Comparison of flow- and streamline-based fibre tracking algorithms using an anisotropic diffusion phantom

J. S. Campbell[1,2], V. Rymar[1], A. S. Sadikot[3], K. Siddiqi[2], G. B. Pike[1]

[1]McConnell Brain Imaging Centre, Montreal Neurological Institute, Montreal, Quebec, Canada, [2]Centre for Intelligent Machines, McGill University, Montreal, Quebec, Canada, [3]Montreal Neurological Institute, Montreal, Quebec, Canada

Introduction There is significant interest in evaluating the performance and reliability of white matter fibre tractography algorithms. While simulated data can provide a gold standard to which tracking results can be compared [2,3,6], it is of additional value to evaluate the results of tracking on real MRI data, in the presence of normal imaging artifacts, noise characteristics, and voxel size limitations. The goal of this study was to evaluate the performance of our flow-based fibre tractography algorithm [1], Tracking Using arbitrary Fibre Orientation Likelihood Distribution (TUFOLD). Flow-based algorithms are of particular interest because of their potential for tracking in regions of subvoxel partial volume averaging of fibre directions, and their applicability to high angular resolution measurements of the diffusion pdf. In TUFOLD, a modification and extension of fast marching tractography [5], tracking is done by surface evolution. As the surface passes through a voxel, its speed of evolution in the direction of unit vector $\boldsymbol{u}$ is given by the value of the fibre orientation likelihood distribution $L_f(\boldsymbol{u})$ in that direction. L_f is estimated from the diffusion orientation distribution function, given by $\int_0^\infty P(r\mathbf{u})dr$, with P the diffusion pdf measured at arbitrarily high angular resolution. This tracking algorithm yields a scalar map of likelihood of connection to a user-defined seed, as well as 3D curves representing fibre tracts. We compare TUFOLD to a well established streamline method, FACT [4]. We do so by first creating a physical anisotropic diffusion phantom with fibre tracts of size comparable to those in the human brain, and subsequently calculating a performance measure for the tractography results.

Methods Two Sprague-Dawley rats were euthanized and their spinal cords surgically excised. The fresh cords were embedded in 2% agar in a configuration designed to have curved, straight, and crossing tracts. The cords were scanned one hour after the surgery with a Siemens 1.5T Sonata MR scanner (Siemens Medical Systems, Erlangen, Germany) using an extremity coil. 90 diffusion weighted images were acquired using isotropically spaced diffusion weighting directions (b=1300 s/mm^2, TR=8s, TE=110ms, 2.5 mm isotropic voxels, 40 slices), as well as 10 T2-weighted images with the same imaging parameters. A 1 mm isotropic resolution T1-weighted scan was also performed (TR=22ms, TE=9.2ms, α=30^o). The diffusion tensor was calculated from this dataset using least-squares linear regression.

The gold standard tracts were defined for the two cords, one straight and one curved, by a combined automatic and manual segmentation of the scalar image of the trace of the diffusion tensor. The two cords crossed in one region, therefore some voxels were identified to lie in both. A gold standard map, *GS*, for a given tract is a binary map with a value of unity in all voxels identified to lie on that tract. In each gold standard tract, 5 evenly spaced seed regions were defined: these consisted of one voxel thick cross-sections of the cord. Tracking was performed using the TUFOLD and FACT algorithms with each separate seed region. For FACT, 20 subvoxel seeds were placed in each voxel of the seed region. For TUFOLD, the initial surface was defined as the edge of the seed region. For each tracking algorithm, we calculated tracts as 3D curves, as well as a scalar likelihood of connection map L_c. For FACT, L_c is a binary map set to unity if a voxel is reached by tracking, whereas for TUFOLD, it is a continuous map that indicates our confidence in each voxel being connected to the seed region. For evaluation, the L_c map for TUFOLD was converted to a binary map by assigning a value of one to all voxels with L_c above an optimized threshold, and zero to all voxels with L_c below this threshold.

We defined the performance measure $\langle d \rangle$ for a scalar map L_c (given a gold standard map *GS*) to be the average of the distances d between pairs of voxels, where the pairs are given by the union of: i) each voxel with *GS*=1 and the nearest voxel with L_c=1; and ii) each voxel with L_c =1 and the nearest voxel with *GS*=1. $\langle d \rangle$ is measured in voxel units. The performance measure in the case of perfect tracking is zero.

Results Fig.1 shows the gold standard maps and tracking results. The plot in Fig.1d shows the mean performance measures over all seeds.

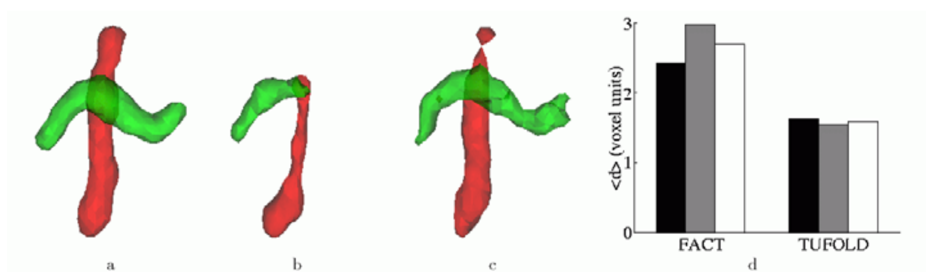

Fig. 1. (a) Voxels identified as gold standard voxels, rendered as a surface: straight (red) and curved (green) tracts. Voxels reached using (b) FACT and (c) TUFOLD with seeding in each of the two tracts. (d) Performance measures $\langle d \rangle$ for the two tracking algorithms: curved tract (black), straight tract (grey), and average over all tracts (white).

Discussion TUFOLD showed a trend of better performance in this dataset, with a mean $\langle d \rangle$ of 1.59 voxels, versus 2.70 voxels for FACT. Statistical confirmation of this trend will require more datasets. Visual inspection showed that FACT was good at piecewise reconstruction of the tracts, but tracts from individual seeds did not reach the ends due to difficulty passing through the region of subvoxel crossing (Fig.1b). This is most likely due to the ambiguity in the direction of the principle eigenvector of the diffusion tensor in regions of partial volume averaging of fibre directions. The seed point dependence of both algorithms was high: the mean (averaged over both tracts) of the standard deviation of the performance measures for all the seeds in a given tract was 2.80 voxels for FACT and 1.40 voxels for TUFOLD. FACT had a higher degree of seed point dependence due to its poor performance when seeds were placed near the end of the straight tract closest to the crossing region (see Fig.1a). This explains its poorer performance in the straight tract (Fig.1d).

We are also evaluating the performance of TUFOLD using high angular resolution diffusion measurements for the estimation of L_f, which should prove to be even more robust to subvoxel averaging of fibre directions. Investigations are also underway to evaluate the performance of the algorithms with more complex fibre architecture, and to determine the dependence of algorithm performance on the signal-to-noise ratio of the base images.

Conclusion We have introduced a method for constructing a physical anisotropic diffusion phantom that allows us to evaluate the performance of fibre tracking algorithms using real MRI data. We have shown our results comparing the TUFOLD and FACT algorithms using one high-quality diffusion tensor dataset. Our results show that flow-based tractography using the fibre orientation likelihood distribution defined for all (θ, ϕ) may be useful for mapping connections in regions of partial volume averaging of fibre directions.

References [1] Campbell *et al. Proc. ISMRM* 2002:1130. [2] Lazar *et al. Neuroimage*, 20(**2**):1140-53, 2003. [3] Lori *et al. Proc. ISMRM* 2000:775. [4] Mori *et al. Ann. Neurol*., 45:265-269, 1999. [5] Parker *et al. IEEE Trans. Med. Imag*., 21(5):505-512, 2002. [6] Tournier *et al. Magn. Reson. Med*., 47(**4**):701-8, 2002.