Fiber tracts of high angular resolution diffusion MRI are easily segmented with spectral clustering.

L. Jonasson[1], P. Hagmann[1], J-P. Thiran[1], V. J. Wedeen[2]

[1]Signal Processing Institute, EPFL, Lausanne, Vaud, Switzerland, [2]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and the Harvard Medical School, Boston, MA, United States

Introduction

Fiber tractography on High Angular Resolution Diffusion Imaging (HARDI) data, such as Diffusion Spectra Imaging (DSI) or q-ball imaging, results in a large set of fiber tracts with a very complex geometry. The higher complexity compared to Diffusion Tensor MRI (DT-MRI) is due the numerous intersections between fibers that can be resolved, that is to say separated, using HARDI. Even though fiber bundles can be well separated in orientation space they intertwine and mix in the 3D position space. Spectral clustering techniques are methods that aim at obtaining new data representations to separate clusters with significant overlaps by creating a new feature space in which the clusters are clearly distinct from each other. This new space is constructed from the eigenvectors of a local affinity matrix representing the data and any classical clustering algorithm can then be applied on the eigenvectors [1, 2]. Considering the character of fiber tract data, spectral clustering seems like a highly appropriate segmentation technique. To test this hypothesis we have constructed a simple algorithm for counting the number of intersections between fibers and run a spectral segmentation algorithm on the co-occurrence matrix obtained from the counting. The proposed method is an unsupervised clustering technique, applicable for large sets of fiber tracks. It is easy to implement and has a low computational cost.

Method

The first step of our algorithm is to create a 3D Euclidean space of an appropriate resolution. An intuitive choice would be to simply return to the voxel space from which the fibers have originally been generated. However, we have chosen a coarser resolution by reducing the image size by half along each axis. As Brun et al. remarks in [3], mapping the fibers into Euclidean space might seem a bit crude but works really well. Every fiber will then be assigned a number of identification and a list of all fibers passing through will be saved in every voxel.

In the second step we create an N by N large co-occurrence matrix, where N is the number of the fibers to cluster. The co-occurrence matrix contains the number of times two fibers share the same voxel. Since the number of fibers that we wish to segment is very large this co-occurrence matrix will also be very large. Due to the nature of the HARDI data the matrix is also sparse since most of the fibers never cross which makes it possible to handle it in Matlab despite its size. The matrix will be made even sparser by removing the influence of fibers that only have a few voxels in common by setting their values in the co-occurrence matrix to zero. It is this simple co-occurrence matrix that will serve as our affinity matrix and represent our data set. The sparse function in Matlab makes it possible to calculate the six largest eigenvalues and their corresponding eigenvectors of the affinity matrix. In the last step we run a k-means clustering on these eigenvectors and the resulting clusters will correspond to fiber bundles.

Material The method has been tested on a set of fibers obtained from DSI data using a tractography method described in Hagmann et al. [4]. The diffusion images were obtained on a healthy volunteer with a 3T Allegra scanner (Siemens, Erlangen, Germany). We used a twice-refocused spin echo EPI sequence with TR/TE/Δ=3000/154/66 ms, bmax = 1700mm2/s and a spatial resolution of 3mm3. Data were acquired using 515 different diffusion encoding directions sampling on a sphere of radius r=5 grid units.

Results

The left image shows an example of unclustered fibers from the corpus callosum and the cortico spinal tract. The fibers are color coded according to their tangent vectors. The algorithm was told to separate the fibers into eight clusters and the second image shows six of these clusters, each color representing a cluster. Of the clusters that are not shown, one of the clusters contains all fibers with less than two neighbors and the second one only contained very few fibers and is therefore not visible.
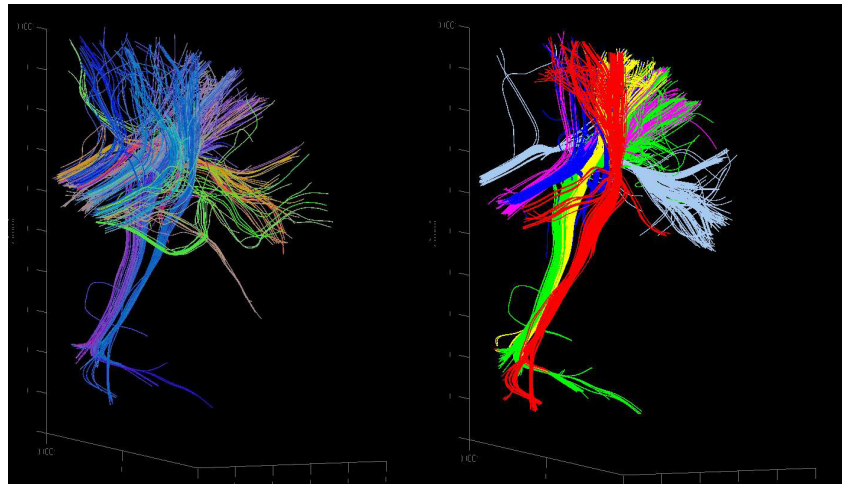
Discussion and conclusion

We have shown a promising example of the capability of spectral clustering to segment fibers that intertwine in position space and showing evidence of their separation in orientation space. This unsupervised segmentation technique is a first step towards a quantitative analysis of the fiber tracts from HARDI data which will improve the comprehension of connectivity in anatomical structures.

An approach related to spectral segmentation has been made for DT-MRI by Brun et al. [3]. First they generate a weighted graph from a pair wise comparison of distance and shape between fibers. This graph can be compared to our affinity matrix. Then they use normalized cuts to partition the graph by using its second smallest eigenvector and then successively subdividing the graph until obtaining the desired number of clusters. Even though authors disagree on which eigenvectors to use and how to derive clusters from them, recent papers on spectral clustering now agree that using more eigenvectors and directly partitioning in the desired number of clusters will improve the results. Weiss et al. [2] argues for the more eigenvectors are used, the better segmentation. However, in our application only the six largest eigenvectors are used. If this is a sufficiently large number of eigenvectors is left to explore but for the moment this is the number we manage to efficiently compute from this large amount data that is our affinity matrix.

The higher orientational resolution that HARDI provides, allows us to use a very simple approach to obtain our affinity matrix. The sparseness of the affinity matrix is a consequence of this and is important for several reasons. Most important, the sparseness gives a sense to our counting procedure. With a lower angular resolution numerous fibers would have several intersecting points in common without belonging to the same tract; this is one reason for which this counting procedure will most probably not work for DT-MRI. Methods for clustering fibers from DT-MRI have previously been presented by a number of authors. They mostly rely on pairwise similarity between fibers which include defining appropriate distances between fibers, both concerning shape and Euclidean distance. Defining these similarity measures is obviously a much more challenging task than simply counting the intersection points.

References [1] Ng et al. On spectral clustering: Analysis and an algorithm, In Advances in Neural Information Processing Systems 14: Proc. 2001.; [2] Weiss, Y. Segmentation using eigenvectors: a unifying view. Proc. IEEE Int. Conf. on Computer Vision p. 975-982 (1999).; [3] Brun, A. et al. Clustering fiber traces using normalized cuts. MICCAI'2004, Rennes, Saint Malo, France, September 26-30, 2004. [4] Hagmann, P. et al. Diffusion Spectrum Imaging tractography in complex cerebral white matter: an investigation of the centrum semiovale. ISMRM 2004.