Hindawi Publishing Corporation
International Journal of Biomedical Imaging
Volume 2007, Article ID 27432, 11 pages
doi:10.1155/2007/27432

Research Article

Total Variation Regularization of Matrix-Valued Images

Oddvar Christiansen,[1] Tin-Man Lee,[2] Johan Lie,[1] Usha Sinha,[2] and Tony F. Chan[3]

[1] *Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Bergen, 5008 Bergen, Norway*
[2] *Medical Imaging Informatics Group, Department of Radiological Sciences, University of California, Los Angeles, CA 90024, USA*
[3] *Division of Physical Sciences, College of letters science, University of California, Los Angeles, CA 90095, USA*

Received 30 October 2006; Accepted 13 March 2007

Recommended by Hongkai Zhao

We generalize the total variation restoration model, introduced by Rudin, Osher, and Fatemi in 1992, to matrix-valued data, in particular, to diffusion tensor images (DTIs). Our model is a natural extension of the color total variation model proposed by Blomgren and Chan in 1998. We treat the diffusion matrix D implicitly as the product $D = LL^T$, and work with the elements of L as variables, instead of working directly on the elements of D. This ensures positive definiteness of the tensor during the regularization flow, which is essential when regularizing DTI. We perform numerical experiments on both synthetical data and 3D human brain DTI, and measure the quantitative behavior of the proposed model.

Copyright © 2007 Oddvar Christiansen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. INTRODUCTION

Image processing methods using variational calculus and partial differential equations (PDEs) have been popular for a long time in the image processing research community. Among popular PDE methods are the anisotropic diffusion method proposed by Perona and Malik [1], the total variation method introduced by Rudin et al. [2], and various methods related to these [3–7]. Many of these methods were originally introduced for scalar-valued (gray-scale) images, and were later generalized to vector-valued (color) images.

During the last decade or so, a new magnetic resonance modality called diffusion tensor imaging (DTI) has been extensively studied [8–13]. Using DTI, it is possible to study anatomical structures like the nerve fibers in the human brain noninvasively. The DTI images are matrix valued. In each voxel of the imaging domain, we construct a *diffusion tensor* (i.e., *diffusion matrix*) D based on a series of K direction-specific MR measurements $\{S_k\}_{k=1}^K$. The matrix $D \in R^{3\times 3}$ is a symmetric, positive definite matrix

$$D = V\Lambda V^{-1}, \tag{1}$$

where V is an orthogonal matrix, and Λ is a diagonal matrix with positive elements. We may look at the diffusion matrix as a hyperellipse where the eigenvectors $\{V_i\}_{i=1}^3$ span the ellipsoid, and the corresponding eigenvalues $\{\lambda_i\}_{i=1}^3$ determine the length of each semiaxis (see Figure 1). It is customary to arrange the eigenvalues in decreasing order. By the diffusion tensor model we assume that the set of images $\{S_k\}_{k=1}^K$ is related to the nonweighted image S_0 by the Stejskal-Tanner equation [14, 15]

$$S_k = S_0 e^{-bg_k^T D g_k}, \quad k = 1, 2, \ldots, K. \tag{2}$$

Here $g_k \in R^3$ denotes the direction associated with S_k, and $b > 0$ is a scalar which among other factors depends on the acquisition time and the strength of the magnetic field [16]. Since $D \in R^{3\times 3}$ is symmetric, it has six degrees of freedom. Thus at least six measurements $\{S_k\}_{k=1}^6$ are required to estimate the tensor, as well as the nonweighted measurement S_0. The tensor D can be estimated from (2). In the case of more than six measurements S_k, we can estimate D by, for example, a least-squares minimization. From the eigenvalue decomposition of the diffusion tensor, we can reveal properties like the dominant diffusion direction and the anisotropy of diffusing water molecules [17]. This information can be used to construct maps of the anatomy of the brain.

From the developments in DTI, a need for robust regularization methods for matrix-valued images has emerged. As far as the authors are aware of, there exists no state-of-the-art method for regularization of tensor-valued images, although many methods have been proposed [18–21].

All measurements $\{S_k\}_{k=1}^K$ contain noise, which degrades the accuracy of the estimated tensor. Compared with conventional MR, direction-sensitive acquisitions have a lower

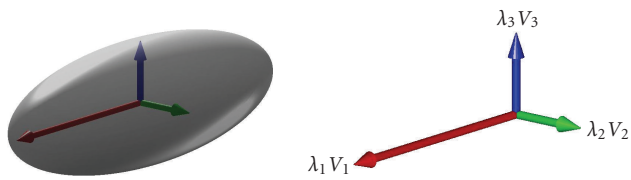


Figure 1: The diffusion matrix D can be represented by a diffusion ellipsoid, where the semiaxes are spanned by the eigenvectors $\{V_i\}_{i=1}^3$ of D, and the length of each semiaxis is given by the eigenvalues $\{\lambda_i\}_{i=1}^3$. In this illustration, the diffusion is anisotropic. The principal diffusion direction is along eigenvector V_1.

signal-to-noise ratio (SNR). Thus the gradient weighted images $\{S_k\}_{k=1}^K$ contain more noise than S_0. There are several ways to increase the accuracy of the estimated tensor. The most intuitive way is to make an average of a series of repeated measurements. Alternatively, we can increase the number of gradient directions. An obvious disadvantage of both of these approaches is the increased scanner time. Perhaps a better way to improve the quality of the tensor is by postprocessing the data. In this paper, we follow this approach, by introducing a regularization method for tensor-valued data.

Since D models diffusion, regularization methods in DTI must preserve the positive definiteness of D. A positive definite matrix has only positive eigenvalues, which is necessary from the physical modeling perspective. In a minimization method for regularization of the tensor data, one possible way to ensure positive definiteness would be to impose a constraint on the minimization problem. Then the constrained problem would have a solution which is on the manifold of positive definite matrices. Regularization of tensor-valued data constrained to manifolds has been studied during the last couple of years, see [22–24]. We however follow a different strategy. Using essentially the same idea as Wang et al. did in a slightly different setting, we treat D implicitly by writing D as the product $D = LL^T$, where L is a lower triangular matrix [18]. Every symmetric positive definite (SPD) matrix has a factorization on this form. We will in this work develop a regularization method for diffusion tensor images by generalizing methods previously developed for scalar- and vector-valued images [2, 25].

Before we go into details of the proposed method, we briefly introduce the total variation (TV) methods for scalar- and vector-valued images. During the last 15 years or so, TV models have undergone extensive studies, initiated by the work of Rudin, Osher, and Fatemi (ROF) [2].

Define the total variation (TV) seminorm for scalar-valued data as

$$\mathrm{TV}[u] = \int_\Omega |\nabla u| dx. \tag{3}$$

Throughout this paper, ∇ denotes the spatial gradient, while $\nabla\cdot$ denotes the divergence operator. In the ROF model, the TV seminorm with an added L_2 fidelity norm is minimized

$$\min_u \left\{ G(u, f, \lambda) = \mathrm{TV}[u] + \frac{\lambda}{2}\|u - f\|_2^2 \right\}. \tag{4}$$

Note that we can write the functional G more abstractly as

$$G(u, f, \lambda) = R(u) + \frac{\lambda}{2} F(u, f), \tag{5}$$

where $R(u)$ is a regularization functional and $F(u, f)$ is a fidelity functional. The regularization term is a geometric functional measuring smoothness of the estimated solution. The fidelity term is a measure of fitness of the estimated solution compared to the input data. It is customary to measure the fidelity in the sense of least squares. Equation (4) has the corresponding Euler-Lagrange equation

$$\partial_u G = -\nabla \cdot \left(\frac{\nabla u}{|\nabla u|} \right) + \lambda(u - f). \tag{6}$$

We can find a minimum of (4) by searching for a steady state of

$$\frac{\partial u}{\partial t} = -\partial_u G, \tag{7}$$

which is the way the ROF model was first formulated [2]. Alternatively, we can directly attack the zero of the Euler-Lagrange equation

$$-\partial_u G = 0, \tag{8}$$

for example, by a fixed-point iteration [26]. This is in general less time consuming than solving the equation using the method of steepest descent, but more tedious to carry out numerically. When we generalize the method to matrix-valued images, we solve the minimization problem by the method of steepest descent. Various methods have been proposed to generalize the ROF model to vector-valued image regularization. Among the successful methods, we find the color TV model developed by Blomgren and Chan [25] and the model by Sapiro [27]. Blomgren and Chan [25] generalized the ROF model to color (vector) image regularization using a set of coupled equations

$$\left\{ \frac{\partial u_i}{\partial t} = \alpha_i \nabla \cdot \left(\frac{\nabla u_i}{|\nabla u_i|} \right) - \lambda(u_i - f_i),\ i = 1, 2, 3 \right\} \tag{9}$$

with

$$\alpha_i = \frac{\mathrm{TV}[u_i]}{\mathrm{TV}[\mathbf{u}]}, \quad i = 1, 2, 3,$$
$$\mathrm{TV}[\mathbf{u}] = \sqrt{\sum_{i=1}^{3} \mathrm{TV}[u_i]^2}. \tag{10}$$

The weight α_i in (9) acts as a coupling between the geometric part of the three image channels. In this work, we extend in a natural way the color TV model of Blomgren and Chan to a matrix TV model. However, the method we propose is not restricted to our choice of regularization functional (TV). For a detailed treatment of TV regularization methods we refer the reader to the recent book by Chan and Shen [5].

In Section 2, we define the minimization problem that we propose to solve, and arrive at the Euler-Lagrange equations

corresponding to this minimization problem. We perform numerical experiments in Section 3, before we finish the paper in Section 4 with a conclusion. Details on the Euler-Lagrange equation and the numerical implementation are given in the appendix at the end of the paper.

2. MINIMIZATION PROBLEM

In this section, we introduce the functional that we minimize in order to regularize tensor-valued data. Let L be a lower triangular matrix. We define D by

$$D = LL^T. \tag{11}$$

This has immediate implications on D: symmetry, positive definiteness, and orthogonality of eigenvectors. These properties are required by the diffusion tensor model. Thus (11) is a natural choice. We define ℓ_{ij} as the element in the ith row and jth column of L. The elements d_{ij} are defined in the same manner.

Let us look at the algebraic equation expressing D as a function of ℓ_{ij}. We derive the expressions for $D \in R^{3\times 3} \subset$ SPD. We explicitly write out the matrix multiplication (11),

$$D = \begin{pmatrix} \ell_{11}^2 & \ell_{11}\ell_{21} & \ell_{11}\ell_{31} \\ \ell_{11}\ell_{21} & \ell_{21}^2 + \ell_{22}^2 & \ell_{21}\ell_{31} + \ell_{22}\ell_{32} \\ \ell_{11}\ell_{31} & \ell_{21}\ell_{31} + \ell_{22}\ell_{32} & \ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2 \end{pmatrix}. \tag{12}$$

In our proposed model, we solve a minimization problem in terms of ℓ_{ij}. For each unique ℓ_{kl}, we minimize

$$\min_{\ell_{kl}} \left\{ \sqrt{\sum_{ij} \mathrm{TV}[d_{ij}(\ell_{kl})]^2} + \frac{\lambda}{2} \sum_{ij} \|d_{ij} - \hat{d}_{ij}\|_2^2 \right\}, \tag{13}$$

where $\{kl\} \in \{11, 21, 22, 31, 32, 33\}$ and $\hat{d}_{ij}$ denotes the elements of the tensor estimated from the noisy data. As in the scalar model, the functional (13) has the abstract form (5). The scalar ROF ($\mathrm{TV} - L_2$) functional is convex. But when we introduce the factorization (11) into the model, we cannot expect the functional (13) to be convex or even quasiconvex. However, from numerical experiments where we used different (random) intial conditions we ended up with almost exactly the same solution. This means that even though we are not able to prove that the functional is convex, we have indications that it is at least quasiconvex.

We note that minimizing the functional (13) is related to the functional used by Wang et al. [18]. Apart from the fact that they simultaneously estimate and regularize the tensor, there are fundamental differences between our proposed regularization functional and the functional proposed by Wang et al. Even though we represent the diffusion matrix on the form of a Cholesky factorization, we regularize the elements of the full diffusion tensor D, while Wang et al. regularize the elements of the lower triangular matrix L. Intuitively, regularizing the elements of D is more direct than regularizing the elements of L. We highlight the difference between Wang's method and our proposed method by a numerical simulation in a simplified setting in Section 3. In addition, the method proposed in this paper has a coupling between all elements of D in the regularization PDE, while the method proposed by Wang et al. does not have such a coupling between the channels.

We also note that the functional (13) is chosen mainly because of the good properties of the corresponding scalar- and vector-valued functionals [2, 25], with edge preservation as the most prominent property. Depending on the application at hand, other functionals might be considered as alternatives. The framework developed in this paper is however applicable for other regularization functionals as well.

2.1. Euler-Lagrange equations

In this section, we derive the Euler-Lagrange equations corresponding to the minimization functional (13). We first differentiate the fidelity functional

$$\begin{aligned} \frac{\partial F}{\partial \ell_{kl}} &= \frac{\partial}{\partial \ell_{kl}} \sum_{ij} \|d_{ij} - \hat{d}_{ij}\|_2^2 \\ &= 2 \sum_{ij} (d_{ij} - \hat{d}_{ij}) \frac{\partial d_{ij}}{\partial \ell_{kl}}. \end{aligned} \tag{14}$$

We differentiate D with respect to ℓ_{kl}, for example

$$\frac{\partial D}{\partial \ell_{11}} = \begin{pmatrix} 2\ell_{11} & \ell_{21} & \ell_{31} \\ \ell_{21} & 0 & 0 \\ \ell_{31} & 0 & 0 \end{pmatrix}. \tag{15}$$

The other derivatives follow the same pattern. Writing out (14), we find the derivative of the fidelity functional,

$$\begin{aligned} \frac{\partial F}{\partial \ell_{kl}} = 2\Big[&(d_{11} - \hat{d}_{11}) \frac{\partial d_{11}}{\partial \ell_{kl}} + 2(d_{21} - \hat{d}_{21}) \frac{\partial d_{21}}{\partial \ell_{kl}} \\ &+ (d_{22} - \hat{d}_{22}) \frac{\partial d_{22}}{\partial \ell_{kl}} + 2(d_{31} - \hat{d}_{31}) \frac{\partial d_{31}}{\partial \ell_{kl}} \\ &+ 2(d_{32} - \hat{d}_{32}) \frac{\partial d_{32}}{\partial \ell_{kl}} + (d_{33} - \hat{d}_{33}) \frac{\partial d_{33}}{\partial \ell_{kl}} \Big], \end{aligned} \tag{16}$$

where $\{d_{ij}\}_{i=1,j=1}^{3}$ denote the elements of the matrix D. We differentiate the regularization functional in (13). Define the total variation norm of a matrix $D \in R^3 \times R^3$ as

$$\begin{aligned} \mathrm{TV}[D] = \Big(&\mathrm{TV}[d_{11}(\ell_{ij})]^2 + 2\mathrm{TV}[d_{21}(\ell_{ij})]^2 \\ &+ \mathrm{TV}[d_{22}(\ell_{ij})]^2 + 2\mathrm{TV}[d_{31}(\ell_{ij})]^2 \\ &+ 2\mathrm{TV}[d_{32}(\ell_{ij})]^2 + \mathrm{TV}[d_{33}(\ell_{ij})]^2 \Big)^{1/2}. \end{aligned} \tag{17}$$

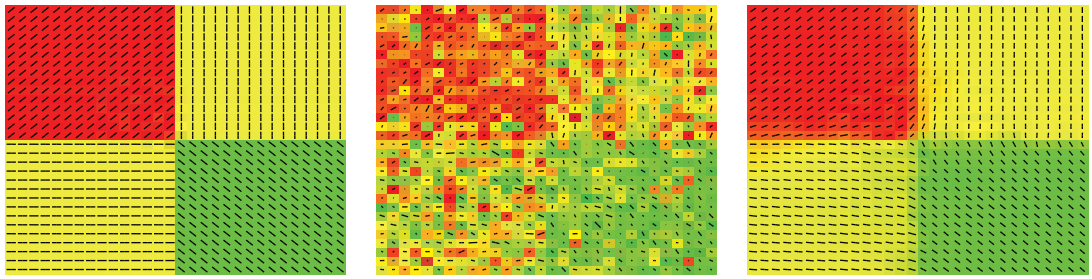
This is a straightforward generalization of the total variation norm of a vector [25].

Using the chain rule, we find the derivatives of the regularization functional to be

$$\frac{\partial R}{\partial \ell_{kl}} = -\sum_{ij} \alpha_{ij} \nabla \cdot \left(\frac{\nabla d_{ij}}{|\nabla d_{ij}|} \right) \frac{\partial d_{ij}}{\partial \ell_{kl}}, \tag{18}$$

with

$$\alpha_{ij} = \frac{\mathrm{TV}[d_{ij}]}{\mathrm{TV}[D]}. \tag{19}$$



(a) (b) (c)

Figure 2: A synthetically produced purely anisotropic tensor field with four distinct regions is degraded with normally distributed noise. The noisy field is then processed with our proposed method: (a) the clean vector field D_0; (b) the noisy field $\hat{D}$; (c) the recovered field D.

Note here that this derivative is essentially similar to the derivative in the color TV model of Blomgren and Chan [25], but with the important difference that we represent the diffusion matrix by its Cholesky factors.

In the next section, we perform numerical simulations using the method proposed in this paper. We give more details on the Euler-Lagrange equations in the appendix, which also contains some details on the numerical implementation of the model.

3. NUMERICAL EXPERIMENTS

In this section we perform numerical experiments on synthetically constructed tensor fields and real tensor fields from a human brain. The numerical implementation of the method is briefly discussed in the appendix.

For the synthetical fields we have constructed clean tensor fields, which are degraded with noise with a prior known distribution. Thus, we are able to measure how well the method performs on synthetical data. For the real human brain DTI, the "true" solution is of course not known in advance. In this case, we measure the performance of the method in terms of a reference solution where a large series of acquisitions are averaged. This is explained in detail in Section 3.3. For the numerical implementation of the method and some of the visualizations, we have used Matlab [28].

3.1. Synthetical tensor fields

In the first numerical experiment, displayed in Figure 2, we test the performance of the proposed method on a simple tensor field. This field is mapping a square domain $\Omega \subset R^2$, with four distinct regions, to $R^{2\times 2}$. We construct the clean tensor-valued data by prescribing the eigenvalues and corresponding eigenvectors. The values of each element of L is in the range $[0, 1]$. Then we add normally distributed noise $\eta(\sigma)$ to each element of the Cholesky factorization of the matrix, that is, $\hat{L} = L + \eta(\sigma)$. Finally, the degraded diffusion tensor is constructed by $\hat{D} = \hat{L}\hat{L}^T$. The noise levels in the simulations in Figures 2 and 4 are given by $\sigma = 0.35$ and $\sigma = 0.25$, respectively. The time step is $\Delta t = 0.001$. Note that the discontinuities in the data are preserved in the solution, that is, the edge preserving property of scalar and vector-valued TV flow is kept in the proposed matrix-valued flow. In the first example, the diffusion is anisotropic in the whole domain. To show how the proposed method differentiates between isotropic and anisotropic regions we show a similar example where one of the four regions is interchanged with an isotropic region. The isotropic region is constructed by considering the orthogonal matrix Q from the QR factorization of a random 2×2 matrix. The columns of the matrix Q are considered to be the eigenvectors of the diffusion tensor. We specify two random numbers in the range $[0, 1]$ as the eigenvalues of the tensor. Thus the diffusion is random in the isotropic region. As we can observe from these two numerical examples on synthetical data, edges are preserved in the regularized images. We observe that even when the noise level is high, we are able to reconstruct an image which is close to the true noise-free image.

From these numerical experiments on synthetical data we see that the proposed method gives encouraging results. Similarly as in the scalar- and vector-valued settings, edges are well preserved. We further investigate the edge preservation in the next experiment.

3.2. Qualitative experiments

To highlight the qualitative differences between regularizing the elements of the tensor D and the elements of the Cholesky factors L, we have constructed a simple numerical example in 1D. We have removed the fidelity measure from the model, thus the method is in this setting merely a diffusion filter. Thus we have simplified the model in such a way that we can study the qualitative behavior of the two regularization filters in the same setting (see Figure 3). From this example, we clearly see that when we regularize D the edges are better preserved than when we regularize L. Note that Wang et al. regularize the Cholesky factors [18].

We also present a numerical example in 2D where we solve the PDEs first as an uncoupled system, that is, by employing the weighting factors $\alpha_{ij} = 1$, and then as a coupled system where we use the weighting factors from (10). We denote the clean field by D, the noisy field by $\hat{D}$, the field regularized with the uncoupled system by D^u and the field regularized with the coupled system by D^c. In Figures 5(a)–5(d),

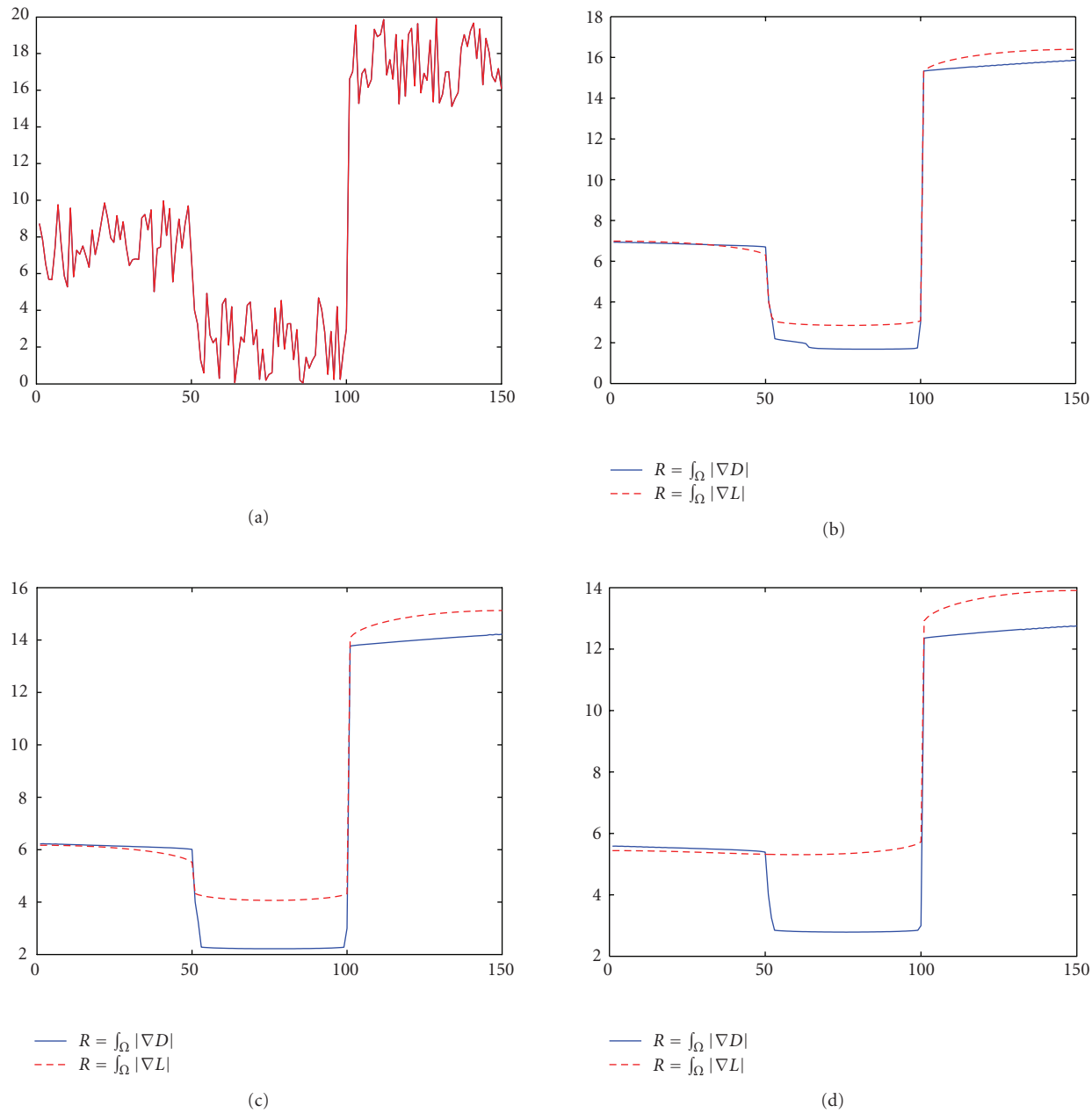


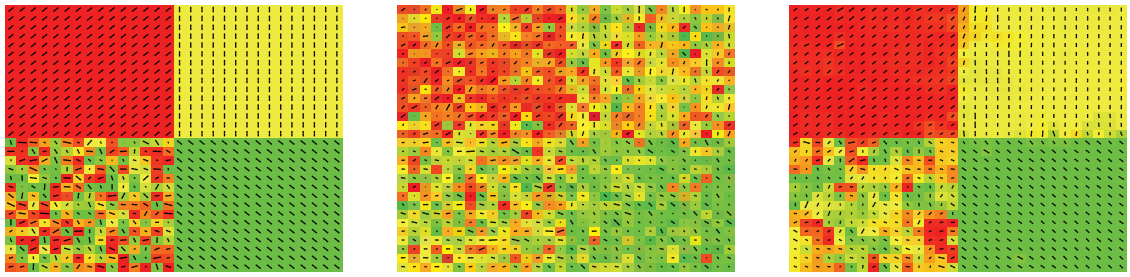
Figure 3: A simple 1D example showing the qualitative behavior of the model for regularizers $\int_\Omega |\nabla L|$ and $\int_\Omega |\nabla D|$. The noisy signal in (a) is processed with both flows. Subfigures (b), (c), and (d) are snapshots during the flow at the three times $t = 8$, $t = 16$, and $t = 24$.

we show the elements D_{11}, $\hat{D}_{11}$, D^u_{11}, and D^c_{11}, respectively. Subindexes denote the elements of the matrix field. Figures 5(e)–5(h) show the elements D_{12}, $\hat{D}_{12}$, D^u_{12}, and D^c_{12}, while Figures 5(i)–5(l) show the element D_{22}, $\hat{D}_{22}$, D^u_{22}, and D^c_{22}. From Figure 5, we observe that the uncoupled system does not discriminate the noise from the weak signal. The coupled system on the other hand better discriminates the noise from the weak signal. A similar 1D example is shown by Blomgren and Chan using the color TV model [25].

In the next section, we go one step further and process real human brain DTMRI.

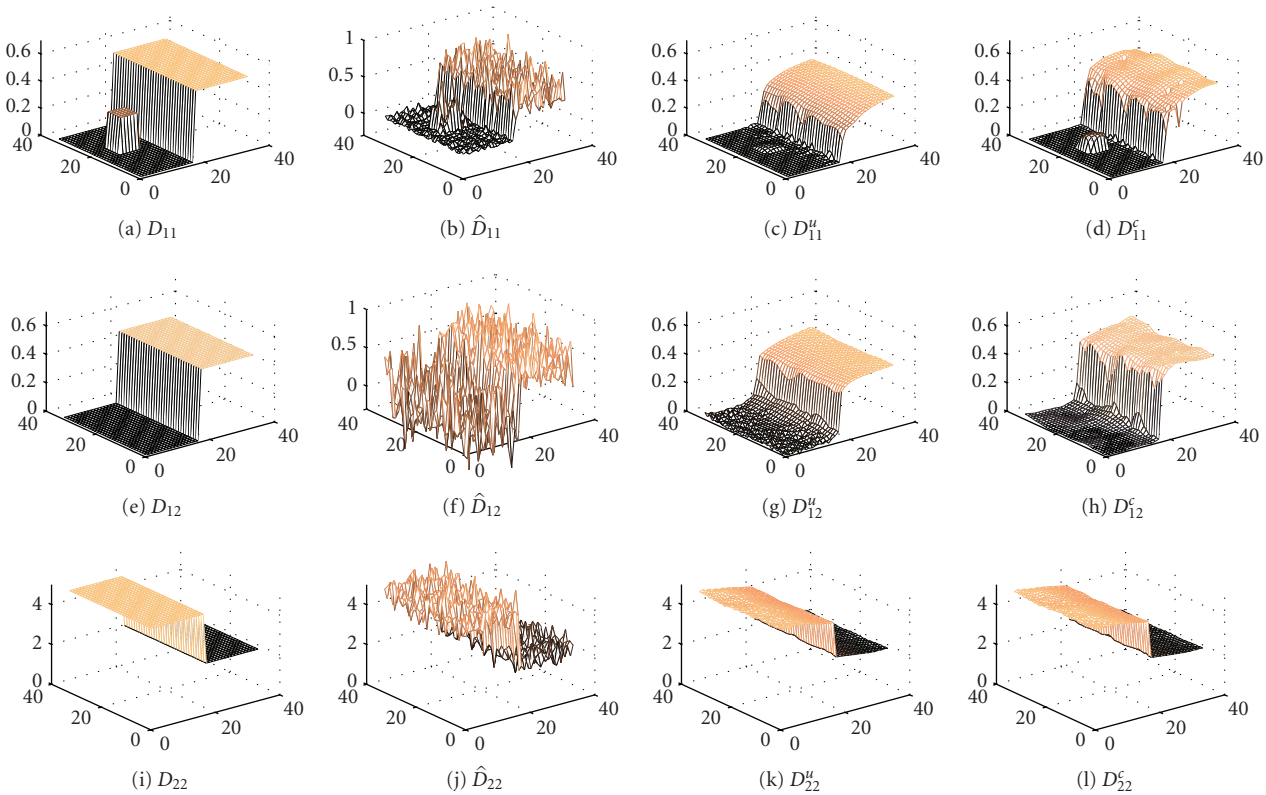
3.3. Human brain DTMRI

We also perform numerical experiments on DTMRI acquisitions of a healthy human brain from a volunteer. The human subject data is acquired using a 3.0 T scanner (Magnetom Trio, Siemens Medical Solutions, Erlangen, Germany) with an 8-element head coil array and a gradient subsystem with the maximum gradient strength of 40 mT/m and maximum slew rate of 200 mT/m/ms. The DTI data is based on spin-echo single-shot EPI acquired utilizing generalized autocalibrating partially parallel acquisitions (GRAPPA)



(a) (b) (c)

Figure 4: Visualization of (a) the true vector field, (b) the noisy field, and (c) the recovered field. In this example, the tensor field is isotropic in the lower-left corner, anisotropic in the other parts.



(a) D_{11} (b) $\hat{D}_{11}$ (c) D^u_{11} (d) D^c_{11}

(e) D_{12} (f) $\hat{D}_{12}$ (g) D^u_{12} (h) D^c_{12}

(i) D_{22} (j) $\hat{D}_{22}$ (k) D^u_{22} (l) D^c_{22}

Figure 5: A noisy 2D tensor field is regularized. In this example, the smallest parts of the signal are not easily discriminated from the noise.

technique with acceleration factor of 2 and 64 reference lines. The DTI acquisition consists of one baseline EPI and six diffusion weighted images (b-factor of 1000 s/mm^2) along the following gradient directions: $G_1 = 1/\sqrt{2}[1,0,1]^T$, $G_2 = 1/\sqrt{2}[-1,0,1]^T$, $G_3 = 1/\sqrt{2}[0,1,1]^T$, $G_4 = 1/\sqrt{2}[0,1,-1]^T$, $G_5 = 1/\sqrt{2}[1,1,0]^T$, $G_6 = 1/\sqrt{2}[1,1,0]^T$. Each acquisition has the following parameters: TE/TR/averages is 91 ms/10000 ms/2, FOV is 256 mm $\times$ 256 mm, slice thickness/gap is 2 mm/0 mm, acquisition matrix is 192 $\times$ 192 pixels, and partial Fourier encoding is 75%.

For validation of the proposed regularization method on real data, we construct a reference solution D^* by averaging 18 replications. Each replication consists of six-direction weighted and one nonweighted acquisitions. This reference solution is compared to solutions where averages of two, four, and six replications are postprocessed with the proposed regularization method. As a measure of the distance between the reference solution and the processed solution, we use the following metric:

$$m(D,D^*) = \Big([d_{11}-d^*_{11}]^2 + 2[d_{12}-d^*_{12}]^2 + 2[d_{13}-d^*_{13}]^2 + [d_{22}-d^*_{22}]^2 + 2[d_{23}-d^*_{23}]^2 + [d_{33}-d^*_{33}]^2\Big)^{1/2}. \quad (20)$$

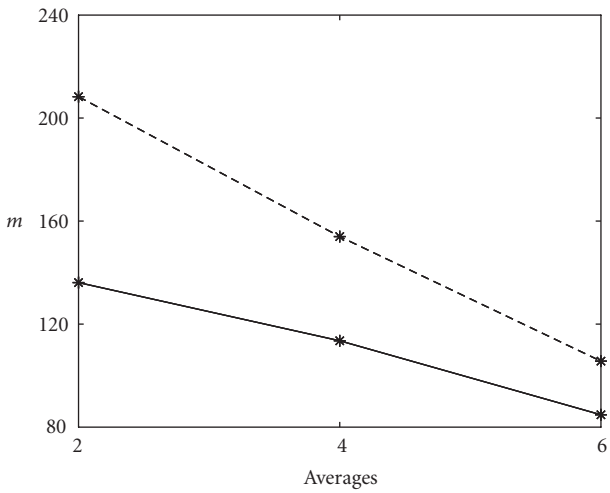


Figure 6: Comparison of $m(D, D^*)$ for the original tensors (dashed) and the regularized tensors (solid) versus the number of averaged acquisitions.

Table 1: The distance $m(D, D^*)$ of the regularized and the nonregularized tensor fields from the numerical examples shown in Figures 7 and 8.

Averages	λ	Reg. $m(D, D^*)$	Nonreg. $m(D, D^*)$
2	9	136.1	208.3
4	13	113.5	154
6	19	84.8	105.6

Table 2: The average deviation angle (ADA) of the noisy data, the processed data (two different regularization parameters), and the reference data.

Data(↓) ADA (→)	Global	ROI 1	ROI 2	ROI 3
Noisy (4 avgs.)	12.32	32.92	41.02	42.87
Denoised, $\lambda = 13$	6.27	11.77	31.50	25.27
Denoised $\lambda = 20$	7.58	13.34	32.88	28.86
Clean image (18 avgs.)	6.65	18.23	24.80	24.80

For each simulation, we report the regularization parameter λ and the metric $m(\cdot, \cdot)$ in Table 1 and in Figure 6. We display the result before and after applying the proposed method on real DTMRI data in Figures 7 and 8. In the figures, we display a 2D slice of an RGB direction encoded fractional - anisotropy (FA) measure defined by

$$\mathrm{FA} = \sqrt{\frac{3}{2}\frac{(\bar{\lambda} - \lambda_1)^2 + (\bar{\lambda} - \lambda_2)^2 + (\bar{\lambda} - \lambda_3)^2}{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}}, \tag{21}$$

where $\bar{\lambda} = (\lambda_1 + \lambda_2 + \lambda_3)/3$. The FA measure is direction-encoded as described by Pajevic and Pierpaoli [29]. We use the DTMRI software DTIStudio to construct the visualizations [30]. In the figures, we show the color-coded FA.

The noise level is different for each simulation due to the varying number of acquisitions. Consequently, the regularization parameter λ is different for each simulation. However, for clinical applications, the regularization parameter is estimated once for each imaging protocol. When this is done, the same regularization parameter can be used for subsequent applications of the same imaging protocol.

3.4. Human brain ROI study

Since our algorithm regularizes the tensor field, we focus on the evaluation of the tensor field, and the derived scalar FA map. However, we note that from the processed tensor field, we may reconstruct the corresponding diffusion weighted images $\{S_i\}_{i=1}^6$ by (2). There are obvious visual improvements in the processed diffusion weighted images compared to the noisy diffusion weighted images. Edges are preserved and noise is suppressed. Quantitatively, the mean and standard deviation at certain homogeneous regions of interests (ROIs) show significant improvements. We will now assess the visual and quantitative improvements in terms of the denoised tensors.

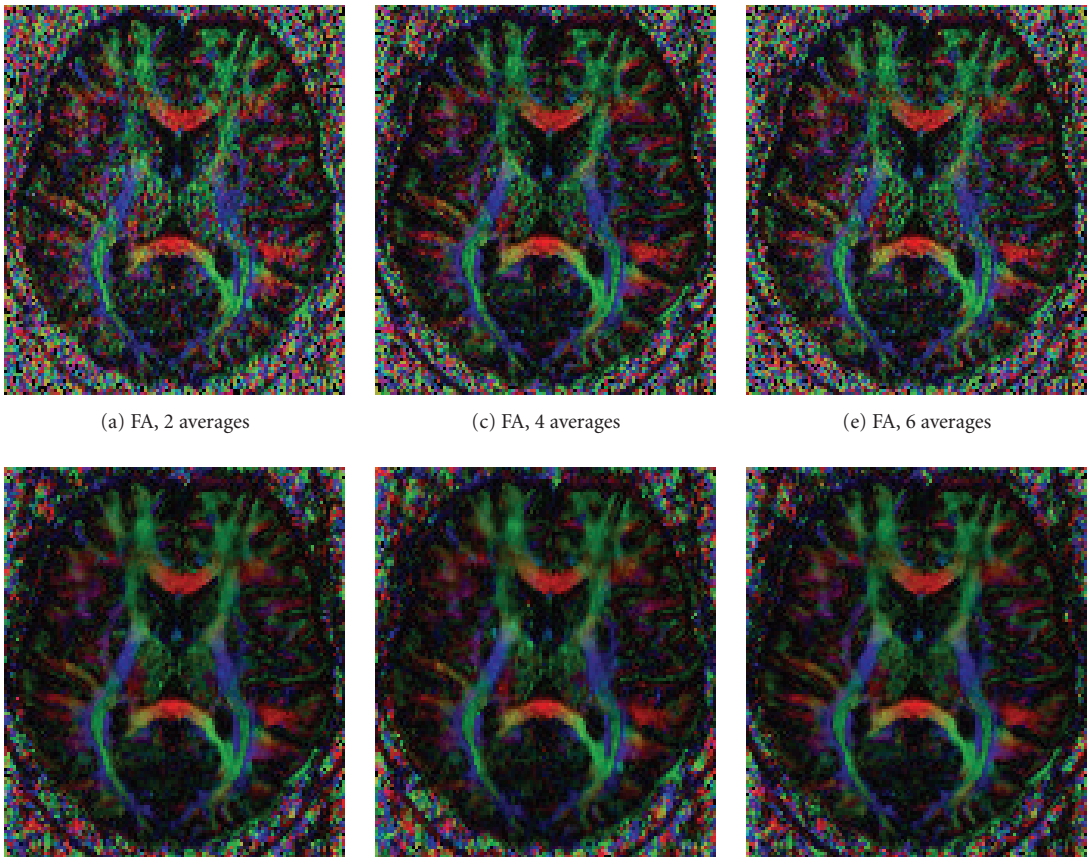
For qualitative evaluation, we select three regions of interest (ROIs) from one slice of the acquired images, with a 15-by-15 voxel size. We plot the 2D projection of the eigenvector corresponding to the major eigenvalue in Figure 9. From Figure 9, we can clearly see that our method preserves discontinuities (edges) in the tensor field, while it smooths the tensor field in homogeneous regions. The denoised tensor field from the 4-average acquisition is close to the tensor field obtained from the 18-average acquisition.

For quantitative measures, we use the average deviation angle (ADA) index of Chen and Hsu to measure if the tensor field undergoes gradual changes or sharp turns [21]. The PDE filtering is performed after the tensors are computed, so we use the angle deviation in adjacent voxels as a measure of its performance instead of normalized magnitude of diffusion tensor error (NMTE) index [21]. Denote the eigenvector corresponding to the largest eigenvalue by V^*. Define the ADA by

$$\mathrm{ADA} = \frac{\Delta\alpha_{i-1} + \Delta\alpha_{i+1} + \Delta\alpha_{j-1} + \Delta\alpha_{j+1} + \Delta\alpha_{k-1} + \Delta\alpha_{k+1}}{6}, \tag{22}$$

where, for example, $\Delta\alpha_{i-1} = \cos^{-1}(|(V^*_{ijk}, V^*_{i-1jk})|)$. We note that we use the absolute value of the inner product $(\cdot, \cdot)$ to accommodate antisense directional vectors. A small change in direction from one voxel to its neighbor gives a small ADA, while a large change in direction gives a large ADA.

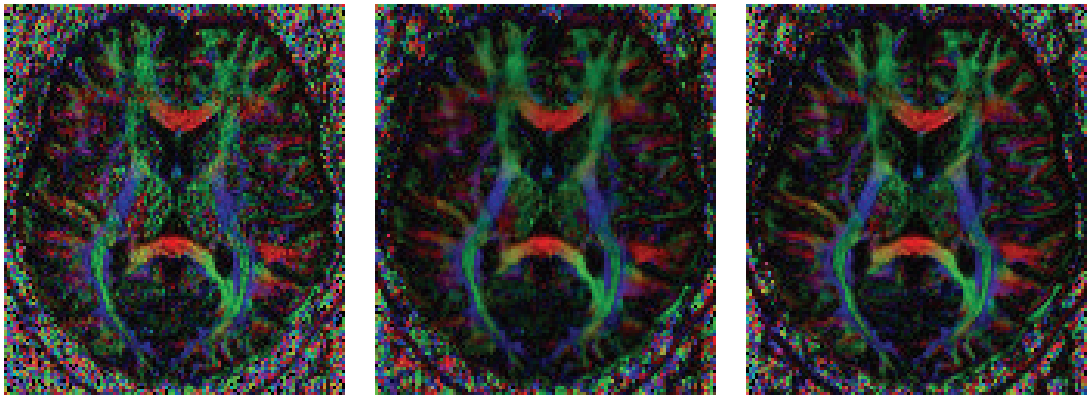
After masking the background, we compute the average ADA within the brain, and call it the global ADA. From Table 2, we see that the global ADA of the data is reduced from 12.31 to 6.27 by the denoising algorithm, whereas the 18-average clean data has an ADA of 6.65. With a higher data fidelity requirement (when λ is larger, e.g., 20), the smoothing is not very aggressive and the ADA is not as close as when $\lambda = 13$. When λ is less than 13 (data not shown here), the smoothing is excessive and the ADA values fall well below the



(a) FA, 2 averages (c) FA, 4 averages (e) FA, 6 averages

(b) FA, 2 averages denoised (d) FA, 4 averages denoised (f) FA, 6 averages denoised

Figure 7: Color-coded fractional anisotropy (FA) maps constructed from averages of two (a), four (c), and six (e) acquisitions, and the corresponding denoised maps (b), (d), and (f).



(a) FA, 4 averages (b) FA, 4 averages denoised (c) FA, 18 averages

Figure 8: The noisy 4-average acquisition (a) is compared with the denoised acquisition (b) and a reference solution at 18 averages.

ADA of the 18-average data. From this information, we conclude that for the current acquisition data, $\lambda = 13$ is the best choice. The ADA at selected ROIs is larger than the global ADA because in those regions, there are obvious edges that contributed to the relatively large ADA values. Compared with the noisy 4-average data, the denoised data show significant improvements. Using the regularization parameter $\lambda = 13$, the ADA is close to the ADA of the 18-average data. The ADAs of all the ROIs are however reduced compared to the noisy data.

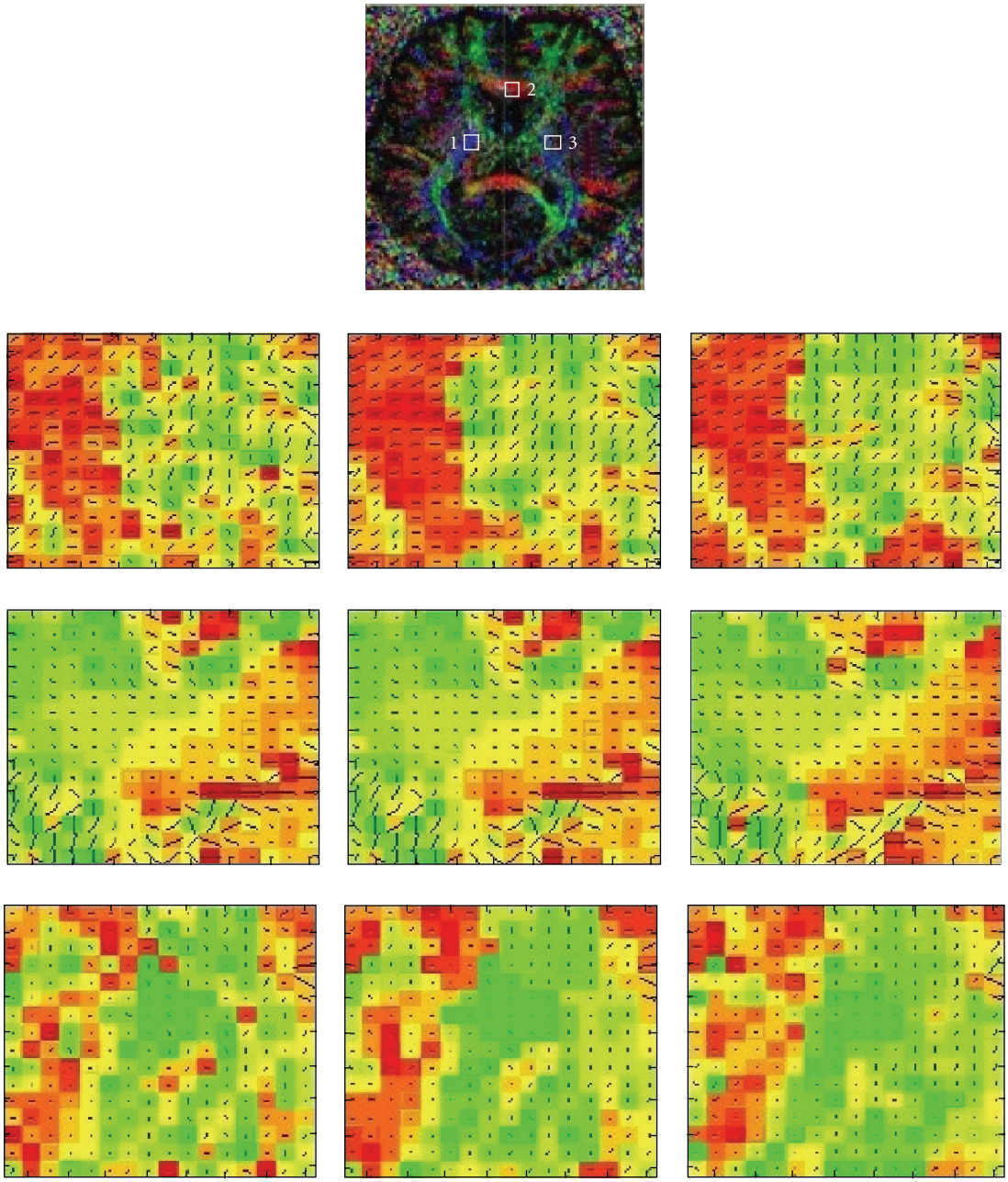


Figure 9: ROI study. Top image shows the ROIs that we use. The second row from left to right: the noisy (4-average) data, denoised data with $\lambda = 13$, and clean (18-average) data of ROI 1. The third row from left to right: the noisy (4-average) data, denoised data with $\lambda = 13$, and clean (18-average) data of ROI 2. The fourth row from left to right: the noisy (4-average) data, denoised data with $\lambda = 13$, and clean (18-average) data of ROI 3.

4. CONCLUSION

In this work, we have generalized the color TV regularization method of Blomgren and Chan [25] to yield a structure preserving regularization method for matrix-valued images. We have shown that the proposed method performs well as a regularization method with the important property of preserving both edges in the data and positive definiteness of the diffusion tensor. Numerical experiments on synthetically produced data and real data from DTI of a human brain of a healthy volunteer indicate good performance of the proposed method.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Siamak Ardekani at UCLA for kindly providing the DRMRI data used in this project. We would also like to thank Dr. Jerome Darbon for

his very useful comments on the paper. The work of Johan Lie has been funded by the Norwegian Research Council, Project BeMatA. The work of Tony F. Chan and Tin-Man Lee has been funded by NSF, ONR, and NIH.

APPENDIX

A. EULER-LAGRANGE EQUATION AND NUMERICAL IMPLEMENTATION

In this appendix, we explicitly write out the Euler-Lagrange equations corresponding to the minimization functional (13). In addition, the numerical scheme used in the simulations in Section 3 is briefly discussed.

Using the short-hand notation

$$p(x_{ij}) = \alpha_{ij}\nabla\cdot\left(\frac{\nabla x_{ij}}{|\nabla x_{ij}|}\right), \tag{A.1}$$

we can write out the derivatives of R and F as

$$\begin{aligned}
\frac{\partial R}{\partial \ell_{11}} &= 2(\ell_{11}p(\ell_{11}^2) + \ell_{21}p(\ell_{11}\ell_{21}) + \ell_{31}p(\ell_{11}\ell_{31})),\\
\frac{\partial R}{\partial \ell_{21}} &= 2(\ell_{11}p(\ell_{11}\ell_{21}) + \ell_{21}p(\ell_{21}^2 + \ell_{22}^2)\\
&\quad + \ell_{31}p(\ell_{21}\ell_{31} + \ell_{22}\ell_{32})),\\
\frac{\partial R}{\partial \ell_{22}} &= 2(\ell_{22}p(\ell_{21}^2 + \ell_{22}^2) + \ell_{32}p(\ell_{21}\ell_{31} + \ell_{22}\ell_{32})),\\
\frac{\partial R}{\partial \ell_{31}} &= 2(\ell_{11}p(\ell_{11}\ell_{31}) + \ell_{21}p(\ell_{21}\ell_{31} + \ell_{22}\ell_{32})\\
&\quad + \ell_{31}p(\ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2)),\\
\frac{\partial R}{\partial \ell_{32}} &= 2(\ell_{22}p(\ell_{21}\ell_{31} + \ell_{22}\ell_{32}) + \ell_{32}p(\ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2)),\\
\frac{\partial R}{\partial \ell_{33}} &= (2\ell_{33}p(\ell_{31}^2 + \ell_{32}^2 + \ell_{33}^2)),\\
\frac{\partial F}{\partial \ell_{11}} &= 4[(d_{11} - \hat{d}_{11})\ell_{11} + (d_{21} - \hat{d}_{21})\ell_{21} + (d_{31} - \hat{d}_{31})\ell_{31}],\\
\frac{\partial F}{\partial \ell_{21}} &= 4[(d_{21} - \hat{d}_{21})\ell_{11} + (d_{22} - \hat{d}_{22})\ell_{21} + (d_{32} - \hat{d}_{32})\ell_{31}],\\
\frac{\partial F}{\partial \ell_{22}} &= 4[(d_{22} - \hat{d}_{22})\ell_{22} + (d_{32} - \hat{d}_{32})\ell_{32}],\\
\frac{\partial F}{\partial \ell_{32}} &= 4[(d_{32} - \hat{d}_{32})\ell_{22} + (d_{33} - \hat{d}_{33})\ell_{32}],\\
\frac{\partial F}{\partial \ell_{31}} &= 4[(d_{31} - \hat{d}_{31})\ell_{11} + (d_{32} - \hat{d}_{32})\ell_{21} + (d_{33} - \hat{d}_{33})\ell_{31}],\\
\frac{\partial F}{\partial \ell_{33}} &= 4[(d_{33} - \hat{d}_{33})\ell_{33}].
\end{aligned} \tag{A.2}$$

By combining each of the equations

$$\frac{\partial G}{\partial \ell_{ij}} = \frac{\partial R}{\partial \ell_{ij}} + \frac{\partial F}{\partial \ell_{ij}}, \quad \{ij\} \in \{11, 21, 22, 31, 32, 33\}, \tag{A.3}$$

we arrive at the Euler-Lagrange equations corresponding to the minimization problem (13). In the numerical simulations, we use the steepest descent method with a fixed time step. Thus, we have the six equations,

$$d_{ij}^{n+1} = d_{ij}^{n} - \Delta t\frac{\partial G^n}{\partial \ell_{ij}}, \quad \{kl\} \in \{11, 21, 22, 31, 32, 33\}, \tag{A.4}$$

where n is the iteration index, and Δt is the time-step parameter. We approximate the gradient $\partial G/\partial \ell_{ij}$ by standard finite difference schemes, see, for example, [4]. We here note that each iteration of the form (A.4) is performed sequentially. Thus, the equations are solved as a coupled system of six PDEs.

REFERENCES

[1] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, 1992.

[3] M. Lysaker, S. Osher, and X.-C. Tai, "Noise removal using smoothed normals and surface fitting," *IEEE Transactions on Image Processing*, vol. 13, no. 10, pp. 1345–1357, 2004.

[4] T. F. Chan and S. Esedoglu, "Aspects of total variation regularized L^1 function approximation," *SIAM Journal on Applied Mathematics*, vol. 65, no. 5, pp. 1817–1837, 2005.

[5] T. F. Chan and J. Shen, *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*, SIAM, Philadelphia, Pa, USA, 2005.

[6] J. Weickert, "A review of nonlinear diffusion filtering," in *Proceedings of the 1st International Conference on Scale-Space Theory in Computer Vision*, vol. 1252 of *Lecture Notes in Computer Science*, pp. 3–28, Utrecht, The Netherlands, July 1997.

[7] J. Weickert and T. Brox, "Diffusion and regularization of vector- and matrix-valued images," Tech. Rep. preprint no. 58, Fachrichtung 6.1 Mathematik, Universitat des Saarlandes, Saarbrücken, Germany, 2002.

[8] P. J. Basser, J. Mattiello, and D. LeBihan, "MR diffusion tensor spectroscopy and imaging," *Biophysical Journal*, vol. 66, no. 1, pp. 259–267, 1994.

[9] D. Le Bihan, J.-F. Mangin, C. Poupon, et al., "Diffusion tensor imaging: concepts and applications," *Journal of Magnetic Resonance Imaging*, vol. 13, no. 4, pp. 534–546, 2001.

[10] C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *Medical Image Analysis*, vol. 6, no. 2, pp. 93–108, 2002.

[11] S. Mori and P. B. Barker, "Diffusion magnetic resonance imaging: its principle and applications," *The Anatomical Record*, vol. 257, no. 3, pp. 102–109, 1999.

[12] S. Mori and P. C. M. van Zijl, "Fiber tracking: principles and strategies—a technical review," *NMR in Biomedicine*, vol. 15, no. 7-8, pp. 468–480, 2002.

[13] R. Bammer, "Basic principles of diffusion-weighted imaging," *European Journal of Radiology*, vol. 45, no. 3, pp. 169–184, 2003.

[14] E. O. Stejskal and J. E. Tanner, "Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient," *The Journal of Chemical Physics*, vol. 42, no. 1, pp. 288–292, 1965.

[15] E. O. Stejskal, "Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow," *The Journal of Chemical Physics*, vol. 43, no. 10, pp. 3597–3603, 1965.

[16] P. Tofts, Ed., *Quantitative MRI of the Brain*, John Wiley & Sons, New York, NY, USA, 2005.

[17] C.-F. Westin, S. E. Maier, H. Mamata, A. Nabavi, F. A. Jolesz, and R. Kikinis, "Processing and visualization for diffusion tensor MRI," *Medical Image Analysis*, vol. 6, no. 2, pp. 93–108, 2002.

[18] Z. Wang, B. C. Vemuri, Y. Chen, and T. H. Mareci, "A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI," *IEEE Transactions on Medical Imaging*, vol. 23, no. 8, pp. 930–939, 2004.

[19] D. Tschumperlé and R. Deriche, "Variational frameworks for DT-MRI estimation, regularization and visualization," in *Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV '03)*, vol. 1, pp. 116–121, Nice, France, October 2003.

[20] C.-F. Westin and H. Knutsson, "Tensor field regularization using normalized convolution," in *Proceedings of the 9th International Conference on Computer Aided Systems Theory (EUROCAST '03)*, R. Moreno-Diaz and F. Pichler, Eds., vol. 2809 of *Lecture Notes in Computer Science*, pp. 564–572, Las Palmas de Gran Canaria, Canary Islands, Spain, February 2003.

[21] B. Chen and E. W. Hsu, "Noise removal in magnetic resonance diffusion tensor imaging," *Magnetic Resonance in Medicine*, vol. 54, no. 2, pp. 393–401, 2005.

[22] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, "Regularizing flows for constrained matrix-valued images," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 147–162, 2004.

[23] Y. Gur and N. Sochen, "Denoising tensors via Lie group flows," in *Proceedings of the 3rd International Workshop on Variational, Geometric, and Level Set Methods in Computer Vision (VLSM '05)*, vol. 3752 of *Lecture Notes in Computer Science*, pp. 13–24, Springer, Beijing, China, October 2005.

[24] D. Groisser, "Some differential-geometric remarks on a method for minimizing constrained functionals of matrix-valued functions," *Journal of Mathematical Imaging and Vision*, vol. 24, no. 3, pp. 349–358, 2006.

[25] P. Blomgren and T. F. Chan, "Color TV: total variation methods for restoration of vector-valued images," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 304–309, 1998.

[26] T. F. Chan, G. H. Golub, and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration," *SIAM Journal of Scientific Computing*, vol. 20, no. 6, pp. 1964–1977, 1999.

[27] G. Sapiro, "Color snakes," Tech. Rep. HPL-95-113, Hewlett Packard Computer Peripherals Laboratory, Palo Alto, Calif, USA, 1995.

[28] The Mathworks, "MatLab, The Language of Technical Computing," http://www.mathworks.com/matlab/.

[29] S. Pajevic and C. Pierpaoli, "Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain," *Magnetic Resonance in Medicine*, vol. 42, no. 3, pp. 526–540, 1999.

[30] Mori and coworkers, "DTI-studio", http://cmrm.med.jhmi.edu/.