Shape Analysis
Using a Point-Based Statistical Shape Model
Built on Correspondence Probabilities

Heike Hufnagel[1,2], Xavier Pennec[1], Jan Ehrhardt[2], Heinz Handels[2], and
Nicholas Ayache[1]

[1] Asclepios Project, INRIA, Sophia Antipolis, France
`heike.hufnagel@sophia.inria.fr`,
[2] Medical Informatics, Universitätsklinikum Hamburg-Eppendorf, Germany

Abstract. A fundamental problem when computing statistical shape
models is the determination of correspondences between the instances of
the associated data set. Often, homologies between points that represent
the surfaces are assumed which might lead to imprecise mean shape and
variability results. We propose an approach where exact correspondences
are replaced by evolving correspondence probabilities. These are the ba-
sis for a novel algorithm that computes a generative statistical shape
model. We developed an unified MAP framework to compute the model
parameters ('mean shape' and 'modes of variation') and the nuisance
parameters which leads to an optimal adaption of the model to the set
of observations. The registration of the model on the instances is solved
using the Expectation Maximization - Iterative Closest Point algorithm
which is based on probabilistic correspondences and proved to be robust
and fast. The alternated optimization of the MAP explanation with re-
spect to the observation and the generative model parameters leads to
very efficient and closed-form solutions for (almost) all parameters. Ex-
perimental results on brain structure data sets demonstrate the efficiency
and well-posedness of the approach. The algorithm is then extended to
an automatic classification method using the k-means clustering and ap-
plied to synthetic data as well as brain structure classification problems.

1 Introduction

One of the central difficulties of analyzing different organ shapes in a statis-
tical manner is the identification of correspondences between the shapes. As
the manual identification of landmarks is not a feasible option in 3D, several
preprocessing techniques were developed to automatically find exact one-to-one
correspondences [1, 2] between surfaces. Some approaches solve this with a search
for the registration transformation using an atlas [3] or the ICP algorithm [4].
Other methods directly combine the search of correspondences and of the sta-
tistical shape model (SSM) [5–7]. However, exact correspondences can only be
determined between continuous surfaces, not between point cloud representa-
tions of surfaces. Thus, using imprecise homologies leads to variability modes
that not only represent the organ shape variations but also artificial variations

whose importance is linked to the local sampling. The SoftAssign algorithm tries to solve this problem with some kind of probabilistic formulation [8]. Another recent approach proposes an entropy based criterion to find shape correspondences, but requires implicit surface representations [9]. Other recent methods combine the shape analysis with the search for correspondences, however, these methods are not easily adaptable to multiple observations of unstructured point sets [10–12] or focus only on the mean shape [13]. In order to build an SSM based on inexact correspondences between point clouds, we pursue a probabilistic concept and base our work on a EM-ICP registration algorithm which proved to be robust, precise, and fast [14]. In section 2, we realize a Maximum a Posteriori (MAP) estimation of the model and observation parameters which lead to a unique criterion. We then compute the mean shape and eigenmodes which best fit the given data set by optimizing the global criterion iteratively with respect to all model and observation parameters. A key part of our method is that we can find a closed-form solution for almost each of the parameters. In particular, the approach solves for the mean shape and the eigenmodes *without* the need of one-to-one correspondences as is usually required by the PCA. Experiments in section 3 demonstrate that the resulting deformation coefficients can be used as an efficient measure to classify each observation.

2 Construction of the Statistical Shape Model

2.1 Model and Observation Parameters

In the process of computing the SSM, we distinguish strictly between **model parameters** and **observation parameters**. The generative SSM is explicitly defined by 4 **model parameters**:

- mean shape $\bar{M} \in \mathbb{R}^{3N_m}$ parametrized by N_m points $m_j \in \mathbb{R}^3$,
- eigenmodes v_p consisting of N_m 3D vectors v_{pj},
- associated standard deviations λ_p which describe - similar to the classical eigenvalues in the PCA - the impact of the eigenmodes,
- number n of eigenmodes.

Using the generative model $\Theta = \{\bar{M}, v_p, \lambda_p, n\}$ of a given structure, the shape variations of that structure can be generated by $M_k = \bar{M} + \sum_{p=1}^{n} \omega_{kp} v_p$ with $\omega_{kp} \in \mathbb{R}$ being the deformation coefficients. The shape variations along the modes follow a Gaussian probability with variance λ_p:

$$p(M_k|\Theta) = p(\Omega_k|\Theta) = \prod_{p=1}^{n} p(\omega_{kp}|\Theta) = \frac{1}{(2\pi)^{n/2} \prod_{p=1}^{n} \lambda_p} \exp\left(-\sum_{p=1}^{n} \frac{\omega_{kp}^2}{2\lambda_p^2}\right). \quad (1)$$

In order to account for the unknown position and orientation of the model in space, we introduce the random (uniform) rigid or affine transformation T_k. A model point m_j can then be deformed and placed by $T_k \star m_{kj} = T_k * (\bar{m}_j + \sum_p \omega_{kp} v_p)$. Finally, we specify the sampling of the model surface: Each sampling (e.g. observation) point s_{ki} is modeled as a Gaussian measurement of a (transformed) model point m_{kj}. The probability of the observation $p(s_{ki}|m_{kj}, T_k)$ knowing the originating model point m_{kj} is given by

$p(s_{ki}|m_{kj}, T_k) = (2\pi)^{-3/2}\sigma^{-1}\exp(-\frac{1}{2\sigma^2}(s_{ki} - T_k \star m_{kj})^T.(s_{ki} - T_k \star m_{kj})$. As we do not know the originating model point for each s_{ki}, the probability of a given observation point s_{ki} is described by a Mixture of Gaussians and the probability for the whole scene S_k becomes:

$$p(S_k|M, T_k) = \prod_{i=1}^{N_k} \frac{1}{N_m} \sum_{j=1}^{N_m} p(s_{ki}|m_{kj}, T_k). \tag{2}$$

We summarize the **observation parameters** as $Q_k = \{\Omega_k, T_k\}$. Notice that the correspondences are hidden parameters that do not belong to the observation parameters of interest.

2.2 Derivation of the Global Criterion Using a MAP Approach

When building the SSM, we deal with the inverse problem of the approach in section 2.1: We have N observations $S_k \in \mathbb{R}^{3N_k}$, and we are interested in the parameters linked to the observations $Q = \{Q_k\}$ as well as the unknown model parameters Θ. In order to determine all parameters of interest, we optimize a MAP on Q and Θ rather than an ML to take into account that Q and Θ are not independent.

$$\text{MAP} = -\sum_{k=1}^{N} \log(p(Q_k, \Theta|S_k)) = -\sum_{k=1}^{N} \log\left(\frac{p(S_k|Q_k, \Theta)p(Q_k|\Theta)p(\Theta)}{p(S_k)}\right). \tag{3}$$

As $p(S_k)$ does not depend on Θ and $p(\Theta)$ is assumed to be uniform, the global criterion integrating our unified framework is the following:

$$C(Q, \Theta) = -\sum_{k=1}^{N} \left(\log(p(S_k|Q_k, \Theta)) + \log(p(Q_k|\Theta))\right). \tag{4}$$

The first term describes the ML criterion (2) whereas the second term is the prior on the deformation coefficients ω_{kp} as described in (1). Dropping the constants, our criterion simplifies to $C(Q, \Theta) \sim \sum_{k=1}^{N} C_k(Q_k, \Theta)$ with

$$C_k(Q_k, \Theta) = \sum_{p=1}^{n} \left(\log(\lambda_p) + \frac{\omega_{kp}^2}{2\lambda_p^2}\right) - \sum_{i=1}^{N_k} \log\left(\sum_{j=1}^{N_m} \exp\left(-\frac{\|s_{ki} - T_k \star m_{kj}\|^2}{2\sigma^2}\right)\right) \tag{5}$$

This equation is the heart of the unified framework for the model computation and its fitting to observations. By optimizing it alternately with respect to the operands in $\{Q, \Theta\}$, we are able to determine all parameters we are interested in. Starting from the initial model parameters Θ, we fit the model to each of the observations (section 2.3). Next, we fix the observation parameters Q_k and update the model parameters (section 2.4). Some terms will recur in the different optimizations, so we introduce the following notation for the derivation of the second term $\xi_{kij}(T_k, \Omega_k, \bar{M}, v_p, \lambda_p) = \log\sum_{j=1}^{N_m} \exp\left(\frac{-\|s_{ki} - T_k \star m_{kj}\|^2}{2\sigma^2}\right)$ with respect to one of the function's parameters (let's say x):

$$\frac{\partial \xi}{\partial x} = \sum_{j=1}^{N_m} \gamma_{kij} \frac{(s_{ki} - T_k \star m_{kj})^T}{\sigma^2} \frac{\partial(s_{ki} - T_k \star m_{kj})}{\partial x} \tag{6}$$

where the weights $\gamma_{ijk} = \exp\left(-\frac{\|s_{ki} - T_k \star m_{kj}\|^2}{2\sigma^2}\right)\left[\sum_{l=1}^{N_m} \exp\left(-\frac{\|s_{ki} - T_k \star m_{kl}\|^2}{2\sigma^2}\right)\right]^{-1}$ are sometimes interpreted as soft labels/correspondences.

2.3 Mapping the Model to the Observations

Optimization with respect to the Transformations As no closed form solution exists for the optimization of criterion (2), we employ an EM algorithm where the correspondence probabilities between S_k and M are modeled as the hidden variable $H_k \in \mathbb{R}^{N_k \times N_m}$. An instance point s_{ki} corresponds to a model point m_j with probability $E(H_{k_{ij}})$. By computing the expectation of the log-likelihood of the complete data distribution with T_k fixed, we find in the *expectation step* $E(H_{k_{ij}}) = \gamma_{kij}$. As defined above, the γ_{kij} represent the weights of each pair (s_{ki}, m_j) in the criterion. Next, $T_k = \{A_k, t_k\}$ is computed in the *maximization step* by maximizing the global criterion in (5) with all γ_{kij} fixed in a closed-form solution. The implementation of the EM-ICP algorithm is realized in a multi-scaling frame regarding the variance [14]. $\sigma_{initial}$ and its decrease rate have to be carefully adapted to the data at hand (σ_{final} should be in the order of the average point distance).

Optimization with respect to the Deformation Coefficients The observation parameter T_k and Θ are fixed, and we compute the ω_{kp} which solve $\partial C_k(Q_k, \Theta)/\partial \omega_{kp} = 0$. This leads to a matrix equation of the form $\Omega_k = (B_k - \sigma^2 \Lambda_{nn})^{-1} d_k$ with $d_{kp} = \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} \gamma_{kij}(s_{ki} - t_k - A_k \bar{m}_j)^T A_k v_{pj}, \; d_{kp} \in \mathbb{R}$ and $b_{kqp} = \sum_{i=1}^{N_k} \sum_{j=1}^{N_m} \gamma_{kij} v_{qj}^T A_k^T A_k v_{pj}, \; b_{kqp} \in \mathbb{R} \;\; b_{kqp} = b_{kpq}$.

2.4 Learning the Model from the Observations

Optimization with respect to the Standard Deviations The computation of the optimal standard deviation λ_p with parameters $\bar{M}, v_p$ and Q_k fixed is simply:

$$\frac{\partial C(Q, \Theta)}{\partial \lambda_p} = \sum_{k=1}^{N} \left(\frac{1}{\lambda_p} - \frac{\omega_{kp}^2}{\lambda_p^3} \right) = 0 \quad \Leftrightarrow \quad \lambda_p^2 = \frac{1}{N} \sum_{k=1}^{N} \omega_{kp}^2. \tag{7}$$

Optimization with respect to the Mean Shape Setting $\partial C(Q, \Theta)/\partial \bar{m}_j$ to 0 and using the general derivation presented in (6), we find

$$\bar{m}_j = \left(\sum_{k=1}^{N} \sum_{i=1}^{N_k} \gamma_{kij} A_k^T A_k \right)^{-1} \sum_{k=1}^{N} \sum_{i=1}^{N_k} \gamma_{kij} A_k^T (s_{ki} - t_k - A_k \sum_{p=1}^{n} \omega_{kp} v_{pj}) \tag{8}$$

Optimization with respect to the Eigenmodes (The parameters $\lambda_p, \bar{M}$ and Q_k are fixed.) Let us first define the matrix $V \in \mathbb{R}^{3N_m \times n}$ containing the eigenmodes $v_p \in \mathbb{R}^{3N_m}$ in its columns. The $v_{pj} \in \mathbb{R}^3$ referred to in the equations are the eigenmode information associated to point $\bar{m}_j$. As we want the eigenmodes to be orthonormal, we add a Lagrange multiplier by introducing the symmetric matrix $Z \in \mathbb{R}^{n \times n}$ to our global criterion in the form: $\Lambda = C + \frac{1}{2} tr \left(Z(V^T V - I_{n \times n}) \right)$. Deriving the Lagrangian with respect to v_{pj} gives in the rigid case

$$\frac{\partial \Lambda}{\partial v_{pj}} = \sum_{q=1}^{n} z_{qp} v_{qj} - \sum_{q=1}^{n} b_{pq} v_{qj} + q_{pj}$$

where $q_{pj} = \frac{1}{\sigma^2} \sum_{k=1}^{N} \sum_{i=1}^{N_k} \gamma_{kij}(s_{ki} - t_k - A_k \bar{m}_j)^T \omega_{kp} A_k, \;\; q_{pj} \in \mathbb{R}^3$ and $b_{pqj} = \frac{1}{\sigma^2} \sum_{k=1}^{N} \sum_{i=1}^{N_k} \gamma_{kij} \omega_{kq} \omega_{kp} I_{3 \times 3} \;\; b_{pqj} \in \mathbb{R}^{3 \times 3}$.

Hence we find $\sum_{q=1}^{n} \boldsymbol{v}_{jq}(z_{qp} + b_{pqj}) = \boldsymbol{q}_{jp}$. We approach the problem regarding each of the N_m bands $[V]_{\{j\}} \in \mathbb{R}^{3 \times n}$ of matrix $V \in \mathbb{R}^{3N_m \times n}$ separately with $[V]_{\{j\}} = [\boldsymbol{v}_{j1}, ..., \boldsymbol{v}_{jq}, ..., \boldsymbol{v}_{jn}]$ and $[V]_{\{j\}} (B_j + Z) = [Q]_{\{j\}}$. We iterate the following two steps until $\|V^{t+1} - V^t\|^2 \leq \epsilon$.

1. For Z known, we compute V: $[V]_{\{j\}} = [Q]_{\{j\}} (B_j + Z)^{-1}$ for all model point indices j. To enforce V to be orthonormal, we apply first a singular value decomposition $V = USR^T$ and then replace V by UR^T.
2. For all $[V]_{\{j\}}$ known, we determine Z: $Z = V^T \tilde{Q}$ with $[\tilde{Q}]_{\{j\}} = [Q]_{\{j\}} - [V]_{\{j\}} B_j$. As Z has to be symmetric, we set $Z \leftarrow \frac{1}{2}(Z + Z^T)$.

3 Experiments and Results

3.1 Validation of Algorithm on Synthetic Data

We generated a data set consisting of ellipsoids with and without bump. Each ellipsoid was transformed using a random affine transformation, and then an uniform noise was added, see Fig. 1a) for some observation examples.

Building the SSM on Ellipsoids The data set contained 18 ellipsoids, half with bump, half without. An initial mean shape was randomly chosen from the data set, Fig. 1d). The results of the alignment as seen in Fig. 1b) were obtained with the EM-ICP registration. The final mean shape and the deformations according to the first eigenmode are depicted in Fig. 1d). As can be seen in Fig. 1c), all observation shapes can be generated using the resulting SSM and the deformation coefficients as $S_k = \bar{M} + \sum_{kp} \omega_{kp} \boldsymbol{v}_p$. Figure 1e) shows the converging values of the global criterion (5) during the iterations of the SSM computation. Since we discard eigenvectors whose standard deviation falls below a certain threshold, n diminishes from 10 to 7 during computation. The results show that the algorithm computes a representative SSM for a given data set.

Classification of Ellipsoids The deformation coefficients Ω computed during the optimization of (5) serve as a classification measure regarding the shape of the observations S_k. We formed feature vectors $\omega_k = (\omega_{k1}, \omega_{k2}, ..., \omega_{kn})$ and used them as input for a k-means clustering. The resulting two classes coincide with the 'bump' and 'without bump' classes with an average Rand index [15] of 0.95. See Fig. 1f) for the values of the 2D feature vectors $(\omega_{k1}, \omega_{k2})$.

3.2 Shape Analysis of Brain Structure

Here, we focus on the putamen. The data set consists of $N = 21$ right and left segmented instances (approximately $20mm \times 20mm \times 40mm$) which are represented by min 994 and max 1673 points, see Fig. 2a),b) for some shape examples. The MR images ($255 \times 255 \times 105$ voxels of size $0.94mm \times 0.94mm \times 1.50mm$) as well as the segmentations were kindly provided by the Hôpital La Pitié-Salpêtrière, Paris, France. The data was collected in the framework of a study on hand dystonia. We chose the following parameters as input: Number of eigenmodes $n = 20$, initial sigma in the EM-ICP $\sigma = 4mm$, EM-ICP iterations
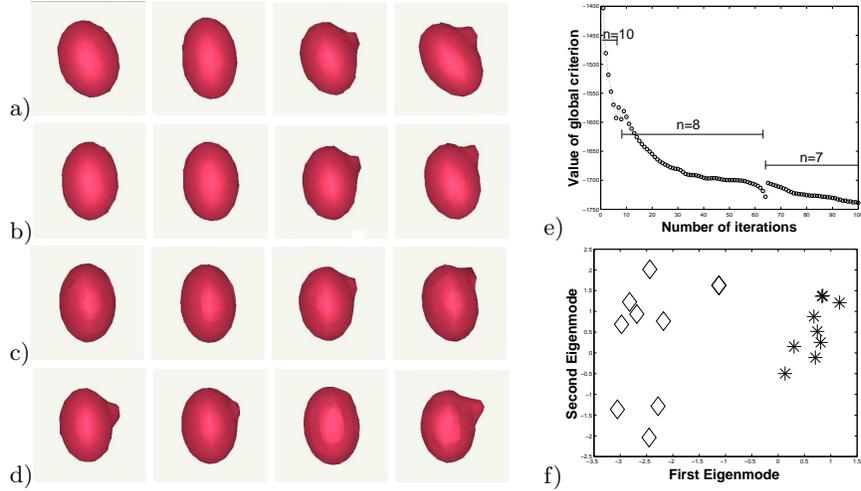
Fig. 1. *SSM on synthetic data set. Row a) shows 4 observation examples. Row b) shows the same observation after being aligned to the mean shape. Row c) shows the shapes generated by using the SSM and the deformation coefficients associated with the observations. Row d) shows (from left to right) the initial mean shape (a randomly chosen observation), the final mean shape, and the mean shape deformed with respect to the first eigenmode ($\bar{M} - 3\lambda_1 \boldsymbol{v}_1$ and $\bar{M} + 3\lambda_1 \boldsymbol{v}_1$). e): Values of global criterion after each iteration. f): 2D deformation coefficient feature vectors (ω_{k1}, ω_{k2}) for the first two eigenmodes, 'with bump' observations (diamonds) and 'without bump' (stars).*
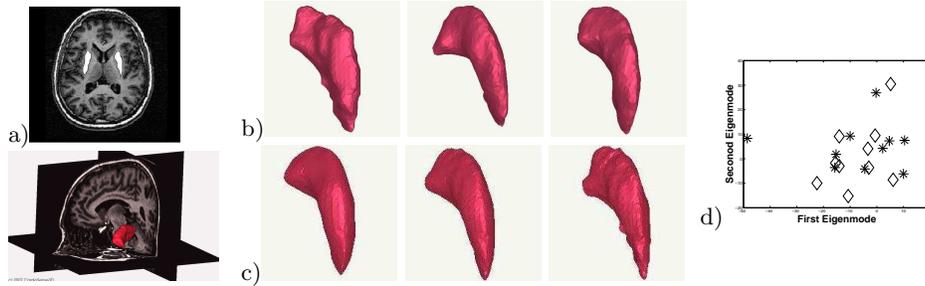

Fig. 2. *Shape analysis of the putamen. a) CT-images with segmented left putamen. b) Observation examples of the data set. c) Mean shape (middle) and its deformations according to the first eigenmode ($\bar{M} - 3\lambda_1 \boldsymbol{v}_1$ and $\bar{M} + 3\lambda_1 \boldsymbol{v}_1$). d) 2D deformation coefficient feature vectors (ω_{k1}, ω_{k2}) for the first two eigenmodes, 'control' observations are represented as diamonds and 'patient' as stars.*

4, variance multi-scaling factor of the EM-ICP 0.7. The computation of the SSM converges after 30 iterations. The resulting smooth mean shape and the deformations according to the first two eigenmodes are shown Fig. 2c). In order to analyse the shapes, again we formed feature vectors $\omega_k = (\omega_{k1}, \omega_{k2}, ..., \omega_{kn})$ and used them as input for a k-means clustering. In this case, no two distinct shape classes were found (Fig.2d)).

3.3 Practical Aspects

Initial Model Parameters As the computation of the observation parameters is based on known model parameters $\Theta = \{\bar{M}, v_p, \lambda_p\}$, we initialize $\bar{M}$ with one of the observations S_k in the given data set, preferably with a typical shape. Next, by applying the EM-ICP registration, we evaluate the resulting correspondence probabilities between $\bar{M}$ and each S_k and determine "virtual" one-to-one correspondences. These are then used as input for the Principal Components Analysis to compute the initial eigenvectors v_p and the initial eigenvalues λ_p. In order to test for the sensibility of our SSM computation with respect to the initialization, we compared the mean shape results which are obtained when using dissimilar initial mean shapes M_1 and M_2 (e.g.the first two shapes in Fig. 2b)). We established that M_1 can be generated based on the SSM found with M_2 with statistically very small deformation coefficients ω_{1p}: $M_1 = M_2 + \sum_p \omega_{1p} \boldsymbol{v}_p$ with e.g. $\omega_{11} = 3.8 << \lambda_{21} = 15.7$.

Model Selection As we want to find a good balance between complexity and simplicity of the model, we reduce the dimension of the eigenvector space during the iterated computation of the parameters. If the standard deviation λ_p becomes "too small", the associated eigenmodes v_p are no longer taken into account. Finally, we want to add the Bayesian Information Criterion [16] to our global criterion with $BIC(n, N, N_k) = C(Q, \Theta) + \frac{n}{2} \log \sum_{k=1}^{N} N_k$. The Bayesian selection approach rates the goodness of a model based on the probability it assigns to the observed data while preferring a more constrained model than the Akaike Information Criterion. This suits our needs as we assume that several of the eigenmodes only represent noise variations.

4 Discussion

We developed a novel algorithm to generate statistical shape models (SSMs) which does not need one-to-one point correspondences but relies solely on point correspondence probabilities for the computation of mean shape and eigenmodes. Therefore, elaborate preprocessing of the observations in the data set to establish correspondences becomes obsolete, no questionable correspondences between point clouds representing surfaces are assumed, and the number of points in the observation may vary. The approach can be used for non-spherical surfaces and can be adapted to applications on data sets with different topologies as the connectivity between points does not play a role. We developed a mathematically sound and unified framework for the computation of model parameters and observation parameters and succeeded in determining a closed form solution for optimizing the associated criterion alternately for all parameters. Experiments showed that our algorithm works well and leads to plausible results. It proved to be robust to different initial mean shape choices and is stable even for a small number of observations. The explicit computation of all parameters involved allows a in-depth analysis of the data set. By evaluating the standard deviation and associated deformation coefficients for each eigenmode and each observation,

a direct automatic classification of the data set is possible as we showed for the synthetic data set. We then performed a shape analysis on a putamen data set and found no statistically significant shape differences between dystonia patients and control group after affine normalizations (which confirms the presumption of the concerned physicians). From a theoretical point of view, a very powerful feature of our method is that we are optimizing a unique criterion. Thus, the convergence is ensured. However, the practical convergence rate has to be investigated more carefully. For instance, a fast decrease of the multi-scale variance σ^2 easily freezes the model in local minima. For further validation, we intend to study other kinds of data (e.g. hippocampus or ganglion) whose shapes are less convex than the putamen. In order to ensure robustness, we will extend the distance measure in the EM-ICP to include the normals.

References

1. Lorenz, C., Krahnstoever, N.: Generation of point-based 3D statistical shape models for anatomical objects. CVIU **77**(2) (2000) 175–191
2. Styner, M., Gerig, G., Lieberman, J., Jones, D., Weinberger, D.: Statistical shape analysis of neuroanatomical structures based on medial models. MedIA (03)
3. Shelton, C.R.: Morphable surface models. IJCV **38(1)** (2000) 75–91
4. Besl, P.J., McKay, N.D.: A method for registration of 3D shapes. IEEE Trans. PAMI (1992) 239–256
5. Davies, R., Twining, C., Cootes, T.: A minimum description length approach to statistical shape modeling. IEEE Trans. Med. Imag. **21(5)** (2002)
6. Heimann, T., Wolf, I., Williams, T., Meinzer, H.: 3D active shape models using gradient descent optimization of description length. In: IPMI. Volume 3565. (2005)
7. Zhao, Z., Theo, E.K.: A novel framework for automated 3D PDM construction using deformable models. In: Medical Imaging 2005, SPIE Proc. (2005)
8. Rangarajan, A., Chui, H., Bookstein, F.L.: The softassign procrustes matching algorithm. In: IPMI. (1997) 29–42
9. Cates, J., Meyer, M., Fletcher, P., Whitaker, R.: Entropy-based particle systems for shape correspondences. In: MICCAI'06. (2006)
10. Tsai, A., Wells, W.M., Warfield, S.K., Willsky, A.S.: An EM algorithm for shape classification based on level sets. MedIA **9** (2005) 491–502
11. Peter, A., Rangarajan, A.: Shape analysis using the Fisher-Rao Riemannian metric: Unifying shape representation and deformation. In: IEEE ISBI. (2006)
12. Kodipaka, S., Vemuri, B., Rangarajan, A., Leonard, C., Schmallfuss, I., Eisenschenk, S.: Kernel fisher discriminant for shape-based classification in epilepsy. Medical Image Analysis **11** (2007)
13. Rangarajan, H.C.A., Zhang, J., Leonard, C.: Unsupervised learning of an atlas from unlabeld point-sets. IEEE Transactions on PAMI **26**(2) (2004)
14. Granger, S., Pennec, X.: Multi-scale EM-ICP: A fast and robust approach for surface registration. In: ECCV 2002. LNCS 2353 (2002) 418–432
15. Rand, W.M.: Objective criteria for the evaluation of clustering methods. American Statistical. Association **66** (1971)
16. Schwarz, G.: Estimating the dimension of a model. Ann. of Stat. **6**(2) (1978)