

Fiber Tract-Oriented Statistics for Quantitative Diffusion Tensor MRI Analysis

Isabelle Corouge^{*},

*Departments of Computer Science and Psychiatry
University of North Carolina, Chapel Hill, USA*

P. Thomas Fletcher,

*Scientific Computing and Imaging Institute
University of Utah, USA*

Sarang Joshi,

*Department of Radiation Oncology
University of North Carolina, Chapel Hill, USA*

Sylvain Gouttard

*Department of Psychiatry
University of North Carolina, Chapel Hill, USA*

Guido Gerig

*Departments of Computer Science and Psychiatry
University of North Carolina, Chapel Hill, USA*

Abstract

Quantitative diffusion tensor imaging (DTI) has become the major imaging modality to study properties of white matter and the geometry of fiber tracts of the human brain. Clinical studies mostly focus on regional statistics of fractional anisotropy (FA) and mean diffusivity derived from tensors. Existing analysis techniques do not sufficiently take into account that the measurements are tensors, and thus require proper interpolation and statistics of tensors, and that regions of interest are fiber tracts with complex spatial geometry. We propose a new framework for quantitative tract-oriented DTI analysis that systematically includes tensor interpolation and averaging, using nonlinear Riemannian symmetric space. A new measure of tensor anisotropy, called geodesic anisotropy (GA) is applied and compared with FA. As a result, tracts of interest are represented by the geometry of the medial spine attributed with tensor statistics (average and variance) calculated within cross-sections. Feasibility of our approach is demonstrated on various fiber tracts of a single data set. A validation study, based on six repeated scans of the same subject, assesses the reproducibility of this new DTI data analysis framework.

Key words: Diffusion tensor interpolation, diffusion tensor statistics, DTI analysis, fiber tract modeling.

* Corresponding author.

Email addresses:

icorouge@gmail.com (Isabelle Corouge),
fletcher@sci.utah.edu (P. Thomas Fletcher),
sjoshi@unc.edu (Sarang Joshi),
sylvain.gouttard@yahoo.fr (Sylvain Gouttard),
gerig@cs.unc.edu (Guido Gerig).

1 Introduction

Diffusion Tensor Imaging (DTI) of brain structures measures diffusion properties by the local probability of self-motion of water molecules. A tensor field characterizes amount and locally preferred directions of local diffusivity. While diffusion can be considered isotropic in fluid it appears highly anisotropic along neural fiber tracts due to inhibition of free diffusion of intra- and extra-cellular fluid [Beaulieu, 2002]. DTI has become the preferred modality to explore white matter properties associated with brain connectivity *in vivo*.

The literature proposes a variety of DTI processing techniques, ranging from tensor field computation to quantitative analysis, and including visualization, regularization, registration, tractography and population statistics. Few of these methods make use of the full tensor information though most would benefit from an appropriate mathematical framework for tensor operations and tensor statistics calculation. For instance, tensor interpolation is of high interest for regularization, which is a crucial component in DTI in view of the high sensitivity to noise and to partial voluming effects. However most approaches proposed so far do not directly regularize the tensor measurements. The diffusion weighted images are smoothed before tensor calculation in [Parker et al., 2000], only the vector field defined by the principal direction diffusion (PDD) is filtered in [Poupon et al., 2001]. In [Coulon et al., 2004], the PDD field and the three eigenvalue maps are restored in a decoupled manner. Registration and spatial normalization [Alexander et al., 2001], [Jones et al., 2002] are another typical examples where tensor interpolation is required. Moreover, tensor statistics calculation also becomes necessary for statistical DTI analysis in population studies. So far, analysis schemes have mostly focused on measuring properties in regions of interest and to a lesser extent along fiber bundles [Ding et al., 2003], [Fillard et al., 2003], [Corouge et al., 2004], [Jones et al., 2006] and they have not made use of the full tensor information. Consequently, clinical studies have mostly been limited to statistics of fractional anisotropy or mean diffusivity maps on a voxel-by-voxel basis [Lim and Helpert, 2002].

In this paper, we design a new framework for quantitative DTI data analysis. First, we use the full tensor information and include tensor interpolation and tensor averaging. We choose the affine-invariant Riemannian metric to define tensor operations and tensor statistics out of the various tensor metrics proposed in the literature. Second, as opposed to voxel-based analysis, we propose an object-oriented approach in which the fiber tracts act as coordinate systems for quantitative DTI analysis. Such a structural approach is superior for data representation of DTI if it is to be used for analysis of functional properties of anatomical structures, in this case white matter fiber tracts. Our concept provides a complete representation of each individual bundle, describing both geometry and diffusion properties. The representation includes model of the geometry of individual bundles and statistics of diffusion tensors to be associated with the geometric model. The tract geometry is modeled by estimation of a prototype shape and characterization of shape variability. Tensor information is integrated across cross-sections and

represented along bundles. Each location along the template curve is attributed with a template tensor (an average tensor), from which we derive diffusion properties.

The next sections motivates our choice for the affine invariant Riemannian metric and summarizes the key principles of the Riemannian framework, which has been fully described elsewhere, for tensor operations and tensors statistics. It also presents a new measure of tensor anisotropy consistent with the chosen tensor metric. Section 3 describes the geometric modeling of fiber tracts and explained how such a theoretical framework is used to attribute the mean geometric model with diffusion tensor statistics. Section 4 illustrates our methodology on a single data set before presenting a validation study.

2 Theoretical Framework

2.1 Motivation

We denote the space of all diffusion tensors, i.e., the space of all 3×3 symmetric, positive-definite matrices, as $PD(3)$. Averaging and interpolation of diffusion tensors can be formulated as a least-squares minimization problem in this space. This definition depends on the choice of metric, or distance, on the space $PD(3)$. Various metrics have been proposed to measure the distance between two tensors. Some of them are based on scalar features extracted from the diffusion tensor, like in [Guimond et al., 2002] where DTI data registration is driven by eigenvalues maps. Such approaches unfortunately ignore the directional information contained in the tensor. Out of several similarity measures based on the full diffusion tensor, the tensor Euclidean distance, or Frobenius norm, is empirically shown to perform the best for matching diffusion tensor images [Alexander et al., 1999]. It compares with the tensor metric proposed by Zhang [Zhang et al., 2004], [Zhang et al., 2005], which is derived from diffusion profiles and expressed as a weighted sum of the Euclidean distance and the trace distance (absolute value of the difference of the tensors' traces). The Frobenius norm is used in [Jones et al., 2002] to compute statistics of a distribution of tensors. The average tensor is defined as the Fréchet mean of a set of tensors and coincides with the linear averaging of the tensors coefficients under the chosen metric. Linear averaging is also applied in [Westin et al., 2002] for Gaussian filtering of a tensor field.

However, tensors with the standard addition and scalar multiplication, i.e., as defined on square matrices, do not form a vector space. For example, the negation of a positive-definite matrix is not positive-definite. Accordingly, standard linear operations and statistics are not appropriate: they do not preserve the natural properties of the tensors. The determinant, resp. the positive-definiteness, of the diffusion tensors is not preserved by linear first, resp. second, order statistics. In particular, linear averages suffer from a "swelling" effect where diffusion tensors with the same determinant will have an average with a larger determinant. This can be thought of as introducing diffusion when averaging, which is not physically acceptable. Linear interpolation of diffusion tensors

suffers from this same effect.

Lately, several groups have overcome these shortcomings by describing the space of diffusion tensors, $PD(3)$, as a curved manifold, or more specifically, a Riemannian symmetric space and by deriving on this space a more natural metric for tensor operations: the affine-invariant Riemannian metric [Fletcher and Joshi, 2004], [Batchelor et al., 2005], [Pennec et al., 2005]. Another family of Riemannian metrics, the Log-Euclidean metrics, has very recently been proposed [Arsigny et al., 2005]. Being Euclidean in the domain of tensor logarithms, these metrics simplify calculations on tensors and lower the computational cost. Averages of a set of tensors with Log-Euclidean and affine-invariant metrics are theoretically and practically very similar. However, the affine-invariant Riemannian metric is the only one with full affine invariance.

In this paper, we adopt the affine-invariant Riemannian metric as its mathematical properties makes it an appropriate choice for computation of tensor operations and tensor statistics. The symmetric space metric does not suffer from the swelling effect of the linear metric, that is, diffusion tensors with the same determinant will have an average with the same determinant. Synthetic examples of weighted averages of tensors are provided in Fig. 1. We use tensor averaging and interpolation methods, first presented in [Fletcher and Joshi, 2004], [Fletcher, 2004], that are based on the notion of geodesic distance within this space.

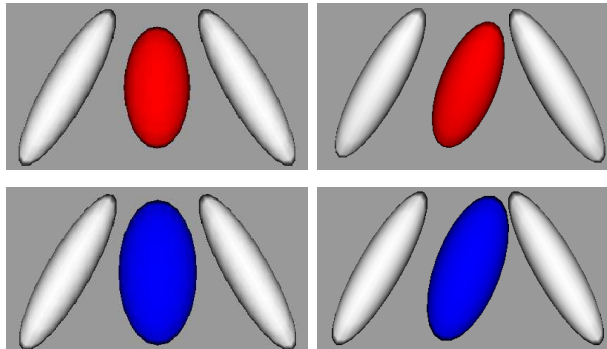


Figure 1. Synthetic examples of weighted averages of tensors. The white ellipsoids average to the red ellipsoid with the geodesic method (top) and to the blue ellipsoid with the linear method (bottom). Left: weights = $\{0.5, 0.5\}$. Right: weights = $\{0.75, 0.25\}$. It can be observed that the linear method does not preserve the determinant.

2.2 Statistics and Interpolation of Diffusion Tensors

In Appendix B, we further develop the notion of symmetric space and the computation of geodesic distance on $PD(3)$. We define statistics, average and variance, of a set of diffusion tensors based on the geodesic distance on $PD(3)$. Interpolation of diffusion tensors follows as an extension to weighted averaging.

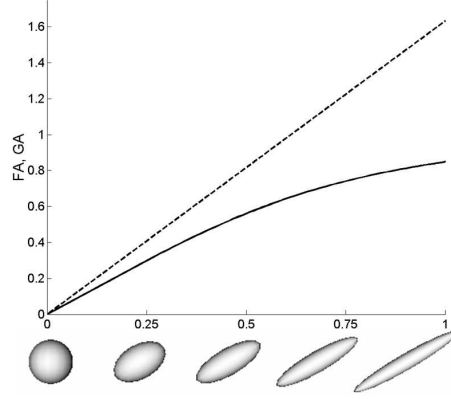


Figure 2. Comparison of FA (solid line) and GA (dashed line) values for the tensors with eigenvalues $\exp(t)$, $\exp(-t)$, $\exp(-t)$.

2.3 Geodesic Anisotropy

An important widely used function for characterizing the anisotropy of a tensor is fractional anisotropy (FA). Although FA characterizes anisotropy, it is not based on the intrinsic geometry of the space of diffusion tensors. We now describe a new anisotropy measure for diffusion tensors, first defined in [Fletcher, 2004], called *geodesic anisotropy* (GA) that is based on the geodesic distance in the symmetric space $PD(3)$. Geodesic anisotropy is intuitively a measure of how far away a diffusion tensor is from being isotropic. Therefore, a natural measurement of the anisotropy of a diffusion tensor $p \in PD(3)$ is the geodesic distance between p and the closest isotropic diffusion tensor. The geodesic anisotropy of p can be written as

$$GA(p) = \left(\sum_{i=1}^3 (\log(\lambda_i) - \overline{\log \lambda})^2 \right)^{\frac{1}{2}}. \quad (1)$$

where λ_i denotes the eigenvalues of p and $\overline{\log \lambda}$ denotes the average of the logs of the eigenvalues (see Appendix B for mathematical details).

Geodesic anisotropy, like FA, is invariant to uniform scaling of a diffusion tensor. Unlike FA, which is in the range $[0, 1]$, the GA is unbounded and can take values in $[0, \infty)$. FA and GA represent a different mapping of the eigenvalues. Equation (1) shows that the geodesic anisotropy is equivalent to the standard deviation of the log of the eigenvalues (times a scale factor). This is similar to how the fractional anisotropy is defined via the standard deviation of the eigenvalues, which are treated as linear entities. The GA is consistent with the thinking of $PD(3)$ as a symmetric space, where the eigenvalues are treated as multiplicative entities rather than linear ones. A comparison of FA and GA values of the one-parameter family of tensors that have eigenvalues $\lambda_1 = \exp(t)$, $\lambda_2 = \lambda_3 = \exp(-t)$ is shown in Fig. 2.

3 Analysis Methodology

Interpolation and averaging of tensors is applied for quantitative fiber tract-oriented analysis of DTI. The geometry of an individual fiber tract is modeled, basically with what is commonly called a point distribution model (PDM) [Cootes et al., 1995]. Diffusion tensor statistics are computed across fiber tract sections and are associated with the mean geometric model, resulting in a compact description of diffusion properties along the fiber tract. An overview of our framework is illustrated in Fig. 3.

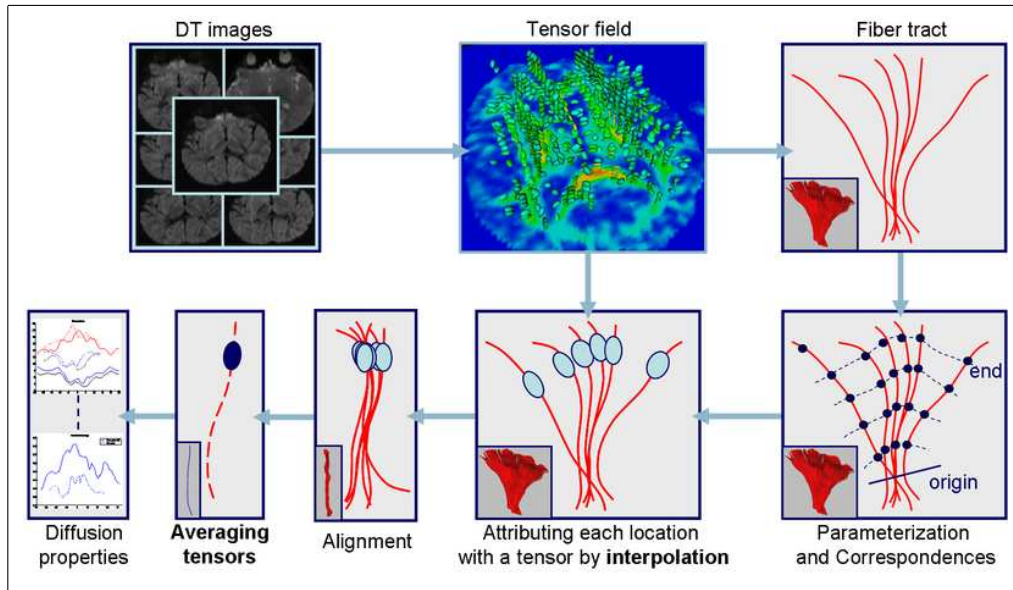


Figure 3. Overview of the DTI analysis framework.

3.1 Preprocessing: Tensor Field Computation and Fiber Extraction

The tensor field, defined by a 3×3 symmetric definite-positive matrix in each voxel, is computed from DTI data by solving the Stejskal-Tanner’s diffusion equation system as described in [Westin et al., 2002]. A tractography algorithm [Fillard and Gerig, 2003] extracts streamlines following the principal diffusion tensor directions between source and target regions of interest. The tracking is performed under local continuity constraints [Xu et al., 2002], backwards, and provides sub-voxel precision. Our latest version of the tractography tool includes tensor interpolation as described in Sec. 2. Except at branching or crossing points, the extracted 3D curves are assumed to represent the most likely pathways through the tensor field. Note that the term “fibers” is used for streamlines in the vector field which do not represent real anatomical fibers. Since the robustness of fiber tracking remains limited at junctions and in noisy low-contrast regions, the extracted fiber set contains outlier curves. We developed an iterative algorithm to reject outliers and to cluster curves to fiber bundles based on pairwise distance metrics measuring position and shape similarity of pairs of fibers [Corouge et al., 2004].

3.2 Geometric Modeling

An individual fiber tract, described by a set of streamlines, acts as a training set from which we estimate a template shape, the mean shape, and statistical deviations by learning its inherent shape variability. Representation and matching of the training set relies on the definition of common end points and on a data reparametrization from which we establish correspondences. Pose parameters are then estimated by a Procrustes analysis [Goodall, 1991]. A principal component analysis is subsequently applied to characterize statistical shape variation.

3.2.1 Parameterization and Correspondences

First, for each fiber tract under analysis, we specify common start and end points, which can be reliably identified across subjects. The start and end points are defined as the intersection of the fiber tract with a plane. Most often, the orientation of the plane is chosen orthogonal to the fiber tract direction. Its position is determined either by anatomical information or by a geometric criterion, like the location where fibers start dispersing towards various cortical regions. Note that multiple cutting planes with different orientations can be used for a given fiber tract. Second, fibers represented as polylines are reparametrized by cubic B-spline curves. This choice is well adapted to model a wide range of curves. It provides each fiber with a continuous representation and ensures a regular sampling along each fiber as well as a consistent sampling for all fibers in the tract. We slightly oversample the observations in order to prevent any loss of shape information but also to avoid any undesirable increase of dimensionality. Finally, points with the same arc length along the fiber tract are defined as homologous. Given this correspondence, the alignment of all curves in the training set is achieved by Procrustes analysis.

3.2.2 Pose Parameter Estimation: Procrustes Analysis

Let $\mathcal{F} = \{\mathbf{F}_n, 1 \leq n \leq N, \mathbf{F}_n \in \mathcal{M}_{k,m}\}$ be a set of N fibers, each defined by a set of k corresponding points in $m = 3$ dimensions, and represented by a $k \times m$ matrix. For $N = 2$, an Ordinary Procrustes Analysis (OPA) gives the optimal similarity transformation parameters in a least squares sense by minimizing

$$d_{\text{OPA}}^2(\mathbf{F}_1, \mathbf{F}_2) = \|\mathbf{F}_2 - (s\mathbf{F}_1\mathbf{R} + \mathbf{1}_k\mathbf{t}^t)\|^2, \quad (2)$$

where $s \in \mathbb{R}^{+*}$ is a scaling parameter, $\mathbf{R} \in SO(m)$ is a rotation, \mathbf{t} is a $m \times 1$ translation vector and $\mathbf{1}_k$ is a $k \times 1$ vector of ones. Minimization of (2) over the similarity group has an algebraic solution when shapes are centered, i.e., $\mathbf{1}_k^t \mathbf{F} = 0$, and normalized to unit size, i.e., $\|\mathbf{F}\| = \sqrt{\text{trace}(\mathbf{F}^t \mathbf{F})} = 1$: $\mathbf{t} = 0$, $\mathbf{R} = \mathbf{U}\mathbf{V}^t$, $s = \text{trace}(\mathbf{D})$ where $\mathbf{V}\mathbf{D}\mathbf{U}^t = \mathbf{F}_2^t \mathbf{F}_1$ is the singular value decomposition of $\mathbf{F}_2^t \mathbf{F}_1$. In the actual case where $N > 2$, a Generalized Procrustes Analysis (GPA) estimates the similarity transformation

parameters which minimize the sum of squared norms of pairwise differences

$$d_{\text{GPA}}^2(\mathbf{F}_1, \dots, \mathbf{F}_N) = \frac{1}{N} \sum_{n=1}^N \sum_{p=n+1}^N \|(s_n \mathbf{F}_n \mathbf{R}_n + \mathbf{1}_k \mathbf{t}_n) - (s_p \mathbf{F}_p \mathbf{R}_p + \mathbf{1}_k \mathbf{t}_p)\|^2. \quad (3)$$

The optimization is performed iteratively:

- (1) Translation. Fibers are centered with respect to their center of mass, \mathbf{g}_n : $\mathbf{F}_n^c = \mathbf{F}_n - \mathbf{g}_n$.
- (2) Scaling. Centered data is normalized to unit size: $\mathbf{F}_n^{\text{cs}} = \mathbf{F}_n^c / \|\mathbf{F}_n^c\|$.
- (3) Rotation. Let $\mathbf{F}_n^{\text{old}} = \mathbf{F}_n^{\text{cs}}$. The N shapes are rotated in turn. For each n , $1 \leq n \leq N$:
 - (a) $\bar{\mathbf{F}}_n = \frac{1}{N-1} \sum_{p \neq n} \mathbf{F}_p^{\text{old}}$,
 - (b) $s_n = 1$, $\mathbf{t}_n = 0$, $\mathbf{R}_n = \arg \min_{\mathbf{R}} d_{\text{OPA}}^2(\mathbf{F}_n^{\text{old}}, \bar{\mathbf{F}}_n)$,
 - (c) $\mathbf{F}_n^{\text{new}} = \mathbf{F}_n^{\text{old}} \mathbf{R}_n$ and $\mathbf{F}_n^{\text{old}} = \mathbf{F}_n^{\text{new}}$.

Step 3 is iterated until the Generalized Procrustes distance $d_{\text{GPA}}^2(\mathbf{F}_1^{\text{old}}, \dots, \mathbf{F}_N^{\text{old}})$ can not be reduced further. The alignment of the training set is achieved by applying the estimated rotations to the centered but non unit-scaled initial shapes \mathbf{F}_n^c , resulting in the set of aligned fibers $\mathcal{F}^A = \{\mathbf{F}_n^A, 1 \leq n \leq N\}$. Indeed, the scaling is needed to optimally estimate the rotation but a size normalization is not desirable since the training fibers belong to the same individual fiber tract.

In summary, let \mathcal{T} be the set of Procrustes estimated transformations, $\mathcal{T} = \{(-\mathbf{g}_n, \mathbf{\Gamma}_n), 1 \leq n \leq N\}$ with \mathbf{g}_n the translation vector defined by the center of mass of the n^{th} fiber and $\mathbf{\Gamma}_n$ the resulting rotation for fiber n : $\mathbf{\Gamma}_n = \Pi_i \mathbf{R}_n^{(i)}$, with $\mathbf{R}_n^{(i)}$ the rotation computed in the i^{th} step 3 iteration of the GPA. Then,

$$\mathcal{F}^A = \{\mathbf{F}_n^A = (\mathbf{F}_n - \mathbf{g}_n) \mathbf{\Gamma}_n, 1 \leq n \leq N\}. \quad (4)$$

3.2.3 Estimation of the Mean Shape

Given the set of aligned shapes, the mean shape $\bar{\mathbf{F}}$ is estimated by averaging the spatial coordinates at each corresponding location over the tract:

$$\bar{\mathbf{F}} = \frac{1}{N} \sum_{n=1}^N \mathbf{F}_n^A. \quad (5)$$

Additionally, statistical shape deviations from this template shape along the tract can be characterized by extracting the principal modes of deformation relative to the mean shape via a principal component analysis.

3.3 *Attributing the Geometric Model with Diffusion Tensor Statistics*

The estimated mean shape models the geometry of the fiber tract. A complete representation of the tract, describing both geometry and diffusion properties, is obtained by attributing each location along the mean curve with statistics of diffusion tensors calculated over cross-sections.

3.3.1 *Computing the mean tensor over fiber tract cross-sections at each location of the tract*

First, each sample point x from the set of reparameterized fibers is assigned a tensor p . Since the tensor field is defined on the discrete voxel grid while x lies on a continuous curve, a geodesic interpolation (see Sec. 2.2 and Appendix B) is required to compute the tensor p at the location x . The tensor p is given by the weighted average of the eight voxel tensor values in the nearest $2 \times 2 \times 2$ neighborhood of x , the weights being defined by trilinear interpolation (see Eq. (B.3)). Let \mathcal{P} be the set of obtained tensors, $\mathcal{P} = \{p_{n,i}\}$ with n indexing the set of reparameterized fibers \mathcal{F} and i the location along each reparameterized fiber. Then, the tensor set \mathcal{P} is aligned by rotation. Using $\Gamma_n \in SO(3)$, the rotation estimated by Procrustes analysis for the reparameterized fiber \mathbf{F}_n , each tensor $p_{n,i}$ lying on \mathbf{F}_n is rotated to the tensor $p'_{n,i}$ by the group action $\Gamma_n^t: p'_{n,i} = \Gamma_n^t p_{n,i} \Gamma_n, \forall i$. Last, at each corresponding location i along the tract, the mean tensor μ_i is computed from the set of aligned tensors, $\{p'_{n,i}\}$, as defined by Eq. (B.1). In addition, cross-sectional tensor diffusion variability can be assessed at each location of the average curve by computing the geodesic standard deviation (see Eq. (B.2)).

Just note that, for visualization purposes, each average tensor is translated to its corresponding average location on the average curve. Since diffusion tensors are invariant to translation, this does not affect any diffusion property.

3.3.2 *Deriving Diffusion Properties*

At each location along the template curve, diffusion properties are derived from the average tensor. We consider the following measures:

- the three eigenvalues, λ_1, λ_2 and λ_3 , of the average diffusion tensor. They represent the diffusivities along the three principal directions of the tensor.
- the mean diffusivity (MD), defined by the first moment of the diffusion tensor eigenvalues.
- the fractional anisotropy (FA). FA is a normalized measure of the shape of the diffusion tensor and defines a distance to isotropy:

$$FA = \frac{\sqrt{3}}{\sqrt{2}} \frac{\sqrt{\sum_{i=1}^3 (\lambda_i - \bar{\lambda})^2}}{\sqrt{\sum_{i=1}^3 \lambda_i^2}} \quad (6)$$

- the geodesic anisotropy (GA), defined in Sec. 2.3.

3.4 *Towards Cross-Population Studies*

So far, the proposed fiber tract modeling applies to an individual bundle. It provides a compact representation of the geometry of a tract and of associated diffusion properties. Ultimately, it aims at being used for inter-subject comparison and statistical analysis. This implies correspondence issues that are currently investigated by arc length parameterization and could include local shape features of curves which have been shown to yield typical patterns along major fiber tracts [Corouge et al., 2004]. Applications of the methodology, particularly for group comparison in clinical studies, would require more advanced statistical techniques, for instance for comparison of probability distributions of tensors and hypothesis testing.

4 Experiments and Results

We first apply our quantitative DTI analysis to a single data set to demonstrate feasibility; second we proceed to a validation study to assess the reproducibility of our framework.

4.1 *Experiments on a single data set*

4.1.1 *Data*

We selected one case out of a 3 Tesla high resolution ($2 \times 2 \times 2 \text{ mm}^3$) DT MRI database of healthy controls and applied tractography. The regions of interest are specified following [Mori et al., 2002] and [Jellison et al., 2004] and defined on the FA image using our SNAP tool [Yushkevich et al., 2006]. Nine fiber tracts were extracted. They represent:

- Three commissural sub-bundles passing respectively through the genu, the splenium and the body of the corpus callosum (BCC),
- Two projection tracts part of the corona radiata, from the internal capsule to superior central cortical areas of the left (LIC) and right hemisphere (RIC),
- Four association bundles: the left and right cingulum (LCG and RCG), and the left and right uncinate fasciculus (LUF and RUF).

The extracted bundles were filtered to remove potential outliers or streamlines impracticable for the subsequent analysis, e.g., clusters composed of very short uncinate fibers hooking around the lateral fissure were discarded.

4.1.2 Average of diffusion tensors in cross-sections along tracts

The geometric model and associated diffusion tensor statistics are computed for each of the nine tracts as described in Sec. 3. Figure 4 illustrates the application of our methodology on tract BCC. Figure 5 shows the nine extracted bundles after start and end points

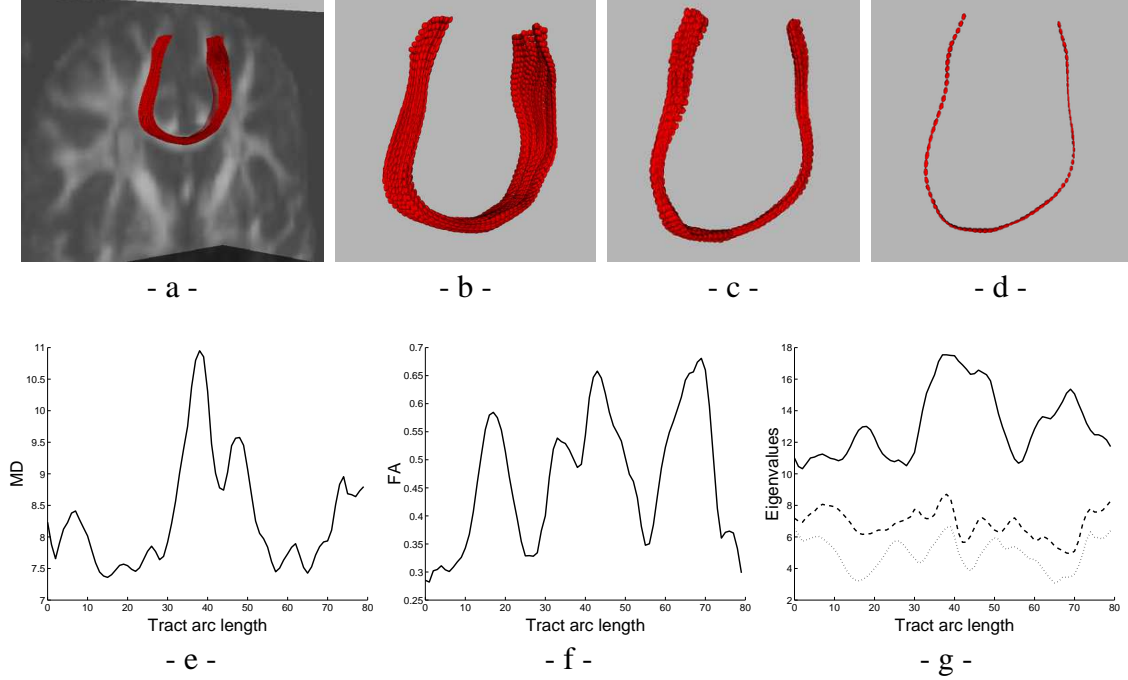


Figure 4. Quantitative analysis is applied to tract BCC. a) DTI data with fiber tract overlaid on a coronal slice of FA image (right-left orientation). b) Streamlines are reparameterized and associated tensors are computed by interpolation. c) The fiber tract is aligned by Procrustes analysis. d) The geometric model is built and tensors statistics are computed along the tract. For visualization purposes, average tensors are displayed along the average curve. e–g) Scalar diffusion properties are derived from the average tensors and are plotted as a function of location along the tract, the start point being set at the superior left side of the brain: e) MD, f) FA, g) eigenvalues: λ_1 (solid lines), λ_2 (dashed lines) and λ_3 (dotted lines).

were placed at locations where the fibers start dispersing towards various cortical areas using cutting planes perpendicular to the bundle directions. Figure 6 shows the mean tensors along the estimated mean shape for each selected fiber tract. For visualization purposes, the mean curve has been translated back to the center of mass of the corresponding bundle in the original coordinate system. To assess the representativity of the estimated mean curve shape, we reconstruct an approximation to the initial fiber tract by applying to the average curve the inverse rigid transformations estimated by Procrustes analysis for each fiber. Given (4), the reconstructed fiber tract $\tilde{\mathcal{F}}$ is defined by

$$\tilde{\mathcal{F}} = \{\tilde{\mathbf{F}}_n = \bar{\mathbf{F}}\mathbf{\Gamma}_n^t + \mathbf{g}_n, 1 \leq n \leq N\} \quad (7)$$

where $\bar{\mathbf{F}}$ is the estimated mean fiber. Let us define the distance $d(\mathbf{F}, \tilde{\mathbf{F}})$ between an original and a reconstructed fiber as the mean point to point distance between the two

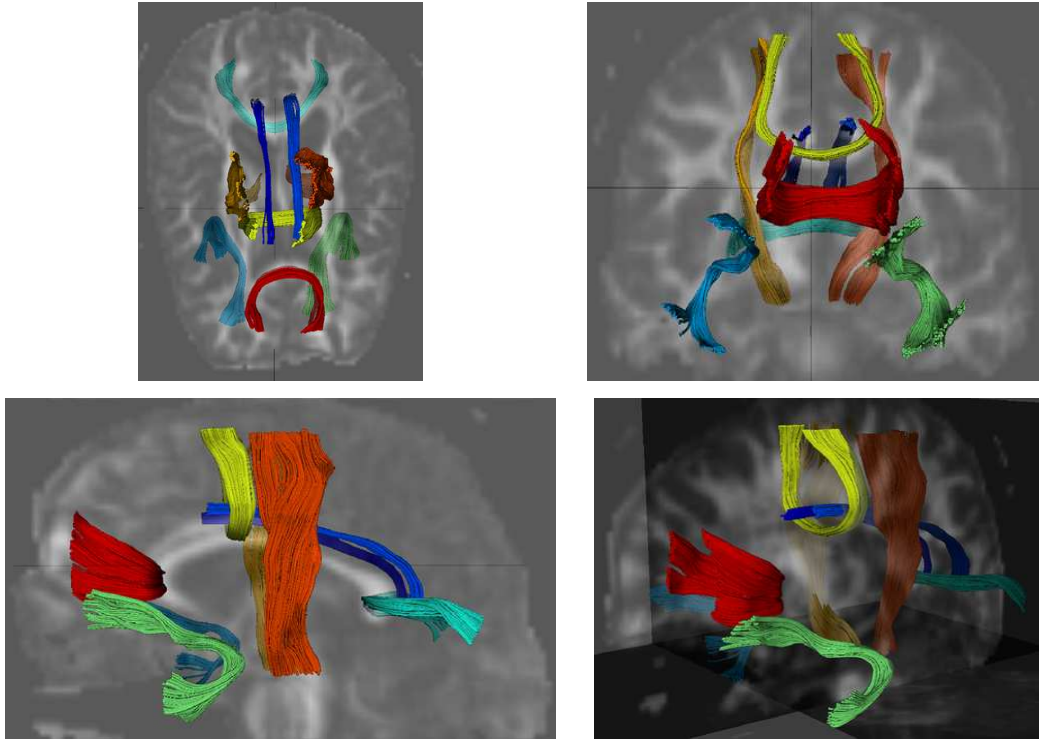


Figure 5. Axial, coronal, sagittal and 3D views of the nine extracted fiber tracts on a single dataset. Yellow: tract BCC, red: genu, cyan: splenium, dark yellow and orange: tracts LIC and RIC, dark cyan and green: LUF and RUF, dark blue and blue: LCG and RCG.

fibers. The mean and standard deviation of $d(\mathbf{F}, \tilde{\mathbf{F}})$ are presented in Table 1. For all tracts, the error is less than one voxel. Figure 7.a shows the geodesic standard deviation

Table 1

Mean and standard deviation of the distance (in voxels) between an original and a reconstructed fiber for the nine tracts of the single data set. Voxel size is $2 \times 2 \times 2 \text{ mm}^3$.

Tract	Average	Standard deviation
BCC	0.33	0.10
Genu	0.53	0.12
Splenium	0.63	0.24
LIC	0.49	0.13
RIC	0.60	0.15
LUF	0.62	0.25
RUF	0.54	0.17
LCG	0.17	0.08
RCG	0.20	0.06

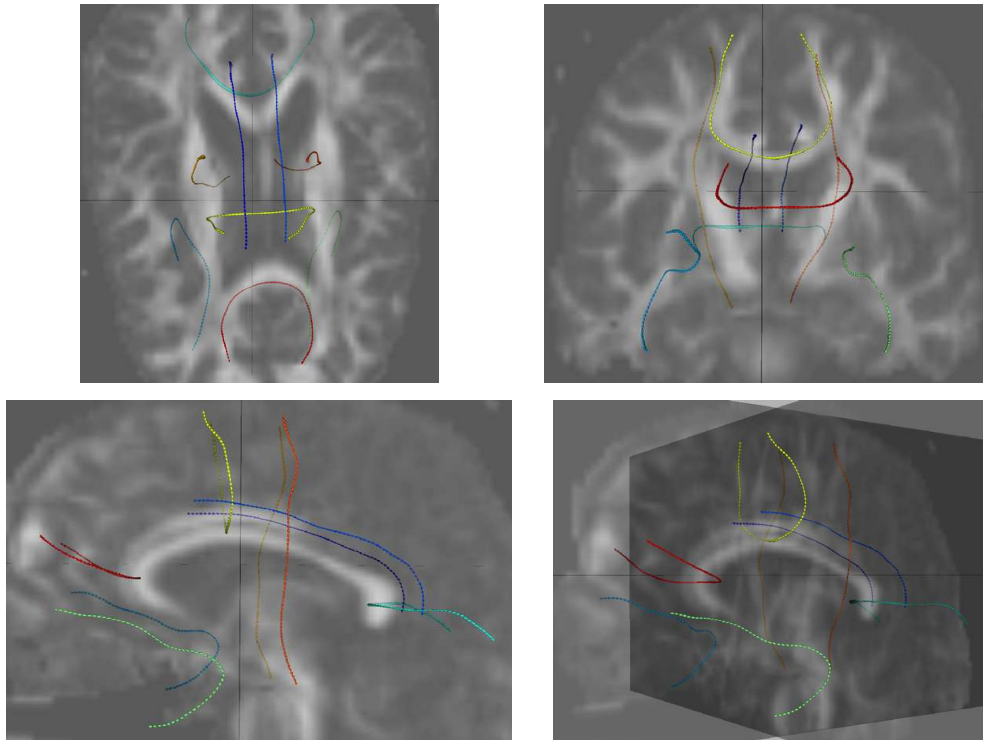


Figure 6. Average tensors calculated in cross-sections displayed along central spine of each bundle. For visualization purposes, each central spine has been translated back in the original coordinate system. Yellow: tract BCC, red: genu, cyan: splenium, dark yellow and orange: tracts LIC and RIC, dark cyan and green: LUF and RUF, dark blue and blue: LCG and RCG.

for all mean locations of tract BCC. Description of variability will be important for hypothesis testing in group studies.

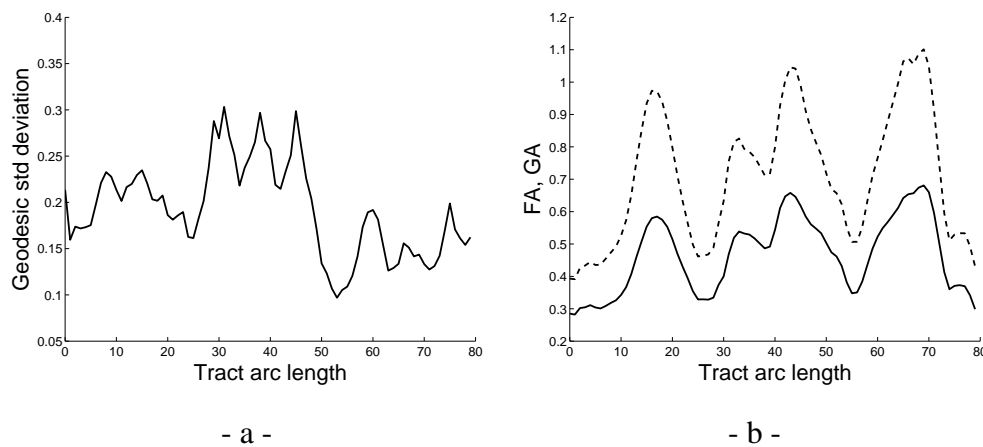


Figure 7. a) Geodesic standard deviation corresponding to tract BCC of the single data set. b) FA (solid line) and GA (dashed line) plot corresponding to tract BCC of the single data set.

4.1.3 Diffusion properties along tracts

The diffusion properties computed from the mean tensors are plotted in Fig. 4 for tract BCC and in Fig. 8 for tract RIC and RUF. Figure 8 top displays FA along the fiber tract for all bundles. The FA plots clearly reflect the pattern shown in the color display. Such

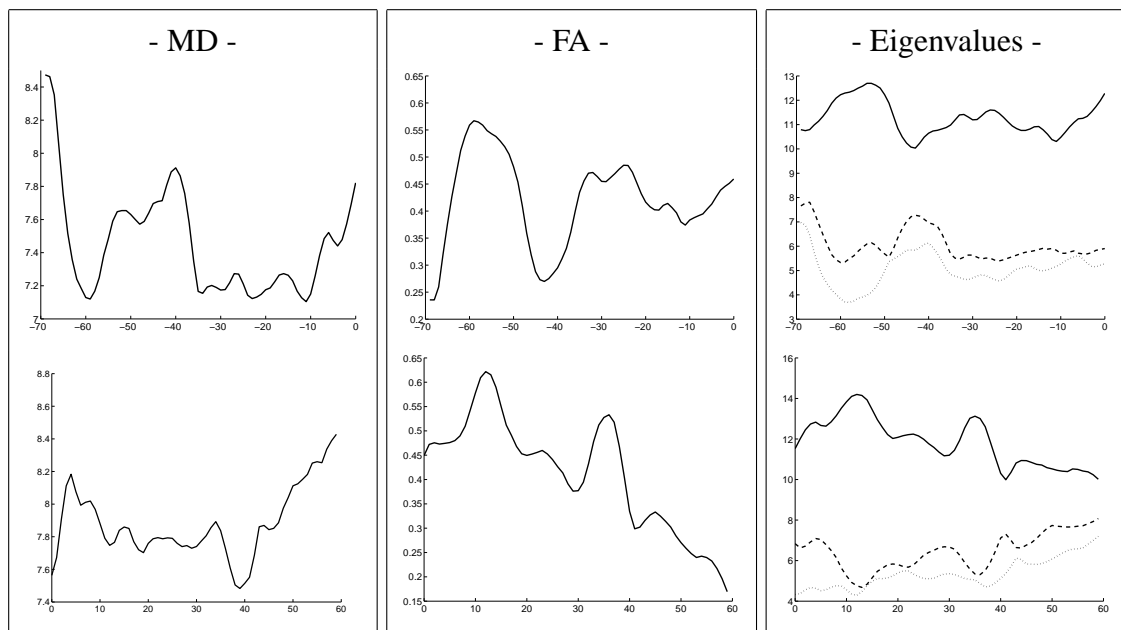
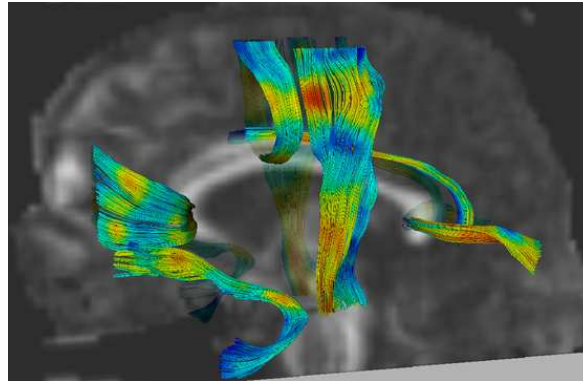


Figure 8. Diffusion properties along fiber tracts. Top: FA is displayed along the nine tracts of the single data sets. Plots: MD, FA and eigenvalues (λ_1 : solid line, λ_2 : dashed line, λ_3 : dotted line) are plotted for tract RIC (top row) and tract RUF (bottom row) as a function of tract arc length. The start point for RIC, resp. RUF, corresponds to the most inferior, resp. anterior, location.

visualizations demonstrate that the diffusion properties vary significantly as a function of location along the tract. This might be explained by the coarse sampling of the underlying macroscopic structures, partial voluming and also natural variation of fiber density.

4.2 Validation study

4.2.1 Data

One subject is imaged six times using slightly different head position. DTI imaging is done on a 3.0 Tesla whole-body MRI system (Trio, Siemens Medical Systems, Malvern, PA, USA) using the 8-channel head coil. Diffusion tensor axial images included 6 diffusion directions with a b value of 1000 sec/mm^2 , plus an acquisition where $b = 0 \text{ sec/mm}^2$, using the parameters of 25.6 cm FOV; 2 mm slice, 0 gap; $\text{Tr} = 10000$, $\text{Te} = 80$; 1345 Hz/pixel bandwidth; 128×128 matrix.

4.2.2 Evaluation framework

An average DTI is computed from all 6 scans after alignment. One scan is arbitrarily chosen as a target. For each of the five other scans, a rigid transformation (i.e., translation and rotation) towards this target is estimated from the baseline images using the RView software [Rueckert, 2002]. The maximal translation and rotation were respectively 3 voxels and 0.5° . Baseline and sensitized to diffusion images are then registered to the target coordinate system and averaged. An average tensor field is then computed.

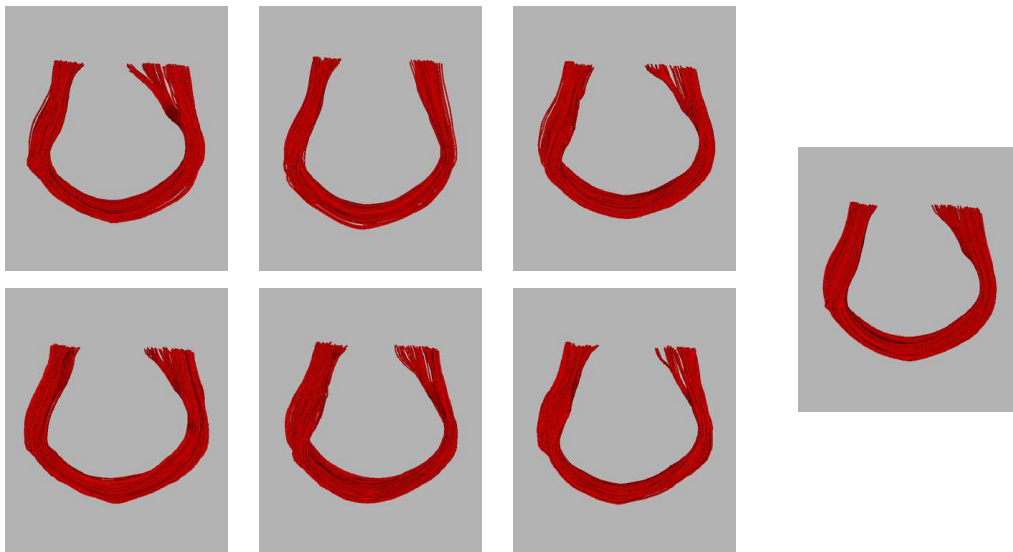


Figure 9. DTI data sets used in the validation study: callosal fiber tracts extracted from the six repeated scans (left) and from the average DTI (right).

For all six scans and the average DTI, a callosal bundle connecting left and right motor cortices is extracted by tractography as shown in Fig. 9. Regions of interest are specified on the target scan and transformed by rigid registration onto the five other scans. The processing pipeline is applied to each single scan and to the average DTI as described in Sec. 3. Number of streamlines, mean shape of the geometric model and diffusion properties derived from the average tensors are compared.

4.2.3 Results

Table 2 lists the number of streamlines obtained for each experiment. Table 3 gives the mean and standard deviation of the distance between the average curves of the six scans, as well as the mean and standard deviation of the distance between the average curve of one scan and of the average DTI. The distance between two average curves is defined as in Sec. 4.1.2, i.e., as the mean point to point distance between the two curves.

Table 2

Number of streamlines obtained for each experiment.

Scans	Scan1	Scan2	Scan3	Scan4	Scan5	Scan6	Average
Number of streamlines	197	165	163	194	219	135	296

Table 3

Left: Mean and standard deviation of the distance between the average curves of the six scans. Right: Mean and standard deviation of the distance between the average curve of one scan and of the average DTI.

	$d(\text{Scan}_i, \text{Scan}_j)$	$d(\text{Scan}_i, \text{Average})$
Mean	0.57	0.40
Std	0.16	0.10

Figure 10 presents the diffusion properties derived from tensors statistics along the bundle. The plots show the mean and standard deviation calculated from the 6 scans as a function of arc length. We also compare the mean values with results from the average DTI image. Results at the center, i.e., at the position of the midsagittal plane, are presented in Table 4.

Table 4

Diffusion properties at the position of the midsagittal for each of the 6 scans and the average DTI.

	Scan1	Scan2	Scan3	Scan4	Scan5	Scan6	Mean	Std	%Std	Average
MD	8.71	9.43	9.39	9.62	10.09	9.90	9.52	0.48	5.05	8.93
FA	0.60	0.58	0.54	0.53	0.51	0.53	0.55	0.03	5.90	0.60
GA	0.90	0.85	0.79	0.78	0.74	0.77	0.80	0.06	7.16	0.88
λ_1	14.54	15.12	15.08	15.58	16.09	15.68	15.35	0.54	3.55	14.97
λ_2	5.37	5.76	6.46	6.65	7.13	6.89	6.38	0.68	10.66	5.48
λ_3	4.68	5.43	5.49	5.80	6.22	5.87	5.58	0.52	9.39	5.01

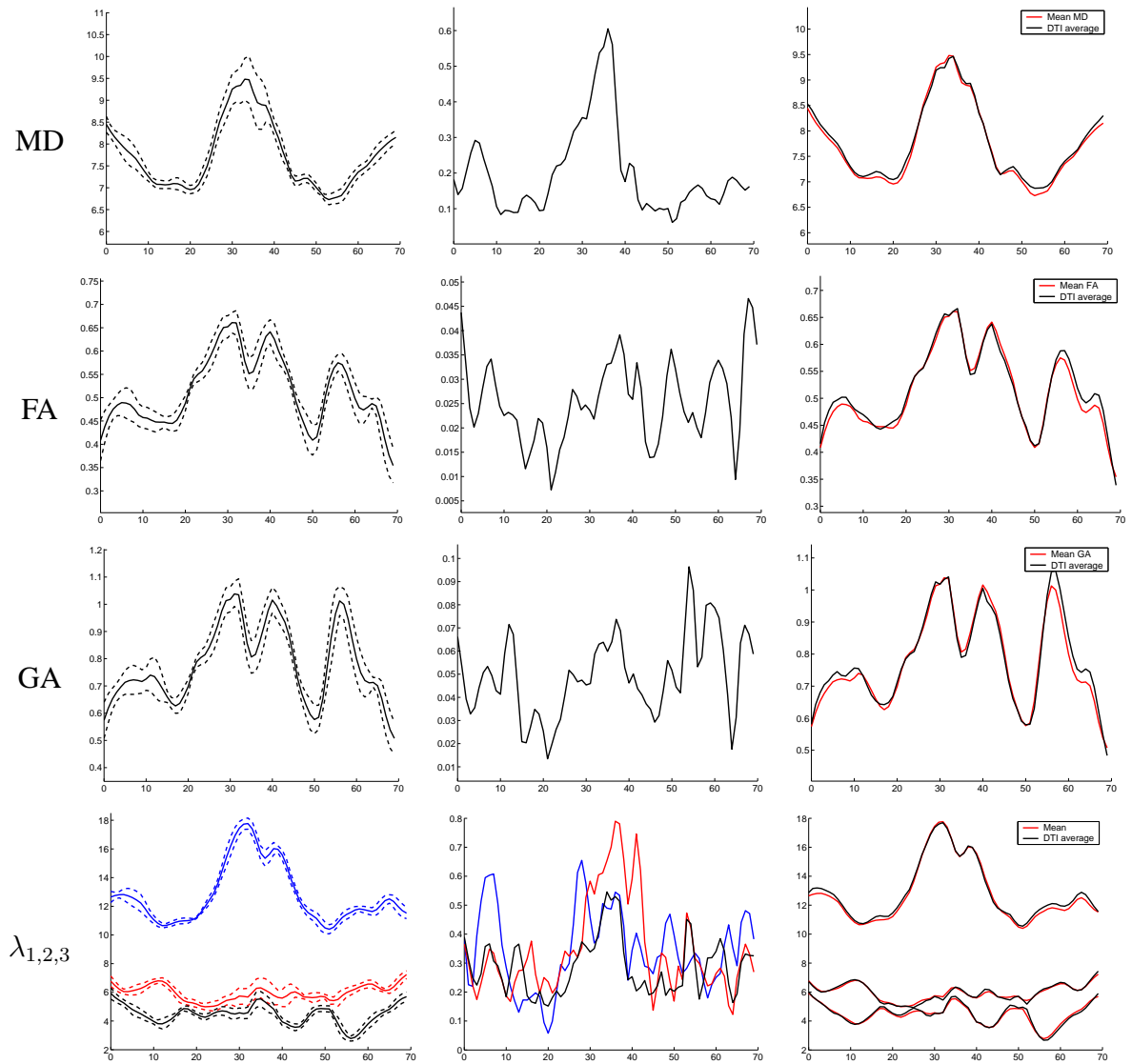


Figure 10. Diffusion properties derived from average tensors plotted as a function of arc length (left-right orientation). From top to bottom: MD, FA, GA, Eigenvalues (λ_1 is in blue, λ_2 in red and λ_3 in black). From left to right: mean and standard deviation calculated from the six scans, standard deviation calculated from the 6 scans, comparison of the values calculated from the six scans and from the average DTI.

4.2.4 Discussion

The number of obtained streamlines distinctly differs from one scan to the other, and is the largest when tractography is applied to the average DTI. This has to be expected since fiber tract reconstruction is an ill-posed problem and so, highly sensitive to noise and to partial voluming effects. Consequently, the shape of the average curve is also more variable: the standard deviation of the distance between average curves appear relatively high in regards to its mean which is small (less than a voxel). However, in spite of the instability of tractography, the average diffusion tensor values are approximately within 5%Std for MD and FA and only 3.5%Std for the first eigenvalue at the position of the

midsagittal plane. It is interesting to note that the second and third eigenvalues are the most variable in the center area. Along the bundles, average diffusion properties across the six scans appear quite stable and are very close to the diffusion properties of the average DTI.

5 Conclusion

We have presented a new framework for fiber-tract oriented quantitative analysis of DTI data. It combines a geometric model of fiber tracts with diffusion tensor statistics. We use non linear statistics for tensor interpolation and averaging. Unlike most other statistical analyses of DTI data, we do not compute statistics on scalar measurements derived from tensors but we do compute statistics on diffusion tensors followed by calculation of tensor properties. The different behaviour of FA and the new geodesic anisotropy, GA, which are both measures of anisotropy, is shown in Fig. 7.b. GA, which is a non-normalized measure, lies in a wider range of values. This indicates a higher sensitivity to anisotropy changes and would suggest a higher discriminative power. This will be further investigated in future work.

Results obtained on a single data set proves the feasibility of our pipeline on various white matter fiber tracts. The validation study demonstrates a good reproducibility of diffusion tensor measurements and statistics in regions of interest defined by fiber tracts. Besides variability of diffusion properties along fiber bundles is clearly shown. It indicates that region of interest analysis is not sufficient and might be very sensitive to the exact definition of cross-sections.

Fiber tract modeling will potentially serve for improved inter-individual registration and comparison of diffusion tensor properties along and across fiber tracts. Clinical research is interested in a quantitative analysis which finally might lead to answer questions in regard to fiber integrity or fiber disruption and its effect on brain connectivity. Moreover, modeling of fiber tracts in healthy controls will help to study geometric and diffusion changes of white matter tracts in the presence of pathology, e.g. tumor and edema or white matter lesions.

6 Acknowledgements

This research is supported by the NIH NIBIB grant P01 EB002779, the NIMH Silvio Conte Center for Neuroscience of Mental Disorders MH064065, and the UNC Neurodevelopmental Disorders Research Center HD 03110. This work is part of the National Alliance for Medical Image Computing (NAMIC), funded by the National Institutes of Health through the NIH Roadmap for Medical Research, Grant U54 EB005149. Information on the National Centers for Biomedical Computing can be obtained from

<http://nihroadmap.nih.gov/bioinformatics>. We acknowledge the Insight Toolkit community for providing the software framework for the DTI analysis algorithms. Dr. Weili Lin, UNC Radiology, and James Mc Fall, Duke University, are acknowledged for providing us with the DT MRI data.

Appendices

A Glossary

$PD(3)$	space of diffusion tensors
$GL^+(3)$	group of positive-determinant matrices
$d(p_1, p_2)$	geodesic distance between two tensors $p_1, p_2 \in PD(3)$
$\det(p)$	determinant of the tensor p
I_3	3×3 identity matrix
$SO(3)$	rotation group for three-dimensional space
$\lambda_1, \lambda_2, \lambda_3$	eigenvalues of a tensor in decreasing order
MD	mean diffusivity
FA	fractional anisotropy
GA	geodesic anisotropy

B Theoretical Framework: Mathematical Details

Symmetric spaces [Helgason, 1978] arise from transformation groups on manifolds. The Riemannian metric is chosen to be invariant under the group transformations. The symmetric space structure of $PD(3)$ arises from transformations by $GL^+(3)$, the group of positive-determinant matrices. The transformation of a diffusion tensor $p \in PD(3)$ by a matrix $g \in GL^+(3)$ is given by $p \mapsto gpg^T$. Because of the algebraic nature of the symmetric space structure, distance and geodesic computations on $PD(3)$ are also algebraic in nature. For instance, the geodesic distance between two tensors $p_1, p_2 \in PD(3)$ can be computed using singular-value decomposition (SVD) as follows:

- Let $p_1 = U\Lambda U^T$ be the SVD of p_1 , set $g = U\sqrt{\Lambda}$.
- Compute the action of g^{-1} on p_2 : $y = g^{-1}p_2(g^{-1})^T$.
- Again using SVD, compute the eigenvalues σ_i of y .
- The geodesic distance is $d(p_1, p_2) = \left(\sum_{i=1}^3 \log(\sigma_i)^2\right)^{\frac{1}{2}}$.

B.1 Statistics of Diffusion Tensors

We now define the mean and variance of diffusion tensors respecting the geometry of the space. Following Fréchet [Fréchet, 1948], we define the average as the minimum mean squared error estimator under the natural Riemannian metric defined above. Given a set of diffusion tensors $p_1, \dots, p_N \in PD(3)$ the mean is defined as

$$\mu = \arg \min_{p \in PD(3)} \sum_{i=1}^N d(p, p_i)^2. \quad (\text{B.1})$$

This minimization problem can be solved using a gradient descent method as described in [Fletcher and Joshi, 2004]. This is analogous to the algorithm for computing the intrinsic mean given by Pennec [Pennec, 1999]. Having defined the mean, we define the sample variance of the data as the expected value of the squared geodesic distances from the mean. Given a set of diffusion tensors $p_1, \dots, p_N \in PD(3)$, we define the sample variance as

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N d(\mu, p_i)^2. \quad (\text{B.2})$$

B.2 Interpolation of Tensors

For developing consistent interpolation between diffusion tensors we extend the definition of the mean defined above to weighted averaging. Using a least-squares criterion, we define the weighted average of diffusion tensors $p_1, \dots, p_N \in PD(3)$ as

$$\text{Ave}(\{w_i\}, \{p_i\}) = \arg \min_{p \in PD(3)} \sum_{i=1}^N w_i d(p, p_i)^2, \quad (\text{B.3})$$

where w_1, \dots, w_N are positive real weights that sum to 1.

For interpolating tensors within a voxel, trilinear weights may be used for the w_i . In this paper, we only focus on trilinear weights although higher order interpolation may be defined using the same concept. This interpolation is a natural generalization of trilinear interpolation of scalar values, i.e., if we replaced the diffusion tensors in the above definitions with real numbers, we would arrive at trilinear interpolation. It follows easily from the use of trilinear weights that the interpolation function does indeed interpolate the corner points. It can also be shown that the interpolation function is continuous on $[0, 1]^3$ (see [Fletcher, 2004] for a proof).

B.3 Geodesic Anisotropy

As introduced in Sec. 2.3, the geodesic anisotropy of a diffusion tensor $p \in PD(3)$ is the geodesic distance between p and the closest isotropic diffusion tensor. Thus we define

the geodesic anisotropy as

$$GA(p) = \min_s d(sI_3, p). \quad (\text{B.4})$$

It turns out that the nearest isotropic diffusion tensor to p is the one with the same determinant as p , i.e., the matrix $(\det(p))^{\frac{1}{3}} \cdot I_3$. With this observation we can explicitly write the GA of the tensor based on the eigenvalues. Let λ_i denote the eigenvalues of p , and let $\overline{\log \lambda}$ denote the average of the logs of the eigenvalues. The geodesic anisotropy of p can be written as

$$\begin{aligned} GA(p) &= d((\det(p))^{\frac{1}{3}} \cdot I_3, p) \\ &= \left(\sum_{i=1}^3 (\log(\lambda_i) - \overline{\log \lambda})^2 \right)^{\frac{1}{2}}. \end{aligned} \quad (\text{B.5})$$

References

- [Alexander et al., 1999] Alexander, D., Gee, J., and Bajcsy, R. (1999). Similarity measures for matching diffusion tensor images. In Pridmore, T. and Elliman, D., editors, *Proc. of the British Machine Vision Conference*, pages 93–102.
- [Alexander et al., 2001] Alexander, D., Pierpaoli, C., Basser, P., and Gee, J. (2001). Spatial transformations of Diffusion Tensor Magnetic Resonance images. *IEEE Trans. Medical Imaging*, 20(11):1131–1139.
- [Arsigny et al., 2005] Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. (2005). Fast and simple calculus on tensors in the Log-Euclidean framework. In *Proc. of Medical Image Computing and Computer-Assisted Intervention*, volume 3749 of *Lecture Notes in Computer Science*, pages 115–122.
- [Batchelor et al., 2005] Batchelor, P., Moakher, M., Atkinson, D., Calamante, F., and Connelly, A. (2005). A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53(1):221–225.
- [Beaulieu, 2002] Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system - a technical review. *NMR in Biomedicine*, 15(7-8):435–455.
- [Cootes et al., 1995] Cootes, T., Taylor, C., Cooper, D., and Graham, J. (1995). Active shape models - their training and application. *Computer Vision and Image Understanding*, 61(1):38–59.
- [Corouge et al., 2004] Corouge, I., Gouttard, S., and Gerig, G. (2004). Towards a shape model of white matter fiber bundles using Diffusion Tensor MRI. In *Proc. IEEE International Symposium on Biomedical Imaging*, pages 344–347.
- [Coulon et al., 2004] Coulon, O., Alexander, D., and Arridge, S. (2004). Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8:47–67.

- [Ding et al., 2003] Ding, Z., Gore, J., and Anderson, A. (2003). Classification and quantification of neuronal fiber pathways using diffusion tensor MRI. *Magnetic Resonance in Medicine*, 49:716–721.
- [Fillard and Gerig, 2003] Fillard, P. and Gerig, G. (2003). Analysis tool for diffusion tensor MRI. In *Proc. of Medical Image Computing and Computer-Assisted Intervention*, volume 2879 of *Lecture Notes in Computer Science*, pages 967–968.
- [Fillard et al., 2003] Fillard, P., Gilmore, J., Lin, W., and Gerig, G. (2003). Quantitative analysis of white matter fiber properties along geodesic paths. In *Proc. of Medical Image Computing and Computer-Assisted Intervention*, volume 2879 of *Lecture Notes in Computer Science*, pages 16–23.
- [Fletcher, 2004] Fletcher, P. T. (2004). *Statistical Variability in Nonlinear Spaces: Application to Shape Analysis and DT-MRI*. PhD thesis, University of North Carolina.
- [Fletcher and Joshi, 2004] Fletcher, P. T. and Joshi, S. (2004). Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In *Proceedings of ECCV 2004 Workshop on Computer Vision Approaches to Medical Image Analysis (CVAMIA)*, volume LNCS 3117, pages 87–98. Springer-Verlag.
- [Fréchet, 1948] Fréchet, M. (1948). Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, (10):215–310.
- [Goodall, 1991] Goodall, C. (1991). Procrustes methods in the statistical analysis of shape. *J.R. Statist. Soc. B*, 53(2):285–239.
- [Guimond et al., 2002] Guimond, A., Guttman, C., Warfield, S., and Westin, C.-F. (2002). Deformable registration of DT-MRI data based on transformation invariant tensor characteristics. In *IEEE International Symposium on Biomedical Imaging*.
- [Helgason, 1978] Helgason, S. (1978). *Differential Geometry, Lie Groups, and Symmetric Spaces*. Academic Press.
- [Jellison et al., 2004] Jellison, B., Field, A., Medow, J., Laar, M., Salamat, M., and Alexander, A. (2004). Diffusion tensor imaging of cerebral white matter: A pictorial review of physics, fiber tract anatomy, and tumor imaging patterns. *Am J Neuroradiol*, 25:356–369.
- [Jones et al., 2006] Jones, D., Catani, M., Pierpaoli, C., Reeves, S., Shergill, S., O’Sullivan, M., Golesworthy, P., McGuire, P., Horsfield, M., Simmons, A., Williams, S., and Howard, R. (2006). Age effects on diffusion tensor magnetic resonance imaging tractography measures of frontal cortex connections in schizophrenia. *Human Brain Mapping*.
- [Jones et al., 2002] Jones, D., Griffin, L., Alexander, C., Catani, M., Horsfield, M., Howard, R., and Williams, S. (2002). Spatial normalisation and averaging of diffusion tensor MRI data sets. *NeuroImage*, 17:592–617.
- [Lim and Helpert, 2002] Lim, K. and Helpert, J. (2002). Neuropsychiatric applications of dti - a review. *NMR in Biomedicine*, 15(7-8):587–593.
- [Mori et al., 2002] Mori, S., Kaufmann, W., Davatzikos, C., Stieltjes, B., Amodei, L., Fredericksen, K., Pearlson, G., Melhem, E., Solaiyappan, M., Raymond, G., Moser, H., and van Zijl P.C.M. (2002). Imaging cortical association tracts in the human brain using diffusion-tensor-based axonal tracking. *Magnetic Resonance in Medicine*, 47:215–223.

- [Parker et al., 2000] Parker, G., Schnabel, J., Symms, M., Werring, D., and Barker, G. (2000). Nonlinear smoothing for reduction of systematic and random errors in Diffusion Tensor Imaging. *J. Magnetic Resonance Imaging*, 11(6):702–710.
- [Pennec, 1999] Pennec, X. (1999). Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In *IEEE Workshop on Nonlinear Signal and Image Processing*.
- [Pennec et al., 2006] Pennec, X., Fillard, P., and Ayache, N. (2006). A Riemannian framework for tensor computing. *Int. J. Computer Vision*, 66(1):41–66.
- [Poupon et al., 2001] Poupon, C., Mangin, J.-F., Clark, C., Frouin, V., Régis, J., Le Bihan, D., and Bloch, I. (2001). Towards inference of human brain connectivity from MR diffusion tensor data. *Medical Image Analysis*, 5:1–15.
- [Rueckert, 2002] Rueckert, D. (2002). Rview. <http://www.doc.ic.ac.uk/~dr/software/>.
- [Westin et al., 2002] Westin, C.-F., Maier, S., Mamata, H., Nabavi, A., Jolesz, F., and Kikinis, R. (2002). Processing and visualization for diffusion tensor MRI. *Medical Image Analysis*, 6:93–108.
- [Xu et al., 2002] Xu, D., Mori, S., Solaiyappan, M., van Zijl, P., and Davatzikos, C. (2002). A framework for callosal fiber distribution analysis. *NeuroImage*, 17:1131–1143.
- [Yushkevich et al., 2006] Yushkevich, P., Piven, J., Cody Hazlett, H., Gimpel Smith, R., Ho, S., Gee, J., and Gerig, G. (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, In print.
- [Zhang et al., 2004] Zhang, H., Yushkevich, P., and Gee, J. (2004). Registration of diffusion tensor images. In *Proc. of Conf. Computer Vision Pattern Recognition*, pages 842–847.
- [Zhang et al., 2005] Zhang, H., Yushkevich, P., and Gee, J. (2005). Deformable registration of diffusion tensor MR images with explicit orientation optimization. In Duncan, J. and Gerig, G., editors, *Proc. of Medical Image Computing and Computer-Assisted Intervention*, volume 3749 of *LNCS*, pages 172–179. Springer-Verlag.