

Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data

P. Thomas Fletcher
Scientific Computing and Imaging Institute
University of Utah
50 S Central Campus Drive Room 3490
Salt Lake City, UT 84112, USA
fletcher@sci.utah.edu

Sarang Joshi
Medical Imaging and Display Analysis Group
University of North Carolina at Chapel Hill
101 Manning Drive, Campus Box 7512
Chapel Hill, NC 27599, USA
sjoshi@unc.edu

Abstract

The tensors produced by diffusion tensor magnetic resonance imaging (DT-MRI) represent the covariance in a Brownian motion model of water diffusion. Under this physical interpretation, diffusion tensors are required to be symmetric, positive-definite. However, current approaches to statistical analysis of diffusion tensor data, which treat the tensors as linear entities, do not take this positive-definite constraint into account. This difficulty is due to the fact that the space of diffusion tensors does not form a vector space. In this paper we show that the space of diffusion tensors is a type of curved manifold known as a Riemannian symmetric space. We then develop methods for producing statistics, namely averages and modes of variation, in this space. We show that these statistics preserve natural geometric properties of the tensors, including the constraint that their eigenvalues be positive. The symmetric space formulation also leads to a natural definition for interpolation of diffusion tensors and a new measure of anisotropy. We expect that these methods will be useful in the registration of diffusion tensor images, the production of statistical atlases from diffusion tensor data, and the quantification of the anatomical variability caused by disease. The framework presented in this paper should also be useful in other applications where symmetric, positive-definite tensors arise, such as mechanics and computer vision.

Keywords: Diffusion tensor MRI, statistics, Riemannian manifolds

1 Introduction

The clinical significance of diffusion tensor magnetic resonance imaging (DT-MRI) [2] derives from its ability to image *in vivo* the structure of white matter fibers in the brain. To fully harness the power of this relatively new modality, statistical tools for handling diffusion tensor data types are needed. A 3D diffusion tensor models the covariance of the Brownian motion of water at a voxel, and as such is required to be a 3×3 , symmetric, positive-definite matrix. The aim of this paper is to provide new methods for the statistical analysis of diffusion tensors that take into account the requirement that the tensors be positive-definite.

Diffusion tensor imaging has shown promise in clinical studies of brain pathologies, such as multiple sclerosis and stroke, and in the study of brain connectivity [5]. Several authors have addressed the problem of estimation and smoothing within a DT image [8, 10, 23]. Further insights might be had from the use of diffusion tensor imaging in intersubject studies. Statistical brain atlases have been used in the case of scalar images to quantify anatomical variability across patients. However, relatively little work has been done towards constructing statistical brain atlases from diffusion tensor images. Alexander *et al.* [1] describe a method for the registration of multiple DT images into a common coordinate frame, however, they do not include a statistical analysis of the diffusion tensor data. Previous attempts [3, 20] at statistical analysis of diffusion tensors within a DT image use a Gaussian model of the linear tensor coefficients.

In this paper we demonstrate that the space of diffusion tensors is more naturally described as a Riemannian symmetric space, rather than a linear space. In our previous work [14] we introduced *principal geodesic analysis* (PGA) as an analog of principal component analysis for studying the statistical variability of Lie group data. Extending these ideas to symmetric spaces, we develop new methods for computing averages and describing the variability of diffusion tensor data. We show that these statistics preserve natural properties of the diffusion tensors, most importantly the positive-definiteness, that are not preserved by linear statistics. We also develop a natural method for 3D interpolation of diffusion tensor images based on the symmetric space formulation. Also, the Riemannian geometry of diffusion tensors leads to a natural definition of anisotropy, called *geodesic anisotropy*, which is based on the geodesic distance to the nearest isotropic tensor. The framework presented in this paper provides the statistical methods needed for constructing statistical atlases of diffusion tensor images.

This work is an expanded version of the material found in [12] and the first author's thesis work [11]. Many of the ideas appearing in this paper have been independently developed by Batchelor *et al.* [4]. In their work they describe the same symmetric space geometry for averaging an arbitrary number of tensors, geodesic anisotropy, and interpolation between two diffusion tensors. Our work has gone somewhat further, being the first to describe second order statistics, namely covariance and modes of variation, in the symmetric space framework. Also, our interpolation scheme handles an arbitrary number of tensors, rather than only two. Thus we are able to define

full 3D interpolation of diffusion tensor images. More recently, the symmetric space framework for averaging and covariance have also been independently developed in [19, 22], with tensor interpolation also being developed in [22]. In this paper and in [11] the theoretical properties of the Riemannian interpolation are described in more detail. The statistical methods and interpolation presented in this paper have recently been applied to the quantification of white matter diffusion along fiber tracts in [9].

The rest of the paper is organized as follows. Section 2 demonstrates the motivation for using a curved geometry for diffusion tensors, rather than a linear one. Section 3 develops the necessary theory of diffusion tensors as a Riemannian symmetric space. Section 4 develops the mean and PGA statistics of diffusion tensors. Section 6 describes the new interpolation method for 3D diffusion tensor images based on symmetric space geometry. We introduce a new anisotropy measure, called geodesic anisotropy, in Section 7.

2 The Space of Diffusion Tensors

Recall that a real $n \times n$ matrix A is symmetric if $A = A^T$ and positive-definite if $x^T A x > 0$ for all nonzero $x \in \mathbb{R}^n$. We denote the space of all $n \times n$ symmetric, positive-definite matrices as $PD(n)$. The tensors in DT-MRI are thus elements of $PD(3)$. The space $PD(n)$ forms a convex subset of \mathbb{R}^{n^2} . One can define a linear average of N positive-definite, symmetric matrices A_1, \dots, A_N as $\mu = \frac{1}{N} \sum_{i=1}^N A_i$. This definition minimizes the Euclidean metric on \mathbb{R}^{n^2} . Since $PD(n)$ is convex, μ lies within $PD(n)$. However, linear averages do not interpolate natural properties. The linear average of matrices of the same determinant can result in a matrix with a larger determinant. Second order statistics are even more problematic. The standard principal component analysis is invalid because the straight lines defined by the modes of variation do not stay within the space $PD(n)$. In other words, linear PCA does not preserve the positive-definiteness of diffusion tensors. The reason for such difficulties is that space $PD(n)$, although a subset of a vector space, is not a vector space; for example, the negation of a positive-definite matrix is not positive-definite.

In this paper we derive a more natural metric on the space of diffusion tensors, $PD(n)$, by viewing it not simply as a subset of \mathbb{R}^{n^2} , but rather as a Riemannian symmetric space. Following Fréchet [15], we define the average as the minimum mean squared error estimator under this metric. We develop the method of principal geodesic analysis to describe the variability of diffusion tensor data. Principal geodesic analysis is the generalization of principal component analysis to manifolds. In this framework the modes of variation are represented as flows along geodesic curves, i.e., shortest paths under the Riemannian metric. These geodesic curves, unlike the straight lines of \mathbb{R}^{n^2} , are completely contained within $PD(n)$, that is, they preserve the positive-definiteness. Principal component analysis generates lower-dimensional subspaces that maximize the projected variance of the data. Thus the development of principal geodesic analysis requires that we generalize the concepts of variance and projection onto lower-dimensional subspaces for data in symmetric spaces.

To illustrate these issues, consider the space $PD(2)$, the 2×2 symmetric, positive-

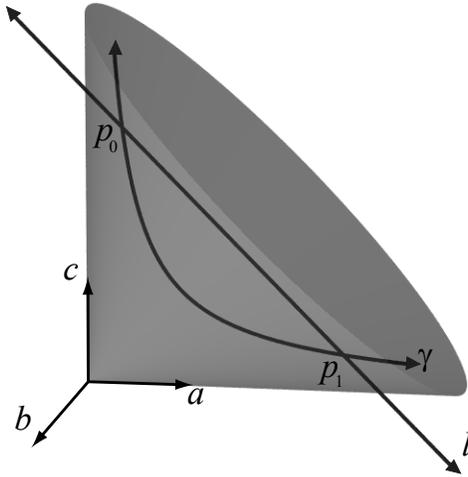


Figure 1: The space $PD(2)$, showing the geodesic γ and the straight line l between the two points p_0 and p_1 .

definite matrices. A matrix $A \in PD(2)$ is of the form

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}, \quad ac - b^2 > 0, \quad a > 0.$$

If we consider the matrix A as a point $(a, b, c) \in \mathbb{R}^3$, then the above conditions describe the interior of a cone as shown in Fig. 1. The two labeled points are $p_0 = (1, 0, 7)$, $p_1 = (7, 0, 1)$. The straight line l between the two points, i.e., the geodesic in \mathbb{R}^{n^2} , does not remain contained within the space $PD(2)$. The curve γ is the geodesic between the two points when $PD(2)$ is considered as a Riemannian symmetric space. This geodesic lies completely within $PD(2)$. We chose $PD(2)$ as an example since it can be easily visualized, but the same phenomenon occurs for general $PD(n)$, i.e., $n > 2$.

3 The Geometry of $PD(n)$

In this section we show that the space of diffusion tensors, $PD(n)$, can be formulated as a Riemannian symmetric space. This leads to equations for computing geodesics that will be essential in defining the statistical methods for diffusion tensors. The differential geometry of diffusion tensors has also been used in [8], where the diffusion tensor smoothing was constrained along geodesic curves. A more thorough treatment of symmetric spaces can be found in [6, 17].

A *symmetric space* is a connected Riemannian manifold M such that for each $x \in M$ there is an isometry σ_x which (1) is involutive, i.e., $\sigma_x^2 = \text{id}$, and (2) has x as an isolated fixed point, that is, there is a neighborhood U of x where σ_x leaves only x fixed. It can be shown that σ_x is the map that reverses all geodesics through the

point x . Riemannian symmetric spaces, and the methods for computing geodesics and distances on them, arise naturally from Lie group actions on manifolds.

3.1 Lie Group Actions

A *Lie group* is an algebraic group G that also forms a differentiable manifold, where the two group operations, multiplication and inversion, are smooth mappings. Many common geometric transformations of Euclidean space form Lie groups. For example, rotations, translations, and affine transformations of \mathbb{R}^n all form Lie groups. More generally, Lie groups can be used to describe transformations of smooth manifolds.

Given a manifold M and a Lie group G , a *smooth group action* of G on M , or *smooth G -action* on M , is a smooth mapping $\phi : G \times M \rightarrow M$ such that for all $g, h \in G$, and all $x \in M$ we have $\phi(e, x) = x$, and $\phi(g, \phi(h, x)) = \phi(gh, x)$, where e is the identity element of G . Consider the Lie group of all $n \times n$ real matrices with positive determinant, denoted $GL^+(n)$. This group acts on $PD(n)$ via

$$\begin{aligned} \phi : GL^+(n) \times PD(n) &\rightarrow PD(n) \\ \phi(g, p) &= gp g^T. \end{aligned} \tag{1}$$

The *orbit* under ϕ of a point $x \in M$ is defined as $G(x) = \{\phi(g, x) : g \in G\}$. In the case that M consists of a single orbit, we call M a *homogeneous space* and say that the G -action is *transitive*. The space $PD(n)$ is a homogeneous space, as is easy to derive from the fact that any matrix $p \in PD(n)$ can be decomposed as $p = gg^T = \phi(g, I_n)$, where $g \in GL^+(n)$ and I_n is the $n \times n$ identity matrix. The *isotropy subgroup* of x is defined as $G_x = \{g \in G : \phi(g, x) = x\}$, i.e., G_x is the subgroup of G that leaves the point x fixed. For $PD(n)$ the isotropy subgroup of I_n is $SO(n) = \{g \in GL^+(n) : gg^T = I_n\}$, i.e., the space of $n \times n$ rotation matrices.

Let H be a closed Lie subgroup of the Lie group G . Then the *left coset* of an element $g \in G$ is defined as $gH = \{gh : h \in H\}$. The space of all such cosets is denoted G/H and is a smooth manifold. There is a natural bijection $G(x) \cong G/G_x$ given by the mapping $g \cdot x \mapsto gG_x$. Therefore, we can consider the space of diffusion tensors, $PD(n)$, as the coset space $GL^+(n)/SO(n)$. An intuitive way to view this is to think of the polar decomposition, which decomposes a matrix $g \in GL^+(n)$ as $g = pu$, where $p \in PD(n)$ and $u \in SO(n)$. Thus, the diffusion tensor space $PD(n) \cong GL^+(n)/SO(n)$ comes from “dividing out” the rotational component in the polar decomposition of $GL^+(n)$.

3.2 Invariant Metrics

A *Riemannian metric* on a manifold M smoothly assigns to each point $x \in M$ an inner product $\langle \cdot, \cdot \rangle_x$ on $T_x M$, the tangent space to M at x . If ϕ is a smooth G -action on M , a metric on M is called *G -invariant* if for each $g \in G$ the map $\phi_g : x \mapsto \phi(g, x)$ is an isometry, i.e., ϕ_g preserves distances on M . The space of diffusion tensors, $PD(n)$, has a metric that is invariant under the $GL^+(n)$ action, which follows from the fact that the isotropy subgroup $SO(n)$ is connected and compact (see [6], Theorem 9.1).

The tangent space of $PD(n)$ at the identity matrix can be identified with the space of $n \times n$ symmetric matrices, $\text{Sym}(n)$. Since the group action $\phi_g : s \mapsto gsg^T$ is linear, its derivative map, denoted $d\phi_g$, is given by $d\phi_g(X) = gXg^T$. If $X \in \text{Sym}(n)$, it is easy to see that $d\phi_g(X)$ is again a symmetric matrix. Thus the tangent space at any point $p \in PD(n)$ is also equivalent to $\text{Sym}(n)$. If $X, Y \in \text{Sym}(n)$ represent two tangent vectors at $p \in PD(n)$, where $p = gg^T, g \in GL^+(n)$, then the Riemannian metric at p is given by the inner product

$$\langle X, Y \rangle_p = \text{tr}(g^{-1}Xp^{-1}Y(g^{-1})^T).$$

Finally, the mapping $\sigma_{I_n}(p) = p^{-1}$ is an isometry that reverses geodesics of $PD(n)$ at the identity, and this turns $PD(n)$ into a symmetric space.

3.3 Computing Geodesics

Geodesics on a symmetric space are easily derived via the group action (see [17] for details). Let p be a point on $PD(n)$ and X a tangent vector at p . There is a unique geodesic, γ , with initial point $\gamma(0) = p$ and tangent vector $\gamma'(0) = X$. To derive an equation for such a geodesic, we begin with the special case where the initial point p is the $n \times n$ identity matrix, I_n , and the tangent vector X is diagonal. Then the geodesic is given by $\gamma(t) = \exp(tX)$, where \exp is the matrix exponential map given by the infinite series

$$\exp(X) = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

For the diagonal matrix X with entries x_i , the matrix exponential is simply the diagonal matrix with entries e^{x_i} .

Now for the general case consider the geodesic γ starting at an arbitrary point $p \in PD(n)$ with arbitrary tangent vector $X \in \text{Sym}(n)$. We will use the group action to map this configuration into the special case described above, i.e., with initial point at the identity and a diagonal tangent vector. Since the group action is an isometry, geodesics and distances are preserved. Let $p = gg^T$, where $g \in GL^+(n)$. Then the action $\phi_{g^{-1}}$ maps p to I_n . The tangent vector is mapped via the corresponding tangent map to $Y = d\phi_{g^{-1}}(X) = g^{-1}X(g^{-1})^T$. Now we may write $Y = v\Sigma v^T$, where v is a rotation matrix and Σ is diagonal. The group action $\phi_{v^{-1}}$ diagonalizes the tangent vector while leaving I_n fixed. We can now use the procedure above to compute the geodesic $\tilde{\gamma}$ with initial point $\tilde{\gamma}(0) = I_n$ and tangent vector $\tilde{\gamma}'(0) = \Sigma$. Finally, the result is mapped back to the original configuration by the inverse group action, ϕ_{gv} . That is,

$$\gamma(t) = \phi_{gv}(\tilde{\gamma}(t)) = (gv) \exp(t\Sigma)(gv)^T.$$

If we flow to $t = 1$ along the geodesic γ we get the Riemannian exponential map at p (denoted Exp_p , and not to be confused with the matrix exponential map), that is, $\text{Exp}_p(X) = \gamma(1)$. In summary we have

Algorithm 1: Riemannian Exponential Map

Input: Initial point $p \in PD(n)$, tangent vector $X \in \text{Sym}(n)$.

Output: $\text{Exp}_p(X)$
 Let $p = u\Lambda u^T$ ($u \in SO(n)$, Λ diagonal)
 $g = u\sqrt{\Lambda}$
 $Y = g^{-1}X(g^{-1})^T$
 Let $Y = v\Sigma v^T$ ($v \in SO(n)$, Σ diagonal)
 $\text{Exp}_p(X) = (gv) \exp(\Sigma)(gv)^T$

An important property of the geodesics in $PD(n)$ under this metric is that they are infinitely extendible, i.e., the geodesic $\gamma(t)$ is defined for $-\infty < t < \infty$. A manifold with this property is called *complete*. Again, Fig. 1 demonstrates that the symmetric space geodesic γ remains within $PD(2)$ for all t . In contrast the straight line l quickly leaves the space $PD(2)$.

The map Exp_p has an inverse, called the Riemannian log map and denoted Log_p . It maps a point $x \in PD(n)$ to the unique tangent vector at p that is the initial velocity of the unique geodesic γ with $\gamma(0) = p$ and $\gamma(1) = x$. Using a similar diagonalization procedure, the log map is computed by

Algorithm 2: Riemannian Log Map

Input: Initial point $p \in PD(n)$, end point $x \in PD(n)$.

Output: $\text{Log}_p(x)$
 Let $p = u\Lambda u^T$ ($u \in SO(n)$, Λ diagonal)
 $g = u\sqrt{\Lambda}$
 $y = g^{-1}x(g^{-1})^T$
 Let $y = v\Sigma v^T$ ($v \in SO(n)$, Σ diagonal)
 $\text{Log}_p(x) = (gv) \log(\Sigma)(gv)^T$

Using the notation of Algorithm 2, geodesic distance between the diffusion tensors $p, x \in PD(n)$ is computed by $d(p, x) = \|\text{Log}_p(x)\|_p = \sqrt{\text{tr}(\log(\Sigma)^2)}$.

3.4 Computations on Noisy Data

Physical laws of diffusion require that the eigenvalues of a diffusion tensor be positive. However, the eigenvalues of tensors in DT-MRI may be nonpositive due to imaging noise and linear estimation of the tensor components from the diffusion weighted images. The computations in the Riemannian exponential and log maps (Algorithms 1 and 2) fail if the tensors involved have any nonpositive eigenvalues. To avoid these problems, we project these invalid noisy tensors to the space of valid tensors by setting their nonpositive eigenvalues to some small positive value. Another option is to simply exclude invalid tensors from the analysis. A more rigorous approach would be to enforce positive eigenvalues during tensor estimation as in the work of Wang *et al.* [23], which also regularizes the tensor field during estimation.

4 Statistics of Diffusion Tensors

Having formulated the geometry of diffusion tensors as a symmetric space, we now develop methods for computing statistics in this nonlinear space.

4.1 Averages of Diffusion Tensors

To define an average of diffusion tensors we follow Fréchet [15], who defines the mean of a random variable in an arbitrary metric space as the point that minimizes the expected value of the sum-of-squared distance function. Consider a set of points $A = \{x_1, \dots, x_N\}$ on a Riemannian manifold M . Then we will be concerned with the sum-of-squared distance function

$$\rho_A(x) = \frac{1}{2N} \sum_{i=1}^N d(\mu, x_i)^2,$$

where d is geodesic distance on M . The *intrinsic mean* of the points in A is defined as a minimum of ρ_A , that is,

$$\mu = \arg \min_{x \in M} \rho_A(x). \quad (2)$$

The properties of the intrinsic mean have been studied by Karcher [18], and Pennec [21] describes a gradient descent algorithm to compute the mean. Since the mean is given by the minimization problem (2), we must verify that such a minimum exists and is unique. Karcher shows that for a manifold with non-positive sectional curvature the mean is uniquely defined. In fact, the space $PD(n)$ does have non-positive sectional curvature, and, thus, the mean is uniquely defined. Also, the gradient of ρ_A is given by

$$\nabla \rho_A(x) = -\frac{1}{N} \sum_{i=1}^N \text{Log}_x(x_i)$$

Thus the intrinsic mean of a collection of diffusion tensors is computed by the following gradient descent algorithm:

Algorithm 3: Intrinsic Mean of Diffusion Tensors

Input: $p_1, \dots, p_N \in PD(n)$

Output: $\mu \in PD(n)$, the intrinsic mean

```

 $\mu_0 = p_1$ 
 $\tau = 1$ , the initial step size
Do
   $X_i = \frac{1}{N} \sum_{k=1}^N \text{Log}_{\mu_i}(p_k)$ 
   $\mu_{i+1} = \text{Exp}_{\mu_i}(\tau X_i)$ 
  If  $\|X_i\| > \|X_{i-1}\|$ 
     $\tau = \tau/2$ ,  $X_i = X_{i-1}$ 
While  $\|X_i\| > \epsilon$ .
```

The parameter τ determines the step size to take in the downhill gradient direction. The gradient descent is guaranteed to converge for *some* value of τ (see [18] for details). For typical data found in DT-MRI the algorithm converges with $\tau = 1$. However, if the tensors being averaged are widely dispersed, smaller step sizes are needed. Algorithm 3 ensures convergence by dynamically adjusting the step size τ if the gradient descent ever starts to diverge.

4.2 Principal Geodesic Analysis

Principal component analysis (PCA) is a useful method for describing the variability of Euclidean data. In our previous work [13] we introduced *principal geodesic analysis* (PGA) as a generalization of PCA to study the variability of data in a Lie group. In this section we review the method of principal geodesic analysis and apply it to the symmetric space of diffusion tensors. We begin with a review of PCA in Euclidean space. Consider a set of points $x_1, \dots, x_N \in \mathbb{R}^d$ with zero mean. Principal component analysis seeks a sequence of linear subspaces that best represent the variability of the data. To be more precise, the intent is to find an orthonormal basis $\{v_1, \dots, v_d\}$ of \mathbb{R}^d , which satisfies the recursive relationship

$$v_1 = \arg \max_{\|v\|=1} \sum_{i=1}^N \langle v, x_i \rangle^2, \quad (3)$$

$$v_k = \arg \max_{\|v\|=1} \sum_{i=1}^N \sum_{j=1}^{k-1} \langle v_j, x_i \rangle^2 + \langle v, x_i \rangle^2. \quad (4)$$

In other words, the subspace $V_k = \text{span}(\{v_1, \dots, v_k\})$ is the k -dimensional subspace that maximizes the variance of the data projected to that subspace. The basis $\{v_k\}$ is computed as the set of eigenvectors of the sample covariance matrix of the data.

Now turning to manifolds, consider a set of points p_1, \dots, p_N on a Riemannian manifold M . Our goal is to describe the variability of the p_i in a way that is analogous to PCA. Thus we will project the data onto lower-dimensional subspaces that best represent the variability of the data. This requires first extending three important concepts of PCA into the manifold setting:

- **Variance.** Following the work of Fréchet, we define the sample variance of the data as the expected value of the squared Riemannian distance from the mean.
- **Geodesic subspaces.** The lower-dimensional subspaces in PCA are linear subspaces. For manifolds we extend the concept of a linear subspace to that of a *geodesic submanifold*.
- **Projection.** In PCA the data is projected onto linear subspaces. We define a projection operator for geodesic submanifolds, and show how it may be efficiently approximated.

We now develop each of these concepts in detail.

4.2.1 Variance

The variance σ^2 of a real-valued random variable x with mean μ is given by the formula $\sigma^2 = \mathcal{E}[(x - \mu)^2]$, where \mathcal{E} denotes expectation. It measures the expected localization of the variable x about the mean. The definition of variance we use comes from Fréchet [15], who defines the variance of a random variable in a metric space as the expected value of the squared distance from the mean. That is, for a random variable x in a metric space with intrinsic mean μ , the variance is given by $\sigma^2 = \mathcal{E}[d(\mu, x)^2]$. Thus in the manifold case, given data points $p_1, \dots, p_N \in M$ with mean μ , we define the sample variance of the data as

$$\sigma^2 = \sum_{i=1}^N d(\mu, p_i)^2 = \sum_{i=1}^N \|\text{Log}_\mu(p_i)\|^2. \quad (5)$$

Notice that if M is \mathbb{R}^n , then the variance definition in (5) is given by the trace of the sample covariance matrix, i.e., the sum of its eigenvalues. It is in this sense that this definition captures the total variation of the data.

4.2.2 Geodesic Submanifolds

The next step in generalizing PCA to manifolds is to generalize the notion of a linear subspace. A geodesic is a curve that is locally the shortest path between points. In this way a geodesic is the generalization of a straight line. Thus it is natural to use a geodesic curve as the one-dimensional subspace, i.e., the analog of the first principal direction in PCA.

In general if N is a submanifold of a manifold M , geodesics of N are not necessarily geodesics of M . For instance the sphere S^2 is a submanifold of \mathbb{R}^3 , but its geodesics are great circles, while geodesics of \mathbb{R}^3 are straight lines. A submanifold H of M is said to be *geodesic at* $x \in H$ if all geodesics of H passing through x are also geodesics of M . For example, a linear subspace of \mathbb{R}^d is a submanifold geodesic at 0. Submanifolds geodesic at x preserve distances to x . This is an essential property for PGA because variance is defined by squared distance to the mean. Thus submanifolds geodesic at the mean will be the generalization of the linear subspaces of PCA.

4.2.3 Projection

The projection of a point $x \in M$ onto a geodesic submanifold H of M is defined as the point on H that is nearest to x in Riemannian distance. Thus we define the projection operator $\pi_H : M \rightarrow H$ as $\pi_H(x) = \arg \min_{y \in H} d(x, y)^2$. Since projection is defined by a minimization, there is no guarantee that the projection of a point exists or that it is unique. However, because $PD(n)$ has non-positive curvature and no conjugate points, projection onto geodesic submanifolds is unique in this case.

Projection onto a geodesic submanifold at μ can be approximated in the tangent space to the mean, $T_\mu M$. If v_1, \dots, v_k is an orthonormal basis for $T_\mu H$, then the

projection operator can be approximated by the formula

$$\text{Log}_\mu(\pi_H(x)) \approx \sum_{i=1}^k \langle v_i, \text{Log}_\mu(x) \rangle_\mu. \quad (6)$$

As with any tangent-space approximation, the error in projection operator approximation gets larger the farther away you get from the mean.

4.3 Computing Principal Geodesic Analysis

We are now ready to define principal geodesic analysis for data p_1, \dots, p_N on a connected Riemannian manifold M . Our goal, analogous to PCA, is to find a sequence of nested geodesic submanifolds that maximize the projected variance of the data. These submanifolds are called the *principal geodesic submanifolds*.

The principal geodesic submanifolds are defined by first constructing an orthonormal basis of tangent vectors v_1, \dots, v_d that span the tangent space $T_\mu M$. These vectors are then used to form a sequence of nested subspaces $V_k = \text{span}(\{v_1, \dots, v_k\})$. The principal geodesic submanifolds are the images of the V_k under the exponential map: $H_k = \text{Exp}_\mu(V_k)$. The first principal direction is chosen to maximize the projected variance along the corresponding geodesic:

$$v_1 = \arg \max_{\|v\|=1} \sum_{i=1}^N \|\text{Log}_\mu(\pi_H(p_i))\|^2, \quad (7)$$

where $H = \exp(\text{span}(\{v\}))$.

The remaining principal directions are then defined recursively as

$$v_k = \arg \max_{\|v\|=1} \sum_{i=1}^N \|\text{Log}_\mu(\pi_H(p_i))\|^2, \quad (8)$$

where $H = \exp(\text{span}(\{v_1, \dots, v_{k-1}, v\}))$.

If we use (6) to approximate the projection operator π_H in (7) and (8), we get

$$v_1 \approx \arg \max_{\|v\|=1} \sum_{i=1}^N \langle v, \text{Log}_\mu(p_i) \rangle_\mu^2,$$

$$v_k \approx \arg \max_{\|v\|=1} \sum_{i=1}^N \sum_{j=1}^{k-1} \langle v_j, \text{Log}_\mu(p_i) \rangle_\mu^2 + \langle v, \text{Log}_\mu(p_i) \rangle_\mu^2.$$

The above minimization problem is simply the standard principal component analysis in $T_\mu M$ of the vectors $\text{Log}_\mu(p_i)$, which can be seen by comparing the approximations above to the PCA equations, (3) and (4). Applying these ideas to $PD(n)$, we have the following algorithm for approximating the PGA of diffusion tensor data:

Algorithm 4: PGA of Diffusion TensorsInput: $p_1, \dots, p_N \in PD(n)$ Output: Principal directions, $v_k \in \text{Sym}(n)$, variances, $\lambda_k \in \mathbb{R}$ $\mu =$ intrinsic mean of $\{p_i\}$ (Algorithm 3) $x_i = \text{Log}_\mu(p_i)$ $\mathbf{S} = \frac{1}{N} \sum_{i=1}^N x_i x_i^T$ (treating the x_i as column vectors) $\{v_k, \lambda_k\} =$ eigenvectors/eigenvalues of \mathbf{S} .

A new diffusion tensor p can now be generated from the PGA by the formula $p = \text{Exp}_\mu \left(\sum_{k=1}^d \alpha_k v_k \right)$, where the $\alpha_k \in \mathbb{R}$ are the coefficients of the modes of variation.

The use of the sample covariance matrix S in the tangent space to the mean leads to an obvious definition for a ‘‘Gaussian’’ distribution in $PD(n)$. This probability density function for this distribution is given by

$$p(x) = \exp \left(-\frac{1}{2} \text{Log}_\mu(x)^T \Sigma^{-1} \text{Log}_\mu(x) \right),$$

where μ is an element of $PD(n)$ and Σ is a covariance matrix in the tangent space $T_\mu PD(n)$. This distribution may be used to describe diffusion tensor data, with μ being the intrinsic mean and $\Sigma = S$ the sample covariance matrix as above. However, care should be taken in using this distribution as it does not have many of the desirable properties of Gaussian distributions in \mathbb{R}^n , such as the Central Limit Theorem. Alternatively, Gaussian distributions may be defined on Lie groups and homogeneous spaces as fundamental solutions to the heat equation, see [16] for details. These distributions do generalize properties such as the Central Limit Theorem, but they are typically in the form of infinite summations, and thus difficult to use in computations. Therefore, the tangent space Gaussian given above may be a useful, more computationally feasible, alternative. It should be noted that the definition of PGA is not dependent on any underlying assumption of the data distribution. In other words, PGA is a valid *descriptive* statistic for any sample of diffusion tensor data.

5 Properties of PGA on $PD(n)$

We now demonstrate that PGA on the symmetric space $PD(n)$ preserves certain important properties of the diffusion tensor data, namely the properties of positive-definiteness, determinant, and orientation¹. This makes the symmetric space formulation an attractive approach for the statistical analysis of diffusion tensor images. We have already mentioned that, in contrast to linear PCA, symmetric space PGA preserves positive-definiteness. That is, the principal geodesics are completely contained within $PD(n)$, and any matrix generated by the principal geodesics will be positive-definite.

¹The orientation of the tensor $p \in PD(n)$ is given by the rotation matrix u in the singular-value decomposition $p = u \Lambda u^T$, modulo rotations of 180 degrees about any axis. For 3D diffusion tensors this is the orientation frame of the ellipsoid representation of p .

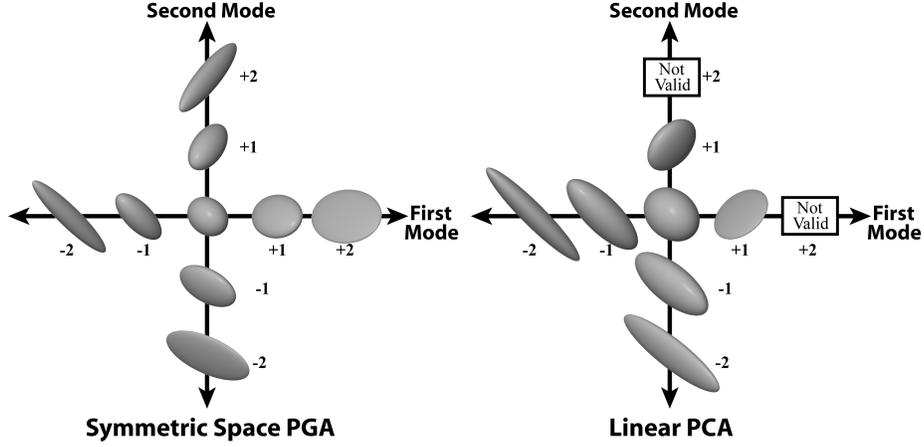


Figure 2: The first two modes of variation of the simulated data: (left) using the symmetric space PGA, and (right) using linear PCA. Units are in standard deviations. The boxes labelled “Not Valid” indicate that the tensor was not positive-definite, i.e., it had negative eigenvalues.

The next two properties we consider are the determinant and orientation. Consider a collection of diffusion tensors that all have the same determinant D . We wish to show that the resulting average and any tensor generated by the principal geodesic analysis will also have determinant D . To show this we first look at the subset of $PD(n)$ of matrices with determinant D , that is, the subset $P_D = \{p \in PD(n) : \det(p) = D\}$. This subset is a *totally geodesic submanifold*, meaning that any geodesic within P_D is a geodesic of the full space $PD(n)$. Notice the difference from the definition of a submanifold geodesic at a point; totally geodesic submanifolds are geodesic at *every* point. Now, the fact that P_D is totally geodesic implies that the averaging process in Algorithm 3 will remain in P_D if all the data lies in P_D . Also, the principal directions v_k in the PGA will lie in the tangent subspace $T_\mu P_D$. Thus any diffusion tensor generated by the principal geodesics will remain in the space P_D .

The same argument may be applied to show that symmetric space averaging and PGA preserve the orientation of diffusion tensors. In fact, the subset of all diffusion tensors having the same orientation is also a totally geodesic submanifold, and the same reasoning applies. Unlike the positive-definiteness and determinant, orientations are also preserved by linear averaging and PCA.

To demonstrate these properties, we simulated random 3D diffusion tensors and computed both their linear and symmetric space statistics. We first tested the determinant preservation by generating 100 random 3D diffusion tensors with determinant 1. To do this we first generated 100 random 3×3 symmetric matrices, with entries distributed according to a normal distribution, $N(0, \frac{1}{2})$. Then, we took the matrix exponential of these random symmetric matrices, thus making them positive-definite diffusion tensors. Finally, we normalized the random diffusion tensors to have determi-

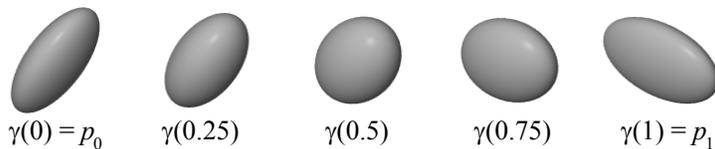


Figure 3: An example of geodesic interpolation of two diffusion tensors. The diffusion tensors at times 0.25, 0.5, and 0.75 were generated along the unique geodesic segment γ between two random tensors p_0 and p_1 .

nant 1 by dividing each tensor by the cube root of its determinant. We then computed the linear average and PCA and symmetric space average and PGA of the simulated tensors. The results are shown in Fig. 2 as the diffusion tensors generated by the first two modes of variation. The linear PCA generated invalid diffusion tensors, i.e., tensors with negative eigenvalues, at ± 2 standard deviations in both the first and second modes. All of the diffusion tensors generated by the symmetric space PGA have determinant 1. The linear mean demonstrates the “swelling” effect of linear averaging. It has determinant 2.70, and the linear PCA tensors within ± 2 standard deviations have determinants ranging from -2.80 to 2.82 . The negative determinants came from the tensors that were not positive-definite. Therefore, we see that the symmetric space PGA has preserved the positive-definiteness and the determinant, while the linear PCA has preserved neither.

Next we tested the orientation preservation by generating 100 random, axis-aligned, 3D diffusion tensors. This was done by generating 3 random eigenvalues for each matrix, corresponding to the x , y , and z axes. The eigenvalues were chosen from a log-normal distribution with log mean 0 and log standard deviation 0.5. Next we generated a random orientation $u \in SO(3)$ and applied it to all of the axis-aligned matrices by the map $p \mapsto upu^T$. Thus each of the diffusion tensors in our test set had eigenvectors equal to the columns of the rotation matrix u . We computed both the symmetric space and linear statistics of the data. As was expected, both methods preserved the orientations. However, the linear PCA again generated tensors that were not positive-definite.

6 Diffusion Tensor Interpolation

The most basic method for resampling a warped image is a nearest neighbor approach. Another possibility is to use trilinear interpolation of the linear tensor coefficients. The tensor interpolation method that we propose is based on the symmetric space averaging method developed in Section 4.1. First, consider the case of two diffusion tensors $p_1, p_2 \in PD(n)$. We would like an interpolation method given by a continuous curve $c : [0, 1] \rightarrow PD(n)$ satisfying $c(0) = p_1$ and $c(1) = p_2$. Given the symmetric space formulation for $PD(n)$ presented above, an obvious choice for c is the unique geodesic curve segment between p_1 and p_2 . This geodesic interpolation is demonstrated between two randomly chosen diffusion tensors in Fig. 3. Geodesic interpolation can be seen as a direct generalization of linear interpolation for scalar or vector data.

Now for 3D images of diffusion tensors an interpolation method can be thought of as a smooth function in a cube, where the tensor values to be interpolated are given at the corners of the cube. In other words, we want a smooth function $f : [0, 1]^3 \rightarrow PD(n)$, where the values $f(i, j, k) : i, j, k \in \{0, 1\}$ are specified. It is tempting to first create f using “tri-geodesic” interpolation, that is, by repeated geodesic interpolation in the three coordinate directions. However, unlike linear interpolation, geodesic interpolation of diffusion tensors does not commute. Therefore, a “tri-geodesic” interpolation would be dependent on the order in which the coordinate interpolations were made. A better method for interpolating diffusion tensors in three dimensions is using a weighted geodesic average.

Weighted averaging of data on a sphere S^n has been studied by Buss and Fillmore [7]. We follow their approach, extending the definition of weighted averages to diffusion tensors. Given a set of diffusion tensors $p_1, \dots, p_N \in PD(n)$ and a set of weights $w_1, \dots, w_N \in \mathbb{R}$, consider the weighted sum-of-squared distances function

$$\rho(p; p_1, \dots, p_N; w_1, \dots, w_N) = \sum_{i=1}^N w_i d(p, p_i)^2.$$

Given a set of non-negative real weights w_1, \dots, w_N with sum equal to 1, the *weighted average* of the p_i with respect to the weights w_i is defined as a minimum of the weighted sum-of-squared distances function, i.e.,

$$\text{Avg}(p_1, \dots, p_N; w_1, \dots, w_N) = \arg \min_{p \in PD(n)} \rho(p; p_1, \dots, p_N; w_1, \dots, w_N). \quad (9)$$

The intrinsic mean definition given in Section 4.1 is equivalent to weighted average definition with all weights set to $w_i = (1/N)$. For vector-valued data $v_1, \dots, v_N \in \mathbb{R}^n$ the weighted average is given by the weighted sum $\text{Avg}(\{v_i\}; \{w_i\}) = \sum_{i=1}^N w_i v_i$.

For diffusion tensor data the weighted average can be computed using a generalization of the intrinsic mean algorithm (Algorithm 3). The gradient of the sum-of-squared distances function is given by $\nabla \rho(p; \{p_i\}; \{w_i\}) = -\sum_{i=1}^N w_i \text{Log}_p(p_i)$. Therefore, the gradient descent algorithm for finding the weighted average of a set of diffusion tensors is given by

Algorithm 5: Weighted Average of Diffusion Tensors

Input: $p_1, \dots, p_N \in PD(n)$ and weights $w_1, \dots, w_N \in \mathbb{R}$

Output: $\mu \in PD(n)$, the weighted average

```

 $\mu_0 = I$ 
 $\tau = 1$ , the initial step size
Do
   $X_i = \sum_{k=1}^N w_k \text{Log}_{\mu_i}(p_k)$ 
   $\mu_{i+1} = \text{Exp}_{\tau \mu_i}(X_i)$ 
  If  $\|X_i\| > \|X_{i-1}\|$ 
     $\tau = \tau/2$ ,  $X_i = X_{i-1}$ 
While  $\|X_i\| > \epsilon$ .
```

Returning to the problem of finding an interpolating function for diffusion tensors in a volume image, we want to define our interpolating function $f : [0, 1]^3 \rightarrow PD(n)$, where the values at the corners are given. Let $A = \{0, 1\}^3$, and let $\alpha = (\alpha_1, \alpha_2, \alpha_3) \in A$ be a multi-index for the eight corners of the unit cube. Let $p_\alpha \in PD(n)$ be a set of diffusion tensors given at the corners of the unit cube. We define the *geodesic weighted interpolation* of the p_α as the function $f : [0, 1]^3 \rightarrow PD(n)$ via a weighted average

$$f(x_1, x_2, x_3) = \text{Avg}(\{p_\alpha\}, \{w_\alpha(x_1, x_2, x_3)\}) \quad (10)$$

where the $w_\alpha : [0, 1]^3 \rightarrow \mathbb{R}$ are weight functions on the unit cube. For example, we may choose the tri-linear weights $w_\alpha(x_1, x_2, x_3) = \prod_{i=1}^3 (1 - \alpha_i + (-1)^{1-\alpha_i} x_i)$. Higher-order weights, e.g., cubic polynomials, are also possible. As is the case with scalar images, higher-order schemes would have higher-order smoothness across voxel boundaries and would require larger areas of influence. Investigation of these higher-order schemes is an area of future work.

The interpolation function $f : [0, 1]^3 \rightarrow PD(n)$ given by (10) is a C^∞ function. The proof of this fact is a direct application of the Implicit Function Theorem and can be found in [11]. The weighted geodesic interpolation function is well-defined for any initial diffusion tensor values p_α , and it does not depend on any arbitrary choice of ordering as did the “tri-geodesic” method. Another important property of weighted geodesic interpolation is that it preserves determinants and orientations of the initial data. That is, if the p_α all have the same determinant (respectively, orientation), then any tensor interpolated by (10) will also have the same determinant (orientation). This follows from the same argument given in the previous section to show that the intrinsic mean preserves these properties. That is, if the data lie in the same totally geodesic submanifold (the submanifold representing diffusion tensors with the same determinant or the same orientation), the weighted average of the data will lie in the same submanifold. Since the weighted geodesic interpolation is defined via weighted averages, it follows that it also preserves determinants and orientations.

An example of the weighted geodesic interpolation is shown in Figure 4 in a region of a coronal slice in a DTI image of the brain. It can be seen that the interpolated image is smooth and does not suffer from swelling of the tensor determinants. Weighted geodesic interpolation does take about 20 times as long as a simple trilinear interpolation of the six tensor components. For applications such as registration for statistical group comparisons this extra computational cost is not critical.

7 Geodesic Anisotropy Measure

We now develop a new anisotropy measure for diffusion tensors based on the geodesic distance on the symmetric space $PD(3)$. 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. It turns out that the nearest isotropic diffusion

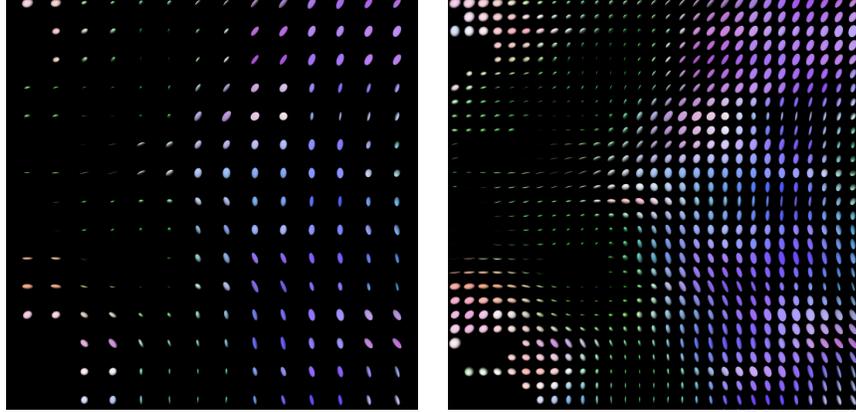


Figure 4: An example of geodesic interpolation from a coronal DTI slice. On the left is the original data, and on the right is the data up-sampled by two using geodesic weighted interpolation.

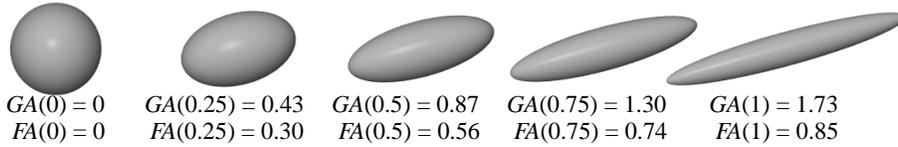


Figure 5: Comparison of GA and FA values (as a function of t) for tensors with eigenvalues $\lambda_1 = \exp(t)$, $\lambda_2 = \lambda_3 = \exp(-t)$.

tensor to p is the one with the same determinant as p , i.e., the matrix $\det(p)^{\frac{1}{3}} \cdot I_3$. Thus we define the *geodesic anisotropy* as

$$GA(p) = d(\det(p)^{\frac{1}{3}} \cdot I_3, p) = \left(\sum_{i=1}^3 \|\log(\lambda_i) - \overline{\log \lambda}\|^2 \right)^{\frac{1}{2}}, \quad (11)$$

where λ_i are the eigenvalues of p and $\overline{\log \lambda}$ is the average of the $\log \lambda_i$. This 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 $PD(n)$ as a symmetric space, where the eigenvalues are treated as multiplicative entities rather than linear ones.

Geodesic anisotropy, like FA, is invariant to scaling of a diffusion tensor. Unlike FA, which is in the range $[0, 1]$, the GA is unbounded and can take values in $[0, \infty)$. Figure 5 shows a comparison of FA and GA values of the one-parameter family of tensors with eigenvalues $\lambda_1 = \exp(t)$, $\lambda_2 = \lambda_3 = \exp(-t)$, for $t \in [0, 1]$. Figure 6

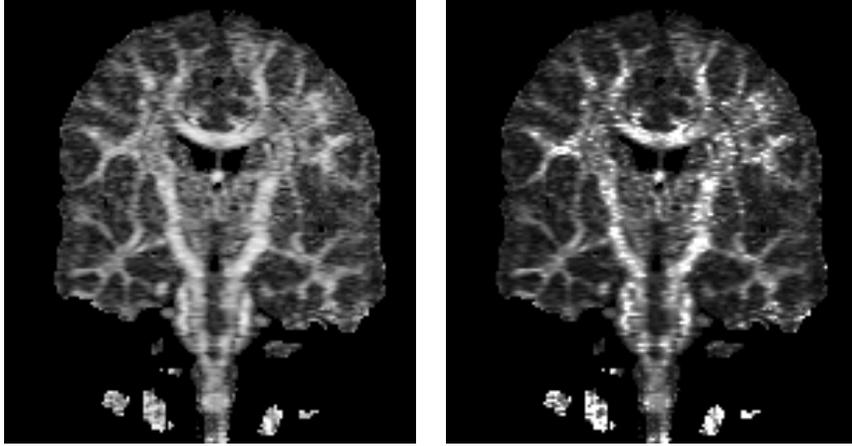


Figure 6: Comparison of a coronal image of FA values (left) with the corresponding image of GA values (right).

compares the FA and GA images from a coronal slice of a DT-MRI. The majority of the GA values fell in the range $[0, 2]$, so the image was scaled by $1/2$ for display in comparison to FA, which is in the range $[0, 1]$. Notice that the GA image has higher contrast, due to the nonlinear stretching of the image range compared to FA. It should be noted that there is not a functional relationship between FA and GA, i.e., there is no real-valued function f , such that $f(FA(p)) = GA(p)$, for all tensors $p \in PD(3)$. In other words, FA and GA differ non-trivially as functions of the eigenvalues. It would be expected then that group comparisons of DT-MRI data based on GA would give different results than tests using FA. Further research is needed to evaluate the differences between the two measures.

8 Conclusion

We have introduced a new geometric framework for computations and statistics of diffusion tensor data. This framework is based on the treatment of the space of positive-definite, symmetric tensors as a Riemannian symmetric space. The advantages of this formulation are that it inherently preserves the positive-definiteness of the tensors and also naturally handles properties such as the determinant and orientation. We developed new methods for computing means and second-order statistics of diffusion tensors and showed that these statistics preserve positive-definiteness that cannot be preserved using linear statistics. We used the Riemannian geometry to define 3D interpolation of diffusion tensor images and to define a new measure called geodesic anisotropy.

Further research in this area may extend these ideas to other applications where positive-definite, symmetric tensors arise. Examples include strain tensors of deformations in mechanics, and the structure tensor arising in optical flow and texture analysis

in computer vision. In DT-MRI applications we plan to incorporate the Riemannian statistics in building statistical atlases of diffusion tensor images. The interpolation method presented in this paper should prove useful in registration of DT-MRI and in fiber tracking applications. Also, we are investigating the use of geodesic anisotropy as a measure in group comparisons of DT-MRI.

Acknowledgments

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>. This work was also supported by NIH grants P01 CA47982 and EB02779. Brain dataset courtesy of Gordon Kindlmann at the Scientific Computing and Imaging Institute, University of Utah, and Andrew Alexander, W. M. Keck Laboratory for Functional Brain Imaging and Behavior, University of Wisconsin-Madison. We would like to thank Dr. Guido Gerig and Dr. Stephen Pizer for useful discussions and suggestions.

References

- [1] D. C. Alexander, C. Pierpaoli, P. J. Basser, and J. C. Gee. Spatial transformations of diffusion tensor MR images. *IEEE Transactions on Medical Imaging*, 20(11):1131–1139, 2001.
- [2] P. J. Basser, J. Mattiello, and D. Le Bihan. MR diffusion tensor spectroscopy and imaging. *Biophysics Journal*, 66:259–267, 1994.
- [3] P. J. Basser and S. Pajevic. A normal distribution for tensor-valued random variables: applications to diffusion tensor MRI. *IEEE Transactions on Medical Imaging*, 22(7):785–794, 2003.
- [4] P. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A rigorous framework for diffusion tensor calculus. *Magnetic Resonance in Medicine*, 53:221–225, 2005.
- [5] D. Le Bihan, J.-F. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.
- [6] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, 2nd edition, 1986.
- [7] S. R. Buss and J. P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Transactions on Graphics*, 20(2):95–126, 2001.
- [8] C. Chéfd’hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Constrained flows of matrix-valued functions: Application to diffusion tensor regularization. In *European Conference on Computer Vision*, pages 251–265, 2002.

- [9] I. Corouge, P. T. Fletcher, S. Joshi, J. H. Gilmore, and G. Gerig. Fiber tract-oriented statistics for quantitative diffusion tensor MRI analysis. In *Proceedings Medical Image Computing and Computer Aided Intervention (MICCAI)*, pages 131–139, 2005.
- [10] O. Coulon, D. C. Alexander, and S. Arridge. Diffusion tensor magnetic resonance image regularization. *Medical Image Analysis*, 8(1):47–68, 2004.
- [11] P. T. Fletcher. *Statistical Variability in Nonlinear Spaces: Application to Shape Analysis and DT-MRI*. PhD thesis, University of North Carolina, August 2004.
- [12] P. T. Fletcher and S. Joshi. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In *Proceedings of Workshop on Computer Vision Approaches to Medical Image Analysis (CVAMIA)*, 2004.
- [13] P. T. Fletcher, C. Lu, and S. Joshi. Statistics of shape via principal geodesic analysis on Lie groups. In *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pages 95–101, 2003.
- [14] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging (to appear)*, 2004.
- [15] M. Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Ann. Inst. H. Poincaré*, (10):215–310, 1948.
- [16] U. Grenander. *Probabilities on Algebraic Structures*. John Wiley and Sons, 1963.
- [17] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. Academic Press, 1978.
- [18] H. Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Math*, 30(5):509–541, 1977.
- [19] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor mri. Technical report, INRIA, 2004.
- [20] S. Pajevic and P. J. Basser. Parametric and non-parametric statistical analysis of DT-MRI. *Journal of Magnetic Resonance*, 161(1):1–14, 2003.
- [21] X. Pennec. Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements. In *IEEE Workshop on Nonlinear Signal and Image Processing*, 1999.
- [22] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 61(1):41–66, 2006.
- [23] Z. Wang, B. C. Vemuri, Y. Chen, and T. Mareci. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from DWI. In *Information Processing in Medical Imaging*, pages 660–671, 2003.