, 1(4):271–294, 1996/7.

[Vaillant and Davatzikos 19967] M. Vaillant and Ch. Davatzikos. Finding parametric represetations of the cortical sulci using an active contour model. *Medical Image Analysis*, 1(4):295–315, 1996/7.

[Vemuri and Radisavljevic 1994] B.C. Vemuri and A. Radisavljevic. Multiresolution Stochastic Hybrid Shape Models with Fractal Priors. *ACM Trans. Graphics*, 13(2):177–200, 1994.

[Vemuri *et al.* 1996] B.C. Vemuri, Y. Guo, Ch.M. Leonard, and Sh.-H. Lai. Fast Algorithms for Fitting Multiresolution Hybrid Shape Models to Brain MRI. In K.H. Höhne and R. Kikinis, editors, *Visualization in Biomedical Computing*, pages 213–222. Springer, September 1996.

[Vemuri *et al.* 19967] B.C. Vemuri, Y. Guo, Ch.M. Leonard, and Sh.-H. Lai. Fast numerical algorithms for fitting multiresolution hybrid shape models to brain MRI. *Medical Image Analysis*, 1(4):343–362, 1996/7.

[Wang and Staib 1998] Y. Wang and L.H. Staib. Physical Model Based Non-Rigid Registration Incorporating Statistical Shape Information. In *To appear in the proceedings of MICCAI'98*, 1998.

[Wyskocil 199798] T Wyskocil. Symmetrien geschlossener Kurven und Flächen. Master's thesis, ETH Zurich, 1997/98.