

Simultaneous Manifold Learning and Clustering: Grouping White Matter Fiber Tracts Using a Volumetric White Matter Atlas

Demian Wassermann^{1,2} Rachid Deriche¹

¹ Odyssée Project Team INRIA/ENPC/ENS

INRIA, Sophia-Antipolis, 2004 Route des Lucioles, 06902 Sophia Antipolis, France

² CS Dept., FCEyN, UBA, Pabellón 1, Ciudad Universitaria, 1428 Buenos Aires, Argentina
{demian.wassermann, rachid.deriche}@sophia.inria.fr

Abstract. We propose a new clustering algorithm. This algorithm performs clustering and manifold learning simultaneously by using a graph-theoretical approach to manifold learning. We apply this algorithm in order to cluster white matter fiber tracts obtained from Diffusion Tensor MRI (DT-MRI) through streamline tractography. Our algorithm is able to perform clustering of these fiber tracts incorporating information about the shape of the fiber and a priori knowledge as the probability of the fiber belonging to known anatomical structures. This anatomical knowledge is incorporated as a volumetric white matter atlas, in this case LONI's ICBM DTI-81.

1 Introduction

Diffusion MRI recovers the in vivo and non invasively effective diffusion of water molecules in histological tissues, thus providing unique biologically and clinically relevant information not available from other imaging modalities. This information can help characterize tissue micro-structure and its architectural organization [1] by modeling the local anisotropy of the diffusion process of water molecules. Once the diffusion information has been recovered within each voxel, it can be assembled through the volume in order to assess brain connectivity in vivo, one of the two main techniques in order to do this is streamline tractography [2], which recovers white matter fiber tracts from a seed voxel by following the principal direction of the diffusion tensor. Multiple seeds can be scattered in the brain resulting in a full brain reconstruction of the fiber tracts as depicted in fig. 1. However, analysis or visualization of the whole fiber ensemble in order to get insight of the brain structure is difficult by the cluttering of the fibers. Thus, there is a need for automatic fiber clustering algorithms in order to perform visualization, anatomical structure identification and group analysis.

There are two main issues to deal with in fiber tract clustering. The first one is how to assess geometric similarity between fibers and the second is how to incorporate *a priori* knowledge of anatomical structures. Several approaches have been proposed to automatically produce fiber tract clusters. The works by [3,4] apply regular clustering algorithms over shape metrics obtaining moderate results due to the complex structure

of fiber tract clusters. To overcome this problem the works [5,6,7] use spectral manifold learning techniques in order to produce a mapping of the fiber tracts to a high-dimensional Euclidean space. In that space regular clustering algorithms are applied incorporating the complex structure of the cluster. At the heart of these algorithms, lies the construction of a huge matrix of fiber-to-fiber similarities and the resolution of a highly memory intensive eigenproblem for which only approximate solutions can be generated. In order to incorporate a priori knowledge to this approach [6] takes the computed clusters and performs expert validation and manual reclustering generating an atlas. Then, the identification of new fibers is done by projection to this atlas. The creation of this high-dimensional atlas needs fine-tuning of several parameters which, in spite that their behaviour is well studied in [6], it is a difficult task with great influence in the final results. Recently, in order to better deal with outliers and artifacts due to tractography and perform quantitative analysis, two works employed expectation maximization techniques: in [8] the EM clustering is performed incorporating the atlas developed by [6] and inheriting its difficulties and in [9] the EM approach requires the number of clusters and focuses totally on the shape of the fibers lacking the incorporation of a priori knowledge. *A priori* information was also incorporated to the same problem in [10] simultaneously with our work.

In this work we present a new algorithm which performs manifold learning and clustering simultaneously by using elements at the heart of two techniques, IsoMap manifold learning [11] and medoidshift clustering [12]. The present approach has several advantages; in the first place, this algorithm does not need the computation of huge matrices nor the resolution of eigenproblems; secondly it automatically finds the number of clusters, thus being able to handle outliers by producing small sized clusters. Moreover, our algorithm uses the shape of each fiber and a priori knowledge in using a volumetric atlas as shown in fig. 1 thus providing anatomically coherent classifications of the fiber tracts.

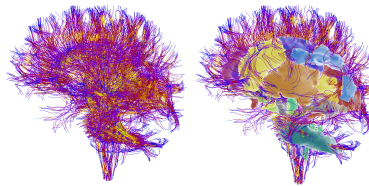


Fig. 1. Whole brain streamline tractography example. The fibers are colored according to the anisotropy of the water diffusion in each segment. The cluttering of the fibers makes difficult to get insight of the data [5]. On the right a volumetric atlas, LONI ICBM DTI-81, representing anatomical knowledge has been superimposed to the tractography.

2 Methods

In this section we describe the two principal parts of our manifold learning and clustering approach: the mode seeking clustering methodology, the incorporation of simultaneous manifold learning and finally the application to DT fiber tracts.

2.1 Mode seeking by mean shifts

Mode seeking is the general name for a family of clustering algorithms which assumes that the data to be clustered can be regarded as an empirical probability distribution function (PDF) where these dense regions correspond to the maxima of the PDF or its *modes*. Once these *modes* are located, clusters associated with each mode can be delineated. Mean shift based mode seeking techniques were introduced in the fields of computer vision and pattern recognition by [13], amongst others. These techniques consist of shifting each element to be clustered towards its corresponding *mode* until the *mode* has been reached. Then the number of clusters is determined by the number of *modes*, that is the points in each cluster to which the shifting procedure converged to.

The mean shift procedure can be formulated as follows [13]. Let \mathcal{D} be a set of elements $\{x_1 \dots x_N\}$ sampled from a manifold \mathcal{M} associated with a distance function $\text{dist}(\cdot, \cdot)$, and let $\Phi(\cdot)$ be a kernel function. Then, the PDF characterizing \mathcal{D} at x can be estimated by the kernel density,

$$f_{\langle \mathcal{D}, d \rangle}(x) = \frac{c}{N} \sum_i \Phi \left(\frac{\text{dist}^2(x_i, x)}{h^2} \right) \quad (1)$$

where h is an application-oriented bandwidth parameter and c is a normalizing constant ensuring that $f_{\langle \mathcal{D}, d \rangle}(\cdot)$, or $f(\cdot)$, integrates to 1. In order to shift an element towards its corresponding *mode*, or local maxima, of $f(\cdot)$: let ϕ be the negated derivative of Φ , $\phi = -\Phi'$. A shifting operation is applied to each element in \mathcal{D} until it reaches a *mode*. In order to avoid the need of an interpolation operation amongst the elements in \mathcal{D} and due to the robustness of the medoid operator, the modes are estimated using the recent *medoid shift* approach [12]: the operation shifting y_t to the *sample medoid* at y_t , y_{t+1} is

$$y_{t+1}^{\text{medoid}} = \arg \min_{y \in \mathcal{D}} \sum_i \frac{\text{dist}^2(x_i, y)}{h^2} \phi \left(\frac{\text{dist}^2(x_i, y_t)}{h^2} \right). \quad (2)$$

As a final remark to this section, the importance of working in the right manifold, i.e. using the right metric between the points, can be emphasised by the following simulated example. Let $\mathcal{T} = \{t_1 \dots t_{|\mathcal{T}|}\}$, where t_i is drawn with equal probability from one of the following three gaussian distributions, $\mathcal{G}(0, \frac{1}{5})$, $\mathcal{G}(2, \frac{1}{2})$ or $\mathcal{G}(5, 1)$, thus producing one cluster from each distribution. Then our dataset is produced as a mapping of \mathcal{T} , $\mathcal{D} = \{x_i = x(t_i) : x(t) = -t(\sin(t); \cos(t))\}$. In this sample, the manifold \mathcal{M} is the curve $s(t) = -t(\sin(t); \cos(t))$, $t \in [0, \inf)$, and the mapping results in three different clusters along a spiral as shown in fig. 2. Taking the kernel $\Phi(\zeta) = e^{-\zeta}$ we perform two kernel density estimations with different distance functions. The first one, uses the natural distance on the euclidean space \mathbb{R}^2 , or *extrinsic*

distance, $\text{dist}_{\mathbb{R}^2}(x_i, x_j) = \|x_i - x_j\|_2$, followed by kernel density estimation being performed by following equation (1). The value of the kernel density estimation over \mathbb{R}^2 , $f_{\langle \mathcal{D}, \text{dist}_{\mathbb{R}^2} \rangle}$, is plotted under \mathcal{D} as contour isolines in fig. 2(a), where red indicates higher and blue lower values, respectively; it can be seen that the *modes* of $f_{\langle \mathcal{D}, \text{dist}_{\mathbb{R}^2} \rangle}$, the red areas, are not aligned with the clusters of \mathcal{D} . The second distance, is chosen to be a distance on the parameter space \mathcal{T} , or *intrinsic distance*, $\text{dist}_{\mathcal{T}}(x_i, x_j) = \|t_i - t_j\|_2$. Then performing kernel density estimation using the function $f_{\langle \mathcal{D}, \text{dist}_{\mathcal{T}} \rangle}$ is equivalent to doing it over the one dimensional curve manifold $\langle \mathcal{M}, \text{dist}_{\mathcal{M}} \rangle$ shown in fig. 2(b), where the different *modes* of the distribution are aligned with the clusters as expected.

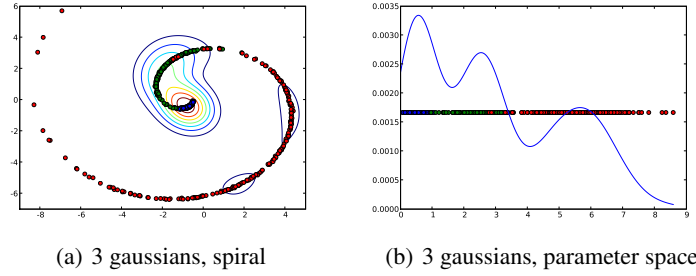


Fig. 2. Importance of working in the right manifold for kernel density estimation. Given 3 clusters with different gaussian distribution embedded in a curve, on the left kernel density estimation on the bidimensional space enclosing the curve are shown as contour plots. Concentration areas do not coincide with clusters. On the right kernel density estimation along the curve: peaks coincide with the clusters.

2.2 Medoidshift with simultaneous Manifold Learning

In this section we describe an algorithm which incorporates manifold learning and medoidshift clustering in order to perform the clustering in the intrinsic space of the dataset. The importance of working in the right manifold has been stressed in section 2.1.

The main tool of this section is the following theorem³, [14, Main Theorem A]

Theorem 1 (Geodesics are approximated by shortest paths on an induced graph). Let $\mathcal{D} = \{x_1, \dots, x_{|\mathcal{D}|}\}$ be a set of finite discrete points properly sampled from the compact manifold $\mathcal{M} \subset \mathcal{R}$, and an extrinsic distance function in \mathcal{R} , $\text{dist}_{\mathcal{R}} : \mathcal{R} \times \mathcal{R} \rightarrow \mathbb{R}_{\geq 0}$. A graph $G(V, k) = \langle V, E(k) \rangle$ can be built where the vertexes are $V = \mathcal{D}$, the edges are defined as $E(k) = \{(x_i, x_j) : \text{dist}_{\mathcal{R}}(x_i, x_j) \leq \text{neigh}^k(x_i)\}$ and $\text{neigh}^k(x_i)$ is the distance $\text{dist}_{\mathcal{R}}(\cdot, \cdot)$ to the k th closest neighbor of x_i .

³ In this work the theorem is stated in a simplified form.

Then, as $|\mathcal{D}| \rightarrow \infty$, the length of the shortest path between two points in the graph $\text{dist}_{G(\mathcal{D},k)}(x_i, x_j)$ asymptotically converges to the geodesic distance on the manifold \mathcal{M} , or intrinsic distance, $\text{dist}_{\mathcal{M}}(x_i, x_j)$.

In order to produce a clustering algorithm capable of working in the learned manifold without solving an eigenproblem as done in [11,5,15], we present an algorithm which does not rely on interpolation and differential operators generalizing the idea of using the IsoMap distance matrix as starting point for medoidshift [12]. The formulation for kernel density estimation equation (1), over the learned manifold can be formulated by using the graph distance, $\text{dist}_G(\cdot, \cdot)$, defined in theorem 1 as $f_{\langle \mathcal{D}, \text{dist}_G \rangle}(x) = \frac{c}{N} \sum_i \Phi\left(\frac{\text{dist}_G^2(x, x_i)}{h^2}\right)$. Due to theorem 1, this algorithm asymptotically approaches an algorithm working on the geodesics of \mathcal{D} . The *modes*, and its associated clusters, on this density estimation can be approximated by modifying the medoidshift formula equation (2),

$$y_{i+1}^{\text{graphmedoid}} = \arg \min_{y \in \mathcal{D}} \sum_i \frac{\text{dist}_G^2(x_i, y)}{h^2} \phi\left(\frac{\text{dist}_G^2(x_i, y)}{h^2}\right). \quad (3)$$

2.3 Application to DT white matter fiber tracts

In this section we explain the use of our algorithm which performs clustering over learned manifolds to cluster a set for dMRI-extracted fiber tracts $\mathcal{F} = \{F_1, \dots, F_{|\mathcal{F}|}\}$ taking in account the shape of the fibers and a priori knowledge about white matter anatomy.

Shape Distances between fiber tracts Once the streamline fiber tracts have been extracted from dMRI we want to be able to measure similarity in order to cluster them. Curve matching is an extensive research field in computer vision and medical imaging. In the particular case of DT-MRI tractography, several metrics have been proposed in order to quantify similarity among fiber tracts [16,17,18,9]. Nevertheless this is still and open problem, specially in the case of large deformations [19]. In this work, we assume that fiber tracts within the same bundle can be characterized by a sequence of small deformations, spanning a manifold on the curve space. Hence, using theorem 1, we are only concerned with quantifying similarity between very similar fibers and then propagating this similarity. In order to quantify these small similarities we resort to the symmetrized Chamfer distance : Given two fiber tracts represented as point sequences like $F = \{\mathbf{p}_1, \dots, \mathbf{p}_{|F|} \in \mathbb{R}^3\}$, the distance metric between F_i and F_j is calculated as,

$$\text{dist}_{\text{Chamfer}}(F_i, F_j) := \frac{\left(\sqrt{\frac{\sum_{\mathbf{p} \in F_i} \min_{\mathbf{p}' \in F_j} (\|\mathbf{p} - \mathbf{p}'\|^2)}{|F_i|}} + \sqrt{\frac{\sum_{\mathbf{p}' \in F_j} \min_{\mathbf{p} \in F_i} (\|\mathbf{p} - \mathbf{p}'\|^2)}{|F_j|}} \right)}{2} \quad (4)$$

Finally, in order to obtain the intrinsic distance between fibers, we apply theorem 1: a directed weighted graph is induced from the distance $\text{dist}_{\text{Chamfer}}$ in order to represent

the fiber shape feature: $GS := G(\mathcal{F}, k_F)$ as stated in theorem 1. Finally the geodesic distance on the fiber manifold is approximated by $\text{dist}_{GS}(F_i, F_j)$, the shortest path from F_i to F_j in the graph GS .

Anatomical knowledge prior information In order to model *a priori* anatomical knowledge we define the probability of a fiber F crossing a label l of an atlas A ,

$\text{Prob}(F \in l|A) = \frac{1}{|F|} \sum_{i=1}^{|F|} \beta(A(\mathbf{p}_i) = l)$ where $\beta(\cdot)$ is the indicator function. Thus each fiber is identified by a discrete probability distribution and the Kullback-Leibler divergence, $D_{KL}(F_i||F_j) = \sum_{l \in \mathcal{L}} \text{Prob}(F_i \in l|A) \log \left(\frac{\text{Prob}(F_i \in l|A)}{\text{Prob}(F_j \in l|A)} \right)$, symmetrized becomes a natural choice of distance function between fibers

$$\text{dist}_{KL}(F_i, F_j) = D_{KL}(F_i||F_j) + D_{KL}(F_j||F_i). \quad (5)$$

Finally in order to approximate the intrinsic distance function of the manifold where each fiber is represented by a discrete probability distribution, we resort to theorem 1: Let the graph $GA = G(\mathcal{F}, k_A)$. Thus, the intrinsic distance is approximated by $\text{dist}_{GA}(F_i, F_j)$, the shortest path from F_i to F_j in the graph GA .

DT Tractography clustering Let $\mathcal{F} = \{F_1, \dots, F_N\}$ be the set of fiber tracts in a tractography, by using the distances $\text{dist}_{GS}(\cdot, \cdot)$ and $\text{dist}_{GA}(\cdot, \cdot)$. A multi-feature kernel density estimation function is derived from equation (1), $f_{\langle \mathcal{F}, \text{dist}_{GS}, \text{dist}_{GA} \rangle}(F) = \frac{c}{N} \sum_i \Phi \left(\frac{\text{dist}_{GS}^2(F_i, F)}{h_S^2} + \frac{\text{dist}_{KL}^2(F_i, F)}{h_A^2} \right)$ where, for each fiber $y_0 \in \mathcal{F}$, the corresponding modal fiber is found by reaching the fixed point of the following equation derived from equation (3)

$$y_{t+1} = \underset{y \in \mathcal{F}}{\text{argmin}} \sum_i \left(\frac{\text{dist}_{GS}^2(F_i, y)}{h_S^2} + \frac{\text{dist}_{GA}^2(F_i, y)}{h_A^2} \right) \phi \left(\frac{\text{dist}_{GS}^2(F_i, y_t)}{h_S^2} + \frac{\text{dist}_{GA}^2(F_i, y_t)}{h_A^2} \right). \quad (6)$$

Clustering algorithm The full algorithm for white matter fiber tracts is as follows:

Algorithm 1 Given a set of fiber tracts $\mathcal{F} = \{F_1, \dots, F_{|\mathcal{F}|}\}$ and the atlas ICBM DTI-81, noted as A , registered to the coordinated space of the fiber set:

1. Generate the graph $GS = G(\mathcal{F}, k_S)$ using the distance shown in equation (4) and taking k_S as the smallest number s.t. GS is connected.
2. Generate the graph $GA = G(\mathcal{F}, k_A)$ using the distance shown in equation (5) and the atlas A . The parameter is taken k_A as the smallest number s.t. GA is connected.
3. For each fiber $f \in \mathcal{F}$, set $y_0 = f$ and iterate equation (6) until convergence.
4. Delineate each cluster as all the fibers $f \in \mathcal{F}$ that converged to the same mode

3 Results

DTI data was acquired by the Magnetic Resonance Imaging Center at the Pitié-Salpêtrière hospital, Paris, France with a 3T Siemens Trio TIM 32 channel system. Diffusion

weighted images were acquired with 50 gradient directions and 60 slices with a FOV of $256mm$, a b-value of $1000 \frac{s}{mm^2}$ and $2mm^3$ isotropic voxels. Diffusion tensor estimation and streamline tractography were performed by using riemannian estimation with the Odyssee Brain Visa toolbox, finally affine registration was performed using the FSL software toolkit (<http://www.fmrib.ox.ac.uk>). The number of neighbors for the graphs GS and GA for each image were chosen as the smallest number such that the graph is connected. Clusters under a specific size were considered outliers and eliminated. Results are shown in fig. 3 where only selected fiber bundles are shown for clarity. In this figure, it can be seen the correct clustering of the left cortical spinal tract(yellow), both cingulum bundles(dark purple and blue), the arcuate fasciculus(light purple) and the inferior longitudinal fasciculus(red), has been performed.

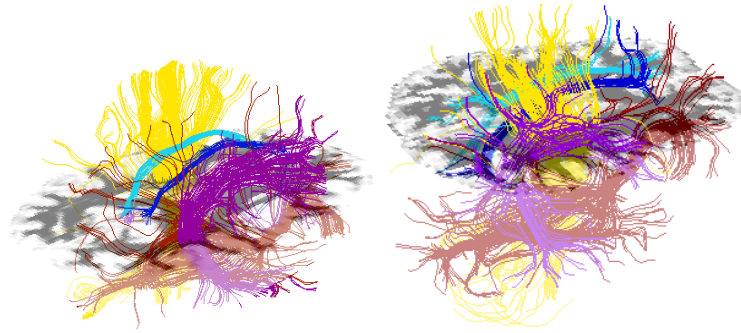


Fig. 3. Clustering results for two sample cases where only selected fiber bundles are shown for clarity. In this figure it can be seen that correct clustering of the left cortical spinal tract(yellow), both cingulum bundles(cyan and blue), the arcuate fasciculus(light purple) and the inferior longitudinal fasciculus(red), has been performed

4 Conclusions

In this paper we present a simultaneous manifold learning and clustering algorithm in order to cluster white matter fiber tracts using the shape of the fibers and the spatial location combined with a volumetric atlas. After the tracts are calculated, the algorithm only requires the tuning of 4 parameters, the two parameters in order to build the graphs which can be obtained automatically and the bandwidth parameters for the clustering process which should be tuned by the user. Finally we have shown that this algorithm successfully identifies major fiber bundles. Further work should study automatic determination of the bandwidth parameters and the use of the clustered results in order to perform group analysis.

References

1. Basser, P., Pierpaoli, C.: Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of Magnetic Resonance B* **111**(3) (1996) 209–219
2. Mori, S., Crain, B.J., Chacko, V.P., Zijl, P.C.M.V.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* **45**(2) (1999) 265–269
3. Shimony, J., Snyder, A., Lori, N., Conturo, T.: Automated fuzzy clustering of neuronal pathways in diffusion tensor tracking. *Proc. Intl. Soc. Mag. Reson. Med* **10** (2002)
4. Ding, Z., Gore, J., Anderson, A.: Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine* **49**(4) (2003) 716–721
5. Brun, A., Park, H., Knutsson, H., Shenton, M.E., Westin, C.: Clustering fiber traces using normalized cuts. In: *Medical Image Computing and Computer-Assisted Intervention MICCAI 2004*. Volume 3216 of *Lecture Notes in Computer Science*, Springer/Verlag (2004) 518–529
6. O'Donnell, L.J., Westin, C.F., Golby, A.J.: Tract-based morphometry. In Ayache, N., Ourselin, S., Maeder, A., eds.: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2007*. Volume 4792 of *LNCS*, Springer (2007) 161–168
7. Tsai, A., Westin, C., Hero III, A., Willsky, A.: Fiber Tract Clustering on Manifolds With Dual Rooted-Graphs. *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on (2007)* 1–6
8. Ziyang, U., Sabuncu, M.R., Grimson, W.E.L., Westin, C.F.: A robust algorithm for fiber-bundle atlas construction. In: *MMBIA, IEEE Computer Society* (2007)
9. Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M.: A unified framework for clustering and quantitative analysis of white matter fiber tracts. *Medical Image Analysis* **12**(2) (2008) 191–202
10. Maddah, M., Zollei, L., Grimson, W.E.L., Westin, C.F., Wells, W.M.: A mathematical framework for incorporating anatomical knowledge in DT-MRI analysis. In: *ISBI*. (2008)
11. Tenenbaum, J., Silva, V., Langford, J.: A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* **290**(5500) (2000) 2319
12. Sheikh, Y.A., Khan, E., Kanade, T.: Mode-seeking by medoidshifts. In: *IEEE International Conference on Computer Vision*. Number 141 (oct 2007)
13. Comaniciu, D., Meer, P.: Mean shift: a robust approach toward feature space analysis. *IEEE PAMI* **24**(5) (may 2002) 603–619
14. Bernstein, M., Silva, V., Langford, J.C., Tenenbaum, J.B.: Graph approximations to geodesics on embedded manifolds. (dec 2000)
15. O'Donnell, L.J., Westin, C.F.: Automatic tractography segmentation using a high-dimensional white matter atlas. *IEEE Transactions on Medical Imaging* **26**(11) (nov 2007) 1562–1575
16. Brun, A., Park, H., Knutsson, H., Westin, C.: Coloring of dt-mri fiber traces using laplacian eigenmaps. In: *Computer Aided Systems Theory - EUROCAST*. Volume 2809 of *Lecture Notes in Computer Science*. Springer (2003) 518–529
17. Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G.: Fiber tract-oriented statistics for quantitative diffusion tensor mri analysis. In: *Medical Image Analysis*. Volume 10. (Oct 2005) 786–798 *The Eighth International Conference on Medical Imaging and Computer Assisted Intervention MICCAI 2005*.
18. O'Donnell, L., Westin, C.F.: High-dimensional white matter atlas generation and group analysis. In Larsen, R., Nielsen, M., Sporring, J., eds.: *MICCAI*. Volume 4191 of *Lecture Notes in Computer Science*. Springer (2006) 243–251
19. Glaunès, J., Qiu, A., Miller, M., Younes, L.: Large deformation diffeomorphic metric curve mapping. *International Journal of Computer Vision* (2008)