

Hierarchical Density-Based Clustering of White Matter Tracts in the Human Brain

Junming Shao, University of Munich Germany

Klaus Hahn, HMGU Helmholtz Center Munich, Germany

Qinli Yang, University of Edinburgh, UK

Afra Wohlschläeger, Technical University of Munich, Germany

Christian Boehm, University of Munich, Germany

Nicholas Myers, Technical University of Munich, Germany

Claudia Plant, Florida State University, USA

ABSTRACT

Diffusion tensor magnetic resonance imaging (DTI) provides a promising way of estimating the neural fiber pathways in the human brain non-invasively via white matter tractography. However, it is difficult to analyze the vast number of resulting tracts quantitatively. Automatic tract clustering would be useful for the neuroscience community, as it can contribute to accurate neurosurgical planning, tract-based analysis, or white matter atlas creation. In this paper, the authors propose a new framework for automatic white matter tract clustering using a hierarchical density-based approach. A novel fiber similarity measure based on dynamic time warping allows for an effective and efficient evaluation of fiber similarity. A lower bounding technique is used to further speed up the computation. Then the algorithm OPTICS is applied, to sort the data into a reachability plot, visualizing the clustering structure of the data. Interactive and automatic clustering algorithms are finally introduced to obtain the clusters. Extensive experiments on synthetic data and real data demonstrate the effectiveness and efficiency of our fiber similarity measure and show that the hierarchical density-based clustering method can group these tracts into meaningful bundles on multiple scales as well as eliminating noisy fibers.

Keywords: DBSCAN, Density-Based Fiber Clustering, Diffusion Tensor Imaging, Dynamic Time Warping, Lower Bounding Distance, OPTICS

INTRODUCTION

Diffusion Tensor imaging (DTI) can explore the organization and integrity of human white mat-

ter tracts in vivo, using water diffusion properties as a probe (Mori, 2007). It measures for every voxel the diffusivity of water molecules within the tissue, and thus gives valuable insight into the orientation of fiber tracts, since water diffusion is strongest along the direction of fibers

DOI: 10.4018/jkdb.2010100101

and restricted in the directions perpendicular to them. Potential pathways of fiber tracts in the white matter of the human brain can be reconstructed by deterministic and stochastic tractography based on the measured diffusion weighted images. These techniques have attracted attention in the study of anatomical connectivity (Hagmann et al., 2008), brain changes (Huang et al., 2006), and various pathologies related to white matter atrophy, such as schizophrenia (Park et al., 2004), Alzheimer's disease (Damoiseaux et al., 2009), or multiple sclerosis (Law & Grossman, 2005).

Performing fiber tracking in the human brain usually results in large sets of tracks. A problem in clinical research is how to segment and interpret this vast amount of information. A frequently used method is to select fiber groups of interest on the basis of expert knowledge by *virtual dissection*. Experts first specify some regions of interest (ROIs) and select then all fibers that pass through these pre-defined ROIs (Catani et al., 2002). This process tends to be inefficient, since the manual handling of ROIs is time-consuming and also limited by the availability of experts. Moreover, manual specification of ROIs may be biased in patient populations. Therefore, automatic clustering of fiber tracts into bundles of similar fibers is preferable for many applications. Two fibers are considered as similar if they have comparable length, similar shape and similar location (Ding, Gore, & Anderson, 2003). A number of approaches have been proposed to automatically cluster white matter tracts. However, before grouping fibers into clusters, it is necessary to specify a fiber similarity measure.

A fiber similarity measure is a function that quantifies the similarity between pairs of fibers. In the early work by Brun et al. (2003), it is assumed that two fiber tracts with similar end points should be considered as similar. The Euclidean distance between the starting and ending points of the two fibers is used to calculate fiber similarity. However, this assumption is not sufficient in some cases, as not all fibers in a bundle start and end in the same region. It also

ignores the shape information contained in all points on the fibers. Ding, Gore, and Anderson (2003) propose a similarity measure by cutting each fiber into corresponding fiber segments and use then the mean Euclidean distance between the segments to define piece-wise similarity. This similarity method is efficient, but may lack effectivity since this measure loses the point-by-point information. Several authors acknowledge that the point-by-point correspondence of the trajectories should be included to define a convenient fiber similarity measure. For instance, in (Zhang, Demiralp, & Laidlaw, 2003; Zhang & Laidlaw, 2005) the distance is defined between two fibers as the average distance from any point on the shorter fiber to the closest point on the longer fiber, where only distances above a certain threshold contribute to this average. Another possibility is proposed by Klein et al. (2007). They introduce a grid-based similarity measure. A box around the fibers is partitioned into identical cells. Each point of a fiber is then assigned to the cells with weights. The similarity between a pair of fibers is calculated by the pairwise weights, summed over all cells including the fibers. The side length of the cell is an adjustable parameter that controls the scale of comparison. Corouge, Gouttard, and Gerig (2004) form point pairs, mapping each point of one fiber to the closest point on the other fiber. The resulting point pairs are then used to define the distance between fiber pairs. They define three similarity distances: closest point distance, mean of closest point distance (MCP) and Hausdorff distance (HDD).

Currently, MCP (Corouge, Gouttard, & Gerig, 2004; O'Donnell et al., 2006; O'Donnell & Westin, 2007; Moberts, Vilanova, & Wijk, 2005; Zhang & Laidlaw, 2005) and HDD (Corouge, Gouttard, & Gerig, 2004; Gerig, Gouttard, & Corouge, 2004; Moberts, Vilanova, & Wijk, 2005; Xia, Turken, Whitfield-Gabrieli, & Gabrieli, 2005) are widely used fiber similarity measures. Their formal definition is the following:

Suppose there are two fibers P and Q, then MCP is defined as:

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the product's webpage:

www.igi-global.com/article/hierarchical-density-based-clustering-white/49547?camid=4v1

This title is available in InfoSci-Journals, InfoSci-Journal Disciplines Medicine, Healthcare, and Life Science.

Recommend this product to your librarian:

www.igi-global.com/e-resources/library-recommendation/?id=2

Related Content

Mining of Medical Trends Using Social Networks

Shruti Kohli and Sonia Saini (2016). *Biomedical Image Analysis and Mining Techniques for Improved Health Outcomes* (pp. 164-182).

www.igi-global.com/chapter/mining-of-medical-trends-using-social-networks/140489?camid=4v1a

Scaling Unsupervised Risk Stratification to Massive Clinical Datasets

Zeeshan Syed and Ilan Rubinfeld (2011). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 45-59).

www.igi-global.com/article/scaling-unsupervised-risk-stratification-massive/52770?camid=4v1a

Incorporating Correlations among Gene Ontology Terms into Predicting Protein Functions

Pingzhao Hu, Hui Jiang and Andrew Emili (2013). *Bioinformatics: Concepts, Methodologies, Tools, and Applications* (pp. 831-850).

www.igi-global.com/chapter/incorporating-correlations-among-gene-ontology/76098?camid=4v1a

Dis2PPI: A Workflow Designed to Integrate Proteomic and Genetic Disease Data

Daniel Luis Notari, Samuel Brando Oldra, Mauricio Adami Mariani, Cristian Reolon and Diego Bonatto (2012). *International Journal of Knowledge Discovery in Bioinformatics* (pp. 67-85).

www.igi-global.com/article/dis2ppi-workflow-designed-integrate-proteomic/77811?camid=4v1a