A comparative study of 16 tractography algorithms for the corticospinal tract: reproducibility and subject-specificity

Emmanuel Caruyer, Luke Bloy, Birkan Tunç, Jérémy Lecoeur, Varsha Shankar, Ragini Verma

A comparative study of 16 tractography algorithms for the corticospinal tract: reproducibility and subject-specificity

Emmanuel Caruyer[1], Luke Bloy[2], Birkan Tunç[1], Jérémy Lecoeur[1], Varsha Shankar[1], and Ragini Verma[1]

[1]*Section of Biomedical Image Analysis, Department of Radiology, University of Pennsylvania, Philadelphia, PA, United States,* [2]*Children's Hospital of Pennsylvania, Philadelphia, PA, United States*

Target audience– We present a comparative study of fiber tractography algorithms on a test-retest dataset. This work and the conclusions it leads to will guide the choice of a tracking algorithm for applications ranging from surgical planning to group-based connectivity analysis.

Purpose– We evaluate 16 different tracking algorithms, using their either freely available implementations or in house software. We compare their ability to produce reproducible tracking, as well as to capture subject-specific tract geometry.

Methods– *Data acquisition*. We scanned 9 healthy subjects (3F/6M, 31.25 ± 4.2y) at three time points separated by a 2 weeks interval. Diffusion-weighted EPI were acquired on a Siemens 3T Verio^{TM} scanner, b = 1000 s/mm², on 64 evenly distributed directions plus 2 b = 0 images.
Pre-processing. The diffusion-weighted images were regularized and corrected for Rician noise[1]. The diffusion tensor was estimated using linear least squares fitting, and fractional anisotropy (FA) maps were computed. Deformation field from and to the JHU Eve template[2] were computed using non-linear registration[3] of the FA images.
Fiber tracking. The fiber tracking of left and right corticospinal tracks (CST) was performed in subject-space. When applicable, the tractography was initiated in the cerebral peduncle (CP), using 5 seeds per voxel in that region of interest (ROI). The step size was set to 0.2mm, and the maximum turning angle to 35°. Tracking was terminated whenever underlying FA fell below 0.15. The tract was a posteriori filtered to keep only streamlines going through the posterior limb of the internal capsule (PLIC), and decimated to keep only 300 streamlines per track. The regions of interest were taken from Eve labels, warped to subject space using the above-computed deformation field. Tracking algorithms are streamline and probabilistic streamline on constrained spherical deconvolution (MRtrix[5]), FACT, second-order Runge-Kutta, streamline and tensorline on tensor (Diffusion toolkit[6]), Gibbs globals tracking (MITK[7]), 4th order Runge-Kutta and tensorline on tensor (Camino[8]), tensorline, streamline, tensor deflection, bootstrap streamline, bootstrap tensor deflection on tensor and streamline on CSA ODF (in-house software), and unscented Kalman filter on multi-tensor model[9].
Statistical analysis. Tract similarity was computed using a Gaussian process framework[4]. To this end, normalized tract probability images were computed in subject space, and warped back to template space. As a summary of the similarity matrix, we computed the ratio of the average inter-subject distance over the intra-subject distances.

Results– All the results reported here were obtained using a specific instance of the tracking parameters for each method. When applicable, the parameters were chosen to be as consistent as possible across the methods being compared.
Reproducibility and specificity. All the methods tested reported an inter-subject average distance at least 3 times higher than its intra-subject counterpart (Fig. 1.a). Among the methods giving a lower score, we can identify two different trends. On one hand, some methods such as MITK Gibbs tracking, are more conservative, and therefore provide a very high intra-subject similarity, at the cost of loosing inter-subject discriminative power. On the other hand, some methods such as Camino Euler tracking are more exploratory, and on average captures differences between subject, at the cost of a lower reproducibility. Interestingly, the probabilistic streamline tracking in MRTrix is comparable with other methods in terms of reproducibility. Finally, the similarity matrix which presents the most well-defined clusters "visually" was obtained with DTK tensorline method.

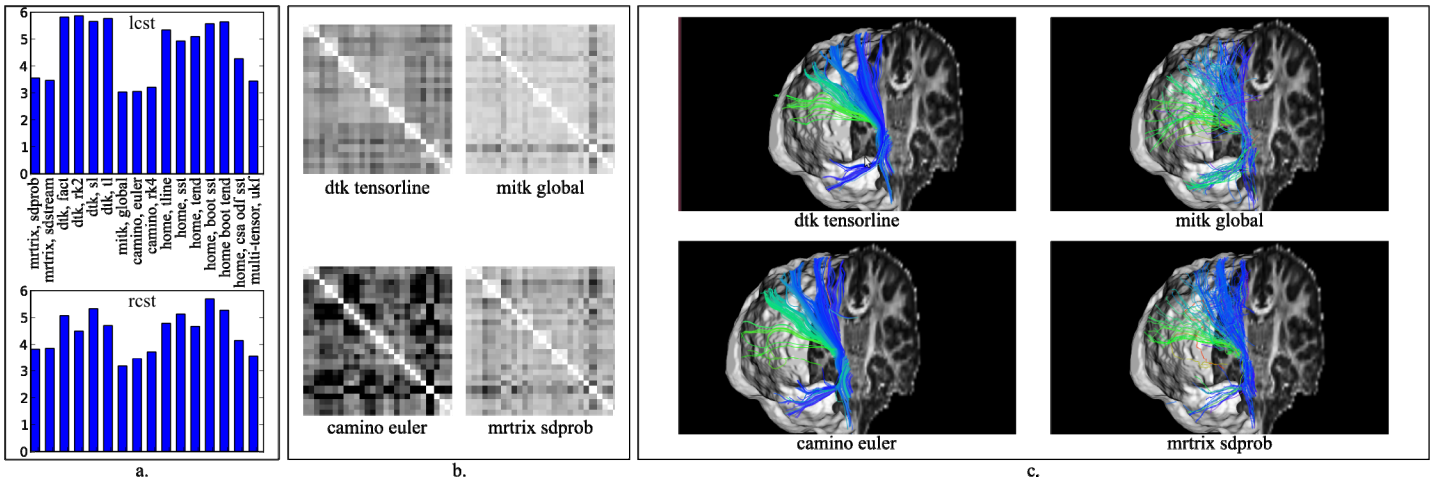

Figure 1. a. Ratio of the average inter-subject distance over the average intra-subject distance. b. similarity matrix for selected methods; repetitions of the same subject are grouped. c. Result of left CST tracking for selected methods.

Conclusions– We compared a total of 16 tracking methods in terms of their ability to capture subject-specific white matter fiber path geometries. For the specific tracking parameters chosen for this study, we can identify 9 methods giving globally superior results (Fig 1.a). A more comprehensive study, including possible variation of the parameters, and extension to a high angular resolution diffusion imaging dataset, will be undertaken in the future.

References– 1. A. Tristán-Vega and S. Aja-Fernández, MedIA 14(2), pp, 205-218 (2010). **2. S. Mori et al.**, Neuroimage 40(2), pp. 570-582 (2008). **3. Y. Ou et al.**, MedIA 15(4), pp. 622-639 (2011). **4. D. Wassermann et al.**, NeuroImage 51(1), pp. 228-241, (2010). **5. J. Tournier et al.**, Int J Imag Syst Tech 22(1), pp. 53-66, 2012. **6. R. Wang et al.**, ISMRM, 2007. **7. M. Reisert et al.**, Neuroimage 54(2), pp. 955-962, 2011. **8. P. A. Cook et al.**, ISMRM, 2006. **9. J. G. Malcom et al.**, MICCAI, pp. 894-909, 2009.