

Bayesian Atlas Estimation for the Variability Analysis of Shape Complexes

Pietro Gori^{1,2}, Olivier Colliot^{1,2}, Yulia Worbe², Linda Marrakchi-Kacem^{1,2,3},
Sophie Lecomte^{1,2,3}, Cyril Poupon³, Andreas Hartmann², Nicholas Ayache⁴,
and Stanley Durrleman^{1,2}

¹ Aramis Project-Team, Inria Paris-Rocquencourt, Paris, France

² CNRS UMR 7225, Inserm UMR-S975, UPMC, CRICM, Paris, France

³ Neurospin, CEA, Gif-Sur-Yvette, France

⁴ Asclepios Project-Team, Inria Sophia Antipolis, Sophia Antipolis, France

Abstract. In this paper we propose a Bayesian framework for multi-object atlas estimation based on the metric of currents which permits to deal with both curves and surfaces without relying on point correspondence. This approach aims to study brain morphometry as a whole and not as a set of different components, focusing mainly on the shape and relative position of different anatomical structures which is fundamental in neuro-anatomical studies. We propose a generic algorithm to estimate templates of sets of curves (fiber bundles) and closed surfaces (sub-cortical structures) which have the same “form” (topology) of the shapes present in the population. This atlas construction method is based on a Bayesian framework which brings to two main improvements with respect to previous shape based methods. First, it allows to estimate from the data set a parameter specific to each object which was previously fixed by the user: the trade-off between data-term and regularity of deformations. In a multi-object analysis these parameters balance the contributions of the different objects and the need for an automatic estimation is even more crucial. Second, the covariance matrix of the deformation parameters is estimated during the atlas construction in a way which is less sensitive to the outliers of the population.

1 Introduction

In the last years statistical analysis of shapes has acquired a central role in medical imaging. One of the main applications is to find morphological differences between a population of controls and one of patients or to highlight the effects of a treatment (i.e. drug) on a group of patients. Most studies focus their attention on a single anatomical structure [6, 8–10, 12]. Others propose multi-object analysis considering only a particular kind of shape, either only surfaces (sub-cortical structures) [2, 11, 13] or only curves (fiber-bundles) [5]. However, brain anatomy consists of an intricate network of white matter fiber bundles and sub-cortical structures which need to be studied together in many neuro-anatomical studies. One example is the study of the neural circuits whose morphological changes are often correlated with neurodevelopmental disorders, such as in Gilles de la

Tourette syndrome (GTS) [3]. To this end, we propose here a new atlas construction method based on the framework of currents which permits to deal simultaneously with curves and surfaces. Given a set of shape complexes, each one is seen as a deformation of a common template complex. Both the template complex and the deformations, together called “atlas”, need to be estimated and they characterize the anatomical invariants and the variability of the population. Every deformation is based on one single diffeomorphism of the whole 3D space which preserves the spatial organization of the objects and prevents shape components to intersect, fold or shear during deformation. In the case of neuro-anatomical studies, this makes possible a more realistic analysis since one can study the brain as a whole and not as a series of independent components. Moreover the metric between shape complexes is derived from the metric on deformations that consistently integrates the variations of each component of the complex. The use of currents allows to estimate such deformations without relying on point or fiber correspondence.

We extend to shape complexes (curves and surfaces) the methodology of [2] proposed only for surfaces and we enrich it. The aspects borrowed from [2] aim not to increase the dimensionality of the deformation parametrization between a single and a multi-object analysis and to make the results more useful for neuro-anatomical studies. The first one is achieved by separating the parametrization of the deformations from the one of the template complex and therefore making it not dependent on the size and number of objects. A second aspect is to construct templates with the same form (topology) of the shapes. This was not the case in previous works based on currents like in [5] where the template was a set of disconnected Dirac delta currents or in [6] where it was the superimposition of warped surfaces. Now templates of closed surfaces have a mesh structure defined by the user while templates of fiber bundles have the form of sets of curves connecting cortical to sub-cortical structures. This makes easier the interpretation of the results since the template complex can be compared with the shapes of the population and it is also possible to study the relations between objects of the template complex which is crucial in neuro-anatomical studies like in [3].

In this paper, the atlas is estimated using the same generative model as in [1, 2]. We propose to estimate the atlas of shape complexes using similar Bayesian priors as in [1] for images. This enables to automatically estimate the trade-off between the data term and the regularity of the deformations instead of being fixed by the user as in [2]. This parameter is important in the context of statistical analysis since a value too small might weight too much the data term leading to a situation of over-fitting. On the contrary, a value too large might penalize the deformations making less accurate the analysis of the variability of the population. Moreover, the situation is even more complicated in multi-object analysis where each object is characterized by its own parameter which weights the contribution of the object in the criterion to be optimized. Objects characterized by a bigger norm than the others should be weighted lesser in order to balance all the contributions for the atlas estimation. The automatic estimation of the different parameters takes into account also this aspect in only one simulation.

Previously [2] it was necessary more than one simulation in order to understand the “best” trade-off and the choice was very subjective.

A second improvement is the estimate of the covariance matrix of the deformation parameters during the atlas construction in a more statistically robust way. This should lead to a PCA less influenced by the outliers in contrast to an analysis based on a covariance matrix computed at the end of the procedure.

2 Bayesian Framework for Atlas Estimation

In the first paragraph we recall how to build diffeomorphic deformations with the new control point scheme defined in [2]. Afterwards, we explain the novelty of our new Bayesian formulation, highlighting the methodological differences with respect to [2] linked to the use of priors. Eventually we introduce an innovative solution to initialize templates of shape complexes (curves and surfaces).

Diffeomorphism. Assume that we have M different structures segmented from structural and diffusion images for N subjects. All structures belonging to subject i can be seen as a shape complex $\mathbf{S}_i = \{S_{ij}\}_{j=1\dots M}$ which is modeled as the deformation of a common template complex $\phi^i(\mathbf{T})$ plus a residual ϵ_i where $\mathbf{T} = \{T_j\}_{j=1\dots M}$ and $\epsilon_i = \{\epsilon_{ij}\}_{j=1\dots M}$. Shape complexes, deformed template complex and residuals are modelled as currents. The deformation ϕ^i depends only on subject i . The whole 3D space is deformed by a single diffeomorphism using the control point formulation presented in [2]. Diffeomorphic deformations are built by integrating a time-varying vector field $v_t(x)$ over the interval $[0, 1]$. Calling $\phi_t(x)$ the position of a point x at time t , its evolution is given by: $\dot{\phi}_t(x) = v_t(\phi_t(x))$ and under some smoothness constraints of v_t satisfied here [7] the set of deformations $\{\phi_t\}_{t \in [0,1]}$ is a flow of diffeomorphisms. The speed vector field v_t is defined by a dynamical system of C_p control points $\mathbf{c} = \{c_k\}$, shared among the whole population and a set of momenta $\alpha^i = \{\alpha_k^i\}$ linked to each control point and specific to each subject: $\dot{x}(t) = v_t(x(t)) = \sum_{p=1}^{C_p} K(x(t), c_p(t)) \alpha_p^i(t)$, where K is an interpolating kernel (i.e. gaussian). This equation defines the motion of all the points in the 3D space: both template and control points. The momenta α^i parametrize the deformation of the template complex towards the shapes of subject i . If we assume that there are no external forces in the system, the total energy is conserved and it is equal to the Hamiltonian $\sum_{k=1}^{C_p} \sum_{p=1}^{C_p} \alpha_k^i(t)^T K(c_k(t), c_p(t)) \alpha_p^i(t)$. Control points and momenta satisfy therefore the Hamiltonian system:

$$\begin{cases} \dot{c}_k(t) = v_t(c_k(t)) = \sum_{p=1}^{C_p} K(c_k(t), c_p(t)) \alpha_p^i(t) \\ \dot{\alpha}_k^i(t) = - \sum_{p=1}^{C_p} \alpha_k^i(t)^T \alpha_p^i(t) \nabla_1 K(c_k(t), c_p(t)) \end{cases} \quad (1)$$

Integrating Eq.1 and $\dot{\mathbf{T}}(t) = v_t(\mathbf{T}(t))$ with $\mathbf{T}(0) = \mathbf{T}$ from $t=0$ to $t=1$, one obtains the deformed template complex $\mathbf{T}(1) = \phi^i(\mathbf{T})$. The last diffeomorphism ϕ^i at time $t=1$ is completely parametrized by the initial conditions of the system. Since the control points are shared among the population, the only subject-specific deformation parameters are the C_p initial momenta $\{\alpha_k^i(0)\} = \alpha_0^i$.

Currents. The aim of our atlas construction is to estimate simultaneously the C_p control points \mathbf{c} , the N sets of initial momenta $\{\alpha_0^i\}$ and the template complex \mathbf{T} optimizing at the same time the residuals $\{\epsilon_i\}$ in a Bayesian framework sense. More formally this can be achieved by maximizing the joint posterior distribution of \mathbf{c} , $\{\alpha_0^i\}$ and \mathbf{T} given $\{\mathbf{S}_i\}$. The space of currents is of infinite dimension and therefore *pdf* are not defined. In order to overcome this problem we fix M finite dimensional spaces $W_{\Lambda_j}^*$, one for each object j , constituted by grids on which shapes and templates are projected (Π) and where *pdf* are defined. As demonstrated in [7], the norm in this space is: $\Pi(S_{ij} - \phi^i(T_j))^T K_W^j \Pi(S_{ij} - \phi^i(T_j))$ which is precisely the numerical scheme used in [7] to compute the continuous norm on currents. K_W^j is the currents kernel sampled at the grid points of $W_{\Lambda_j}^*$ and it is defined as a block matrix whose blocks are 3D gaussian kernels characterized by the same standard deviation. In this way, both the α_0^i and the residuals ϵ_{ij} can be modelled with Gaussian distributions. Assuming independence of observations, their likelihoods are: $p(\alpha_0^i | \Gamma_\alpha) \propto \frac{1}{|\Gamma_\alpha|^{1/2}} \exp[-\frac{1}{2}(\alpha_0^i)^T \Gamma_\alpha^{-1} \alpha_0^i]$ and $p(\epsilon_{ij} | \sigma_{\epsilon_j}^2) \propto \frac{1}{|\sigma_{\epsilon_j}^2|^{\Lambda_j/2}} \exp[-\frac{1}{2\sigma_{\epsilon_j}^2} \|(S_{ij} - \phi^i(T_j))\|_{W_{\Lambda_j}^*}^2]$, where Λ_j is the number of points of the j -th grid. Γ_α and $\sigma_{\epsilon_j}^2 (K_W^j)^{-1}$ are two covariance matrices. The scalar $\sigma_{\epsilon_j}^2$ depends on the object j and it is the variance of $\epsilon_j = \{\epsilon_{ji}\}_{i=1\dots N}$.

Bayesian Framework. As in [1] for images, it is possible to estimate both Γ_α and $\sigma_{\epsilon_j}^2$ in a Bayesian framework using the standard conjugate prior of the Gaussian distribution, the Inverse Wishart: $\Gamma_\alpha \sim \mathcal{W}^{-1}(P_\alpha, \frac{w_\alpha}{N})$ and $\sigma_{\epsilon_j}^2 \sim \mathcal{W}^{-1}(P_{\epsilon_j}, \frac{w_{\epsilon_j}}{N})$ where the matrix P_α and the scalars $w_\alpha, \{P_{\epsilon_j}\}, \{w_{\epsilon_j}\}$ are new hyper-parameters fixed by the user. Using an uniform prior distribution for the template complex \mathbf{T} and for the control points \mathbf{c}_0 and assuming all random variables independent, it is possible to write explicitly the posterior distribution F of $\mathbf{T}, \{\alpha_0^i\}, \mathbf{c}_0, \Gamma_\alpha, \{\sigma_{\epsilon_j}^2\}$ given the shapes $\{\mathbf{S}_i\}$. Maximizing F is equivalent to minimize $E = -\log(F) =$

$$\begin{aligned} & \frac{M}{2} \sum_{i=1}^N (\alpha_0^i)^T \Gamma_\alpha^{-1} \alpha_0^i + \sum_{j=1}^M \sum_{i=1}^N \frac{1}{2\sigma_{\epsilon_j}^2} \left(\|(S_{ij} - \phi^i(T_j))\|_{W_{\Lambda_j}^*}^2 + \frac{P_{\epsilon_j} w_{\epsilon_j}}{N} \right) + \\ & \frac{M}{2} (w_\alpha + N) \log(|\Gamma_\alpha|) + \frac{M}{2} w_\alpha \text{tr}(\Gamma_\alpha^{-1} P_\alpha) + \sum_{j=1}^M \frac{1}{2} (w_{\epsilon_j} + \Lambda_j N) \log(\sigma_{\epsilon_j}^2) \end{aligned} \tag{2}$$

where the first two terms were present also in [2]. The use of two conjugate priors makes possible to compute the optimal values for Γ_α and $\sigma_{\epsilon_j}^2$ in a closed form:

$$\hat{\Gamma}_\alpha = \frac{\sum_{i=1}^N [(\alpha_0^i)(\alpha_0^i)^T] + w_\alpha P_\alpha}{(w_\alpha + N)} \quad \hat{\sigma}_{\epsilon_j}^2 = \frac{\sum_{i=1}^N \|(S_{ij} - \phi^i(T_j))\|_{W_{\Lambda_j}^*}^2 + w_{\epsilon_j} P_{\epsilon_j}}{(w_{\epsilon_j} + N \Lambda_j)}$$

$\hat{\Gamma}_\alpha$ is equal to a weighted sum between the sample covariance matrix and the prior. A good choice for the prior seems to be: $P_\alpha = K_V^{-1}$, where K_V is a block matrix whose blocks are 3D gaussian kernels between two different control points. If all α_0^i are equal to zero, $\hat{\Gamma}_\alpha \sim K_V^{-1}$ and this means that the “deformation

regularity” part in Eq.2 becomes $\sum_{i=1}^N (\alpha_0^i)^T K_V \alpha_0^i$, which is the sum of the geodesic distances from the template complex to all the shapes as in [2].

The second parameter $\hat{\sigma}_{\epsilon_j}^2$ is equal to a weighted sum between the data-term of the j -th object and the prior and it was previously fixed by the user. The results were highly dependent on its value as shown in the next section. Now it is automatically estimated in a way which balances the contributions of the different objects in order that objects characterized by bigger norms do not stand above the smaller ones. With this new technique the sample covariance matrix penalizes the deformations of the template complex towards “outliers” at each iteration. Thus adjusting also the residual variance $\sigma_\epsilon^2 = \{\sigma_{\epsilon_j}^2\}_{j=1\dots M}$, contrary to [2] where both Γ_α and σ_ϵ^2 were fixed during optimization.

Gradient Descent. For the other parameters $\mathbf{T}, \{\alpha_0^i\}, \mathbf{c}_0$ there is not a closed form and they are computed using a gradient descent algorithm. Their gradients are: $\nabla_{T_k} E = \sum_{i=1}^N \nabla_{T_k} \left[\frac{1}{2\sigma_{\epsilon_k}^2} D_{ik} \right]$, $\nabla_{\alpha_0^s} E = \sum_{j=1}^M \nabla_{\alpha_0^s} \left[\frac{1}{2\sigma_{\epsilon_j}^2} D_{sj} \right] + M \Gamma_\alpha^{-1} \alpha_0^s$ and $\nabla_{\mathbf{c}_0} E = \sum_{i=1}^N \sum_{j=1}^M \nabla_{\mathbf{c}_0} \left[\frac{1}{2\sigma_{\epsilon_j}^2} D_{ij} \right]$. The differentiation of the data term $D_{ij} = \|(S_{ij} - \phi^i(T_j))\|_{W_{\Lambda_j}^*}^2$ is exactly the same as in [2]. The gradient of the prior P_α with respect to \mathbf{c}_0 is not taken into account since its norm is negligible. The use of a gradient descent method implies the choice of an initial template. Its “form” (topology) is preserved during the minimization process. We propose to initialize templates of 3D closed surfaces as centred and scaled ellipsoids. Templates of fiber bundles are initialized in two steps. First, it is selected randomly 10% of the fibers of each subject bundle from its most dense part. After, all the fibers are grouped and it is used a greedy approximation method based on the framework of currents to select the H most representative fibers, where H is the average number of fibers in the subject bundles.

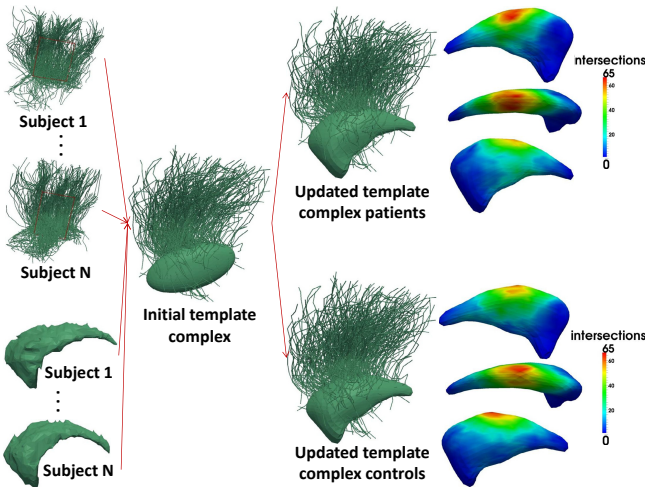


Fig. 1. Left: template update process for left caudate (LC) and left caudate bundle (LCB) in patients and controls. Center: updated template complexes starting from the same initial template complex but using only the population of patients or controls. Right: number of intersections between fibers and surface.

3 Experiments and Discussion

In the following experiments the data set is composed of left caudate (LC) and left caudate bundle (LCB) of 10 controls and 10 patients with GTS. The segmentation of the shapes was performed from T1-weighted MRI and DWI respectively [3, 4]. We use the non-oriented currents metric, also called varifold metric [14], which allows not to orient shapes in a consistent way across population.

The initial template process ends with two updated template complexes for the population of controls and patients respectively (Fig.1). These two template complexes reveal the common anatomical features of their populations. In Fig.1-right are highlighted the number of fibers intersecting the surface. Fibers seem to be more spread on the template surface of patients with respect to controls.

In Fig.2 we show the updated template complexes deformed according to the first mode of PCA for both groups using the estimated covariance matrix. In the set of controls there is mainly a variability in the distribution of fibers along the upper part of the surface. For patients there is especially a spread/concentration of the fibers on the surface towards the sides of the nucleus. There is also an elongation/shortening of the surface but it is common to both populations.

Eventually we evaluate the robustness of our method w.r.t. the hyperparameter values and we compare its performance with the same optimization scheme using $\Gamma_\alpha = K_V^{-1}$ and different fixed values for σ_ϵ^2 , leading to the same results as with the method in [2]. In Fig.3 we show the norm of the difference between the updated template complex and a reference template complex which we choose arbitrarily equal to the one with a fixed σ_ϵ^2 equal to 0.01 for all the

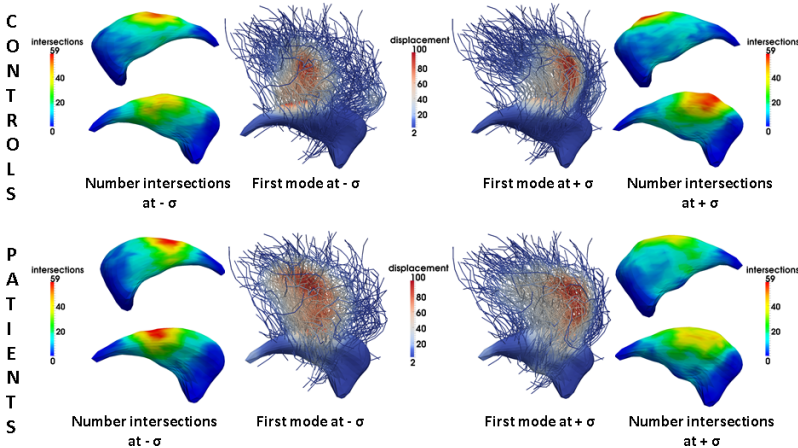


Fig. 2. PCA based on the covariance matrix of the deformation parameters. The central panels show the deformed updated template complexes at $\pm \sigma$ along the first principal direction. The colors refer to the magnitude of the displacement of the points from the template complex. The lateral panels show the number of intersections between the bundle and the surface at $\pm \sigma$.

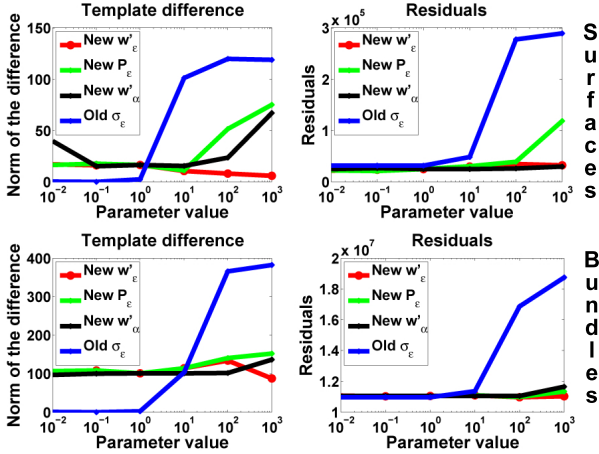


Fig. 3. Comparison of the robustness of the results w.r.t. hyper-parameters values (“New”) and previously manually set trade-off (“Old”). Each result for the “New” method comes from an atlas estimation where it has been changed only one of the hyper-parameters fixing the others equal to 1. Both w_ϵ and w_α have been normalised ($w_\epsilon = w'_\epsilon NA$, $w_\alpha = w'_\alpha N$) so that the results can be compared using the same range of values for both the hyper-parameters and the trade-off.

objects. We show also the residuals ϵ obtained at the end of the atlas procedure. In both cases we show results using different normalised hyper-parameter values or different fixed values for σ_ϵ^2 . It is possible to conclude that the results with a fixed σ_ϵ^2 (as in [2]) lead to results more variable than with the hyper-parameters which can therefore have a “universal” value, for example equal to 1. We obtained similar results for different sets of parameters using only surfaces, only bundles or both.

4 Conclusions

We have presented here a new multi-object atlas construction method based on a Bayesian framework and we have compared its performance with the one of a previous technique described in [2]. We have shown that this new formulation permits to have results less sensitive to the parameters fixed by the user. Moreover the covariance matrix of the deformation parameters is estimated throughout the atlas construction penalizing the contributions of the outliers. This new method can be applied simultaneously to fiber bundles (curves) and sub-cortical structures (closed surfaces) for which we propose also a generic template initialization procedure. Initial templates show the same “form” of the shapes which is also preserved during the atlas estimation. Moreover this technique allows to preserve the underlying organization of the structures under examination. This permits to study the relative positions of different objects in the template complex and how they interact. In this way it is possible, for example, to highlight precisely the anatomical differences between a population of controls and

patients. Our preliminary results in a GTS study show differences in neuronal connexions which still need to be discussed in regards to the hypothesis put forth in [3]. The use of currents, which minimizes the need of user intervention, gives the possibility to apply full brain morphometry to large data sets.

References

1. Allasonnière, S., Amit, Y., Trouvé, A.: Toward a Coherent Statistical Framework for Dense Deformable Template Estimation. *J. R. Statist. Soc. B* 69(1), 3–29 (2007)
2. Durrleman, S., Prastawa, M., Korenberg, J.R., Joshi, S.C., Trouvé, A., Gerig, G.: Topology Preserving Atlas Construction from Shape Data without Correspondence using Sparse Parameters. In: *Proc. Med. Image Comput. Comput. Assist. Interv.* (2012)
3. Worbe, Y., Gerardin, E., Hartmann, A., Valabrègue, R., Chupin, M., Tremblay, L., Vidailhet, M., Colliot, O., Lehericy, S.: Distinct structural changes underpin clinical phenotypes in patients with Gilles de la Tourette syndrome. *Brain* 133, 3649–3660 (2010)
4. Marrakchi-Kacem, L., Delmaire, C., Guevara, P., Poupon, F., Lecomte, S., Tucholka, A., Roca, P., Yelnik, J., Durr, A., Mangin, J., Lehericy, S., Poupon, C.: Mapping Cortico-Striatal Connectivity onto the Cortical Surface: A New Tractography-Based Approach to Study Huntington Disease. *PLoS One* 8(2), e53135 (2013)
5. Durrleman, S., Fillard, P., Pennec, X., Trouvé, A., Ayache, N.: Registration, Atlas Estimation and Variability Analysis of White Matter Fiber Bundles Modeled as Currents. *NeuroImage* 55(3), 1073–1090 (2011)
6. Glaunès, J., Joshi, S.C.: Template estimation from unlabeled point set data and surfaces for Computational Anatomy. In: *Proc. International Workshop on the Mathematical Foundations of Computational Anatomy* (2006)
7. Durrleman, S.: Statistical models of currents for measuring the variability of anatomical curves, surfaces and their evolution. PhD thesis, University of Nice (2010)
8. Hufnagel, H., Pennec, X., Ehrhardt, J., Ayache, N., Handels, H.: Computation of a Probabilistic Statistical Shape Model in a Maximum-a-posteriori Framework. *Methods Inf. Med.* 48(4), 314–319 (2009)
9. Kurtek, S., Klassen, E., Ding, Z., Avison, M.J., Srivastava, A.: Parameterization-invariant shape statistics and probabilistic classification of anatomical surfaces. *Inf. Process. Med. Imaging* 22, 147–158 (2011)
10. Joshi, S.H., Klassen, E., Srivastava, A., Jermyn, I.: A Novel Representation for Riemannian Analysis of Elastic Curves in \mathbb{R}^n . In: *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 1–7 (2007)
11. Chen, T., Rangarajan, A., Eisenschenk, S.J., Vemuri, B.C.: Construction of a neuroanatomical shape complex atlas from 3D MRI brain structures. *NeuroImage* 60(3), 1778–1787 (2012)
12. Davies, R.H., Twining, C.J., Cootes, T.F., Taylor, C.J.: Building 3-D Statistical Shape Models by Direct Optimization. *IEEE Trans. Med. Imag.* 29(4), 961–981 (2010)
13. Gorczowski, K., Styner, M., Jeong, J.Y., Marron, J.S., Piven, J., Hazlett, H.C., Pizer, S.M., Gerig, G.: Multi-object analysis of volume, pose, and shape using statistical discrimination. *IEEE Trans. Pattern. Anal. Mach. Intell.* 32(4), 652–661 (2010)
14. Charon, N., Trouvé, A.: Functional currents: a new mathematical tool to model and analyse functional shapes. *CoRR* abs/1206.3564 (2012)