

Diffusion-Based Population Statistics Using Tract Probability Maps

Demian Wassermann¹, Efstathios Kanterakis², Ruben C. Gur³,
Rachid Deriche¹, and Ragini Verma²

¹ Athena project-team, INRIA Sophia Antipolis-Mediterran  e, 2004 rt des Lucioles, 06902, FR

² Section of Biomedical Image Analysis, Radiology, UPENN, PA 19104, USA

³ Department of Psychiatry, University of Pennsylvania Medical Center, PA 19104, USA

Abstract. We present a novel technique for the tract-based statistical analysis of diffusion imaging data. In our technique, we represent each white matter (WM) tract as a tract probability map (TPM): a function mapping a point to its probability of belonging to the tract. We start by automatically clustering the tracts identified in the brain via tractography into TPMs using a novel Gaussian process framework. Then, each tract is modeled by the skeleton of its TPM, a medial representation with a tubular or sheet-like geometry. The appropriate geometry for each tract is implicitly inferred from the data instead of being selected *a priori*, as is done by current tract-specific approaches. The TPM representation makes it possible to average diffusion imaging based features along directions locally perpendicular to the skeleton of each WM tract, increasing the sensitivity and specificity of statistical analyses on the WM. Our framework therefore facilitates the automated analysis of WM tract bundles, and enables the quantification and visualization of tract-based statistical differences between groups. We have demonstrated the applicability of our framework by studying WM differences between 34 schizophrenia patients and 24 healthy controls.

1 Introduction

Diffusion tensor imaging (DTI) has become of modality of choice for studying the white matter (WM) pathology due to its unique *in vivo* characterization of WM microstructure [1]. A wide range of paradigms have been developed for studying group differences between healthy subjects and a population with pathology using DTI. Of these, the tract-based spatial statistics (TBSS) [2] made a break from voxel-wise techniques by performing statistics on the skeletonization of the full WM¹. This technique has rapidly gained popularity for finding group differences in WM. However, TBSS relies on a whole-brain approach disregarding anatomical specificity. Recently, techniques tried to alleviate this problem by analysing WM tracts individually by representing each tract as a lower-dimensional structure, using tubular geometries [3,4,5] or sheet-like parametrizations [6,7]. However, there is an issue in these tract-specific techniques: the need to *a priori* choose between a tubular or a sheet-like representation of the WM tracts which introduces an artificial bias as some fasciculi as inferior-fronto

¹ In this work we follow the same definition of skeleton as in Smith et al. [2].

occipital have a spreading structure which is not accurately represented by a tube and the cingulum are tube-like and not well represented by a sheet.

In this paper, we propose a technique to perform tract-specific statistical analyses overcoming the previously mentioned limitations. For this, we model each WM tract as a function mapping any point in space to the probability of its belonging to such a tract. This model is called a tract probability map (TPM) and can be robustly calculated using a recently proposed mathematical framework for WM bundles based on Gaussian processes (GPs) [8]. Within this framework, we develop a skeleton-based statistical analysis [2,6,7] technique for finding differences among populations. In developing this technique, we present two contributions: first, a general approach for tract-based statistical analysis which is not bound to a sheet-or-tube representation; and second, the applicability of the mathematical framework for the segmentation of WM structures presented by [8] for statistical analysis.

2 Computing Tract Probability Maps

The first step in developing our statistical analysis technique is the calculation of tract probability maps for a set of WM tracts obtained from whole brain DTI-tractography [9,8]. DTI-tractography is a technique to trace pathways followed by axons in the WM, representing them as smooth three dimensional trajectories at super-voxel resolution [1]. These trajectories provide the means to work on elements which are more anatomically-oriented than isolated voxels and can be grouped into WM tracts like the cingulum or the fornix. We show examples of TPMs for WM tracts in fig. 1.

In order to calculate TPMs for WM tracts from these trajectories while preserving super-voxel resolution, we use a GP framework [8]. The advantages of this representation are two-fold: first, the TPMs calculated using GPs are continuous functions which can be sampled at any desired resolution; second, being a parametrical representation the GP framework allows us to work robustly on its parameter space instead of performing operations in image space. In the remainder of this section we detail the calculation of TPM for a WM tract: we start by describing the representation of a trajectory within

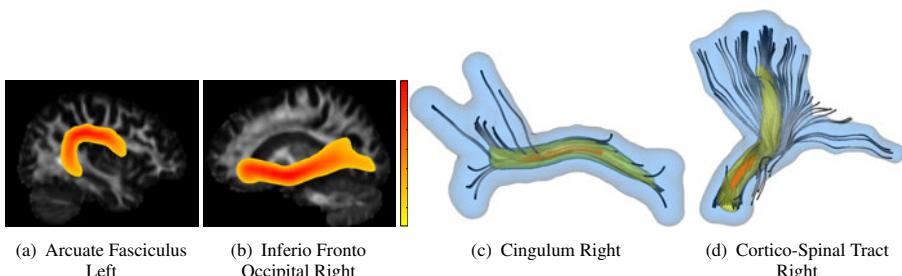


Fig. 1. Population-averaged tract probability maps for four WM tracts. These maps are shown over FA (a-b) and as iso-probability surfaces (c-d). The probability at each voxel is calculated using eq. (2). Color code for fig. (c-d) is as follows, Blue: $\text{TPM}(\mathbf{p}) = .01$, Yellow: $\text{TPM}(\mathbf{p}) = .2$, Red: $\text{TPM}(\mathbf{p}) = .6$.

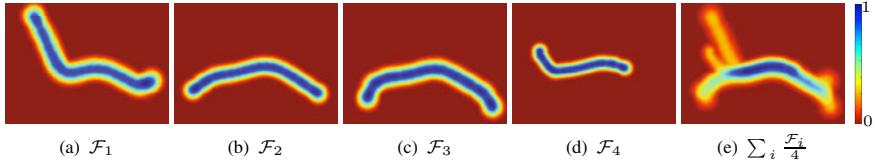


Fig. 2. Mean indicator function for four fiber tracts (a-d) and mean indicator function for the bundle formed by averaging them according to our framework (e). Color code ranges from blue when it is likely that a voxel belongs to the bundle of fibres to red when it does not belong.

the GP framework; next, we show how to go from the GP representation of a single trajectory to the representation of a bundle of trajectories; and finally, using the GP framework, we describe the calculation of the TPM for a bundle of trajectories.

We model each trajectory \mathcal{F} as a blurred indicator function $y(\cdot)$, such that if the three dimensional point \mathbf{p} belongs to \mathcal{F} , the value of $y(\mathbf{p})$ is 1 and it decays to 0 as \mathbf{p} is further away from \mathcal{F} . Several examples of blurred indicator for a trajectory are shown in fig. 2 (a-d). Formulating the hypothesis that the least curved a blurred indicator function is the better it represents a smooth trajectory, it is possible to model the indicator function $y(\cdot)$ for a trajectory \mathcal{F} as a GP [8]. Then, using the properties of the GPs, the value of $y(\cdot)$ for a trajectory can be characterized at each point in space \mathbf{p} as an univariate Gaussian: $y(\mathbf{p}) \sim \mathcal{G}(y^*(\mathbf{p}), \sigma^2(\mathbf{p}))$. Particularly, the mean and the variance of this univariate Gaussian are inferred from a sampling of \mathcal{F} , $\mathbf{f} = \{\mathbf{f}_1 \dots \mathbf{f}_N\} \subset \mathcal{F}$:

$$y^*(\mathbf{p}) = [\mathbf{C}_{\mathbf{f}}(\mathbf{p})]^T \mathbf{C}_{\mathbf{ff}}^{-1} \mathbf{1} \quad \text{and} \quad \sigma^2(\mathbf{p}) = c(\mathbf{p}, \mathbf{p}) - \mathbf{C}_{\mathbf{f}}(\mathbf{p})^T \mathbf{C}_{\mathbf{ff}}^{-1} \mathbf{C}_{\mathbf{f}}(\mathbf{p}) \quad (1)$$

where $[\mathbf{C}_{\mathbf{f}}(\mathbf{p})]_i = [c(\mathbf{f}_i, \mathbf{p})]_i$, $[\mathbf{C}_{\mathbf{ff}}]_{ij} = [c(\mathbf{f}_i, \mathbf{f}_j)]_{ij}$, $\mathbf{1} = [1 \dots 1]^T$ and

$$c(\mathbf{p}, \mathbf{p}') := \psi(\|\mathbf{p} - \mathbf{p}'\|), \quad \psi(r) = \begin{cases} 2|r|^3 - 3Rr^2 + R^3 & r \leq R \\ 0 & r > R \end{cases}.$$

Once we have represented single trajectories as GPs, obtaining the representation for a bundle of trajectories $\mathcal{B} = \{\mathcal{F}_1, \dots, \mathcal{F}_N\}$ is straightforward: the fact that at each point the value of the indicator function for each trajectory \mathcal{F} , $y_{\mathcal{F}}(\mathbf{p})$, has a univariate Gaussian distribution, enables us to calculate the mean and covariance of the Gaussian distribution characterizing the indicator function for \mathcal{B} at a point \mathbf{p} as [8]

$$y_{\mathcal{B}}^*(\mathbf{p}) = \frac{1}{N} \sum_{i=1}^N y_{\mathcal{F}_i}^*(\mathbf{p}) \quad \text{and} \quad \sigma_{\mathcal{B}}(\mathbf{p}) = \frac{1}{N^2} \sum_{\mathcal{F} \in \mathcal{B}} \sigma_{\mathcal{F}}^2(\mathbf{p}).$$

The result of the combination of the indicator functions representing various trajectories into a bundle is illustrated in fig. 2(e).

Finally, according to [8], the TPM for a WM tract formed by a bundle of trajectories \mathcal{B} is calculated using the GP framework as

$$\text{TPM}(\mathbf{p}) := \text{prob}\{\mathbf{p} \in \mathcal{B}\} \propto \mathbb{E}[y_{\mathcal{B}}(\mathbf{p}) = 1] = \frac{1}{2\sqrt{\pi(1 + \sigma_{\mathcal{B}}^2(\mathbf{p}))}}, \quad (2)$$

where $\mathbb{E}[\cdot]$ stands for the expected value of a random variable. Figure 1 illustrates the *tract probability map* calculated from a GP for a bundle using color-coded surfaces and probability maps overlaid on the FA image of the spatially normalized subject.

In the above, we have introduced our GP-based framework for representing white matter fiber bundles and its two operations: combination of fibers into a bundle and calculation of the TPM. These tools are fundamental to performing a tract-based statistical analysis of the WM, which we develop in the next section.

3 Skeleton-Based WM Tract Analysis

Our statistical analysis framework is designed to help identify precise differences in tracts. The algorithm to analyse a particular fibre bundle starts by generating its TPM as in section 2. By construction, the $\text{TPM}(\cdot)$ function has a ridge (or valley) of high probability at the centre of the tract that decays going outwards. Consequently, the main idea is to find a representation of the tract as a sheet or a line and to project diffusivity information on that representation. For this, we adapt the thinning algorithm of [2] which is used extensively in recent literature. However, our work can be easily restated using a different thinning algorithm. Then, we project a scalar diffusivity measure (e.g. the FA) on the skeleton and perform group analysis on it. This process is shown in fig. 4.

Skeleton Calculation. The first step in the algorithm is to obtain the skeleton. For this, we start by identifying the voxels contained in such a skeleton. We compute the centre of gravity (CofG) of a sphere of radius r around the voxel \mathbf{p} , $S(\mathbf{p}, r) \subset \mathbb{R}^3$,

$$\text{CofG}(\mathbf{p}; r) = \frac{\int_{\mathbf{q} \in S(\mathbf{p}, r)} \mathbf{q} \text{TPM}(\mathbf{q}) d\mathbf{q}}{\int_{\mathbf{q} \in S(\mathbf{p}, r)} \text{TPM}(\mathbf{q}) d\mathbf{q}}. \quad (3)$$

If the centre of gravity is sufficiently close to \mathbf{p} , it is reasonable to assume that \mathbf{p} is very close the ridge or valley which we will call skeleton from now on. This is illustrated in fig. 3(a) where \mathbf{p} is a cross, $S(\mathbf{p}, r)$ a circle and the $\text{CofG}(\mathbf{p}, r)$ a square. Then, to select a voxel in the skeleton, we start from each voxel that is sufficiently close to its centre of gravity and move in a direction perpendicular to the skeleton until we reach a local maximum. The voxels containing such a maximum are considered to be in the skeleton.

As we have seen previously, in order to extract the skeleton, we need to calculate the direction that is perpendicular to it. If the voxels are sufficiently far away from it, this direction is provided by the gradient of the TPM. This is due to the fact that the probability of a voxel belonging to the underlying tract of the TPM decreases when we move away from the skeleton. However, if we are on the skeleton, meaning on a ridge or a valley, the gradient will have a 0 norm and no defined orientation. In these cases it is reasonable to assume that the gradient changes more rapidly in the direction perpendicular to the skeleton [10]. Thus, we take the eigenvector corresponding to the largest eigenvalue of the Hessian of the TPM, as the direction perpendicular to the skeleton. This is illustrated in fig. 3. Then, the direction perpendicular to the skeleton at \mathbf{p} is given by

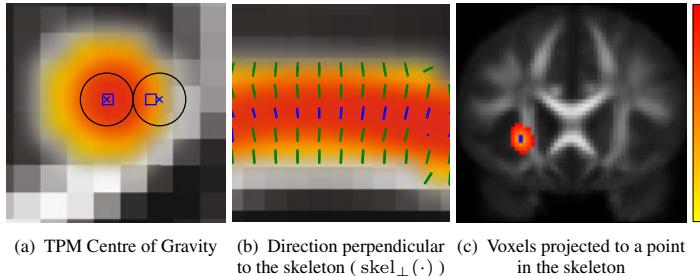


Fig. 3. Finding the direction perpendicular to the tract probability map skeleton. Fig.(a): the deviation of the TPM's centre of gravity (square) from the centre of a neighbourhood (cross) indicates the distance from the skeleton. Fig.(b): when the voxel is far enough from the skeleton, the direction of the gradient (green) is perpendicular; when it is close enough, the gradient norm is close to zero, in these cases the principal direction of the Hessian (blue) indicates the perpendicular direction. Fig. (c): skeleton of a WM tract (blue) and the voxels which are closest to it.

$$\text{skel}_\perp(\mathbf{p}) := \begin{cases} \nabla \text{TPM}(\mathbf{p}) / \| \nabla \text{TPM}(\mathbf{p}) \| & \text{if } \| \text{CofG}(\mathbf{p}; r) - \mathbf{p} \| > \epsilon \\ \mathbf{e}_1 (\nabla^2 \text{TPM}(\mathbf{p})) & \text{if } \| \text{CofG}(\mathbf{p}; r) - \mathbf{p} \| \leq \epsilon \end{cases}. \quad (4)$$

In this equation ϵ is taken to be a tenth of the voxel width and $\mathbf{e}_1(M)$ is the eigenvector corresponding to the largest eigenvalue of M , and ∇ and ∇^2 are the gradient and the Hessian operators. Moreover, due to the properties of our GP framework, the gradient and Hessian of the TPM (eq. (2)) can be calculated analytically.

Finally, it is important to emphasize that this is just one technique to calculate the skeleton. The statistical analysis algorithm is not dependent on it, but is general and can be applied to any skeleton calculated from a TPM. Most of the ridge and valley calculation algorithms are based on the gradient and Hessian of the analysed function [10], hence it is possible to reformulate them in terms of the GP framework for white matter bundles.

Voxel Projections and Statistical Analysis. Having created the skeleton for a tract, we project all the voxels within the thresholded TPM to their closest point on the skeleton. This is achieved by starting at every voxel and following the direction perpendicular to the skeleton at that voxel, calculated in eq. (4), until the skeleton is reached. Then, we calculate the expected scalar diffusivity measure at every point of the skeleton. In fig. 3(c), we show a point on the inferio-fronto occipital tract (IFOF) skeleton in green and the voxels in the image which are closest to it. The values of these voxels are projected to the skeleton and averaged according to their probability of being in the bundle. This procedure for the IFO and the fractional anisotropy (FA), is as follows: for each point in the skeleton of IFO, \mathbf{s} , we calculated the set of voxels $\mathcal{N}(\mathbf{s})$ within the TPM having \mathbf{s} as the closest on the skeleton. This is done by following the direction perpendicular to the skeleton calculated in the previous section. Then, the expected value of FA on each skeleton voxel \mathbf{s} , is the expectation of the value of the FA on $\mathcal{N}(\mathbf{s})$ projected to \mathbf{s} :

$$E_{\text{FA}}^{\text{IFOF}}(\mathbf{s}) = \frac{\int_{\mathcal{N}(\mathbf{s})} \text{TPM}^{\text{IFOF}}(\mathbf{p}) \text{FA}(\mathbf{p}) d\mathbf{p}}{\int_{\mathcal{N}(\mathbf{s})} \text{TPM}^{\text{IFOF}}(\mathbf{p}) d\mathbf{p}} \quad (5)$$

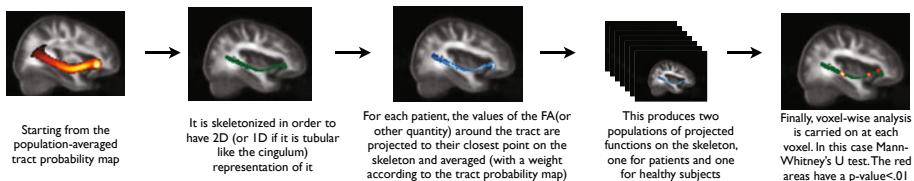


Fig. 4. Skeleton-based tract-specific statistical analysis on a WM tract

Finally, for each voxel in the skeleton, we perform voxel-based analysis. In this case we use the Mann-Whitney U test to find voxels where the distribution of the fractional anisotropy or the axial or radial diffusivities among patients and controls is different, and we report trends indicated by the analysed t-score.

4 Data

Whole-brain DWI datasets were acquired on 58 volunteers (35.60 ± 10.48 years, 28 Male, 24 healthy, 34 diagnosed with schizophrenia) on a Siemens Trio 3T scanner with $1.71 \times 1.71 mm^2$ in-plane resolution, $2mm$ slice thickness. Six unweighted B0 images and 64 diffusion weighted images ($b=1000 s/mm^2$) were acquired with non-collinear diffusion sensitizing gradients.

Data Preparation. DTI images for each subject were computed and deformably registered, using DTI-DROID [11], to a DTI atlas [9]. Full brain tractography was performed by seeding in sub-voxel resolution by taking every voxel with linear anisotropy higher than .3, dividing it into $0.25 \times 0.25 \times 0.25 mm^3$ sub-voxels and seeding from each sub-voxel. On average, 10,000 fibre tracts were obtained for each subject.

Identification of WM Tracts. In order to identify white matter tracts, the clustering algorithm presented by [8] was applied to every subject individually. *Population-averaged tract probability maps* for major tracts were obtained for each extracted white matter tract: on both hemispheres the arcuate, inferio-fronto occipital, uncinate fasciculi and cingulum bundle along with the fornix; also, the frontal and posterior forceps were extracted. In order to perform the statistical analysis, we extracted the skeleton of each *population-averaged tract*.

Statistical Analysis of Tracts. To characterize diffusion properties of the white matter structures, four scalar quantities derived from the diffusion tensors were used: Fractional Anisotropy (FA) [1], axial diffusivity ($\lambda_{||} = \lambda_1$) and radial diffusivity ($\lambda_{\perp} = \frac{1}{2}(\lambda_2 + \lambda_3)$). The FA measure was used due to its biological interpretation as a measure of tract integrity and its popularity in the literature. We chose the other two measures because there are several reports relating differences on axial and radial diffusivity with differences in myelination and axonal calibre [12].

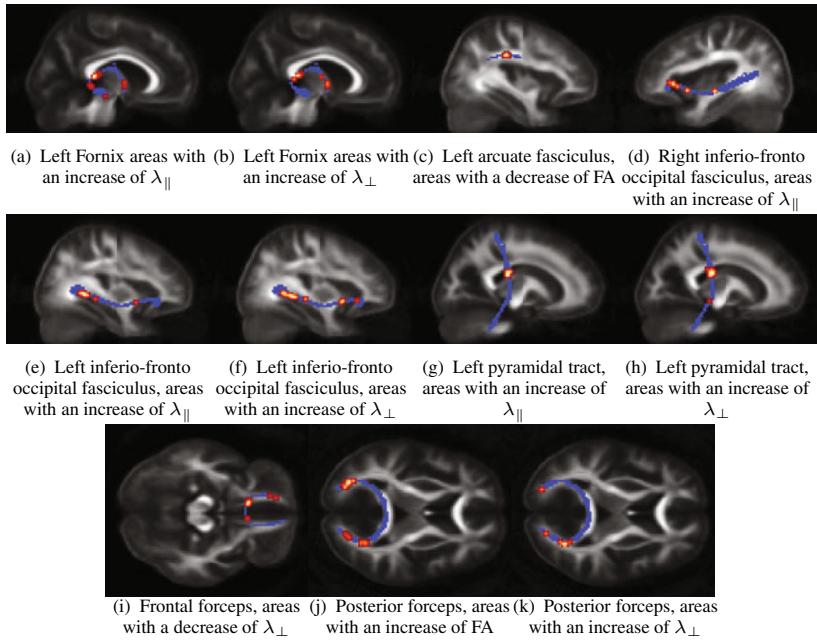


Fig. 5. Differences between healthy and schizophrenic populations in tracts obtained using the statistical analysis described in section 3. We show tracts where more than 10% of the number of voxels in the whole skeleton, shown in blue, have a p-value < 0.001. For the regions defined by these voxels, shown in red, we report trend of the difference between patients and controls indicated by the t-statistic.

5 Results

Results of our skeleton-based analysis are shown in fig. 5. In this figure we show tracts where more than 10% of the number of voxels in the whole skeleton have a p-value < 0.001. For the regions of the tracts defined by these voxels, we report the trend of the difference between patients and controls indicated by the t-statistic in order to characterize the detected change in diffusivity properties. In the left fornix, there are regions with an increase of λ_{\parallel} and λ_{\perp} ; the right arcuate fasciculus shows regions with a decrease of FA; the right inferio-fronto occipital fasciculus shows areas of decreasing λ_{\parallel} ; in the left inferio-fronto occipital fasciculus there are areas of increased axial and λ_{\perp} ; the frontal forceps shows areas of decreased λ_{\perp} ; the posterior forceps has areas of increased FA and λ_{\perp} ; and the left pyramidal tract exhibits an increase of λ_{\parallel} and λ_{\perp} .

6 Discussion

Our results are in agreement with current literature. The predominance of differences in tracts within the left frontal and temporal lobes has been reported by a recent meta-analysis of DTI studies in schizophrenia [13] and a detailed survey [14]. Particularly, the

decrease of the FA on the left arcuate fasciculus, shown by our analysis was reported by [15]. Regarding our studies analysing FA changes in tract skeletons, we found an increased FA in left IFO which was previously shown by [16] and decreased FA in the fornix is consistent with [17]. This agreement with literature demonstrates that our statistical analysis is suited for white matter studies, therefore encouraging us to carry on with a more comprehensive study on larger databases.

Regarding the analyses of parallel (λ_{\parallel}) and perpendicular (λ_{\perp}) diffusivities, even if recent work shows that these measures are easier to correlate with biological changes [12], there are no studies in schizophrenia using DTI that consider them [14,13]. However, our results agree with [13] in that the differences found are predominantly in the temporal and frontal lobes of the left hemisphere. In future work, we expect that an in-depth study of the differences in these quantities will shed some light into the biological characteristics of white matter changes in schizophrenia.

Finally, the method presented is not bound to a single scalar diffusivity measure, it is even generalizable to full-tensor studies as done, for instance, by [6]. Moreover, it is possible to correlate these changes with age and cognitive benchmarks in order to better analyse the relation between white matter changes and the different degrees of disease progression.

Acknowledgements. This work was supported by NIH grants #R01MH079938 and #R01MH060722, the French ANR NucleiPark and the France-Parkinson Association.

References

1. Pierpaoli, C., Barnett, A., Pajevic, S., Chen, R., Penix, L., Virta, A., Basser, P.: Water diffusion changes in Wallerian degeneration and their dependence on white matter architecture. *Neuroimage* 13(6), 1174–1185 (2001)
2. Smith, S.M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T.E., Mackay, C.E., Watkins, K.E., Ciccarelli, O., Cader, M.Z., Matthews, P.M., Behrens, T.E.J.: Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31(4), 1487–1505 (2006)
3. O'Donnell, L.J., Westin, C.F., Golby, A.J.: Tract-based morphometry for white matter group analysis. *NeuroImage* 45(3), 832–844 (2009)
4. Maddah, M., Grimson, W.E.L., Warfield, S.K., Wells, W.M.: A unified framework for clustering and quantitative analysis of white matter fiber tracts. In: MIA, pp. 191–202 (2008)
5. Corouge, I., Fletcher, P.T., Joshi, S., Gouttard, S., Gerig, G.: Fiber tract-oriented statistics for quantitative diffusion tensor mri analysis. *MIA* 10(5), 786–798 (2006)
6. Yushkevich, P.A., Zhang, H., Simon, T.J., Gee, J.C.: Structure-specific statistical mapping of white matter tracts. *NeuroImage* 41(2), 448–461 (2008)
7. Zhang, H., Awate, S., Das, S., Woo, J., Melhem, E., Gee, J., Yushkevich, P.: A Tract-Specific Framework for White Matter Morphometry Combining Macroscopic and Microscopic Tract Features. In: Yang, G.-Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (eds.) MICCAI 2009. LNCS, vol. 5762, pp. 141–149. Springer, Heidelberg (2009)
8. Wassermann, D., Bloy, L., Kanterakis, E., Verma, R., Deriche, R.: Unsupervised white matter fiber clustering and tract probability map generation: Applications of a gaussian process framework for white matter fibers. *NeuroImage* 51(1), 228–241 (2010)

9. Hua, K., Zhang, J., Wakana, S., Jiang, H., Li, X., Reich, D.S., Calabresi, P.A., Pekar, J.J., van Zijl, P.C.M., Mori, S.: Tract probability maps in stereotaxic spaces: Analyses of white matter anatomy and tract-specific quantification. *NeuroImage* 39(1), 336–347 (2008)
10. Kindlmann, G., Estepar, R., Smith, S., Westin, C.: Sampling and Visualizing Creases with Scale-Space Particles. *IEEE Trans. on Viz. and Comp. Graph.* 15(6), 1415–1424 (2009)
11. Yang, J., Shen, D., Davatzikos, C., Verma, R.: 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)
12. Beaulieu, C.: The Biological Basis of Diffusion Anisotropy. In: *Diffusion MRI*. Academic Press, London (2009)
13. Ellison-Wright, I., Bullmore, E.: Meta-analysis of diffusion tensor imaging studies in schizophrenia. *Schizophrenia Research* 108(1-3), 3–10 (2009)
14. Kubicki, M., Shenton, M.E.: Diffusion Tensor Imaging and Its Application to Schizophrenia and Related Disorders. In: *Diffusion MRI*. Academic Press, London (2009)
15. Burns, J., Job, D., Bastin, M.E., Whalley, H., Macgillivray, T., Johnstone, E.C., Lawrie, S.M.: Structural disconnectivity in schizophrenia: a diffusion tensor magnetic resonance imaging study. *The British journal of psychiatry: the journal of mental science* 182, 439 (2003)
16. Ashtari, M., Cottone, J., Ardekani, B., Cervellione, K., Szeszko, P., Wu, J., Chen, S., Kumra, S.: Disruption of white matter integrity in the inferior longitudinal fasciculus in adolescents with schizophrenia as revealed by fiber tractography. *Arch. of Gen. Psych.* 64(11), 1270 (2007)
17. Fitzsimmons, J., Kubicki, M., Smith, K., Bushell, G., Estepar, R., Westin, C., Nestor, P., Niznikiewicz, M., Kikinis, R., McCarley, R., et al.: Diffusion tractography of the fornix in schizophrenia. *Schizophrenia Research* 107(1), 39–46 (2009)