

Diffusion Tensor Field Registration in the Presence of Uncertainty

Mustafa Okan Irfanoglu^{1,2}, Cheng Guan Koay², Sinisa Pajevic²,
Raghu Machiraju¹, and Peter J. Basser²

¹ The Ohio State University, Columbus OH 43210, USA

² National Institutes of Health, NICHD, Bethesda, MD 20892, USA

Abstract. We propose a novel method for deformable tensor-to-tensor registration of Diffusion Tensor Imaging (DTI) data. Our registration method considers estimated diffusion tensors as normally distributed random variables whose covariance matrices describe uncertainties in the mean estimated tensor due to factors such as noise in diffusion weighted images (DWIs), tissue diffusion properties, and experimental design. The dissimilarity between distributions of tensors in two different voxels is computed using the Kullback-Leibler divergence to drive a deformable registration process, which is not only affected by principal diffusivities and principal directions, but also the underlying DWI properties. We in general do not assume the positive definite nature of the tensor space given the pervasive influence of noise and other factors. Results indicate that the proposed metric weights voxels more heavily whose diffusion tensors are estimated with greater certainty and exhibit anisotropic diffusion behavior thus, intrinsically favoring coherent white matter regions whose tensors are estimated with high confidence.

1 Introduction

Accurate registration of tensor fields is of great relevance in various stages of Diffusion Tensor Imaging (DTI) analysis [1]. Because of the complex nature of DTI data, cross-registration of DTI population data needed for longitudinal and multi-site studies, and brain atlas creation requires specialized mathematical tools. An accurate tensor interpolation scheme and a tensor dissimilarity metric reflecting the tensor's principal diffusivities and directions and uncertainty due to noise are needed considering the large variability among DTI data and experimental designs. Early registration approaches used derived scalar fields such as Apparent Diffusion Coefficients (ADC), Fractional Anisotropy (FA), or individual tensor components. Next-generation registration methods operated on the tensor manifolds and employed a metric to compute tensor distances such as the Riemannian [2] or Log-Euclidean [3]. Zhang *et al.* proposed a locally affine registration algorithm based on diffusion profiles, as a function of spatial direction [4]. Another work is from Cao *et al.* where the authors realize the registration by optimizing for geodesics on the space of diffeomorphisms connecting two diffusion tensor images [5]. The use of full tensor information for registration, along

with metrics powerful enough to capture shape and direction information, has been shown to lead to better registration accuracy [6] [7]. However, all these approaches consider diffusion tensors as independent from the original DWIs. It is crucial to note that diffusion tensors are obtained through an optimization process on the DWIs and do not only reflect the underlying diffusion properties, but also depend on the noise in the DWIs and gradient information.

In this paper, we propose a method that uses a dissimilarity metric that not only makes use of the full estimated tensor data, but also uses the uncertainty present in typical clinical DWIs. This causes the registration to favor directionally more informative, more anisotropic and less noisy regions. To our knowledge, this property of diffusion tensors has never been investigated and employed in a registration procedure. For each voxel, a tensor-variate Gaussian distribution is constructed with a mean and a covariance matrix obtained from the tensor fitting function itself; the mean tensor provides the best estimate of the diffusion tensor while the covariance matrix quantifies the uncertainty of estimated mean diffusion tensors. The main contributions of this work are:

- using the uncertainty information present in DWIs in tensor distributions to help the registration automatically favor brain regions with high anisotropy and fiber coherence forming an internal skeleton that guides the registration.
- incorporating an initial segmentation for a tissue adaptive registration.
- providing analytically derived error differentials for faster convergence.

2 Registration Framework

The Kullback-Leibler (KL) divergence dissimilarity for tensor-variate Gaussian distributions is used as a voxelwise dissimilarity metric in a hierarchical registration framework that starts with a coarse, rigid registration, continues with affine, and finishes with a finely resolved B-splines deformable registration. A 6×6 covariance matrix is computed from the invariant Hessian of the non-linear tensor fitting function along with each mean estimated diffusion tensor to construct a Gaussian tensor-variate distribution. Figure 1 depicts the workflow of the proposed framework.

Positive Definiteness and Distributions of Diffusion Tensors. Diffusion tensors are predicted to have non-negative eigenvalues, representing the real molecular water diffusion. However, in DTI, the diffusion tensors are obtained through a physical setup not only affected by real water diffusion but also the scan parameters. This results in generally non positive-definite tensors in typical DTI scans (unconstrained fitting), especially in highly anisotropic regions such as Corpus Callosum. In his work, Pasternak et al. considers diffusion tensors as Cartesian physical quantities and shows that the Euclidean space is better suited for diffusion tensors than affine-invariant Riemannian manifolds [8]. Pajevic et al. also shows through Monte-Carlo simulations mimicking physical imaging setups, that tensor coefficients can be modelled with a Gaussian distribution over

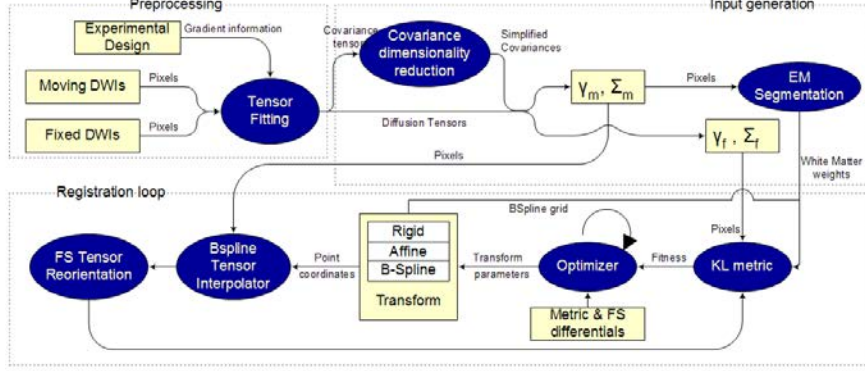


Fig. 1. The flow of the proposed registration algorithm

a wide range of SNR and the number of DWIs acquired. [9]. Aiming to cope with uncertainties such as noise and artifacts in *practical* settings, it was preferable to use a tensor-variate Gaussian distribution in our framework, instead of a Wishart distribution, which conserve positive definiteness.

3 Methodology

3.1 Tensor Fitting and Covariance Tensor Estimation

In a typical DTI experiment, the measured signal in a single voxel has the form [1], $s = S_0 \exp(-bg^T \mathbf{D}g)$, where the measured signal, s , depends on the diffusion encoding vector, g , the diffusion weight, b , the reference signal, S_0 , and the diffusion tensor \mathbf{D} . Given $n \geq 7$ sampled signals derived from six non-collinear gradient directions and at least one sampled reference signal, the diffusion tensor estimate can be found with non-linear regression with the following objective function $f_{NLS}(\gamma) = \frac{1}{2} \sum_{i=1}^n \left(s_i - \exp \left[\sum_{j=1}^7 \mathbf{W}_{ij} \gamma_j \right] \right)^2$. The symbol γ represents the vectorized version of diffusion tensor entries, s_i is the measured DW signal corrupted with noise, $\hat{s}_i(\gamma) = e^{\sum_{j=1}^7 (\mathbf{W}_{ij} \gamma_j)}$ is the predicted DW signal evaluated at γ , and \mathbf{W} the experimental design matrix is presented in [10].

The f_{NLS} function in Equation introduces the variability in the signal as explained in the design matrix, \mathbf{W} . In [11], it is shown that the diffusion tensors at each voxel can be considered as a normally distributed random variable with the covariance matrix being a function of the Hessian matrix at the optimum solution. Thus according to [11], the Hessian matrix can be computed as $\nabla^2 f_{NLS}(\gamma) = \mathbf{W}^T (\hat{\mathbf{S}}^2 - \mathbf{R}\hat{\mathbf{S}}) \mathbf{W}$, where \mathbf{S} and $\hat{\mathbf{S}}$ are diagonal matrices whose diagonal elements are the observed and the estimated DW signals, respectively, and $\mathbf{R} = \mathbf{S} - \hat{\mathbf{S}}$. Then, the covariance matrix of a diffusion tensor can be estimated as in [10]: $\Sigma_\gamma = \sigma_{DW}^2 [\nabla^2 f_{NLS}(\hat{\gamma})]^{-1}$, where σ_{DW}^2 represents the variance of the noise in the DWIs [10].

The covariance matrix is therefore a function of DWI noise, σ_{DW}^2 , the gradient magnitudes and directions (embodied in the design matrix, \mathbf{W}) and the tissue's

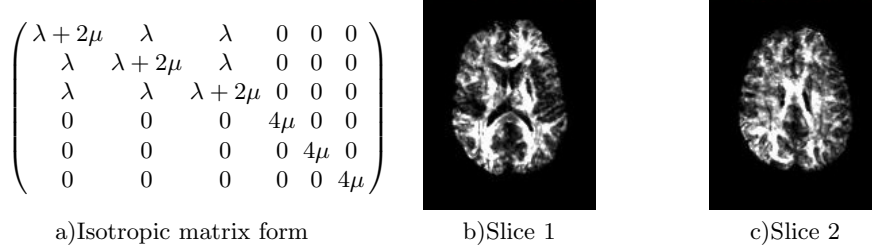


Fig. 2. Isotropic covariance structure in *a*) requires only 2 parameters. Deviations of the original covariance matrices from the isotropic form are displayed in *b*) and *c*). The images show that covariance matrices tend to be more anisotropic in WM regions.

underlying diffusion properties. The anisotropy, the norm, and the shape of this 6×6 matrix all provide insights on the reliability of the optimum diffusion tensor solution and the tissue properties. Figure 2 *b*) and *c*) display maps of the deviations from these matrices from the pure isotropic designs of Section 3.4, thus displaying a measure of the shape of these matrices w.r.t. the tissue type.

3.2 Dissimilarity Metric

In this work, we propose a new metric function, F , for diffusion tensor field registration. This metric uses the distribution of diffusion tensors obtained in each voxel, arising from noise and tissue properties. F is based on the symmetric KL divergence and can be described as:

$$F(I_f, I_m, \Theta) = \frac{1}{N} \sum_{p \in \Omega} w_p(I_f, I_m) \begin{pmatrix} \text{tr}(\Sigma_m^{-1} \Sigma_f) + (\gamma'_m - \gamma_f)^T \Sigma_m^{-1} (\gamma'_m - \gamma_f) \\ + \\ \text{tr}(\Sigma_f^{-1} \Sigma_m) + (\gamma_f - \gamma'_m)^T \Sigma_f^{-1} (\gamma_f - \gamma'_m) \end{pmatrix} \quad (1)$$

In Equation 1, γ_f signifies $\gamma_f(p)$, the diffusion tensor on the fixed image at a physical voxel location p ; similarly Σ_f signifies the covariance at voxel location p , i.e., $\Sigma_f(p)$, and Θ symbolizes the transformation parameters. For the moving image I_m , the covariance matrix is obtained through interpolation so Σ_m corresponds to $\Sigma_m(T(p, \Theta))$. Interpolation is done through a continuous B-splines approximation framework [12]. Deforming a diffusion tensor, $\gamma_m(p)$, with a (locally) affine transformation matrix, \mathbf{A} , involves tensor interpolation followed by reorientation. In this work, we follow the Finite Strain model proposed in [13] then the interpolated and rotated diffusion tensor $\gamma'_m(p)$ can be found to be $\gamma'_m(p) = \mathbf{R}^T \gamma_m(T(p, \Theta)) \mathbf{R}$. \mathbf{R} is the rotation component extracted from the affine matrix, \mathbf{A} , and can be found to be $\mathbf{R} = (\mathbf{A} \mathbf{A}^T)^{-1/2} \mathbf{A}$. For the elastic registration case, \mathbf{A} is not constant throughout the image and can be locally estimated from the displacement field, u , as $\mathbf{A}(p) = \mathbf{I} + \mathbf{J}(u(p))$ where \mathbf{I} is the identity matrix, and $\mathbf{J}(u(p))$ is the Jacobian of the deformation field at p .

Equation 1 is the Kullback-Leibler (KL) divergence symmetrized with respect to both distributions. When the first part of the equation is examined, $\frac{1}{2}(\text{tr}(\Sigma_m^{-1} \Sigma_f) + (\gamma'_m - \gamma_f)^T \Sigma_m^{-1} (\gamma'_m - \gamma_f))$, it can be seen that the first term

in the summation, $tr(\Sigma_m^{-1}\Sigma_f)$, measures the similarities between the two covariance matrices; the second term is the standard Mahalanobis distance. The overall metric for the registration is the weighted ($w_p(I_f, I_m)$) summation over the KL metrics on all voxels, normalized by the number of voxels used.

3.3 Error Metric Differentials

Registration is mainly an optimization procedure, where the optimizers generally require partial differentials of the error metric with respect to the transformation parameters. Most of the DTI registration frameworks suffer from using numerical approximations to these gradients [14], such as centered differences. The problem with this approach is that it requires two metric computations per transform parameter. For deformable registrations with very large parameter space dimensionality, this approach is infeasible and an analytical solution for the differential is required. In this section, we will analytically derive the error metric gradient so that each partial differential involved has a simple form and is easy to compute numerically. This way the metric evaluations are minimized and the gradient computations are more accurate and faster. Let us have a closer look at the first term of the error metric:

$$F = \frac{1}{2}(tr(\Sigma_m^{-1}(T(p, \Theta))\Sigma_f(p)) + (\gamma'_m(p) - \gamma_f(p))^T \Sigma_m(T(p, \Theta))^{-1}(\gamma'_m(p) - \gamma_f(p)))$$

Let f be the trace term, $f = tr(\Sigma_m^{-1}(T(p, \Theta))\Sigma_f(p))$, and g be the Mahalanobis term, $g = (\gamma'_m(p) - \gamma_f(p))^T \Sigma_m(T(p, \Theta))^{-1}(\gamma'_m(p) - \gamma_f(p))$. The differential can be expressed as $\partial F/\partial\Theta_i = \partial f/\partial\Theta_i + \partial g/\partial\Theta_i$. From the chain rule, it follows that:

$$\partial f/\partial\Theta_i = \sum_{j=1}^6 \sum_{k=1}^6 \frac{\partial tr(\Sigma_f \Sigma_m^{-1}(T(p, \Theta)))}{\partial \Sigma_m^{-1}\{kj\}} \sum_{x,y,z} \frac{\partial \Sigma_m^{-1}(T(p, \Theta))}{\partial T_{x,y,z}} \frac{\partial T_{x,y,z}}{\partial \Theta_i} \quad (2)$$

- The first differential term, $\frac{\partial tr(\Sigma_f \Sigma_m^{-1}(T(p, \Theta)))}{\partial \Sigma_m^{-1}\{kj\}}$, is just $\Sigma_f\{kj\}$ from the symmetry of covariance matrices and the derivative of traces w.r.t the matrices. Also note that the inverses of the covariance matrices are stored and used as images, cancelling the need for the inverse operation for the differential. Additionally, as explained in Section 3.4 the isotropic covariance matrix $\Sigma_m^{-1}(T(p, \Theta))$ is obtained only using interpolation but not reorientation due to rotational invariance assumption, yielding a simpler formula [9].

The second partial in Equation 2 represents the image gradient of the maps of each covariance components w.r.t. imaging directions. These gradients need to be computed once at the beginning of the registration.

- The last term $\frac{\partial T_{x,y,z}}{\partial \Theta_i}$ corresponds to the Jacobian of the transformation and needs to be computed once per iteration.

The Mahalanobis part of the function F , i.e., the function g , has a more complicated differential due to the rotation of diffusion tensors $\gamma_m(T(p, \Theta))$ into $\gamma'(p)$ if an affine or deformable registration scheme is employed. Let a be $a = \gamma'_m - \gamma_f$,

then the Mahalanobis part g can be rewritten as $g = \sum_j \sum_k a_j a_k \Sigma_m^{-1}\{jk\}$. Then the differential can be rewritten as:

$$\frac{\partial g}{\partial \Theta_i} = \sum_j \sum_k \frac{\partial a_j}{\partial \Theta_i} a_k \Sigma_m^{-1}\{jk\} + \sum_j \sum_k a_j \frac{\partial a_k}{\partial \Theta_i} \Sigma_m^{-1}\{jk\} + \sum_j \sum_k a_j a_k \frac{\partial \Sigma_m^{-1}\{jk\}}{\partial \Theta_i}$$

The differential in the last term, $\frac{\partial \Sigma_m^{-1}\{jk\}}{\partial \Theta_i}$ is the same as the one used in Equation 2, i.e., $\sum_{x,y,z} \frac{\partial \Sigma_m^{-1}(T(p,\Theta))}{\partial T_{x,y,z}} \frac{\partial T_{x,y,z}}{\partial \Theta_i}$. With the finite strain model, a can be described as $a = R^T \gamma_m(T(p,\Theta)) R - \gamma_f$. Then the first differential becomes:

$$\frac{\partial a_i}{\partial \Theta_z} = \frac{\partial R^T}{\partial \Theta_z} \gamma_m(T(p,\Theta)) R + R^T \frac{\partial \gamma_m(T(p,\Theta))}{\partial \Theta_z} R + R \gamma_m(T(p,\Theta)) \frac{\partial R}{\partial \Theta_z} \quad (3)$$

The second partial in Equation 3 can be found similarly to the covariance matrix case and is $\sum_{x,y,z} \frac{\partial \gamma_m(T(p,\Theta))}{\partial T_{x,y,z}} \frac{\partial T_{x,y,z}}{\partial \Theta_i}$. In the case of an affine transformation, where the parameters Θ_i corresponds to the entries in the affine matrix, A , the partial derivative of the rotation matrix, R , with respect to the transformation parameter, Θ_z , comes from the chain rule, $\frac{\partial R}{\partial \Theta} = \frac{\partial R}{\partial A}$. For the elastic registration case, the local affine matrix is estimated from the displacement field, u , and the differential becomes $\frac{\partial R}{\partial \Theta_z} = \sum_j \sum_k \frac{\partial R}{\partial u_{jk}} \frac{\partial u_{jk}}{\partial \Theta_z}$. For B-splines registration of order 3, the displacement field can be written as, $u(p,\beta) = \sum_i \sum_j \sum_k \beta_{ijk} b_{i,3}(p_x) b_{j,3}(p_y) b_{k,3}(p_z)$, where β_{ijk} are B-splines weights corresponding to parameters Θ and $b_{i,3}$ are 3^{rd} order spline basis functions. Then the second partial derivative, $\frac{\partial u_{jk}}{\partial \Theta_z}$, is just $b_{i,3}(p_x) b_{j,3}(p_y) b_{k,3}(p_z)$ for $\Theta_z = \beta_{ijk}$. The first term, $\frac{\partial R}{\partial u_{jk}}$, can be found in in [14].

3.4 Covariance Matrix Dimensionality Reduction

Independent components of diffusion tensors (6) and covariance matrices (21) generally yield a total of 27 dimensions, which poses problems in terms of memory and speed during registration. Being a function of the matrix W , the covariance matrix's form depends on the number of gradients and the direction of gradients used. In [9], it is shown that with sufficient number of diffusion gradients sampling the unit icosahedron densely enough, the $4D$ covariance tensor ($3 \times 3 \times 3 \times 3$) corresponding to the $2D$ covariance matrix tends to be isotropic and rotationally invariant. These isotropic covariance matrix yield a specific 6×6 matrix structure, with the block matrix form shown in Figure 2 a).

3.5 Tissue Segmentation

The tensor covariance matrix provides additional information on the tissue type. To further use this additional information for a more robust and faster registration, we perform a classical Expectation–Maximization (EM) segmentation initialized with K -means clustering, with the distance function originating from

our KL-metric and tensor-variate distributions derived from the mean and covariance matrices. This procedure is used in the registration initialization. For each moving image, first a segmentation is carried out. The probability of a voxel being a WM voxel obtained from the EM segmentation is used as the weighting factor $w_p(I_f, I_m)$ in Equation 1. Additionally, the segmentation labels are used to build a multi-level grid for B-splines registration. A coarser B-splines transformation grid is placed on CSF locations to decrease the computational complexity, whereas a denser grid is used for WM.

4 Experiments and Results

We acquired data from six healthy subjects with DTI parameters, $b=1000s/mm^2$, 72 diffusion gradient directions. Matrix sizes for all images were 128×157 with 114 axial slices and $1.5mm$ isotropic voxel resolution. One of the images was chosen to be the fixed image and the other five were used as moving images. For comparison, we implemented a benchmark multi-channel registration algorithm with *six* channels for tensor components, including one channel for FA and one channel for ADC. The benchmark method followed the same vector image registration steps. Standard deviation maps of the FA maps were also computed from the registered images. For the described dataset, the proposed registration pipeline with rigid, affine, and B-splines transformations (maximum grid size $20 \times 20 \times 20$), takes on the average 30 minutes per image on a modern computer.

4.1 Segmentation Outputs

We segmented a brain image by using three different levels of information: the isotropic covariance matrices (trace part of the error metric), only the full covariance matrices, and full covariance matrices along with the diffusion tensors as described by the error metric. Figure 3 displays the result of these segmentations. The images in Figure 3 show that the tensor covariance information brings additional information about the tissue type. With increasing complexity of the covariance matrix structure, tissue layers can better be discriminated. The use of full covariance matrix along with the diffusion tensor further improves the segmentation (Figure 3 (d)).

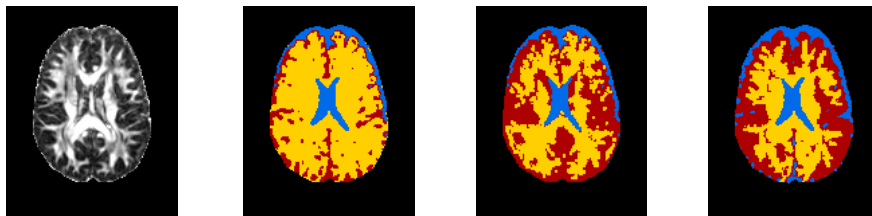


Fig. 3. Segmentation results. (a) Fa image. (b) Segmentation with only isotropic covariance matrix. (c) Segmentation using full covariance matrix. Segmentation of white matter improves using full covariance matrix and diffusion tensor information (d).

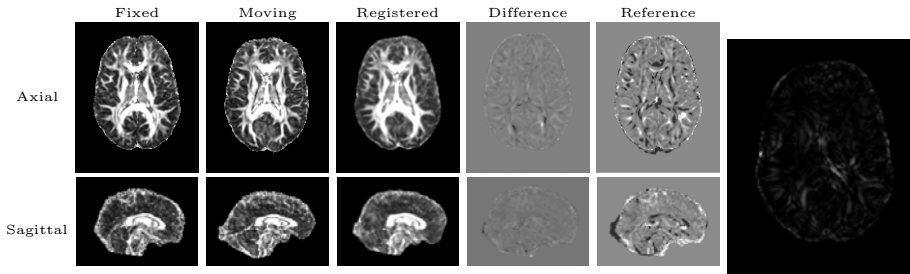


Fig. 4. Output of the registration algorithm. The fixed, moving, registered moving, difference images, and FA standard deviation displayed in different columns.

4.2 Registration Results

Figure 4 displays the output of the registration algorithm. The metric proves to perform well on white matter regions, as can be observed from the similarity of the images in the first and third columns. The difference image of the registered moving image and the fixed image is displayed on the fourth column, where it is visible that the metric performs significantly better than the benchmark method on the Corpus Callosum. The difference image for the benchmark method is displayed in the fifth column. The sixth column, displays the standard deviation of the FA maps of the five images registered with the proposed method. Note that the performance of the algorithm on white matter is clearly visible.

5 Discussions and Conclusions

In this work, we proposed a novel, robust and fast approach for tensor-to-tensor registration for Diffusion Tensor Image data, suitable for group analysis and tensor atlasing problems. The proposed metric captures the uncertainty of the diffusion tensors with a tensor-variate Gaussian distribution. Our future research directions include the analysis of the shape and isotropy characteristics of covariance matrices and testing of the algorithm with a larger population.

References

1. Basser, P.J., Jones, D.K.: Diffusion-tensor MRI: theory, experimental design and data analysis - a technical review. *NMR Biomed.* 15(7-8), 456–467 (2002)
2. Fletcher, P., Joshi, S.: Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In: Sonka, M., Kakadiaris, I.A., Kybic, J. (eds.) *CVAMIA/MMBIA 2004*. LNCS, vol. 3117, pp. 87–98. Springer, Heidelberg (2004)
3. Arsigny, V., Fillard, P., Pennec, X.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine* 56, 411–421 (2006)
4. Zhang, H., Yushkevich, P.: Deformable registration of diffusion tensor MR images with explicit orientation optimization. In: Duncan, J.S., Gerig, G. (eds.) *MICCAI 2005*. LNCS, vol. 3749, pp. 172–179. Springer, Heidelberg (2005)

5. Cao, Y., Miller, M., Mori, S., Winslow, R.L., Younes, L.: Diffeomorphic matching of diffusion tensor images. In: CVPR Workshop, vol. 22(17), pp. 67–67 (2006)
6. Jinzhong, Y., Dinggang, S., Ragini, V.: Diffusion tensor image registration using tensor geometry and orientation features. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI 2008, Part II. LNCS, vol. 5242, pp. 905–913. Springer, Heidelberg (2008)
7. Zhang, H., Avants, B.B., Yushkevich, P.A., Woo, J.H., Wang, S., McCluskey, L.F., Elman, L.B., Melhem, E.R., Gee, J.C.: High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: An example study using amyotrophic lateral sclerosis. *IEEE Trans. Med. Imaging* 26(11), 1585–1597 (2007)
8. Pasternak, O., Verma, R., Sochen, N., Basser, P.J.: On what manifold do diffusion tensors live. In: MICCAI Workshop (2008)
9. Basser, P.J., Pajevic, S.: A normal distribution for tensor-valued random variables: applications to diffusion tensor mri. *IEEE TMI* 202, 785–794 (2003)
10. Koay, C.G., Chang, L., Pierpaoli, C., Basser, P.J.: Error propagation framework for diffusion tensor imaging via diffusion tensor representations. *IEEE Transactions on Medical Imaging* 26(8), 1017–1034 (2007)
11. Koay, C.G., Chang, L., Pierpaoli, C., Basser, P.J.: A unified theoretical and algorithmic framework for least square methods of estimation in diffusion tensor imaging. *Journal of Magnetic Resonance* 182, 115–125 (2006)
12. Pajevic, S., Aldroubi, A., Basser, P.J.: Continuous tensor field approximation of diffusion tensor mri data. *Visualization and Processing of Tensor Fields* (2006)
13. Alexander, D.C., Pierpaoli, C., Basser, P.J., Gee, J.C.: Spatial transformations of diffusion tensor magnetic resonance images. *TMI* 20(11), 1131–1139 (2001)
14. Yeo, B., Vercauteren, T., Fillard, P., Pennec, X., Ayache, N., Clatz, O.: DTI registration with exact finite-strain differential. In: *IEEE ISBI* (2008)