A Statistical Model for Smooth Shapes in Kendall Shape Space*

Akshay V. Gaikwad, Saurabh J. Shigwan, and Suyash P. Awate

Computer Science and Engineering Department,
Indian Institute of Technology (IIT) Bombay
suyash@cse.iitb.ac.in

Abstract. This paper proposes a novel framework for learning a statistical shape model from image data, automatically without manual annotations. The framework proposes a *generative model* for image data of individuals within a group, relying on a model of group shape variability. The framework represents shape as an equivalence class of pointsets and models group shape variability in Kendall shape space. The proposed model captures a novel shape-covariance structure that incorporates *shape smoothness*, relying on Markov regularization. Moreover, the framework employs a novel model for data likelihood, which lends itself to an inference algorithm of low complexity. The framework infers the model via a novel expectation maximization algorithm that *samples* smooth shapes in the Riemannian space. Furthermore, the inference algorithm normalizes the data (via similarity transforms) by optimal *alignment* to (sampled) individual shapes. Results on simulated and clinical data show that the proposed framework learns better-fitting compact statistical models as compared to the state of the art.

Keywords: Statistical shape model, Kendall shape space, geodesic distance, Markov model, shape sampling, shape alignment, image data normalization.

1 Introduction and Related Work

The typical notion of object shape [4,13] is an equivalence class of object boundaries / silhouettes, where the equivalence relation is given by a similarity transformation. These geometric invariance properties make shape space non-Euclidean. Following Kendall [8], representing shapes via pointsets with known correspondences, we model (i) preshape space as a subset of Euclidean space, i.e., an intersection of the unit hypersphere (for scale invariance) with a hyperplane through the origin (for translation invariance) and (ii) shape space with an additional rotational-invariance structure. While typical pointset-based shape models ignore this structure, the proposed method adapts shape modeling and inference to this Riemannian structure.

Unlike typical pointset-based shape models, the proposed model incorporates prior information that real-world objects have *smooth* boundaries. Such information is useful during (i) model learning: as regularization to counter the noise in image data (even errors in manual landmark placement) to produce more compact models; and (ii) model

* Thanks to the Royal Academy of Engineering 1314RECI076 and IIT Bombay 14IRCCSG010.

N. Navab et al. (Eds.): MICCAI 2015, Part III, LNCS 9351, pp. 628–635, 2015.
DOI: 10.1007/978-3-319-24574-4_75

application (e.g., segmentation with shape priors): to more effectively regularize the learned covariance, having several near-zero eigenvalues, to enforce smoother shapes.

Many early methods for statistical shape analysis rely on manually defined landmarks on image data [3] or templates [11]. Later methods optimize point positions relying on model covariance or information-theoretic criteria [2,4], but do *not* propose generative models. Medial representations [13] lead to sophisticated Riemannian statistical shape modeling approaches [6], but are limited to non-branching objects.

Some advanced approaches capture shape variability through nonlinear diffeomorphic warps of object boundaries [1,5,7,14]. Unlike our approach, such approaches entail joint statistical analysis in the very large dimensional Riemannian spaces of nonlinear diffeomorphisms coupled with the aligned boundaries. Some such approaches [5] enforce shape smoothness via current distance [14], which is a metric, but is computationally expensive as compared to the shape dissimilarity in our approach. Others [1] models object boundaries as missing data by treating deformations as hidden variables, somewhat analogous to our approach in Kendall shape space.

We propose a novel generative statistical model that relies on Kendall shape space and pointset-based shape representation. It introduces a novel prior on shape smoothness. It treats each individual shape pointset as a random variable, thereby modeling its distribution that can lead to more robust inference via expectation maximization (EM). The proposed EM relies on a novel algorithm for sampling smooth shapes.

2 Methods

We present a shape model in Kendall shape space including a Markov prior for shape smoothness, a computationally-efficient likelihood model for the observed image data given individual shapes, and EM inference involving sampling smooth shapes.

2.1 Statistical Shape Model Using Riemannian Distances and Smoothness Priors

We model each shape as an equivalence class of pointsets, where equivalence is defined via translation, rotation, and isotropic scaling. During model learning, given the correspondences between points in the mean shape and points in each individual shape, we infer point locations within each shape. The known correspondences allow us to order the points in each pointset and thereby represent each pointset as a vector.

Consider I individuals labeled $i = 1, \cdots, I$ with 3-dimensional anatomical shapes represented using pointsets of cardinality N. Let the *hidden* random vector Z_i model the shape for individual i, where an instance $z_i \in \mathbb{R}^D, D = 3N$. Let the *parameters* $\{\mu, C, \beta\}$ model the probability density function (PDF) of individual *shape pointsets*, i.e., $P(Z_i|\mu, C, \beta)$, where $\mu \in \mathbb{R}^D$ represents the group *mean*, $C \in \mathbb{R}^{D\times D}$ represents the unregularized group *covariance*, and $\beta \in \mathbb{R}^+$ represents the *smoothness* of the spatial variation of point locations in each shape. Without loss of generality, we constrain the mean shape μ and the individual shapes z_i to (i) have the centroid at the coordinate-space origin and (ii) be of unit norm. Thus, μ and $\{z_i\}_{i=1}^I$ lie in preshape space.

We propose to model a PDF on preshape space to capture the covariance structures of population shape variability and shape smoothness, by relying on the (i) approximate Normal law [10] that models a Gaussian PDF in the hypersphere's tangent space

at mean μ and (ii) Markov-based Gibbs energy that penalizes deviation of each point's location from its neighbors' locations. We extend this PDF to shape space, enforcing rotation invariance. The Normal law relies on the logarithmic map $\text{Log}_\mu(\cdot)$ and the exponential map $\text{Exp}_\mu(\cdot)$. Let $z_{in} \in \mathbb{R}^3$ be the n-th point in the shape pointset for the i-th individual. The Markov model relies on a neighborhood system $\mathcal{N} := \{\mathcal{N}_n\}_{n=1}^N$, where $\mathcal{N}_n$ gives the set of neighbors of the n-th point. We enforce the same $\mathcal{N}$ for all pointsets. We propose to model shape variability, for all individuals i, via $P(z_i|\mu, C, \beta) :=$

$$\frac{1}{\eta(\mu, C, \beta)} \exp\left(-\frac{1}{2}(\text{Log}_\mu(z_i))^T C^{-1} \text{Log}_\mu(z_i) - \frac{\beta}{2} \sum_n \sum_{m \in \mathcal{N}_n} \|z_{in} - z_{im}\|_2^2 \right), \quad (1)$$

where $\eta(\mu, c, \beta)$ is the normalization constant.

2.2 Efficient Model of Likelihood of Image Data Given Object Shape

Let random vector X_i model the observed image data for individual i. We propose a novel generative model for the observed individual data $\{x_i\}_{i=1}^I$ given individual shapes $\{z_i\}_{i=1}^I$. This paper considers the data to represent the object boundary by the set of pixels on the zero crossing of a level set fitted to the 0.5-valued isosurface of the fuzzy segmentation. Each *data pointset* x_i can have a different number of points M_i.

The proposed model relies on a novel dissimilarity measure $\delta(x_i, z_i)$ between two pointsets x_i and z_i of different cardinality, where the pointset z_i represents shape (in shape space). We design this dissimilarity measure by measuring the dissimilarity between pointsets, modulo translation t, rotation R, and scale s, as follows:

$$\delta(x_i, z_i) := \min_{s,R,t} \left(\sum_{n=1}^N \min_m \|x_{im}^{sRt} - z_{in}\|_2^2 + \sum_{m=1}^M \min_n \|x_{im}^{sRt} - z_{in}\|_2^2 \right), \quad (2)$$

where data pointset x_i comprises M points $\{x_{im}\}_{m=1}^M$, shape pointset z_i comprises N points $\{z_{in}\}_{n=1}^N$, $s \in \mathbb{R}$ represents isotropic scaling, $R \in \mathbb{R}^{D \times D}$ represents rotation as an orthogonal matrix with determinant 1, $t \in \mathbb{R}^D$ represents translation, and $x_m^{sRt} := sRx_m + t$ is the location of the similarity-transformed point x_m.

The dissimilarity $\delta(\cdot, \cdot)$ is non-negative and symmetric. Ensuring pointsets x_i and z_i to be devoid of coincident points, which is straightforward, $\delta(x_i, z_i) = 0$ if and only if $x_i = z_i$. Thus, except the triangular inequality, $\delta(\cdot, \cdot)$ satisfies all metric properties. Moreover, unlike methods [5,15] using current distance having quadratic complexity in either pointset cardinality, $\delta(x_i, z_i)$ can be well approximated efficiently using algorithms of complexity close to $O(M + N)$ instead of $O(MN)$, by maintaining a list of nearest inter-pointset neighbors and updating it intermittently (see Section 2.3).

Thus, we model $P(x_i|z_i) := \exp(-\delta(x_i, z_i))/\gamma$, with normalization constant γ.

2.3 Parameter Inference Using Expectation Maximization

We fit the model to the data $\bar{\mathbf{x}} := \{x_i\}_{i=1}^I$ by computing the maximum-likelihood estimate (MLE) for the parameters. In this paper, β is a user-defined parameter. Denoting $\theta := \{\mu, C\}$, we solve for the MLE $\arg\max_\theta P(\bar{\mathbf{x}}|\theta)$, treating the shape pointsets $\bar{\mathbf{z}} := \{z_i\}_{i=1}^I$ as hidden / latent variables, via iterative EM optimization.

At iteration t, given parameter estimates $\theta^t := \{\mu^t, C^t\}$, the *E step* defines

$$Q(\theta;\theta^t) := E_{P(\bar{\mathbf{Z}}|\bar{\mathbf{x}},\theta^t)}[\log P(\bar{\mathbf{x}},\bar{\mathbf{Z}}|\theta)],\ \text{where } P(\bar{\mathbf{x}},\bar{\mathbf{z}}|\theta) = \prod_{i=1}^{I} P(x_i|z_i)P(z_i|\theta). \quad (3)$$

The expectation is difficult to evaluate analytically and, hence, we approximate it using Monte-Carlo simulation of shape pointsets $\bar{\mathbf{z}}$ (see Section 2.4) as follows:

$$Q(\theta;\theta^t) \approx \widehat{Q}(\theta;\theta^t) := \frac{1}{S}\sum_{s=1}^{S} \log P(\bar{\mathbf{x}},\bar{\mathbf{z}}^s|\theta),\ \text{where } \bar{\mathbf{z}}^s \sim P(\bar{\mathbf{Z}}|\bar{\mathbf{x}},\theta^t). \quad (4)$$

The *M step* obtains parameter updates $\theta^{t+1} := \arg\max_\theta \widehat{Q}(\theta;\theta^t)$. Given data $\bar{\mathbf{x}}$ and the sample individual shapes $\{\bar{\mathbf{z}}^s\}_{s=1}^S$, we alternatingly optimize the shape-distribution parameters μ, C and the internal parameters $\{s_i, R_i, t_i\}_{i=1}^I$, until convergence.

Update Mean μ: Given C and $\{\bar{\mathbf{z}}^s\}_{s=1}^S$, the optimal shape mean estimate $\widehat{\mu}$ is:

$$\arg\min_{\mu} \sum_{s=1}^{S}\sum_{i=1}^{I} (\mathrm{Log}_\mu(z_i^s))^T C^{-1} \mathrm{Log}_\mu(z_i^s). \quad (5)$$

The gradient of the objective function is the vector $C^{-1}\sum_s\sum_i \mathrm{Log}_\mu(z_i^s)$ in the tangent space at μ. At a local minimum, the gradient magnitude must be zero. For all real-world data with finite variance, the null space of C^{-1} contains only the zero vector. So, the gradient magnitude can be zero only when $\sum_s\sum_i \mathrm{Log}_\mu(z_i^s)$ is zero. We iteratively seek μ satisfying the aforementioned condition, via projected gradient descent (projection via C) where the iterative update, with a step size $\tau \in (0,1)$, is: $\widehat{\mu} \leftarrow \mathrm{Exp}_{\widehat{\mu}}((\tau/SI)\sum_{s=1}^S\sum_{i=1}^I \mathrm{Log}_{\widehat{\mu}}(z_i^s))$. We set $\tau = 0.5$ and find that a few iterations suffice for convergence.

Update Covariance C: We observe that the Markov regularization term enforcing shape smoothness, i.e., $\sum_n\sum_{m\in\mathcal{N}_n} 0.5\beta \parallel z_{in} - z_{im} \parallel_2^2$, is a quadratic function of z_i and can be rewritten as $0.5 z_i^T \Omega z_i$ where Ω is a sparse precision matrix that has (i) all diagonal elements $= 2\beta$, (ii) off-diagonal elements in row n and column $m \neq n$ as $(-\beta)$ when points n and m are neighbors, and (iii) all other off-diagonal elements $= 0$.

Given shape mean μ and shape-smoothness parameter β, the optimal covariance is

$$\arg\min_{C} \sum_{s=1}^{S}\sum_{i=1}^{I} \frac{1}{2}(\mathrm{Log}_\mu(z_i^s))^T C^{-1} \mathrm{Log}_\mu(z_i^s) + \frac{1}{2}(z_i^s)^T \Omega z_i^s + \log\eta(\mu, C, \beta). \quad (6)$$

The normalization constant $\eta(\mu, C, \beta)$ is difficult to evaluate analytically (even for the Normal law [10] alone). Nevertheless, we propose a novel scheme to achieve a good approximation for practical utility that is evident from the results in this paper. We assume that the distribution of shapes $P(Z_i|\theta)$ has sufficiently low variance such that the Gaussian PDF on $\mathrm{Log}_\mu(z_i)$, in the tangent space at μ, can be approximated by a Gaussian PDF on $z_i - \mu$ in the tangent space. That is, we approximate locations z_i to the tangent-space locations $\breve{z}_i := \mu + \mathrm{Log}_\mu(z_i)$. This observation allows us to treat $P(Z_i|\theta,\beta)$

as a product of two Gaussian PDFs in $\mathbb{R}^D$, i.e., (i) a Gaussian PDF $P(Z_i|\theta)$ with mean μ and covariance C restricted to the tangent space at μ, capturing the population shape variability, and (ii) a Gaussian PDF $P(Z_i|\beta)$ enforcing shape smoothness via a Markov-based Gibbs energy, with mean 0 and covariance Ω^{-1}. We know that the product of two Gaussian PDFs is also a Gaussian PDF where the normalization constant is easy to compute. In our case, the product Gaussian $P(Z_i|\theta)P(Z_i|\beta) \approx P(Z_i|\theta,\beta)$ has mean μ and covariance $C_{\text{reg}} = (C^{-1}+\Omega)^{-1}$ that is restricted to the tangent space at μ, which makes the normalization constant $\eta(\mu, C, \beta) \approx (2\pi)^{D/2}|C_{\text{reg}}|^{1/2}$. Thus, the optimal covariance $\widehat{C}_{\text{reg}}$ is the sample covariance of $\breve{z}_i^s$ in the tangent space: $\widehat{C}_{\text{reg}} = (1/(IS))\sum_{s=1}^{S}\sum_{i=1}^{I}\text{Log}_\mu(z_i^s)\text{Log}_\mu(z_i^s)^T$. The optimal covariance $\widehat{C} = (\widehat{C}_{\text{reg}}^{-1} - \Omega)^{-1}$.

Update Similarity Transforms: Given sample shapes $\{\overline{\mathbf{z}}^s\}_{s=1}^S$, we optimize the similarity transforms modeled by $\{s_i, R_i, t_i\}_{i=1}^I$ independently for each individual i via

$$\arg\min_{s_i,R_i,t_i} \sum_{s=1}^{S}\left(\sum_{n=1}^{N}\min_m \parallel x_{im}^{s_iR_it_i} - z_{in} \parallel_2^2 + \sum_{m=1}^{M}\min_n \parallel x_{im}^{s_iR_it_i} - z_{in} \parallel_2^2\right) \quad (7)$$

that optimally aligns the data pointset x_i to the set of sampled shape pointsets $\{z_i^s\}_{s=1}^S$. We optimize $\{s_i, R_i, t_i\}_{i=1}^I$ using gradient descent that relies on an efficient gradient computations. The objective function depends on mappings between each point in one pointset and the closest point in the other pointset (say, $\phi(\cdot)$ and $\psi(\cdot)$, detailed next). While the data pointset x_i has a large number of points densely distributed along the object boundary, the shape pointset z_i^s has far fewer points distributed sparsely along the object boundary (sufficient to model object shape). Assuming the parameter updates to be sufficiently small, (i) for most points x_{im} in the data pointset, the nearest point $z_{i\phi(m)}^s$ in the shape pointset before and after the update is unchanged and (ii) for most points z_{in}^s in the shape pointset, the displacement vector to the nearest point $x_{i\psi(n)}$ in the data pointset before and after the update is unchanged. This leads to a good approximation of the gradient of the objective function. For parameter t_i, the gradient is

$$\sum_{s=1}^{S}\left(\sum_{n=1}^{N}2(x_{i\psi(n)}^{s_iR_it_i} - z_{in}^s)^T\frac{\partial x_{i\psi(n)}^{s_iR_it_i}}{\partial t_i} + \sum_{m=1}^{M}2(x_{im}^{s_iR_it_i} - z_{i\phi(m)}^s)^T\frac{\partial x_{im}^{s_iR_it_i}}{\partial t_i}\right), \quad (8)$$

where $\partial x_{ij}^{s_iR_it_i}/\partial t_i = \mathbb{I}$, the identity matrix. Similarly, for parameter s_i, $\partial x_{ij}^{s_iR_it_i}/\partial s_i = R_i x_{ij}$. Equating each gradient to zero gives closed-form updates for s_i and t_i. To update the rotation matrix R_i, we perform projected gradient descent on the manifold on rotation matrices, with the gradient in ambient space as

$$\sum_{s=1}^{S}\left(\sum_{n=1}^{N}2s_ix_{i\psi(n)}(x_{i\psi(n)}^{s_iR_it_i} - z_{in}^s)^T + \sum_{m=1}^{M}2s_ix_{im}(x_{im}^{s_iR_it_i} - z_{i\phi(m)}^s)^T\right). \quad (9)$$

2.4 Sampling Smooth Shapes in Kendall Shape Space

The E step in Section 2.3 relies on Monte-Carlo approximation to the expectation by sampling, for all individuals i, shape pointsets $\{z_i^s\}_{s=1}^S$ from the posterior PDF

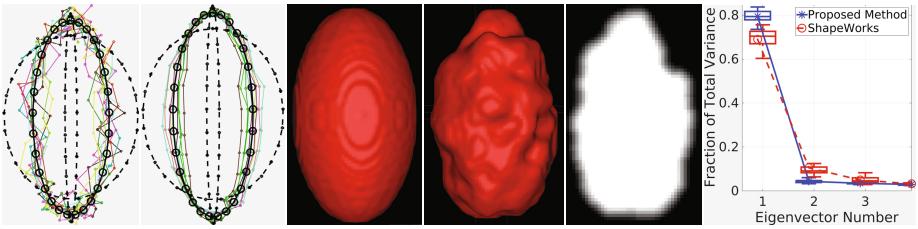

Fig. 1. Smooth-Shape Sampling using HMC. HMC-sampled shapes (colored), **(a)** without shape smoothness ($\beta = 0$) and **(b)** with shape smoothness ($\beta > 0$), from a distribution (on 2D ellipses, for easy viewing); solid black ellipse $\equiv$ mean pointset (point locations $\equiv$ small circles) and the principal mode of variation $\equiv$ dotted black ellipses. **Results on** 40 **Simulated Ellipsoids. (c)** Uncorrupted segmentation. **(d)** Corrupted segmentation with **(e)** a 2D slice of its distance transform (thresholded beyond $[-2, 2]$ mm for viewing). **(f)** Box plot showing fraction of total variance modeled by eigenvectors of estimated shape covariance.

$P(Z_i|\theta^t, \beta, x_i)$ that is proportional to $P(x_i|Z_i)P(Z_i|\theta^t, \beta)$. We sample using a novel adaptation of Hamiltonian Monte Carlo (HMC) [9] to the Riemannian space.

Following the analysis in Section 2.3 (used for updating C), which models the prior $P(Z_i|\theta^t, \beta)$ as a flat Gaussian on the tangent space of the hypersphere at μ^t, we consider the (i) likelihood $P(x_i|Z_i)$ as a non-flat PDF in $\mathbb{R}^D$ and (ii) posterior $P(Z_i|\theta^t, \beta, x_i)$ as a flat PDF on the tangent space at μ^t. Thus, we sample shapes z_i^s by (i) first sampling tangent vectors t^s in the tangent space at μ from the posterior PDF that is restricted to the tangent space and (ii) then producing shape samples $\{z_i^s\}_{s=1}^S$ by taking the exponential map $\text{Exp}_\mu(t^s)$. HMC sampling requires the gradient of $\log P(Z_i|\theta^t, \beta, x_i)$ with respect to Z_i, upto an additive constant. Because we represent the flat posterior PDF (of dimensionality $< D$) in the ambient D-dimensional space, a naive gradient computation can result in a gradient with a component orthogonal to the tangent space of shape space. Hence, within HMC, we replace the naive gradient by a (i) projected gradient onto the tangent space of preshape space followed by (ii) a rotational alignment with the shape mean μ. Figures 1(a)–(b) clearly shows that enforcing shape smoothness ($\beta > 0$) enables modeling realistic shape variations leading to realistic shape samples.

3 Results

We compare the proposed method with ShapeWorks [2,12] that optimizes point locations within shape pointsets, but restricts the points to the object-boundary isosurface in the (fuzzy) segmentation and neither enforces shape smoothness nor realigns data. Both methods use (i) $N = 64$ points in the shape model and (ii) similarity-transform-aligned distance transforms of the object boundaries as input. We initialize shape mean μ, and the neighborhood system $\mathcal{N}$, using the vertices and edges of a triangular mesh fitted to the object surface in one of the input images. After getting optimal θ^* via EM, we solve for optimal (i) individual pointsets z_i^* as $\arg\max_{z_i} P(z_i|x_i, \theta^*, \beta)$ and (ii) aligned data pointsets $x_i^{s_i^*, R_i^*, t_i^*}$ to aid evaluation (described next).

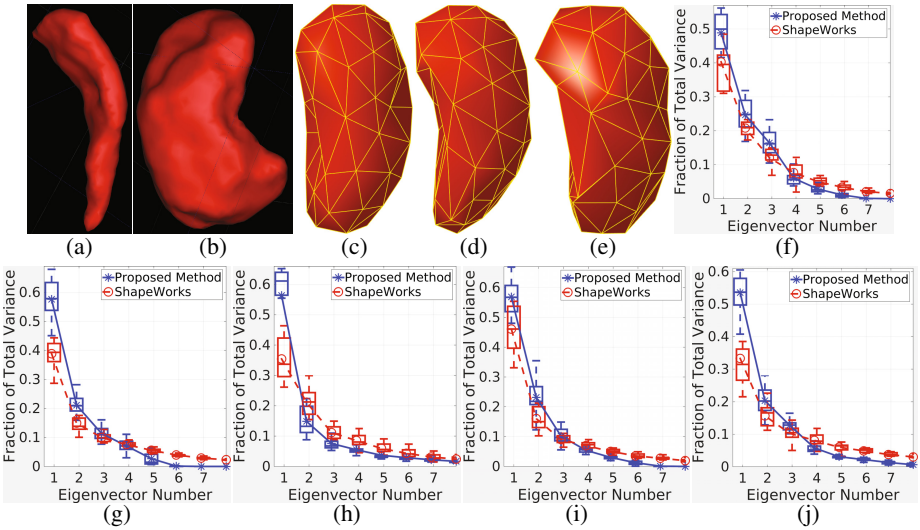

Fig. 2. Results on 186 **Clinical Brain MR Images.** A segmentation for **(a)** caudate, **(b)** thalamus. **(c)–(e)** Sampled thalamic shapes. Fraction of total variance modeled by eigenvectors of estimated shape covariance for **(f)** caudate, **(g)** globus pallidus, **(h)** hippocampus, **(i)** putamen, **(j)** thalamus.

Validation on Simulated Data: We generate 40 images of ellipsoids (Figure 1(c)) by introducing 1 major mode of variation, keeping 2 of the ellipsoid axis lengths to be fixed (40 mm) and varying the third (between 40–80 mm). The proposed method with $\beta = 0$ (without enforcing shape smoothness) and ShapeWorks were both able to learn this variation correctly, producing only 1 non-zero eigenvalue for the learned covariance matrix (in our case, C_{reg}). The optimal point locations z^*_{in}, within shape pointsets, were virtually exactly on the object boundary seen in the input images. We then *corrupt* the segmentation data with random perturbations to each object boundary (Figures 1(d)–(e)), get optimal shape-model parameter estimates θ^*, and repeat 20 times. With the corrupted data, the proposed method correctly detects a single mode of variation (Figure 1(f)). We set $\beta = 2$ to be sufficiently small to ensure that, within the optimal shape pointsets, point locations z^*_{in} were still within about 1 mm of the *uncorrupted* object boundary (mean 0.9 mm, standard deviation 0.2 mm). ShapeWorks incorrectly detected ≥ 2 modes of variation (Figure 1(f)), which is a less compact statistical model.

ShapeWorks restricts the shape pointset to lie on the segmented object boundary, asssuming it to be devoid of errors from, e.g., random noise or inhomogeneity. ShapeWorks may treat coarse- or fine-scale random spatial perturbations in object boundary's segmentation as part of the true signal, leading to inflated covariance. We prevent this by allowing a small deviation of the shape pointset from segmented object boundary.

Evaluation on Clinical Data: We used 186 T1-weighted magnetic resonance (MR) brain images (voxel size 1 mm^3 isotropic) with expert segmentations for the caudate, globus pallidus, hippocampus, putamen, and thalamus (Figures 2(a)–(e)) provided by the National Alliance for Medical Image Computing (`www.na-mic.org`).

The proposed method used $\beta = 2$ for all structures except the hippocampus where we set $\beta = 0.2$ (weaker smoothness enforced) because of the more complex hippocampal shape. We set β to be sufficiently small to ensure that distances between (i) points z^*_{in} in each optimal shape pointset and (ii) the closest point $x^{s^*_i, R^*_i, t^*_i}_{i\phi(n)}$ in the realigned data pointset were about the same as the distances in case of $\beta = 0$. To evaluate the robustness of the proposed method, we bootstrap sample 40 brains from the 186 brains, estimate the shape model, and repeat 20 times. Compared to ShapeWorks, the proposed method is able to learn more compact models (Figures 2(f)–(j)) by capturing a larger fraction of the total variance in fewer eigenvectors / dimensions.

Conclusions. The results show that the proposed method produces more compact shape models consistently for several anatomical structures, using novel contributions in the form of (i) Riemannian statistical shape modeling, (ii) Markov regularity for shape smoothness, (iii) optimal realignment of data during model learning, and (iv) a generative model that effectively deals with corrupted data via EM inference. We also propose a novel method for sampling smooth shapes in Kendall shape space. The proposed generic regularized shape model can be used for other applications.

References

1. Allassonniere, S., Kuhn, E., Trouve, A.: Construction of Bayesian deformable models via a stochastic approximation algorithm: A convergence study. Bernoulli 16(3), 641–678 (2010)
2. Cates, J., Fletcher, P.T., Styner, M., Shenton, M., Whitaker, R.: Shape modeling and analysis with entropy-based particle systems. In: Info. Proc. Med. Imag., pp. 333–345 (2007)
3. Cootes, T., Taylor, C., Cooper, D., Graham, J.: Active shape models - their training and application. Comp. Vis. Image Understanding 61(1), 38–59 (1995)
4. Davies, R., Twining, C., Taylor, C.: Statistical Models of Shape: Optimisation and Evaluation. Springer (2008)
5. Durrleman, S., Pennec, X., Trouve, A., Ayache, N.: Statistical models of sets of curves and surfaces based on currents. Med. Imag. Anal. 13(5), 793–808 (2009)
6. Fletcher, T., Lu, C., Pizer, S., Joshi, S.: Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans. Med. Imag. 23(8), 995–1005 (2004)
7. Glasbey, C.A., Mardia, K.V.: A penalized likelihood approach to image warping. J. Royal Statistical Society, Series B 63(3), 465–514 (2001)
8. Kendall, D.: A survey of the statistical theory of shape. Statist. Sci. 4(2), 87–99 (1989)
9. Neal, R.: MCMC using Hamiltonian dynamics. In: Handbook of Markov Chain Monte Carlo, pp. 113–162. Chapman and Hall, CRC Press (2010)
10. Pennec, X.: Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. J. Mathematical Imaging and Vision 25(1), 127–154 (2006)
11. Rangarajan, A., Coughlan, J., Yuille, A.: A Bayesian network framework for relational shape matching. In: Int. Conf. Comp. Vis., pp. 671–678 (2003)
12. SCI Institute: ShapeWorks: An open-source tool for constructing compact statistical point-based models of ensembles of similar shapes that does not rely on specific surface parameterization (2013). `http://www.sci.utah.edu/software/shapeworks.html`
13. Siddiqi, K., Pizer, S.: Medial Representations: Mathematics, Algorithms and Applications. Springer, Heidelberg (2008)
14. Vaillant, M., Glaunes, J.: Surface matching via currents. In: Info. Proc. Med. Imag., pp. 381–392 (2005)
15. Yu, Y., Fletcher, P., Awate, S.: Hierarchical Bayesian modeling, estimation, and sampling for multigroup shape analysis. In: Med. Image Comput. Comp. Assist. Interv., pp. 9–16 (2014)