Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

A Kernel-Based Approach for User-Guided Fiber Bundling using Diffusion Tensor Data

Raúl San José Estépar   Marek Kubicki   Martha Shenton   Carl-Fredrik Westin

Abstract— **This paper describes a novel user-guided method for grouping fibers from diffusion tensor MRI tractography into bundles. The method finds fibers, that passing through user-defined ROIs, still fit to the underlying data model given by the diffusion tensor. This is achieved by filtering the data and the ROIs with a kernel derived from a geodesic metric between tensors. A standard approach using binary decisions defining tracts passing through ROIs is critically dependent on ROIs that includes all trace lines of interest. The method described in this paper uses a softer decision mechanism through a kernel which enables grouping of bundles driven less exact, or even single point, ROIs. The method analyzes the responses obtained from the convolution with a kernel function along the fiber with the ROI data. Results in real data shows the feasibility of the approach to fiber bundling.**

I. INTRODUCTION

Diffusion tensor imaging (DTI) provides a noninvasive way to measure water diffusion in the brain. The diffusion of the water is based on the structural organization of the fibers that restrict the moment of the water molecules along the fiber orientation. This argument has motivated the use of DTI as a way of finding *in vivo* organizations of the nervous pathways by means of tractography. From a neuroscience point of view, DTI does not provide a plausible model for brain connectivity; however recent studies as those in [1] have shown the correlation between main neural pathways, i.e. fiber tracts, and the principal direction of diffusion (PDD) given by the main eigenvector of the diffusion tensor. Taking the validity of this assumption, tractography methods yield a set of pathways, i.e. fibers, by following the PDD.

Once a fiber set has been obtained by means of a tractography method [2], [3], the next step to achieve clinical relevance is to resolve significant neuropathways that form homonymous fiber bundles. In other words, we want to perform a classification of a fiber tree according to a connectivity criterion so that the organization can be described quantitatively. This is a daunting task that deals with not only the complexity of the problem itself but also with the deficiencies of the tractography methods and the diffusion tensor model. In recent years, significant efforts have been carried out to tackle this problem. Two main approaches have been followed:

- User-guided approaches: these methods try to identify neuropathways based on prior knowledge provided by an expert in terms of regions of interest (ROI). The expert delineates several ROIs that correspond to the pathways he/she wants to extract. Then, the method selects the fibers that pass through the ROIs from the fiber set previously obtained. Methods of this kind have been previously reported in [4], [5], [6], [7], [8]. The main advantage of these methods is that they provide results that are more robust compared to tractography based on seeding in a single ROI as well as providing anatomically relevant connectivity information due to prior user knowledge. The main disadvantages are that they are not automatic.
- Clustering approaches: these methods try to *learn* organizations in the fiber sets from their own data [9]. The main advantages of these methods compared to the former ones is that they do not require user intervention so the results are not biased toward variabilities in the ROI delineation. On the other hand, the main shortcoming is that the cluster yielded by these methods does not necessarily correspond to a defined neuropathway.

The method proposed in this paper could be classified as a user-guided approach for fiber clustering. Our approach is unique in the sense that it tries to deal with some of the disadvantages of the user-guided approaches previously reported in the literature, namely, dependency on the accuracy of the ROI delineation. Those methods at heart followed the same approach: fiber bundles are obtained by finding those fibers that, in a geometric sense, cross through the selected ROIs. This means that the underlying diffusion data, i.e., the diffusion tensor is only used in the tractography and the bundling is left in hands of the precision of the ROI delineation. Our method tries to overcome this dependency by presenting a method based on ROIs but driven by the diffusion tensor data. Our method favors fibers that, in addition to going through the ROIs, fit the underlying tensor data. This is achieved by analyzing the responses along the fiber after convolving along the fiber trajectory a kernel derived from the tensor data with the ROIs labelmap.

The paper is outlined as follows: Section 2 describes the convolution along fibers used in this paper. Section 3 describes our method to perform fiber bundling using ROIs and the diffusion data. In section 4 we will show results that reflect the relevance of the proposed method. Finally, section 5 provides a discussion of the results.

R.S.J.E. and C.-F.W. are with Laboratory of Mathematics in Imaging. {rjosest}@bwh.harvard.edu
MK and MS are with Psychiatry Neuroimaging Laboratory.
All within the Brigham and Women's Hospital, Boston, MA.

This work has been supported by grant P41-RR13218 and the Morphometry BIRN (Biomedical Informatics Research Network, http://www.nbirn.net), a National Center for Research Resources Project(U24-RR021382)

II. FIBER CONVOLUTION

A fiber, Γ, can be defined as parametric set of points such that

$$\Gamma(\gamma) = (x(\gamma), y(\gamma), z(\gamma))^T \qquad \text{with} \quad \gamma \in [0,1]. \tag{1}$$

Then a fiber describes a trajectory in the 3D space.

Now, we introduce the definition of *fiber convolution* denoted by the operator $\circledast_\Gamma$. The convolution along a fiber Γ of a signal field $\mathbf{f}(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}^m$ with a kernel function $h(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}$ is given by

$$\mathbf{g}(\Gamma(\gamma)) \triangleq f(\mathbf{x}) \circledast_\Gamma h(\mathbf{x}) = \int_{\mathbb{R}^3} \mathbf{f}(\mathbf{x}) h(\Gamma(\gamma) - \mathbf{x}) d\mathbf{x} \tag{2}$$

where $\mathbf{g}(t) : \mathbb{R} \rightarrow \mathbb{R}^m$ is a vector-valued function whose support is the fiber Γ. This operation is not more than a standard convolution of a field with a kernel but along the support defined by the fiber. This is an elegant way of finding the responses that a 3D field f will yield along the trajectory defined by the fiber. The convolution operator $\circledast_\Gamma$ satisfies the same properties than a standard convolution [10].

A. Geodesic kernel

The kernel function h defines how the values of $\mathbf{f}$ are locally combined to give a response. In this case, we can be more explicit in the requirements that h should fulfill in terms of *fiber convolution*. We cannot forget that a fiber Γ is the result of an integration of a diffusion tensor field $\mathbf{D}(\mathbf{x})$. Therefore, the fiber Γ can be seen as a geodesic path in the tensor space. A proper convolution along the fiber should take into account the metric that imposes the tensor field that generated the fiber. For example, the convolution in a given point $\mathbf{x}_0$ implies integrating the function f in a local neighborhood of $\mathbf{x}_0$ using the function h as a weight. It is not adequate to weight neighbor locations of $\mathbf{x}_0$ the same way that those that have a diffusion tensor totally different that the diffusion in $\mathbf{x}_0$.

Consequently, a more formal way of defining the fiber convolution is by means of a kernel function h that depends on a metric in the diffusion tensor space such that $h(\Gamma(\gamma) - \mathbf{x}) = h\left(d(\mathbf{D}(\Gamma), \mathbf{D}(\mathbf{x})\right)$, where $d(\cdot, \cdot)$ is a metric in the diffusion tensor space that defines the geodesic distance between $\mathbf{D}(\Gamma)$ and $\mathbf{D}(\mathbf{x})$.

The space of diffusion tensors, i.e. the space of symmetric positive-definite matrices $P(3)$, form a manifold known as a Riemannian symmetric space rather than a linear space [11]. In [11], Fletcher et al. derive a natural metric on the space of diffusion tensor based on an exponential map. This metric defines the distance between two tensor, $\mathbf{D}_1$ and $\mathbf{D}_2$, in a geodesic way as

$$d_G(\mathbf{D}_1, \mathbf{D}_2) = \mathrm{Tr}\left\{log(\Sigma)^2)\right\} \tag{3}$$

where Σ is diagonal matrix with the eigenvalues of the symmetric matrix $\mathbf{Y} = \mathbf{G}^{-1}\mathbf{D}_2(\mathbf{G}^{-1})^T$. $\mathbf{G}$ is a group action that takes the tensor $\mathbf{D}_1 = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ to the identity and defined as $\mathbf{G} = \mathbf{U}\sqrt{\mathbf{\Lambda}}$.

The metric defined by Fletcher is a natural candidate to define a kernel function based on that metric. Another metric

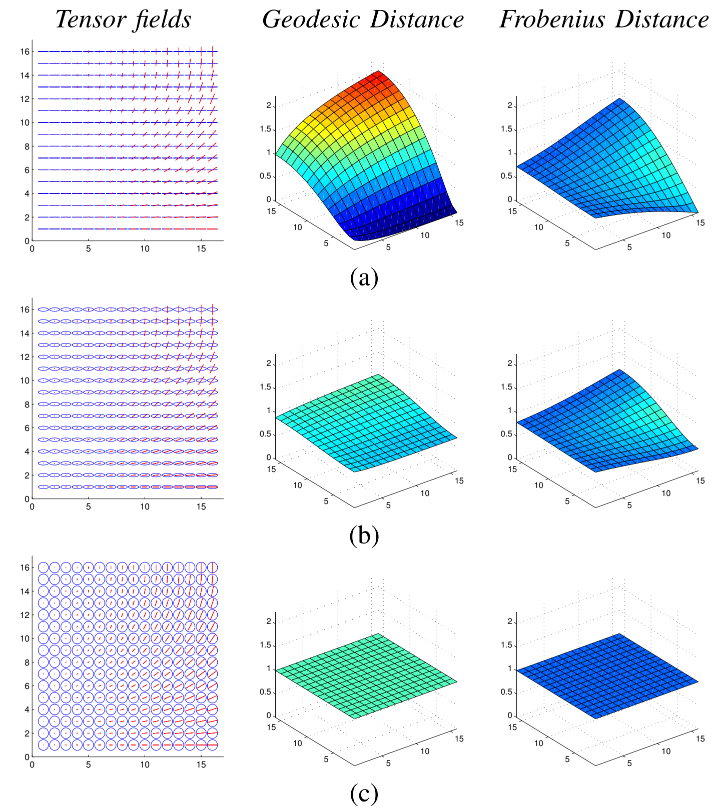


Fig. 1. Comparison between Geodesic metric and Frobenius metric. The distance between a pure anisotropic tensor with different orientations (red) and a tensor (blue) with variable anisotropy (from top to bottom) is studied.

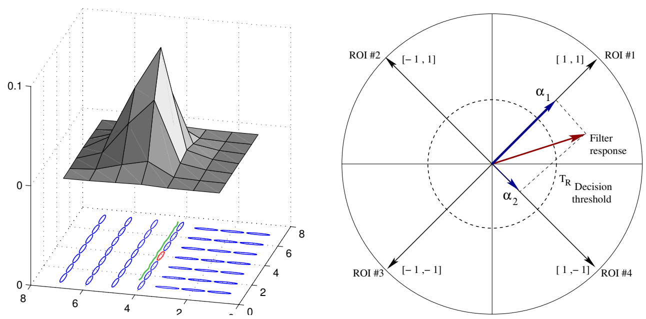
that has been used in the field of diffusion tensor is, for example, the Frobenius metric defined as $d_F(\mathbf{D}_1, \mathbf{D}_2) = \|\mathbf{D}_1 - \mathbf{D}_2\|_F$. A comparison between this two metric is shown in figure 1. These results reveal that the geodesic metric provides a better discrimination in the scape of tensors than the Frobenius one. It is interesting to realize how, in a geodesic sense, two anisotropic tensor with different orientations have a higher distance (figure 1a right corner) than one anisotropic tensor and a pure isotropic one (figure 1c right corner). This is very relevant in our application due to the fact that we want to stress the discrimination between two fibers with different orientations. The metrics use in figure 1 have been normalized with respect to the identity tensor such as

$$\hat{d}_G = \frac{d_G(\mathbf{D}_f, \mathbf{D})}{d_G(\mathbf{D}_f, \mathbf{I})}, \qquad \hat{d}_F = \frac{d_F(\mathbf{D}_f, \mathbf{D})}{d_F(\mathbf{D}_f, \mathbf{I})} \tag{4}$$

Among all the possible choices for h, the Gaussian kernel is one of the most used in signal and image processing due to its localization both in spatial domain and the Fourier domain. A Gaussian kernel based on the metric space given by eq. (3) is

$$h\left(d_G(\mathbf{D}(\Gamma), \mathbf{D}(\mathbf{x}))\right) = \frac{1}{(2\pi)^{\frac{3}{2}} \sigma d_G(\mathbf{D}(\Gamma), \mathbf{I})} e^{-\left(\frac{\hat{d}_G(\mathbf{D}(\Gamma), \mathbf{D}(\mathbf{x}))^2}{2} + \frac{\|\mathbf{x}\|^2}{2\sigma^2}\right)}, \tag{5}$$

where σ defines the spatial extent of the kernel and $\mathbf{D}(\Gamma)$ is the diffusion tensor along the fiber. An example of this kernel



(a) (b)

Fig. 2. (a) Convolution kernel associated to a fiber tract. The underlying diffusion tensor is shown underneath the kernel. A clear interface between diffusion direction in the x-axis and y-axis is presented.(b) Constellation of ROI values.

for the 2D case is shown in figure 2a. In this example, the underlying tensor structure corresponds to a clear interface between a tract along the x-axis and a tract along the y-axis. If we have a fiber along the x-axis, the fiber convolution with the kernel given by eq. (5) precludes using the part of the signal where the tensor field is oriented in the y-axis.

III. METHOD

The method proposed for user-guided fiber bundling uses the fiber convolution operator described in the previous section to find responses along the fiber tracts by convolving the fiber with a field derived from the ROI labelmap delineated by an expert. From this responses, a decisor takes care of selecting the fibers that pass through a given set of ROI labels.

A. Tractography

Fibers are obtained from the diffusion tensor data $\mathbf{D}(\mathbf{x})$ by means of a tractography algorithm. The tractography has been carried out using a hyperstreamline algorithm based on integrating the fiber trajectory along the PDD and implemented in the open source platform *3D Slicer*. As a byproduct of the integration, the interpolated tensor for each fiber location is also computed and stored. We will denote this tensor field along the fiber as D_f.

B. Fiber convolution

Once a set of fibers have been extracted, responses along the set of fibers are found by the fiber convolution method introduced in Section II. For each point along the fiber a function field $\mathbf{f}$ is convolved with a kernel function h. The function field $\mathbf{f}$ is obtained from the ROI labelmaps that have been previously delineated by an expert. We will elaborate on the definition of this function in the next section.

The kernel function is given by eq. (5). The kernel is parameterized by σ and the diffusion tensor for a given fiber location, $\mathbf{D}(\Gamma)$. This diffusion tensor is obtained as a linear combination of the tensor obtained from the tractography process, $\mathbf{D}_f$, and the outer product of the fiber tangent, such that

$$\mathbf{D}(\Gamma) = \frac{1}{2}\mathbf{D}_f + \frac{\|\mathbf{D}_f\|}{2}\frac{\partial\Gamma}{\partial\gamma}\left(\frac{\partial\Gamma}{\partial\gamma}\right)^T \tag{6}$$

C. ROI constellation

The ROIs labelmap, $R(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{N}^+$ is a scalar function with integer values, defined by an expert, to provide prior knowledge about the locations of a specific pathway. The goal is to obtain responses that are correlated with the ROI that the fiber goes through and therefore a decision can be made.

A straightforward approach is to obtain responses by doing a fiber convolution on R such that $f = R$. A fiber passes through a given ROI with value r, if the response along the filter after the convolution process yields a value equal to r. However, having a value equal to r is not always possible due to either

- the fiber crosses tangent to the ROI or
- discrepancies of the diffusion tensor of the fiber, $\mathbf{D}(\Gamma)$, with the neighbors tensors.

This approach is clearly inefficient if we want to simplify the decision stage. It can be easily seen that two different ROI values could yield the same responses making the decision stage harder.

An alternative approach is to increase the dimensionality of the ROI space to overcome the aforementioned problem. In this case, each ROI value of the labelmap R is associated to a m-D prototype vector. Therefore, the ROI labelmap is represented in this case by a vector value function $\mathbf{R}(\mathbf{x}) : \mathbb{R}^3 \rightarrow \mathbb{R}^m$. The set of all vector prototypes, $\mathbf{R}_n$, for the ROIs label defines a *constellation* in the m-D space. This concept is illustrated in figure 2b. The constellation can be defined by evenly spreading the vector in the space depending on the number of ROI values.

D. ROI intersection

The last stage of our method is deciding whether a fiber *passes through* a set of ROI values from those defined in the constellation. A fiber is said to pass through a ROI with prototype $\mathbf{R}_i$ if the norm of the responses along the fibers, projected on the prototype, are above a threshold (see figure 2b). The projected response for the prototype $\mathbf{R}_i$, α_i, is given by

$$\alpha_i(\Gamma) = \| < \mathbf{g}(\Gamma), \mathbf{R}_i > \|. \tag{7}$$

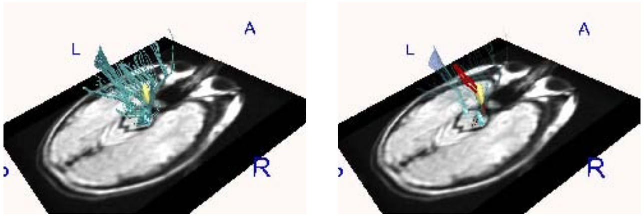
After a median filtering, $M(\cdot)$, to taper off outlier responses, a binary variable is defined as

$$\beta_i = \begin{cases} 1 & M(\alpha_i(\Gamma)) \geq T_R \\ 0 & M(\alpha_i(\Gamma)) < T_R \end{cases} \tag{8}$$

Finally a binary operator decides if the fiber passes through a set of N ROI values with indexes $\{s(j)\}_{j=1}^N$ if $\bigcap_{k=s(j)} \beta_k = 1$

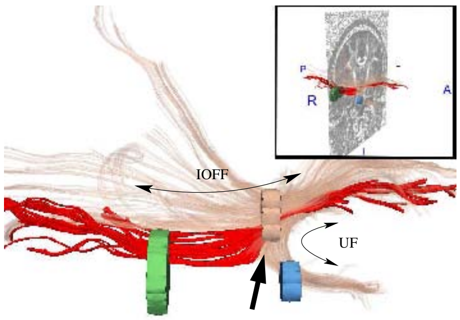
IV. RESULTS

The proposed method has been tested on two real diffusion weighted MRI. The first case corresponds to a fiber bundling of the brain stream based on two ROI and the result is shown in figure 3. For the second case, an expert has been asked to delineate three ROIs that would define 2 fiber tracts —uncinate fasciculus (UF) and inferior occipito-frontal

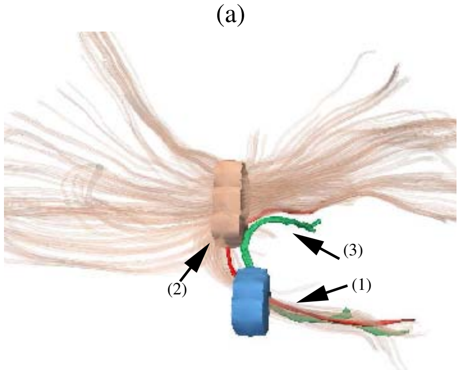


(a) (b)

Fig. 3. Bundling of the brain steam.(a) Tractography before applying the method and (b) resulting bundle (highlighted in red)



(a)



(b)

Fig. 4. (a) Bundling of the Inferior Occipital Frontal Fasciculus (IOFF) and (b) the Uncinate Fasciculus (UF) using the proposed method. Computed bundles are highlighted in red.

fasciculus (IOFF)— as shown in figure 4a. The delineation of the ROI has been undertaken based on prior knowledge of the connectivity anatomy but without taking special care in the definition of boundaries that cover the whole bundle. Fibers have been obtained by seeding points in the middle ROI (temporal stem) where both tracts merge. The extraction of the connection between the leftmost and the middle ROIs is shown in figure 4a. The extraction of the UF bundle is shown in figure 4b.

For both cases connectivity between the ROIs have been obtained using $\sigma = 1.5$ with a kernel sampled in a $7 \times 7 \times 7$ grid. A 2D ROI constellation has been used as shown in figure 2b. Finally, the decision threshold has been $T_R = 0.25$.

V. DISCUSSION

The first case, figure 3, is a clear example of the behavior of our method. The diffusion in the inferior-superior direction imposed a constrain in the fibers that are selected and only very straight fibers, a subset of those that cross the ROI, are selected by our method. We believe this is a desired behavior when extracting well-defined paths. An extreme situation that our method runs into is shown in figure 4b. In this case only a fiber is chosen (arrow 1), however there are other "U"-shape fibers that passes through both ROIs (arrow 2). Although this is not a desire behavior it shows how the kernel based on the tensor metric restricts the resulting bundle. The UF bundle is defined by more fibers (arrow 3) that those obtained in the initial tractography. This additional fibers have been manually seeded to show the extent of the UF bundle. When the arrow-2 fibers are convolved with the kernel in the middle ROI, the crossing with the IOFF decreases the response, due to an increasing of the distance between the neighbor tensors, below the threshold. This results can be improved by reducing the threshold T_R.

Our results show that by using the underlying diffusion tensor data, a conservative bundling is achieved. However, this behavior may be crucial if the method wants to be applied to find connectivity in a population based on ROI defines from atlases. We believe that in this situation our method is a much better candidate that standard ROI methods that do not use the tensor data. This may have a strong impact in population based neuropsychiatry research, where we expect to find disruptions in fiber tract organization.

REFERENCES

[1] C.-P. Lin, W.-Y. Tseng, H.-C. Cheng, and J.-H. Chen, "Validation of diffusion tensor magnetic resonance axonal fiber imaging with registered manganese-enhanced optic tracts," *NeuroImage*, vol. 14, pp. 1035–1047, 2001.

[2] S. Mori, B. J. Crain, V. P. Chacko, and P. C. van Zijl, "Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging," *Ann. Neurol.*, vol. 45, pp. 265–269, 1999.

[3] P. J. Basser and D. P. et al., "In vivo fiber tractography using DT-MRI data," *Magnetic Resonance in Medicine*, vol. 44, pp. 625–632, 2000.

[4] B. Stieltjes and W. E. K. et al., "Diffusion tensor imaging and axonal tracking in the human brainstem," *NeuroImage*, vol. 14, pp. 723–735, 2001.

[5] S. Mori and W. E. K. et al., "Imaging cortical association tracts in human brain," *Magnetic Resonance in Medicine*, vol. 47, pp. 215–223, 2002.

[6] M. Catani, R. J. Howard, S. Pajevic, and D. K. Jones, "Virtual in vivo interactive dissection of white matter fasciculi in the human brain," *NeuroImage*, vol. 17, pp. 77–94, 2002.

[7] P. Fillard and G. Gerig, "Analysis toll for diffusion tensor mri," in *Medical Image Computing and Computer-Assisted Intervention*, ser. Lecture Notes in Computer Science, R. E. Ellis and T. M. Peters, Eds., vol. 2879. Springer Verlag, 2003, pp. 967–968.

[8] H. Huang, J. Zhang, P. C. M. van Zijl, and S. Mori, "Analysis of noise effects on dti-based tractography using the brute-force and multi-roi approach," *Magnetic Resonance in Medicine*, vol. 52, pp. 559–565, 2004.

[9] A. Brun, H. Knutsson, H. J. Park, M. E. Shenton, and C.-F. Westin, "Clustering fiber tracts using normalized cuts," ser. Lecture Notes in Computer Science, vol. 3216, 2004, pp. 368–375.

[10] R. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed. McGraw-Hill, 1999.

[11] P. T. Fletcher and S. Joshi, "Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors," ser. Lecture Notes in Computer Science, vol. 3117. Springer Verlag, 2004, pp. 87–98.