A MATHEMATICAL FRAMEWORK FOR INCORPORATING ANATOMICAL KNOWLEDGE IN DT-MRI ANALYSIS

Mahnaz Maddah[1,3]*, Lilla Zöllei*[2]*, W. Eric L. Grimson*[1]*, Carl-Fredrik Westin*[1,3]*, William M. Wells*[1,3]

[1] Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
[2] Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Charleston, MA 02129, USA.
[3] Surgical Planning Laboratory, Brigham and Women's Hospital, Boston, MA 02115, USA.

ABSTRACT

We propose a Bayesian approach to incorporate anatomical information in the clustering of fiber trajectories. An expectation-maximization (EM) algorithm is used to cluster the trajectories, in which an atlas serves as the prior on the labels. The atlas guides the clustering algorithm and makes the resulting bundles anatomically meaningful. In addition, it provides the seed points for the tractography and initial settings of the EM algorithm. The proposed approach provides a robust and automated tool for tract-oriented analysis both in a single subject and over a population.

Index Terms— Diffusion Tensor MRI, Clustering, Anatomical Information, Tract-Oriented Quantitative Analysis.

1. INTRODUCTION

Diffusion tensor MR imaging (DT-MRI) has increasingly attracted attention in the neurosurgical community to identify white matter pathologies, minimize post-operative neurological deficit, and study brain development and aging. Pathways of fiber tracts are often extracted using a tractography algorithm to aid the visualization of brain connectivity. However, most clinical studies to date have focused on the analysis of the local parameters measured in a manually defined region of interest (ROI) [1]. Most recently, tract-oriented quantitative analysis has emerged as a complementary tool in clinical studies, mitigating the user-dependence and uncertainty in defining the ROIs [2, 3]. However, tractography methods are also prone to the subjectivity and sensitivity resulting from the selection of seed points [4]. Furthermore, due to the presence of noise and image imperfections, outliers are often generated in tractography.

In order to benefit from the tract-oriented analysis and to eliminate the above mentioned problems of the tractography step, automated clustering of the fiber trajectories is necessary, so that the tractography can be seeded from the whole brain or a sufficiently large ROI. Different methods, mostly unsupervised [2, 5, 6], have thus been proposed to group trajectories into clusters. However, a supervised clustering that benefits from anatomical information not only yields a more robust clustering, that are less sensitive to the presence of outliers and imperfections in DT data, but also produces clusters that are anatomically meaningful. The latter makes the correspondence between clusters across different subjects automatically known.

Earlier attempts to use anatomical information in fiber clustering are limited to an atlas-based clustering where trajectories are grouped based on their distance to the trajectories in a reference subject (atlas) [7] and associating groups of fiber trajectories to anatomical structures after they are clustered and mapped into an embedded space [6]. 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 the quantitative parameters [8]. This approach relies exclusively on the atlas data and hence does not take into account the coherence among the trajectories in each bundle. Furthermore, the quantitative analysis is limited to averages over the entire ROI or slices perpendicular to the main axes of the atlas coordinate system. This is due to the fact that, unlike tract-oriented methods, a curve is not defined as the cluster center.

While the existing methods provide valuable quantitative information about the integrity of fiber tracts, our experience shows that they are either limited to specific fiber bundles or prone to be adversely affected by inaccurate setting of user-specified parameters. Given that an atlas of fiber tracts [9] is available, we propose a probabilistic method to rigorously use the atlas as an anatomical prior. This is an extension of our earlier work, in which an expectation-maximization method is used to cluster the fiber trajectories in a mixture model context [3].

2. DEFINITIONS AND GRAPHICAL MODEL

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, $\mathbf{r}_i$, a vector $\mathbf{d}_i = [d_{i1}, ..., d_{iJ}]$ is calculated, where d_{ij} is the distance between the trajectory and the jth cluster center as a function of their point coordinates and correspondences [10]. J is a user-defined number of clusters. We assume that each element of the vector $\mathbf{d}_i$ follows the distribution

$$d_{ij} \sim \begin{cases} \text{Gamma}(\Theta_j) & j = z_i \\ \text{Uniform}([0, d_0]) & j \neq z_i \end{cases}, \quad (1)$$

where z_i denotes the label assignment of trajectory i, and d_0 is a sufficiently large number. This means that when inferring the parameters for the jth cluster, we only look at the distances to the

The authors would like to thank Dr. S. Mori for the parcellation map of fiber tracts, Dr. M. Shenton and Dr. M. Kubiki for providing the DT images. This work is supported in part by NIH grants P41 RR013218, U54 EB005149, U41 RR019703, P30 HD018655, R01 RR021885, R03 CA126466, P41-RR14075, BIRN002, U24 RR021382, and R21 MH067054 and NIH R01 MH074794, and grant RG 3478A2/2 from the NMSS.

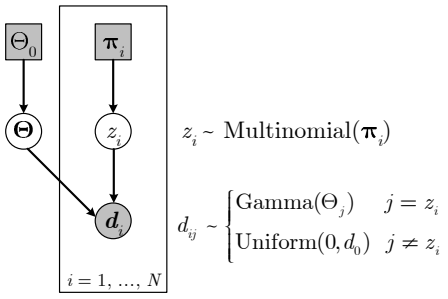


Fig. 1. A Bayesian network that describes the dependencies between different variables. $\boldsymbol{\pi}$: atlas prior, z: trajectory label, $\mathbf{d}$: distance between trajectory and cluster centers, $\boldsymbol{\Theta}$: parameters of the gamma distributions, and Θ_0: an upper bound on these parameters. Shaded nodes represent the observed data.

corresponding cluster center, i.e., the jth component of the vectors $\mathbf{d}_i$. The uniform distributions are added to the definition just to have a complete probability model for $\mathbf{d}_i$'s. The gamma distribution associated with cluster j is defined by $\Theta_j = \{\alpha_j, \beta_j\}$, i.e., its shape and inverse scale parameters. As we previously showed in [10], the gamma distribution well models the nature of the distance between the trajectories and the cluster centers. One could define hyper-parameters Θ_0 to introduce prior information over $\boldsymbol{\Theta}$. Here, we assume a uniform distribution since such information is not available.

In this work, we further assume that the labels follow a multinomial distribution, $z_i \sim \text{Multinomial}(\boldsymbol{\pi}_i)$, where $\boldsymbol{\pi}_i$'s are supplied by an atlas, as described in Section 4. Specifically, $\Pr(z_i = j|\boldsymbol{\pi}_i) = \pi_{ij}$, where $\sum_{j=1}^{J} \pi_{ij} = 1$.

To better understand the dependencies between the variables, Fig.1 shows the directed graphical model of the problem setup. Note that this model can be extended to define a prior distribution for $\boldsymbol{\pi}_i$'s in order to control the influence of the atlas on the clustering.

The goal is to estimate the unknown parameters $\boldsymbol{\Theta}$ by:

$$\hat{\boldsymbol{\Theta}} = \arg\max_{\boldsymbol{\Theta}} \log p(Z, \boldsymbol{\Theta}|D, \Pi), \tag{2}$$

where D and Π are the observed data and the prior information, i.e. the collection of $\mathbf{d}_i$'s and $\boldsymbol{\pi}_i$'s, respectively. We propose to perform the optimization using the expectation-maximization method as described in Section 3.

3. METHOD

Once the trajectories are extracted from the DT data, they are mapped into the atlas coordinate system. Each trajectory $\mathbf{r}_i$ then takes a membership probability $\boldsymbol{\pi}_i = [\pi_{i1}, ..., \pi_{iJ}]$, where each π_{ij} element denotes the atlas-specified membership of $\mathbf{r}_i$ to cluster j. The number of clusters, J, is a subset of anatomical bundles in the atlas. The details of this procedure are given in Section 4. Initial cluster centers are provided by the atlas and the distance vectors are computed [3]. Expectation-maximization [11] is then used to infer the unknown model parameters, $\boldsymbol{\Theta}$, and the membership probabilities in two steps.

In the expectation step, given the current estimates of the parameters, $\boldsymbol{\Theta}^t$, the membership probabilities are calculated.

$$p_{ij} = \Pr(z_i = j|\mathbf{d}_i, \boldsymbol{\Theta}^t) = \frac{\Pr(\mathbf{d}_i|z_i = j, \boldsymbol{\Theta}^t)\Pr(z_i = j|\boldsymbol{\Theta}^t)}{\sum_k \Pr(\mathbf{d}_i|z_i = k, \boldsymbol{\Theta}^t)\Pr(z_i = k|\boldsymbol{\Theta}^t)}.$$

Assuming that Z and $\boldsymbol{\Theta}^t$ are independent, the class assignment is independent of the model parameters:

$$\Pr(z_i = j|\boldsymbol{\Theta}^t) = \Pr(z_i = j) = \pi_{ij},$$

which is supplied by the atlas (See Section 4). So,

$$p_{ij} = \Pr(z_i = j|\mathbf{d}_i, \boldsymbol{\Theta}^t) = \frac{\text{Gamma}(d_{ij}; \Theta_j^t)\pi_{ij}}{\sum_{k=1}^{J} \text{Gamma}(d_{ik}; \Theta_k^t)\pi_{ik}}. \tag{3}$$

The new estimates of the model parameters are updated in the maximization step:

$$\begin{aligned} \boldsymbol{\Theta}^{t+1} &= \arg\max_{\boldsymbol{\Theta}} \left(E_{Z|D,\boldsymbol{\Theta}^t}[\log p(D, Z|\boldsymbol{\Theta})]\right) \\ &= \arg\max_{\boldsymbol{\Theta}} \left(E_{Z|D,\boldsymbol{\Theta}^t}[\log p(D|Z, \boldsymbol{\Theta}) + \log p(Z|\boldsymbol{\Theta})]\right) \end{aligned}$$

Since we assume Z to be independent of $\boldsymbol{\Theta}$, the last term can be eliminated from the maximization. Also, $\mathbf{d}_i$'s are independent and each $\mathbf{d}_i$ only depends on z_i. Hence,

$$\begin{aligned} \boldsymbol{\Theta}^{t+1} &= \arg\max_{\boldsymbol{\Theta}} \sum_{i=1}^{N} E_{Z|D,\boldsymbol{\Theta}^t}[\log p(\mathbf{d}_i|\Theta, z_i)] \\ &= \arg\max_{\boldsymbol{\Theta}} \sum_{i=1}^{N}\sum_{j=1}^{J} \Pr(z_i = j|d_i, \boldsymbol{\Theta}^t) \log p(\mathbf{d}_i|\boldsymbol{\Theta}, z_i = j) \end{aligned} \tag{4}$$

where $\Pr(z_i = j|d_i, \boldsymbol{\Theta}^t)$ is given by (3), calculated in the expectation step. The maximization expression (4) is similar to what we previously derived in the absence of the atlas prior [10] as the prior only appears in the expectation step. So, to update the gamma parameters we have:

$$\alpha_j \approx \frac{3 - x_j + \sqrt{(x_j - 3)^2 + 24x_j}}{12x_j}, \tag{5}$$

where

$$x_j = \log\left(\frac{\sum_i p_{ij} d_{ij}}{\sum_i p_{ij}}\right) - \frac{\sum_i p_{ij} \log(d_{ij})}{\sum_i p_{ij}}. \tag{6}$$

and

$$\beta_j = \alpha_j \sum_i p_{ij} / \sum_i p_{ij} d_{ij}. \tag{7}$$

Once the EM algorithm converges, we update the cluster centers and recompute the distance vectors. The outliers are identified in the

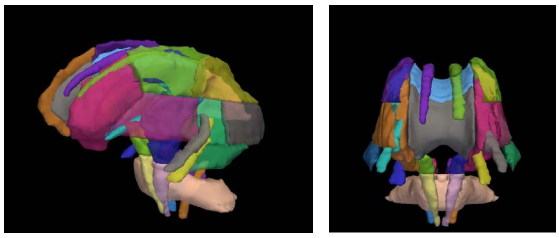


Fig. 2. Visualization of some of ROIs outlined by the atlas. These ROIs correspond to the major anatomical fiber tracts.

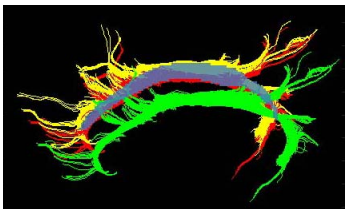


Fig. 3. Demonstration of the registration process for the trajectories of superior cingulum in one of the subjects. Original trajectories (green) and trajectories mapped into the atlas space with either an affine registration (red) or a congealing registration to a common space for all subjects and then by an affine registration into the atlas space (yellow) are shown. The spatial extent of the cingulum specified by the atlas is also shown for comparison.

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 is controlled. The larger the threshold is, the more compact the resulting bundles are, and consequently the greater the number of unclustered trajectories.

4. EXPERIMENTS AND RESULTS

We use the atlas constructed by Mori *et al.* (http://lbam.med.jhmi.edu), which consists of 48 labeled regions that correspond to major anatomical bundles of fiber tracts in the human brain. Visualization of some of these regions in the 3D Slicer (`www.slicer.org`) is shown in Fig. 2. To allow a probabilistic assignment at the region boundaries, we apply a Gaussian kernel with a $3 \times 3 \times 3$ window with standard deviation of 2 to each region.

EPI DT-MR images were acquired from healthy volunteers as well as Schizophrenia patients on a 3T scanner. DT data were reconstructed from 5 baseline and 51 gradient images and a spatial resolution of $0.93 \times 0.93 \times 1.7$ mm.

Trajectories are extracted for each subject using a streamline tractography method [12] and mapped into the MNI atlas space. Seed points for tractography are provided by the mapped and dilated regions from the atlas to each subject's space. Registration is performed on the corresponding maps of the fractional anisotropy (FA) to normilize for brain geometry, and then the obtained transformation is applied to the trajectories. An affine registration [13] usually gives satisfactory results as reflected in Fig. 3. However, for population studies we opted to first map the subjects into a common space using the congealing algorithm [14] followed by the affine registration to the MNI space. We decided to use this approach, as opposed to a series of pair-wise subject-template registration, in order to avoid introducing bias in the population analysis. Figure 3 shows the results of registering the trajectories from the superior cingulum to the atlas space for one of the subjects.

With the trajectories projected to the atlas space, the membership probability for each trajectory, π_{ik}'s, 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. This insures that the results are not biased towards large tracts, such as the corpus callosum. The membership probabilities of each trajectory are then normalized, so that $\forall i, \sum_{j=1}^{K} \pi_{ij} = 1$. Note that with this implementation, the atlas

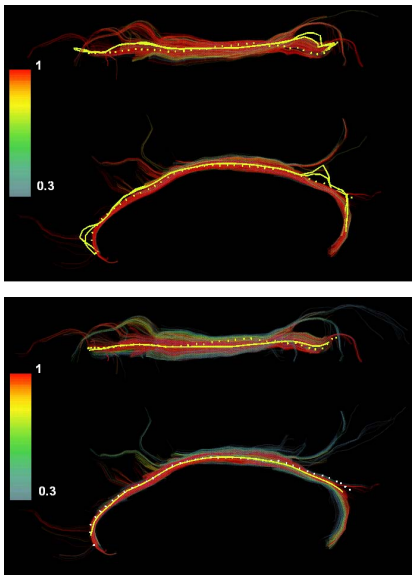
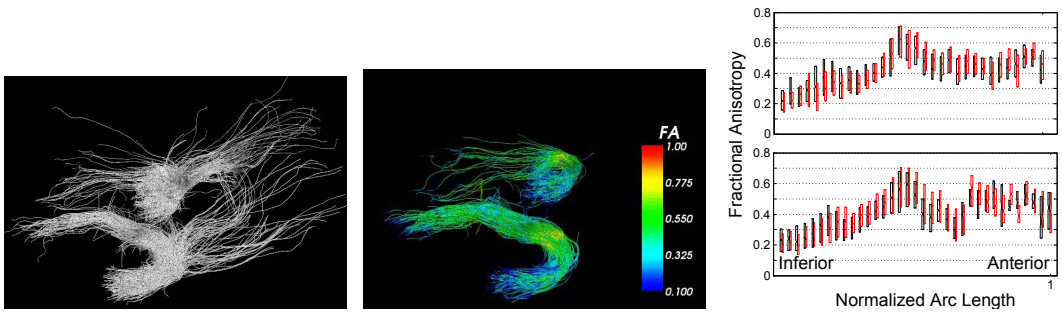


Fig. 4. A comparison of the clustering results without (top) and with (bottom) atlas. Axial and sagittal views are shown for superior cingulum. Trajectories are colored by their assigned membership probability. The cluster centers at consecutive EM iterations are shown in yellow, and dotted lines represent the initial centers. Without the atlas, and with improper setting of the clustering parameters the cluster centers drift and its extent increases as the algorithm proceeds. Less sensitivity to parameter setting and robustness is achieved with the atlas incorporated.

ROIs provide the spatial prior, while the information about the shape and orientation of the tracts are captured by a representative curve for each bundle, used as the initial center. Unlike a voxel-based method in which individual voxels (of each trajectory) receive their own membership probability [8], in our approach the probability is assigned to the trajectory and hence the method is less sensitive to local errors in registration.

Figure 4 compares the clustering results obtained with and without incorporating the atlas prior for the superior cingulum. The trajectories are colored based on their membership probability in each case. To emphasize the effect of the atlas, a worst-case scenario is presented in which the parameters that control the extent of the clusters are set such that the algorithm only excludes those trajectories that receive very small membership probability. Without the atlas, the algorithm gives moderate membership probability to those trajectories that are not very close to the initial center. However, as the algorithm proceeds, the cluster center drifts as shown in Fig. 4 (a), so that these trajectories receive higher and higher membership probabilities. Even though the clustering results might be still acceptable, the cluster center is deformed, introducing significant error in the quantitative analysis. If the parameters are set in the correct range, the extent of the cluster can be controlled in order to prevent the inclusion of outliers and excessive drift of the cluster center. A trade-off should be made between the homogeneity of the clustered trajectories and the number of unclustered trajectories. With the atlas prior included, more robust results are obtained, less sensitive to incorrect user settings.

Figure 5 shows an example of a population study that consists of clustering and quantitative analysis performed on the uncinate fas-



(a) (b) (c)

Fig. 5. A 3D view of the uncinate fasciculi (UNC) trajectories, left and right, from 26 subjects mapped into the atlas space and an example of the quantitative analysis. (a) Tractography results seeded from UNC ROI. (b) Clustering results colored by the local fractional anisotropy. (c) A box-plot showing the FA variation along the tract for healthy (black) and schizophrenia (red) cases and for the right (up) and left (down) UNC. Note that (c) quantifies the trend of FA variations, visualized in (b).

ciculus (UNC) fiber tracts of 13 healthy and 13 schizophrenia subjects. The tractography on each subject was seeded from the dilated UNC ROI, mapped from the atlas to the subject's space. Figure 5(a) shows the trajectories from all cases mapped back to the atlas space. The clustering results are shown in Fig. 5(b), where the trajectories are colored based on their local FA. With the point correspondence between the trajectories and the cluster centers, the tract-oriented analysis is easily performed by calculating a weighted average of the quantitative parameters along the cluster center. An example is shown in Fig. 5(c), which shows box-plots of the FA variation along the cluster center for the left and right UNC and for healthy and schizophrenia cases. Only the superior part of the UNC shows significant difference between the left and right structures. No significant difference is seen between control and pathological cases.

5. CONCLUSIONS

We have proposed a novel probabilistic framework that incorporates an atlas of fiber tracts as anatomical prior in clustering of fiber trajectories. To the best of our knowledge, such a framework has not been previously reported. The supervision of anatomical information combined with sophistication of the similarity-based clustering yields a robust method for accurate tract-oriented quantitative analysis. An important aspect of our method is that it does not only take into account the spatial prior, but also the shape of the fiber bundles. An example was presented to demonstrate the applicability of the proposed method for population studies.

6. REFERENCES

[1] H. J. Park, C.-F. Westin, M. Kubicki, S. E. Maier, M. Niznikiewicz, A. Baer, M. Frumin, R. Kikinis, F. A. Jolesz, R. W. McCarley, and M. E. Shenton, "White matter hemisphere asymmetries in healthy subjects and in schizophrenia: A diffusion tensor MRI study," *NeuroImage*, vol. 24, pp. 213–223, 2004.

[2] I. Corouge, P. T. Fletcher, S. Joshi, S. Gouttard, and G. Gerig, "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *Med. Image Anal.*, vol. 10, pp. 786–798, 2006.

[3] M. Maddah, W. M. Wells, S. K. Warfield, C-F Westin, and W. E. L. Grimson, "Probabilistic clustering and quantitative analysis of white matter fiber tracts," in *IPMI*, 2007, pp. 372–383.

[4] J. D. Clayden, A. J. Storkey, and M. E. Bastin, "A probabilistic model-based approach to consistent white matter tract segmentation," *IEEE Trans. Medical Imaging*, vol. 26, pp. 1555–156, 2007.

[5] A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C.-F. Westin, "Clustering fiber tracts using normalized cuts," in *MICCAI*, 2004, pp. 368–375.

[6] L. O'Donnell and C.-F. Westin, "White matter tract clustering and correspondence in populations," in *MICCAI'05*, 2005, Lecture Notes in Computer Science 3749, pp. 140–147.

[7] M. Maddah, A. Mewes, S. Haker, W. E. L. Grimson, and S. K. Warfield, "Automated atlas-based clustering of white matter fiber tracts from DTMRI," in *MICCAI*, 2005, pp. 188–195.

[8] K. Hua, J. Zhang, S. Wakana, H. Jiang, X. Li, D. S. Reich, P. A. Calabresi, J. J. Pekar P. C.M. van Zijl, and S. Mori, "Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification," *NeuroImage*, vol. 39, pp. 336–347, 2008, to be published.

[9] S. Wakana, H. Jiang, L. M. Nagae-Poetscher, P. C. M. van Zijl, and S. Mori, "Fiber tract-based atlas of human white matter anatomy," *Radiology*, vol. 230, pp. 77–87, 2004.

[10] M. Maddah, W.E.L. Grimson, S.K. Warfield, and W.M. Wells, "A unified framework for clustering and quantitative analysis of white matter fiber tracts," *Med. Image Anal.*, in press.

[11] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society, Series B*, vol. 39, pp. 138, 1977.

[12] P.J. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, "In vivo fiber tractography using DT-MRI data," *Magn. Reson. Med.*, vol. 44, pp. 625–632, 2000.

[13] W.M. Wells, P. Viola, H. Atsumi, S. Nakajima, and R. Kikinis, "Multi-modal volume registration by maximization of mutual information," *Medical Image Analysis*, vol. 1, pp. 35–51, 1996.

[14] L. Zöllei, E. Learned-Miller, E. Grimson, and W.M. Wells, "Efficient population registration of 3D data," in *Workshop on Computer Vision for Biomedical Image Applications: Current Techniques and Future Trends*, 2005.