

# DTI Fiber Clustering and Cross-subject Cluster Analysis

S. Zhang<sup>1</sup>, D. H. Laidlaw<sup>1</sup>

<sup>1</sup>Computer Science Department, Brown University, Providence, RI, United States

**Introduction** – We report on unsupervised clustering of tractography paths and automated matching of the resulting clusters across two brains. DTI tractography methods generate paths that correlate well with large WM structures. However, it is not clear how reliably they can identify anatomical tracts. Recently, several approaches have been proposed to cluster these paths [1,2,3]. To avoid errors, our approach aims at reliably identifying only relatively large and unambiguous sets of paths; we remove paths that are ambiguous before clustering and only find matches between clusters when they are relatively compelling.

**Methods** – A Siemens Symphony 1.5T scanner was used for imaging. Three slice packets of DWIs were acquired sagittally and interleaved to acquire a data volume of 128×128×90 with a voxel size of 1.7×1.7×1.7 mm. The Siemens MDDW protocol was used, with three b values (0, 500, 1000) in 12 directions. Two subjects were scanned in this study and DTIs were calculated.

Paths were calculated within each DTI dataset starting at seed points jittered slightly from a regular grid with 0.85 mm spacing [4]. The paths were integrated in both directions through the first eigenvector field using 2<sup>nd</sup>-order Runge-Kutta; integration stopped when the linear anisotropy value dropped below 0.15, the path exited the dataset, the path curved excessively, or the signal-to-noise ratio (SNR) in the B0 image dropped below 50% of the SNR of WM. The paths were culled to remove short paths, redundant paths, and paths that did not project into gray matter according to classification of the B0 image with FAST [5].

An agglomerative hierarchical clustering method [6] was used to cluster the paths. It started from a set of singleton clusters and iteratively merged the current two nearest clusters until the intercluster distance reached a threshold of 2.5 mm. The distance function between paths is defined as:

$$D_i = \int_{t_0}^{t_1} \max(\text{dist}(s) - T_i, 0) ds / \int_{t_0}^{t_1} \max((\text{dist}(s) - T_i) / |\text{dist}(s) - T_i|, 0) ds$$

and is described in more detail in [4]. Intercluster distance is the minimum distance between any pair of paths with one from each cluster.

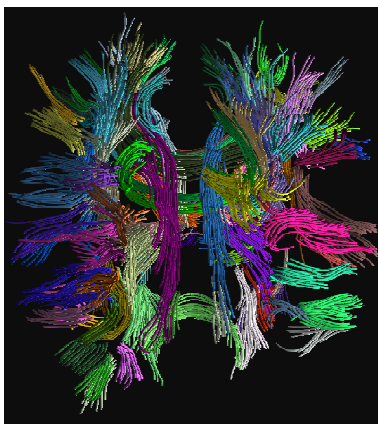
We matched clusters across datasets as follows. First, we roughly registered the datasets by matching a bounding box surrounding the whole brain WM. Only translation and scaling along each axis were used. For each path cluster, the centroids of the starting points, middle points, and end points were calculated and concatenated to form a nine-valued feature vector. Path clusters from the two subjects were then matched up according to the Euclidean distance between their feature vectors. Matching clusters were required to be mutually closest in the feature space as well as nearer than 40 units.

**Results and Discussion** – Figure 1 shows the result of clustering on one subject. Only clusters that have more than 10 paths are shown. A single color was randomly chosen for each cluster. The two cingulum bundles are clearly identified as independent clusters, the corpus callosum is represented by several coherent clusters. Figure 2 shows the result of matching clusters across two subjects. Each matching pair share the same color. Unmatched clusters are not shown. A number of structures clearly match well, including the cingulum bundles. Note that many structures are not represented due to uncertainty in either the clustering or matching stage. We believe that this starting point for clustering and matching will provide a framework for resolving the more difficult and ambiguous cases.

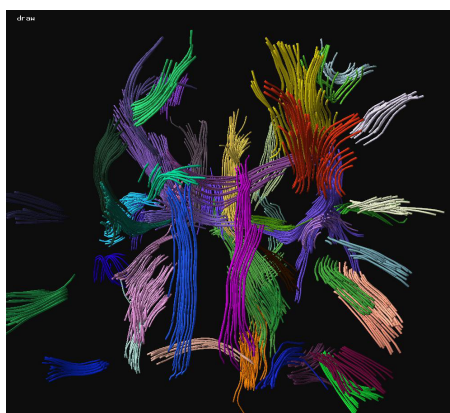
**Conclusions** – The unsupervised clustering of relatively unambiguous paths in DTI datasets provides a possibly approach for identifying relevant features automatically. The ability to match such clusters across subjects automatically suggests that automating comparison of DTI datasets using clustering results may be feasible.

**Acknowledgments** - Support from NSF CCR-0086065, Ittleson Fund at Brown, and the Brain Science Program at Brown.

**References** – [1] Ding et al. 2003 MRM, 49:716-721. [2] Corouge et al. ISBI, 2004, 344-347 [3] Anders et al. MICCAI 2004, 368-375. [4] Zhang et al. 2003 IEEE TVCG, 9, 454-462. [5] Zhang et al. 2001, 20(1):45-57. [6] Duda et al. 2000. Pattern Classification.



**Figure 1** Top view of the path clusters. Paths within the same cluster share the same color.



**Figure 2:** Path clusters across the two subjects are matched up according to the distance between their feature vectors. The left cingulum bundles and right cingulum bundles, among other features, match up well.