NeuroImage 130 (2016) 63–76

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Training shortest-path tractography: Automatic learning of spatial priors

Niklas Kasenburg [a,*], Matthew Liptrot [a,b], Nina Linde Reislev [c], Silas N. Ørting [a], Mads Nielsen [a], Ellen Garde [c], Aasa Feragen [a]

[a] Department of Computer Science, University of Copenhagen, Denmark
[b] DTU Compute, Technical University of Denmark, Denmark
[c] DRCMR, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Denmark

ARTICLE INFO

Article history:
Received 18 August 2015
Accepted 12 January 2016
Available online 22 January 2016

Keywords:
Tractography
Diffusion MRI
Graph theory
Prior information

ABSTRACT

Tractography is the standard tool for automatic delineation of white matter tracts from diffusion weighted images. However, the output of tractography often requires post-processing to remove false positives and ensure a robust delineation of the studied tract, and this demands expert prior knowledge. Here we demonstrate how such prior knowledge, or indeed any prior spatial information, can be automatically incorporated into a shortest-path tractography approach to produce more robust results. We describe how such a prior can be automatically generated (learned) from a population, and we demonstrate that our framework also retains support for conventional interactive constraints such as waypoint regions. We apply our approach to the open access, high quality Human Connectome Project data, as well as a dataset acquired on a typical clinical scanner. Our results show that the use of a learned prior substantially increases the overlap of tractography output with a reference atlas on both populations, and this is confirmed by visual inspection. Furthermore, we demonstrate how a prior learned on the high quality dataset significantly increases the overlap with the reference for the more typical yet lower quality data acquired on a clinical scanner. We hope that such automatic incorporation of prior knowledge and the obviation of expert interactive tract delineation on every subject, will improve the feasibility of large clinical tractography studies.

Introduction

Diffusion weighted imaging (DWI) of the human brain provides local estimates of water diffusion summarized as voxel-wise diffusion orientation distribution functions (dODFs) (Hagmann et al., 2006). These can be transformed into fiber orientation distribution functions (fODFs), representing estimates of the fiber directions within a voxel. Subsequently, tractography attempts to delineate the underlying anatomical *tracts* connecting brain regions by inferring inter-voxel connectivity from these fODFs.

Over the last decade it has become increasingly clear how critical the integrity of these white matter (WM) tracts is to the healthy functioning of the brain (Iwata et al., 2011; Abhinav et al., 2014; Connally et al., 2014; Steinbach et al., 2015). Therefore, techniques are needed which allow evaluation and monitoring of microstructural tissue properties to gain insight into the mechanisms underlying brain development, aging and pathology. Metrics of such properties include those derived from DWI data, such as the simple fractional anisotropy (FA) (Xia et al., 2012; McGrath et al., 2013; Whitford et al., 2015), generalized FA (Tang et al., 2010), diffusivity measurements (Davis et al., 2009; Benedetti et al., 2011; Galantucci et al., 2011; Wozniak et al., 2013) or other more complex estimates of microstructure (Assaf et al., 2013; Golestani et al., 2014). Additionally, such tract-based integrity measures can also be derived from other magnetic resonance sequences such as T1 relaxometry or magnetization transfer imaging (Alexander et al., 2011).

When comparing tract-specific features across subjects, it is important that the tracts from which they originate are robustly and reliably reconstructed. Voxel-wise FA values (Iwata et al., 2011; Whitford et al., 2015) or white matter tissue probabilities (Iturria-Medina et al., 2007) are often used to either prune results or guide the tractography. The post-processing of tractography outputs using waypoint or exclusion regions (Connally et al., 2014; Benedetti et al., 2011; Galantucci et al., 2011; Rojkova et al., 2015) is also often necessary to ensure consistent and accurate tract delineation. It would therefore be beneficial, in terms of manual effort and reliability, if domain knowledge about the connection or subject could be automatically included in the tract delineation. This can be achieved by integrating prior information regarding tract location into the tractography framework.

Jbabdi et al. (2007) present a Bayesian framework for global probabilistic tractography, which aims to find the optimal tract between two regions, incorporating prior information. However, the possible priors only include knowledge about the existence or absence of a connection and do not include any prior information about the tract

* Corresponding author at: DIKU North Campus, Universitetsparken 5, DK-2100 København Ø, Denmark.
E-mail address: niklas.kasenburg@di.ku.dk (N. Kasenburg).

http://dx.doi.org/10.1016/j.neuroimage.2016.01.031
1053-8119/

location. Furthermore, due to the large complexity of the problem, the optimal solution is intractable and the tracts are estimated by heuristically sampling from the posterior distribution. Building upon this global tractography framework, Yendiki et al. (2011) include prior anatomical information about the spatial location of a fixed number of segments along each fiber for whole-brain tractography. The prior is computed from training data comprised of manually labeled and verified tracts. Although this reduces the search space and is able to guide the tractography reliably on unseen data without manual intervention, it still requires that domain experts post-process the tractography training data and it may need to be repeated for any future novel tasks or derived features.

We present an extension to a shortest-path tractography framework that can include any type of prior information about the spatial location of a tract. Such prior information could, for example, consist of per-voxel white matter probabilities, which guide tractography through white matter, or anatomical knowledge in the form of a tract atlas. Our algorithm also allows priors that are generated from expert annotation, similar to Yendiki et al. (2011). In addition, we present a method for automatically learning such spatial priors from previous tractography results. We demonstrate, in particular, how a prior learned on independent, high quality data, where tract delineation is easier and more accurate, is able to improve the performance of tractography on lower quality data.

Our framework employs a shortest-path tractography (SPT) approach, which finds the globally optimal path connecting two voxels. Like the framework described by Jbabdi et al. (2007), SPT has the advantage of being less susceptible to local noise in the data because it evaluates all possible connections. Moreover, because the discretization into a graph allows the use of optimal graph-based shortest-path algorithms, graph-based SPT methods (Iturria-Medina et al., 2007; Zalesky, 2008; Sotiropoulos et al., 2010; Vorburger et al., 2013) are guaranteed to find the best path connecting any pair of voxels. In contrast, their continuous counterparts (Lenglet et al., 2004; O'Donnell et al., 2002; Fuster et al., 2014; Schober et al., 2014; Hauberg et al., 2015), as well as the probabilistic approach by Jbabdi et al. (2007), require a good initialization to avoid local optima. While existing graph-based SPT algorithms often impose strict assumptions upon the form that the f/dODF may take, our framework gives full modeling flexibility by permitting any form of fODF. We obtain a Bayesian SPT algorithm by interpreting spatial priors as soft or hard constraints on tract location. As existing graph-based tractography methods do not provide algorithmic solutions to constrained tractography problems (Iturria-Medina et al., 2007; Sotiropoulos et al., 2010), we furthermore derive intuitive, exact and efficient algorithmic solutions to incorporate prior information from multiple sources into our tractography framework.

In addition to determining the most probable path for the tract connecting two voxels, our SPT algorithm also returns a confidence score which provides a quantitative measure of how well a shortest path is supported by both the underlying fODFs of all component voxels and by the prior information. This "importance" evaluation of any path provides a numerical score that permits our framework to automatically learn a tract prior from training data without requiring expert interaction.

In the next section we briefly review graph-based shortest-path tractography and how it can be applied for region to region global tractography. We then describe how we integrate prior information in the "Shortest-path tractography with spatial priors" section. In the "Data" section we describe the two datasets used throughout this study, how they were pre-processed and how the tractography experiments were performed. We also describe the reference used for the evaluation of tractography results. In the "Tractography results" section, we show the results of the tractography, first using a simple subject-specific prior given by white matter probability, second using study-specific or independent learned priors, and finally using both simple and learned priors in combination with a binary waypoint prior. We show, in particular, the results for tractography on the dataset acquired on a typical clinical scanner with a prior learned from the high quality dataset. We conclude with a discussion of the results and a brief conclusion.

Revisiting graph-based shortest-path tractography

In this section we review graph-based tractography and phrase its solution as a shortest-path problem, which will allow us to efficiently integrate spatial priors into the tractography algorithm in the following.

From the DWI data of a brain we extract an undirected brain-graph $G=(V,E,w_E)$ whose node set V contains all the DWI voxels within the brain, excluding those classified as cerebrospinal fluid (CSF) by prior tissue segmentation. Each node is connected by an edge $e\in E$ to all white matter voxels in its $3\times3\times3$ neighborhood on the 3D image grid. Each edge e is assigned a weight $w_E(e)\in[0,1]$ reflecting the probability of a fiber bundle connecting its two endpoint nodes; this process is described in the next section.

A path $\pi_{v,v'}$ connecting nodes (or voxels) $v\in V$ and $v'\in V$ in G is defined as a sequence of nodes $\pi_{v,v'}=[v_1,v_2,\ldots,v_n]$, where $v_1=v$, $v_n=v'$ and $(v_i,v_{i+1})\in E$ for all $i=1,\ldots,n-1$. The cardinality $|\pi_{v,v'}|=n$ of a path is given by the number of nodes in the path. The likelihood of the path $\pi_{v,v'}$ is defined as the product of all edge weights $w_E(e)$ encountered along the path:

$$\mathcal{L}(\pi_{v,v'})=\prod_{i=1}^{n-1} w_E(v_i,v_{i+1}). \tag{1}$$

From fODF to edge probability

For each voxel, assume that the diffusion information from the DWI is summarized in an fODF $f:S^2\to R_+$ associating to any given direction θ on the unit sphere S^2 a probability that there exists a fiber along that direction. We are interested in the 26 directions θ_i with $i=1,\ldots,26$, pointing from the center of a voxel toward its 26 neighboring voxels.

We model the connectivity $w(\theta_i)$ along an edge from the voxel center in the direction $\theta_i\in S^2$ by integrating the fODF over the set C_i of all directions $\theta\in S^2$ pointing out of the voxel that are closer to θ_i than to any other of the 26 directions θ_j, $j\neq i$. The set C_i is called a Voronoi cell (Voronoi, 1908). Since computing integrals over Voronoi cells on the sphere is computationally hard, we numerically approximate the integral through sampling. The weight $w(\theta_i)$ describes the probability of connection in the direction θ_i and is defined and approximated as follows:

$$w(\theta_i)=\int_{C_i} f(\theta)\ d\theta \approx \sum_{\tilde{\theta}_k\in S_i}\left(f\left(\tilde{\theta}_k\right)\cdot\frac{\mathrm{Vol}\left(S^2\right)}{N}\right), \tag{2}$$

where the set $S=\{\tilde{\theta}_k\in S^2, k=1:N\}$ is a uniform sample of N directions, $S_i=S\cap C_i$ is the set of direction samples belonging to C_i and $\mathrm{Vol}(S^2)/N$ is the average volume corresponding to a sample direction $\tilde{\theta}_k$. As $w(\theta_i)$ depends on its source node, the edge weight $w_E(v,v')$ is defined as the average:

$$w_E(v,v')=1/2\cdot(w(v\to v')+w(v'\to v)), \tag{3}$$

where $v\to v'$ is the direction from v to v'. This yields an undirected graph with symmetric edge weights: $w_E(v,v')=w_E(v',v)$.