Robust Automated White Matter Pathway Reconstruction for Large Studies

Jalmar Teeuw, Matthan W.A. Caan, and Silvia D. Olabarriaga*

Academic Medical Center Amsterdam, Amsterdam, The Netherlands
jalmar@me.com

Abstract. Automated probabilistic reconstruction of white matter pathways facilitates tractography in large studies. TRACULA (TRActs Constrained by UnderLying Anatomy) follows a Markov-chain Monte Carlo (MCMC) approach that is compute-intensive. TRACULA is available on our Neuroscience Gateway (NSG), a user-friendly environment for fully automated data processing on grid computing resources. Despite the robustness of TRACULA, our users and others have reported incidents of partially reconstructed tracts. Investigation revealed that in these situations the MCMC algorithm is caught in local minima. We developed a method that detects unsuccessful tract reconstructions and iteratively repeats the sampling procedure while maintaining the anatomical priors to reduce computation time. The anatomical priors are recomputed only after several unsuccessful iterations. Our method detects affected tract reconstructions by analyzing the dependency between samples produced by the MCMC algorithm. We extensively validated the original and the modified methods by performing five repeated reconstructions on a dataset of 74 HIV-positive patients and 47 healthy controls. Our method increased the rate of successful reconstruction in the two most prominently affected tracts (forceps major and minor) on average from 74% to 99%. In these tracts, no group difference in FA and MD was found, while a significant association with age could be confirmed.

1 Introduction

White matter bundles are known to degenerate with aging [5] and may also be affected by neurodegenerative diseases. They are therefore extensively studied in population imaging and comparative studies. Different approaches exist to reconstruct pathways in Diffusion Weighted Images (DWI) [4]. Deterministic streamline tractography models a path as a one-dimensional curve, rendering it sensitive to noise. Probabilistic tractography takes the uncertainty in the principal diffusion orientation into account but might miss a sparse pathway that is dominated by another denser pathway. Global tractography defines the tract end-points and searches the space for all possible connections.

* The AGEhIV Study Group and Co-morbidity in Relation to AIDS (COBRA) Study Group.

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 101–108, 2015.
DOI: 10.1007/978-3-319-24553-9_13

Tractography relies on anatomical landmarks that define the pathway. Conventionally, these landmarks are manually annotated by an operator, which is a time-consuming and subjective procedure. This may be overcome by adopting a white matter parcellation atlas, built by populating tract probability maps of a cohort of healthy subjects [6]. Such an atlas averages out inter-subject morphological variation. Instead, landmarks may be retrieved from an atlas, transformed to subject space and used to obtain subject-specific tractography [10]. Additional subject-specific prior information may be obtained from an automated segmentation and parcellation of a high-resolution structural scan. This allows automatic delineation of 18 tracts constrained by underlying anatomy, in a method coined TRACULA (TRActs Constrained by UnderLying Anatomy) [9], which alleviates the need of manual tracing. TRACULA is one of the few methods providing an automated probabilistic method to segment individual pathways.

The adoption of TRACULA in practice imposes two challenges. First, it needs to be robust to be adopted in large imaging studies. For example, Figure 1 shows partial reconstruction in a subset of tracts [2] requiring manual intervention, which would not be feasible for large studies. Second, processing the data is compute-intensive, estimated to one CPU-year for a study of 100 subjects. To facilitate the adoption of such methods in clinical research, in our hospital we offer them as service available on our Neuroscience Gateway (NSG, https://neuro.ebioscience.amc.nl) [7]. From a web interface, the user browses the scans based on metadata, and runs applications such as TRACULA on the selected scans. The computations are performed seamlessly on a grid infrastructure, including the transport of data from the data server to the grid storage, and of results back to the data server. In the case of TRACULA, however, we

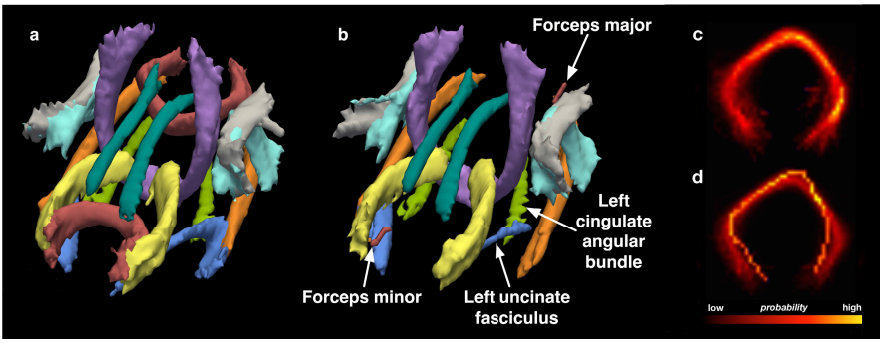

Fig. 1. Example 3D models (isosurface at 20% of maximum probability) of all tracts and maximum intensity projection along the inferior-superior axis of the probability density (PD) maps of forceps major for two repeated runs of TRACULA on the same subject scan. (***a***) successful reconstruction of all 18 tracts. (***b***) partial reconstruction of forceps major, forceps minor, left uncinate fasciculus, and left cingulate angular bundle. (***c***) PD map of successful reconstruction of forceps major. (***d***) PD map of partial reconstruction of forceps major showing artificially high probability for the track sample where the MCMC algorithm is caught in a local minimum.

observed that human interaction would be necessary in a large number of cases to deliver a complete reconstruction of all tracts; therefore we tried to find a solution to reduce this manual intervention.

In this paper we propose a robust and automated white matter pathway reconstruction method that detects unsuccessful tract reconstructions. It iteratively repeats the sampling procedure of TRACULA and updates anatomical priors only if needed. The method is implemented on the NSG, enabling it to be run by clinical researchers on their data.

2 Method

Automated Tract Reconstruction. TRACULA builds upon a Bayesian framework for global tractography [3] in which a pathway $\mathcal{F}$ is estimated from DWIs $\mathbf{Y}$ via the posterior distribution $p(\mathcal{F} \mid \mathbf{Y}) \propto p(\mathbf{Y} \mid \mathcal{F})\, p(\mathcal{F})$. The likelihood $p(\mathbf{Y} \mid \mathcal{F})$ encapsulates the variability in the measured data given the pathway, assuming a "ball-and-stick" diffusion model [8]. The prior $p(\mathcal{F})$ is informed by an anatomical segmentation map of a structural scan [1], providing information of intersecting and adjacent structures along the pathway. The posterior distribution $p(\mathcal{F} \mid \mathbf{Y})$ is estimated via the Markov Chain Monte Carlo (MCMC) algorithm, which after a burn-in period runs a set of sampling iterations. A sample produced by the MCMC algorithm consists of perturbed spatial coordinates (x, y, z) for a set of control points that define a path through the tract by spline interpolation. The perturbation can be accepted or rejected, and in the latter case the perturbation is undone before the next iteration. Practically, separate steps can be run in TRACULA as follows: *a*) preparation (`trac-all -prep`), includes computing anatomical priors; *b*) reconstruction (`trac-all -path`), this runs the MCMC algorithm and produces statistics on the reconstructed tracts; and *c*) preparation step *a* with reinitialization of initial control points (`set reinit = 1` option).

Detecting Local Minima. By default the MCMC algorithm is run for 7500 iterations, but only every 5th iteration produces a sample in order to reduce the dependency between samples, resulting in 1500 samples. We observed that in partially reconstructed tracts a high number of duplicate samples were produced. We suspect that this occurs for example near the edge of a tract, where a random perturbation of control points likely leads to a rejected state outside the white matter. In such cases, $p(\mathcal{F} \mid \mathbf{Y})$ will then be biased to the edge of the tract, resulting in a 'narrow' pathway $\mathcal{F}$. We therefore try to detect when the MCMC algorithm was caught in such local minima by analyzing the presence of duplicate samples in the logs of TRACULA. We consider two samples duplicate when all x, y, z values of corresponding control points are identical. A sample is considered unique when it occurs only once in the set of 1500 samples. A local minimum can then be detected when *a*) there are less than A unique samples in the set; or *b*) the most frequently occurring duplicate has more than B samples in the set.

Robust Pathway Reconstruction. We propose an augmented sampling strategy that aims to automatically recover partially reconstructed tracts. The proposed method consist of three stages. In the first stage we obtain an initial reconstruction of the tracts by running step a followed by step b. In the second stage, partially reconstructed tracts are detected as described above. For these tracts, we attempt to increase the likelihood of obtaining a fully sampled tract by repeating step b maximally k times. In the third stage, if necessary, the anatomical priors $p(\mathcal{F})$ are recomputed and a different set of initial control points is chosen than in previous runs. This is done by running step c followed by step b a maximum of m times. This three-stage approach minimizes computing time by only recalculating partially reconstructed tracts and recomputing anatomical priors only when necessary.

3 Experiments

We performed four experiments to analyze the performance of our method using a large dataset of 121 subjects from the AGEhIV Cohort Study, consisting of 74 HIV positive patients and 47 healthy controls. Subjects were scanned on a 3.0 Tesla Philips Intera system. T1-weighted structural scans were acquired at a resolution of 1mm^3 and DWIs at 2mm^3 along 64 gradient directions with $b = 1000s/mm^2$ and 6 non-diffusion weighted images ($b = 0s/mm^2$). The structural MRI scans are processed with Freesurfer v5.3 [1] to obtain brain segmentations that are used to compute the anatomical priors in TRACULA. The DWI scans are preprocessed with an in-house developed pipeline. The preprocessed DWI scans are then processed with FSL BedpostX [8] v5.0.5 to obtain probabilistic diffusion parameters using a two-compartment model. The output of all three applications is merged and used as input for TRACULA. Default parameters are used for all three software packages. All the experiments were performed using the NSG within one week and consumed an estimated amount of 8228 computing hours (almost a full year) and 726 GB of data.

The *first experiment* assesses reproducibility of the problem of partial reconstructed tracts and validates our approach to detect the occurrence of this problem. We applied the original TRACULA implementation five times to reconstruct all 18 tracts in all 121 subjects. The detection method described in Section 2 was applied off-line to determine the number of unsuccessful reconstructions per scan and individually for each of the tracts in all 605 subject reconstructions. The results of the reconstructions for the first 55 subjects were verified by visual inspection. All 18 tracts were checked to determine if the result of the detection method is a true positive (TP), i.e. correctly detected unsuccessful reconstructions; false positive (FP), i.e. incorrectly detected successful reconstructions which our method superfluously recomputes; false negative (FN), i.e. incorrectly detected unsuccessful reconstructions representing missed cases of partial reconstruction; or true negative (TN), i.e. correctly detected successful reconstructions.

The *second experiment* investigates the behavior of the parameters A and B used by the proposed method. The parameters were varied within plausible

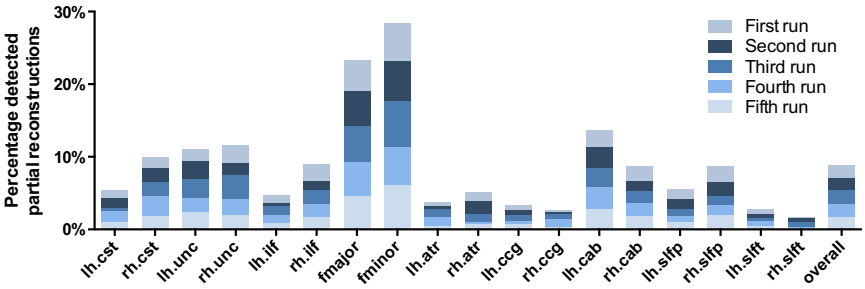

Fig. 2. Detected partial reconstructions of Tracula for five runs (different shadings).

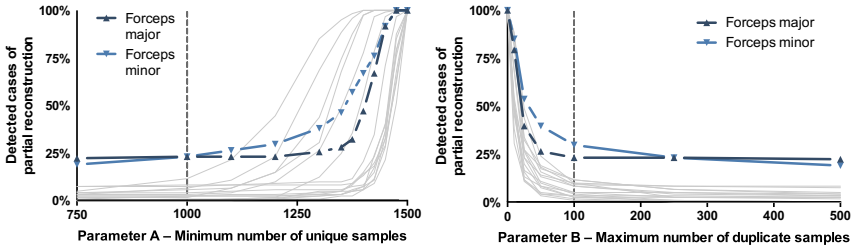

Fig. 3. Response curves per tract as percentage of detected cases for values of parameters A (left) and B (right). Dashed lines indicate values chosen for experiments.

range and the method was repetitively applied to a single run of TRACULA on all 121 subjects. We again applied the detection method from Section 2 and determined the number of failed reconstructions individually for each tract.

The *third experiment* aims at assessing the performance of the proposed method. Thresholds $A = 1000$ and $B = 100$ were set based on results from the second experiment. The other parameters were set empirically to $k = 7$ and $m = 3$, for a maximum of $n = k + m = 10$ iterations. Based on the results of the 605 subject reconstructions with our method we quantified the incidence rate of partially reconstructed tracts. We also assessed the behavior of our method by analyzing the number of iterations necessary to generate a successful reconstruction. We compared the reported volume and the weighted average fractional anisotropy (FA) of the tracts to the results obtained in the first experiment.

Finally, in the *fourth experiment*, a general linear model was fitted to the weighted average FA in the forceps major with age, group and reconstruction method as predictors.

4 Results

Figure 2 illustrates the number of unsuccessful reconstructions for each tract in all the five runs of Tracula from the first experiment. Note that there are variations between the runs, and that some tracts (e.g. the forceps major and

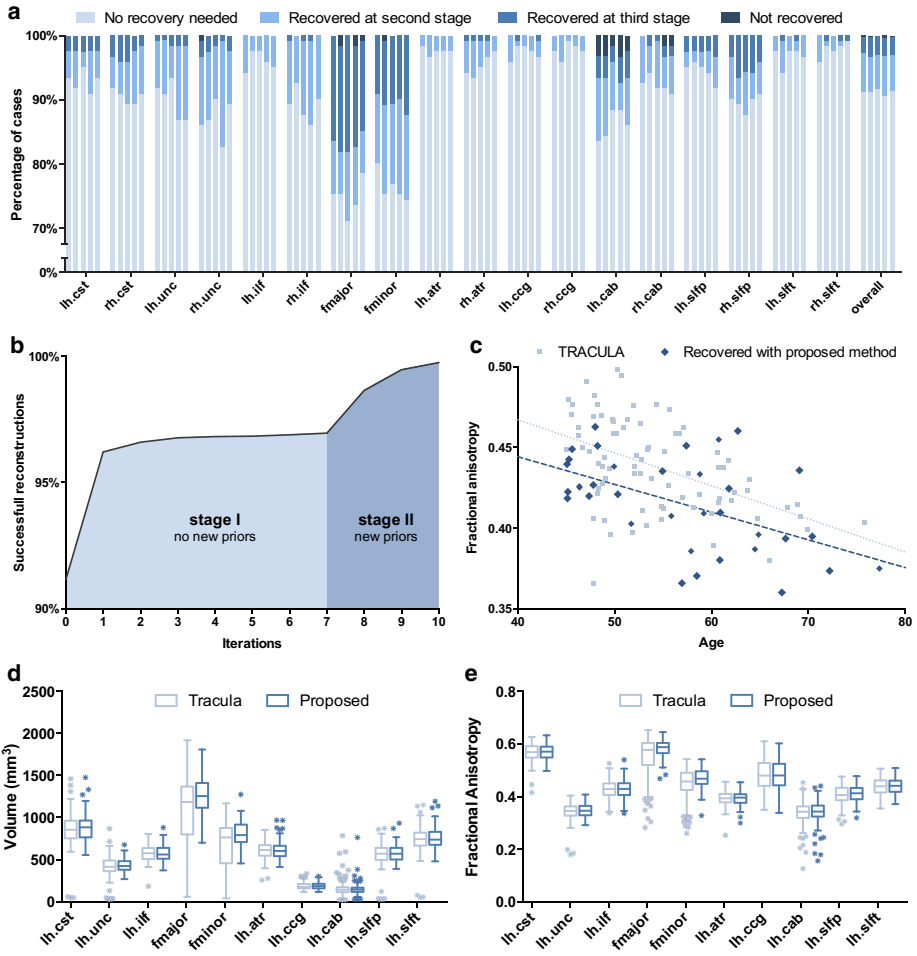

Fig. 4. (***a***) Percentage of successfully reconstructed tracts in each stage of our method (light blue=first, medium blue=second, dark blue=third). In darkest blue are tracts that failed after the maximum number of iterations. The overall (average of all tracts) is shown in the last column. (***b***) Percentage of successfully reconstructed tracts at each iteration of our method. (***c***) Scatter plot of FA as a function of age for TRACULA and proposed method, with trend lines for each group. There is a significant effect of age ($p = 0.015$) and reconstruction method (TRACULA versus proposed method, $p = 0.0002$). (***d***) Box plot of the segmented volumes per tract for TRACULA and the proposed method. Right hemisphere tracts show comparable results to the left hemisphere and are omitted. (***e***) Box plot of the weighted average FA per tract for TRACULA and the proposed method.

minor) are more prone to failure than others. The results of the second experiment are presented in Figure 3. When parameters A, B are too strict, too many reconstructions are detected as partial reconstruction. We chose $A = 1000$ and $B = 100$ as optimal values to be used in the remaining experiments.

Figure 4a synthesizes the results obtained for the original method and the proposed robust method. On overall, 8.8% of the tracts were identified by our method as partial reconstruction. The variability in success rate between repeated runs of the initial reconstruction is low at $\leq 5\%$. Differences between hemispheres are negligible. Figure 4b reveals that 65% of the detected cases were successfully reconstructed during stage two of our method, and another 32% during stage three, i.e. after recomputing anatomical priors. The overall success rate increases from 91.2% to 99.7%. The few failed cases were detected in the forceps major and minor and cingulum angular bundles. From this experiment we can optimize the number of iterations for stage 2 and set $k = 2$ in the future.

Visual inspection of the first 55 subject reconstructions, for a total of 990 tract reconstructions, produced by the first run of TRACULA produced the following results: 5.86% of the detected cases were true positive, 2.63% false positive, 0.71% false negative, and 90.81% true negatives. The cingulum angular bundle and cingulumcingulate gyrus, both thin elongated tracts, are most often wrongly detected as partial reconstruction (i.e. detected as FP).

The effect of the recovery strategy is most prominently seen in the volume measures of the tracts (Figure 4d). The partially reconstructed tracts no longer appear as outliers with small volume when using the proposed method. The 25-75 percentile volume range of the forceps major and minor are reduced by approximately a factor of two and the median volume is marginally increased. Similar effects can be witnessed for the weighted average FA of the tracts (Figure 4e). This supports our suspicion that partially reconstructed tracts are biased toward the edge of the tract, resulting in a lower average FA.

Comparing FA values in the forceps major, no group difference was found, while a significantly association with age could be confirmed ($p = 0.015$, Figure 4c). FA was significantly lower among the subgroup recovered using the proposed method compared to the set reconstructed with TRACULA ($p = 0.0002$).

5 Discussion and Conclusions

The high curvature of the forceps major and forceps minor could have caused a lower success rate compared to the other tracts. Future work may consider increasing the number of control points for these highly curved tracts. In general, optimizing the parameter settings per tract individually, including the number of control points and number of samples, would be recommended.

When applying the method to a dataset of HIV positive patients and healthy controls, the weighted average FA in the successfully reconstructed forceps major using the proposed method was significantly lower than those obtained using TRACULA. The uncertainty in the principal diffusion orientation is known to

be inversely related to FA. This makes tractography more challenging in subjects with low FA and may explain the higher failure rate. Failing to include these subjects in the analysis may bias the statistical comparison between patients and controls, illustrating the importance of robust pathway reconstruction in comparative studies.

We presented a method for robust and automated white matter pathway segmentation. The proposed method provides efficient recovery of partially reconstructed tracts by repeated MCMC sampling, and only having to recompute the anatomical priors when necessary. The method is effective in improving the success rate for the two most prominently affected tracts (i.e. the forceps major and minor) from 74% to 99%, and on overall from 91.2% to 99.7%.

References

1. Fischl, B.: FreeSurfer. NeuroImage 62(2), 774–781 (2012)
2. Hatton, S.N., Lagopoulos, J., Hermens, D.F., Hickie, I.B., Scott, E., Bennett, M.R.: White matter tractography in early psychosis: clinical and neurocognitive associations. Journal of Psychiatry & Neuroscience: JPN 39(6), 417–427 (2014)
3. Jbabdi, S., Woolrich, M.W., Andersson, J.L.R., Behrens, T.E.J.: A Bayesian framework for global tractography. NeuroImage 37(1), 116–129 (2007)
4. Lazar, M.: Mapping brain anatomical connectivity using white matter tractography. NMR in Biomedicine 23(7), 821–835 (2010)
5. Lebel, C., Gee, M., Camicioli, R., Wieler, M., Martin, W., Beaulieu, C.: Diffusion tensor imaging of white matter tract evolution over the lifespan. NeuroImage 60(1), 340–352 (2012)
6. Mori, S., Oishi, K., Jiang, H., Jiang, L., Li, X., Akhter, K., Hua, K., Faria, A.V., Mahmood, A., Woods, R., Toga, A., Pike, G., Neto, P., Evans, A., Zhang, J., Huang, H., Miller, M., van Zijl, P., Mazziotta, J.: Stereotaxic white matter atlas based on diffusion tensor imaging in an ICBM template. NeuroImage 40(2), 570–582 (2008)
7. Shahand, S., Benabdelkader, A., Jaghoori, M.M., al Mourabit, M., Huguet, J., Caan, M.W.A., van Kampen, A.H.C., Olabarriaga, S.D.: A Data-Centric Neuroscience Gateway: Design, Implementation, and Experiences. Concurrency and Computation: Practice and Experience 27(2), 489–506 (2015)
8. Woolrich, M.W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T., Beckmann, C., Jenkinson, M., Smith, S.M.: Bayesian analysis of neuroimaging data in FSL. NeuroImage 45(1), S173–S186 (2009)
9. Yendiki, A., Panneck, P., Srinivasan, P., Stevens, A., Zöllei, L., Augustinack, J., Wang, R., Salat, D., Ehrlich, S., Behrens, T., Jbabdi, S., Gollub, R., Fischl, B.: Automated probabilistic reconstruction of white-matter pathways in health and disease using an atlas of the underlying anatomy. Frontiers in Neuroinformatics 5(23), 1–12 (2011)
10. Zhang, W., Olivi, A., Hertig, S., Zijl, P.V., Mori, S.: Automated fiber tracking of human brain white matter using diffusion tensor imaging. Neuroimage 42(2), 771–777 (2008)