Tensor Field Regularization using Normalized Convolution and Markov Random Fields in a Bayesian Framework

Chapter · January 2006

DOI: 10.1007/3-540-31272-2_24

CITATIONS
10

READS
53

5 authors, including:

Carl-Fredrik Westin
Harvard University
454 PUBLICATIONS **18,154** CITATIONS
SEE PROFILE

Marcos Martin-Fernandez
Universidad de Valladolid
172 PUBLICATIONS **1,342** CITATIONS
SEE PROFILE

Carlos Alberola-López
Universidad de Valladolid
240 PUBLICATIONS **2,833** CITATIONS
SEE PROFILE

Juan Ruiz-Alzola
Universidad de Las Palmas de Gran Canaria
94 PUBLICATIONS **1,229** CITATIONS
SEE PROFILE

24

Tensor Field Regularization using Normalized Convolution and Markov Random Fields in a Bayesian Framework

Carl-Fredrik Westin, Marcos Martin-Fernandez, Carlos Alberola-Lopez, Juan Ruiz-Alzola, and Hans Knutsson

Laboratory of Mathematics in Imaging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA
{westin,marcma}@bwh.harvard.edu

Summary. This chapter presents two techniques for regularization of tensor fields. We first present a nonlinear filtering technique based on normalized convolution, a general method for filtering missing and uncertain data. We describe how the signal certainty function can be constructed to depend on locally derived certainty information and further combined with a spatially dependent certainty field. This results in reduced mixing between regions of different signal characteristics, and increased robustness to outliers, compared to the standard approach of normalized convolution using only a spatial certainty field. We contrast this deterministic approach with a stochastic technique based on a multivariate Gaussian signal model in a Bayesian framework. This method uses a Markov random field approach with a 3D neighborhood system for modeling spatial interactions between the tensors locally. Experiments both on synthetic and real data are presented. The driving tensor application for this work throughout the chapter is the filtering of diffusion tensor MRI data.

24.1 Introduction

Using conventional MRI, we can easily identify the functional centers of the brain (cortex and nuclei). However, with conventional anatomical MRI techniques, the white matter of the brain appears to be homogeneous without any suggestion of the complex arrangement of fiber tracts. Diffusion Tensor MRI (DT-MRI) is a relatively recent imaging modality that measures the diffusion of water in biological tissue. A common first order model of anisotropic diffusion is the Gaussian model that gives ellipsoidal isoprobability surfaces describing the diffusion. This then naturally leads to the tensor representation through the analogy between symmetric 3×3 tensors to the ellipsoidal representation. Within white matter, the mobility of the water is restricted by the axons that are oriented along the fiber tracts. Hence, the demonstration of

anisotropic diffusion in the brain by MRI has paved the way for non-invasive exploration of the structural anatomy of the white matter in vivo [1, 12]. Alexander presents a good introduction to DT-MRI and the diffusion process in Chap. 5.

The diffusion weighted MRI images are corrupted by noise. There are many factors that contribute to the noise. One source is associated with the receiving coil resistance, others come from inductive losses. The major source of noise will depend on the strength of the static magnetic field and the volume sample size. In addition, the final image noise can also depend on other factors like the voxel size, the receiver bandwidth and the number of time averages in the acquisition process [4]. Hahn et al. discuss the origin of the noise and its impact in Chap. 6. By imposing smoothness or regularization constraints on the tensor field the amount of noise can be reduced. In this chapter we will present two different types of regularization approaches: one deterministic and one stochastic. An overview of other recent regularization methods can be found in [15].

24.2 Normalized Convolution

In this section we outline how normalized convolution (NC) can be used for regularizing scalar, vector, and higher order tensor fields. The method presented here closely follows the description in [17].

NC was introduced as a general method for filtering missing and uncertain data [7, 16]. This method can be viewed as locally solving a weighted least squares (WLS) problem. A local description of a signal, $\mathbf{f}$, can be defined using a weighted sum of basis functions stored as columns, $\mathbf{B}$. Minimizing[1]

$$||\mathbf{W}(\mathbf{B}\boldsymbol{\theta} - \mathbf{f})|| \tag{24.1}$$

with respect to the weights $\boldsymbol{\theta}$, which are the coordinates of the signal in the basis $\mathbf{B}$, results in the WLS solution [5]

$$\mathbf{f}_0 = \mathbf{B}\boldsymbol{\theta} = \mathbf{B}(\mathbf{B}^*\mathbf{W}^T\mathbf{W}\mathbf{B})^{-1}\mathbf{B}^*\mathbf{W}^T\mathbf{W}\mathbf{f}\,, \tag{24.2}$$

where $\mathbf{B}^*$ is the conjugate transpose of $\mathbf{B}$.

In NC the weights $\mathbf{W}$ are split into two parts. One that belongs to the basis functions and usually referred to as the applicability function a, a scalar windowing function that deals with spatial localization of the operators in $\mathbf{B}$. This provides an alternative to traditional windowing with the advantage of not changing the function values.

The second part belongs to the signal and is referred to as the signal certainty function c describing the credence of the signal samples. Missing

[1] The norm of a matrix $\mathbf{A}$ is the Frobenius norm given by $||\mathbf{A}|| = \sqrt{tr(\mathbf{A}^T\mathbf{A})}$, where $tr(\mathbf{A})$ is the trace and $\mathbf{A}^T$ the transpose of $\mathbf{A}$.

samples are handled by setting this function to zero. Note that handling missing data in a more traditional way by for example setting the signal values to zero would introduce artificial structures in the image. Further, the certainty function is usually set to zero outside the signal border reducing the impact of traditional edge effects.

The weight matrix $\mathbf{W}$ can be constructed by multiplying two diagonal matrices containing the applicability and the certainties in the diagonal respectively, $\mathbf{W}_a$ and $\mathbf{W}_c$. To keep in line with established notation [16] we define $\mathbf{W}^T\mathbf{W} = \mathbf{W}_a\mathbf{W}_c$ which will avoid squaring the weights in the calculations. Inserting this into equation (24.2) gives

$$\mathbf{f}_0 = \mathbf{B}\boldsymbol{\theta} = \mathbf{B}(\mathbf{B}^*\mathbf{W}_a\mathbf{W}_c\mathbf{B})^{-1}\mathbf{B}^*\mathbf{W}_a\mathbf{W}_c\mathbf{f}\,. \tag{24.3}$$

24.2.1 Certainty Measures

There are several ways to define signal certainty function c. For simplicity we here will use

$$c = c_v c_s \tag{24.4}$$

where c_v is the voxel certainty and c_s is the similarity certainty. The spatial voxel certainty measure c_v is defined by the input data. Further, the similarity certainty measure c_s has been constructed as:

$$c_s = c_m c_a\,, \tag{24.5}$$

where c_m and c_a are the magnitude and angular similarity measures, respectively. For the magnitude certainty c_m the following Gaussian magnitude function has been used

$$c_m = \exp\left[-\left(\frac{||\mathbf{T}_0|| - ||\mathbf{T}||}{\sigma}\right)^2\right], \tag{24.6}$$

where the norm of a tensor is given by $||\mathbf{T}|| = \sqrt{tr(\mathbf{T}^T\mathbf{T})}$, and $\mathbf{T}_0$ is a tensor calculated from a local neighborhood as explained in Sect. 24.2.3. The angular similarity measure, c_a, is based on the inner product between the normalized tensors and is given by

$$c_a = \langle\hat{\mathbf{T}}_0, \hat{\mathbf{T}}\rangle^\alpha = tr(\hat{\mathbf{T}}_0^T\hat{\mathbf{T}})^\alpha \tag{24.7}$$

where $\hat{\mathbf{T}} = \mathbf{T}/||\mathbf{T}||$. In general the spatial voxel certainty function, c_v, will be based on prior information about the data. The voxel certainty is set to zero outside the signal extent to reduce unwanted border effects. If no specific local information is available the voxel certainty is set to one. As described above, the second certainty component, c_s, is defined locally based on neighboring information. The idea here is to reduce the impact of outliers, where an outlier is defined in terms of the local signal neighborhood, and to reduce the blurring across interfaces between regions having very different signal characteristics.

24.2.2 Applicability Functions

The applicability function define the localization of the operator. The family of applicability functions used in the examples in this chapter is given by

$$a = \begin{cases} r^{-\alpha} \cos^{\beta}\left(\frac{\pi}{2} \frac{r}{r_{\max}}\right) & r < r_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{24.8}$$

where r denotes the distance from the neighborhood center and α, β and $r_{\max}$ are positive constants.

24.2.3 Simple Local Neighborhood Model

The simplest possible model in the NC framework is when using only one constant basis function, simplifying the expression for the NC (24.3) to a ratio of convolutions[2] [7]:

$$\mathbf{T}_0(m,n,p) = \frac{\sum_{i,j,k} a(m+i,j+n,k+p) c_v(i,j,k) \mathbf{T}(i,j,k)}{\sum_{i,j,k} a(m+i,j+n,k+p) c_v(i,j,k)} \tag{24.9}$$

where (m,n,p) are the indexes that correspond to each voxel. To focus on the power of introducing the signal/model similarity certainty measure, this simple local neighborhood model is used in our examples below.

24.2.4 Scalar Field Regularization

Before describing the tensor case (Sect. 24.2.5), we will first present a scalar example to show the effect of the the voxel and magnitude certainty functions. This concept can be seen as generalization of bilateral filtering [14] into the signal-certainty framework of normalized convolution.

Figure 24.1 shows the result of filtering a scalar signal using the proposed technique. Figure 24.1(a) shows the original scalar signal: a noisy step function. Figure 24.1(b) shows the result using standard NC demonstrating that reduction of noise is achieved at the expense of unwanted mixing of features from adjacent regions. The amount of border blurring can be controlled effectively by including the new magnitude certainty measure, c_m, given by equation (24.6). The smaller the σ value, the smaller the inter region averaging. The result for $\sigma = 2$ is shown in Fig. 24.1(c).

[2] Using correlation or convolution is a matter of preference since the applicability function is in general symmetric.

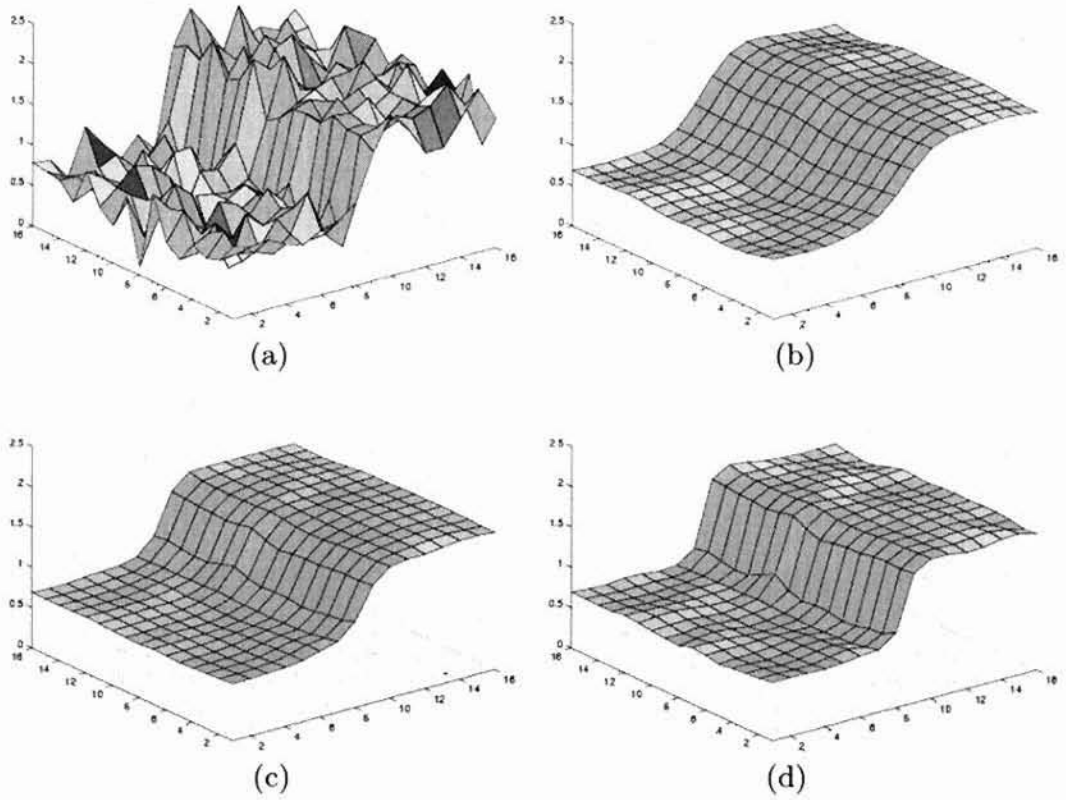
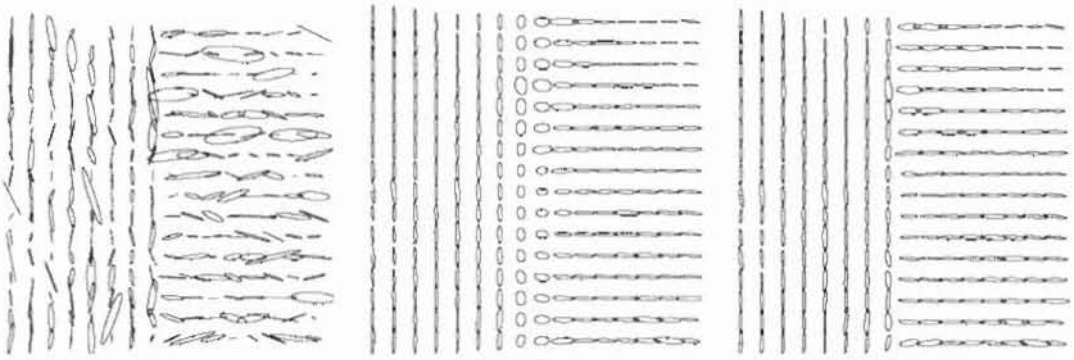


Fig. 24.1. Original scalar field (*step edge*) with added noise (**a**), the result without using the magnitude certainty measurement c_m (**b**), results using c_m with $\sigma = 1$ (**c**) and with $\sigma = 0.5$ (**d**). See appendix for color plates

24.2.5 Tensor Field Regularization

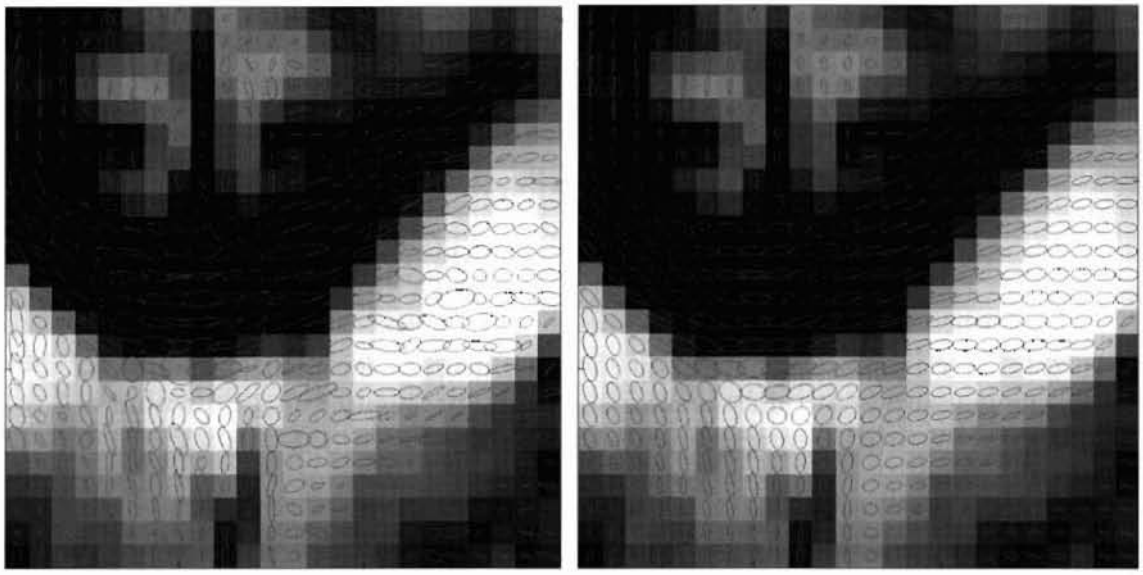
Figure 24.2 shows the result of filtering a synthetic 2D tensor field visualized using ellipses. The original tensor field with added noise is shown in Fig. 24.2(a) and the result in Fig. 24.2(b). In this example, the voxel certainty measure, c_v, was set to one except outside the signal extent where it was set to zero. When filtering tensor data, the angular measure c_a is important since it can be used to reduce mixing of information from regions having different orientations. This is demonstrated in Fig. 24.2(c) using $\alpha = 0$ and in Fig. 24.2(d) with $\alpha = 2$. Notice how the degree of mixing depends on the angular similarity measure.

Figure 24.3 shows a region of interest (ROI) for tensor field generated from DT-MRI data. Figure 24.3(b) shows the result of filtering the DT-MRI tensor field using the proposed method. In this example, the voxel certainty measure, c_v, was set to one except outside the signal extent where it was set to zero. An alternative to this is to use for example Proton Density MRI data defining where the MR signal is reliable. For the angular certainty function, c_a, $\alpha = 4$ was used.



(a) (b) (c)

Fig. 24.2. Original synthetic tensor field with added noise (**a**) and the result using the proposed method using $\alpha = 0$ (**b**) and $\alpha = 2$ (**c**)



(a) (b)

Fig. 24.3. Tensor field generated from noisy DT-MRI data (**a**) and result of filtering the tensor field using the NC method (**b**). The ROI is the corpus callosum, the dark bow-shaped region. See appendix for color plates

24.3 Bayesian Regularization using Multivariate Gaussian Markov Random Fields

The use of MRFs for regularization of tensor fields has gained interest in the last few years [13]. We will discuss a new approach for regularization of DT-MRI data using Markov Random Fields (MRFs) in a Bayesian framework [8, 9, 10]. The MRFs that are going to be described have Gaussian distribution [11]. A description of multivariate Gaussian distributed noise for tensors is presented in [2]. The concept of 'isotropic' noise is defined resulting in a distribution whose covariance has two parameters. Our noise model is not constrained to be isotropic so a non-constrained model for the covariance will be used. In this chapter we define a Bayesian model where MRFs are used for

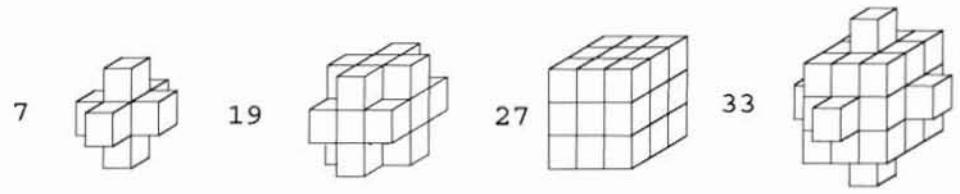


Fig. 24.4. Examples of 3D neighborhood systems $\boldsymbol{\delta}(m,n,p)$ with different spatial extent. The numbers indicate the number of voxels L belonging to each system

achieving a global result from local interactions described by a simple neighborhood model [19]. We will first present a scalar version of the model and later extend this to the tensor case.

24.3.1 Prior Probability Model

Let the (non-observable) scalar volume that we want estimate be described by a random vector $\mathbf{X}$ with K elements. This vector is formed by rearranging the original data with dimensions $M \times N \times P$, and thus giving $K = MNP$ elements.

We assume that the prior probability density function (PDF) of the non-observable scalar field $\mathbf{X}$ is a K-dimensional multivariate Gauss-MRF [11] defined by a $K \times 1$ mean vector $\boldsymbol{\mu}$, and a $K \times K$ covariance matrix $\mathbf{C}$.

As the 3D spatial dependencies are assumed to be strictly local, the covariance matrix $\mathbf{C}$ will be sparse, and thus it will be more practical to work directly with the local characteristics of the field. This is achieved by defining a 3D neighborhood system $\boldsymbol{\delta}(m,n,p)$ for each voxel. Each $\boldsymbol{\delta}(m,n,p)$ is a set of L triplet indices. In Fig. 24.4, four different neighborhoods are shown. Using each index set (one for each voxel site) we can define the sets of neighboring random vectors $\boldsymbol{\delta}X(m,n,p)$ with L elements as

$$\boldsymbol{\delta}X(m,n,p) = \Big\{X(m',n',p'), (m',n',p') \in \boldsymbol{\delta}(m,n,p)\Big\} \tag{24.10}$$

and the same for the squared image values $\boldsymbol{\delta}X^2(m,n,p)$,

$$\boldsymbol{\delta}X^2(m,n,p) = \Big\{X^2(m',n',p'), (m',n',p') \in \boldsymbol{\delta}(m,n,p)\Big\}\,. \tag{24.11}$$

Under these assumptions the local characteristic of the field is given by the following conditional Gauss-MRF[3] [11]

$$p\big(X\big|\boldsymbol{\delta}X\big) = \frac{1}{\sqrt{2\pi}\,\sigma_X} \exp\left\{-\frac{(X-\mu_X)^2}{2\sigma_X^2}\right\} \tag{24.12}$$

[3] For simplicity we are omitting describing the dependence of the voxel indices (m,n,p), unless explicitly specified.

with

$$\mu_X = \frac{1}{L}\sum \delta X \qquad \sigma_X^2 = \frac{1}{L}\sum \delta X^2 - \mu_X^2 \tag{24.13}$$

the prior local mean and variance, respectively, both estimated with the maximum likelihood (ML) method [5]. The summations in (24.13) is over the L elements of the sets δX and δX^2 as defined in (24.10) and (24.11), respectively.

24.3.2 Likelihood Model

In this section we will define a likelihood model for the observed (noisy) signal. Let the observed scalar signal be defined by

$$\mathbf{Y} = \mathbf{X} + \mathbf{N} \tag{24.14}$$

where $\mathbf{X}$ is the wanted (non-observable) signal, and where $\mathbf{N}$ denotes the noise. For simplicity we assume that the noise is independent of the signal $\mathbf{X}$ to be estimated. The likelihood model (also called transition model) is given by the conditional PDF of the observed signal $\mathbf{Y}$ given $\mathbf{X}$, which is assumed to be a K-dimensional multivariate Gauss-MRF with given $K \times 1$ mean vector $\mathbf{X}$ and $K \times K$ covariance matrix $\mathbf{C}_N$.

In the likelihood model we will exploit the spatial dependencies of the MRF to determine the local likelihood characteristic of the scalar data. This is achieved by assuming the following local conditional independence property for the PDF

$$p(Y|X, \delta X) = p(Y|X) \tag{24.15}$$

which is here described by a conditional Gauss-MRF

$$p(Y|X) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left\{ -\frac{(Y-X)^2}{2\sigma_N^2} \right\}, \tag{24.16}$$

where the local noise variance $\sigma_N^2 = \sigma_{Y|X}^2$ is assumed to be homogeneous, that is, independent of the spatial indices (m, n, p). This variance has to be estimated from the observed scalar volume $\mathbf{Y}$. The estimator for the noise variance used in the examples below is

$$\sigma_N^2 = \lambda\sigma_{N_{\text{ave}}}^2 + (1-\lambda)\sigma_{N_{\text{min}}}^2 , \tag{24.17}$$

where $0 \leq \lambda \leq 1$ is a free parameter defining the degree of regularization (low λ gives low regularization). The average local variance $\sigma_{N_{\text{mean}}}^2$ is given by

$$\sigma_{N_{\text{ave}}}^2 = \frac{1}{K}\sum_{m=1}^{M}\sum_{n=1}^{N}\sum_{p=1}^{P}\sigma_Y^2(m,n,p) , \tag{24.18}$$

and where $\sigma_Y^2(m,n,p)$ is given by the ML method following the (24.13) replacing all the X variables with Y. The minimum local variance $\sigma_{N_{\text{min}}}^2$ is given by

$$\sigma^2_{N_{\min}} = \min_{m,n,p} \sigma^2_Y(m,n,p) \tag{24.19}$$

The reasoning for (24.17) is that as the signal and the noise are independent, and the total variance is given by adding the noise and the signal variance. A possible estimator for the noise variance would then be given by a total variance estimator in a region for which the signal is zero. If we consider the local variance as a realization for the total variance, then the minimum among this local variances will be a possible estimator for the noise variance (24.19). However, due to the presence of outliers and that the local variance has significant variability, this minimum will have a bias towards zero. An alternative estimator for the noise variance is given by the average of the local variances (24.18). In this case, due to the presence of signal components in the local variances, this estimator will have a bias towards infinity. A trade-off between these two estimators can be defined by introducing a regularization parameter λ ranging in the interval $(0,1)$ reducing the effect of these biases (24.17). The two estimators in (24.18) and (24.19) naturally define two bounds that can be used to define the maximum and minimum meaningful amount of regularization, and the parameter λ between 0 and 1 can be used as a regularization parameter to tune the desired amount of regularization.

24.3.3 Posterior Probability Modeling

This section describes how to combine the prior probability model for the estimated signal with the likelihood model for the observed signal. Bayes' theorem lets us write the posterior probability density function as

$$p(\mathbf{X}|\mathbf{Y}) = \frac{p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})}{p(\mathbf{Y})}\,, \tag{24.20}$$

where $p(\mathbf{Y})$ depends only on the known signal $\mathbf{Y}$.

In the Gaussian case the maximum a posteriori (MAP) estimation is equal to the minimum mean square error (MMSE) estimation [5] given both by

$$\mathbf{X}_{\text{MAP}} = \arg\max_{\mathbf{X}} p(\mathbf{X}|\mathbf{Y}) = \mathbf{X}_{\text{MMSE}} = E[\mathbf{X}|\mathbf{Y}] = \boldsymbol{\mu}_{\mathbf{X}|\mathbf{Y}} \tag{24.21}$$

with $\boldsymbol{\mu}_{\mathbf{X}|\mathbf{Y}}$ defined by

$$\boldsymbol{\mu}_{\mathbf{X}|\mathbf{Y}} = \mathbf{C_N}\Big(\mathbf{C_X} + \mathbf{C_N}\Big)^{-1}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{C_X}\Big(\mathbf{C_X} + \mathbf{C_N}\Big)^{-1}\mathbf{Y}\,. \tag{24.22}$$

In general, it is not feasible to explicitly determine $\mathbf{C_X}$, $\mathbf{C_N}$ and $\boldsymbol{\mu}_{\mathbf{X}}$, and we will resort to either the Gibbs sampler (GS) algorithm to iteratively find a solution for the MMSE estimation or the simulated annealing (SA) algorithm for the MAP estimation [3, 19]. Those algorithms are based on iteratively visiting all the sites (voxels) by sampling the posterior local characteristic of the field directly (temperature parameter $T = 1$) for the GS algorithm and

using a logarithmic cooling schedule defined by a temperature parameter T for the SA algorithm. The MMSE solution is given by the average of the solutions achieved after each iteration of the GS algorithm and the MAP solution is given by the last iteration of the SA algorithm (low temperature) [19]. In practice, 20 to 50 iterations are enough to get reasonable MMSE solutions and 10 to 20 for MAP solutions.

For determining the posterior local characteristic we will again resort to the Bayes' theorem:

$$p(X|Y,\delta X) = \frac{p(Y|X)p(X|\delta X)}{p(Y|\delta X)} \tag{24.23}$$

where $p(Y|\delta X)$ depends only on the known signal Y. This posterior local characteristic is a Gauss-MRF given by [11]

$$p(X|Y,\delta X) = \frac{1}{\sqrt{2\pi T}\,\sigma_{X|Y}} \exp\left\{ -\frac{(X-\mu_{X|Y})^2}{2T\sigma_{X|Y}^2} \right\} \tag{24.24}$$

with the following posterior local mean and variance

$$\mu_{X|Y} = \frac{\sigma_N^2\,\mu_X + \sigma_X^2 Y}{\sigma_X^2 + \sigma_N^2} \qquad \sigma_{X|Y}^2 = \frac{\sigma_N^2\,\sigma_X^2}{\sigma_X^2 + \sigma_N^2}\,. \tag{24.25}$$

These expressions show that in the Gaussian case both the posterior local mean $\mu_{X|Y}$ and the posterior local variance $\sigma_{X|Y}^2$ can be obtained by closed form expressions, which greatly simplifies the implementation of the method, as well as reduces the computational burden. The proof for (24.25) is rather involved. Some hints on how to prove it can be found in [5], the core of the proof is that one quadratic form is to be constructed from a *summation* of two quadratic forms coming from the prior and likelihood PDFs.

To determine the performance of the described method a synthetic data volume was constructed. The volume is defined as a sphere with a one period sawtooth radial profile. Figure 24.5(a) shows the middle slice for this data, Fig. 24.5(b) the same volume with added noise, Fig. 24.5(c) the estimated MAP solution and Fig. 24.5(d) the MMSE solution. The regularization parameter λ was set to 0.5. Notice that as the method is designed for continuous signals, it cannot deal well with discontinuities. This is noticeable in Fig. 24.5(c) and 24.5(d) where the voxels close to the boundary have not been filtered. The use of an edge model as proposed in [19] may be helpful. Such models has the potential to filter the boundary voxels along the boundaries without inter region blurring. Comparing Fig. 24.5(c) to Fig. 24.5(d), we can see that the variability of the MAP solution is lower than the variability of the MMSE. However the bias of the MAP solution is greater than the bias for the MMSE. In general, the MAP method converges faster than the MMSE. To quantitatively compare these results, we have computed the signal to noise ratio (SNR). For the noisy data shown in Fig. 24.5(b) the SNR is 4.1, for the

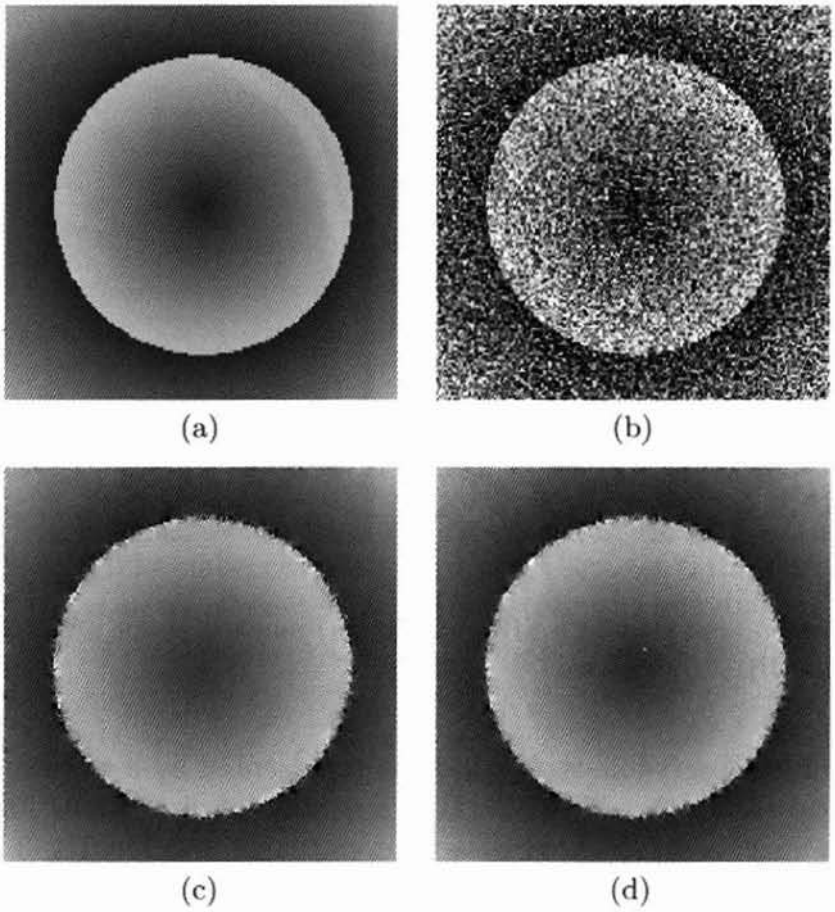


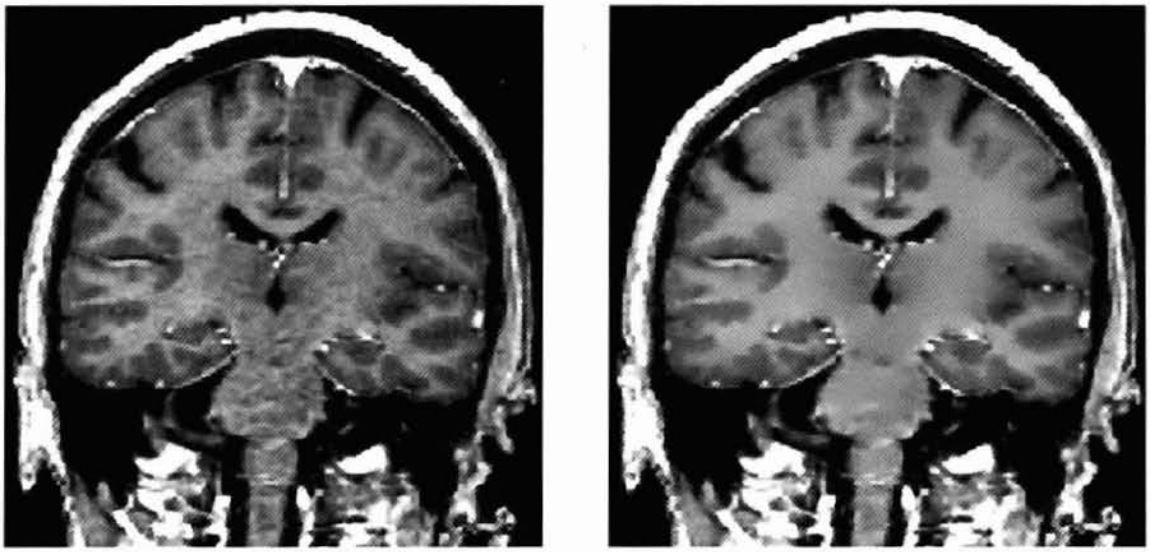
Fig. 24.5. Synthetic data volume (**a**), with added noise (**b**), the result using the MAP estimator (**c**) and the MMSE estimator (**d**)

MAP solution shown in Fig. 24.5(c) the SNR is 150.1 and for the MMSE solution shown in Fig. 24.5(d) the SNR is 184.6. Measuring the SNR only close to the discontinuity gives for the noisy data 4.8, for the MAP 18.6, and for the MMSE 21.7. The SNR values in the continuous areas are as follows: for the noisy 4.0, for the MAP 674.0 and for the MMSE 949.1. The SNR is better for the MMSE than for the MAP estimator due to the bias of the latter. As expected, the model performs better in the homogeneous areas.

Figure 24.6(a) shows a coronal view for MRI data set of a human brain and in Fig. 24.6(b) the result after applying the proposed regularization scheme. The regularization parameter λ was set to 0.2. The noise is reduced in the homogeneous regions without smoothing the boundaries.

24.3.4 Multi-Component Model Extension

In this section we extend the scalar case to multi-component case. For DT-MRI we have a second-rank tensor at each voxel position which is being represented by a symmetric 3×3 tensor matrix.



(a) (b)

Fig. 24.6. A coronal slice of a MRI volume (**a**) and the result after applying the 3D Gaussian MRF scheme (**b**)

Let the non-observable tensor field be represented by a random matrix $\mathbf{X}$ with dimensions $3 \times 3 \times M \times N \times P$ of a volume of dimensions $M \times N \times P$, where each voxel is represented by a random tensor given by the 3×3 matrix

$$\mathbf{X}(m,n,p) = \begin{pmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{pmatrix} . \tag{24.26}$$

Each tensor matrix $\mathbf{X}(m,n,p)$ is symmetric so that $X_{ij} = X_{ji}$, and has 6 different random variables. The total number of different random variables is thus $K = 6MNP$. We further define a rearrangement matrix operator **LT** (lower triangular part) in order to extract the 6 different elements at each voxel position and to regroup them as a 6×1 column vector, $\mathbf{X}_{LT}(m,n,p)$ as

$$\mathbf{X}_{LT}(m,n,p) = \mathbf{LT}\big[\mathbf{X}(m,n,p)\big] = (X_{11}\ X_{21}\ X_{31}\ X_{22}\ X_{32}\ X_{33})^T . \tag{24.27}$$

By applying the rearrangement operator to each tensor, a 4D random matrix $\mathbf{X}_{LT}$ with dimensions $6 \times M \times N \times P$ is obtained without repeated elements. In order to formulate the probability density function of the tensor field we need to rearrange the tensor field $\mathbf{X}_{LT}$ as a column vector, we define a second rearrangement operator **CV** (column vector) as

$$\mathbf{X}_{CV} = \mathbf{CV}\big[\mathbf{X}_{LT}\big] \tag{24.28}$$

making $\mathbf{X}_{CV}$ a $K \times 1$ random vector which represents the whole tensor field.

In order to calculate the MAP or the MMSE estimator for the tensor field $\mathbf{X}_{CV}$ we need to extend the prior model presented in Sect. 24.3.1, the likelihood model in Sect. 24.3.2 and the posterior model in Sect. 24.3.3 by using

6-dimensional multivariate Gaussian distribution instead of the 1-dimensional Gaussian distribution explained for the scalar case. This should be done using the local characteristics of the field defined for each 6-component vector variable $\mathbf{X}_{LT}(m,n,p)$. A short list of what is needed to be estimated in the method follows:

- For the prior PDF we need to estimate the vector means $\boldsymbol{\mu}_{\mathbf{X}_{LT}}(m,n,p)$ and the covariance matrices $\mathbf{C}_{\mathbf{X}_{LT}}(m,n,p)$ by using the neighboring sites.
- For the likelihood PDF we need to estimate the noise covariance matrix $\mathbf{C}_{\mathbf{N}_{LT}}$, which does not depend on the site indices (m,n,p) as the noise is assumed to be homogeneous. This noise covariance matrix can be estimated in a similar way as the one proposed in Sect. 24.3.2 for the scalar case, but modified accordingly to deal with covariance matrices instead of variances.
- For the posterior PDF using the prior parameters $\boldsymbol{\mu}_{\mathbf{X}_{LT}}(m,n,p)$ and $\mathbf{C}_{\mathbf{X}_{LT}}(m,n,p)$, the noise covariance matrix $\mathbf{C}_{\mathbf{N}_{LT}}$ and the observed noisy field $\mathbf{Y}_{LT}(m,n,p)$ we can determine the posterior means $\boldsymbol{\mu}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}(m,n,p)$ and the posterior covariances $\mathbf{C}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}(m,n,p)$ by means of[4]

$$\boldsymbol{\mu}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}} = \mathbf{C}_{\mathbf{N}_{LT}}\left(\mathbf{C}_{\mathbf{X}_{LT}} + \mathbf{C}_{\mathbf{N}_{LT}}\right)^{-1}\boldsymbol{\mu}_{\mathbf{X}_{LT}} + \mathbf{C}_{\mathbf{X}_{LT}}\left(\mathbf{C}_{\mathbf{X}_{LT}} + \mathbf{C}_{\mathbf{N}_{LT}}\right)^{-1}\mathbf{Y}_{LT} \tag{24.29}$$

and

$$\mathbf{C}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}} = \mathbf{C}_{\mathbf{X}_{LT}}\left(\mathbf{C}_{\mathbf{X}_{LT}} + \mathbf{C}_{\mathbf{N}_{LT}}\right)^{-1}\mathbf{C}_{\mathbf{N}_{LT}} \,. \tag{24.30}$$

These equations are a vector generalization of the ones given by (24.25) in Sect. 24.3.3. In [5] we can also find some hints on how to prove that equations. In the vector case, special care has to be taken, however the procedure is similar to the scalar case.

For the SA algorithm we need to sample the posterior local characteristic, that is a 6-dimensional multivariate Gaussian distribution with parameters $\boldsymbol{\mu}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$ and $\mathbf{C}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$. This can be done by first generating a 6×1 sample vector $\mathbf{U}$ whose elements are independent and have Gaussian distribution with zero mean and unit standard deviation, and then transforming this sample according to

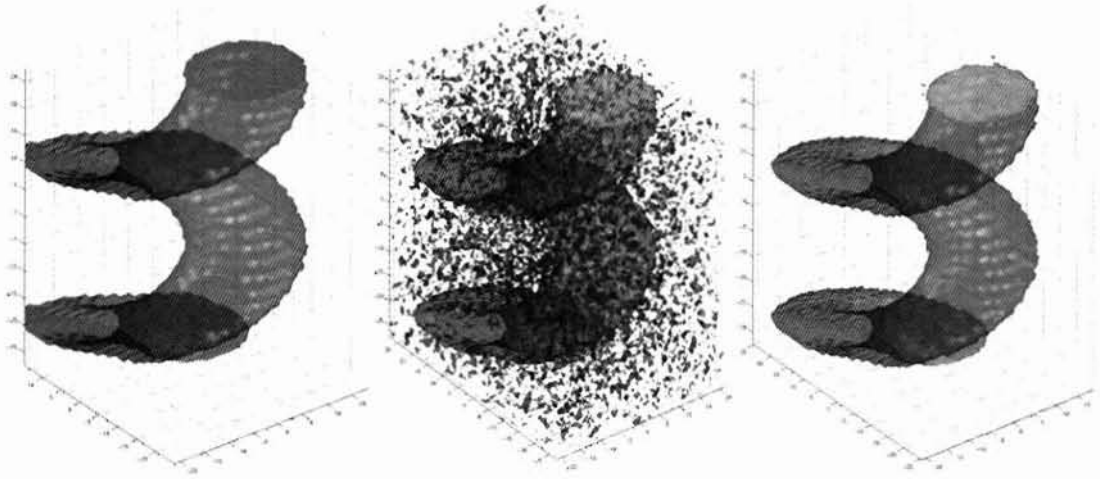
$$\mathbf{X}_{LT} = \sqrt{T}\,\mathbf{D}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}\mathbf{U} + \boldsymbol{\mu}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}} \tag{24.31}$$

where $\mathbf{D}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$ is a matrix given by[5]

$$\mathbf{D}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}} = \mathbf{Q}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}\sqrt{\boldsymbol{\Lambda}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}} \tag{24.32}$$

[4] We omit the indices (m,n,p) to enhance the readability of the equations.

[5] The square root of a diagonal matrix is a diagonal matrix whose elements are the square root of the corresponding original matrix.



(a) (b) (c)

Fig. 24.7. Original synthetic tensor field (**a**), with noise added (**b**) and regularization result (**c**)

where the matrix $\mathbf{Q}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$ has in its columns the eigenvectors of the covariance matrix $\mathbf{C}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$. The matrix $\boldsymbol{\Lambda}_{\mathbf{X}_{LT}|\mathbf{Y}_{LT}}$ is diagonal having the corresponding eigenvalues in the principal diagonal.[6]

In order to assure the positive semidefinite condition of the diffusion tensors, the SA algorithm is modified as follows: after visiting a voxel, the condition is tested; if the test is not passed then the tensor is discarded and sampled again until the condition is satisfied.

We have generated a 3D helix synthetic tensor field for which the internal tensors are anisotropic and the external tensors are isotropic. Figure 24.7(a) shows the surface corresponding to the boundary separation between the two tensor classes. We have added noise to that tensor field and repeated the same visualization obtaining the Fig. 24.7(b). The random spots outside the helix surface correspond to anisotropic tensors and the internal random holes to isotropic tensors. After applying the proposed regularization method, although the field is not completely homogeneous, the two tensor classes are recovered as shown in Fig. 24.7(c). The regularization parameter λ was set to 0.5.

We have also regularized a DT-MRI data volume of a monkey brain. The results can be shown in Figs. 24.8, 24.9 and 24.10. The regularization parameter λ was set to 0.05. Figure 24.8(a) shows an axial view of the original noisy tensor described by the trace of the tensors, and Fig. 24.8(b) the result after regularization. Note that the main structures are maintained while the background noise in the flat regions have been removed. Figure 24.8(c) shows the fractional anisotropy (FA) measure [1] for the original data and Fig. 24.8(d)

[6] A real symmetric positive semidefinite matrix $\mathbf{C}$ can be factorized as $\mathbf{C} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T = \mathbf{D}\mathbf{D}^T$, where $\mathbf{D} = \mathbf{Q}\sqrt{\boldsymbol{\Lambda}}$. A covariance matrix is always real, symmetric and positive semidefinite.

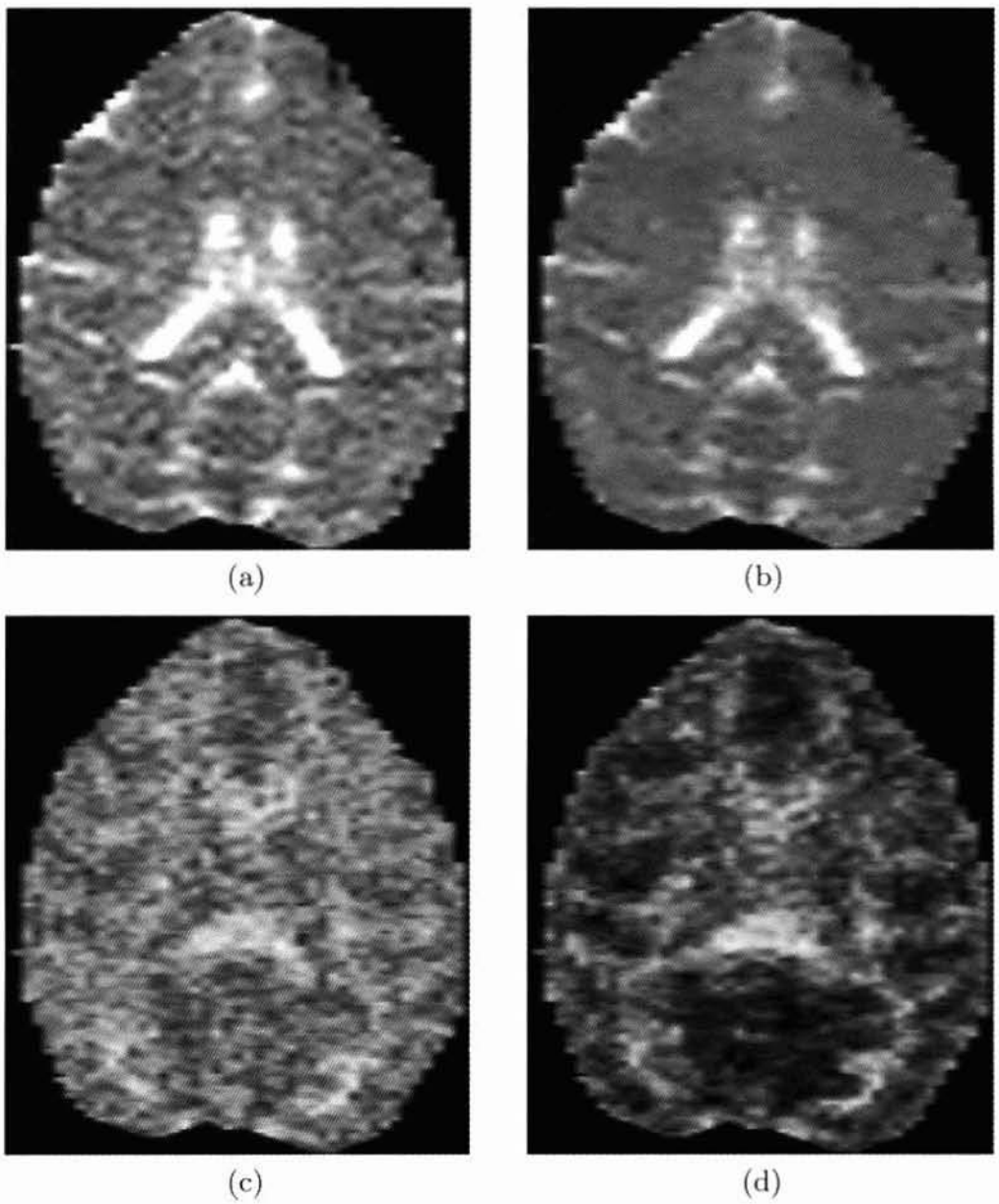


Fig. 24.8. Axial slice of the tensor trace for the original noisy DT-MRI data from a monkey brain (**a**), trace after regularization (**b**), axial slice of the FA measure for the original noise DT-MRI data (**c**) and FA measure after regularization (**d**)

after regularization. The higher the FA value, the more likely the presence of white matter fiber tracts. Notice that after regularization the white matter structures are better defined. In Fig. 24.9(a) a glyph visualization using superquadratics [6] for same data is shown. A description of these glyphs can be found in Chap. 7. Figure 24.9(b) shows the results after regularization. Notice that fiber tracts can be better distinguished in the filtered data set. To further explore the result of regularization we present results from tractography in the corpus callosum. Figure 24.10(a) displays the result of tractography in the original data and Fig. 24.10(b) in the regularized data. Notice that the fiber tracts are more continuous and better grouped in the filtered data.

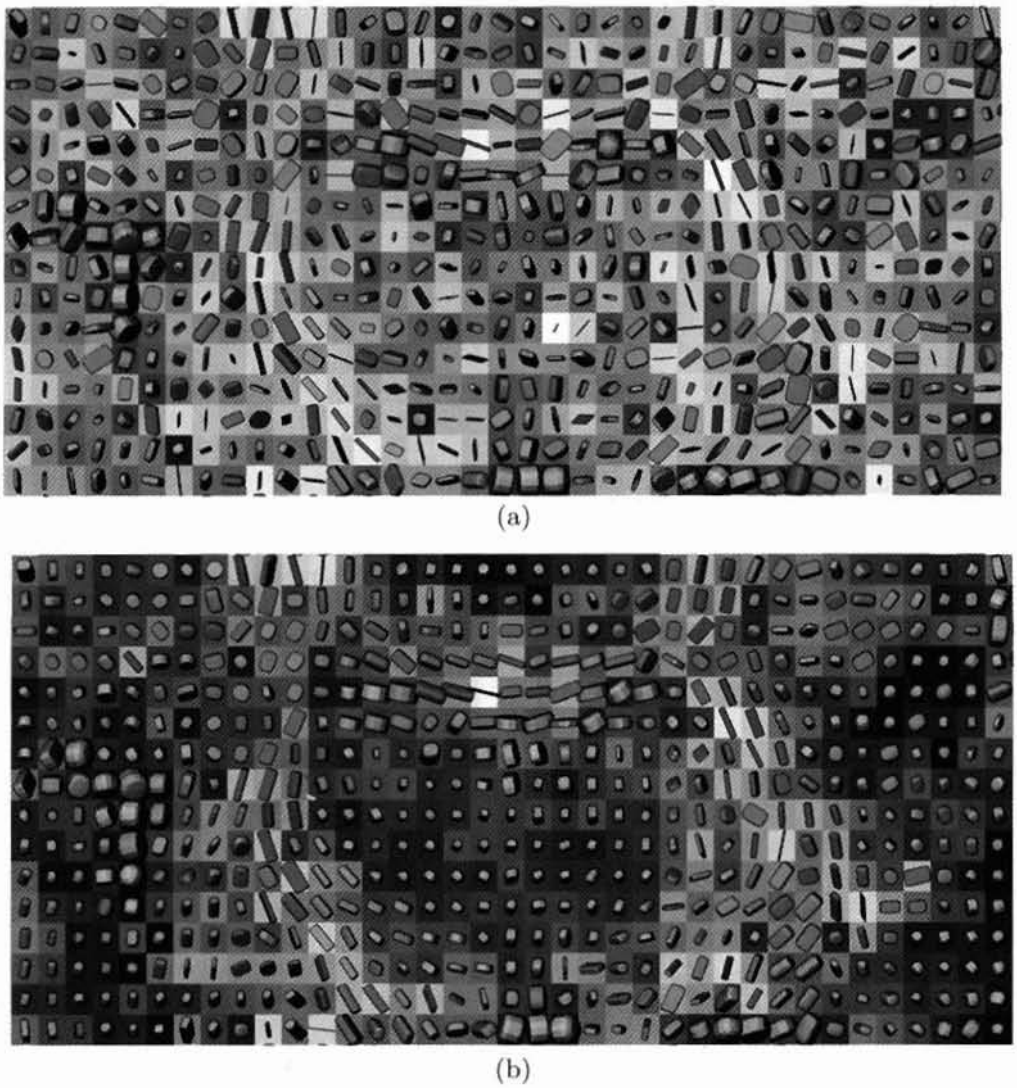


Fig. 24.9. Glyph visualization for a coronal slice for the original noisy DT-MRI data from a monkey brain (**a**) and the same visualization after regularization (**b**). See appendix for color plates

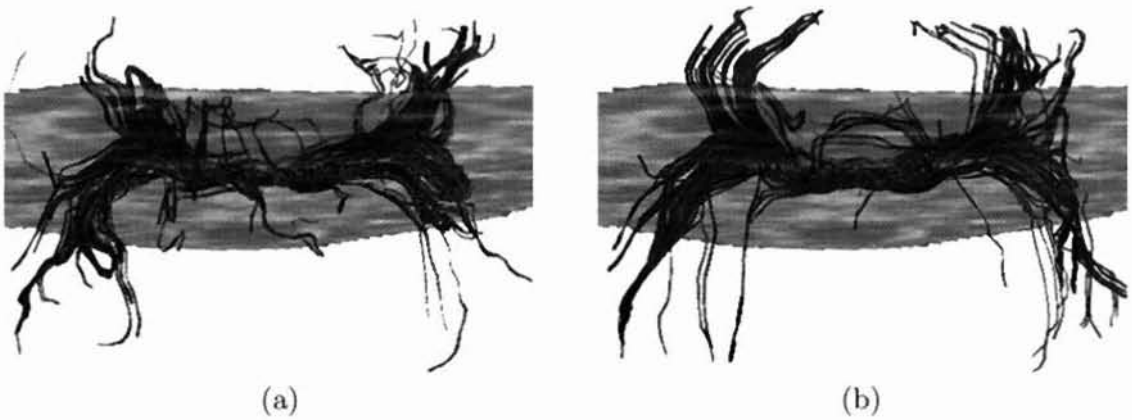


Fig. 24.10. Tractography in the corpus callosum for the original noise DT-MRI data from a monkey brain (**a**) and the same after regularization (**b**). See appendix for color plates

24.4 Conclusion

In this chapter we have described two different regularization methods. First we described NC, a general method for filtering missing and uncertain data. This filtering method solves a WLS problem by applying filters in a basis set which are assumed to span the signal locally. In NC, a certainty is defined for the signal, and localization of the basis functions is achieved by an applicability function, the corresponding certainty function for the basis filters. The effect of this is that spatial localization of the filter is achieved by a certainty window and not by changing the filter coefficients as in traditional windowing. The strengths of NC are the power of signals/certainty framework, and the use of basis functions that can be defined to span a subspace locally that fits the signal.

Second, we described a purely stochastic approach to regularization based on Gaussian MRF modeling of the signal. The strength of this approach is the Bayesian framework which enables an efficient control of the trade-off of applying the model (here the Gaussian model) where it fits, and not applying it in areas where it does not. The presented method can be improved by introducing boundary models for the prior at an expense of increasing the computational load.

Future work combining the best of the two approaches presented in this chapter will likely provide interesting and useful results.

Acknowledgments

This work was funded by NIH grants P41-RR13218 and R01-MH 50747, CIMIT, Spanish CICYT for the research grant TIC2001-3808-C02-02, the European Commission for the funding associated to the Network of Excellence SIMILAR FP6-5076609. The second author acknowledges the Fulbright Commission for the postdoctoral grant FU2003-0968.

References

1. P. J. Basser. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine*, 8:333–344, 1995.
2. P. J. Basser and S. Pajevic. A normal distribution for tensor-valued random variables: Application to diffusion tensor MRI. *IEEE Trans. Medical Imaging*, 22:785–794, 2003.
3. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 6:721–741, 1984.
4. H. Gudbjartsson and S. Patz. The rician distribution of noisy msri data. *Magnetic Resonance in Medicine*, 24:910–914, 1995.

5. S. M. Kay. *Fundamentals of statistical signal processing: estimation theory.* Prentice-Hall, 1993.
6. G Kindlmann. Superquadric tensor glyphs. In *IEEE TVCG/EG Symposium on Visualization*, pp. 147–154, 2004.
7. H. Knutsson and C.-F. Westin. Normalized and differential convolution: Methods for interpolation and filtering of incomplete and uncertain data. In *Computer Vision and Pattern Recognition*, pp. 515–523, 1993.
8. M. Martin-Fernandez, C. Alberola-Lopez, J. Ruiz-Alzola, and C. F. Westin. Regularization of diffusion tensor maps using a non-gaussian markov random field approach. In *Medical Image Computing and Computer-Assisted Intervention*, volume 2879 of *Lecture Notes in Computer Science*, pp. 92–100, 2003.
9. M. Martin-Fernandez, R. San Jose-Estepar, C. F. Westin, and C. Alberola-Lopez. A novel gauss-markov random field approach for regularization of diffusion tensor map. In *Computer Aided Systems Theory*, volume 2809 of *Lecture Notes in Computer Science*, pp. 506–517, 2003.
10. M. Martin-Fernandez, C.-F. Westin, and C. Alberola-Lopez. 3d bayesian regularization of diffusion tensor mri using multivariate gaussian markov random fields. In *Medical Image Computing and Computer-Assisted Intervention*, volume 3216 of *Lecture Notes in Computer Science*, pp. 351–359, 2004.
11. J. M. F. Moura and S. Goswami. Gauss markov radom fields (gmrf) with continuous indices. *IEEE Trans. Information Theory*, 43:1560–1573, 1997.
12. C. Pierpaoli, P. Jezzard, P. J. Basser, A. Barnett, and G. Di Chiro. Diffusion tensor MR imaging of the human brain. *Radiology*, 201:637, 1996.
13. C. Poupon, C. A. Clark, F. Frouin, J. Régis, I. Bloch, D. Le Bihan, I. Bloch, and J.-F. Mangin. Regularization of diffusion-based direction maps for the tracking brain white matter fascicles. *NeuroImage*, 12:184–195, 2000.
14. C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *International Conference on Computer Vision*, pp. 839–846, 1998.
15. D. Tschumperlé and R. Deriche. Dt-mri images: Estimation, regularization and application. In *Computer Aided Systems Theory*, volume 2809 of *Lecture Notes in Computer Science*, pp. 530–541, 2003.
16. C.-F. Westin. *A Tensor Framework for Multidimensional Signal Processing.* PhD thesis, Linköping University, Sweden, 1994.
17. C.-F. Westin and H. Knutsson. Tensor field regularization using normalized convolution. In *Computer Aided Systems Theory*, volume 2809 of *Lecture Notes in Computer Science*, pp. 564–572, 2003.
18. C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis. Processing and visualization of diffusion tensor MRI. *Medical Image Analysis*, 6:93–108, 2002.
19. G. Winkler. *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods, Applications of Mathematics, Stochastic Modelling and Applied Probability.* Springer Verlag, 2003.

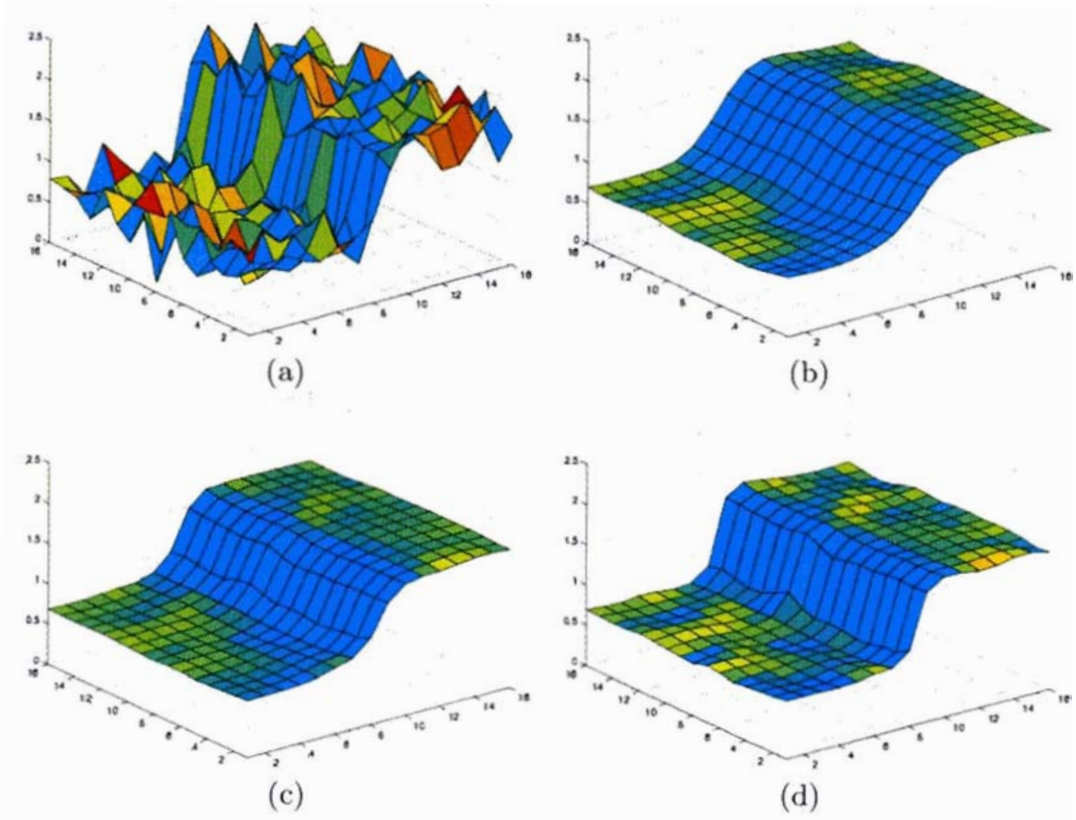
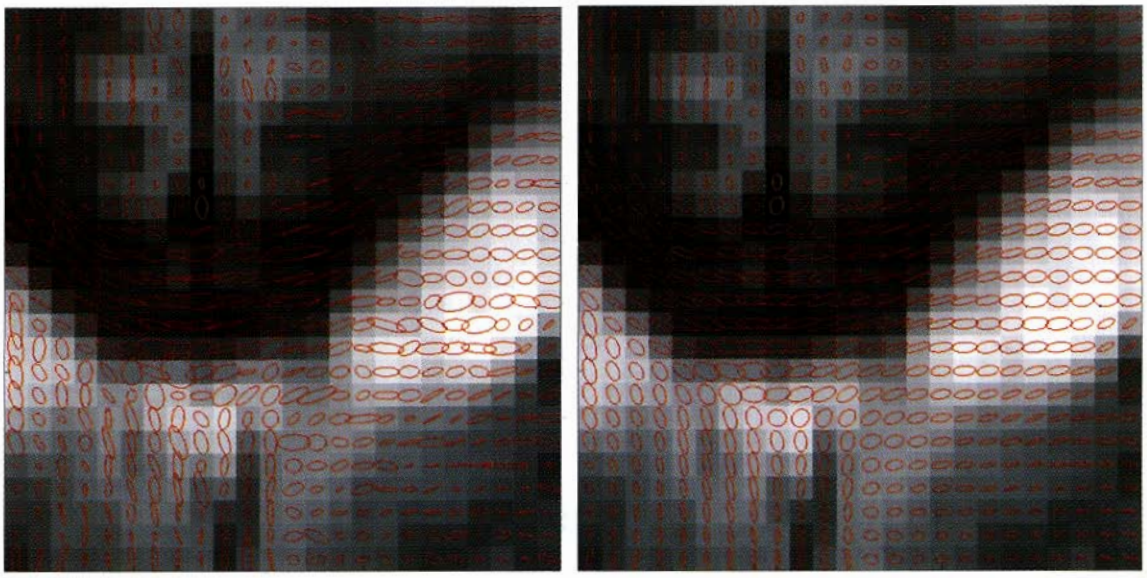
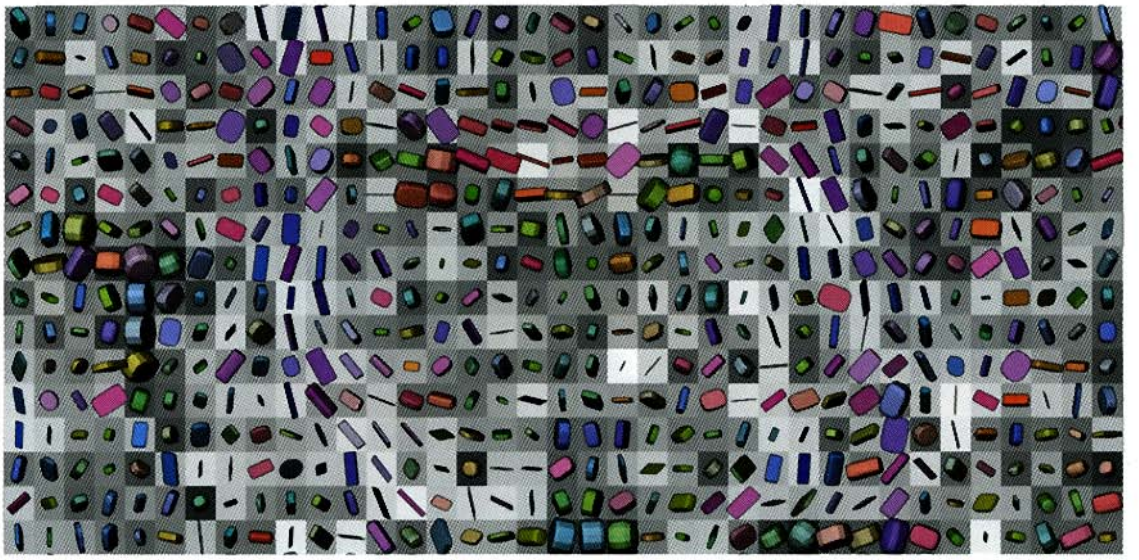


Fig. 24.1. Original scalar field (step edge) with added noise (a), the result without using the magnitude certainty measurement c_m (b), results using c_m with $\sigma = 1$ (c) and with $\sigma = 0.5$ (d)

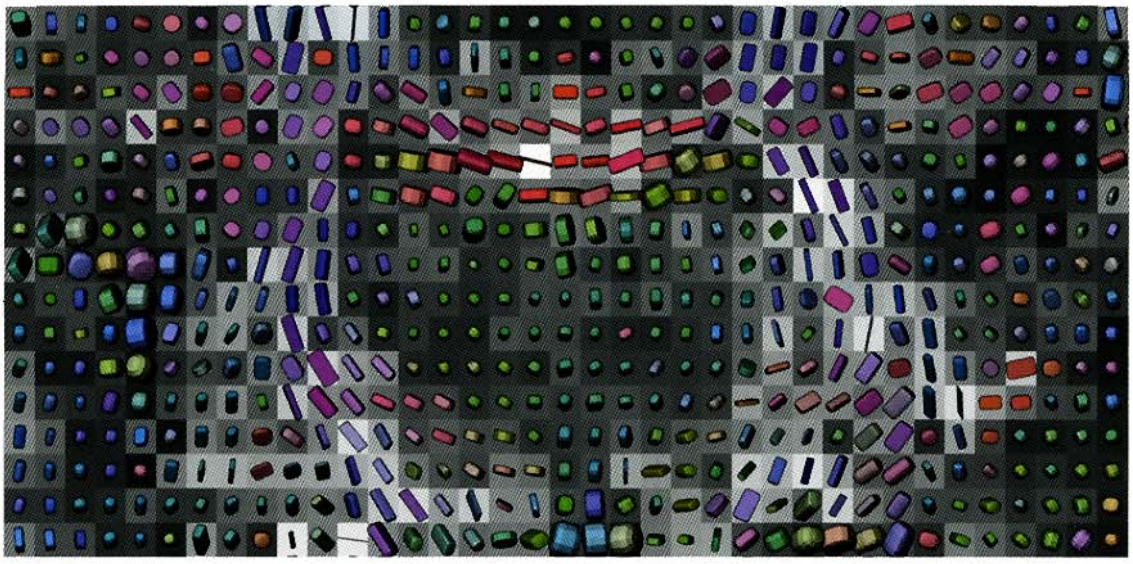


(a) (b)

Fig. 24.3. Tensor field generated from noisy DT-MRI data (**a**) and result of filtering the tensor field using the NC method (**b**). The ROI is the corpus callosum, the dark bow-shaped region

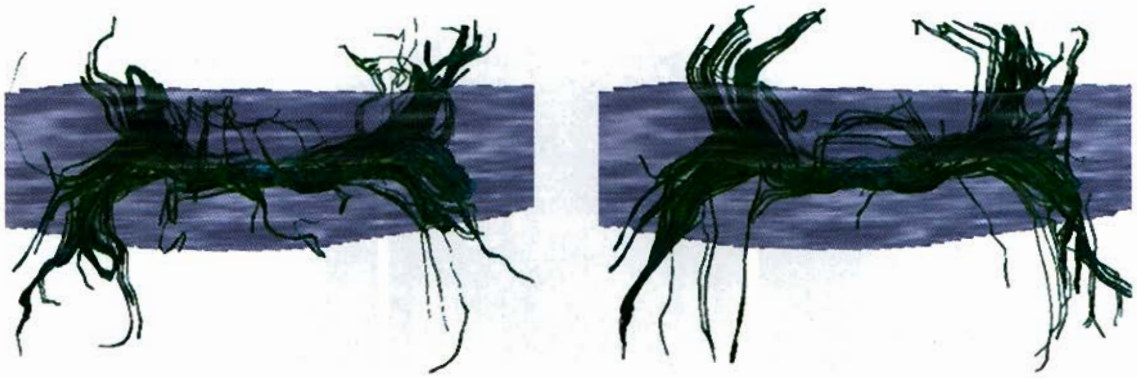


(a)



(b)

Fig. 24.9. Glyph visualization for a coronal slice for the original noisy DT-MRI data from a monkey brain (**a**) and the same visualization after regularization (**b**)



(a) (b)

Fig. 24.10. Tractography in the corpus callosum for the original noise DT-MRI data from a monkey brain (**a**) and the same after regularization (**b**)