DEPARTMENT OF STATISTICS
University of Wisconsin
1210 West Dayton St.
Madison, WI 53706

TECHNICAL REPORT NO. 1104
June 20, 2005

Anisotropic Kernel Smoothing in Diffusion Tensor Imaging: Theoretical Framework

[1,2,3]Moo K. Chung, [3,4]Jee Eun Lee, [3,4]Andrew L. Alexender

[1]Department of Statistics
[2]Department of Biostatistics and Medical Informatics
[3]Waisman Laboratory for Brain Imaging and Behavior
[4]Department of Medical Physics
University of Wisconsin-Madison
Madison, WI 53706

Anisotropic Kernel Smoothing in Diffusion Tensor Imaging: Theoretical Framework

[1,2,3]Moo K. Chung, [3,4]Jee Eun Lee, [3,4]Andrew L. Alexender

[1]Department of Statistics, [2]Department of Biostatistics and Medical Informatics
[3]Waisman Laboratory for Brain Imaging and Behavior
[4]Department of Medical Physics
University of Wisconsin-Madison
Madison, WI 53706

Abstract

We present a theoretical framework for smoothing diffusion tensor images. We formulate smoothing data along the white fiber tracks as iterated anisotropic Gaussian kernel smoothing. The formulation is derived from the construction of the transitional probability of diffusion process. The smoothing is performed in such a way that it smooth data more along the most likely direction of water molecule diffusion direction. Statistical properties of our smoothing are also presented.

1 Introduction

Diffusion tensor imaging (DTI) is a new magnetic resonance imaging technique technique that provides the directional information of water diffusion in the white matter of the brain. The directional information is usually represented as a symmetric positive definite 3×3 matrix which is usually termed as the *diffusion tensor* $D = (d_{ij})$. The water diffusion is assumed to follow diffusion equation

$$\frac{\partial C}{\partial t} = \nabla \cdot (D \nabla C), \tag{1}$$

where C is the concentration of water molecules [10].

The previous approaches in smoothing DTI data are usually formulated as the anisotropic diffusion or the Laplace-Beltrami flow [3, 13, 17, 18, 21]. In anisotropic diffusion framework, the white fiber connectivity metric is given as a solution to equation (1) with the initial condition where every voxel is zero except a seed region where it is given the value one [3]. Then the value 1 at the seed region is diffused though the brain and the numerical value that should be between 0 and 1 is taken as a probability of connection. Mathematically it is equivalent as the Monte-Carlo random walk simulation with no restriction [11].

Our kernel smoothing approach is different from [3] and [11]. In our kernel-based approach, we construct anisotropic kernel that is the transition probability density of a continues diffusion process and perform iterative anisotropic Gaussian kernel smoothing. This formulation is simpler than solving equation (1) numerically. Further, unlike

the Monte-Carlo random walk simulation approach, it always gives the deterministic results that does not change from one run to another. Due to the simplistic nature of our smoothing formulation, it is also easier to derive its statistical properties.

As applications of our technique, we will present two examples. In the first example, the transitional probability from a region in the corpus callosum will be computed. In the second example, the fractional anisotropy (FA) map will be smoothed out to show that our smoothing will preserve FA while increasing the signal-to-noise ratio (SNR).

2 Transition probability

Let $P_t(\mathbf{p}, \mathbf{q})$ be the *transition probability density* of a particle going from $\mathbf{p}$ to $\mathbf{q}$ under anisotropic diffusion process. This is the probability density of the particle hitting $\mathbf{q}$ at time t when the particle is at $\mathbf{p}$ at time 0. So we have $\int_{\mathbb{R}^n} P_t(\mathbf{p}, \mathbf{x})\ d\mathbf{x} = 1$ and $P_t(\mathbf{p}, \mathbf{q}) = P_t(\mathbf{q}, \mathbf{p})$. The transition probability serves as a mathematical analogue of the water molecule concentration C in equation (1). The solution to equation (1) with the initial condition $p(0, \mathbf{x}) = \delta(\mathbf{x})$, the Dirac delta function, has been used as the probabilistic representation of white fiber track connectivity in [3].

When D is constant, it can be shown that $P_t(\mathbf{p}, \mathbf{q}) = K_t(\mathbf{p} - \mathbf{q})$, the n-dimensional anisotropic Gaussian kernel defined as follows. First, let $\mathbf{x} = (x_1, \cdots, x_n)' \in \mathbb{R}^n$. The n-dimensional isotropic Gaussian kernel is defined as $K(\mathbf{x}) = (2\pi)^{-n/2} e^{-\mathbf{x}'\mathbf{x}/2}$ which is the joint multivariate normal probability density function of n independent standard normal random variables. Note that $\int_{\mathbb{R}^n} K(\mathbf{x})\ d\mathbf{x} = 1$. Then we define anisotropic kernel K_t via transformation $h : \mathbf{x} \to (2\sqrt{t}D)^{-1/2}\mathbf{x}$ such that

$$K_t(\mathbf{x}) = \det{}^{-1}\Big(\frac{\partial h}{\partial \mathbf{x}}\Big) K \circ h(\mathbf{x}) = \frac{1}{(4\pi t)^{n/2} \det{}^{1/2} D} \exp\Big(-\frac{\mathbf{x} D^{-1} \mathbf{x}}{4t}\Big).$$

The transformation matrix $2\sqrt{t}D^{-1/2}$ is called the bandwidth matrix and controls the concentration of the anisotropic kernel weights. K_t is a multivariate normal density with the mean zero and the covariance matrix $2tD$.

For long distance transition probability, the computation is based on the Chapman-Kolmogorov equation [14]. The transition probability density of a particle going from $\mathbf{p}$ to $\mathbf{q}$ is the total sum of the probabilities of going from $\mathbf{p}$ to $\mathbf{q}$ through all possible intermediate points $\mathbf{x} \in \mathbb{R}^n$:

$$P_t(\mathbf{p}, \mathbf{q}) = \int_{\mathbb{R}^n} P_s(\mathbf{p}, \mathbf{x}) P_{t-s}(\mathbf{x}, \mathbf{q})\ d\mathbf{x} \tag{2}$$

for any $0 < s < t$. The equation still hold in the case when s is either 0 or t, since in that case one of the probability in the integral will become the Dirac-delta function and in turn the integral collapse to the probability on the left side. For $\mathbf{x} \in B_{\mathbf{p}}$, small neighborhood centered around $\mathbf{p}$, D can be considered constant and hence $P_t(\mathbf{p}, \mathbf{q}) \doteq K_t(\mathbf{p} - \mathbf{q})$. Therefore, the Chapman-Kolmogorov equation (2) can be written as

$$P_t(\mathbf{p}, \mathbf{q}) \doteq \frac{\int_{B_{\mathbf{p}}} K_s(\mathbf{p} - \mathbf{x}) P_{t-s}(\mathbf{x}, \mathbf{q})\ d\mathbf{x}}{\int_{B_{\mathbf{p}}} K_s(\mathbf{p} - \mathbf{x})\ d\mathbf{x}}. \tag{3}$$

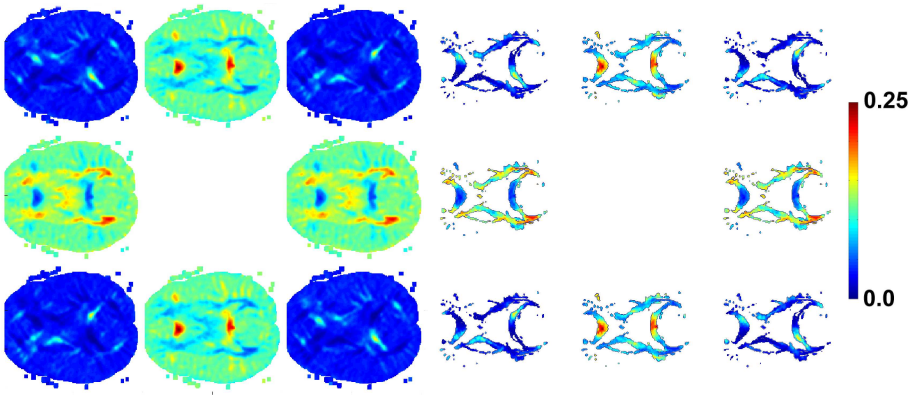

Fig. 1. The transition probabilities from the center voxel to neighboring voxels. They form the anisotropic kernel weights. Left: transition probability without restriction. Right: transition probability within FA > 0.4.

The denominator is introduced to correct for the truncation error. Note that when $s \to 0$ or B_p becomes large, the equation (3) converges to equation (2). Let us define a truncated Gaussian kernel

$$\tilde{K}_t(\mathbf{p}-\mathbf{x}) = \frac{K_t(\mathbf{p}-\mathbf{x})\mathbf{1}_{B_\mathbf{p}}(\mathbf{x})}{\int_{B_\mathbf{p}} K_t(\mathbf{p}-\mathbf{x})\, d\mathbf{x}}$$

where $\mathbf{1}_{B_\mathbf{p}}(\mathbf{x})$ is the index function taking value 1 if $\mathbf{x} \in B_\mathbf{p}$ and 0 otherwise. Rewriting the Chapman-Kolmogorov equation, for each fixed $\mathbf{p}$, we have

$$P_t(\mathbf{p},\mathbf{q}) \doteq \tilde{K}_s * P_{t-s}(\mathbf{p},\mathbf{q}).$$

The convolution is with respect to the first argument $\mathbf{p}$. By letting $t = k\Delta t$ and $s = \Delta t$, we have iterative relationship: $P_{k\Delta t}(\mathbf{p},\mathbf{q}) = \tilde{K}_{\Delta t} * P_{(k-1)\Delta t}$. Solving this we have k-fold convolution

$$P_{k\Delta t}(\mathbf{p},\mathbf{q}) = \underbrace{\tilde{K}_{\Delta t} * \tilde{K}_{\Delta t} * \cdots * \tilde{K}_{\Delta t}}_{k \text{ times}} * P_0(\mathbf{p},\mathbf{q}), \tag{4}$$

where $P_0(\mathbf{p},\mathbf{q}) = \delta(\mathbf{p}-\mathbf{q})$. For numerical implementation, we take the Kroneker-delta function rather than the Dirac-delta function. We will denoting the k-fold convolution as $\tilde{K}^{(k)}_{\Delta t}$. Note that

$$\int_{\mathbb{R}^n} P_{k\Delta t}(\mathbf{p},\mathbf{q})\, d\mathbf{q} = 1$$

for all k.

The computation for the long distance transition probability can be viewed as smoothing the initial probability distribution $P_0(\mathbf{p}-\mathbf{q}) = \delta(\mathbf{p}-\mathbf{q})$. In the numerical implementation, the initial probability 1 at the seed will be smoothed over all voxels.

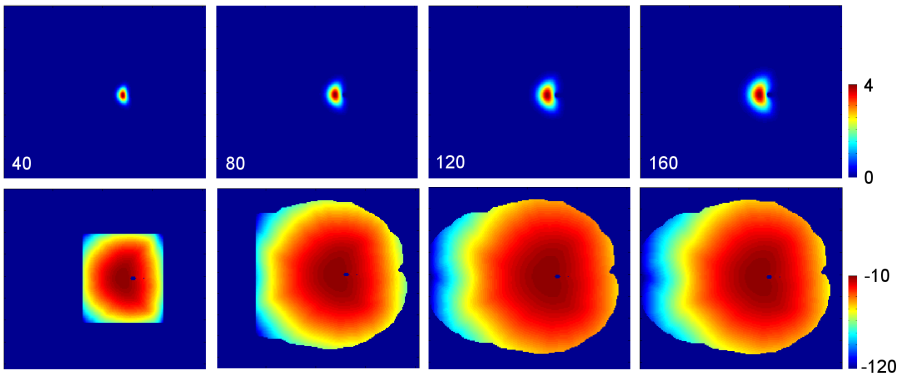

Fig. 2. The long distance transition probability from the seed voxel in the corpus callosum with $\Delta t = 0.1$ and $k = 40, 80, 120, 160$ iterations. The scale in the top images is 10^{-3}. The bottom images are in natural log scale.

Figure 1 shows the transition probability from one voxel to the neighboring voxels computed from the anisotropic Gaussian kernel. In the left figure, the transition probability is computed without constrain while the right figure shows the transition probability computed within FA > 0.4. Figure 2 shows the long distance transition probability. The bottom figure shows the transition probability represented in log-scale, i.e. $\rho = \log P_t(\mathbf{p}, \mathbf{q})$. $\Delta t = 0.1$ with $k = 40, 80, 120, 150$ iterations are used.

3 Anisotropic kernel smoothing

Instead of smoothing the initial probability, the same approach can be used to smooth noisy data in such a way that we smooth more along the direction the white matter tracks and less across the tracks. Consider observation Y which will be assumed to be

$$Y(\mathbf{x}) = f(\mathbf{x}) + \epsilon(\mathbf{x}),$$

where f is the mean function and ϵ is a zero mean Gaussian random field. We assume that ϵ is an anisotropic field. We are interested in estimating f via kernel smoothing. If the signal is defined along the white matter tracks, it is necessary to incorporate the directional information of the white fiber into the shape of kernel. In computing the transitional probability, we smoothed the initial probability distribution by the iterated kernel smoothing framework. Instead smoothing probability, we will smooth observation Y along the white fiber tracks. We formulate *anisotropic Gaussian kernel smoothing* as the convolution of observation Y and kernel K_t:

$$\hat{f}(\mathbf{x}) = K_t * Y(\mathbf{x}) = \int_{\mathbb{R}^n} K_t(\mathbf{x} - \mathbf{y}) Y(\mathbf{y})\, d\mathbf{y}. \tag{5}$$

It can be shown that the above convolution is the unique solution to (1) with the initial condition $C(\mathbf{x}, 0) = Y(\mathbf{x})$ after time t for fixed D.

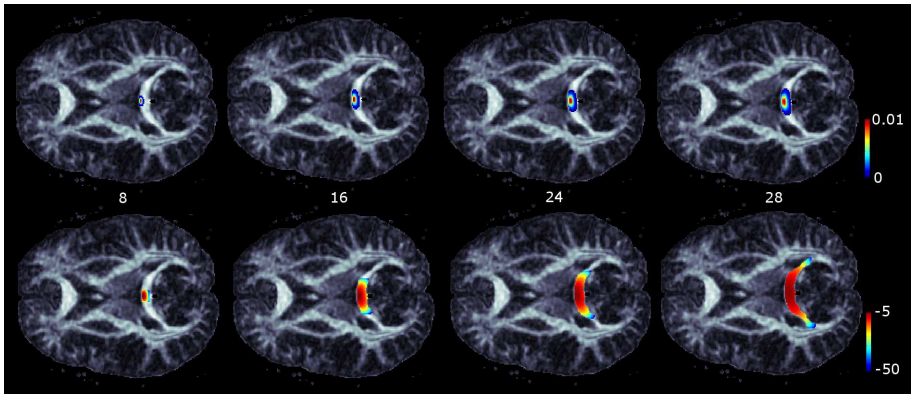

Fig. 3. Transition probability from a single point in the corpus callosum with $t = 0.1$ and $k = 8, 16, 24, 28$ iterations. The scale in the top images is 10^{-3}. The bottom images are in natural log scale.

The anisotropic kernel smoothing can be viewed as a local regression. To see this, consider partial differential operator $\mathcal{L}\psi = \nabla \cdot (D\nabla\psi)$ and its eigenvalue problem $\mathcal{L}\psi = \lambda\psi$. Since D is positive definite, we have ordered positive eigenvalues $0 = \lambda_0 \leq \lambda_1 \leq \lambda_2 \leq \cdots$ and corresponding eigenfunctions $\psi_0, \psi_1, \cdots$. The eigenfunctions ψ_j form an orthonormal basis of $L^2(\mathbb{R}^n)$, the L^2 space of functions defined on $\mathbb{R}^n$. Then it can be shown that

Theorem 1.

$$K_t * Y(\mathbf{x}) = \arg\min_{h \in L^2(\mathbb{R}^n)} \int K_t(\mathbf{x}-\mathbf{y})\Big[Y(\mathbf{y}) - h\Big]^2 \, d\mu(\mathbf{y}).$$

By expanding the integral as a quadratic in h: $h^2 - 2hK_t * Y + K_t * Y^2$ and differentiating with respect to h, we obtains the minimum when $h(\mathbf{x}) = K_t * Y(\mathbf{x})$. This theorem is also true for the l-dimensional finite subspace $\mathcal{H}_l \subset L^2(\mathbb{R}^n)$ that is spanned by the finite basis functions $\psi_0, \psi_1, \cdots, \psi_l$.

The following theorem shows the total amount of measurements is invariant under smoothing.

Theorem 2 (Conservation of total signal).

$$\int_{\mathbb{R}^n} K_t^{(k)} * Y \; d\mathbf{x} = \int_{\mathbb{R}^n} Y \; d\mathbf{x}.$$

It can be easily seen by interchanging the order of integral

$$\int_{\mathbb{R}^n} K_t * Y \; d\mathbf{x} = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K_t(\mathbf{x}-\mathbf{y}) Y(\mathbf{y}) \; d\mathbf{y} \; d\mathbf{x}$$

and noting K_t is a probability distribution. Then the statement follows inductively. The bias of the heat kernel smoothing $\mathbb{E}\hat{f} - f = K_t * Y - f \to 0$ as $t \to 0$. The total bias

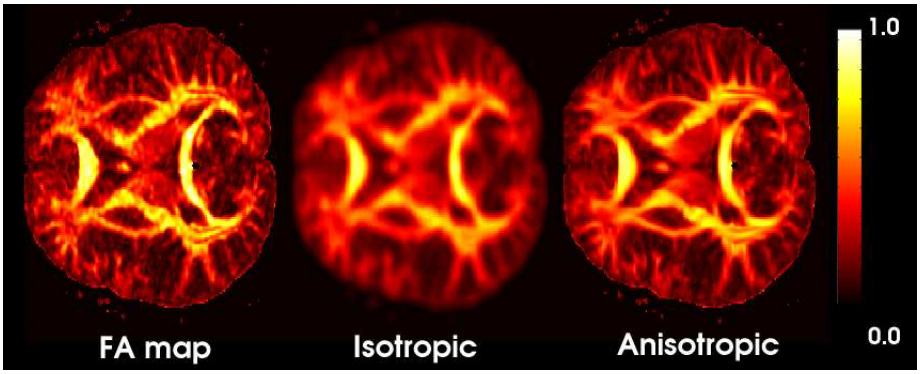

Fig. 4. Left: FA image. Middle: isotropic Gaussian kernel smoothing with 12mm FWHM. Right: anisotropic Gaussian kernel smoothing with the equivalent filter size showing the preservation of directional information. The advantage of the anisotropic smoothing is that it increase the signal-to-noise ratio without sacrificing much of the directional characteristic of DTI.

always vanishes: $\int_{\mathbb{R}^n}[\mathbb{E}\hat{f}(\mathbf{x}) - f(\mathbf{x})]\ d\mathbf{x} = \int_{\mathbb{R}^n}[K_t * f(\mathbf{x}) - f(\mathbf{x})]\ d\mathbf{x} = 0$. This is the consequence of Theorem 2.

We have applied our smoothing method on each component of diffusion tensor $D = (d_{ij})$ for increasing the signal-to-noise ratio in the FA measure. The following theorem guarantee the positive definiteness is preserved under smoothing.

Theorem 3 (Conservation of positive-definiteness).
If $D = (d_{ij})$ is positive definite, $K_H * D = (K_t * d_{ij})$ is also positive-definite.

Note that D is positive-definite if $\mathbf{x}'D\mathbf{x} = \sum_{ij} x_i x_j d_{ij} > 0$ for all $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$. Then $\mathbf{x}'(K_H * D)\mathbf{x} = \sum_{ij} x_i x_j K_t * d_{ij} = K_t * \sum_{ij} x_i x_j d_{ij} > 0$ proving the statement.

Figure 4 shows the FA images before and after smoothing. For comparison, we performed both isotropic and anisotropic Gaussian kernel smoothing with the similar bandwidth. The standard unit of isotropic smoothing bandwidth is the full width at the half maximum (FWHM) of the Gaussian kernel. For anisotropic smoothing, we have taken the average full width at the half maximum along the three principal axis of the kernel ellipsoid.

4 Conclusions and Discussions

Gössl *et al.* [8] applied an isotropic Gaussian kernel smoothing with small filter size to get the smooth, continuous representation of DTI data using a linear state space model. However, the isotropic smoothing tends to smooth out the directional characteristic of the DTI data. Motivated by this, we have introduced a new method for smoothing data along the white fiber tracks to get the continuous and smooth representation of the data while preserving the directional characteristic of DTI by performing spatially adaptive iterative anisotropic kernel smoothing. The kernel is constructed using the concept of

the transitional probability of the diffusion process that directly corresponds to the water molecule diffusion in DTI.

Compared to the previous approaches of a Monte-Carlo simulation [11] or solving a diffusion equation [5, 13, 21], our kernel method is simpler to implement numerically and statistical properties can be investigated easily. To speed up the kernel smoothing, one may use the decomposition scheme of Geusebroek [9].

An additional advantage of our smoothing formulation is that it can be interpreted differential geometrically. Consider vector field $\mathbf{V} = (V_1, \cdots, V_n)'$ that is the principal eigenvector of D. Suppose we wish to smooth data along the field in such a way that we smooth more along the direction of the field. Let the flow $\mathbf{x} = \psi(t)$ corresponding to the field is given by

$$\frac{d\psi}{dt} = \mathbf{V} \circ \psi(t).$$

This system of ordinary differential equations give a family of integral curves whose tangent vector is $\mathbf{V}$. The line element is

$$d\psi^2 = V_1^2\ dx_1^2 + \cdots + V_n^2 dx_n^2$$

and the Riemannian metric tensor is given by $G = (g_{ij}), g_{ij} = V_i^2 \delta_{ij}$. By matching the bandwidth matrix of a general anisotropic kernel to the Riemannian metric tensor proportionally, we get

$$K_t(\mathbf{x}) = (4\pi t)^{-n/2} \prod_{j=1}^{n} \frac{1}{|V_j|} \exp\Big(- \frac{x_j^2}{4tV_j^2}\Big).$$

By introducing the scaling parameter t, we left a room for adjusting the amount of smoothing. The advantage of using only the principal eigenvectors would be the simplicity of the implementation while the drawback is that the method does not completely utilize the full diffusion tensor D. To use D directly, we note that D gives a natural Riemannian metric tensor in the diffusion process, i.e. $G = D$ [6]. This results in the same anisotropic kernel we have introduced in computing the transition probability. A similar Riemannian metric tensor approach in connection with DTI is given in [4, 12]. In particular [12] matched the diffusion tensor D to the inverse of the metric tensor G and applied it in computing the tensor-warped distances in the white fibers.

References

1. Basser, P.J., and Pierpaoli, C. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *J. Magn. Reson. B* **111:**209-219. 1996.
2. Basser, P.J., Pajevic, S., Pierpaoli, C., Duda, J., Aldroubi, A. In Vivo Tractography using DT-MRI data. *Magn. Res. Med.* **44**:625-632. 2000.
3. Batchelor, P.G., Hill, D.L.G., Atkinson, D., Calamanten, F., Atkinson, D. Study of Connectivity in the Brain Using the Full Diffusion Tensor from MRI, IPMI 2001, *Lecture Notes in Computer Science* **2082**:121-133. 2001.
4. Chefd'hotel, C., Tschumperlé, D., Deriche, R., Faugeras, O. Constrained flows of matrix-valued functions: Application to diffusion tensor regularization. *Lecture Notes in Computer Science* **2350**:251-265, Springer-Verlag, Berlin. 2002.

5. Conturo, T.E., Lori, N.F., Cull, T.S., Akbudak, E., Snyder, A.Z., Shimony, J.S., McKinstry, R.C., Burton, H., Raichle, M.E. Tracking Neuronal Fiber Pathways in the Living Human Brain, *Proc. Natl. Acad. Sci.* **96**:10422-10427. 1999.
6. De Lara, M.C. Geometric and Symmetric Properties of a Nondegenerate Diffusion Process, *Annals of Probability*. **23**:1557-1604. 1995
7. Hagmann, P., Thiran, J.-P., Vandergheynst, P., Clarke, S., Meuli, R. Statistical fiber tracking on DT-MRI data as a potential tool for morphological brain studies. *ISMRM Workshop on Diffusion MRI: Biophysical Issues*. 2000.
8. Gössl, C., Fahrmeir, L., Pütz, B., Auer, L.M., and Auer, D.P. Fiber tracking from DTI using linear state space models: delectability of the pyramidal tract. *NeuroImage* **16:**378-388. 2002.
9. Geusebroek, J.-M., Smeulders, A. W. M., van de Weijer, J. Fast Anisotropic Gauss Filtering, *Lecture Notes in Computer Science* **2350**:99-112, Springer-Verlag, Berlin. 2000.
10. Kang, N. and Zhang, J. and Carlson, E.S., and Gembris, D. White matter fiber tractography via anisotropic diffusion simulation in the human brain. Technical report. 410-04. Department of Computer Science. University of Kentucky. Lexington.
11. Koch, M.A., Norris, D.G., and Hund-Georgiadis, M. An investigation of functional and anatomical connectivity using magnetic resonance imaging. *NeuroImage* **16**:241-250. 2002.
12. O'Donnell, L., Haker, S., Westin, C.-F., New Approaches to Estimation of White Matter Connectivity in Diffusion Tensor MRI: Elliptic PDEs and Geodesics in a Tensor-Warped Space. Fifth International Conference on Medical Image Computing and Computer-Assisted Intervention - MICCAI 02. 2002.
13. Parker, G.J.M., Schnabel, J.A., Symms, M.R., Werring, D.J, Barker, G.J. Nonlinear Smoothing for Reduction of Systematic and Random Errors in Diffusion Tensor Imaging. *Journal of magnetic Resoance Imaging.* **11**:702-710. 2000.
14. Paul, W., Baschnagel, J., *Stochastic Processes from Physics to Finance*. Springer-Verlag Berlin. 2000.
15. Mori, S., Crain, B., Chacko, V.P., van Zijl, P.C.M. Three Dimensional Tracking of Axonal Projections in the Brain by Magnetic Resonance Imaging. *Ann. Neurol.* **45**:265-269. 1999.
16. Stevens, C.F. *The six core theories of modern physics*, The MIT Press. 1995.
17. Spira, A., Kimmel, R., Sochen, N. Efficient Beltrami Flow using a Short Time Kernel. *Lecture Notes in Computer Science.* **2695**: 511-522. 2003.
18. Tang, B., Sapiro, G., Caselles, V. Direction Diffusion. ICCV 1245-1252. 1999.
19. Tench, C.R., Morgan, P.S., Blumhardt, L.D., Constantinescu, C. Improved White Matter Fiber Tracking Using Stochastic Labelling, *Magn. Res. Med.* **48**:677-683. 2002.
20. Wang, Z., Vemuri, B.C., Chen Y, Mareci, T., Simultaneous Smoothing and Estimation of the Tensor Field from Diffusion Tensor MRI. CVPR. 2003.
21. Weinstein, D.M., Kindlmann, G.,Lundberg, E. Tensorlines: Advection-Diffusion Based Propagation Through Diffusion Tensor Fields. *IEEE Visualization* 249-253. 1999.