Automated Atlas-Based Clustering of White Matter Fiber Tracts from DTMRI

Mahnaz Maddah[1,2], Andrea U. J. Mewes[2], Steven Haker[2,3], W. Eric L. Grimson[1], and Simon K. Warfield[1,2,3,4]

[1] Computer Science and Artificial Intelligence Laboratory, Massachussets Institute of Technology, Cambridge, MA 02139, USA.
[2] Computational Radiology Laboratory, Brigham and Women's Hospital, Boston, MA 02115, USA.
[3] Surgical Planning Lab., Brigham and Women's Hospital, Boston, MA 02115, USA.
[4] Children's Hospital and Brighham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. *
{mmaddah, welg}@mit.edu, {mewes, haker, warfield}@bwh.harvard.edu

Abstract. A new framework is presented for clustering fiber tracts into anatomically known bundles. This work is motivated by medical applications in which variation analysis of known bundles of fiber tracts in the human brain is desired. To include the anatomical knowledge in the clustering, we invoke an atlas of fiber tracts, labeled by the number of bundles of interest. In this work, we construct such an atlas and use it to cluster all fiber tracts in the white matter. To build the atlas, we start with a set of labeled ROIs specified by an expert and extract the fiber tracts initiating from each ROI. Affine registration is used to project the extracted fiber tracts of each subject to the atlas, whereas their B-spline representation is used to efficiently compare them to the fiber tracts in the atlas and assign cluster labels. Expert visual inspection of the result confirms that the proposed method is very promising and efficient in clustering of the known bundles of fiber tracts.

1 Introduction

Diffusion tensor MR imaging (DT-MRI) non-invasively measures the diffusivity of water molecules within the tissue [1]. It thus provides some knowledge about the direction and density of the fiber tracts in the brain, as the water diffusion is restricted in the direction normal to the fibers. While anisotropy measures are being used to assess the density of the fiber tracts and or the degree of myelination in different regions of the brain, tractography methods have been

* The authors would like to thank Dr. Polina Golland and Dr. Kilian Pohl for their helpful discussions and suggestions. This investigation was supported in part by NSF ITR 0426558, a research grant from the Whitaker Foundation, a research grant from CIMIT, grant RG 3478A2/2 from the NMSS, and by NIH grants R21 MH67054, R01 LM007861, R01 CA109246, P41 RR13218, 1U54 EB005149, and P01 CA67165.

developed to track the fibers (e.g. [2]), extract the fiber bundles, and estimate the interconnectivity of different regions of the brain [3].

To date, significant work has been devoted to providing tools for processing the images [4], tractography schemes [2, 4], visualization [5], and recently quantitative analysis of the extracted tracts [6]. However, limited work has been done to address the inter-subject similarity/variability of the fiber tracts in the human brain, which is our ultimate goal. Toward this goal and to quantitatively analyze the fiber tracts in a population, the first step is to identify the fiber tract clusters in each case. This process would be very inefficient if an expert were needed to specify all regions of interest in every case under study. An alternative approach is to extract all of the tracts in the whole brain and then automatically cluster the desired bundles.

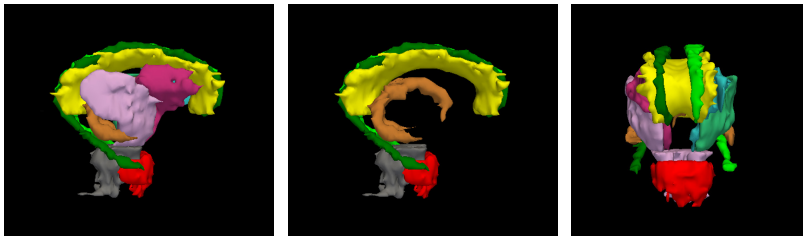
Clustering of fiber tracts has attracted significant attention recently. The tracts have been successfully clustered into bundles using spectral clustering [7, 8], normalized cuts [1], agglomerative hierarchical clustering [9], fuzzy c-means [10], k-nearest neighbors [11]. However, in the absence of any expert knowledge or atlas to assign the actual anatomical structures to the clusters, there is no guarantee that a particular clustering algorithm yields the desired bundles. The algorithm might easily over- or under-cluster the fiber tracts. So, a clustering scheme is desired that is able to automatically cluster large numbers of tracts to the bundles of interests, and takes into account available anatomical knowledge. Such an algorithm should be also flexible enough to cluster either to the whole anatomical structure or into sub-structures. Accuracy and time-efficiency are also required as in any other application.

In this paper, we present a new method which satisfies the above requirements. We propose to use a labeled atlas of the fiber tracts of the whole brain to perform clustering on the subjects under study. The essence of the proposed method is as follows: First, all tracts in the white-matter tissue are extracted and then projected to the atlas. The applied transformation matrix is obtained by registering the baseline image of the subject (the images obtained without applying any gradient in the magnetic field) to that of the atlas. The final step is to assign a cluster label to each projected tract. This is done by taking the cluster label of the most similar tract in the atlas. The details of each step will be described in the next sections. The preliminary results show the efficiency of the proposed method. The next section describes how we develop the atlas, while the clustering method is explained in Section 3 along with the results obtained at each step. Conclusions and future work are discussed in Section 4.

2 Constructing the Fiber Tract Atlas

2.1 Specifying the Regions of Interest

High-resolution DT-MRI data were achieved from healthy volunteers on a 3T General Electrics Signa scanner (Milwaukee, WI) with repetition time/echo time of 2500/70 ms, b-factors 50 of 5 and 750 s/mm2, and six gradient directions. Raw



(a) (b) (c)

Fig. 1. Sagittal (a),(b) and anterior (c) view of the ROIs for some of the fiber tract bundles specified by an expert. The internal capsula are removed in (b) to show the inner structures.

data were converted to derive tensor information, and were loaded into 3D slicer (www.slicer.org). The region of interests (ROIs) were outlined by an expert based on information from the map of the fractional anisotropy (FA), and a map, which color-codes the direction of the main eigenvector, both generated in the slicer. The expert specified major association, projection and commissural tracts as the ROIs which are used as starting points for tractography (See Fig. 1).

2.2 Tractography

The tractography approach we implemented is categorized as a stochastic approach. Various stochastic tractography methods have been proposed to alleviate the shortcomings of deterministic approaches in crossing regions by adding some randomness to the deterministic tractography. Instead of defining an ad-hoc stochastic component, we used the assumption of a Gaussian distribution for finding a particle at a distance $\mathbf{r}$ from its initial position when diffusing in an anisotropic medium, as described by Basser [12]. If we represent the diffusion tensor at each voxel, $\mathbf{r}$, by $\mathbf{D}(\mathbf{r})$, with eigenvalues and corresponding eigenvectors denoted by λ_i and $\mathbf{E}_i$ for $i = 1, 2, 3$, we get a Gaussian displacement distribution [12]:

$$\lim_{\mathbf{r}\to\infty} \Pr(\mathbf{r}, \Delta) = \frac{1}{\sqrt{|\mathbf{D}|(4\pi\Delta)^3}} \exp(-\frac{\mathbf{r}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{r}}{4\Delta}) \tag{1}$$

where Δ is the diffusion time. Diagonalization of the tensor $\mathbf{D}$ gives:

$$\lim_{\mathbf{r}\to\infty} \Pr(\mathbf{r}, \Delta) = \frac{1}{\sqrt{\lambda_1\lambda_2\lambda_3(4\pi\Delta)^3}} \exp(-\frac{r_1^2}{4\Delta\lambda_1} - \frac{r_2^2}{4\Delta\lambda_2} - \frac{r_3^2}{4\Delta\lambda_3}) \tag{2}$$

where the r_i's are the components of $\mathbf{r}$ in the base described by the $\mathbf{E}_i$'s. This means that the stochastic term is composed of three Gaussian distribution functions along each $\mathbf{E}_i$, whose variance is equal to the corresponding λ_i.

Using the proposed tractography method, we extract all the tracts starting from all voxels of the ROIs obtained in section 2.1. Fig. 2 shows the extracted

(a) (b) (c)

Fig. 2. The constructed atlas of fiber tracts shown for the obtained bundles in the sagittal (a),(b), and anterior view. Internal capsula's tracts have been removed in (b) to show the inner bundles.

fiber tracts, colored based on the cluster to which they belong. This data, i.e. the labeled fiber tracts, is used as the reference model or atlas throughout the work.

3 Clustering Framework

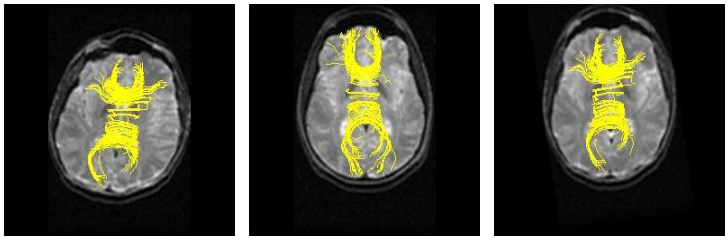
Having constructed the atlas, the whole brain fiber tracts are extracted from DT-MRI of each subject and then will be clustered using the atlas as will be described in this section. The method was applied to 10 DT-MRI data sets acquired from healthy volunteers with voxel size of $1.054 \times 1.054 \times 2$ mm.

3.1 Whole Brain Tractography

The white matter of the whole brain is segmented using a k-NN algorithm [13], and used as the starting point for the tractography. The tracts are then extracted through the stochastic approach described in Section 2.2. From each starting point, two tracts are extracted in opposite directions and connected together to form the complete fiber tract. The tractography algorithm stops for each tract when an FA value of less than 0.2 is reached.

3.2 Projection

The second step is to project the extracted tracts to the reference subject (atlas). To obtain the transformation matrix, we register the baseline images of the subject to that of the atlas. Fig. 3 shows the results of affine registration using a block matching algorithm [14] for one of the subjects under study.



(a) (b) (c)

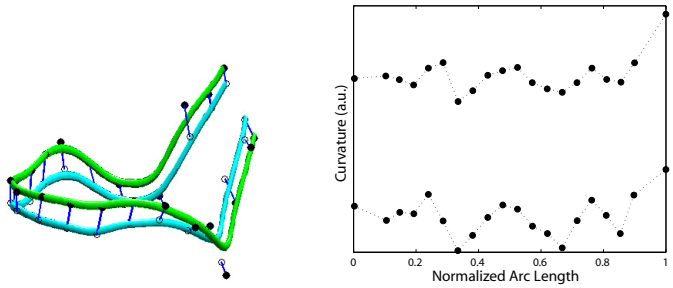
Fig. 3. (a) A set of fiber tracts extracted from one of the subjects under study, overlaid on an axial slice of its baseline image. (b) Same as (a) for the reference subject. (c) The projected tracts overlaid on the registered baseline image of the subject. The pairwise comparison is performed between tracts in (b) and (c) to assign cluster labels to the tracts in the subject.

3.3 Clustering

The main and final step is to assign each projected tract to the proper cluster, considering the fact that the tracts of the atlas are already labeled based on the ROI to which they belong (Sec. 2.2). There are several possibilities: The simplest idea is to use the nearest neighborhood technique such that for each projected tract, the nearest tract in the atlas is identified and its cluster label is assigned to the projected tract. The nearest tract can be defined as the tract which minimizes either the closest point distance, the mean of the closest distance, or the Hausdorff distance between the two tracts [6]. Such methods are usually computationally expensive. An alternative approach is to use a curvature and torsion of the 3D curves (tracts), which has the advantage of being invariant to translation and rotation [15]. Such approaches in fact compare a shape model of the curves and neglect the spatial distance between the two curves. Thus, this is not be the best choice for our purpose where spatial location of the tracts is also important. Also, computing the curvature and torsion requires 3rd order differentiation of the data, which is prone to noise and quantization errors. Quintic splines are sometimes used to smooth out the data and to perform the differentiation through convolution with spline kernels instead of simple differentiation [15].

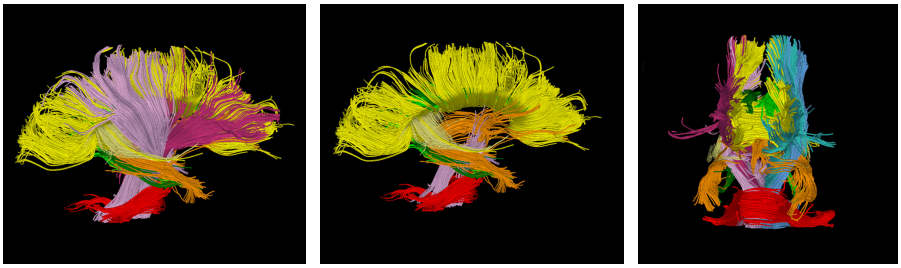
In our clustering approach, we use the B-spline representation [16] of the fiber tracts for pairwise comparison of the fiber tracts extracted from the subject to those from the atlas. The spline representation contains information of both spatial location and shape of the tracts. We used quintic splines to preserve the information needed to calculate the curvature and torsion of the 3D curves. A 20-knot representation proved sufficient. Figure 4(a) shows the spline representation of two neighboring tracts from the corpus-callosum, while Figure 4(b) shows the corresponding curvatures computed along the tracts.

Once the spline representation is computed, each tract from the subject is compared to the tracts in the atlas. In doing so, a hierarchical scheme was used



(a) (b)

Fig. 4. (a) Two neighboring fiber tracts and their corresponding 3D spline representation. Knot points are marked with circles and connected with a short line to guide the eye. (b) The corresponding curvature along each of the two tracts calculated based on their spline representation. The curvature values are shifted with an arbitrary offset for ease of comparison.



(a) (b) (c)

Fig. 5. Sagittal (a),(b) and anterior (c) view of the clustered fiber tracts in of one of the subjects under study.

so that the projected tracts are first compared with the major bundles in the atlas and then with the smaller ones. For the atlas shown in Fig. 2, this leads to a significant drop in computation time once the corpus-callosum and capsula-interna bundles are clustered. The maximum number of bundles clustered is equal to the number of labeled bundles in the atlas. Also, note that since our atlas does not include all of the structures, some tracts are left unclustered. These tracts can be further clustered with a non-atlas based algorithm to reveal those structures not present in the atlas, or the atlas can be extended to include more structures. Fig. 5 shows different views of the fiber tracts from a subject clustered to the bundles based on the atlas of Fig. 2. Finally, Fig. 6 shows the sagittal view of the clustered corpus-callosum bundle in four different subjects.

In our framework it is also possible to include shape information, such as curvature and torsion when clustering the bundles. For example, for a structure

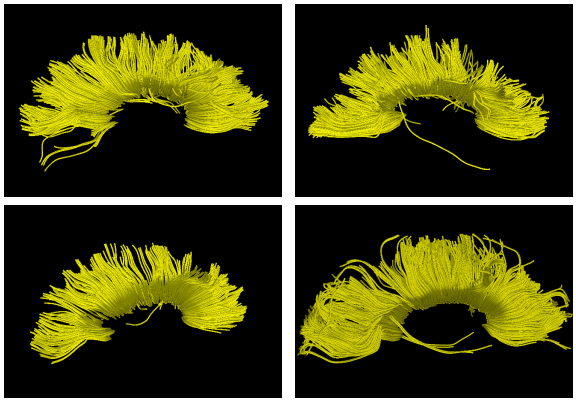
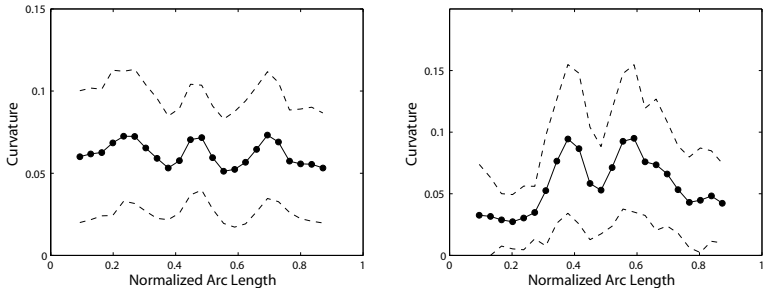


Fig. 6. Sagittal view of the clustered corpus callosum in four different subjects.



(a) (b)

Fig. 7. The average and variation of the curvature along the tracts for the (a) corpus callosum and (b) pons. Dotted lines represent the average ± the standard deviation.

such as the corpus callosum that can be described by a single shape model [17], the tracts with such shape characteristics can be clustered based on their shape representation. This can be performed even without the need to project the subject tracts to the atlas and by comparing only to a single shape representation obtained as the average shape of the tracts for that particular bundle in the atlas. Fig. 7 shows the curvature along the tract for the corpus callosum and the pons.

4 Conclusions and Future Work

In this paper, we presented for the first time an atlas-based clustering framework for fiber tracts of the whole human brain. Preliminary results prove the efficiency of the proposed method to cluster the fiber tracts into anatomically known bundles. The correspondence of clusters in different subjects is defined by default thus unlike most clustering methods proposed for this application, no

post-processing is required. The proposed framework has also the flexibility to use different similarity measures, such as spatial distance or shape similarity or a combination of them, for different structures.

Since the atlas has a significant impact on the clustering results, modification and improvement of the atlas, in terms of accuracy of ROIs, and the inclusion of more cases when constructing the atlas are of great importance. Quantitative analysis and study of the variability of particular fiber tract bundles in a population is currently underway.

References

1. Brun, A., Knutsson, H., Park, H.J., Shenton, M.E., Westin, C.F.: Clustering fiber traces using normalized cuts. In: MICCAI. (2004) 368–375
2. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A.: In vivo fiber tractography using DT-MRI data. Magn. Reson. Med. **44** (2000) 625–632
3. Behrens, T.E.J., et *al*.: Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature Neuroscience (2003) 750–757
4. Westin, C.F., Maier, S.E., Mamata, H., Nabavi, A., Jolesz, F.A., Kikinis, R.: Processing and visualization of diffusion tensor MRI. Med. Image Anal. **6** (2002) 93–108
5. Fillard, P., Gerig, G.: Analysis tool for diffusion tensor MRI. In: MICCAI. (2003) 967–968
6. Gerig, G., Gouttard, S., Corouge, I.: Analysis of brain white matter via fiber tract modeling. In: Proc. IEEE Int. Conf. EMBS. (2004)
7. Jonasson, L., Hagmann, P., Thiran, J.P., Wedeen, V.J.: Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering. In: Proc. ISMRM. (2005) 1310
8. O'Donnell, L., Westin, C.F.: White matter tract clustering and correspondence in populations. In: MICCAI. (2005) to be published.
9. Zhang, S., Laidlaw, D.H.: DTI fiber clustering in the whole brain. In: IEEE Visualization. (2004) 28p
10. Shimony, J.S., Snyder, A.Z., Lori, N., Conturo, T.E.: Automated fuzzy clustering of neur-sonal pathways in diffusion tensor tracking. In: Proc. ISMRM. (2002) 10
11. Ding, Z., Gore, J.C., Anderson, A.W.: Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. Mag. Reson. Med. **49** (2003) 716–721
12. Basser, P.J.: Relationships between diffusion tensor and q-space MRI. Magn. Reson. Med. (2002) 392–397
13. Warfield, S.K.: Fast k-NN classification for multichannel image data. Pattern Recognition Lett. **17** (1996) 713–721
14. Ourselin, S., Roche, A., Prima, S., Ayache, N.: Block matching: A general framework to improve robustness of rigid registration of medical images. In: MICCAI. (2000) 557–566
15. Kehtarnavaz, N., deFigueiredo, R.J.P.: A 3-D contour segmentation scheme based on curvature and torsion. IEEE Trans. Pattern Analysis and Machine Intell. **10** (1988) 707–713
16. Cohen, F.S., Huang, Z., Yang, Z.: Invariant matching and identification of curves using B-splines curve representation. IEEE Trans. Image Proc. (1995) 1–10
17. Corouge, I., Gouttard, S., Gerig, G.: Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In: Int. Symp. Biomed. Imag. (2004) 344–347