A probabilistic model-based approach to consistent white matter tract segmentation

J. D. Clayden[1], A. J. Storkey[2], and M. E. Bastin[3]

[1]Neuroinformatics Doctoral Training Centre, School of Informatics, University of Edinburgh, Edinburgh, United Kingdom, [2]Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, Edinburgh, United Kingdom, [3]Medical and Radiological Sciences (Medical Physics), University of Edinburgh, Edinburgh, United Kingdom

Introduction

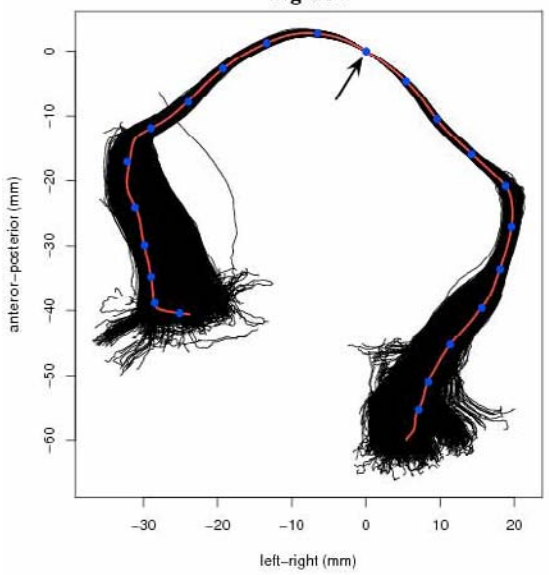
Tractography algorithms have advantages, as tools for segmentation of white matter structures from diffusion MRI (dMRI) data, over more established region of interest (ROI) approaches. In particular, they are capable of automatically segmenting irregularly shaped structures that would be difficult and error-prone for a human observer to isolate. The main problem with tractography based segmentation is that algorithms require a seed point as a starting location. Since this point is typically placed by a human observer, and the segmentation can be very sensitive to its placement, a strong element of subjectivity remains in the results. We have recently demonstrated proof of concept for an approach to automated seed point placement in which a set of points are each used to generate a "candidate" tract, and the single seed point is chosen whose corresponding tract matches best to a predefined reference tract [1]. In that case, each candidate seed point is treated as a hypothesis, and the hypothesis with the best evidence to support it—in terms of tract similarity—is chosen. In the present work we take this approach further, developing a formal probabilistic model for the shape and length relationships between tracts, which resolves many of the shortcomings of the previous method.

Methods

Previously acquired volunteer data, with 51 diffusion weighting directions and a *b*-value of 1000 s mm^{-2}, were used for this study (cf. [1]). The dMRI data were preprocessed to remove skull data and eddy current distortions from the images, using FMRIB Software Library tools (FMRIB, Oxford, UK).

For each seed point, the BEDPOST/ProbTrack tractography algorithm [2] generates a set of 5000 "probabilistic streamlines", each describing a path in two directions from the seed point. Rather than work with this full distribution of streamlines, we instead calculate a median line, which is assumed to have shape and length properties that are representative of the set. This median line is then parameterised as a uniform B-spline curve such that one of the knot points falls on the seed point. Fig. 1 shows the original set of streamlines (black), the median line (red), and the internal knot points of the B-spline (blue), in axial projection. The origin of coordinates is the seed point, which is indicated with an arrow. The knot point spacing is chosen to reduce the residual standard error of the fit below a threshold in the reference tract, and is then fixed for all other tracts. We define a latent variable, z_i, for each of N candidate tracts, i in $\{1,2,...,N\}$, which indicates whether the tract matches the reference tract ($z_i = 1$) or not ($z_i = 0$). We additionally introduce the possibility $z_0 = 1$ if there is no match among the candidate tracts. We then model the shapes and lengths of the tracts for the matching and nonmatching cases separately. We fit the model parameters using maximum likelihood estimation, and then calculate the posterior probability of each tract representing the best match to the reference.

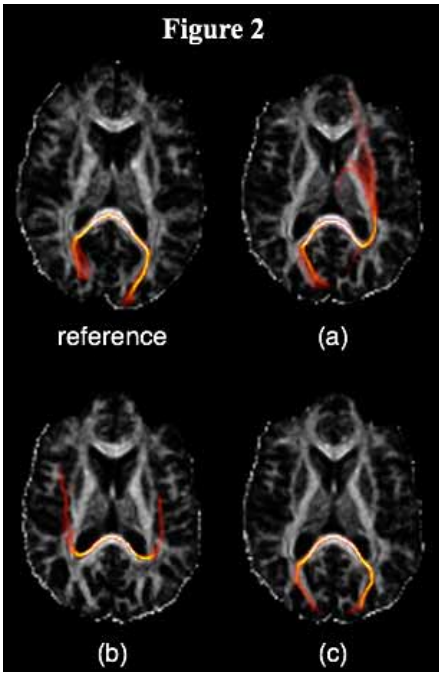
Figure 1



Results

Fig. 2 shows an example of an application of the model to tract selection. Subfigure (a) shows tractography output from a single seed point placed in the splenium region of a standard brain map and registered to the subject's native space. Fig. 2(b) is the candidate tract chosen by applying neighbourhood tractography (NT; see [1]) in the 7x7x7 voxel region around the original seed location, using the reference tract shown, which is from a different individual. In this case the NT method chooses a poor match to the reference tract, but the tract with highest posterior probability across the same seed region in our current model (c) represents a far better match. Its posterior probability was 0.016—this is not very large as there are likely to be a number of plausible matches to the reference, but it's much larger than the posterior null match probability of 0.000070.

Figure 2



Discussion

The B-spline tract parameterisation and formal probabilistic model that we have introduced here improves upon our previous approach to consistent tract segmentation using a reference tract, in a number of ways. One significant practical benefit is that it provides an indication of the likelihood of a good match to the reference tract being wholly absent from a data set, something that it was previously not possible to gauge. Secondly, the numeric measures of fit that it produces have a more concrete meaning—as probabilities that each tract was drawn from the matching model rather than the nonmatching one—than those generated by our previous tract similarity measure. Finally, algorithmic pathological cases that can occur when working with a field of connection likelihood values will not affect the B-spline representation of tracts, making it less likely that poor matches such as that in Fig. 2(b) will be selected by the new method. We believe that in this improved form, reference tract based segmentation could be a robust approach for clinical studies interested in group differences between specific tracts of interest.

References

[1] Clayden *et al.* (2006). *NeuroImage* **33**(2):482–492. **[2]** Behrens *et al.* (2003). *Magn Reson Med* **50**(5):1077–1088.