ANATOMICAL PRIORS FOR GLOBAL PROBABILISTIC DIFFUSION TRACTOGRAPHY

Anastasia Yendiki, Allison Stevens, Jean Augustinack, David Salat, Lilla Zollei, Bruce Fischl

HMS/MGH/MIT Athinoula A. Martinos Center for Biomedical Imaging,
149 13th St., Charlestown, MA 02129, USA
ayendiki@nmr.mgh.harvard.edu

ABSTRACT

We investigate the use of anatomical priors in a Bayesian framework for diffusion tractography. We compare priors that utilize different types of information on the white-matter pathways to be reconstructed. This information includes manually labeled paths from a set of training subjects and anatomical segmentation labels obtained from T1-weighted MR images of the same subjects. Our results indicate that the use of prior information increases robustness to end-point ROI size and yields solutions that agree with expert-drawn manual labels, obviating the need for manual intervention on any new test subjects.

Index Terms— diffusion tractography, magnetic resonance imaging, statistical reconstruction

1. INTRODUCTION

Diffusion tractography is the reconstruction of cerebral white-matter fiber bundles from magnetic resonance (MR) images acquired with diffusion weighting, where image intensity at each voxel depends on the orientation(s) along which water molecules diffuse at that voxel. This is a challenging problem, not only because of the size of the solution space, but also because of the uncertainty introduced by imaging noise and multiple true diffusion directions at every voxel.

Many of the proposed solutions to this problem are *local,* in the sense that the tractography algorithm considers the diffusion-weighted (DW) image data at one location at a time to determine the most likely fiber bundle orientation at that location, then steps in that direction to the next location (*e.g.*, [1]). These algorithms, which can be deterministic or probabilistic, are suitable for exploring all possible connections from one brain region, which is used as a starting point, to any other region. However, they may encounter difficulties in isolating a known pathway between two given brain

regions, particularly when the connection to be identified is a weak one and thus dominated by larger pathways originating from either of the regions of interest.

Global tractography methods have been introduced to address these problems by estimating the entire pathway between two regions at once, rather than step-by-step [2, 3]. Although the pathway is typically parameterized in some way to contain the size of the solution space, searching through that space remains cumbersome and sensitive to initialization.

Many tractography methods, whether local or global, attempt to facilitate the search by introducing different types of constraints. Common examples are bounds on the bending angle or the total length of the pathway. Implicitly these constraints are based on prior knowledge of the shape of the pathway of interest in a typical brain. Usually, however, these methods do not offer a mechanism for training these parameters explicitly, instead leaving it up to the user to tune them by trial and error. In addition to tuning parameters, manual intervention is sometimes required to define waypoints to guide the algorithm, particularly when tracing weaker connections. The need for manual processing can make tractography methods especially inefficient for large population studies.

We recently proposed a fully automated method for probabilistic tractography that obviates the need for manual intervention by introducing prior information on the shape of the pathways of interest, derived from a set of training subjects [4]. The method utilizes a likelihood model for global tractography introduced by Jbabdi *et al.* [2]. This model involves a representation of the diffusion process in terms of multiple diffusion compartments, thus allowing for more than one major diffusion direction per voxel [5], and a parameterization of the unknown pathways as splines with a small number of control points which are to be optimized.

In this work we investigate priors that, in addition to a manual labeling of the paths in the training subjects, also utilize information about surrounding anatomical structures in the training subjects and the test subject. This information is derived from automatic segmentations of T1-weighted images of the same subjects, produced by the FreeSurfer software [6]. We present preliminary results from a comparison of paths obtained by the global probabilistic framework using different prior models on a set of ten healthy subjects.

2. BACKGROUND

Let $\mathcal{F}$ be the unknown path, to be estimated from the measured diffusion-weighted (DW) images $\boldsymbol{Y}$ by maximizing the posterior probability:

$$p(\mathcal{F}, \boldsymbol{\Omega}|\boldsymbol{Y}) \propto p(\boldsymbol{Y}|\mathcal{F}, \boldsymbol{\Omega}) \, p(\mathcal{F}, \boldsymbol{\Omega}), \qquad (1)$$

where $\boldsymbol{\Omega}$ are the other unknown parameters in the model.

2.1. Data likelihood model

Following [2], we model the likelihood of the images $\boldsymbol{Y}$ given the path $\mathcal{F}$ as independent samples of a Gaussian distribution with unknown covariance $\boldsymbol{\Sigma}$:

$$p(\boldsymbol{Y}|\mathcal{F}, \boldsymbol{\Omega}) = p(\boldsymbol{Y}|\boldsymbol{\Theta}, \boldsymbol{\Phi}, \boldsymbol{s}_0, \boldsymbol{d}, \boldsymbol{f}, \boldsymbol{\Sigma}) \sim \mathcal{N}(\boldsymbol{\mu}; \boldsymbol{\Sigma}), \quad (2)$$

where the mean $\boldsymbol{\mu}$ involves a multi-compartment forward model of the diffusion process at each voxel [5]. This model represents the expected intensity at the jth voxel in the DW image acquired with the ith diffusion-encoding direction as

$$\mu_{ij} = s_{0j} \Bigg\{ \Big(1 - \sum_{i=1}^{n_F} f_j^l\Big) \exp(-b_i d_j) +$$

$$\sum_{i=1}^{n_F} f_j^l \exp\big(-b_i d_j \boldsymbol{r}_i^T \boldsymbol{R}(\theta_j^l, \phi_j^l) \boldsymbol{A} \boldsymbol{R}^T(\theta_j^l, \phi_j^l) \boldsymbol{r}_i\big) \Bigg\} (3)$$

where the first summand inside the braces represents an isotropic diffusion compartment and the remaining n_F summands represent perfectly anisotropic fiber components with orientations defined by the angles (θ_j^l, ϕ_j^l). The diffusion-encoding weight and direction, b_i and $\boldsymbol{r}_i$ respectively, are known acquisition parameters. The non-DW image intensity s_{0j}, the diffusivity d_j, the anisotropic compartment volume fractions f_j^l, $l = 1, \ldots, n_F$, and the anisotropic compartment orientations (θ_j^l, ϕ_j^l), $l = 1, \ldots, n_F$, are unknown parameters to be estimated. Finally, $\boldsymbol{A}$ is the outer product of a unit vector along the left-right axis and $\boldsymbol{R}(\theta_j^l, \phi_j^l)$ is the rotation matrix that applies a rotation by θ_j^l around the anterior-posterior axis and by ϕ_j^l around the inferior-superior axis.

The prior term in (1) is given by

$$p(\mathcal{F}, \boldsymbol{\Omega}) = p(\boldsymbol{s}_0) p(\boldsymbol{d}) p(\boldsymbol{\Sigma}) p(\boldsymbol{f}|\mathcal{F}) p(\boldsymbol{\Theta}, \boldsymbol{\Phi}|\mathcal{F}) p(\mathcal{F}), \quad (4)$$

where $p(\cdot)$ denotes the prior distribution of its argument. The DW image intensities $\boldsymbol{Y}$ are assumed independent and the unknown variances in $\boldsymbol{\Sigma}$ are nuisance parameters whose prior is integrated out of the posterior; most other terms in (4) are chosen to be non-informative priors, as detailed in [2].

2.2. Path prior model

Once the two regions of interest (ROIs) where the endpoints of the path $\mathcal{F}$ lie are defined, the path is modeled as a piecewise cubic Catmull-Rom spline connecting the endpoints,

with a small number of intermediate control points. The path estimation problem then becomes to position the endpoints and control points of the spline to maximize the posterior (1). Even with a small number of control points and small end ROIs, however, the space of all possible paths $\mathcal{F}$ is still very large, making the optimization problem difficult. Thus, we constrain the solution space further via the prior $p(\mathcal{F})$.

In [2] the only information on $\mathcal{F}$ included in $p(\mathcal{F})$ is whether the two end ROIs are known *a priori* to be connected to each other or not. That is, conditional on the ROIs being connected, all paths that connect them are assumed to be equally probable.

In [4] we proposed to estimate the prior distribution of $\mathcal{F}$ from a set of training subjects where the paths of interest were labeled manually by an expert. Here we formalize the definition of the priors and expand their treatment to include information on labels from an anatomical segmentation. Let $\mathcal{F}_1, \ldots, \mathcal{F}_n$ be splines fitted to the manually labeled control points from each of the n training subjects. Assuming spatial independence and using basic probability theory, it can be shown that the path prior is given by

$$\begin{aligned} p(\mathcal{F}) &= p(\mathcal{F}|\mathcal{F}_1, \ldots, \mathcal{F}_n) \\ &= \prod_{j \in \mathcal{F}} p(j \in \mathcal{F}|\mathcal{F}_1, \ldots, \mathcal{F}_n) \\ &\quad \cdot \prod_{j \notin \mathcal{F}} p(j \notin \mathcal{F}|\mathcal{F}_1, \ldots, \mathcal{F}_n) \\ &= \prod_{j \in \mathcal{F}} \frac{n_j + 1}{n + 2} \prod_{j \notin \mathcal{F}} (1 - \frac{n_j + 1}{n + 2}), \qquad (5) \end{aligned}$$

where j indexes all voxels in the volume and n_j is the number of training paths that include the voxel j. This is similar to estimating the probability of a voxel j belonging to the path from a histogram of the number of training paths that include j, except that it assigns non-zero probability to voxels that aren't included in any of the training paths, thus resulting in less bias towards the training set.

Similarly, conditional on segmentation maps $\mathcal{A}_1, \ldots, \mathcal{A}_n$ for the training subjects and $\mathcal{A}$ for the test subject, as well as the training paths $\mathcal{F}_1, \ldots, \mathcal{F}_n$, the path prior is given by:

$$\begin{aligned} p(\mathcal{F}) &= p(\mathcal{F}|\mathcal{A}, \mathcal{A}_1, \ldots, \mathcal{A}_n, \mathcal{F}_1, \ldots, \mathcal{F}_n) \\ &= \prod_{j \in \mathcal{F}} \frac{n_{j\mathcal{A}(j)} + 1}{n_{j\mathcal{A}(j)} + \bar{n}_{j\mathcal{A}(j)} + 2} \\ &\quad \cdot \prod_{j \notin \mathcal{F}} \frac{\bar{n}_{j\mathcal{A}(j)} + 1}{n_{j\mathcal{A}(j)} + \bar{n}_{j\mathcal{A}(j)} + 2}, \qquad (6) \end{aligned}$$

where $n_{j\mathcal{A}(j)}$ and $\bar{n}_{j\mathcal{A}(j)}$ are the numbers of training subjects whose segmentation label at voxel j is the same as $\mathcal{A}(j)$ and whose paths do or don't include, respectively, the voxel j.

3. PRELIMINARY RESULTS

We performed tractography on a data set provided by the Mental Illness and Neuroscience Discovery (MIND) Institute [7]. The data was collected on 10 healthy volunteers scanned at 1.5T at the Massachusetts General Hospital. The scans included DW images with resolution 2x2x2 mm and 60 gradient directions, and T1-weighted images with resolution 1x1x1 mm. Each subject's DW images were aligned to the individual's T1 image by affine registration. The DW images of all subjects were also aligned to each other in Talairach space by affine registration.

Using each subject's fractional anisotropy and primary eigenvector maps, experts labeled the corticospinal tract (CST), the three subcomponents of the superior longitudinal fasciculus (SLF1, SLF2, SLF3) and the cingulum. The labeling involved placing control points along each path. A spline was then fit to the points to obtain one training path per tract per subject.

For each subject we used the other nine as a training set. All path posteriors were estimated by MCMC. We tested the following approaches: *(i)* Using no prior information from training subjects, *(ii)* Using the manual path labels only to initialize the MCMC algorithm but not in a prior, *(iii)* A prior like (5) based on the manually labeled paths only, *(iv)* A prior like (6) where $\mathcal{A}(j)$ includes the anatomical segmentation labels at j and its six nearest voxels, and *(v)* A prior like (6) where $\mathcal{A}(j)$ includes the six nearest segmentation labels (that are different from the label at j) in the LR, AP, IS directions.

We obtained the seed and target ROIs by finding the centroids of the corresponding manually labeled end points of the nine training subjects. To assess the robustness of each method to ROI size, we dilated the seed and target ROIs simultaneously and repeated the posterior estimation with ROIs of diameter equal to 1, 3, 5, 9, 17, and 33 voxels.

We computed the modified Hausdorff distance of the path posteriors produced by each of the five approaches to the "ground truth," represented by the training path from the test subject itself. We define this distance between two sets of points as the mean of the minimum distance from a given point in the first set to any point in the second set. One such distance measure was obtained for each subject and ROI size combination and is plotted in Fig. 1.

Fig. 2 shows examples of path posteriors obtained using methods *(ii)-(v)*. For each method, the estimated posteriors for the SLF1 of a single subject with the 6 different end ROI sizes have been superimposed. For comparison, the manually labeled points of the SLF1 for the same subject have also been superimposed on the path posteriors.

4. DISCUSSION AND FUTURE WORK

Our results indicate that incorporating path priors yields solutions with better agreement with manual path labels and bet-
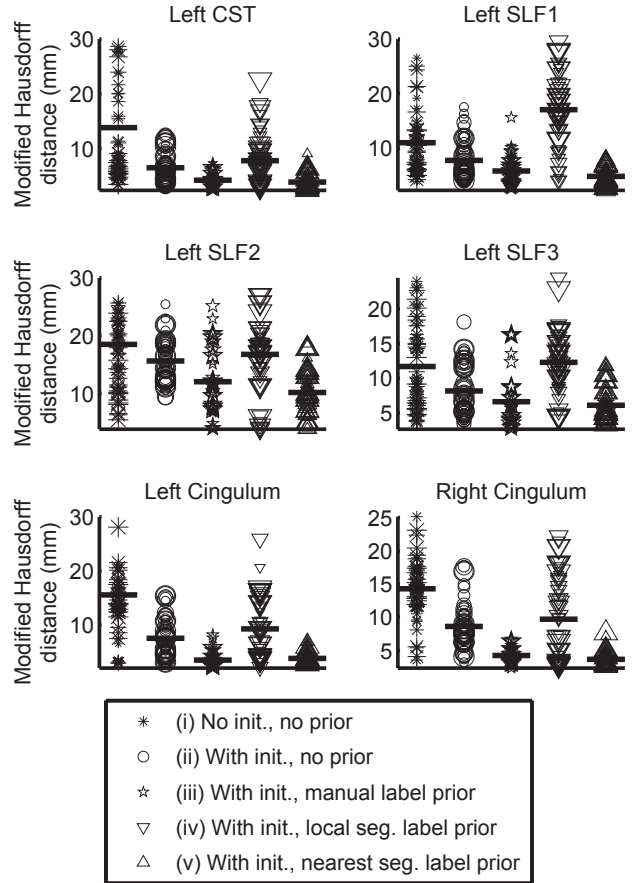

Fig. 1. Modified Hausdorff distance between the estimated path posteriors and the spline fitted to the same subject's manually labeled points. The scatter plots show individual subject/ROI size combinations, with larger markers corresponding to larger end ROIs. The horizontal lines show the average for each method.

ter reliability with respect to end ROI selection. The prior that incorporated information on surrounding anatomical segmentation labels performed better than other priors overall, as illustrated by smaller distances to the manually labeled paths (seen in Fig. 1) and better overlap between posteriors obtained with different end ROI sizes (seen in Fig. 2).

Between the two priors that used anatomical segmentations, results were better when the prior incorporated information on surrounding labels, rather than labels at the local neighborhood of each voxel. This is to be expected, as the local neighborhood along the pathway includes mostly white matter voxels, so it doesn't provide much additional information to constrain the path. The surrounding (mostly non-white-matter) labels, on the other hand, constrain the possible locations of the path more successfully.

Interestingly, the prior that incorporated information on surrounding anatomical segmentation labels performed very
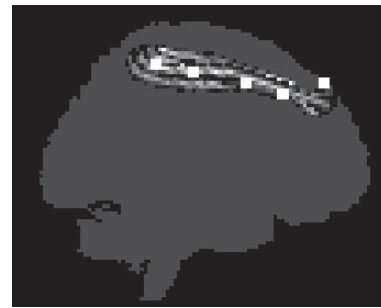
similarly to a prior that used only manually defined labels of the paths of interest. However, compared to the manual label prior, the segmentation prior includes many more parameters whose prior probability has to be estimated from the training data. Thus, the small training set of 9 subjects that was used here may not be sufficient for the segmentation priors to perform optimally. We are currently investigating the use of larger training sets, the sensitivity of the method to inter- and intra-rater variability, as well as improved inter-subject alignment by elastic registration.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Mori, S., Crain, B.J., Chacko, V.P., van Zijl, P.C.: Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. Ann. Neurol. 45, 265–9 (1999).

[2] Jbabdi, S., Woolrich, M.W., Andersson, J.L., Behrens, T.E.J.: A Bayesian framework for global tractography. NeuroImage 37, 116–29 (2007)

[3] Melonakos, J., Mohan, V., Niethammer, M., Smith, K., Kubicki, M., Tannenbaum, A.: Finsler Tractography for White Matter Connectivity Analysis of the Cingulum Bundle. Proc. Intl. Conf. MICCAI, 10(Pt 1):36–43, 2007.

[4] Yendiki, A., Stevens, A., Jbabdi, S., Augustinack, J., Salat, D., Zollei, L., Behrens, T., Fischl, B.: Probabilistic Diffusion Tractography with Spatial Priors. Proc. Intl. Conf. MICCAI Workshop on Computational Diffusion MRI, pp. 54–61, 2008.

[5] Behrens, T.E.J., Johansen-Berg, H., Jbabdi, S., Rushworth, M.F.S., Woolrich, M.W.: Probabilistic diffusion tractography with multiple fibre orientations: What can we gain? NeuroImage 34, 144–155 (2007)

[6] Fischl, B., Salat, D., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M.: Whole brain segmentation. Automated labeling of neuroanatomical structures in the human brain. Neuron, 33(3):341–355 (2002).

[7] Mental Illness and Neuroscience Discovery Institute. http://themindinstitute.org/

(a) Initialization only

(b) Prior using manual labels

(c) Prior using local segmentation labels

(d) Prior using nearest segmentation labels

Fig. 2. SLF1 path posteriors from a single subject with different end ROI sizes, reconstructed with different types of priors. The points from the subject's manual labeling are shown as squares. The manually labeled points from the other 9 subjects were used as a training set.