Comparison and Application of Selected Statistical Shape Models in Medical Imaging

Anke Neumann[1] and Cristian Lorenz[2]

[1] University of the Federal Armed Forces Hamburg, D-22039 Hamburg, Germany
[2] Philips Research Laboratories Hamburg, Röntgenstr. 24–26
D-22335 Hamburg, Germany

Abstract. This paper reviews several kinds of 2D shape representations by a set of parameters based on labeled points, Fourier descriptors and wavelet descriptors, resp.. Shape models are derived by statistical analysis of parameters corresponding to a set of example shapes. Each model consists of a parameter vector describing mean shape and a set of modes of variation for parameters characterizing shape variability. Seven shape models, some of them differing in parameter normalization, for axial slices of spinal vertebra are compared with respect both to their compactness in parameter space and to their scope in corresponding space of shapes. A model based method for segmenting 2D gray level images is developed by formulating boundary finding as an optimization problem with respect to parameters varying according to the modes of variation. Our method includes an easy and fast interactive improvement of segmentation outcome.

1   Introduction

Segmentation of objects found in medical images, e.g. organs, parts of organs, bones, without or with a small amount of user interaction is an outstanding task in medical image analysis. Most segmentation methods use mainly the image gray values and derived local properties, e.g. contrast, texture, to apply certain local criteria. Those methods most often require a favorable initialization and are sensitive to poor image quality. To improve results additional object knowledge must be incorporated into the segmentation procedure. This can be done by human expert interaction, which is often tedious and time consuming, and by appropriately modeling global object information. The overall object shape represents a powerful property for separating an object and its surrounding. Since global shape is too variable to be adequately represented by a single shape template, statistically based techniques for building shape models, which capture the natural shape variation, were developed ([7], [8], [9], [10]).

Although it is obvious, that the suitability of a shape model depends on its application and on the kind of variability in the shape, we found neither a guidance for selecting a model nor a comparison between several models.

This work aims at representing shape of 2D objects and segmenting 2D objects from 2D images, focusing on spinal vertebra in CT (Computer Tomography) images as an example. The vertebra was chosen because of its medical importance, e.g. in orthopaedy, and of the diversity of shape samples.

2 Three statistical shape models

2.1 Representing shape

Many contour representations have been developed([4], [1]). Taylor et al.[9] represented geometric entities by sets of distinct, labeled points, which we will call key points. Fourier descriptors representing closed contours were used by Staib and Duncan[7] and Székely et al.[8]. Herein, we present a third parameterization based on wavelet decomposition of closed contours.

The use of a set of N labeled contour points (key points) (x_0, y_0), (x_1, y_1), ..., (x_{N-1}, y_{N-1}) yields the shape describing parameter vector

$$p_{PDM} = (x_0, y_0, x_1, y_1, \ldots, x_{N-1}, y_{N-1})^T \quad \in \quad \mathbb{R}^{2N} \quad . \tag{1}$$

Let the closed curve $(x(t), y(t))$, $t \in [0, 2\pi)$, represent the contour of a simply connected object. Using complex notation, $r(t) = x(t) + iy(t)$, gives the following Fourier series expansion restricted to the first N degrees $r_N(t) = \sum_{k=0}^{N-1} c_k e^{-ikt}$, where $c_k \in \mathbb{C}$, $k = 0, 1, \ldots, N-1$, are referred to as Fourier descriptors. Now the shape is parameterized by the vector

$$p_{FSM} = (c_0, c_1, \ldots, c_{N-1})^T \quad \in \quad \mathbb{C}^N \quad . \tag{2}$$

The wavelet transform extracts spatial information in addition to spectral information which is also given by Fourier decomposition. The wavelet parameterization of a contour represented by equally sampled points (x_0, y_0), (x_1, y_1), ..., (x_{N-1}, y_{N-1}) is given by the results of separate wavelet representations of x- and y-coordinates. In this work we restrict ourselves to the often used Daubechies wavelet filter of order 3 ([2],[6]).

2.2 Building the shape model

We are given a training set of n shape examples. Let $p_i \in \mathbb{R}^{2N}$ be the parameterization of the ith example. The mean parameter vector $\bar{p}$ is calculated using $\bar{p} = \frac{1}{n} \sum_{i=1}^{n} p_i$. Since the parameter vector uniquely determines a shape, $\bar{p}$ represents a mean shape. Notice, that both the mean shape and the concentration of overall variability on certain parameter elements depend on the kind of parameterization as illustrated in section 5 (Fig. 2 and Fig. 3). A principal component analysis of the empirical covariance matrix $S = \frac{1}{n} \sum_{i=1}^{n} (p_i - \bar{p}) (p_i - \bar{p})^T$ yields the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_{2N} \geq 0$ and the corresponding eigenvectors q_i, $i = 1, 2, \ldots, 2N$. The eigenvectors corresponding to the largest eigenvalues describe the most significant modes of variation in the parameter space. A large amount of variation is probably explained by a small number of modes, say by the first m eigenvectors. They represent together with the mean parameter vector $\bar{p}$ the shape model. We will denote the shape models as *point distribution model* (PDM), *Fourier shape model* (FSM) and *wavelet shape model* (WSM) depending on whether the shape parameterization is based on coordinates of a set of labeled points, on Fourier descriptors or on wavelet descriptors, resp..

The shape model allows to generate new shape samples by first adding a weighted sum of modes to the mean parameter vector, i.e. $p = \bar{p} + Qb$ with $Q = (q_1, q_2, \ldots, q_m)$ the matrix of the first m modes and $b = (b_1, b_2, \ldots, b_m)$ the weight vector, and than deriving the corresponding shape for p. Any shape parameterized by p can be approximated by calculating the corresponding weight vector using $b = Q^T(p - \bar{p})$. Due to the modal analysis the training set of shapes gives empirical covariances between b_i and b_j, $i \neq j$, which are equal to zero, and empirical variances λ_i of b_i. Hence, varying weight vector b within an ellipsoid described by

$$D^2(b) := \sum_{i=1}^{m} \frac{b_i^2}{\lambda_i} = C \qquad (3)$$

for some constant C produces shapes similar to those found in the training set. $D(b)$ is called Mahalanobis distance of b (from $0 \in \mathbb{R}^m$).

We trained spinal vertebra models on 45 manually segmented example shapes according to distinct slices of a 3D CT data set. We refer to section 5 for illustrations of the resulting shape models.

3   Interactive segmentation using shape models

We formulate segmenting as a minimization problem with respect to the weights for the extracted modes of variation, in order to force the segmentation procedure to explore shapes that 1) fit the image, 2) are consistent with the model, and 3) include of interactively marked points to the segmented contour.

We assume, that we can find some contrast between the object and its background in the image, at least for a large portion of the object boundary. So we apply differential edge detection based on *Roberts cross operator*[4] to the image.

For a given vector of weights $b = (b_1, b_2, \ldots, b_m)$, the parameter vector $p(b) = \bar{p} + Qb$ is computed. $p(b)$ determines a set of points $(x_i, y_i)(b)$, $i = 0, 1, \ldots, N-1$, that represent a shape sample. In the jth, $j \geq 1$, iteration step, point $(\tilde{x}_j, \tilde{y}_j)$ is marked by the user as a point that should belong to the boundary of the shape. To guide the segmentation procedure, we define a cost function $M(b)$ by

$$M(b) = -\sum_{i=0}^{N-1} E((x_i, y_i)(b)) + c \cdot D(b) + d \cdot \sum_{i=1}^{j} \underbrace{e\left((\tilde{x}_i, \tilde{y}_i), \{(x_l, y_l)(b)\}_{l=0,\ldots,N-1}\right)}_{=:e_i}$$

for constants $c, d \in \mathbb{R}^+$, where $E((x, y))$ denotes the result of edge detection at point (x, y), $D(b)$ is defined by (3), and e_i denotes the smallest Euclidean distance between $(\tilde{x}_i, \tilde{y}_i)$ and the points $(x_l, y_l)(b)$, $l = 0, \ldots, N-1$.

Figure 1 shows an example gray level image, the initialization by the mean shape, and the result of the minimization procedure (*downhill simplex method* [6]) with underlying *Fourier shape model*(FSM) with $N = 64$ and using the largest $m = 12$ modes of variation. The automatic segmentation algorithm (j=0) did not find the lower vertebra process in this case, but it is included into the segmented object by a small amount of user interaction.
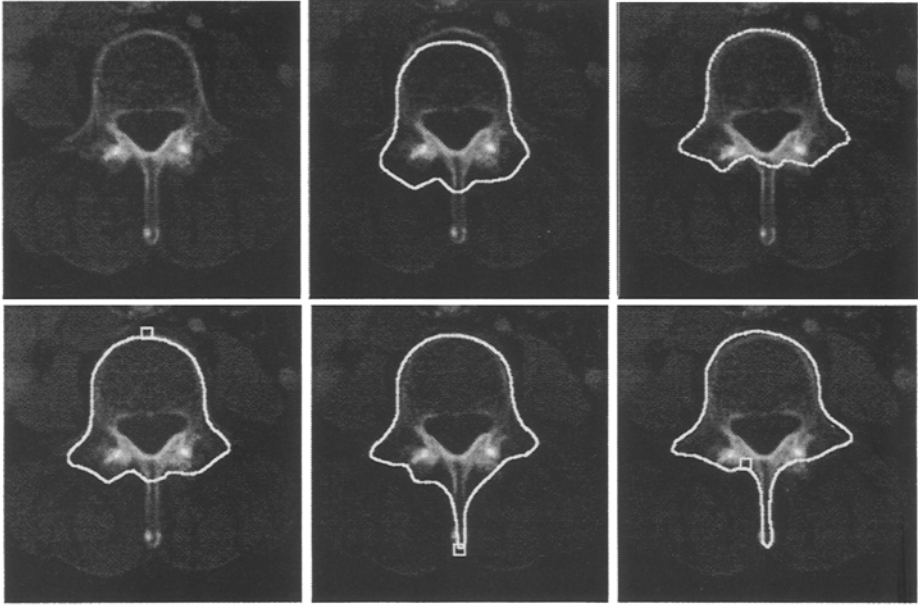
Fig. 1. Segmentation example. The first row shows the gray value image, the initialization, and the result of automatic segmentation, followed by intermediate results of interactive segmentation each indicating the interactively marked point.

4 Invariant shape models

Shape description should be invariant to transformations resulting from change in position, orientation, and scale. Herein, we assume that image format and resolution are standardized. Hence, natural changes in scale are incorporated into the shape model. It turns out, that the performance of the models depends strongly on the kind of normalization being incorporated.

Invariant parameterization based on key points: axial symmetric outlines (PDM1). Since vertebrae in the images of interest are approximately axial symmetric, we determine a coordinate system dependent on the key point set by the center of gravity of key points as origin and the orthonormal eigenvectors of the momentum matrix of key points as basis vectors. A binning procedure is used to select among the two eigenvectors that one which corresponds to the symmetry axis of the outline.

Invariant parameterization based on key points: general outlines (PDM2). In order to achieve invariance with respect to translation and rotation Dryden and Mardia [3] premultiply the $N \times 2$ matrix of the N key points by matrix H, where H is the $N \times N$ Helmert matrix without the first row, and than perform a QR decomposition of the resulting matrix.

Parameterization based on Fourier descriptors: invariance with respect to translation and rotation for axial symmetric outlines (FSM1). The Fourier descriptor c_0 is set equal to zero. The sum

$$c_1 e^{-i\frac{2\pi}{N}k} + c_{N-1} e^{-i\frac{2\pi}{N}(N-1)k} = c_1 e^{-i\frac{2\pi}{N}k} + c_{N-1} e^{i\frac{2\pi}{N}k}$$

with $k = 0, 1, \ldots, N-1$ gives points that lie on an ellipse. We rotate the principal axes of this ellipse into canonical position, where that axis becomes y-axis, which is closest to the symmetry axis of the outline.

Parameterization based on Fourier descriptors: invariance with respect to translation, rotation, starting point, and scaling (FSM2). Pavlidis[5] does the following: c_0 is set equal to zero and all other Fourier descriptors c_i, $i = 1, \ldots, N-1$, are multiplied by a factor $s \cdot e^{i\theta + b}$, where scale parameter s, rotation parameter θ and starting point parameter b are chosen such that c_1 and c_{N-1} become purely imaginary numbers and their sum has modulus 1.

The derived invariant shape models we refer to as PDM1, PDM2, FSM1, and FSM2, resp..

5 Comparing shape models

In order to compare the seven shape models we represent each contour by a fixed number $N = 64$ of equally spaced points, that are key points in the representation (1) for PDM and its normalized versions PDM1 and PDM2, and that are sample points for the parameterization (2) for FSM, its modifications for FSM1 and FSM2, and for shape description by wavelet representation as in WSM.

5.1 Modes of variation

In Figure 2 the influence of the first four modes of variation in the parameter vector to the derived shapes are visualized for each of the seven shape models. The rows correspond to PDM, PDM1, PDM2, FSM, FSM1, FSM2, WSM, resp..The displayed shapes were generated by varying the first four weights for the modes separately, while keeping all other weights zero. The weight values were $-1.5\sqrt{\lambda_i}$ (dark line), 0 and $1.5\sqrt{\lambda_i}$ (bright line), i=1,2,3,4, resp. (1st, 2nd, 3rd, and 4th column, resp.).

Notice, that the invariant parameterizations on the one hand remove translation influence and on the other hand may introduce artificial variability as for PDM1 and FSM1 due to the delicate symmetry axis finding procedure and for FSM2 due to the normalization procedure that selects non-corresponding starting points for Fourier parameterization. This results in artificial scaling expressed by the first mode of variation. The lack of correspondence between the parameter sets of distinct shapes, e.g. due to non-corresponding points selected as starting points in FSM2, deteriorates the statistical model.

5.2 Reconstruction of shape samples

For a fixed number of modes we determined the relative approximation error for each shape in the training set and computed the average value. As expected, using all modes of variation for the approximation allows correct reconstruction of shape samples as shown in Figure 3. Moreover, it turns out, that all shape
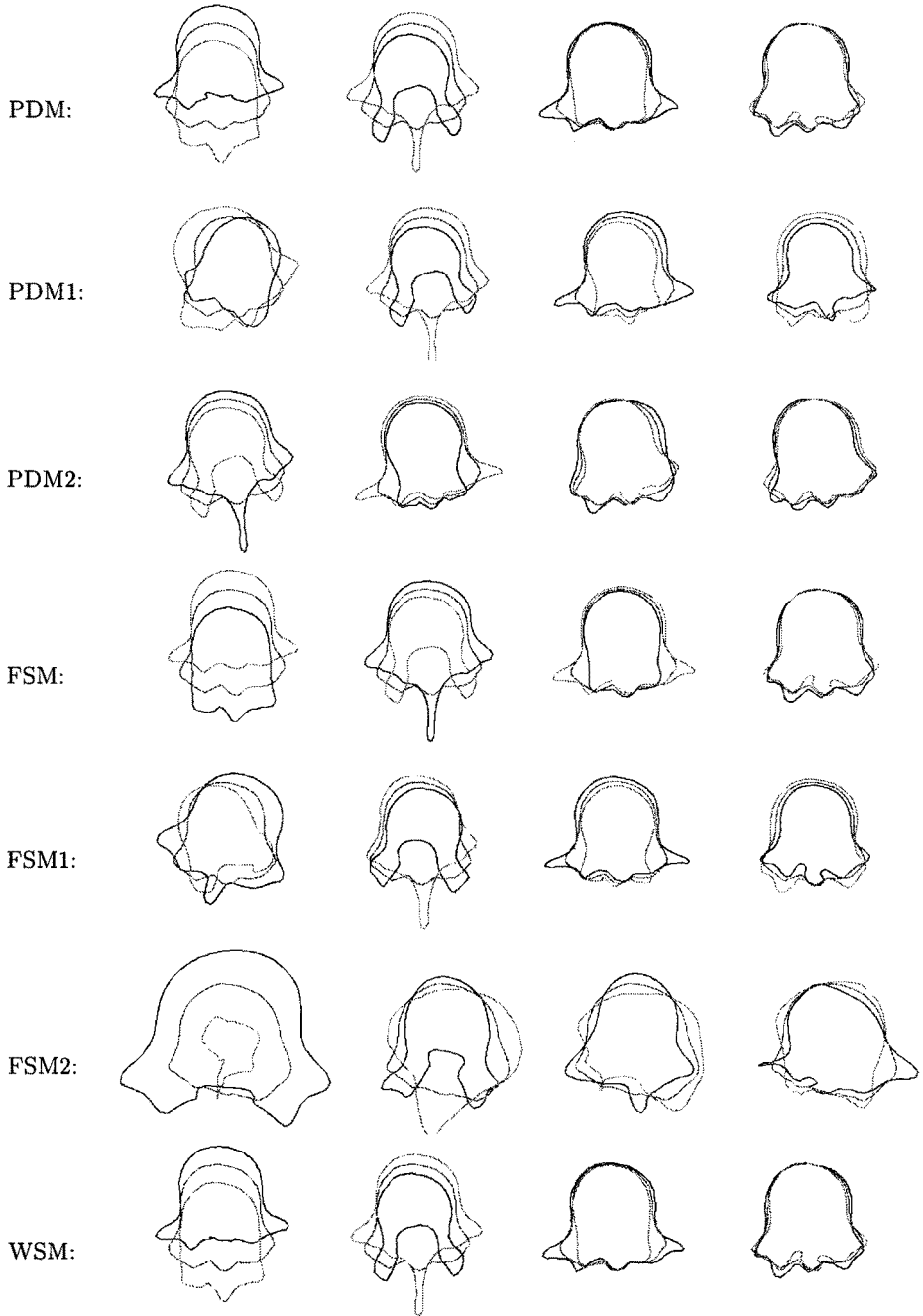
PDM:

PDM1:

PDM2:

FSM:

FSM1:

FSM2:

WSM:

Fig. 2. Shapes generated by varying the first four weights for modes separately.

models, except for FSM2, are similar with respect to reconstruction performance, whereas FSM and WSM are capable to approximate the training shapes with the least number of modes.
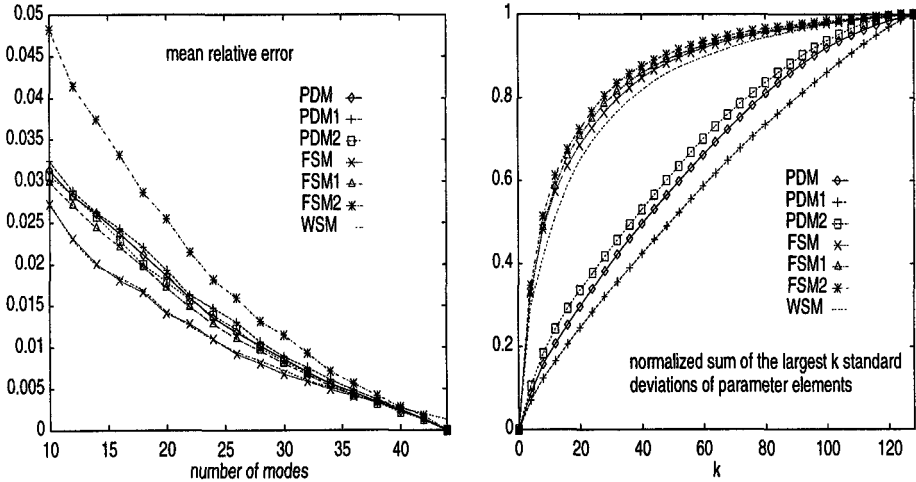

Fig. 3. Comparison of shape models. Left: Mean relative approximation error dependent on a fixed number of modes. Right: Portion of the k parameter elements with largest standard deviations, $k \in \{1, 2, \ldots, 2N\}$, with respect to the overall standard deviation.

5.3 Scope of the models

In order to examine the scope of the shape models, we approximated with all $n - 1$ modes of variation a set of 50 *new shape examples* of vertebra derived from another 3D data set. For each resulting weight vector b we determined the Mahalanobis distances $D(b)$ according to (3). The value $D(b)/\sqrt{n-1}$ can be interpreted as that factor to scale the smallest ellipsoid with, which contains all weights of the *training set*, in order to derive an (variability) ellipsoid that contains the new weight vector b. We obtained the following mean scaling factors and corresponding standard deviations depending on the shape model:

model	PDM	PDM1	PDM2	FSM	FSM1	FSM2	WSM
mean	2.1641	1.8247	1.7468	1.9118	1.7282	1.6403	1.7837
std.dev.	0.5621	0.6194	0.3682	0.3410	0.9155	0.5577	0.3435

A small mean scaling factor combined with a small standard deviation indicates a model that is sufficiently general to capture the whole set of new vertebra shapes. With respect to this point of view the model PDM2 is the best one. PDM1, FSM1, and FSM2 yield very different values for model consistency of the new samples due to the sensitive normalization procedure.

6   Conclusion

Statistical shape models allow a concise and powerful representation of global shape information. In this work seven statistical shape models were built for 2D axial slices of spinal vertebrae based on shape representation by a set of labeled points, Fourier descriptors, or wavelet descriptors. For the first two parameterizations we developed normalization procedures that yield invariance with respect to translation and rotation and are dedicated to symmetric contours. Furthermore, we investigated two normalization methods proposed in the literature([3],[5]). Our experiments showed that the kind of normalization is crucial, since one method resulted in strong artificial shape variations (FSM2), whereas we found also one suitable normalization method (PDM2). The shape model based on the latter normalized shape parameterization also turned out to be most capable to capture new shape samples.

We applied exemplary the shape model based on Fourier descriptors to a segmentation procedure that makes iterative interactive improvement of the result possible by balancing between the image fit, model consistency and fit to the interactively marked points. The procedure was found to be relatively insensitive with respect to blurred boundaries and spurious edges. Though we got promising segmentation results, we still need to spend effort into an improved minimization procedure.

Future work will include incorporating non-linear dependencies between the parameters by using higher order moments or appropriate probability distributions and the investigation of shape models for 3D objects.

References

1. P.J. Besl. *Geometric Modeling and Computer Vision.* Proc.IEEE 76(8), Aug. 1988, pp. 936-958
2. C.K. Chui. *An Introduction to Wavelets.* Academic Press, San Diego, 1992
3. I.L. Dryden, K.V. Mardia. *Multivariate Shape Analysis.* Sankhyā, vol. 55, series A, pt. 3, 1993, pp. 460-480
4. R.O. Duda, P.E. Hart. *Pattern Classification and Scene Analysis.* John Wiley & Sons, New York, 1973
5. T. Pavlidis. *Structural Pattern Recognition.* Springer-Verlag, 1977
6. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Programming.* Cambridge University Press, 1992
7. L.H. Staib, J.S. Duncan. *Boundary Finding with Parametrically Deformable Models.* IEEE PAMI 14(11), 1992, pp. 1061-1075
8. G. Székely, A. Kelemen, Ch. Brechbühler, G. Gerig. *Segmentation of 3D Objects from MRI Volume Data Using Constrained Elastic Deformations of Flexible Fourier Surface Models.* Proc. Computer Vision, Virtual Reality and Robotics in Medicine, Nice, April 1995, pp. 495-505
9. C.J. Taylor, T.F. Cootes, A. Hill, J. Haslam. *Medical Image Segmentation Using Active Shape Models.* in Medical Imaging, L. Beolchi, M.H. Kuhn (eds.), IOS Press, Amsterdam, 1994, pp. 121-144
10. B.C. Vemuri, A. Radisavljevic. *Multiresolution Stochastic Hybrid Shape Models with Fractal Priors.* ACM Trans. Graphics, 13(2), April 1994, pp. 177-207