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

R. E. Smith1,2, J-D. Tournier1,2, F. Calamante1,2, and A. Connelly1,2

1Brain Research Institute, Florey Neuroscience Institutes, Heidelberg West, Victoria, Australia, 2Department 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 [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 pairwise 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 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, based upon a paradigm of finding localised structures (Fig. 1.4). Bundles are identified 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.

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/mm²). Fibre orientation distributions were estimated by Constrained Spherical Deconvolution [3], and 10,000,000 probabilistic streamlines were generated by 2nd Order Integration over Fibre Orientation Distributions [4], using in-house software based upon the MRtrix software package [5].

Results & Discussion: The high angular resolution track density image was produced at 0.5mm isotropic resolution, on 79 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 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

Figure 2. Example structures identified by the algorithm: arcuate fasciculus (purple), cingulum bundles (cyan), fornix (pink), anterior comissure (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 2.5 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.