K-CONFIDENCE: ASSESSING UNCERTAINTY IN TRACTOGRAPHY USING K OPTIMAL PATHS

Colin J. Brown, Brian G. Booth, Ghassan Hamarneh

Medical Image Analysis Lab, School of Computing Science
Simon Fraser University
{cjbrown, bgb2, hamarneh}@sfu.ca

ABSTRACT

Tractography algorithms are susceptible to errors due to poor image quality. As a result, having some measure of confidence in the anatomical accuracy of a tractography result is a desirable goal. We propose that such a measure of confidence can be obtained from the spatial distribution of likely paths between two fixed endpoints. We present for the first time a k optimal paths tractography algorithm for determining the k most likely fiber tract trajectories between two fixed regions. We further examine the spatial spread of the k optimal fiber paths to obtain a measure of confidence that two regions are connected by an axonal fiber. We show on both synthetic and real data that the uniformity of the spread of the k optimal paths shows good correspondence with anatomically known connectivity.

Index Terms— Tractography, Diffusion MRI, k Optimal Paths, Uncertainty

1. INTRODUCTION

In recent years, diffusion MRI (dMRI) has emerged as an important imaging modality for non-invasively examining connectivity in the brain. Diffusion MRI captures the direction and magnitude of water diffusion at each voxel in a target volume. In white matter brain tissue, this diffusion flows preferentially along neural fibers. Tractography algorithms use this fact to attempt to identify and delineate these fibers from dMRI scans, thereby discovering neuronal connectivity in white matter regions of the brain.

Performing tractography accurately is a difficult problem for various reasons, the most significant being that neural fibers are orders of magnitude smaller than the resolution of any available dMRI scan [1]. A single voxel in an image often contains multiple paths, potentially crossing over one another in different directions [2, 3]. The inevitable introduction of noise and other imaging artifacts can corrupt the already low-resolution dMRI data, leading tractography algorithms off of the true directions of axonal pathways [1]. As a result of these imaging limitations, connectivity between any two regions of the brain can be uncertain. Thus, given two particular regions in a dMR image of the brain, our goal here is way of identifying ambiguity in whether a connecting pathway between the two regions exists.

To date, uncertainty in tractography has not been thoroughly examined. Existing streamline tractography algorithms generate paths by following the strongest direction of diffusion at each point, but give us no measure of uncertainty of each generated path [4]. Tractography methods which propagate a front outward, like fast marching and shortest-path methods [5, 6], generate globally optimal paths, however they typically only generate one possible path between any pair of points. Due to the imaging limitations listed above, we can not guarantee that a globally optimal path is in fact the most anatomically correct one. Further, we can make no guarantees that this found path is uniquely optimal. Probabilistic tractography can give us some measure of likelihood of starting at one point and reaching another, typically by counting what fraction of randomly sampled paths from the seed point pass through the selected end point [4]. Unfortunately, probabilities returned from probabilistic tractography decrease as a function of path length. Further, these probabilities are measured from only a *single* fixed endpoint, and thus information about the connectivity between any two regions is incomplete. In contrast, we require a model which captures the variability in path trajectory when *both* endpoints are fixed. Only when fixing both endpoints will any remaining variability then be the result of path ambiguity.

With the above ideas in mind, we present a k optimal paths approach to examine the uniqueness and trajectory of paths connecting two regions. Our contributions are two-fold. First, we present for the first time a k optimal path tractography algorithm based on Yen's 1971 globally optimal k shortest paths algorithm [7]. Second, we use the spatial distribution of these k optimal paths to obtain a measure of confidence in whether our two endpoint regions are anatomically connected. Specifically, we show on both synthetic and real diffusion MRI data that the uniformity of the spread of the k optimal paths suggests that there are no sections of the tract where the tractography algorithm "gets confused". This uniformity of path spread is therefore a good indicator of anatom-

ical connectivity. To our knowledge, quantifying uncertainty in tractography through the spatial position of likely paths has not been previously examined.

2. METHODS

Our proposed method proceeds in three steps; path discovery, statistical modeling, and confidence measuring. We assume as input a dMR image of the brain $\mathcal{I} : \Omega \rightarrow \mathcal{M}$, represented as a mapping from the voxels in image space, Ω, to the manifold, $\mathcal{M}$, of possible fiber orientation distribution functions (fODF). An fODF, $\phi_x : \mathcal{S} \rightarrow [0, 1]$, at some point, $x \in \Omega$, maps each point on the sphere, $\mathcal{S}$, to a probability of neuronal connectivity. An fODF image can be computed from a set of diffusion weighted MR images using a variety of different methods (*e.g.* [2]).

2.1. Path Discovery

In order to examine the ambiguity in the optimal fiber paths between two regions, we must first discover the k most likely fiber paths. In [6], Zalesky efficiently discovers the path of maximum probability by using Dijkstra's algorithm on a graph weighted by the fODF probabilities between neighboring voxels in the original dMR image. Here, we extend this approach, using Yen's algorithm [7], to find the k most probable paths. For this task, we require a graph-based representation of our dMRI scan.

Given $\mathcal{I}$, we construct an undirected graph, $G(V, E)$, such that each node $v \in V$ corresponds to a voxel in $\mathcal{I}$ and edges $E \subset V \times V$ connect each voxel, $v \in V$, to its set of 26 spatially closest neighbours in 3D. Similar to other graph based tractography methods [8, 9], edge weights are derived from the probability of a neural fiber tract connecting neighboring voxels. For neighboring voxels $u, v \in \Omega$, the probability, p_{uv}, of an outgoing edge $(u, v) \in E$ is calculated by integrating over the fODF at that voxel in a conical region, β, surrounding the edge direction

$$p_{uv} = \frac{1}{Z} \int_{\beta} \phi_u \left(\frac{v-u}{|v-u|} \right) dS + \int_{\beta} \phi_v \left(\frac{u-v}{|u-v|} \right) dS. \quad (1)$$

Like in [8, 9], we have selected a solid angle of $\beta = 4\pi/26$ steradians for our region of integration surrounding the edge direction. The normalization constant Z is set as described in [8]. Equation (1) is computed analytically using [10].

2.1.1. k Optimal Paths Tractography

Given two regions of interest in the image, we choose two cliques, $A, B \subset V$, that correspond most closely to the two regions of interest, respectively. The k optimal paths between the source region A and the destination region B can be obtained by extending Yen's algorithm [7] to the realm of tractography. The resulting k optimal path tractography algorithm is given in Algorithm 1.

Algorithm 1 k Optimal Paths Tractography

INPUT : $G = (V, E)$, SEED $A \subset V$,
TARGET $B \subset V$, NUMBER OF PATHS k
OUTPUT : PATHS P
$examinedNodes \leftarrow \emptyset$
$P\{1\} \leftarrow Dijkstra(G, A, B)$
$candidatePaths \leftarrow \{P\{1\}\}$
for $i = 1$ **to** $k - 1$ **do**
  for $j = 1$ **to** $|P\{i\}|$ **do**
    if $P\{i\}(j) \in examinedNodes$ **then**
      continue
    end if
    $examinedNodes \leftarrow examinedNodes \cup P\{i\}(j)$
    $root \leftarrow P\{i\}(1..j)$
    $SR \leftarrow SameRoot(P, root)$ # Paths incl. $root$
    $V' \leftarrow V/\{u \in root\}$
    $E' \leftarrow E/\{(u, v) \in E | u \notin V' \vee v \notin V'\}$
    $E' \leftarrow E'/\{(P\{i\}(j), w) \in E |$
      $(P\{i\}(j), w) \in SR\}$
    $G' \leftarrow (V', E')$
    $spur \leftarrow Dijkstra(G', P\{i\}(j), B)$
    $candidatePaths \leftarrow candidatePaths \cup$
      $(root + spur)$
  end for
  $P\{i + 1\} \leftarrow MinDist(candidatePaths)$
  $candidatePaths \leftarrow candidatePaths/P\{i + 1\}$
end for

The algorithm begins by calling Dijkstra's algorithm to find the shortest path from A to B. Let P represent the set of all previously discovered shortest paths. Yen's algorithm recursively finds each next shortest path by traversing the nodes $v_j \in P$. At each node, v_j, on an existing shortest path, the algorithm calls Dijkstra's on a modified graph G'. This modified graph is a copy of G with edges removed to prevent the reuse of any previously discovered path to B. In particular, it finds all previously discovered paths with the same root as the current path (*SameRoot(...)*) and removes their edges outgoing from v_j. Thus, we find the shortest path from v_j to B that deviates from all other found shortest paths. In a given iteration, we refer to the path from A to v_j as the root, v_j as the deviation node, and the path from v_j to B as the spur. A path joining the current root and spur is added to a list of candidate paths. Once the traversal of all nodes $v_j \in P$ is complete, the shortest of the candidate paths (*MinDist(...)*) is added to P. This loop is repeated k times to obtain the k shortest paths.

To perform the Dijkstra's algorithm step, we employ Zalesky's minimal path tractography algorithm [6]. In Zalesky's algorithm, the negative-log of the edge weights in (1) is used

to ensure that the minimal path is also the most probable fiber path. Further, Zalesky's algorithm employs a curvature constraint to ensure that no steps with greater than 90° curvature appear in the resulting paths.

2.2. Path Trajectory Model

We now aim to analyze the ambiguity in optimal path trajectory. To do so, we need to be able to compare the trajectory of paths and perform statistical analysis on the overall distribution of trajectories. Each path is given as a sequence of points in $\mathbb{R}^3$ (*i.e.* the spatial locations of each voxel/node on the path). Since a path may traverse a different number of edges, paths are re-sampled such that the Euclidean distance between any two sequential nodes is constant within each path and each path has exactly n nodes. The j^{th} point in the i^{th} path will be denoted $\mathbf{x}_j^i$.

2.2.1. *Path Distribution Model Fitting*

For the sake of simplicity, we assume each point on each of our k optimal paths is a sample from a multivariate normal distribution. Under this assumption, we characterize our set of k optimal paths by a mean path $\overline{\mathbf{x}} = (\overline{\mathbf{x}}_1, \cdots, \overline{\mathbf{x}}_n)$, and a vector of point-wise distances $\overline{\mathbf{d}} = (\overline{d}_1, \cdots, \overline{d}_n)$, where,

$$\overline{\mathbf{x}}_j = \frac{1}{k}\sum_{i=1}^{k} \mathbf{x}_j^i, \qquad \overline{d}_j = \frac{1}{k}\sum_{i=1}^{k} \|\mathbf{x}_j^i - \overline{\mathbf{x}}_j\|.$$

Note that the distances d_j effectively capture, in scalar form, the spread in our data set at each point j. A similar statistical treatment of paths has been discussed in [11] except without two fixed endpoints, and so in that work paths had to be aligned first.

2.3. k-Confidence

With the above statistical model, we can reason as to how the values of the above model parameters vary under different connectivity situations. While equating uncertainty to the magnitudes of the point-wise distances may be an intuitive step to take, those point-wise distances are dependent on the ratio between path length and k. Since paths may not fully overlap, for a fixed k, closer end-points will tend to have a wider spread of shortest paths than end-points further apart. Instead, we examine the change in path spread. We hypothesize that if the k optimal paths traverse an area of high uncertainty, then we expect these paths to show significantly higher variability in this area, resulting in a 'bulge'. In these cases, we should be less confident in the overall connectivity between end-points.

To capture this intuition, we propose k-confidence as the inverse variance in point-wise distances along the mean path:

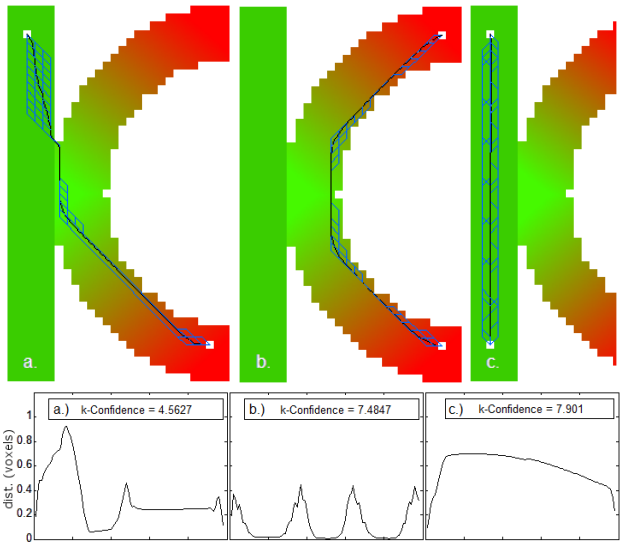


Fig. 1. Three k optimal paths tractography results on a phantom from [9]. Leftmost result is not anatomically correct while the other two are. Point-wise distance plots show greater variance in the leftmost example than in the other two cases, leading to lower k-confidence.

$$c = \sum_{j=1}^{n} (\overline{d}_j - \overline{d}_{avg})^{-2} \tag{2}$$

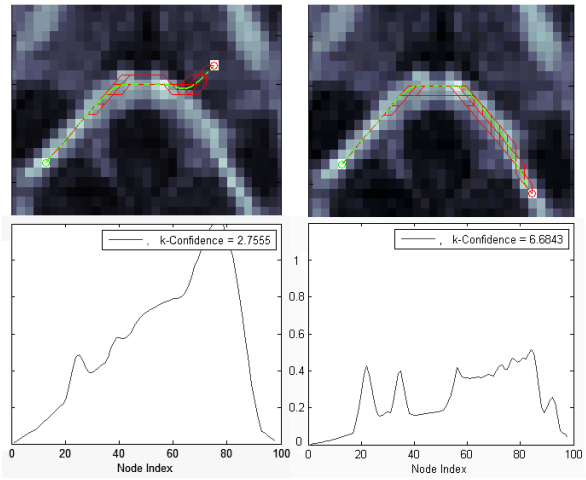
where $\overline{d}_{avg} = 1/n \sum_j \overline{d}_j$. In general, we have higher confidence that an axonal fiber underlies the paths if the set of k optimal paths stay, on average, a uniform distance away from the mean path.

3. RESULTS

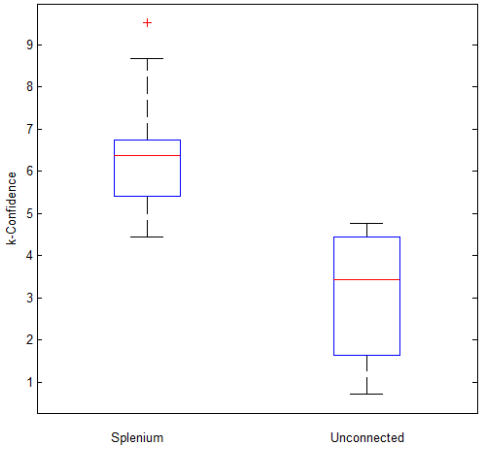
In this section, we motivate our k-confidence measure further by empirically demonstrating its use on both phantoms and real data. In all cases, we set the lone parameter $k = 500$.

3.1. Phantom Data

The top of Figure 1 shows examples of the k optimal paths (blue) computed between points on a DT-MRI phantom of kissing vertical and c-shaped tracts. The resulting mean path is shown in black. Note in the leftmost example that a large 'bulge' of paths occurs in the vertical tract above where the mean path cuts across fiber bundles. This bulge is further evident when we plot the point-wise distances in the lower panel of Figure 1. Note that while the magnitudes of these distances depend on path length and k, the variance in path distance remains a solid indicator of path confidence. The leftmost example, where paths cut across fiber bundles, shows a k-confidence score over one-third less than the other two true connectivity examples (4.56 *vs.* 7.48 and 7.90, respectively).



(a) Representative k optimal paths and k-confidence results



(b) k-confidences over whole dataset

Fig. 2. a) k optimal paths results (in red) and mean paths (green) for the splenium and for two anatomically unconnected points. Note the greater variability in path position in areas of low fractional anisotropy. This focused variability leads to lower k-confidence for the two unconnected points than for the known splenium connection.

3.2. Real DT-MRI Data

We further examine our k-confidence measure on 12 real brain DT-MR images sampled from the John Hopkins MILBA MRI *In Vivo* Human Database [12]. On all 12 images, we perform k optimal path tractography within the splenium (a known strong neuronal connection) as well as between two anatomically unconnected points: one in the left end of the splenium and one in the right ventricle.

Figure 2(a) shows a representative example of k optimal paths obtained in both the splenium and for the two unconnected points. The resulting point-wise distances show high variance in areas with low fractional anisotropy. This results in a much lower k-confidence value for the two unconnected points. Figure 2(b) displays the resulting k-confidence values for both cases over all 12 images. Note that k-confidence values are significantly lower for the unconnected points, suggesting that k-confidence is an effective measure of our confidence that two regions are anatomically connected.

4. CONCLUSIONS AND FUTURE WORK

We presented a novel k optimal path tractography algorithm to extract the k most likely fiber tracts connecting two endpoints. We further proposed that the spatial variability of path position can be an effective measure of our confidence that two brain regions are connected by an axonal fiber tract. Finally we've shown that using these k optimal paths, we can come up with a measure called k-confidence that corresponds well to our known confidence in the connections between two regions. Future work will focus on examining how to set the lone parameter k and its effect on k-confidence values.

5. ACKNOWLEDGEMENTS

CJB and GH were partially supported by NSERC and BGB by IODE Canada and the Government of Alberta.

6. REFERENCES

[1] R. Bammer, B. Acar, and M. E. Moseley, "In vivo MR tractography using diffusion imaging," *J. Radiology*, vol. 45, no. 3, pp. 223, 2003.

[2] M. Descoteaux, R. Deriche, TR Knosche, and A. Anwander, "Deterministic and probabilistic tractography based on complex fibre orientation distributions," *IEEE Trans. Med. Imag.*, vol. 28, no. 2, pp. 269–286, 2009.

[3] S. Mori and P. C. M. van Zijl, "Fiber tracking: principles and strategies - a technical review," *NMR Biomed*, vol. 15, no. 7, pp. 468–480, 2002.

[4] O. Friman, G. Farneback, and C. F. Westin, "A bayesian approach for stochastic white matter tractography," *IEEE Trans. Med. Imag.*, vol. 25, pp. 965–978, 2006.

[5] Geoffrey J.M. Parker, Claudia A.M. Wheeler-Kingshott, and Gareth J. Barker, "Estimating distributed anatomical connectivity using fast marching methods and diffusion tensor imaging," *IEEE Trans. Med. Imag.*, vol. 21, pp. 505–512, 2002.

[6] A. Zalesky, "DT-MRI fiber tracking: A shortest paths approach," *IEEE Trans. Med. Imag.*, vol. 27, no. 10, pp. 1458–1471, 2008.

[7] J. Y. Yen, "Finding the k shortest loopless paths in a network," *Mgmt. Sci.*, vol. 17, no. 11, pp. 712–716, 1971.

[8] Y. Iturria-Medina, E.J. Canales-Rodríguez, L. Melie-García, P.A. Valdés-Hernández, E. Martínez-Montes, Y. Alemán-Gómez, and J.M. Sánchez-Bornot, "Characterizing brain anatomical connections using diffusion weighted MRI and graph theory," *NeuroImage*, vol. 36, no. 3, pp. 645–660, 2007.

[9] B.G. Booth and G. Hamarneh, "Global multi-region competitive tractography," in *Proc. of IEEE MMBIA*, 2012, pp. 73–78.

[10] B. G. Booth and G. Hamarneh, "Exact integration of diffusion orientation distribution functions for graph-based diffusion MRI analysis," in *Proc. of ISBI*, 2011, pp. 935–938.

[11] I. Corouge, P.T. Fletcher, S. Joshi, S. Gouttard, G. Gerig, et al., "Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis," *Medical Image Analysis*, vol. 10, no. 5, pp. 786–798, 2006.

[12] S. Mori, "John Hopkins Medical Institute: Laboratory of Brain Anatomical MRI, in vivo human database," http://lbam.med.jhmi.edu/, accessed February 2013.