Model-based Initialisation for Segmentation

Johannes Hug, Christian Brechbühler, and Gábor Székely

Swiss Federal Institute of Technology, ETH Zentrum
CH-8092 Zürich, Switzerland
{jhug | brech | szekely}@vision.ee.ethz.ch

Abstract. The initialisation of segmentation methods aiming at the localisation of biological structures in medical imagery is frequently regarded as a given precondition. In practice, however, initialisation is usually performed manually or by some heuristic preprocessing steps. Moreover, the same framework is often employed to recover from imperfect results of the subsequent segmentation. Therefore, it is of crucial importance for everyday application to have a simple and effective initialisation method at one's disposal. This paper proposes a new model-based framework to synthesise sound initialisations by calculating the most probable shape given a minimal set of statistical landmarks and the applied shape model. Shape information coded by particular points is first iteratively removed from a statistical shape description that is based on the principal component analysis of a collection of shape instances. By using the inverse of the resulting operation, it is subsequently possible to construct initial outlines with minimal effort. The whole framework is demonstrated by means of a shape database consisting of a set of corpus callosum instances. Furthermore, both manual and fully automatic initialisation with the proposed approach is evaluated. The obtained results validate its suitability as a preprocessing step for semi-automatic as well as fully automatic segmentation. And last but not least, the iterative construction of increasingly point-invariant shape statistics provides a deeper insight into the nature of the shape under investigation.

1 Introduction

The advent of the "Active Vision" paradigm in the 1980s came along with the idea of using model-based prior knowledge to simplify and stabilise the treatment of a specific vision problem. Since then, all kinds of active shape models have emerged in many application areas in various forms such as Snakes [10], deformable templates [23] or active appearance models [2]. The amount of prior knowledge included in these models varies from simple general smoothness assumptions to very detailed knowledge about the shape and the image data to be expected. In the field of medical imaging, the usage of statistical shape models has found widespread use [3, 21, 12, 13], since the notion of biological shape seems to be best defined by a statistical description of a large population.

Even though these statistical methods have proven to be fairly stable and reliable, there are cases where they fail completely in finding at least an approximation of the correct object boundary. If a certain application asks for absolutely

flawless segmentations, alternative or supplemental frameworks must be applied to compensate for the missing functionality. On the one hand, we could employ semi-automatic [6, 10, 17, 8] or manual segmentation tools that rely on a human operator providing the missing information. On the other hand, we may initialise the fully automatic procedure such that the correct solution is just nearby the initial configuration. Since almost all semi-automatic methods rely on suitable initialisations as well, the provision of a reasonable starting point seems to be a valuable extension of both approaches. Our main goal is therefore to provide a possibly interactive initialisation method that still takes into account the prior knowledge of the shape as far as possible.

In order to keep the amount of required user input as small as possible, simple and intuitive interaction metaphors are of crucial importance for the design of such a tool. Since the most simple and probably most feasible interaction metaphor is still the adjustment of individual points lying on the boundary of the object under investigation, we are subsequently looking for a small number of points describing the overall shape of the object to be segmented — analogous to coarse control polygons of hierarchical shape descriptions that have recently been proposed in the field of modelling and animation [7, 24]. Such a "coarse control polygon" should capture as much prior shape knowledge as possible. And there should be a way to calculate the most "natural" fine scale shape given the correct arrangement of the control vertices.

Our shape database should therefore be able to answer the following three questions: Which points along the object boundary are best suited for a compact and robust description of the shape? How many control vertices must be included in the coarsest control polygon? And how should the full resolution object be predicted so as to provide a reasonable initial outline?

In search of answers to these questions, we have decided to pursue the following strategy: Using statistical shape analysis, we examine the remaining variability of shape, if the variation coded by the position of individual points is progressively subtracted. The coarsest control polygon necessary to capture the main shape characteristics is complete as soon as the remaining variability is small with respect to the working range of the subsequent segmentation method. And the most probable shape for a given control polygon can then be calculated by just inverting the process of subtracting the variation of control vertices.

2 Experimental Set-up

In order to have a compact statistical shape description at our disposal, we employ a representation that is based on a principal component analysis (PCA) of all object instances in our database. This approach, first proposed by Cootes and Taylor in [4], has the very useful property to reflect the shape variations occurring within the population by a complete set of basis vectors. These basis vectors span a linear shape space containing all the instances of our collection. This enables us to apply the whole framework of linear algebra to make the statistic

point-wise invariant. Furthermore, the properties of such a shape description are well understood and appropriately documented [11].

In addition to a statistical description method, we need a population of several object instances representing our model-based foreknowledge. The PCA is therefore applied to a collection of 71 hand segmented outlines of the corpus callosum on mid-sagittal MR-slices. Five randomly selected examples of this database are illustrated in Fig. 1. All aspects of the model building process regarding this population are described in detail in [22].

Since we aim at working with a vertex-based control polygon at interactive speed, the original representation based on elliptic Fourier descriptors [20, 14, 21] has been converted to a polygonal representation by equidistantly sampling the parameter space of the outline. For the following analysis, we assume that the underlying arc-length based curve parameterisation with normalised parameter starting point provides a sufficiently good correspondence between the individual specimen. All experiments we performed suggest that the achieved correspondence is not faultless but sufficiently precise for our intentions (see also [11]). In order to normalise the model contours, we represented the vertex positions as usual with respect to an anatomical coordinate system given by the AC-PC line. Experience shows that these anatomical landmarks can easily be located and are very stable with respect to the corpus callosum.

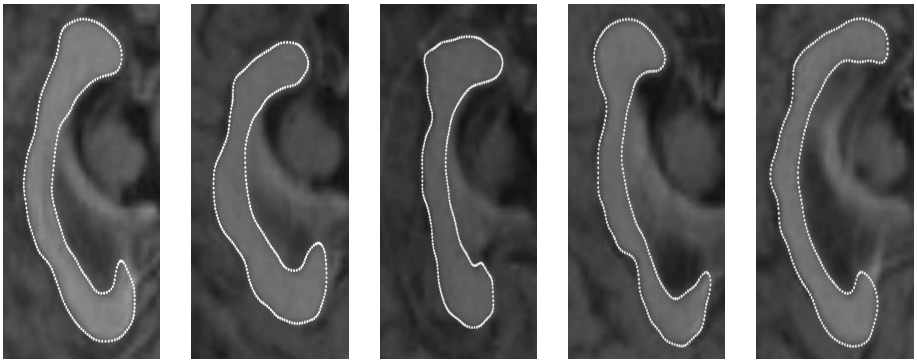


Fig. 1. Five randomly selected corpora callosa from our collection that consists of 71 examples.

In the following, Section 3 reviews shortly the statistical shape analysis using principal components and fixes the mathematical notation. In Section 4, we discuss in detail the aforementioned progressive subtraction of variation, and Section 5 describes subsequently the inversion of this operation. The initialisation procedure founding on the presented framework is evaluated in Section 6 for both interactive and fully automatic mode of operation. Finally, Section 7 concludes this report and outlines the next steps towards a highly robust initialisation oracle.

3 Shape Analysis using Principal Components

The basic idea of statistical shape analysis using principal components consists in separating and quantifying the main variations of shape that occur within a population of several instances exemplifying the same object. More precisely, a PCA defines a linear transformation that decorrelates the parameter signals of the original shape population by projecting the objects into a linear shape space spanned by a complete set of orthogonal basis vectors. If the parameter signals are highly correlated, then the coarse scale variations of shape are described by the first few basis vectors, whereas fine details are captured by the remaining ones. Furthermore, if the joint distribution of the parameters describing the shape is Gaussian, then a reasonably weighted linear combination of the basis vectors results in a shape that is similar to the existing ones. On the other hand, if the joint distribution of the parameters is highly non-Gaussian or if the dependencies of the parameter signals are non-linear, then other decomposition methods such as the independent component analysis [9] should be employed.

As already mentioned, the considered population consists of $N+1=71$ corpus callosum instances, given as polygonal models $\mathbf{p}_i = [x_i^{[1]}, y_i^{[1]}, \ldots, x_i^{[M]}, y_i^{[M]}]^T$ with $M = 256$ points. Since we will later compare statistic-based initialisations to the ground truth given by one object instance, we always exclude this particular instance from the statistic for cross-validation. To simplify the formalism, we centre the parameter signals of the shapes beforehand by calculating an average model $\overline{\mathbf{p}}$ and an instance specific difference vector $\Delta\mathbf{p}_i$:

$$\overline{\mathbf{p}} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{p}_i, \qquad \Delta\mathbf{p}_i = \mathbf{p}_i - \overline{\mathbf{p}}, \qquad \Delta P = [\Delta\mathbf{p}_1 \ \cdots \ \Delta\mathbf{p}_N] \tag{1}$$

Note, the $N = 70$ difference vectors span only a 69-dimensional space; the missing dimension obviously originates from the linear dependence $\sum_{i=1}^{N}\Delta\mathbf{p}_i = 0$. The corresponding covariance matrix $\Sigma \in \mathbb{R}^{2M\times 2M}$ is consequently rank-deficient. As has been pointed out in [5], this circumstance can be exploited to speed up the calculation of the 69 valid eigenvalues and eigenvectors: Instead of calculating the full eigensystem of the covariance matrix Σ, the multiplication of the eigenvectors of a smaller matrix $\breve{\Sigma}$ with ΔP leads to the correct principal components:

$$\begin{gathered}\breve{\Sigma} = \frac{1}{N-1}\,\Delta P^T\Delta P \overset{\text{PCA}}{=} \breve{U}\Lambda'\breve{U}^T, \quad \Lambda' = \operatorname{diag}(\lambda_1,\ldots,\lambda_{N-1},0) \\ U' = [\mathbf{u}_1 \ \cdots \ \mathbf{u}_{N-1}\,\mathbf{u}_N] = \psi\left(\Delta P\,\breve{U}\right), \quad \psi(A) = \text{Normalise columns of A}\end{gathered} \tag{2}$$

As an alternative that is not equally fast but conceptually more elegant, we propose to work in a subspace with a complete set of basis vectors to find the eigensystem of our data. To do so, we project the difference vectors $\Delta\mathbf{p}_i$ into a lower dimensional space whose basis M is constructed by the Gram-Schmidt orthonormalisation χ:

$$M = [\mathbf{m}_1 \ \cdots \ \mathbf{m}_{N-1}] = \chi(\Delta\mathbf{p}_1,\ldots,\Delta\mathbf{p}_{N-1}), \qquad \Delta\tilde{\mathbf{p}}_i = M^T\,\Delta\mathbf{p}_i \tag{3}$$

Note, one arbitrary $\Delta\mathbf{p}_i$ must be dropped for the construction of M, and $\Delta\tilde{\mathbf{p}}_i$ denotes the projection of $\Delta\mathbf{p}_i$ into the subspace spanned by M. The covariance matrix $\tilde{\Sigma}$ and the resulting PCA given by the eigensystem of $\tilde{\Sigma}$ can subsequently be calculated according to:

$$\tilde{\Sigma} = \frac{1}{N-1}\sum_{i=1}^{N} \Delta\tilde{\mathbf{p}}_i\, \Delta\tilde{\mathbf{p}}_i^T \overset{\text{PCA}}{=} \tilde{U}\Lambda\tilde{U}^T, \qquad \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_{N-1}) \tag{4}$$

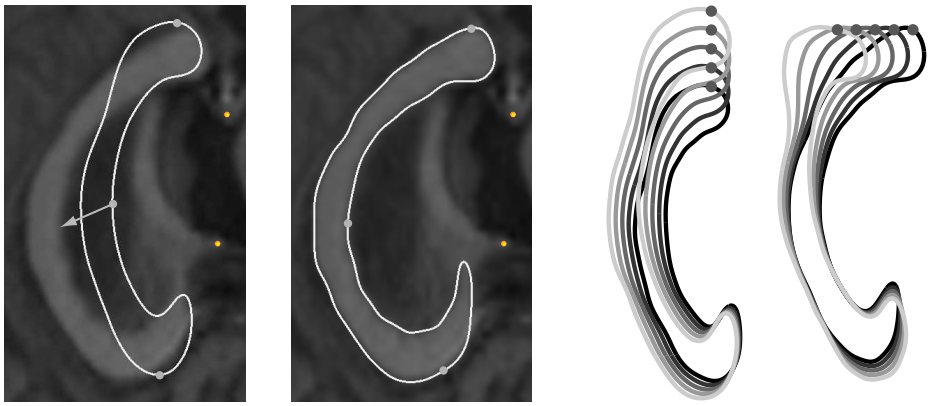
The principal components defining the eigenmodes *in shape space* are then given by back-projecting the eigenvectors $\tilde{U}$: $U = [\mathbf{u}_1 \cdots \mathbf{u}_{N-1}] = M\tilde{U}$. Each object instance can be represented as a linear combination $\mathbf{p}_i = \overline{\mathbf{p}} + U\mathbf{b}_i$ of these eigenmodes, where $\mathbf{b}_i = [b_i^{[1]}, \ldots, b_i^{[N-1]}]^T$. In order to calculate the uncorrelated coordinates $\mathbf{b}_i$ of each object instance, we project the difference vectors $\Delta\mathbf{p}_i$ into the eigenspace: $\mathbf{b}_i = U^T\Delta\mathbf{p}_i$.

The first four eigenmodes resulting from the PCA of our population are displayed in Fig. 3(a). The shapes representing the first eigenmode on the left are calculated by adding the weighted first eigenvector $\mathbf{u}_1$ to the average model $\overline{\mathbf{p}}$. The following three shape variations to the right of the first one are calculated correspondingly.

4 Progressive Elimination of Variation

Given a statistical analysis as defined above, we consider the following situation: After having defined the shape coordinate system by locating the AC-PC line, the initialisation of a new object instance starts with the average model $\overline{\mathbf{p}}$, as illustrated in Fig. 2(a) on the left. Let us assume for the moment that the aforementioned coarse control polygon consists of the three marked vertices on the outline of the mean shape. To generate an initial approximation of the object, we define now a set of boundary conditions for the global shape by moving the control vertices to an approximately correct position. Given these constraints and our prior knowledge of the shape, we wish to choose that outline for initialisation which is most natural in that case. In the following two sections we will show, how this most probable outline can be found.

Since we hope that some control vertices carry more shape information than others, we approach the whole problem iteratively. In a first step, we calculate the most probable shape that satisfies only the boundary conditions provided by the most important control vertex. For the second most important control point we use subsequently the resulting outline as initial configuration. This process is then repeated until we can satisfy all the boundary conditions. Since we do not yet know how to determine the most important control vertex, we will first investigate the computation of the most probable shape given the position of an *arbitrary* point. This will be the subject of the next subsection. The problem of finding the points carrying most shape information will be discussed afterwards in subsection 4.3.



(a) Initial average model (left) and correct segmentation (right)

(b) Basis vectors R_j

Fig. 2. (a) Boundary conditions for an initial outline are established by prescribing a position for each coarse control vertex. (b) Shape variations caused by adding the two basis vectors R_j to the average model, inducing x- and y-translations of point j, respectively. The various shapes are obtained by evaluating $\overline{\mathbf{p}} + \omega\, U\, \mathbf{r}_k$ with $\omega \in \{-2, \ldots, 2\}$ and $k \in \{x_j, y_j\}$.

4.1 Shape-based Basis Vectors for one Point

To start with, we must translate our conceptual goal into mathematical terms. Since the most probable shape is given by the mean model $\overline{\mathbf{p}}$ in the context of PCA, we can reinterpret the notion of "choosing the most probable outline" as "choosing the shape with minimal deviation from the mean". And this means nothing else than choosing the model with minimal Mahalanobis-distance D_m, the common metric in eigenspaces.

The key idea enabling the solution of our first problem can now be summarised as follows: We must find two vectors *in the space of variation* that describe decoupled x- and y-translations of a given point j with minimal variation, respectively. In other words, these two vectors should cause a unit translation of vertex j in either x- or y-direction, and they should have minimal Mahalanobis-length D_m. If we have found them, we can satisfy all possible boundary conditions caused by one vertex with minimal variation by just adding the two appropriately weighted "basis" vectors to the mean. This problem gives rise to the following constrained optimisation:

Let $\mathbf{r}_{x_j}$ and $\mathbf{r}_{y_j}$ denote the two unknown basis vectors causing unit x- and y- translation of point j, respectively. The Mahalanobis-length D_m of these two vectors is then given by:

$$D_m(\mathbf{r}_k) \;=\; (\tilde{U}\mathbf{r}_k)^T \tilde{\Sigma}^{-1} \tilde{U}\mathbf{r}_k \;=\; \mathbf{r}_k^T \Lambda^{-1} \mathbf{r}_k \;=\; \sum_{e=1}^{N-1} \frac{\left(r_k^{[e]}\right)^2}{\lambda_e}, \qquad k \in \{x_j, y_j\} \quad (5)$$

Taking into account that x_j and y_j depend only on two rows of U, we define the sub-matrix U_j according to the following expression:

$$\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} \overline{x}_j \\ \overline{y}_j \end{bmatrix} + \begin{bmatrix} u_{2j-1\,\circ} \\ u_{2j\,\circ} \end{bmatrix} \mathbf{b} = \begin{bmatrix} \overline{x}_j \\ \overline{y}_j \end{bmatrix} + U_j\,\mathbf{b}, \qquad u_{j\,\circ} = j^{\text{th}} \text{ row of } U \tag{6}$$

In order to minimise the Mahalanobis-distance D_m subject to the constraint of a separate x- or y-translation by one unit, we establish — as is customary for constrained optimisation — the Lagrange function L:

$$L(\mathbf{r}_k, \mathbf{l}_k) = \sum_{e=1}^{N-1} \frac{\left(r_k^{[e]}\right)^2}{\lambda_e} \;-\; \mathbf{l}_k^T \left[U_j \mathbf{r}_k - \mathbf{e}_k\right], \qquad \mathbf{e}_{\{x_j, y_j\}} = \{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \} \tag{7}$$

The vectors $\mathbf{l}_{x_j}$ and $\mathbf{l}_{y_j}$ contain as usual the required Lagrange multipliers. To find the minimum of $L(\mathbf{r}_k, \mathbf{l}_k)$, we calculate the derivatives with respect to all elements of $\mathbf{r}_{x_j}$, $\mathbf{r}_{y_j}$, $\mathbf{l}_{x_j}$, and $\mathbf{l}_{y_j}$ and set them equal to zero:

$$\left.\begin{array}{l} \dfrac{\delta L(\mathbf{r}_k, \mathbf{l}_k)}{\delta \mathbf{r}_k} \stackrel{!}{=} 0 \\ \dfrac{\delta L(\mathbf{r}_k, \mathbf{l}_k)}{\delta \mathbf{l}_k} \stackrel{!}{=} 0 \end{array}\right\} \quad \begin{bmatrix} \frac{2}{\lambda_1} & & \vdots & \\ & \ddots & \vdots & -U_j^T \\ & & \frac{2}{\lambda_{N-1}} \vdots & \\ \cdots & \cdots & \cdots & \cdots \\ & U_j & \vdots & 0 \end{bmatrix} \begin{bmatrix} & \vdots & \\ & \mathbf{r}_{x_j} \vdots \mathbf{r}_{y_j} & \\ & \vdots & \\ \cdots & \cdots & \cdots \\ & \mathbf{l}_{x_j} \vdots \mathbf{l}_{y_j} & \end{bmatrix} = \begin{bmatrix} & \vdots & \\ & \mathbf{0} \vdots \mathbf{0} & \\ & \vdots & \\ \cdots & \cdots & \cdots \\ & \mathbf{e}_{x_j} \vdots \mathbf{e}_{y_j} & \end{bmatrix} \tag{8}$$

If the basis vectors and the Lagrange multipliers are combined according to $R_j = [\mathbf{r}_{x_j}\ \mathbf{r}_{y_j}]$ and $L_j = [\mathbf{l}_{x_j}\ \mathbf{l}_{y_j}]$, Eq. (8) can be rewritten as two linear matrix equations:

$$2\Lambda^{-1} R_j = U_j^T L_j \tag{9}$$

$$U_j R_j = I \tag{10}$$

The two basis vectors $\mathbf{r}_{x_j}$ and $\mathbf{r}_{y_j}$ resulting from simple algebraic operations (resolve (9) for R_j and replace R_j in (10) by the result, use to resulting equation to find $L_j = 2[U_j \Lambda U_j^T]^{-1}$ and substitute for L_j in (9)) are then given by:

$$R_j = \begin{bmatrix} \mathbf{r}_{x_j} & \mathbf{r}_{y_j} \end{bmatrix} = \Lambda\, U_j^T \left[U_j\, \Lambda\, U_j^T\right]^{-1} \tag{11}$$

While $\mathbf{r}_{x_j}$ describes the translation of x_j by one unit with constant y_j and minimal shape variation, $\mathbf{r}_{y_j}$ alters y_j correspondingly. The resulting effect caused by adding these shape-based basis vectors to the average model is illustrated in Fig. 2(b). The most probable shape $\check{\mathbf{p}}$ given the displacement $[\Delta x_j, \Delta y_j]^T$ of control vertex j is consequently determined by

$$\check{\mathbf{p}} = \overline{\mathbf{p}} + U R_j \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix} . \tag{12}$$

Another possibility to find the two basis vectors R_j consists in exploiting the least-squares property of the Moore-Penrose pseudo-inverse. The basic idea in this context is to solve the highly under-determined linear system $U_j \mathbf{r}_k = \mathbf{e}_k$ (representing the prescribed constraints) by calculating the pseudo-inverse $U_j^{\#}$. Since we are not looking for the normal least-squares solution but for the one with minimal Mahalanobis-distance, we have to introduce a weighting of the rows of U_j, in order to map the problem into normal Euclidean space, where the minimal solution is then given by the generalised inverse. As shown in Appendix A, this approach leads to the same basis vectors R_j and validates Eq. (11), since there is only one unique element in the hyper-plane of all solutions that has minimum Mahalanobis-norm.

4.2 Point-wise Subtraction of Variation

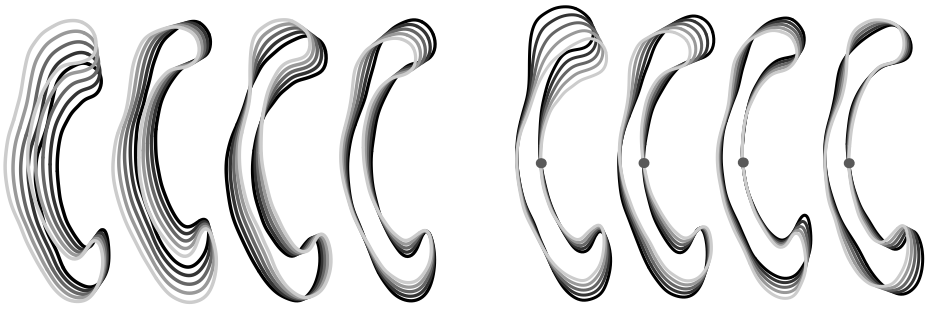
In the previous subsection we have seen how to choose the most probable shape given the position of one specific control vertex j. Before we can now proceed to the next control point, we must ensure that subsequent shape modifications will not alter the previously adjusted vertex j. To do so, we must remove those components from the statistic that cause a displacement of this point. Unfortunately, we cannot apply a projection for this purpose, since the basis vectors R_j are not orthogonal in the shape space. Therefore, we propose to subtract the variation coded by the point j from each instance i, and to rebuild the statistic afterwards. For the first part of this operation, we must subtract the basis vectors R_j weighted by the example-specific displacement $[\Delta x_j, \Delta y_j]_i^T$ from the parameter representation $\mathbf{b}_i$ of each instance i:

$$\mathbf{b}_i^{\hat{j}} = \mathbf{b}_i - R_j \begin{bmatrix} \Delta x_j \\ \Delta y_j \end{bmatrix}_i = \mathbf{b}_i - R_j\, U_j\, \mathbf{b}_i = (I - R_j\, U_j)\ \mathbf{b}_i, \quad \forall i \in \{1, .., N\} \tag{13}$$

Doing so for all instances, we obtain a new description of our population which is invariant with respect to the point j (denoted by $\circ^{\hat{j}}$). The variability in this point-normalised population is expected to be smaller compared to the original collection. In order to verify this assumption and to rebuild the statistic, we apply anew a PCA to the normalised set of instances $\{\mathbf{b}_i^{\hat{j}} \mid i \in \{1, \ldots, N\}\}$. Note, the eigenspace shrinks by two dimensions since we removed two degrees of freedom. The resulting principal components, denoted by $U^{\hat{j}}$, confirm the expected behaviour and validate also the removal of the variation of point j. The first four one-point invariant eigenmodes are illustrated in Fig. 3(b).

4.3 Point Selection Strategy

The point-wise elimination of variability presented above can subsequently be repeated for several points, until the remaining variability is small enough with respect to the working range of the subsequent segmentation algorithm. In order to achieve optimal results and to find the most compact control polygon, we should now explore the strategy for the selection of control points. Since we



(a) Original eigenmodes (b) One-point invariant eigenmodes

Fig. 3. (a) The first four eigenmodes of 70 corpus callosum instances. The various shapes are obtained by evaluating $\overline{\mathbf{p}}+\omega\sqrt{\lambda_k}\mathbf{u}_k$ with $\omega \in \{-2,\ldots,2\}$ and $k \in \{1,\ldots,4\}$. (b) The first four one-point invariant eigenmodes after subtracting the first principal landmark. The various shapes are obtained by evaluating $\overline{\mathbf{p}} + \omega\sqrt{\lambda_k^{\hat{j}}}\mathbf{u}_k^{\hat{j}}$ with $\omega \in \{-2,\ldots,2\}$ and $k \in \{1,\ldots,4\}$.

aim to choose those vertices that carry as much shape information as possible, we should select the points according to their "reduction potential". A control vertex holds a large reduction potential, if the remaining variability after its elimination is small.

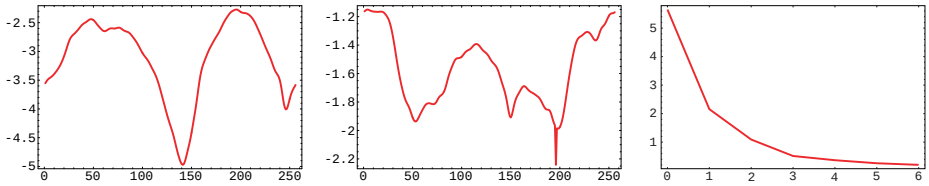
To make the following formalism as precise as possible, we introduce some additional definitions at this point: Firstly, we will subsequently refer to the k^{th} point being removed from the statistic as the k^{th} *principal landmark*. Secondly, let the sequence $\hat{s}_k = \{\hat{j}_1,\ldots,\hat{j}_k\}$ denote the set of point-indices of those k principal landmarks that have been removed from the statistic in the given order. And last but not least, the superscript $\circ^{\hat{s}_k}$ is used for the value of $\circ$, if the principal landmarks $\hat{s}_k$ have been removed.

Using this formalism, the reduction potential P of vertex j_k, being a candidate to serve as the k^{th} principal landmark, can be defined as follows:

$$P(j_k) = -\sum_{l=1}^{N-1-2(k-1)} \left(\tilde{\sigma}_l^2\right)^{\hat{s}_k} = -\text{tr}(\tilde{\Sigma}^{\hat{s}_k}) = -\text{tr}(\Lambda^{\hat{s}_k}), \quad \hat{s}_k = \{\hat{j}_1,\ldots,\hat{j}_k\} \quad (14)$$

Figure 4(a) shows the reduction potential for all the points of the original model. In order to remove as much variation as possible, we choose consequently that point as the first principal landmark that holds the largest reduction potential: $j_1 = \max_j[P(j)]$. The selected vertex and the resulting point-invariant statistic after its elimination have already been shown in Fig. 3(b).

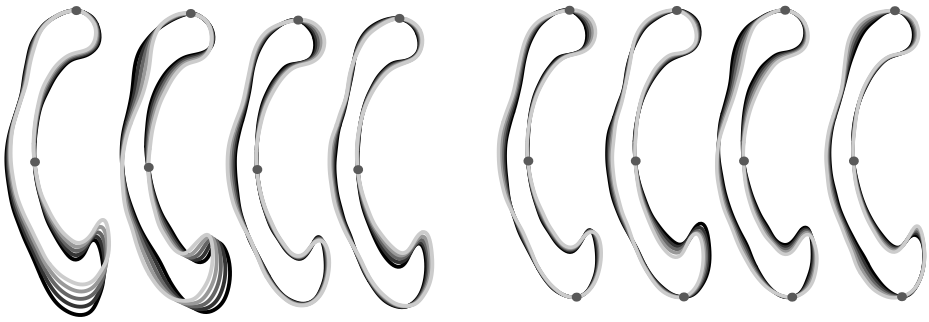
If we apply this selection and elimination step twice again, we end up with the second and third principal landmark. The corresponding eigenmodes and the selected points are depicted in Fig. 5. The decreasing deviations from the mean indicate that the variation within the population is progressively reduced by this operation. The observation of the overall variance subject to progressive



(a) (b) (c)

Fig. 4. (a) & (b) Reduction potentials for the selection of (a) the first and (b) the second principal landmark. For each point j in the abscissa, the reduction potential $P(j)$ is displayed. Note, the first principal landmark $j_1 = 196$ has minimal reduction potential in (b), because subtracting the same point twice has no effect at all. (c) The overall variance $\mathrm{tr}(\tilde{\Sigma}^{\hat{s}_k})$ of the population depending on the number of subtracted principal landmarks.

point removal (see Fig. 4(c)) verifies this hypothesis and shows that the variability decreases surprisingly fast in the beginning. Later on, after three vertices have been processed, the decline levels out and the benefit of each additional principal landmark becomes fairly small. This finding suggests that the main shape characteristics of a corpus callosum can be captured by only three or four principal landmarks.



(a) Two-point invariant eigenmodes (b) Three-point invariant eigenmodes

Fig. 5. Remaining variability after vertex elimination of (a) two and (b) three principal landmarks.

5 Initial Shapes for Segmentation

The progressive application of the point selection and removal process enables now the construction of the most compact *principal control polygon*, consisting of the first few principal landmarks. Analogous to traditional parametric curve representations, each control point has two associated *principal basis functions*

(UR_j) that are globally supported. The final outline $\check{\mathbf{p}}_l$ based on a principal control polygon with l vertices is then given by the inverse operation of the construction, that is, by the combination of the mean shape $\overline{\mathbf{p}}$ and all the weighted principal basis functions $R_{j_k}^{\hat{s}_{k-1}}$:

$$\check{\mathbf{p}}_l = \overline{\mathbf{p}} + \sum_{k=1}^{l} U^{\hat{s}_{k-1}} R_{j_k}^{\hat{s}_{k-1}} \begin{bmatrix} \Delta x_{j_k} \\ \Delta y_{j_k} \end{bmatrix} \tag{15}$$

Note that the weights $[\Delta x_{j_k}, \Delta y_{j_k}]^T$ for the basis vectors $U^{\hat{s}_{k-1}} R_{j_k}^{\hat{s}_{k-1}}$ depend on the shape defined by the previous principal landmarks $\hat{s}_{k-1}$. Therefore, if any control point j_k is modified and the less important landmarks shall remain in their position, the weights $\{[\Delta x_{j_{k+1}}, \Delta y_{j_{k+1}}]^T, \ldots, [\Delta x_{j_l}, \Delta y_{j_l}]^T\}$ must be recalculated in the correct order of vertex removal. To emphasise the hierarchical structure of our formalism and to simplify the algorithmic implementation, we recommend to use the following recursive definition instead of equation (15):

$$\check{\mathbf{p}}_0 = \overline{\mathbf{p}} \ , \qquad \check{\mathbf{p}}_k = \check{\mathbf{p}}_{k-1} + U^{\hat{s}_{k-1}} R_{j_k}^{\hat{s}_{k-1}} \begin{bmatrix} \Delta x_{j_k} \\ \Delta y_{j_k} \end{bmatrix} \tag{16}$$

With this shape-based curve representation $\check{\mathbf{p}}$, the last piece has fallen into place. By utilising a minimal principal control polygon with associated basis functions, we are now able to fulfil all our original objectives: The initialisation of a new shape instance results in the simple adjustment of a small number of points, taking into account all our prior knowledge of the shape.

In order to validate the quality of the proposed method, we will subsequently show some results of cross-validation experiments that have been performed with each shape instance in our database. It goes without saying that the test instance has always been removed from the statistic. The initialisations to be presented have been generated by moving the principal landmarks into the positions of the corresponding points on the outline of the respective test object. A selection of the results of these experiments is illustrated in Fig. 6 and can be summarised as follows: The initial average model in Fig. 6(a) converges efficiently towards an approximation of the correct shape whilst the control vertices are adjusted. In most of our examples, only three or four principal landmarks are necessary to provide a reasonably good initialisation. The consideration of more than five or six points does not significantly improve the quality of the initial shape. In some cases, the initialisation even deteriorates slightly, if too much control vertices are employed. This behaviour may also indicate some deficiencies of the underlying correspondence function. To show a representative cross-section of the achieved results, Fig. 6(b) displays four truly randomly chosen experiments, where four principal landmarks have been adjusted.

6 Interactive and Automatic Initialisation

The major question remaining to be answered is, whether the proposed framework proves its worth in practical application as well. Although we have not yet

(a) Progressive initialisation (b) Four different initialisations

Fig. 6. (a) Generation of an initial outline for segmentation; shape instance in black and fitted initialisations in gray with an increasing number of fitted principal landmarks. (b) Initial shapes with four adjusted principal landmarks for the segmentation of four randomly chosen instances.

gained any experience in everyday clinical application, our experimental results are fairly convincing. Tests have been performed for both interactive and fully automatic initialisation.

The interactive approach simply uses the underlying shape basis as a highly specialised curve representation. The required adjustments of the principal landmarks must be provided by a human operator. Since the recalculation of the outline can be done at interactive speed, the instant feedback supports the operator in finding an appropriate initialisation within a few seconds. In most of the cases, three principal landmarks are sufficient to define a coarse initialisation. At most three additional control vertices can then be used to refine the characteristic details of the shape. Figure 7(a) shows one possible initialisation based on six manually adjusted principal landmarks.

By exploiting the statistical prior knowledge of the shape once again, we can even eliminate the remaining interaction: For each principal landmark j_k, we calculate the covariance matrix $\Sigma_{j_k}^{\hat{s}_{k-1}}$ in order to determine its positional variability. On the assumption that a landmark is Gaussian distributed, we can then compute a confidence ellipse that contains the corresponding control point with probability $\chi_2^2(c) = P(|\boldsymbol{\omega}| < c)$ (see e.g. [1]). The new auxiliary variable $\boldsymbol{\omega}$ is, as usual in this context, a standardised random vector with normal distribution: $\omega_i \sim \mathcal{N}(0,1) \wedge \boldsymbol{\omega} \sim \mathcal{N}(0,I)$. Since it is well known that $\chi_2^2(3) = P(|\boldsymbol{\omega}| < 3) \approx 99\%$, we can construct the main axes a_{j_k} and b_{j_k} of the confidence ellipse that contains the principal landmark with a probability of 99% by the following linear transformation of $\boldsymbol{\omega}$:

$$a_{j_k} = \sqrt{\Sigma_{j_k}^{\hat{s}_{k-1}}}\boldsymbol{\omega}_{\parallel}, \quad b_{j_k} = \sqrt{\Sigma_{j_k}^{\hat{s}_{k-1}}}\boldsymbol{\omega}_{\perp}; \qquad \boldsymbol{\omega}_{\parallel} = 3\begin{bmatrix}1\\0\end{bmatrix}, \quad \boldsymbol{\omega}_{\perp} = 3\begin{bmatrix}0\\1\end{bmatrix} \tag{17}$$

Figure 7(b) shows these confidence ellipses for all considered control vertices. As expected, the length of the axes a_{j_k} and b_{j_k} declines with increasing k, according to the smaller variances in the underlying statistics.

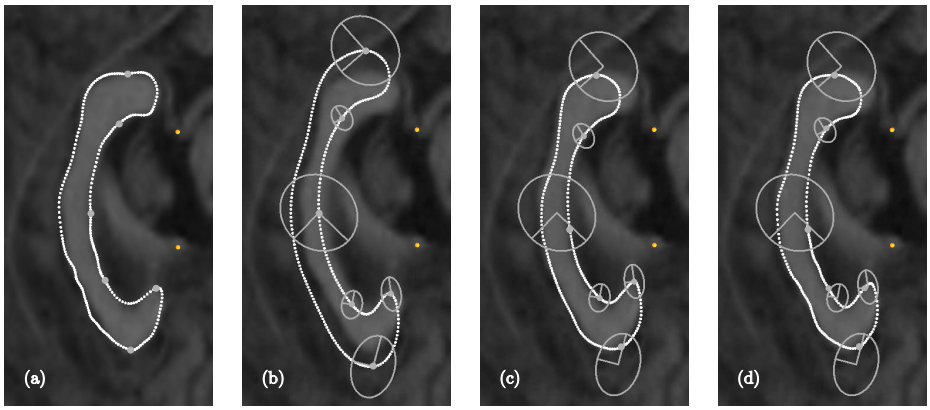


Fig. 7. (a) Interactive initialisation by manual adjustment of six principal landmarks. (b) Initial average model with the confidence ellipses of the control vertices. (c) & (d) Automatic initialisation by the sequential optimisation of the matching function G_k for (c) three and (d) six principal landmarks.

For an automatic initialisation, we can subsequently use these confidence intervals as the region of interest with respect to an optimisation of the fit. The goal function of such an optimisation should measure the correspondence between the shape $\check{\mathbf{p}}_k$ to be optimised and the actual image data I. In order to simplify and accelerate the optimisation process, we propose to fit only one principal landmark at a time, analogous to manual initialisation. By employing a very popular matching function based on the image gradient ∇I, we end up with the following goal function G_k:

$$G_k(\Delta x_{j_k}, \Delta y_{j_k}) = \sum_{e=1}^{M} \| \nabla I[\, \check{\mathbf{p}}_k^{[e]}(\Delta x_{j_k}, \Delta y_{j_k})\,] \|, \qquad I: \text{ Image data} \tag{18}$$

Note that G_k depends on the results of the previously optimised principal landmarks $\hat{s}_{k-1}$, since the centre of the confidence ellipse k is given by $\check{\mathbf{p}}_{k-1}^{[j_k]}$. A closer inspection of the goal functions G_k within the confidence ellipse k shows that the most important goal functions G_1, G_2, and G_3 exhibit several local minima and maxima. But apart from this minor difficulty, their overall behaviour is fairly smooth and regular. However, due to the hierarchical dependencies, it is essential to reliably locate the global maximum. Therefore, we propose the following simple optimisation scheme: In a first step, we sample the goal function within the bounding box of the confidence ellipse on a coarse grid, in order to find the local neighbourhood of the global optimum. Having done so, we apply the Newton-Raphson method to find the proper optimum. Since the computation of a Newton-Raphson iteration includes the calculation of first and second derivatives, we recommend to fit a bivariate Taylor polynom of fourth degree around the estimated optimum, instead of relying on discrete derivatives.

This optimisation scheme has proven to be robust and it finds reliably the global optimum with high sub-pixel accuracy. The time taken to optimise one principal landmark amounts to about one second on an SGI O_2. Figure 7(d) shows the result of the initialisation, if the optimisation is sequentially applied to six principal landmarks. With the exception of the Splenium of the corpus callosum, the resulting outline is very close to the optimum. A comparison between the final outline and the manual initialisation in Fig. 7(a) shows two differences: On the one hand, the automated method obviously detects the border of the shape with higher precision. On the other hand, manual initialisation seems to be superior with respect to an overall fit to the image data. Although we could speculate that the better estimate induces a higher distortion of the correspondence function, we suspect that the superior performance has another reason: The manual approach simply finds a better solution regarding the problem of optimising the position of all principal landmarks at once. The automatic optimisation of this problem is much more difficult due to the dependencies of the goal functions G_k and has not yet been investigated.

7 Conclusion and Future Research

In search of a stable initialisation oracle that is based on a small number of points, we presented a new way to make a statistical shape description point-wise invariant. The inverse of the resulting operation generates initial configurations for subsequent segmentation by choosing the most probable shape given the estimated control polygon. The whole framework has been evaluated by means of a shape population consisting of 71 corpus callosum instances. To demonstrate its practical benefit, we implemented both an interactive and a fully automatic initialisation method. The achieved results are satisfying and validate its suitability for our initialisation purposes. Furthermore, we gained a deeper insight into the nature of the shape under investigation by finding the most compact shape description given by the principal control polygon with associated principal basis functions.

Additional work has to be done in order to evaluate and improve the practical application of the proposed shape analysis. In the context of interactive initialisation, we must explore the influence of the point selection strategy on the user's ability to locate the prescribed vertices in the image. Since we choose the principal landmarks purely on the basis of a statistical measure, problems may arise in locating the correct position of the points in the image to be segmented. Hence, another point selection strategy could be based on the analysis of *local* shape and image characteristics. Control vertices with salient local curve features or locations with stable image characteristics could serve as landmarks well suited for automatic or interactive localisation. Such point selection oracles should be combined with our statistical selection strategy. Moreover, the automatic initialisation should be improved by optimising all the control vertices at once.

Last but not least, model-based initialisations for surfaces should be provided as well, in order to overcome the limitations imposed by the two-dimensional segmentation approach, if three dimensional data sets are available. And if we broaden the horizons beyond the borders of computer vision, we surmise that our framework could be of great value for the interactive animation of various natural objects.

8 Acknowledgements

We would like to express our sincere gratitude to the BIOMORPH consortium (BIOMORPH, EU-BIOMED2 project no BMH4-CT96-0845) for providing the data and the manual segmentation of the 71 corpora callosa.

A Derivation of the Basis Vectors $\mathbf{R}_j$ by Means of the Pseudo-Inverse

Another approach to find the two basis vectors $\mathbf{r}_{x_j}$ and $\mathbf{r}_{y_j}$ causing unit translation in x- and y-direction with minimal Mahalanobis-distance D_m involves the calculation of the generalised Moore-Penrose pseudo-inverse [16, 18]. To derive a solution with this concept, we use only the prescribed constraints as a starting point:

$$U_j R_j = I \tag{19}$$

Since U_j is a $2 \times (N-1)$ matrix, the linear system of equations in (19) is highly *under-determined*. Such a system either has no solution or there will be an $(N-3)$ dimensional family of solutions. In the second case, one can show that there is a unique element in the hyper-plane of all solutions which has minimum 2-norm [19]. It is well known that this least-squares solution can be found by calculating the generalised Moore-Penrose pseudo-inverse $U_j^{\#}$. The resulting vectors R_j with minimal Euclidean norm are then given by

$$R_j = U_j^{\#} I = U_j^{\#} \ . \tag{20}$$

Unfortunately, we are not looking for the solution with minimal 2-norm but for the one with minimal Mahalanobis-distance D_m. For this reason, we introduce the Mahalanobis-norm $\| \circ \|_m$ that can be expressed in terms of the traditional Euclidean norm:

$$\|\mathbf{x}\|_m = \|\sqrt{\Lambda}^{-1}\mathbf{x}\|_2 \tag{21}$$

If we are able to calculate the least-squares solution with respect to this Mahalanobis-norm, we have automatically found the solution with minimal Mahalanobis-distance, since

$$\|\mathbf{x}\|_m^2 = \|\sqrt{\Lambda}^{-1}\mathbf{x}\|_2^2 = \left(\sqrt{\Lambda}^{-1}\mathbf{x}\right)^T \left(\sqrt{\Lambda}^{-1}\mathbf{x}\right) = \mathbf{x}^T \Lambda^{-1}\mathbf{x} = D_m \ . \tag{22}$$

By exploiting relation (21), we can map the minimisation of the Mahalanobis-norm to the normal least-squares problem with respect to the Euclidean norm. The required transformation results in a weighting of the columns of U_j by the square root of the corresponding eigenvalues Λ:

$$\min_{R_j} \|R_j\|_m : \quad U_j R_j = I \qquad \longmapsto \qquad \min_{\breve{R}_j} \|\breve{R}_j\|_2 : \quad \left(U_j \sqrt{\Lambda}\right) \breve{R}_j = I \tag{23}$$

$$R_{j_{\min}} = \sqrt{\Lambda} \breve{R}_{j_{\min}} = \sqrt{\Lambda} \left(U_j \sqrt{\Lambda}\right)^{\#} \tag{24}$$

As illustrated in Eq. (24), the minimal m-norm vectors $R_{j_{\min}}$ are then given by the scaled version of the least-squares solution $\breve{R}_{j_{\min}}$ that is uniquely determined by the pseudo-inverse of $(U_j \sqrt{\Lambda})$. By exploiting subsequently the relation $A^{\#} = A^T \left[AA^T\right]^{-1}$ that holds for $m \times n$ matrices A with $(m < n)$ and $\text{rank}(A) = m$ (see [15]), we end up with a fairly familiar result:

$$R_j = \sqrt{\Lambda} \left(U_j \sqrt{\Lambda}\right)^{\#} = \sqrt{\Lambda} \left(\sqrt{\Lambda} U_j^T \left[U_j \Lambda U_j^T\right]^{-1}\right) \tag{25}$$

References

1. Andrew Blake and Michael Isard. *Active Contours*. Springer, first edition, 1998. ISBN 3-54-076217-5.
2. T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *Proceedings 5^{th} European Conference on Computer Vision*, volume 2, pages 484–498. Springer, 1998. ISBN 3-540-64613-2.
3. T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam. The use of active shape models for locating structures in medical images. In H. H. Barrett and A. F. Gmitro, editors, *Information Processing in Medical Imaging*, pages 33–47. Springer, June 1993.
4. T. F. Cootes and C. J. Taylor. Active shape models - "Smart Snakes". In *Proceedings British Machine Vision Conference*, pages 266–275, Leeds, 1992. Springer. ISBN 3-540-19777-X.
5. T. F. Cootes and C. J. Taylor. Statistical models of appearance for computer vision. Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, Manchester M13 9PT, United Kingdom, September 1999. http://www.wiau.man.ac.uk.
6. M. A. Fischler, J. M. Tenenbaum, and H. C. Wolf. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer Graphics and Image Processing*, 15:201–233, 1981.
7. David R. Forsey and Richard H. Bartels. Hierarchical B-spline refinement. In *Computer Graphics Proceedings SIGGRAPH 1988*, pages 205–212. Addison-Wesley, 1988.
8. J. Hug, C. Brechbühler, and G. Székely. Tamed Snake: A Particle System for Robust Semi-automatic Segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI'99*, number 1679 in Lecture Notes in Computer Science, pages 106–115. Springer, September 1999.

9. Aapo Hyvärinen. Independent component analysis by minimization of mutual information. Technical Report A46, Helsinki University of Technology, Department of Computer Science and Engineering, Laboratory of Computer and Information Science, Rakentajanaukio 2 C, FIN-02150 Espoo, Finland, August 1997. http://www.cis.hut.fi/~aapo.
10. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. In *Proceedings* 1^{st} *International Conference on Computer Vision*, pages 259–268, June 1987.
11. András Kelemen. *Elastic Model-Based Segmentation of 2-D and 3-D Neuroradiological Data Sets*. PhD thesis, Swiss Federal Institute of Technology Zürich, Switzerland, 1998. ETH Dissertation No. 12907.
12. András Kelemen, Gábor Székely, and Guido Gerig. Three-dimensional model-based segmentation. In *IEEE International Workshop on Model-Based 3D Image Analysis*, pages 87–96, January 1998.
13. András Kelemen, Gábor Székely, and Guido Gerig. Elastic Model-Based Segmentation of 3-D Neuroradiological Data Sets. *IEEE Transactions on Medical Imaging*, 18(10):828–839, October 1999.
14. F. P. Kuhl and C. R. Giardina. Elliptic Fourier features of a closed contour. *Computer Vision, Graphics, and Image Processing*, 18:236–258, March 1982.
15. Frieder Kuhnert. *Pseudoinverse Matrizen und die Methode der Regularisierung*. Teubner, first edition, 1976.
16. E. H. Moore. On the reciprocal of the general algebraic matrix (abstract). *Bulletin of the American Mathematical Society*, 26:394–395, 1920.
17. E. N. Mortensen, B. Morse, and W. A. Barret. Adaptive boundary detection using live-wire two-dimensional dynamic programming. *Computers in Cardiology*, pages 635–638, October 1992.
18. R. Penrose. A generalized inverse for matrices. *Proceedings Cambridge Philosophical Society*, 51:406–413, 1955.
19. R. Penrose. On best approximate solutions of linear matrix equations. *Proceedings Cambridge Philosophical Society*, 52:17–19, 1956.
20. E. Persoon and K. S. Fu. Shape discrimination using Fourier descriptors. *IEEE Transactions on Systems, Man and Cybernetics*, 7(3):170–179, March 1977.
21. Lawrence H. Staib and James S. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 14(11):1061–1075, November 1992.
22. Gábor Székely, András Kelemen, Christian Brechbühler, and Guido 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.
23. A. Yuille, D. Cohen, and P. Hallinan. Feature extraction from faces using deformable templates. In *Proceedings Conference on Computer Vision and Pattern Recognition*, pages 104–109, 1989.
24. Denis Zorin, Peter Schröder, and Wim Sweldens. Interactive multiresolution mesh editing. In *Computer Graphics Proceedings SIGGRAPH 1997*, pages 259–268. Addison-Wesley, 1997.