Analysis of White Matter Fiber Tracts via Fiber Clustering and Parametrization

Sylvain Gouttard\textsuperscript{2,4}, Isabelle Corouge\textsuperscript{1,2}, Weili Lin\textsuperscript{3}, John Gilmore\textsuperscript{2}, and Guido Gerig\textsuperscript{1,2}

\textsuperscript{1} Departments of Computer Science, \textsuperscript{2}Psychiatry and \textsuperscript{3}Radiology, University of North Carolina, Chapel Hill, NC27599, USA
\textsuperscript{4}ESCPE Lyon, 69100 Villeurbanne, France
\texttt{gouttard@cpe.fr, corouge@unc.edu, gerig@cs.unc.edu}

\textbf{Abstract.} White matter fiber bundles of the human brain form a spatial pattern defined by the anatomical and functional architecture. Tractography applied to the tensor field in diffusion tensor imaging (DTI) results in sets of streamlines which can be associated with major fiber tracts. Comparison of fiber tract properties across subjects requires comparison at corresponding anatomical locations. Clinical analysis studying fiber tract disruption and connectivity requires analysis along tracts and within cross-sections, which is hard to accomplish by conventional region of interest and voxel-based analysis. We propose a new framework via tractography, fiber clustering and parametrization. The resulting fiber tract models are parametrized with arc length and carry local properties derived from diffusion tensors, e.g. fractional anisotropy FA and apparent diffusion coefficient ADC, local shape characteristics derived from the Frénet frame (curvature, torsion) and area and shape of cross-sections. Comparison of transversal tracts and projection tracts between adults and neonates are shown. This extended set of features might lead to an improved understanding of diffusion properties and its association to brain development and eventually changes due to pathology.

1 Introduction

Diffusion Tensor Imaging (DTI) measures local probability distribution of the self-motion of water molecules. Restricted motion of extra- and intra-cellular fluid within brain white matter fibers correlates local diffusion patterns, in its simplest form represented by local tensors, with the orientation and density of white matter fibers. The potential of DTI to present white matter integrity, disruption and pathology long before visible changes occur in structural imaging makes it the preferred modality to study white matter diseases. The two

\texttt{* This research is supported by the NIH NIBIB grant P01 EB002779, the NIH Conte Center MH064065, and the Stanley Medical Research Institute. We acknowledge Matthieu Jomier and Julien Jomier, both UNC-Chapel Hill, for their precious help, fruitful discussions and support with software development.
tensor measures most commonly used in clinical analysis are the “apparent diffusion coefficient” (ADC, trace of tensor) and the “fractional anisotropy” (FA, normalized elongation). Previous research has been mostly focused on the robust calculation of the tensor field, regularization [1], scientific visualization [2] and fiber tracking [3,4,5,6,7,8]. Atlas building and comparison of subject groups have been studied by non-linear registration of DTI images [9] and by combining tractography and spatial normalization [10].

Clinical analysis of DTI mostly followed the common concept of aligning DTI images via affine transformation to a template and statistical parametric mapping (SPM) for voxel-wise group difference tests [11,12,13]. Despite the use of manually defined regions of interest, discussion of results used the terms “connectivity” and “fiber disruption”, properties that might require a modeling of whole tracts. Kubicki et al. [12,13] report about differences of cross-sectional area of the uncinate fasciculus and the cingulate, measured in orthogonal sections of the MRI acquisition. These studies show that clinical analysis of DTI could benefit from improved tools for reliable extraction of fiber tracts of interest, for establishing homology across subjects, and for measuring and comparing geometric and diffusion properties along tracts. This paper significantly extends earlier work initiated by Fillard et al. [8] and by Corouge et al. [14]. We will not describe tractography but focus on the subsequent quantitative analysis.

2 Quantitative analysis of white matter fiber tracts

2.1 Tractography

The extraction of sets of streamlines from the diffusion tensor field is performed with a method originally developed by Mori et al. [15] and further improved by Fillard et al. [8]. Selection of source and target regions is similar to the concept outlined in [15]; we use backtracking from target to source and use large target regions (the full brain or the whole portion superior to the corpus callosum, e.g.) to generate seed points to be traced back to the source region. Tracking is regularized by two parameters controlling local coherence and smoothness of streamlines. Results for a set of major tracts are shown in Fig. 2a.

2.2 Clustering of curves to bundles

Tractography results in sets of streamlines connecting target to source regions. As our tracking is not guided by geometric constraints, this set can be noisy and can include multiple paths (see Fig. 3a,b and d). We developed a clustering scheme that uses various curve distance metrics to remove outliers and to combine curves to bundles. Closely related previous work has been proposed by Ding et al. [16]. Their bundling algorithm relies on the concept of subdivision into curve segments and the use of Euclidean distance to define similarity centered around a core curve. We extend this notion of cable-like bundles in order to represent ribbon

1 Fiber tracking tool download at http://www.cs.unc.edu/~fillard
cables and even bundles represented by sweeping a template curve to form a manifold [14], using the full set of pairwise distances. This model is motivated by the observation that white matter bundles form thin manifolds of quasi-parallel fibers that disperse towards the cortex (see Fig. 2a).

The fiber tracking process provides us with a set $F$ of 3D curves, $F_i$, each represented by a set of 3D points $p_k$, $F = \{F_i | F_i = \{p_k\}\}$. Given a pairwise distance $d$ and a fiber $F_i$, $d$ is computed between $F_i$ and $F_j$ for all $F_j$ in $F$, $j \neq i$. $F_i$ and $F_j$ are decided to be in the same class if $d(F_i, F_j) < t$ where $t \in \mathbb{R}$ is a threshold to be chosen. Clusters of very low cardinality (e.g. containing less than 10% of initial fibers) are considered as outliers and rejected. Thus, for each fiber $F_i$ within a class $C$, at least one fiber $F_j$, $j \neq i$ in $C$ is such that $d(F_i, F_j) < t$.

After calculating a matrix of pairwise distances, the algorithm propagates labels from neighbouring fiber to neighbouring fiber and benefits from a “transitivity property”. Moreover, only one parameter, the threshold $t$, has to be selected. A large value of $t$ results in a small number of classes, whereas a smaller value will increase the number of classes. The optimal parameter $t$ depends on the data set under examination and on the choice of the distance metric. We compute the histogram of the number of classes as a function of $t$ to study the sensitivity of each metric in regard to this parameter and to guide users to come up with a meaningful choice. For example, users can select the number of sought clusters instead of the parameter itself, which is used in Figure 3d to separate left and right cortico-spinal tracts.

Three pairwise distances between curves $F_i$ and $F_j$ have been implemented:

1. Closest point distance, $d_c$:
   \[
   d_c(F_i, F_j) = \min_{p_k \in F_i, p_l \in F_j} \| p_k - p_l \|, \tag{1}
   \]
   \| . \| being the Euclidean norm;

2. Mean $d_M$ of closest distances, defined as:
   \[
   d_M(F_i, F_j) = \text{mean}(d_m(F_i, F_j), d_m(F_j, F_i))
   \]
   with $d_m(F_i, F_j) = \min_{p_k \in F_i} \min_{p_l \in F_j} \| p_k - p_l \|$, \tag{2}

3. Hausdorff distance, $d_H$:
   \[
   d_H(F_i, F_j) = \max(d_h(F_i, F_j), d_h(F_j, F_i))
   \]
   with $d_h(F_i, F_j) = \max_{p_k \in F_i, p_l \in F_j} \min \| p_k - p_l \|$. \tag{3}

Additionally, we use shape-based distances by extracting geometric characteristics from fibers such as length, center of mass and second order moments. The principle of the clustering algorithm remains the same when using first or second order moments. In the former case, $d$ is the Euclidean distance between centers of mass, called $d_G$, whereas in the latter case it represents orientation similarity of the first principal directions.

The overall closest distance $d_c$ can not be expected to have a good discrimination power between fibers since a single closest point pair does not encode shape similarity. On the contrary, $d_M$ provides a global similarity measure since it integrates closest distances along the whole curve. The Hausdorff distance is a worst-case distance, it is a useful metric to reject outliers and prevents the
algorithm from clustering curves with high dissimilarity. Centers of mass are an appropriate feature to measure coarse similarity of pose since they are a first order complete representation of a fiber, whereas the second order moment metric has difficulties to discriminate dense fiber sets because of its noise sensitivity.

The interactively guided clustering allows a multi-criteria based classification. For example, outliers can be first rejected based on length and Hausdorff distance whereas left/right bundles might be separated by comparing the center of mass. We will develop guidelines for major tracts and thus standardize the procedure aiming towards an automated clustering scheme.

2.3 Attributing bundles with diffusion properties

It is now interesting to study local diffusion not in small regions of interest but as a function of the geometry of the whole tract. Our tractography tool [8] simplifies this task since fibers are stored as standardized ITK polylines attributed with the full tensor and derived properties. Visualization of the splenium (Fig. 2b) clearly demonstrates that the fractional anisotropy varies significantly as a function of location along the tract but also within cross-sections. The histogram representing the mid-sagittal cross-section (Fig. 2c) clearly shows that values range from 0.1 up to 0.9, representing the whole range from nearly isotropic up to highly anisotropic diffusion. We assume that this broad range is a function of the coarse sampling of the underlying macroscopic fiber structures (here \(2 \times 2 \times 2 \text{mm}^3\)), partial voluming, but also natural variation of fiber density and myelination sheath. However, it demonstrates that region of interest analysis is not sufficient and might be very sensitive to the exact definition of cross-sections.

DTI properties of large bundles (corpus callosum and cortico-spinal) are illustrated in the first column of Fig. 4. Whereas apparent diffusion (ADC) is mostly constant across the whole bundle, an interesting pattern is observed in fractional anisotropy (FA). The FA values change significantly and form regular patterns. The dark blue on top is easily explained by fiber dispersion towards the cortex and thus isotropic diffusion. The “stripe” patterns are induced by neighboring tracts either running in parallel or perpendicular. These views again indicate that regions of interest analysis might not capture this natural variation and even might be very sensitive to selecting locations along tracts.

2.4 Parametrization of fiber bundles

Summarizing diffusion properties within a bundle requires parametrization. First, we define a common origin for the set of fibers in each cluster. The choice of this origin might be based on geometric criteria, e.g. a cross-section with minimal area, or based on anatomical information, like the symmetry plane of the interhemispheric fissure. The polyline of each curve is parametrized by a cubic B-spline curve. The set of splines is re-sampled and diffusion measurements at each point are obtained by interpolation. The two attributes ADC and FA can now be integrated across cross-sections and expressed as a function of arc-length.
Fig. 4 displays average and standard deviations of ADC and along the fiber directions, i.e. from inferior to superior direction. Results are illustrated for two adults and one neonate case. The comparison between the adult and neonate cases clearly shows higher ADC (more diffuse) and lower FA (less structured) of the newborn which confirms our earlier single-slice region of interest findings [17]. The curves representing FA clearly reflect the “stripes” shown in the color displays. Note that values close to the top of the brain (right in the graphs) are less stable due to partial voluming of tissue and sulcal fluid.

2.5 Attributing bundles with local shape properties

A comparison of fiber tracts across subjects requires measurements at corresponding locations. Brain mapping by non-linear registration of volumetric DTI is one option [9], posing the difficult problem of interpolation and warping of tensor data. Here, we would like to explore a different option via geometric characterization of fiber bundles. Given a differentiable parametrization $r(t)$ of a curve $C$, the Frénet frame $(\vec{T}, \vec{N}, \vec{B})$ in each point $p$ of the curve is defined by:

\[
\begin{align*}
\vec{T} &= \frac{r'(t)}{||r'(t)||} & \text{where} & \quad r'(t) = \frac{d}{dt}r(t) \\
\vec{N} &= \frac{\vec{T}'}{||\vec{T}'||} & \text{where} & \quad \vec{T}' = \frac{d\vec{T}}{dt} \\
\vec{B} &= \vec{T} \wedge \vec{N} & \text{where} & \quad \wedge \text{ denotes the vector product.}
\end{align*}
\]

The vector $\vec{T}$ is the unit tangent vector, $\vec{N}$ the unit normal vector and $\vec{B}$ supplements the frame so that it is orthonormal. At each point $p$, this frame allows the calculation of local features such as curvature $\kappa$ and torsion $\tau$. These measurements are given by the Serret-Frénet formulæ:

\[
\frac{d\vec{T}}{ds} = \kappa \vec{N} \quad \text{and} \quad \frac{d\vec{B}}{ds} = -\tau \vec{N}
\]  

where $s$ is the curvilinear abscissa of $C$.

We use the parametrization by cubic B-splines as discussed earlier to establish point correspondence between sets of fibers. Points with the same curvilinear abscissae across the fiber set, i.e. with the same arc-length, are defined as homologous. Curvature and torsion are computed along each curve as described above. Pointwise mean and standard deviation on these features characterize local shape properties of the whole bundle. Results for curvature of the corpus callosum bundles of two adults and one neonate (see Fig. 4) are presented in Fig. 1. The mesh representations and the mean and standard deviation of curvature reflect the high similarity of the U-shaped curves close to the origin (midsagittal plane) and the increased variability towards the cortex (left and right in graphs) as expected due to fiber dispersion. Both adult cases represent a similar pattern of 3 curvature peaks (see arrows), suggesting that these locations might be used as features for curve matching. These features are even reflected in the neonate curvature pattern, potentially giving us a possibility to compare subjects across the whole age range, e.g. to study brain development.
3 Conclusion

This paper presents work in progress towards a novel concept of DTI analysis which compares individual fiber tracts rather than volumetric images. This approach is motivated by clinical needs to study diffusion properties along tracts and to compare properties across subjects. Hypotheses of changes of fiber integrity and fiber disruption measured via local diffusion inspired our representation of bundles attributed with diffusion tensor properties. The results demonstrate the significant variability of tensor properties along bundles but also within cross-sections. The statistical analysis presented here uses mean and standard deviation of ADC and FA, but ongoing research will explore non-parametric statistics via quantiles to account for the asymmetric distribution of measurements within cross-sections.

The analysis of local shape properties (curvature and torsion) is currently used to develop a fiber-matching scheme for improved intra-subject matching of curves to bundles but also for comparison of bundles across subjects. The few cases illustrated in this paper are part of two much larger studies which a) explore normal variability within a set of 15 healthy adults and b) study early brain development of newborns at risk (age 2 weeks) with follow-up after 1 and/or 2 years. Models on healthy controls will help us to measure and quantify geometric and diffusion changes of fiber tracts due to pathology.

References

eds.: Medical Image Computing and Computer Assisted Intervention [MICCAI02].
615–622
mapped volume rendering. In: Medical Image Computing and Computer Assisted
294–302
and visualization for diffusion tensor MRI. Medical Image Analysis 6 (2002) 93–108
magnetic resonance images. In Insana, M., Leahy, R., eds.: Proc. Information
Processing in Medical Imaging. Volume 2082 of Lecture Notes in Computer Science.,
human brain using DT-MRI data. IEICE Trans. on Information and Systems
E85-D (2002) 15–21
7. Zhukov, L., Barr, A.: Oriented tensor reconstruction: tracing neural pathways from
Notes in Computer Science. (2003) 967–968
Vision and Image Understanding 77 (2000) 233–250
fiber bundles using diffusion tensor MRI. In: IEEE Proceedings Int. Symposium
15. Mori, S., Kaufmann, W., Davatzikos, C., Stietjes, B., Amodei, L., Frederiksen,
716–721
white matter diffusion in healthy neonates and adults using a 3T head-only MR
Fig. 2. DTI fiber tracts: a) Sets of tracts obtained by tractography, b) coloring FA properties along and across bundle (range $[0 \cdots 1]$ represented from blue to red) and c) histogram of FA properties within mid-sagittal cross-section.

Fig. 3. Clustering of sets of streamlines to fiber bundles. Callosal fibers obtained by tractography (a and b), filtering and clustering using distance metric $D_G$ (c), cortico-spinal tract before (d) and after clustering (e).

Fig. 4. Analysis of diffusion properties along major fiber tracts. Top and middle row represent ADC and FA values for cortico-spinal tracts, bottom row FA for the callosal tract. Columns (a), (b) and (c) illustrate color-coded 3D tracts for adult 1 and graphs of bundle properties for adult 1 and adult 2. Columns (d) and (e) show the same properties for a neonate case. Graphs left to right represent properties along inferior-superior (bottom-top) tract orientation.