Simultaneous Smoothing and Estimation of DTI via Robust Variational Non-local Means

Meizhu Liu, Baba Vemuri, Rachid Deriche

Simultaneous Smoothing & Estimation of DTI via Robust Variational Non-local Means *

Meizhu Liu [1], Baba C. Vemuri [1], Rachid Deriche [2]

[1] Dept. of CISE, University of Florida, Gainesville, FL, 32611, US
[2] Project-Team ATHENA, INRIA, Sophia Antipolis - Méditerranée, 06902, France

Abstract. Regularized diffusion tensor estimation is an essential step in DTI analysis. There are many methods proposed in literature for this task but most of them are neither statistically robust nor feature preserving denoising techniques that can simultaneously estimate symmetric positive definite (SPD) diffusion tensors from diffusion MRI. One of the most popular techniques in recent times for feature preserving scalar-valued image denoising is the non-local means filtering method that has recently been generalized to the case of diffusion MRI denoising. However, these techniques denoise the multi-gradient volumes first and then estimate the tensors rather than achieving it simultaneously in a unified approach. Moreover, some of them do not guarantee the positive definiteness of the estimated diffusion tensors. In this paper, we propose a novel and robust variational framework for the simultaneous smoothing and estimation of diffusion tensors from diffusion MRI. Our variational principle makes use of a recently introduced total Kullback-Leibler (tKL) divergence, which is a statistically robust similarity measure between diffusion tensors, weighted by a non-local factor adapted from the traditional non-local means filters. For the data fidelity, we use the nonlinear least-squares term derived from the Stejskal-Tanner model. We present experimental results depicting the positive performance of our method in comparison to competing methods on synthetic and real data examples.

1 Introduction

Diffusion MRI is a technique that uses diffusion sensitizing gradients to non-invasively image anisotropic properties of tissue. Diffusion tensor imaging (DTI) introduced by Basser et al. [1], approximates the diffusivity function by a symmetric positive definite tensor of order two. There is abundant literature on DTI analysis including but not limited to denoising & tensor field estimation [2–9], DTI registration, fiber tractography etc and all of these latter tasks will benefit from denoising and estimation of smooth diffusion tensors.

* This work was partly supported by the NIH grant NS066340 to Vemuri, the University of Florida Allumni Fellowship to Liu, the INRIA Internships program, the Association France Parkinson and the French National Research Agency (ANR, Neurodegenerative and Psychiatric Diseases).

In most of the existing methods, the diffusion tensors (DTs) are estimated using the raw diffusion weighted echo intensity image (DWI). At each voxel of the 3D image lattice, the diffusion signal intensity S is related with its diffusion tensor $\mathbf{D} \in \mathbf{SPD}(3)$ [1] via the Stejskal-Tanner equation [10]

$$S = S_0 \exp(-b\mathbf{g}^T\mathbf{D}\mathbf{g}), \tag{1}$$

where S_0 is the signal intensity without diffusion sensitizing gradient, b is the diffusion weighting and $\mathbf{g}$ is the direction of the diffusion sensitizing gradient.

Estimating the DTs from DWI is a challenging problem, since the DWI is normally corrupted with noise [7–9]. Therefore, a statistically robust DTI estimation method which is able to perform feature preserving denoising is desired. There are various methods [3, 8, 9, 11–15] that exist in the literature to achieve this goal of estimating $\mathbf{D}$ from S. A very early one is direct tensor estimation [16]. Though time efficient, it is sensitive to noise because only 7 gradient directions are used to estimate $\mathbf{D}$ and S_0. Another method is the minimum recovery error (MRE) estimation or least squares fitting [1] which minimizes the error when recovering the DTs from the DWI. MRE is better than direct estimation but it does not enforce spatial regularization or the SPD constraint resulting in possible inaccuracies.

Bearing these deficiencies in mind, researchers developed variational framework (VF) based estimation [3–5]. These approaches take into account the SPD constraint on the diffusion tensors. The smoothing in all these approaches involves some kind of weighted averaging over neighborhoods which define the smoothing operators resulting from the variational principles. These smoothing operators are locally defined and do not capture global geometric structure present in the image.

More recently, some denoising frameworks have been proposed according to the statistical properties of the noise. They assume the noise follows the Rician distribution [17,18], and the DWI is denoised using maximum likelihood estimation. After denoising the DWI, one can use other techniques to estimate the DTI. Besides, the popular NLM based method has been adapted by many to denoise DTI data sets [19–21]. In the NLM based approaches, one first needs to denoise the DWI field and then estimate DTI from the denoised DWI. Alternatively, the diffusion tensors are first estimated and then denoised using Riemannian approach [22] or the NLM framework incorporating a Log-Euclidean metric [23]. The drawback of such two-stage processes is that the errors might accumulate from one stage to the other.

To overcome the aforementioned problems, we propose a novel statistically robust variational non-local approach for simultaneous smoothing and tensor estimation from the raw DWI data. This approach combines the VF, NLM and an intrinsically robust regularizer on the tensor field. The main contributions of this approach are three-fold. First, we use total Bregman divergence (specifically, the tKL) as a measure to regularize the tensor field. Combined with the Cholesky decomposition of the diffusion tensors, this automatically ensures the

[1] $\mathbf{SPD}(3)$ represents the space of 3×3 symmetric positive definite matrices

positive definiteness of the estimated diffusion tensors, which overcomes the common problem for many techniques [3] that need to manually force the tensor to be positive definite. Second, it preserves the structure of the tensor field while denoising via an adaptation of the NLM framework. Finally, it allows for simultaneous denoising and DT estimation, avoiding the error propagation of a two stage approach described earlier. Besides, this method can be easily extended to higher order tensor estimation. We will explain these points at length in the rest of the paper.

The rest of the paper is organized as follows. In Section 2, we introduce our proposed method and explore its properties, followed by the empirical validation in Section 3. Finally we conclude in Section 4.

2 Proposed method

The simultaneous denoising and estimation of the DTI is achieved by minimizing the following energy function:

$$\begin{aligned}\min_{S_0,\,\mathbf{D}\in\mathbf{SPD}} \mathbf{E}(S_0,\mathbf{D}) = \lambda \sum_{\mathbf{x}\in\Omega}\sum_{i=1}^{n} \left(S_i(\mathbf{x}) - S_0(\mathbf{x})\exp\{-b\mathbf{g}_i^T\mathbf{D}(\mathbf{x})\mathbf{g}_i\}\right)^2 \\ +(1-\lambda)\sum_{\mathbf{x}\in\Omega}\sum_{\mathbf{y}\in\mathbb{V}(\mathbf{x})} w(\mathbf{x},\mathbf{y})\left[(S_0(\mathbf{x})-S_0(\mathbf{y}))^2 + \delta(\mathbf{D}(\mathbf{x}),\mathbf{D}(\mathbf{y}))\right],\end{aligned} \tag{2}$$

where Ω is the domain of the image, $w(\mathbf{x},\mathbf{y})$ is the similarity between voxels $\mathbf{x}$ and $\mathbf{y}$, $\mathbb{V}(\mathbf{x})$ is the user specified search window at $\mathbf{x}$, and $\delta(\mathbf{D}(\mathbf{x}),\mathbf{D}(\mathbf{y}))$ is the total Kullback-Leibler (tKL) divergence proposed in [24,25] between tensors $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$. tKL is defined in [24] and has been used in DTI segmentation [24] and classification [26]. We will redefine it later in Section 2.2 for the sake of completeness. The first term of (2) minimizes the non-linear fitting error, the second term enforces smoothness constraints on S_0 and $\mathbf{D}$ via a non-local means regularizer. λ is the regularization constant balancing the fitting error and the smoothness. Note that S_i, S_0 and $\mathbf{D}$ by default represent the values at voxel $\mathbf{x}$, unless specified otherwise.

2.1 Computation of the Weight $w(\mathbf{x},\mathbf{y})$

The weight $w(\mathbf{x},\mathbf{y})$ is regularizes the similarity between $S_0(\mathbf{x})$ and $S_0(\mathbf{y})$, as well as $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$. Usually, one requires the similarity to be consistent with the similarity between their corresponding diffusion signals. Therefore, we define $w(\mathbf{x},\mathbf{y})$ based on the diffusion signal intensities of the two voxels' neighborhoods.

Let $N(\mathbf{x})$ and $N(\mathbf{y})$ denote the neighborhoods of $\mathbf{x}$ and $\mathbf{y}$ respectively. If $\mathbf{y}\in\mathbb{V}(\mathbf{x})$, then $w(\mathbf{x},\mathbf{y})$ is defined as,

$$w(\mathbf{x},\mathbf{y}) = \frac{1}{Z(\mathbf{x})}\exp\left(-\|S(N(\mathbf{x})) - S(N(\mathbf{y}))\|^2/h^2\right), \tag{3}$$

where h is the user specified filtering parameter and Z is the normalization constant. $S(N(\mathbf{y}))$ represents the signal intensities of the neighborhood at $\mathbf{x}$ and $\|S(N(\mathbf{x})) - S(N(\mathbf{y}))\|^2 = \sum_j^m \|S(\mu_j) - S(\nu_j)\|^2$, where μ_j and ν_j are the j^{th} voxels in the neighborhoods respectively, and m is the number of voxels in each neighborhood.

From (2), we can see that when $w(\mathbf{x}, \mathbf{y})$ is large, then $S_0(\mathbf{x})$ and $S_0(\mathbf{y})$ as well as $\mathbf{D}(\mathbf{x})$ and $\mathbf{D}(\mathbf{y})$ respectively are similar. In other words, if the signal intensities for the neighborhoods of two voxels are similar, their corresponding $\mathbf{D}s$ and $S_0 s$ should also be similar. Though having very good accuracy, NLM is known for its high time complexity. To reduce the time cost, we use two tricks. One is to decrease the number of computations by selecting only those voxels whose signal intensity is similar to that of the voxel under consideration. This is specified by

$$w(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{Z(\mathbf{x})} \exp\left(-\|S(N(\mathbf{x})) - S(N(\mathbf{y}))\|^2 / h^2\right), & \text{if } \frac{\|S(N(\mathbf{x}))\|^2}{\|S(N(\mathbf{y}))\|^2} \in [\tau_1, \tau_2] \\ 0, & \text{Otherwise} \end{cases}.$$

We choose $\tau_1 = 0.5$ and $\tau_2 = 2$ in our experiments. This prefiltering process greatly decreases the number of computations.

The other trick is using parallel computing, where we divide the computations into several parts, and assign the computation parts to several processors. In our case, we divide the volumes into 8 subvolumes, and assign each subvolume to one processor, and a desktop with 8 processors is used. This multi-threading technique greatly enhances the efficiency.

2.2 Computation of the tKL Divergence

Motivated by earlier use of KL divergence as a similarity measure between DTs in literature [3], we use the recently introduced tKL [24] to measure the similarity between tensors. tKL has the property of being intrinsically robust to noise and outliers, yields a closed form formula for computing the median (an ℓ_1-norm average) for a set of tensors, and is invariant to special linear group transformations (denoted as $SL(n)$) – transformations that have determinant one [24].

Note that order-2 SPD tensors can be seen as covariance matrices of zero mean Gaussian probability density functions (pdf) [3]. Let $\mathbf{P}$, $\mathbf{Q} \in \mathbf{SPD}(l)$, then their corresponding pdf are

$$p(\mathbf{t}; \mathbf{P}) = \frac{1}{\sqrt{(2\pi)^l \det \mathbf{P}}} \exp\left(-\frac{\mathbf{t}^T \mathbf{P}^{-1} \mathbf{t}}{2}\right),$$
$$q(\mathbf{t}; \mathbf{Q}) = \frac{1}{\sqrt{(2\pi)^l \det \mathbf{Q}}} \exp\left(-\frac{\mathbf{t}^T \mathbf{Q}^{-1} \mathbf{t}}{2}\right),$$

and the tKL between them (in closed form) is given by,

$$\delta(\mathbf{P}, \mathbf{Q}) = \frac{\int p \log \frac{p}{q} d\mathbf{t}}{\sqrt{1 + \int (1 + \log q)^2 q d\mathbf{t}}} = \frac{\log(\det(\mathbf{P}^{-1}\mathbf{Q})) + tr(\mathbf{Q}^{-1}\mathbf{P}) - m}{2\sqrt{c_1 + \frac{(\log(\det \mathbf{Q}))^2}{4} - c_2 \log(\det \mathbf{Q})}},$$

where $c_1 = \frac{3l}{4} + \frac{l^2 \log 2\pi}{2} + \frac{(l \log 2\pi)^2}{4}$ and $c_2 = \frac{l(1+\log 2\pi)}{2}$.

Moreover, the minimization of the third term in (2)

$$\min_{\mathbf{D}(\mathbf{x})} \sum_{\mathbf{x}\in\Omega} \sum_{\mathbf{y}\in\mathbb{V}(\mathbf{x})} \delta(\mathbf{D}(\mathbf{x}), \mathbf{D}(\mathbf{y}))$$

leads to an ℓ_1-norm average which was shown to have closed form expression [24]. In [24], this was called the t-center and was shown to be invariant to transformations from the $SL(n)$ group, i.e., $\delta(\mathbf{P}, \mathbf{Q}) = \delta(\mathbf{A}^T\mathbf{P}\mathbf{A}, \mathbf{A}^T\mathbf{Q}\mathbf{A})$, where $\det \mathbf{A} = 1$.

Furthermore, given a set of SPD tensors $\{\mathbf{Q}_i\}_{i=1}^m$, its t-center $\mathbf{P}^*$ is given by [24]

$$\mathbf{P}^* = \arg\min_{\mathbf{P}} \sum_i^m \delta(\mathbf{P}, \mathbf{Q}_i), \tag{4}$$

and $\mathbf{P}^*$ is explicitly expressed as

$$\mathbf{P}^* = (\sum_i \frac{a_i}{\sum_j a_j} \mathbf{Q}_i^{-1})^{-1},\ a_i = \left(2\sqrt{c_1 + \frac{(\log(\det \mathbf{Q}_i))^2}{4} - c_2 \log(\det \mathbf{Q}_i)}\right)^{-1}. \tag{5}$$

The t-center for a set of DTs is the weighted harmonic mean, which is in closed form. Moreover, the weight is invariant to $SL(n)$ transformations, i.e., $a_i(\mathbf{Q}_i) = a_i(\mathbf{A}^T\mathbf{Q}_i\mathbf{A}), \forall \mathbf{A} \in SL(n)$. The t-center after the transformation becomes $\hat{\mathbf{P}}^* = (\sum a_i(\mathbf{A}^T\mathbf{Q}_i\mathbf{A})^{-1})^{-1} = \mathbf{A}^T\mathbf{P}^*\mathbf{A}$. This means that if $\{\mathbf{Q}_i\}_{i=1}^m$ are transformed by some member of $SL(n)$, the t-center will undergo the same transformation. It is also found that the t-center is statistically robust to noise in that the weight is smaller if the tensor has more noise [24].

2.3 The SPD Constraint

We now show how to guarantee the positive definiteness constraint on the diffusion tensors to be estimated from the DWI data. It is known that if a matrix $\mathbf{D} \in \mathbf{SPD}$, their exists a unique lower diagonal matrix $\mathbf{L}$ with its diagonal values all positive, and $\mathbf{D} = \mathbf{L}\mathbf{L}^T$ [27]. This is the Cholesky factorization theorem. Many researchers [3] have used Cholesky factorization to ensure the positive definiteness. They first compute $\mathbf{L}$, enforcing the diagonal values of $\mathbf{L}$ to be positive, and consequently, $\mathbf{L}\mathbf{L}^T$ will be positive. Unlike this technique, we use Cholesky decomposition and tKL divergence to regularize the tensor field, and this automatically ensures the diagonal values of $\mathbf{L}$ to be positive. The points can be validated as follows.

Substituting $\mathbf{D} = \mathbf{L}\mathbf{L}^T$ into (4), we get

$$\delta(\mathbf{L}(\mathbf{x}), \mathbf{L}(\mathbf{y})) = \frac{\sum_{i=1}^3 (\log \mathbf{L}_{ii}(\mathbf{y}) - \log \mathbf{L}_{ii}(\mathbf{x})) + tr(\mathbf{L}^{-T}(\mathbf{y})\mathbf{L}^{-1}(\mathbf{y})\mathbf{L}(\mathbf{x})\mathbf{L}^T(\mathbf{x})) - 1.5}{\sqrt{c_1 + \frac{(\sum_{i=1}^3 \log \mathbf{L}_{ii}(\mathbf{y}))^2}{4} - c_2 \sum_{i=1}^3 \log \mathbf{L}_{ii}(\mathbf{y})}}. \tag{6}$$

Because of using the log computation, Eq. (6) automatically ensures $\mathbf{L}_{ii}s$ to be positive, therefore we do not need to add the SPD constraint manually.

2.4 Numerical Solution

In this section, we present the numerical solution to the variational principle (2). The partial derivative equations of (2) with respect to S_0 and $\mathbf{L}$ can be computed explicitly and are,

$$\begin{aligned}
\frac{\partial \mathbf{E}}{\partial S_0(\mathbf{x})} =& -2\lambda \sum_{i=1}^{n} \left(S_i - S_0 \exp\{-b\mathbf{g}_i^T \mathbf{L}\mathbf{L}^T \mathbf{g}_i\}\right) \exp\{-b\mathbf{g}_i^T \mathbf{L}\mathbf{L}^T \mathbf{g}_i\} \\
& - 2(1-\lambda) \sum_{\mathbf{y}\in\mathbb{V}(\mathbf{x})} w(\mathbf{x},\mathbf{y})(S_0(\mathbf{x}) - S_0(\mathbf{y})), \\
\frac{\partial \mathbf{E}}{\partial \mathbf{L}(\mathbf{x})} =& 4\lambda \sum_{i=1}^{n} \left(S_i - S_0 \exp\{-b\mathbf{g}_i^T \mathbf{L}\mathbf{L}^T \mathbf{g}_i\}\right) S_0 \exp\{-b\mathbf{g}_i^T \mathbf{L}\mathbf{L}^T \mathbf{g}_i\} b\mathbf{L}^T \mathbf{g}_i \mathbf{g}_i^T \\
& - 2(1-\lambda) \sum_{\mathbf{y}\in\mathbb{V}(\mathbf{x})} \frac{w(\mathbf{x},\mathbf{y})(\mathbf{L}^{-1}(\mathbf{x}) - \mathbf{L}^T(\mathbf{x})\mathbf{L}^{-T}(y)\mathbf{L}^{-1}(y))}{\sqrt{c_1 + \frac{(\sum_{i=1}^{3} \log \mathbf{L}_{ii}(\mathbf{y}))^2}{4} - c_2 \sum_{i=1}^{3} \log \mathbf{L}_{ii}(\mathbf{y})}}.
\end{aligned} \tag{7}$$

To solve (7), we use the limited memory quasi-Newton method described in [28]. This method is useful for solving large problems with a lot of variables, as is in our case. This method maintains simple and compact approximations of Hessian matrices making them require, as the name suggests, modest storage, besides yielding linear rate of convergence. Specifically, we use L-BFGS [28] to construct the Hessian approximation.

3 Experimental results

We evaluate our method on both synthetic datasets with various levels of noise, and on real datasets. We compared our method with other state-of-the-art techniques including the techniques VF [5], NLMt [21] and NLM [21] respectively. We also present the MRE method for comparison since several software packages in vogue use this technique due to its simplicity. We implemented VF and NLMt by ourselves since we did not find any open source versions on the web. For the NLM, we used existing code [2] for DWI denoising and used our own implementation of the least squares fitting to estimate DTI from the denoised DWI. To ensure fairness, we tuned all the parameters of each method for every experiment, and chose the set of parameters yielding the best results. The visual and numerical results show that our method yields better results than competing methods.

3.1 DTI estimation on synthetic datasets

The synthetic data is a 16×16 tensor field with two homogeneous regions as shown in Fig. 1(a). We let $S_0 = 5$, $b = 1500s/mm^2$, and $\mathbf{g}$ be 22 uniformly-spaced directions on the unit sphere starting from $(1, 0, 0)$. Substituting the

[2] https://www.irisa.fr/visages/benchmarks/

DTs, S_0, b, $\mathbf{g}$ into the Stejskal-Tanner equation, we generate a $16 \times 16 \times 22$ DWI $\mathbf{S}$. One representative slice of $\mathbf{S}$ is shown in Fig. 1(b). Then following the method proposed in [17], we add Rician noise to $\mathbf{S}$ and get $\tilde{\mathbf{S}}$, using the formula, $\tilde{\mathbf{S}}(\mathbf{x}) = \sqrt{(\mathbf{S}(\mathbf{x}) + n_r)^2 + n_i^2}$, where n_r and $n_i \sim N(0, \sigma)$. Fig. 1(c) shows the slice in Fig. 1(b) after adding noise (SNR=10). By varying σ, we get different levels of noise and therefore a wide range of signal to noise ratio (SNR). The estimated DTI from using MRE, VF, NLMt, NLM, and the proposed method are shown in Fig. 1. The figure shows that our method can estimate the tensor field more accurately.

To quantitatively evaluate the proposed model, we compared the average of the angle difference ϵ_θ between the principle directions of the estimated tensor field and the ground truth tensor field, and the difference ϵ_{S_0} between the estimated and ground truth S_0. The results are shown in Table 1, from which it is evident that our method outperforms others and the significance in performance is more evident at higher noise levels. The average CPU time taken to converge for our method on a desktop computer with Intel 8 Core 2.8GHz, 24GB of memory, GNU Linux and MATLAB (Version 2010a) is 7.03s, whereas, NLM requires 10.52s (note both methods are executed using parallel computing).

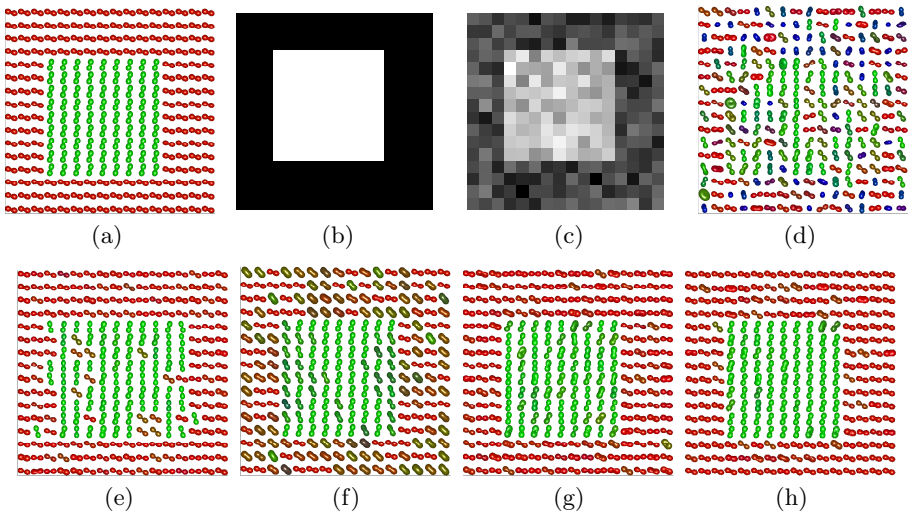


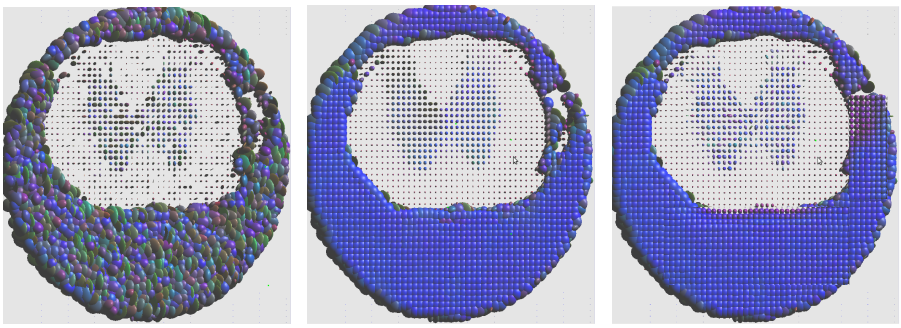
Fig. 1. (a) Ground truth synthetic DTI field, (b) the original DWI, (c) the Rician noise affected DWI, estimation using (d) MRE, (e) VF, (f) NLMt, (g) NLM, and (h) the proposed technique.

SNR	error	MRE	VF	NLMt	NLM	proposed
50	ϵ_θ	20.1 ± 12.1	8.4 ± 10.2	8.9 ± 10.7	6.0 ± 10.1	5.8 ± 7.3
	ϵ_{S_0}	0.54 ± 0.09	0.66 ± 0.05	0.64 ± 0.08	0.31 ± 0.04	0.28 ± 0.01
40	ϵ_θ	22.1 ± 12.5	12.1 ± 13.2	15.7 ± 14.2	7.2 ± 12.5	6.1 ± 8.6
	ϵ_{S_0}	0.75 ± 0.17	0.75 ± 0.31	0.94 ± 0.41	0.64 ± 0.30	0.53 ± 0.27
30	ϵ_θ	22.3 ± 12.9	19.5 ± 13.9	18.3 ± 14.7	7.6 ± 12.7	6.8 ± 9.7
	ϵ_{S_0}	2.24 ± 2.16	1.03 ± 1.22	1.03 ± 1.31	1.02 ± 1.21	0.81 ± 0.69
15	ϵ_θ	28.3 ± 17.1	27.2 ± 15.1	25.6 ± 16.2	14.7 ± 16.1	8.2 ± 10.3
	ϵ_{S_0}	3.81 ± 2.24	1.91 ± 2.02	1.86 ± 1.87	1.85 ± 1.77	1.02 ± 0.87
8	ϵ_θ	43.2 ± 23.4	32.9 ± 25.8	28.2 ± 20.6	20.2 ± 18.5	8.7 ± 11.0
	ϵ_{S_0}	5.29 ± 4.36	2.48 ± 2.72	2.29 ± 2.32	2.24 ± 2.19	1.09 ± 0.92

Table 1. Error in estimated DTI and S_0, using different methods, from synthetic DWI with different levels of noise.

3.2 DTI estimation on real datasets

We also did DTI estimation on a $124 \times 96 \times 40$ 3D rat spinal cord DWI. The data was acquired using a PGSE technique with TR=1.5s, TE=28.3ms, bandwidth=35Khz, 22 diffusion weighted images with b-value about $1014s/mm^2$. The estimated tensor field is shown in Fig. 2. We compared our method with all aforementioned methods, however due to space limit, we only present the results of MRE and NLM. From these figures, we can see that our proposed method can estimate a smoother tensor field which preserves the structure much better compared with other methods.



(a) (b) (c)

Fig. 2. The figures are the one slice of the estimated tensor fields using (a) MRE, (b) NLM, and (c) proposed method respectively.

We also did DTI estimation on a $100 \times 80 \times 32$ 3D rat brain DWI. The data was acquired using a Bruker scanner under the Spin Echo technique with TR=2s, TE=28ms, 52 diffusion weighted images with a b-value of $1334s/mm^2$.

The FA and the principal eigenvectors of the estimated tensor field are shown in Fig. 3, which showed the comparison of our method with MRE and NLM. The results illustrate that our proposed method can estimate the tensor field more accurately compared with others.

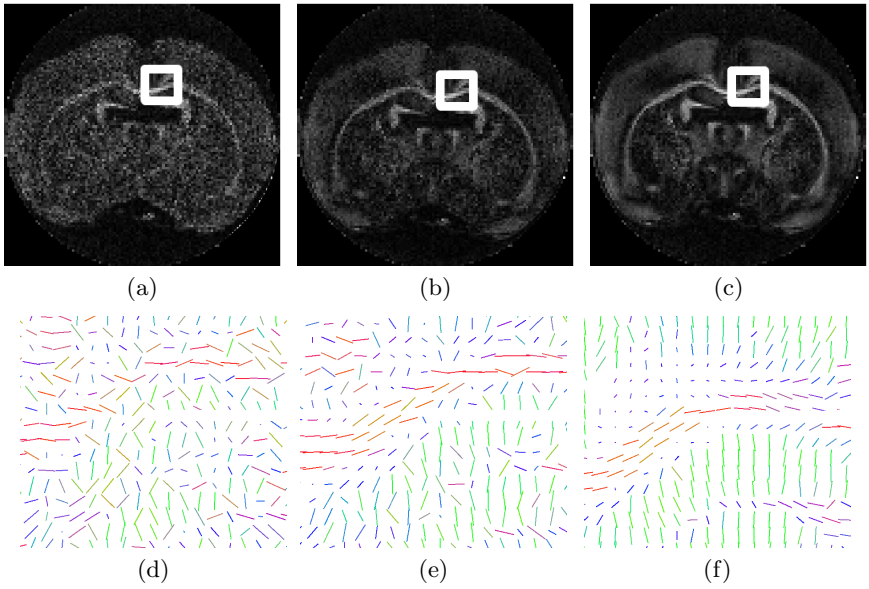


Fig. 3. The figures in the first row are the FA of the estimated tensor fields using (a) MRE, (b) NLM, (c) proposed method, and (d)(e)(f) show the principal eigenvectors in the ROI indicated with rectangles in (a), (b), and (c) respectively.

4 Conclusions

We proposed a robust variational non-local means approach for simultaneous denoising and DTI estimation. The proposed method combines the variational framework, non-local means and an intrinsically robust smoothness constraint. In the variational principle, we used non-linear diffusion tensor fitting term, along with a combination of non-local means and the tKL based smoothness for denoising. To speed up the NLM method, we prefiltered the voxels in the search window to reduce the number of computations and made use of parallel computing to decrease the computational load. This variational non-local approach was validated with synthetic and real data and shown to be more accurate than competing methods in the literature. We did not however compare with many methods in literature that were already compared to the NLM technique in [21]. For future work, we plan to develop a GPU-based implementation to better the

computation time. We will also explore other quantitative measures of validation such as method noise defined for tensors. After getting a more comprehensive tensor estimation technique, we will utilize it as a preprocessing step for the applications of fiber tracking and DTI segmentation.

References

1. Basser, P., Mattiello, J., LeBihan, D.: Estimation of the effective self-diffusion tensor from the NMR spin echo. Journal of Magnetic Resonance **103**(3) (1994) 247–254
2. Vemuri, B.C., Chen, Y., Rao, M., McGraw, T., Wang, Z., Mareci, T.: Fiber tract mapping from diffusion tensor MRI. Proceedings of IEEE Workshop Variational and Level Set Methods (2001)
3. Wang, Z., Vemuri, B.C., Chen, Y., Mareci, T.H.: A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. IEEE Transactions on Medical Imaging **23**(8) (2004) 930–939
4. Chefd'hotel, C., Tschumperlé, D., Deriche, R., Faugeras, O.: Regularizing flows for constrained matrix valued images. Journal of Mathematical Imaging and Vision **20**(1-2) (2004) 147–162
5. Tschumperlé, D., Deriche, R.: Variational frameworks for DT-MRI estimation, regularization and visualization. IEEE International Conference on Computer Vision **1** (2003)
6. Mishra, A., Lu, Y., Meng, J., Anderson, A.W., Ding, Z.: Unified framework for anisotropic interpolation and smoothing of diffusion tensor images. NeuroImage **31**(4) (2006) 1525–1535
7. Poupon, C., Roche, A., Dubois, J., Mangin, J.F., Poupon, F.: Real-time MR diffusion tensor and Q-ball imaging using Kalman filtering. Medical Image Analysis **12** (2008) 527–534
8. Tang, S., Fan, Y., Zhu, H., Yap, P., Gao, W., Lin, W., Shen, D.: Regularization of diffusion tensor field using coupled robust anisotropic diffusion. Mathematical Methods in Biomedical Image Analysis (MMBIA) (2009) 52–57
9. Tristán-Vega, A., Aja-Fernández, S.: DWI filtering using joint information for DTI and HARDI. Medical Image Analysis **14**(2) (2010) 205–218
10. Stejskal, E.O., Tanner, J.E.: Spin diffusion measurements: Spin echoes in the presence of a timedependent field gradient. Journal of Chemical Physis **42** (1965)
11. Salvador, R., Pea, A., Menon, D.K., Carpenter, T.A., Pickard, J.D., Bullmore, E.T.: Formal characterization and extension of the linearized diffusion tensor model. Human brain mapping **24**(2) (2005) 144–155
12. Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F., Connelly, A.: A rigorous framework for diffusion tensor calculus. Magnetic Resonance in Medicine **53**(1) (2005) 221–225
13. Fillard, P., Pennec, X., Arsigny, V., Ayache, N.: Clinical DT-MRI estimation, smoothing, and fiber tracking with Log-Euclidean metrics. IEEE Transactions on Image Processing **26**(11) (2007) 1472–1482
14. Hamarneh, G., Hradsky, J.: Bilateral filtering of diffusion tensor magnetic resonance images. IEEE Transactions on Image Processing **16**(10) (2007) 2463–2475
15. Pennec, X., Fillard, P., Ayache, N.: A Riemannian framework for tensor computing. International Journal of Computer Vision **66**(1) (2006) 41–66

16. Westin, C.F., Maier, S.: A dual tensor basis solution to the Stejskal-Tanner equations for DT-MRI. International Society of Magnetic Resonance in Medicine (ISMRM) (2002)
17. Koay, C.G., Basser, P.J.: Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. Journal of Magnetic Resonance **179**(2) (2006) 317–322
18. Descoteaux, M., Wiest-Daessl, N., Prima, S., Barillot, C., Deriche, R.: Impact of rician adapted non-local means filtering on HARDI. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (2008) 122–130
19. Coupé, P., Yger, P., Barillot, C.: Fast non local means denoising for 3d MR images. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) (2006) 33–40
20. Wiest-Daesslé, N., Prima, S., Coupé, P., Morrissey, S.P., Barillot, C.: Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) **10**(2) (2007) 344–351
21. Wiest-Daesslé, N., Prima, S., Coupe, P., Morrissey, S.P., Barillot, C.: Rician noise removal by non-local means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI. International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) **11**(2) (2008) 171–179
22. Castano-Moraga, C., Lenglet, C., Deriche, R., Ruiz-Alzola, J.: A Riemannian approach to anisotropic filtering of tensor fields. Signal Processing **87**(2) (2007) 217–352
23. Fillard, P., Arsigny, V., Pennec, X., Thompson, P.M., Ayache, N.: Extrapolation of sparse tensor fields: Application to the modeling of brain variability. Information Processing in Medical Imaging **19** (2005) 27–38
24. Vemuri, B.C., Liu, M., Amari, S.I., Nielsen, F.: Total Bregman divergence and its applications to DTI analysis. IEEE Transactions on Medical Imaging **30**(2) (2011) 475–483
25. Liu, M., Vemuri, B.C., Amari, S.I., Nielsen, F.: Total bregman divergence and its applications to shape retrieval. IEEE Conference on Computer Vision and Pattern Recognition (2010) 3463–3468
26. Liu, M., Vemuri, B.C.: Robust and efficient regularized boosting using total bregman divergence. IEEE Conference on Computer Vision and Pattern Recognition (2011)
27. Golub, G.H., Loan, C.F.V.: Matrix Computations. Johns Hopkins University Press (1996)
28. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer (2000)