DTI Fiber Clustering in the Whole Brain

Song Zhang David H. Laidlaw

{sz,dhl}@cs.brown.edu
Department of Computer Science, Brown University, Providence, RI

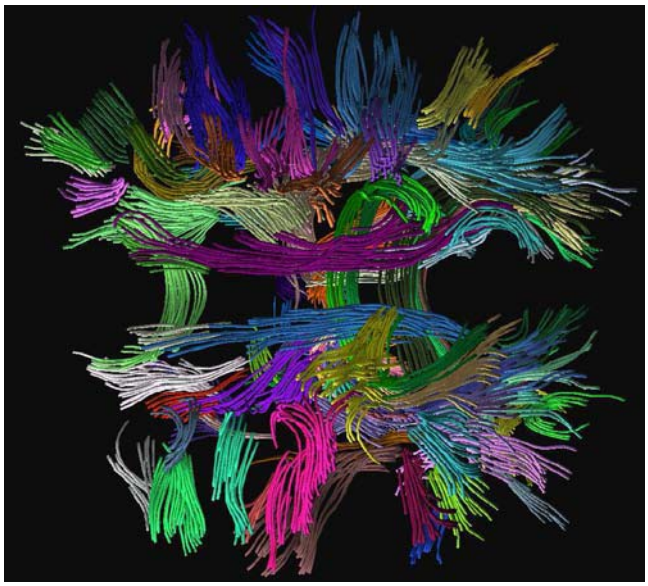

Figure 1: Top view of the streamline clusters on the whole brain. Streamlines within the same cluster share the same color. From the picture, the cingulum bundles can be easily identified in two clusters. Neural fibers along the corpus callosum are clustered into coherent bundles. Some of the U-fibers also form clusters.

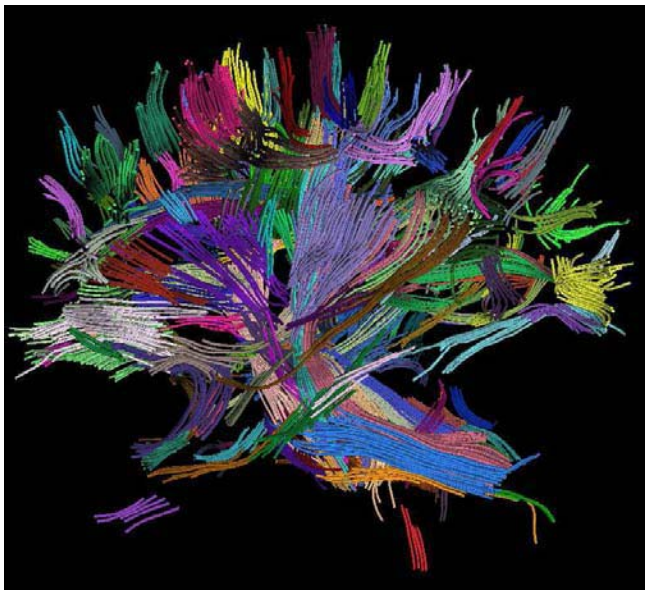

Figure 2: Side view of the same fiber bundles as in Fig. 1. Internal capsule is clustered into several coherent bundles.

Introduction

DTI tractography methods generate streamlines that correlate with the orientations of the neural fibers in the white matter. Researchers studying neuroanatomy label white matter with fiber tracts comprised of groups of neural axons running in proximity to one another. These tracts provide a useful abstraction of the white matter structures; they can be used for identification and quantification of neural fibers. However, due to variation across subjects in white matter, labeling these fiber tracts by hand can be time-consuming and error-prone. Taking advantage of the proximity property, we cany automatically cluster the fiber tracts from the path set generated from a brain DTI data set. The task becomes more challenging when we cluster the paths over the entire brain instead of a region of interest since no human input of the tract information will be implied. We present our experiments in unsupervised clustering of a dense set of fiber paths generated from a whole brain DTI data set.

Related Work

Ding *et al.* [1] proposed DTI fiber classification and quantification. They defined a corresponding segment ratio and employed that ratio together with the mean distance over the corresponding segments to delineate the similarity between two streamlines. The similar streamlines whose seeding points are near the original streamline seeding point are grouped into a bundle.

Corouge *et al.* [2] proposed a clustering algorithm that propagate the cluster to the neighboring fibers. They also employed three distance measures based on the point sets on the pair of streamlines.

We used a similar clustering algorithm on a dense set of paths over the whole brain. Early results showed that a lack of paths within a tract or spurious paths between tracts can lead to possible false classifications. We propose strategies in setting seeding points, path constraints, culling distance and setting distance metrics to minimize the misclassifications.

Method

Generating paths

The head of a normal volunteer was imaged in a Siemens Symphony 1.5T scanner. Three slice packets were acquired sagittally and interleaved to acquire a data volume of $128 \times 128 \times 90$ with a voxel size of $1.7 \times 1.7 \times 1.7$ mm. The Siemens MDDW protocol was used, with three b values (0, 500, 1000) in 12 directions.

We put seeding points on a regular grid every 0.85 mm with small jittering. When the seeding distance is below 1.3 mm, doubling the number of the seeding points results in an almost constant increase in the number of paths. This indicates that our sampling of the paths in the white matter is sufficiently dense to avoid missing links between paths in the same tracts.

We used the streamtube algorithm [3] to generate streamlines from these seed points. The integration stops when the linear

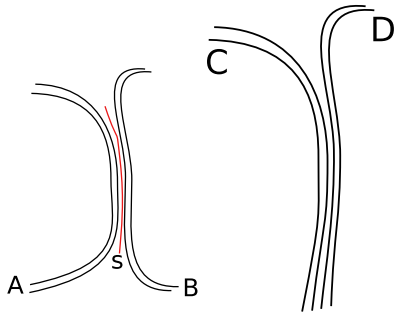

Figure 3: Two potential cases of fiber tract misclassifications using simple clustering schemes. On the left picture, fiber tracts A and B are linked together by short path s, which stops inside an incoherent white matter region. On the right, fiber tracts C and D are clustered together because they are close for a large portion of their length. We design our path culling process and distance metric to avoid misclassifications in these cases.

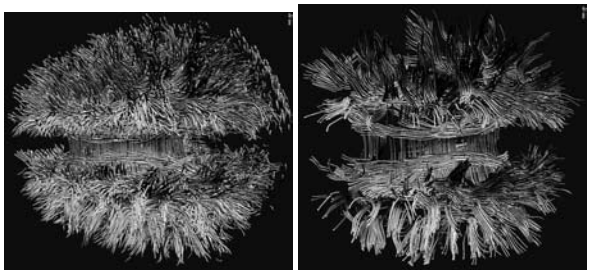

Figure 4: Preprocessing of the streamlines for clustering. On the left, a dense path set representing all regions of linear anisotropy without constraint. On the right, the path set after culling the paths that have low average linear anisotropy, that are too similar, and that do not project into the gray matter.

anisotropy value is below 0.15, the streamline transcends the data boundary, or the signal-to-noise ratio provided by T2-weighted image is below a certain threshold. The initial path set is shown in Fig. 4 left. Since some of the paths might cause artificial linking as shown in Fig. 3 left, we set a threshold on the average linear anisotropy along the path in order to remove paths that are in ambiguous regions. We also remove the paths that are too similar. Some paths stop within ambiguous regions in the white matter. They can be considered spurious paths and might cause artificial linking between anatomically unrelated tracts. We cull again to remove those paths that do not project into the gray matter. The paths after the culling process are shown in Fig. 4 right.

Clustering streamlines

We used agglomerative hierarchical clustering methods [4] to cluster the streamlines. The general algorithm starts with n singleton clusters and forms the sequence by successively merging clusters. The algorithm works as follows:

1. Given a set of n singleton clusters.
2. Merge the two nearest clusters
3. Repeat 2. until the specified number of clusters are generated.

If we set the specified number of clusters to 1, we produce a dendrogram. To determine the "nearest" cluster, we have to define the distance between two paths using the following equation:

$$D_l = \frac{\int_{s_0}^{s_1} \max(\mathrm{dist}(s) - T_l, 0)\mathrm{d}s}{\int_{s_0}^{s_1} \max(\frac{\mathrm{dist}(s)-T_l}{|\mathrm{dist}(s)-T_l|}, 0)\mathrm{d}s},$$

where s parameterizes the arc length of the longer trajectory, and $\mathrm{dist}(s)$ is the shortest distance from location s of the shorter trajectory to the longer trajectory. T_l ensures that we label two trajectories as different if they differ significantly over any portion of the arc length. This property creates a discriminating distance between the two tracts in Fig. 3 right.

The minimum distance between any two paths from two clusters is used for the distance between these clusters, the method is called the nearest-neighbor cluster algorithm, or minimum algorithm.

Results

Figure 4 shows the streamlines before clustering. The left picture shows the path set without the endpoints constraint. There are about $1,1000$ paths, many of which are short paths that stop within the ambiguous white matter regions, which could potentially create artificial links between fiber bundles. The right picture shows the $6,000$ paths after the culling process. Fiber bundles in Fig. 1 and Fig. 2 are clustered on this path set. There are 600 bundles; we only visualize around 100 bundles that contain 10 or more paths.

From the top view of our model (Fig. 1), the two cingulum bundle tracts are clearly indicated as two clusters; the corpus callosum form several bundles that run into the outer brain without much divergence within the bundles. Groups of U fibers also form distinct bundles. From the side view (Fig. 2), the internal capsule is clustered into several bundles that are coherent along the pathways.

Conclusion

We present a clustering process for the fiber paths from a whole brain DTI data set. We propose a path culling process and a similarity metric for the clustering algorithm to avoid misclassifications. The results show that in most places, paths are grouped together in coherent bundles.

Acknowledgments

This work was supported in part by NSF (CCR-0086065) and the Brain Science Program at Brown. Thanks to Dr. Steve Correia from Brown Medical School for the human brain DTI data set.

References

[1] Zhaohua Ding, John C. Gore, and Adam W. Anderson. Classification and quantification of neuronal fiber pathways using diffusion tensor mri. *Magnetic Resonance in Medicine*, 49:716–721, 2003.

[2] Isabelle Corouge, Sylvain Gouttard, and Guido Gerig. Towards a shape model of white matter fiber bundles using diffusion tensor MRI. In *International Symposium on Biomedical Imaging*, pages 344–347, 2004.

[3] Song Zhang, Cagatay Demiralp, and David H. Laidlaw. Visualizing diffusion tensor MR images using streamtubes and streamsurfaces. *IEEE Transactions on Visualization and Computer Graphics*, 9(4):454–462, October 2003.

[4] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley-Interscience Publication, 2000.