A novel paradigm for automated segmentation of very large whole-brain probabilistic tractography data sets

R. E. Smith[1,2], J-D. Tournier[1,2], F. Calamante[1,2], and A. Connelly[1,2]
[1]Brain Research Institute, Florey Neuroscience Institutes, Heidelberg West, Victoria, Australia, [2]Department of Medicine, The University of Melbourne, Melbourne, Victoria, Australia

Introduction: Although there is a vast wealth of structural connectivity data contained within whole-brain diffusion MRI tractography data, this information has remained relatively untapped due to the difficulties associated with the extraction of interesting structures and the interpretation of differences between subjects. Much research has been undertaken into the automated segmentation, or 'clustering' of the data, to identify known anatomical bundles and enable more logical comparisons between data sets (e.g. [1, 2]). These methods typically only extract either very large or very small structures, scale poorly with respect to the number of tracks (or have an upper limit on the number of tracks they can process), or rely upon *ad hoc* parameters to partition the data set into more manageable sizes. In addition, most techniques thus far have been demonstrated only upon deterministic streamlines data, which underestimates the complexity of the connectome. Here we present a novel automated tractography segmentation technique, based upon a paradigm of finding localized bound coherent bundles of tracks, rather than the typical approach of grouping tracks according to pair wise similarities. It is capable of identifying coherent white matter structures at a wide range of physical scales, from probabilistic streamlines tractography data sets of an order of magnitude greater size than could be handled using any previously published technique, with no *a priori* bias. Furthermore, the algorithm does not simply group together tracks into clusters; as the extraction of structures is region-based, and most tracks will be attributed to multiple regions, the connectivity between each pair of identified regions is inherently quantified during the clustering process, providing an intrinsic quantitative metric of structural connectivity throughout the brain.

Method: Using whole-brain fibre tracking data (Fig. 1.1), the algorithm begins by producing a global map of fibre structure in the form of a modified super-resolution track density image [3] (Fig. 1.2) with high angular resolution (tracks contribute to voxels according to the local tangent of the track, such that the total density in each voxel is a symmetric spherical function). Track density "lobes" are produced by agglomerating the high angular resolution track densities (Fig. 1.3). These lobes are matched between neighbouring voxels, and lobes for which a match does not exist in a neighbouring voxel are designated as the edges of white matter structures (Fig. 1.4). Bundles are located by seeding upon these edges throughout the brain, and traversing orthogonally to the lobe density direction along the edges in search of bound coherent paths (e.g. see yellow outline in Fig. 1.5). These loops are filled volumetrically (Fig. 1.6), and used as thin regions of interest to select those tracks passing through each identified region (Figs. 1.7, 1.8). Regions for which the track membership listings are sufficiently similar (effectively the same subset of tracks passes through both regions) are merged to form volumetric regions of interest; unique regions are discarded as spurious. Note that this process achieves whole-brain coverage, and is fully automated.

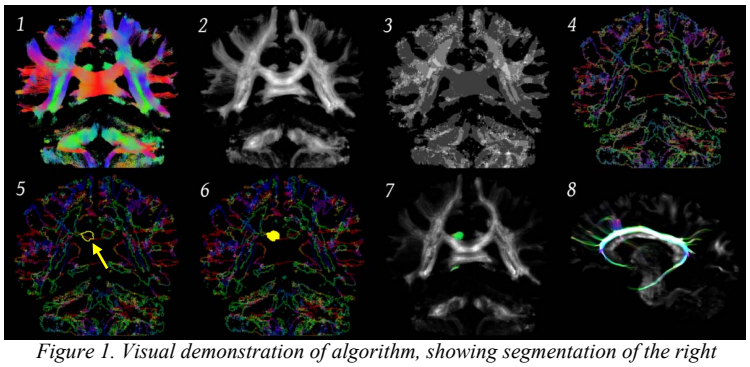

Figure 1. Visual demonstration of algorithm, showing segmentation of the right cingulum bundle superior to the splenium of the corpus callosum; (1.1 - 1.7) coronal slice; (1.8) sagittal track projection. Red: left-right, Green: anterior-posterior, Blue: inferior-superior, Yellow: automated region determination

Data acquisition: Diffusion-weighted images were acquired from a healthy volunteer on a 3T Siemens Tim Trio (2.3mm isotropic resolution, 150 diffusion-sensitization directions, b = 3,000 s/mm^2). Fibre orientation distributions were estimated by Constrained Spherical Deconvolution [4], and 10,000,000 probabilistic streamlines were generated by 2nd Order Integration over Fibre Orientation Distributions [5], using in-house software based upon the MRtrix software package [6].

Results & Discussion: The high angular resolution track density image was produced at 0.5mm isotropic resolution, with 129 directions on the unit hemisphere. 86,036 planar regions of interest were identified by the algorithm, and reduced to 8,274 volumetric regions of interest. Processing was performed on a 2.8GHz processor with 8GB RAM. A large number of well-known white matter structures are identified (Fig. 2), at a very wide range of physical scales (Fig. 3). In addition, the connectivity of the brain can be interrogated more thoroughly by analysing the connection relationships between different regions; this region-based definition of bundles permits Boolean logic to be applied to extract specific connections of interest, without the need for further targeted tracking or clustering of individual fascicle track data sets (Fig. 4). The massive number of tracks in the whole-brain data preserves the fine detail within each structure after clustering.

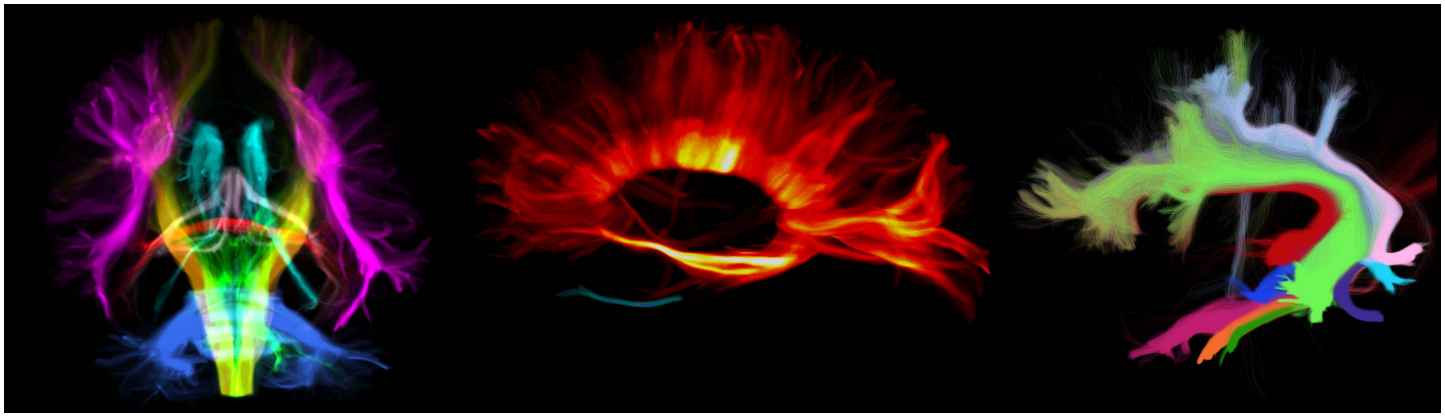

Figure 2. Example structures identified by the algorithm: arcuate fasciculi (purple), cingulum bundles (cyan), fornix (pink), anterior commissure (red), corticospinal tracts (orange), superior cerebellar peduncles (green) and middle cerebellar peduncle (blue); coronal Track Density Image (TDI) Maximum Intensity Projections (MIP)

Figure 3. Regions as large as the corpus callosum (hot) and as small as the oculomotor nerve (cool) are identified by a single execution of the algorithm, highlighting the scale invariance of the technique; sagittal TDI MIP (note: the extracted region corresponding to the corpus callosum contains over 2.3 million tracks)

Figure 4. Connections from the left arcuate fasciculus to a number of temporal lobe gyral projections (all extracted from whole-brain data), individually colour-coded such that their paths through the arcuate to the frontal lobe can be traced; sagittal track projection

Conclusion: We have presented a new algorithm for fully automated segmentation of massive probabilistic tractography data sets, which overcomes many of the fundamental limitations associated with previously published techniques. It enables many qualitative and quantitative methods for the analysis of brain structural connectivity.

References: [1] Guevara et al., NeuroImage (in press) 2010 [2] Visser et al., NeuroImage 54:303-312 (2011) [3] Calamante et al., NeuroImage 53:1233-1243 (2010) [4] Tournier et al., NeuroImage 35: 1459-1472 (2007) [5] Tournier et al., ISMRM 18:1670 (2010) [6] MRtrix, www.brain.org.au/software