Published in final edited form as:
Proc Workshop Math Methods Biomed Image Analysis. 2008 July 15; 2008: 1–7. doi:10.1109/CVPRW.2008.4563003.

Modeling of Anatomical Information in Clustering of White Matter Fiber Trajectories Using Dirichlet Distribution

Mahnaz Maddah[1,2], **Lilla Zöllei**[1,3], **W. Eric L. Grimson**[1,2], and **William M. Wells**[1,2]

Mahnaz Maddah: mmaddah@mit.edu

[1] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

[2] Surgical Planning Laboratory, Brigham and Women's Hospital, Boston, MA 02115, USA

[3] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA 02129, USA

Abstract

In this work, we describe a white matter trajectory clustering algorithm that allows for incorporating and appropriately weighting anatomical information. The influence of the anatomical prior reflects confidence in its accuracy and relevance. It can either be defined by the user or it can be inferred automatically. After a detailed description of our novel clustering framework, we demonstrate its properties through a set of preliminary experiments.

1. Introduction

In recent years, diffusion tensor MR imaging (DT-MRI) has emerged as a powerful tool to identify white matter pathologies, minimize post-operative neurological deficit, and study brain development and aging. Tractography algorithms are often used to extract pathways of fiber tracts in order to aid the visualization of brain connectivity. However, most clinical studies to date have focused on the analysis of scalar diffusion parameters measured in a manually or semi-automatically defined region of interest (ROI) [12]. It is well-known that such ROI-based methods suffer from user-dependence and uncertainty in defining these ROIs, which sometimes lead to inconsistency in the results [6]. Several tract-oriented quantitative analysis algorithms have thus been proposed to mitigate the subjectivity in defining the ROIs [3,9]. However, tractography methods themselves are prone to the subjectivity and sensitivity resulting from the selection of seed points [2]. Furthermore, due to the presence of noise and image imperfections, outliers are often generated in tractography.

It is expected that clustering of fiber trajectories into groups can greatly improve the quality of tract-oriented analysis and eliminate the problems associated with imperfections in the tractography step. Clustering, especially if the groups correspond to anatomically known bundles, ensures that the measurements are performed on a given tract in all subjects rather than on a somewhat arbitrary region. It also allows the tractography process to be seeded from the whole brain or a sufficiently large ROI, which makes it less dependent on the quality of the user input. Although a variety of clustering algorithms have been proposed in the literature to group fiber trajectories, they are mostly unsupervised methods [1,3,10]. We believe that a supervised clustering, that benefits from anatomical information, not only produces anatomically meaningful clusters, but also yields more robust results that are less sensitive to the presence of outliers and imperfections in the DT-MRI data.

Earlier attempts to use anatomical information in fiber clustering are limited to methods where trajectories are grouped based on their distance to the trajectories in a reference subject (atlas) [8] and where groups of fiber trajectories are associated to anatomical structures after they are clustered and mapped into an embedded space [11]. The former relies solely on the atlas information and treats each trajectory individually, while the latter does not use the atlas in the clustering phase. Most recently, an atlas-based quantitative analysis of white matter fiber tracts has been proposed in which a probabilistic parcellation map of the tracts is used to obtain the weighted average of a set of parameters [5]. This approach relies exclusively on the atlas data and hence does not take into account the coherence among the trajectories in each bundle. Furthermore, it is limited to averages over the entire ROI or slices perpendicular to the main axes of the atlas coordinate system.

While existing methods provide valuable quantitative information about the integrity of fiber tracts, they are either limited to specific fiber bundles or prone to be adversely affected by inaccurate settings of user-specified parameters. In this work, we introduce a probabilistic method that rigorously incorporates anatomical prior information from an atlas of fiber tracts [13]. This framework is an extension of our earlier work, in which we described an expectation-maximization solution to cluster the fiber trajectories in a mixture model context [9]. The influence of the anatomical prior in the current model reflects confidence in its accuracy and relevance. It can either be defined by the user or it can be inferred automatically. After a detailed description of the model, we demonstrate its properties through a set of preliminary experiments.

2. Method

We treat each trajectory as a 3-D curve, uniformly sampled along its arc length. For each cluster, a center is defined, which is a sampled curve similar to the trajectories. For each trajectory, a vector $\mathbf{d}_i = [d_{i1},\ldots, d_{iK}]$ is calculated, where d_{ik} is the distance between the trajectory and the kth cluster center as a function of their point coordinates and correspondences [7], and K is the user-defined number of clusters (See Figure 2).

Once the trajectories are extracted from the DTI data using a tractography algorithm, they are mapped into the atlas coordinate system. Each trajectory then takes a membership probability $\boldsymbol{q}_i^{'}=[q_{i1}^{'},\ldots,q_{iK}^{'}]$, where each $q_{ik}^{'}$ element denotes the atlas-specified membership of each trajectory i to cluster k and is calculated by summing up the probabilities of its overlapping voxels with the probability maps of the fiber tracts in the atlas, and normalizing with the volume of each tract in the atlas. The membership probabilities of each trajectory are then normalized, so that $\forall i, \sum_{k=1}^{K} q_{ik}^{'}=1$.

We denote the unknown label of each trajectory by z which is assumed to follow a single-trial multinomial distribution,

$$z_i \sim \text{Multinomial}(\pi_i), \tag{1}$$

$$\Pr(z_i=k|\pi_i)=\pi_{ik}, \tag{2}$$

and $\boldsymbol{\pi}_i$'s follow a Dirichlet distribution, with parameters controlled by the atlas. Specifically, $\sum_{k=1}^{K}\pi_{ik}=1$ and

$$\pi_i \sim \text{Dirichlet}(\mathbf{q}_i) = \frac{\Gamma(q_{i0})}{\prod_{j=1}^{K}\Gamma(q_{ij})}\prod_{k=1}^{K}\pi_{ik}^{(q_{ik}-1)}, \tag{3}$$

where $\Gamma(.)$ is the gamma function, and $q_{i0} = \sum_{k=1}^{K} q_{ik}$. To control the influence of the atlas to some extent, $\boldsymbol{q}_i$ can be set as $\boldsymbol{q}_i = \gamma \boldsymbol{q}_i'$ where $\gamma > 0$ is a weight factor. Doing so, the expectation of the Dirichlet distribution for each trajectory is set to the value it gets from the atlas, and the variance of distribution is controlled by γ. However, once the expectations, i.e. q_{ik}''s, are set, the variation of the Dirichlet distribution is limited. Better control is achieved if a weighted average of the vector $\boldsymbol{q}_i'$ and the unity vector is used, i.e.

$$\boldsymbol{q}_i = a\gamma \boldsymbol{q}_i' + \mathbf{1}. \tag{4}$$

The weight, a, controls the degree of influence of the atlas on the clustering and can be set by the user. In the limiting case when $a = 0$, the Dirichlet distribution becomes a uniform distribution and hence the atlas does not have any control on the clustering. When $a = 1$ the model reduces to $\pi_i \sim \text{Dirichlet}(\gamma \boldsymbol{q}_i' + \mathbf{1})$, which is slightly different from what we introduced earlier to guarantee that the parameters of the distribution are larger than unity and avoid a U-shaped distribution. Alternatively, the weight a can be inferred from the atlas data if multiple atlases are available. One possible application of such a setting is to set the weight proportional to the correlation between the atlases, i.e., set a close to 1 when all atlases agree with each other and close to 0 when they disagree. Specifically, suppose that M atlases, $Q'^{(1)}, \ldots, Q'^{(M)}$, are available, where $Q'^{(m)} = \{\mathbf{q}_i'^{(m)}\}_{m=1}^{M}$. A plausible setting for a can be obtained by forming the correlation matrix of $Q^{(m)}$'s and taking the average of its non-diagonal elements. The prior Q' is calculated as:

$$\boldsymbol{q}_i' = \frac{\sum_{m=1}^{M} \boldsymbol{q}_i^{(m)}}{M} \tag{5}$$

A more sophisticated possibility could combine multiple atlases, for example, if we had a collection of voxelwise labellings (perhaps generated by hand), we could use STAPLE [14] to generate voxelwise label probabilities as input as described at the beginning of this section. When STAPLE combines labellings, it weights them according to how consistent they are with the others.

To understand the dependencies between the variables, Fig. 2 shows the directed graphical model of the problem setup. In this case $\Theta = \{\mathbf{g}, \boldsymbol{\pi}\}$, i.e., the collection of the parameters of the gamma distribution functions, $\mathbf{g}$ as well as parameters of the multinomial distribution, $\boldsymbol{\pi}$.

The goal is to estimate the membership likelihood of each curve to each cluster based on the values of the d_{ik}'s and q_{ik}'s. Note that calculating the probability density of the points on the curves is not straightforward as they are not statistically independent. Therefore, we use the distance metric defined between the trajectories and cluster centers, d_{ik}'s, to build the probability model. As detailed in [7], we assume that distances for each cluster follow a Gamma distribution with shape, α_k, and inverse scale parameters, β_k:

$$\text{Gamma}(d|\alpha_k,\beta_k)=d^{\alpha_k-1}\frac{\beta_k^{\alpha_k}e^{-\beta_k d}}{\Gamma(\alpha_k)}\text{for } d\in[0,\infty] \tag{6}$$

where Γ(.) is the gamma function. We assume that each element of the vector $\mathbf{d}_{\underline{i}}$ follows the distribution

$$d_{ik}\sim\begin{cases}\text{Gamma}(g_k) & k=z_i\\ \text{Uniform}([0,d_0]) & k\neq z_i\end{cases}, \tag{7}$$

where, g_k is the parameter set of the Gamma distribution, $\{\alpha_k, \beta_k\}$, for each cluster and U(x; 0, d_0) is the uniform distribution function over [0, d_0] with d_0 a large enough constant.

3. Expectation Maximization

Expectation Maximization (EM), introduced by [4] is frequently used for the inference with missing data or parameters. To apply the EM approach in clustering, an indicator variable is usually defined to represent the hidden variable which here is the cluster memberships. We denote this variable with Z, where $z_i = k$ represents the membership of the data point i to the cluster k. We denote the complete data likelihood by $p(\mathbf{d}, \mathbf{z} \mid \Theta)$. The goal of our EM approach is to iteratively find the MAP estimates of the parameter Θ by maximizing the expectation of log posterior, or equivalently the joint, of the complete data at each iteration:

$$\Theta^{t+1}=\arg\max_{\Theta}\text{E}_{\mathbf{z}|\mathbf{d},\Theta^t}[\log p(\mathbf{z},\mathbf{d},\Theta)]$$

where superscript t denotes the set of parameters at iteration t. The maximization step takes the form:

$$\begin{aligned}\Theta^{t+1}&=\arg\max_{\mathbf{g},\pi}\text{E}_{\mathbf{z}|\mathbf{d},\mathbf{g}^t,\pi^t}[\log p(\mathbf{z},\mathbf{d},\mathbf{g},\pi;\mathbf{q})]\\ &=\arg\max_{\mathbf{g},\pi}\text{E}_{\mathbf{z}|\mathbf{d},\mathbf{g}^t,\pi^t}[\log p(\mathbf{d}|\mathbf{z},\mathbf{g})+\log p(\mathbf{z}|\pi;\mathbf{q})+\log p(\pi;\mathbf{q})+\log p(\mathbf{g})]\end{aligned} \tag{8}$$

We assume a uniform distribution over **g**, the parameters of the Gamma distribution, therefore the last term in the above equation is a constant and can be omitted from the maximization expression.

$$\begin{aligned}\Theta^{t+1}&=\arg\max_{\mathbf{g},\pi}\text{E}_{\mathbf{z}|\mathbf{d},\mathbf{g}^t,\pi^t}[\sum_{i=1}^{N}\log p(\mathbf{d}_i|z_i,\mathbf{g})+\sum_{i=1}^{N}\log p(z_i|\pi;\mathbf{q})+\log p(\pi;\mathbf{q})]\\ &=\arg\max_{\mathbf{g},\pi}\log p(\pi;\mathbf{q})+\sum_{i=1}^{N}\sum_{k=1}^{K}p_{ik}(\log p(\mathbf{d}_i|z_i=k,\mathbf{g})+\log\pi_{ik})\end{aligned} \tag{9}$$

With the assumption of independence of π_i's, we write $\log p(\pi;\mathbf{q})=\sum_{k=1}^{K}\log p(\pi_i;q_i)$ where π_i's follow a Dirichlet distribution as specified in Equ. (3).

3.1. Expectation Step

p_{ik}'s are computed in the expectation step using Bayes rule as follows:

$$\begin{aligned} p_{ik} &= p(\mathbf{z}_i = k | \mathbf{d}_i, \mathbf{g}_k^t, \pi_i^t) \\ &= \frac{p(\mathbf{d}_i | \mathbf{z}_i = k, \mathbf{g}_k^t) p(\mathbf{z}_i = k | \pi_i^t)}{\sum_{j=1}^{K} p(\mathbf{d}_i | \mathbf{z}_i = j, \mathbf{g}_j^t) p(\mathbf{z}_i = j | \pi_i^t)} \\ &= \frac{p(\mathbf{d}_i | \mathbf{z}_i = k, \mathbf{g}_k^t) \pi_{ik}}{\sum_{j=1}^{K} p(\mathbf{d}_i | \mathbf{z}_i = j, \mathbf{g}_j^t) \pi_{ij}}. \end{aligned} \tag{10}$$

3.2. Maximization Step

Now we only look at the i'th component of Θ:

$$\Theta_i^{t+1} = \arg\max_{\mathbf{g}, \pi_i} \left(\sum_{k=1}^{K} (q_{ik} - 1) \log \pi_{ik} + \sum_{k=1}^{K} (p_{ik} \log p(\mathbf{d}_i | z_i = k, \mathbf{g}) + p_{ik} \log \pi_{ik}) \right) \tag{11}$$

Θ_i is the collection of parameter of the multinomial distribution $\boldsymbol{\pi}_i$ and parameter of the Gamma distribution $\mathbf{g}$. We do the maximization with respect to each parameter separately as follows:

3.2.1 Updating parameters of the multinomial distribution—In order to get the updated parameters of the multinomial distribution:

$$\begin{aligned} \pi_i^{t+1} &= \arg\max_{\pi_i} \sum_{k=1}^{K} ((q_{ik} - 1) \log \pi_{ik} + p_{ik} \log \pi_{ik}) \\ &= \arg\max_{\pi_i} \sum_{k=1}^{K} ((q_{ik} + p_{ik} - 1) \log \pi_{ik}), \end{aligned} \tag{12}$$

under the constraint that

$$\sum_{k=1}^{K} \pi_{ik} = 1 \; \forall i \tag{13}$$

To do the maximization, we add the Lagrange multiplier, λ, to Equ.(12) and differentiate with respect to π_{ik}

$$\pi_i^{(t+1)} = \arg\max_{\pi_i} \sum_{k=1}^{K} ((q_{ik} + p_{ik} - 1) \log \pi_{ik}) - \lambda \left(\sum_{k=1}^{K} \pi_{ik} - 1 \right) \tag{14}$$

or

$$\frac{q_{ik} - 1 + p_{ik}}{\pi_{ik}} - \lambda = 0 \tag{15}$$

This results in:

$$\pi_{ik}=\frac{q_{ik}-1+p_{ik}}{\lambda}. \tag{16}$$

λ is obtained by summing up the Equ.(16) over the clusters:

$$\lambda=\frac{\sum_{k=1}^{K}(q_{ik}-1+p_{ik})}{\sum_{k=1}^{K}\pi_{ik}}=q_{i0}-K+1. \tag{17}$$

Substituting the values from Equ. (4) for q_{ik} in the above equation results in the following expression which explicitly shows the relation to the controlling parameters of atlas influence:

$$\pi_{ik}=\frac{a\gamma q_{ik}^{'}+p_{ik}}{a\gamma+1}. \tag{18}$$

In the limiting case when $a = 0$, the above equation reduces to $\pi_{ik} = p_{ik}$, and when $a\gamma \gg 1$ it becomes $\pi_{ik}=q_{ik}^{'}$.

3.2.2 Updating parameters of the gamma distribution—The updated cluster parameter, **g**, is obtained from:

$$\mathbf{g}^{t+1}=\arg\max_{\mathbf{g}}\sum_{i=1}^{N}\sum_{k=1}^{K}p_{ik}\log p(\mathbf{d}_i|z_i=k,\mathbf{g}) \tag{19}$$

The maximization equations are obtained by differentiating the following expression from the EM formulation with respect to each parameter.

$$Q(\mathbf{g})=\sum_{i=1}^{N}\sum_{k=1}^{K}p_{ik}\left(\log\text{Gamma}(d_{ik};\mathbf{g}_k)+\log\prod_{j\neq k}\text{U}(d_{ij};0,d_0)\right). \tag{20}$$

The inverse scale parameter is obtained by $\partial Q(\mathbf{g})/\partial\beta_k = 0$, which simply results in:

$$\sum_{i=1}^{N}p_{ik}\left(\frac{\alpha_k}{\beta_k}-d_{ik}\right)=0 \tag{21}$$

or

$$\beta_k=\alpha_k\frac{\sum_{i=1}^{N}p_{ik}}{\sum_{i=1}^{N}p_{ik}d_{ik}}. \tag{22}$$

The shape parameter is given by $\partial Q(\mathbf{g})/\partial\alpha_k = 0$, which thus:

$$\sum_{i=1}^{N} p_{ik}\left[\log d_{ik}+\log\alpha_k+1+\log\frac{\sum_{i=1}^{N}p_{ik}}{\sum_{i=1}^{N}p_{ik}d_{ik}}+d_{ik}\frac{\sum_{i=1}^{N}p_{ik}}{\sum_{i=1}^{N}p_{ik}d_{ik}}-\psi(\alpha_k)\right]=0 \quad (23)$$

where $\psi(.) = \Gamma'(.)/\Gamma(.)$ is the digamma function. The resulting equation

$$\log\alpha_k-\psi(\alpha_k)=-\log\frac{\sum_{i=1}^{N}p_{ik}}{\sum_{i=1}^{N}p_{ik}d_{ik}}-\frac{\sum_{i=1}^{N}p_{ik}d_{ik}}{\sum_{i=1}^{N}p_{ik}}, \quad (24)$$

does not have a closed-form solution. However, good approximations can be obtained by noting that

$$\log\alpha_k-\psi(\alpha_k)\approx\frac{1}{\alpha_k}\left(\frac{1}{2}+\frac{1}{12\alpha_k+2}\right). \quad (25)$$

This gives:

$$\alpha_k\approx\frac{3-x+\sqrt{(x-3)^2+24x}}{12x}, \quad (26)$$

where

$$x=\log\left(\frac{\sum_i p_{ik}d_{ik}}{\sum_i p_{ik}}\right)-\frac{\sum_i p_{ik}\log(d_{ik})}{\sum_i p_{ik}}. \quad (27)$$

Once the EM algorithm converges, we update the cluster centers and recompute the distance vectors. The outliers are identified in the expectation step. If the membership likelihoods of a trajectory in all clusters are less than a user-specified threshold, that trajectory is identified as an outlier and is removed from further data processing. In fact, with this threshold, the heterogeneity of the trajectories within each cluster can be controlled. The larger the threshold is, the more compact the resulting bundles are, and consequently the greater the number of unclustered trajectories.

4. Results

To show that the proposed method is able to effectively control the influence of the prior information on the clustering, we apply it on some simulated data. For illustration purposes, we consider two clusters and construct **d** by drawing samples from gamma and uniform distributions as specified in Equ.(7). Two scenarios were examined: In the first one, the atlas priors, q'_{ik}s, were drawn from a uniform distribution in the [0.8, 1] range for each data point *i* that belongs to the *k*th cluster. This models a case where the atlas prior is in agreement with the membership probability inferred from the data only. In the second case, q'_{ik}s, were drawn from a uniform distribution in the [0, 0.2] range if the data point *i* belongs to the *k*th cluster,

i.e. an extreme case where the atlas priors oppose the membership probability inferred form the data.

Figure 3 shows a histogram of the posterior probabilities, colored based on the true clustering and for the case where the atlas prior is in agreement with the membership likelihood inferred from the data. Clustering results for two different values of $a = 0$, i.e. no atlas, and $a = 1$, i.e. full atlas control are shown. The ideal clustering is defined as a case where all data points are assigned to their true clusters, in other words when all data points from each cluster appear in either the upper or lower half of the histogram. While without the atlas roughly 15% of the data points are mis-clustered, only a few data points are mis-clustered once the atlas is used.

Figure 4 shows the case where the atlas disagrees with the membership likelihoods inferred from the data and for different values of the weight a. In this case the number of mis-clustered data points increases as the atlas weight is increased. Fig. 5 summarizes the impact of the atlas weight on the number of the mis-clustered data points. When the atlas agrees with the data, the mis-clustering ratio decreases as the atlas weight is increased and when it disagrees with the atlas the mis-clustering ratio increases.

We believe that the flexibility that our method offers on controlling the influence of the anatomical atlas has important applications in clinical studies. One such application is in analyzing pathological cases where fiber tracts might deviate significantly from normal cases and thus the anatomical atlas. Since cluster centers are able to evolve during the course of the EM algorithm, the method is still able to cluster the fiber trajectories reasonably well when a small weight is given to the atlas, while the algorithm might fail if the atlas is imposed too strongly. Another example where it is important to control the atlas weight is when registration errors are present. To demonstrate the effectiveness of our method in successfully clustering the fiber trajectories in such a case, Fig. 6 shows the clustering results for trajectories from the cingulum. When the atlas is imposed strongly some of the trajectories that belong to cingulum, i.e. are similar in shape and are located close to other trajectories, are rejected only because they do not have enough overlap with the atlas ROI. With a proper choice of the atlas weight these trajectories are also included in the results.

5. Conclusions

A novel method was introduced to incorporate and control the influence of the anatomical knowledge in clustering of white matter fiber trajectories. The Dirichlet distribution as a conjugate to the multinomial distribution was used to rigorously model the prior information in an EM setup. The influence of the atlas can be controlled by a parameter which can be either set by the user or inferred from the atlas information if more than one atlas is present to reflect the degree of confidence in the prior knowledge. Preliminary results presented in this paper demonstrate the effectiveness of the proposed approach. To our knowledge this is the first implementation which offers control over the strength of the anatomical information for the clustering of fiber trajectories.

Acknowledgments

This work is supported in part by U54EB005149, R01NS0518261, U41RR019703-01A2, NIH 5 P41 RR13218, P41-RR14075, BIRN002, U24 RR021382, R01EB006758 and R01NS052585-01.

References

1. Brun A, Knutsson H, Park HJ, Shenton ME, Westin C-F. Clustering fiber tracts using normalized cuts. Int Conf Medical Image Computing and Computer-Assisted Intervention. 2004:368–375.

2. Clayden JD, Storkey AJ, Bastin ME. A probabilistic model-based approach to consistent white matter tract segmentation. IEEE Trans Medical Imaging. 2007; 26:1555–156.

3. Corouge I, Fletcher PT, Joshi S, Gouttard S, Gerig G. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. Med Image Anal. 2006; 10:786–798. [PubMed: 16926104]

4. Dempster A, Laird N, Rubin D. Maximum likelihood from incomplete data via the em algorithm. Journal of the Royal Statistical Society, Series B. 1977; 39:138.

5. Hua K, Zhang J, Wakana S, Jiang H, Li X, Reich DS, Calabresi PA, van Zijl JJPPC, Mori S. Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. NeuroImage. 2008; 39:336–347. to be published. [PubMed: 17931890]

6. Kanaan R, Shergill S, Barker G, Catani M, Ng V, Howard R, McGuire P, Jones D. Tract-specific anisotropy measurements in diffusion tensor imaging. Psychiatry Res. 2006; 146(1):73–82. [PubMed: 16376059]

7. Maddah M, Grimson W, Warfield S, Wells W. A unified framework for clustering and quantitative analysis of white matter fiber tracts. Med Image Anal. 2008:191–202. [PubMed: 18180197]

8. Maddah M, Mewes A, Haker S, Grimson WEL, Warfield SK. Automated atlas-based clustering of white matter fiber tracts from DTMRI. MICCAI. 2005:188–195. [PubMed: 16685845]

9. Maddah M, Wells WM, Warfield SK, Westin C-F, Grimson WEL. Probabilistic clustering and quantitative analysis of white matter fiber tracts. IPMI. 2007:372–383.

10. O'Donnell, L.; Westin, C-F. White matter tract clustering and correspondence in populations. Lecture Notes in Computer Science; Eighth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'05); Palm Springs, CA, USA. October 2005; p. 140-147.

11. O'Donnell LJ, Westin CF. Automatic tractography segmentation using a high-dimensional white matter atlas. IEEE Transactions on Medical Imaging. 2007; 26(11):1562–1575. [PubMed: 18041271]

12. Park HJ, Westin CF, Kubicki M, Maier SE, Niznikiewicz M, Baer A, Frumin M, Kikinis R, Jolesz FA, McCarley RW, Shenton ME. White matter hemisphere asymmetries in healthy subjects and in schizophrenia: A diffusion tensor MRI study. NeuroImage. 2004; 24:213–223. [PubMed: 15325368]

13. Wakana S, Jiang H, Nagae-Poetscher LM, van Zijl PCM, Mori S. Fiber tract-based atlas of human white matter anatomy. Radiology. 2004; 230:77–87. [PubMed: 14645885]

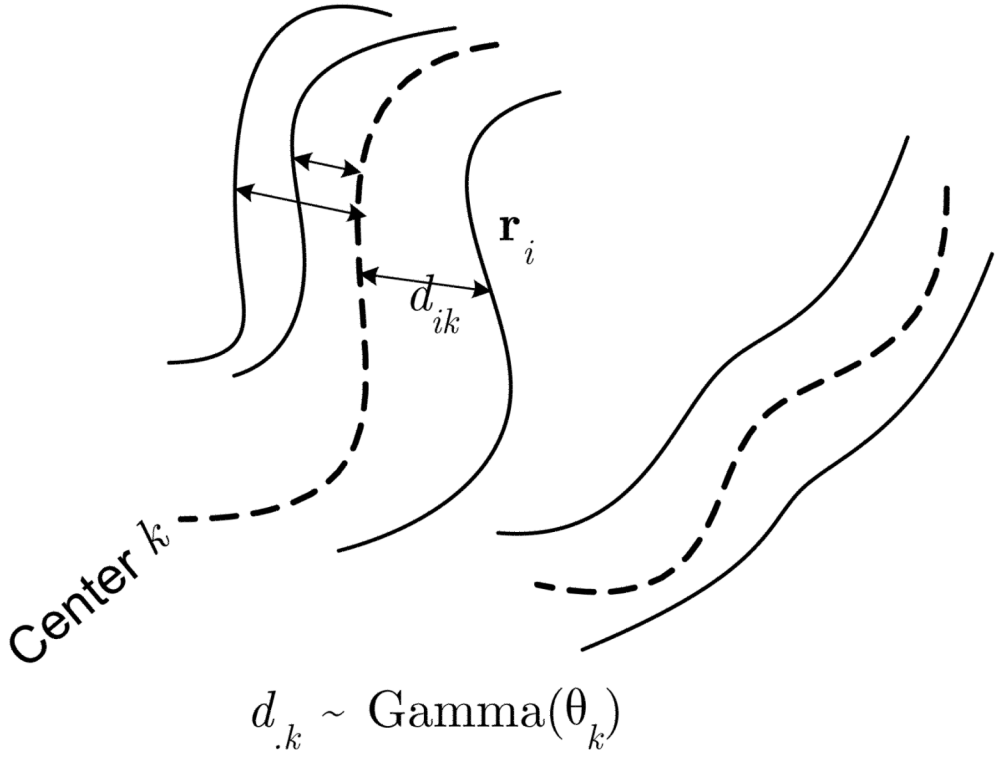


Figure 1.
In our approach the probabilistic model is constructed on distances between the curves(trajectories) and the cluster centers (dashed curves), not on point coordinates.

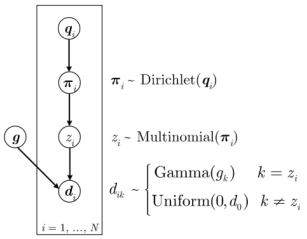
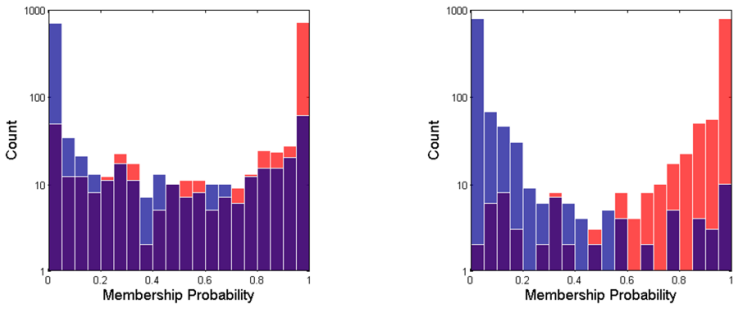


Figure 2.

A graphical model that shows the inference of the priors and cluster parameters form the data when an atlas prior is available that describes the dependencies between different variables. $\boldsymbol{\pi}$: atlas prior, z: trajectory label, **d**: data and **g**: parameters of gamma distributions.



(a) (b)

Figure 3.

A comparison of the clustering results with two degrees of atlas influence and with an atlas prior that agrees with the membership likelihood inferred from the mixture model: (a) $a = 0$, i.e., no atlas, and (b) $a = 1$. Histograms of the membership probability p_{ik}, are shown for synthetic data colored by the true cluster assignment. Misclassification is indicated by the presence of red samples on the left side of the histograms or blue samples on the right. The ratio of mis-clustered data points decreases from 15% to less than 2% when the atlas is used. Note that the y-axis is in logarithmic scale.

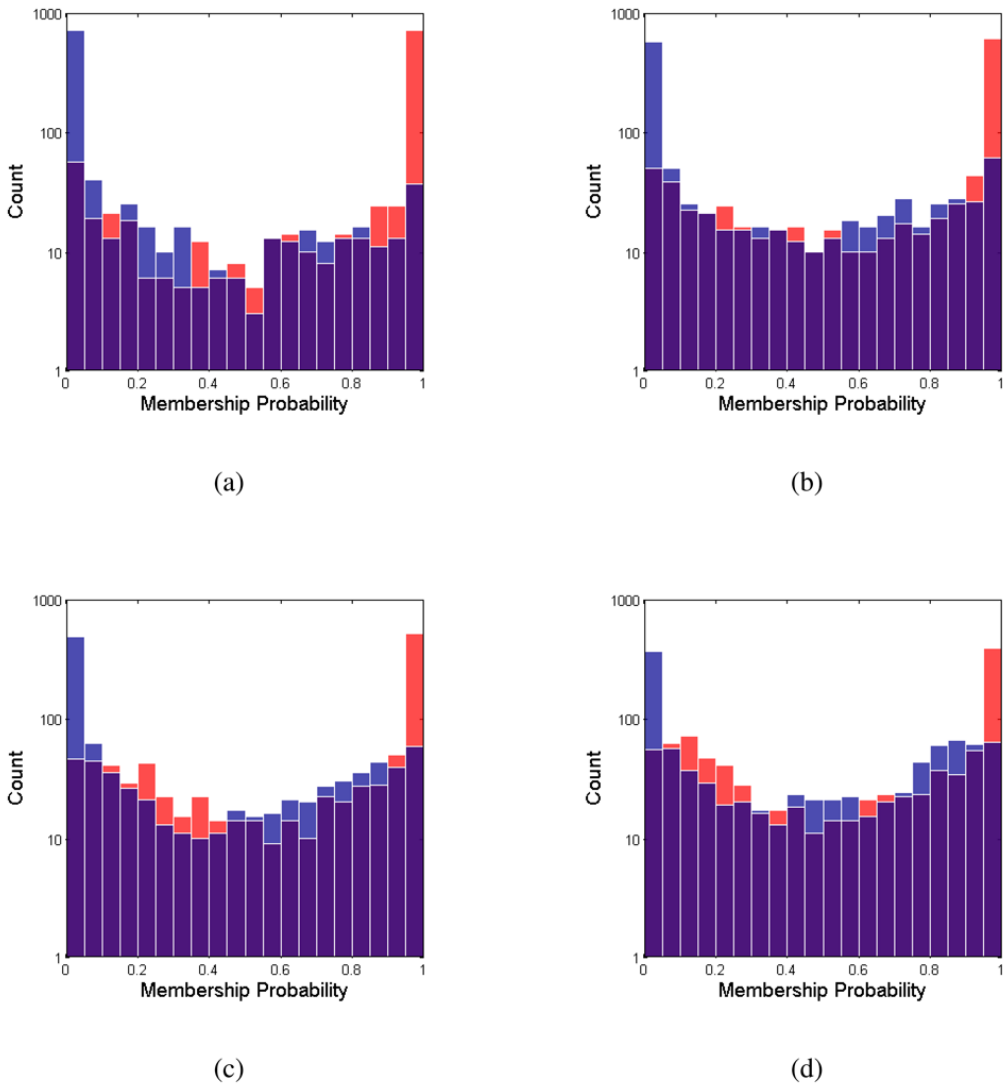


Figure 4.

A comparison of the clustering results with different degrees of atlas influence, but with an atlas prior that disagrees with the membership likelihood inferred from the mixture model: (a) $a = 0$, i.e., no atlas, (b) $a = 0.25$, (c) $a = 0.5$, and (d) $a = 1$. Histograms of the membership probability, p_{ik}s, are shown for synthetic data colored by the true cluster assignment. Misclassification is indicated by the presence of red samples on the left hand side of the histograms. Misclustering ratio increases as the influence of the atlas is increased. Note that the y-axis is in logarithmic scale.

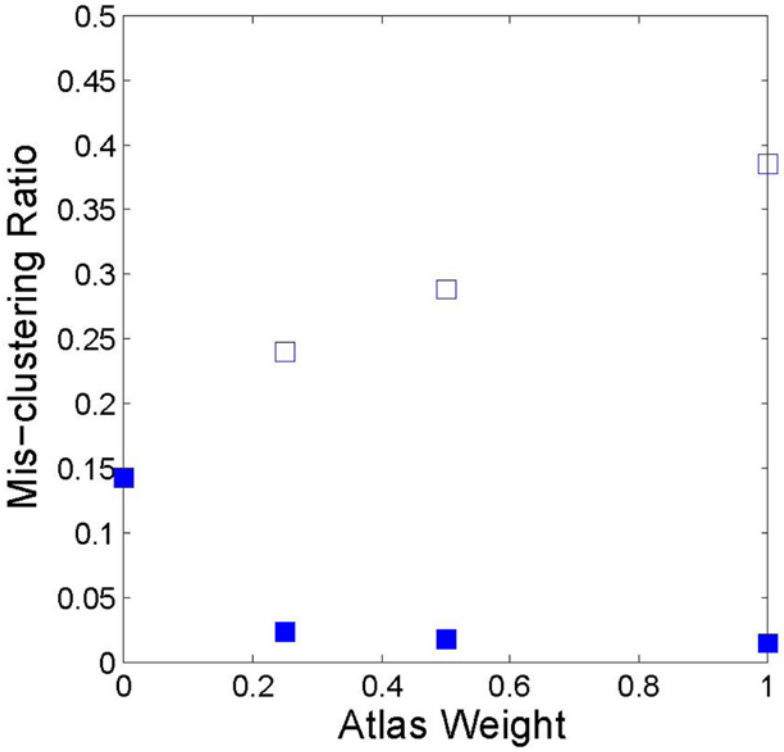


Figure 5.
Effect of atlas weight, *a*, on the mis-clustering ratio and when the atlas prior agrees (filled symbols) of disagrees (open symbols) with the membership likelihood inferred from the mixture model.

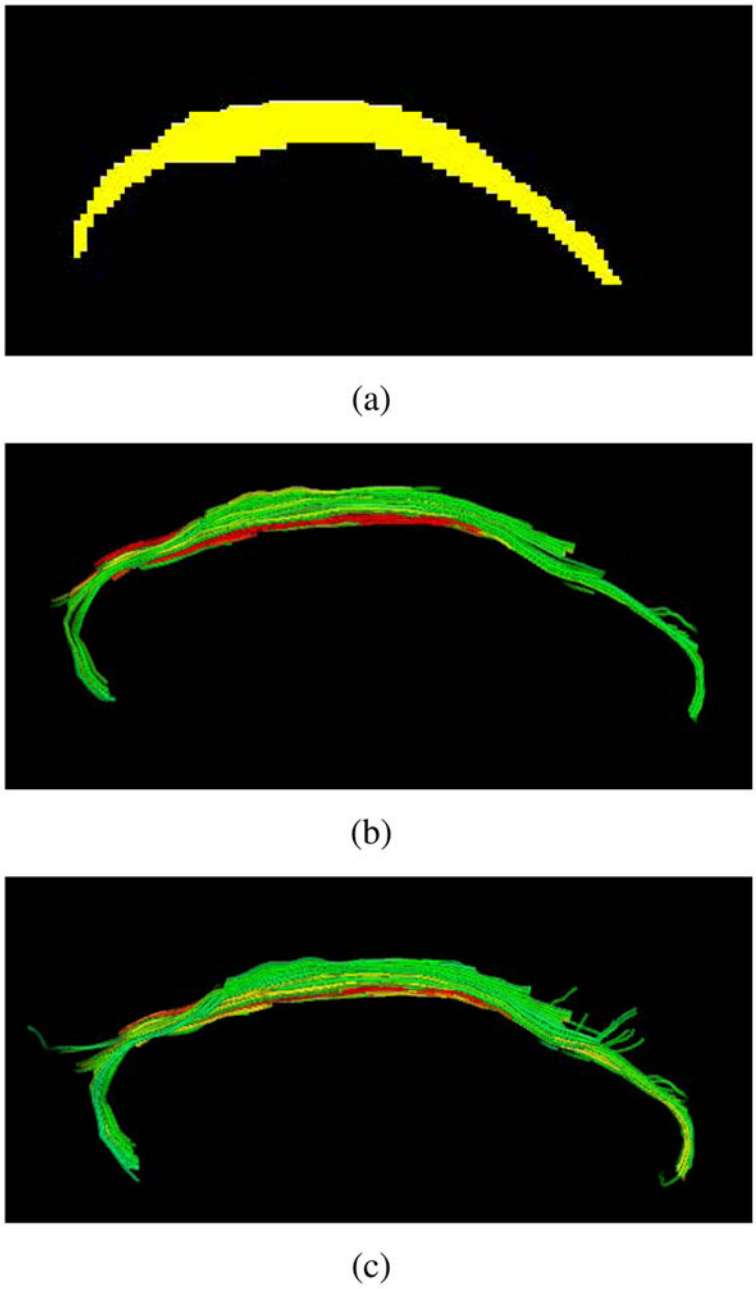


Figure 6.
The influence of atlas weight on the clustering of cingulum trajectories: (a) atlas-specified ROI for cingulum and clustering results with (b) $a = 1$ and (c) $a = 0.5$. When the atlas is imposed strongly some of the trajectories that belong to the cingulum are rejected only because they do not have enough overlap with the atlas ROI.