HHS Public Access

Author manuscript

Magn Reson Med. Author manuscript; available in PMC 2014 October 08.

Published in final edited form as:

Magn Reson Med. 2014 March ; 71(3): 1272–1284. doi:10.1002/mrm.24728.

Denoising Diffusion-Weighted Magnitude MR Images using Rank and Edge Constraints

Fan Lam[1,2], **S. Derin Babacan**[2], **Justin P. Haldar**[3], **Michael W. Weiner**[4], **Norbert Schuff**[4], and **Zhi-Pei Liang**[1,2]

[1]Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois

[2]Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, Illinois

[3]Ming Hsieh Department of Electrical Engineering, University of Southern California, Los Angeles, California

[4]Department of Radiology and Biomedical Imaging, University of California, San Francisco, California

Abstract

Purpose—To improve signal-to-noise ratio (SNR) for diffusion-weighted MR images.

Methods—A new method is proposed for denoising diffusion-weighted magnitude images. The proposed method formulates the denoising problem as an *maximum a posteriori* estimation problem based on Rician/noncentral χ likelihood models, incorporating an edge prior and a low-rank model. The resulting optimization problem is solved efficiently using a half-quadratic method with an alternating minimization scheme.

Results—The performance of the proposed method has been validated using simulated and experimental data. Diffusion-weighted images and noisy data were simulated based on the diffusion tensor imaging (DTI) model and Rician/noncentral χ distributions. The simulation study (with known gold standard) shows substantial improvements in SNR and diffusion tensor estimation after denoising. In-vivo diffusion imaging data at different *b*-values were acquired. Based on the experimental data, qualitative improvement in image quality and quantitative im-provement in diffusion tensor estimation were demonstrated. Additionally, the proposed method is shown to outperform one of the state-of-the-art non-local means based denoising algorithms, both qualitatively and quantitatively.

Conclusion—The SNR of diffusion-weighted images can be effectively improved with rank and edge constraints, resulting in an improvement in diffusion parameter estimation accuracy.

Correspondence to: Fan Lam, Beckman Institute for Advanced Science and Technology, 405 N. Mathews Ave, Urbana, IL 61801 USA, fanlam1@illinois.edu.

Keywords

Diffusion-weighted imaging; diffusion tensor imaging; Rician distribution; noncentral χ distribution; low-rank approximation; edge constraints

Introduction

Low SNR has been a major limiting factor for high resolution diffusion-weighted (DW) imaging and is often problematic for the quantification of diffusion properties of tissues (1, 2). Significant efforts have been made to address this problem. Straightforward approaches, including signal averaging and reducing *k*-space coverage, have found some practical use, but they have noticeable limitations, such as long data acquisition time or low spatial resolution.

An alternative approach is to acquire high-resolution (noisy) data, followed by denoising. Denoising can be applied to either complex-valued or magnitude DW images. The method in (3) assumes a diffusion tensor model and estimates the tensor parameters directly from complex-valued DW images, incorporating smoothness constraints on both the DW images and the diffusion tensor field. The method in (4) imposes no specific diffusion model and performs joint image reconstruction and denoising, utilizing an edge constraint that exploits the correlation of edge structures in DW image series for effective edge-preserving smoothing. Although denoising complex data is generally advantageous (e.g., smaller bias and easier characterization), it is desirable to effectively denoise magnitude data for a number of practical reasons (such as availability and free of phase artifacts (5)).

A key challenge in denoising magnitude images lies in the handling of signal-dependent, non-Gaussian noise (6,7). While the measurement noise in complex *k*-space data can be modeled as signal-independent Gaussian noise, noisy magnitude images often follow Rician distribution (8, 9) or noncentral χ distribution (10–12). Based on these distributions, several meth-ods have been proposed for denoising DW magnitude data. The methods in (13, 14) assume a specific parametric diffusion model and directly estimate the diffusion parameters from noisy magnitude data using Rician likelihood model. The methods in (15–19) do not assume any diffusion model and apply Rician/noncentral χ distribution based image denoising followed by a separate diffusion parameter estimation. In order to denoise DW magnitude data effectively, a variety of constraints that exploit prior information about DW images have been used in the existing literature. They include local smoothness (20–24), image sparsity (25) and non-local similarities (26–29). The reader can refer to (4) for a more complete list of related references.

This paper presents a new method to jointly denoise a sequence of DW magnitude images, combining the following modeling components. First, we model the noisy magnitude data using Rician/noncentral χ distribution, providing flexibility to process images from a range of reconstruction schemes (11). Second, we impose a low-rank structure on the desired DW image series, which helps separate the underlying signal from noise, assuming that the true signal lies in a low-dimensional subspace. The benefits of low-rank modeling have previously been demonstrated in various MRI applications, including reconstruction from

highly undersampled *k-t* space data (30–33) and denoising (34–36). Third, we incorporate a special edge constraint. This constraint exploits the correlation of edge structures in a DW image series for effective edge-preserving smoothing (4). We formulate the denoising problem within an *maximum a posteriori* (MAP) estimation framework that integrates these three components and solve the associated optimization problem using an efficient alternating minimization algorithm. The proposed method is the first attempt to incorporate noisy signal modeling, low-rank modeling and edge-preserving prior in a unified mathematical framework for denoising magnitude MR image series. The proposed method is described in detail in the subsequent section. Its performance will be demonstrated using both simulated and experimental data in the Results section, followed by discussions and conclusions.

Theory

In a typical diffusion imaging experiment, a sequence of DW images is acquired for quantifying diffusion properties at each voxel. Given a sequence of Q noisy DW magnitude images $Y = [y_1, y_2, \ldots, y_Q]$, where the vector $y_i \in R^M$ represents an individual image frame with M pixels, we jointly denoise all the images in Y using the following MAP estimation formulation

$$\hat{A} = \arg\max_A p(A|Y) = \arg\min_A -\log p(Y|A) - \log p(A), \quad [1]$$

where $A = [a_1, a_2, \ldots, a_Q]$ represents the underlying noise-free image series expressed in matrix form with $a_i \in R^M$ being the i^{th} image frame, such that

$$A = [a_1, a_2, \ldots, a_Q] = \begin{bmatrix} a_{11} & \cdots & a_{1Q} \\ \cdots & \ddots & \cdots \\ a_{M1} & \cdots & a_{MQ} \end{bmatrix}. \quad [2]$$

We use $\hat{A}$ to denote the denoised image series, $p(Y|A)$ to represent the likelihood function capturing the distribution of the noisy data, and $p(A)$ to represent the image prior. We next describe $p(Y|A)$ and $p(A)$, and how to solve Eq. [1] with a low-rank constraint on A.

Noise Model

It is generally accepted that the real and imaginary parts of a complex-valued image from any linear reconstruction scheme follow Gaussian distributions with the same noise variance. There-fore, the magnitude image intensity at each voxel, computed as $y = \sqrt{y_R^2 + y_I^2}$, with y_R and y_I being the real and imaginary parts respectively, can be modeled by the Rician distribution (5) as

$$p(y|a) = \frac{y}{\sigma^2} \exp\left(-\frac{y^2 + a^2}{2\sigma^2}\right) I_0\left(\frac{ay}{\sigma^2}\right), \quad [3]$$

where a is the corresponding noise-free intensity, $I_0(\cdot)$ is the zeroth-order modified Bessel function of the first kind, and σ^2 is the variance of the Gaussian noise in the real and imaginary parts.

Assuming independent noise distribution at each voxel and spatially invariant noise variance, the joint likelihood of all voxels p (Y|A) can be expressed as a product of Rician distributions as

$$p(Y|A) = \prod_{q=1}^{Q} \prod_{m=1}^{M} \frac{y_{mq}}{\sigma^2} \exp\left(-\frac{y_{mq}^2 + a_{mq}^2}{2\sigma^2}\right) I_0\left(\frac{a_{mq} y_{mq}}{\sigma^2}\right), \quad [4]$$

where m and q are indices for voxel positions and image frames, respectively.

If the magnitude image is obtained by sum-of-squares combination of images from multiple coils as $y = \sqrt{\sum_{c=1}^{C} (y_{R_c}^2 + y_{I_c}^2)}$ with c being the coil index, the intensity can be modeled by the noncentral χ distribution (10) as

$$p(y|a) = \frac{1}{\sigma^2} \left(\frac{y^C}{a^{C-1}}\right) \exp\left(-\frac{y^2 + a^2}{2\sigma^2}\right) I_{C-1}\left(\frac{ay}{\sigma^2}\right), \quad [5]$$

assuming uncorrelated noise and uniform variance across different coils. C is the number of coils and $I_C(\cdot)$ is the Cth-order modified Bessel function of the first kind. Accordingly, the joint likelihood function can be extended to sum-of-squares images, using a product of noncentral χ distributions as

$$p(Y|A) = \prod_{q=1}^{Q} \prod_{m=1}^{M} \frac{1}{\sigma^2} \left(\frac{y_{mq}^C}{a_{mq}^{C-1}}\right) \exp\left(-\frac{y_{mq}^2 + a_{mq}^2}{2\sigma^2}\right) I_{C-1}\left(\frac{a_{mq}\, y_{mq}}{\sigma^2}\right). \quad [6]$$

If the correlation between different coils needs to be considered, the noncentral χ model can still be used with an effective noise variance σ_{eff}^2 and an effective number of coils C_{eff}, as described in (12). The likelihood models in Eqs. [4] and [6] can also be readily extended to the case of spatially varying noise by substituting σ with a voxel dependent σ_m.

Image Model

The proposed image model has two key features: a low-rank model for A and a special edge prior p (A). First, we express the image series A as (31,37)

$$A = \text{UV}, \quad [7]$$

where $\text{U} \in R^{M \times L}$ and $\text{V} \in R^{L \times Q}$ are low-rank matrices with $L \ll Q < M$. This representation is motivated by and based on the assumption that there is a strong spatial correlation of tissue diffusion properties, whether the diffusion properties are described by a tensor (38) or higher-order models (39–41). This correlation (referred to as spatial-diffusion

correlation here) can be well captured by the Partial Separability (PS) model (30,31), which leads to the low-rank representation of the image series A as expressed in Eq. [7].

Second, we use the following prior for edge-preserving smoothing (the log form of this prior is equivalent to the regularization penalty proposed in (4))

$$p(A) \propto \exp\left(-\lambda \sum_{m=1}^{M} \sum_{n \in \Omega_m} H\left(\sqrt{\sum_{q=1}^{Q} \omega_q^2 |a_{mq} - a_{nq}|^2}\right)\right). \quad [8]$$

In Eq. [8], $H(\cdot)$ can be any edge-preserving regularization function, w_q are the edge weights for the q^{th} image frame, Ω_m represents a neighborhood of the m^{th} voxel, which contains its adjacent voxels along the horizontal and vertical directions, and λ is a regularization parameter. It was demonstrated in (4) that DW image series have highly correlated edge structures that can be enforced via the prior in Eq. [8]. The edge weights, w_q can be chosen according to different denoising preferences. For example, assigning larger weights to the images acquired at higher *b*-values will help preserve image features revealed with heavier diffusion weightings. For more comprehensive discussions on choosing w_q, see (4,42).

Joint Denoising Formulation

For DW images following Rician distribution, we integrate the Rician likelihood model in Eq. [4], the low-rank model in Eq. [7] and the edge prior in Eq. [8] into the MAP formulation in Eq. [1], yielding the following optimization problem

$$\hat{U}, \hat{V} = \arg\min_{U,V} \sum_{q=1}^{Q} \sum_{m=1}^{M} \left[\frac{y_{mq}^2 + (\mathrm{UV})_{mq}^2}{2\sigma^2} - \log I_0\left(\frac{y_{mq}(\mathrm{UV})_{mq}}{\sigma^2}\right)\right] + \lambda R(U,V). \quad [9]$$

where $\hat{\mathrm{U}}$, $\hat{\mathrm{V}}$ are estimates of U and V. The regularization term $R(\cdot)$ is given by

$$R(U,V) = \sum_{m=1}^{M} \sum_{n \in \Omega_m} H\left(\sqrt{\sum_{q=1}^{Q} \omega_q^2 |(\mathrm{UV})_{mq} - (\mathrm{UV})_{nq}|^2}\right). \quad [10]$$

The denoised image series is computed as $\hat{\mathrm{A}} = \hat{\mathrm{U}}\hat{\mathrm{V}}$.

For DW images following noncentral χ distribution (sum-of-squares combination), the likelihood function is replaced by the noncentral χ distribution model in Eq. [6], which yields the following optimization problem

$$\hat{U}, \hat{V} = \arg\min_{U,V} \sum_{q=1}^{Q} \sum_{m=1}^{M} \left[(C-1)\log(\mathrm{UV})_{mq} + \frac{y_{mq}^2 + (\mathrm{UV})_{mq}^2}{2\sigma^2} - \log I_{C-1}\left(\frac{y_{mq}(\mathrm{UV})_{mq}}{\sigma^2}\right)\right] + \lambda R(U,V). \quad [11]$$

The proposed formulations in Eqs. [9] and [11] can be adapted to handle various special cases. For instance, the estimation of a full rank A from noncentral χ distributed data can be reformulated without explicit rank modeling as follows:

$$\hat{A}=\arg\min_{A}\sum_{q=1}^{Q}\sum_{m=1}^{M}\left[(C-1)\log(a_{mq})+\frac{y_{mq}^2+a_{mq}^2}{2\sigma^2}-\log I_{C-1}\left(\frac{y_{mq}a_{mq}}{\sigma^2}\right)+\lambda\sum_{m=1}^{M}\sum_{n\in\Omega m}H\left(\sqrt{\sum_{q=1}^{Q}\omega_q^2|a_{mq}-a_{\mathrm{nq}}^2|}\right)\right], \quad [12]$$

where $C = 1$ for Rician distributed data. In this case, only the spatial edge constraint is utilized.

Algorithm

A multiplicative half-quadratic method (4,43) combined with an alternating minimization scheme is used to solve the problems in Eq. [9] and Eq. [11]. Specifically, we use the Huber function (4,44) for $H(\cdot)$ in this paper, and the proposed algorithm performs an iterative reweighted ℓ_2 regularization by first solving

$$\hat{U},\hat{V}=\underset{U,V}{\operatorname{argmin}}\left\{-\log p(Y|A=\mathrm{UV})+\lambda\sum_{q=1}^{Q}\sum_{m=1}^{M}\sum_{n\in\Omega_m}l_{nm}\omega_q^2\left|(\mathrm{UV})_{mq}-(\mathrm{UV})_{nq}\right|^2\right\}, \quad [13]$$

followed by updating the weights l_{nm} using

$$l_{nm}=\begin{cases}1 & \text{if } t_{nm}<\alpha\\ \alpha/t_{nm} & \text{if } t_{nm}\geq\alpha\end{cases}, \quad [14]$$

where α is a manually chosen threshold and

$$t_{nm}=\sqrt{\sum_{q=1}^{Q}\omega_q^2\left|(\hat{U}\hat{V})_{mq}-(\hat{U}\hat{V})_{nq}\right|^2}. \quad [15]$$

The steps in Eqs. [13] and [14] are alternated until convergence. As can be seen, in each iteration, the prior in Eq. [8] leads to a local edge penalty modulated by the line variables l_{nm} estimated using edge information from all the image frames. At spatial regions where l_{nm} is large, heavy smoothing is applied, while at regions with small l_{nm} (edge structure present), weaker smoothing is applied.

For a fixed l_{nm}, U and V are estimated alternatively. At the k^{th} iteration, we update U with $\mathrm{V}^{(\hat{k})}$ by solving

$$\hat{U}^{(k)} = \arg \min_{U \in R^{M \times L}} \left\{ -\log p\left(Y | A = U\hat{V}^{(k)}\right) + \lambda \sum_{q=1}^{Q} \omega_q^2 \| \mathrm{WB}\left(U\hat{V}^{(k)}\right)_q \|_2^2 \right\}; \quad [16]$$

where B is a finite difference operator and W is a diagonal weighting matrix composed of l_{nm}.

For a fixed $\hat{U}^{(k)}$, optimizing V can be done by solving the following Q decoupled sub-problems (a similar decomposition can be found in (45)):

$$\hat{V}_q^{(k+1)} = \arg \min_{v_q \in R^L} \left\{ -\log p\left(y_q | a_q = \hat{U}^{(k)} v_q\right) + \lambda \omega_q^2 \| \mathrm{WB}\hat{U}^{(k)} v_q \|_2^2 \right\}, \text{ for } q = 1 \ldots Q. \quad [17]$$

Each sub-problem has a very low dimensionality L and is easy to solve. Due to the Rician/noncentral χ likelihood, nonlinear optimization should be used to solve the alternating minimization problems. To solve the problem in Eq. [16], we used the limited-memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) method (46), which was designed for large scale problems where the approximated Hessian cannot be saved explicitly. The standard BFGS method (46) was used for solving each sub-problem in Eq. [17]. We have also derived explicit expressions for the gradients of the optimization cost functions in Eqs. [16] and [17], which lead to efficient computations using matrix-vector multiplications (see the Appendix for derivations).

The algorithm is initialized by applying a singular value decomposition (SVD) on Y and extracting the first L eigenvectors as initial estimates $\hat{U}^{(0)}$ and $\hat{V^{(0)}}$. L was selected empirically by examining the singular value distribution of Y and choosing a proper cut-off threshold. The noise variance σ^2 was estimated from the background regions with negligible signals (5,15). Notice that when $a_{mq} = 0$, the Rician distribution becomes the Rayleigh distribution with a second moment $E\left(y_{mq}^2\right) = 2\sigma^2$ (5). Therefore, σ^2 can be estimated as half of the mean square data intensity in a background region. For noncentral χ distributed data, the second moment for the background signal can be computed similarly as $E\left(y_{mq}^2\right) = 2C\sigma^2$ with C the number of coils (10). Thus, we can again use the mean square intensity in the background to estimate σ^2. Assuming that all images within the series have similar noise levels, the accuracy of this estimation can be improved by averaging the estimates of σ^2 from all the frames. For the edge prior, we assigned slightly larger w_q for images with higher b-values. The regularization parameter λ was chosen based on visual inspection of both the denoised images and the estimated diffusion parameter maps.

Methods

Simulations

DW images were simulated using an adult mouse diffusion tensor atlas from the Biomedical Informatics Research Network (BIRN) Data Repository. The image intensity a_{mq} of the m^{th} voxel in the q^{th} frame was generated using the tensor model as follows:

$$a_{mq} = a_{m0} \exp(-b_q u_q^T D_m u_q),\ m=1,\ldots M,\ q=1,\ldots,Q, \quad [18]$$

where a_{m0} is the intensity without diffusion weighting, b_q is the b-value for the q^{th} frame, u_q specifies the diffusion encoding direction, and D_m is the diffusion tensor for the m^{th} voxel. Forty-five images with 256 × 256 pixels were generated, including one image with $b = 0$ and 44 images with $b = 2000 s/mm^2$ in different diffusion encoding directions. Independent Rician noise was simulated by adding white Gaussian noise to the real and imaginary parts of the images. The noise standard deviation was chosen as 1/10 of the mean intensity in the center region of the first frame (the image with b = 0). We then applied four processing schemes to the same noisy image series to illustrate the effect of each constraint in the proposed method. They are: (1) no denoising; (2) denoising with Gaussian distribution and edge constraints (Eq. [12] with Gaussian likelihood); (3) denoising with Rician distribution and edge constraints (Eq. [12] with C = 1); and (4) denoising with Rician distribution and both low-rank model and edge constraints (Eq. [9]). The low-rank model order was $L = 8$. The weightings for the spatial prior were $w_1 = 1$ for the first frame, and $w_q = 1.5$ ($q \geq 2$) for the rest of the series. The regularization parameters λ for different schemes were chosen such that the noise variances from the background were similar after denoising. The tensors and FA estimates for all cases were produced for comparison using the standard least-squares approach in (47) (same for all the results presented in this paper).

The proposed method was also compared with one of the state-of-the-art image denoising methods: the multi-component non-local means (MNLM) algorithm (29). The algorithm applies NLM filter (26) to the images using filtering kernels defined on image blocks, which is based on a completely different model compared to the proposed method. The original algorithm works for Gaussian noise and needs modification for Rician case. Specifically, noting that the squared intensity of the simulated noisy images follows a chi-square distribution with expectation $E(y^2) = E(a^2) + 2\sigma^2$, we applied the MNLM algorithm to the squared magnitude images and the output intensities were then shifted by $2\sigma^2$ to obtain the estimated squared intensities, as described in (26).

The same noise-free DW image series was also used to generate noncentral χ distributed data to simulate noisy sum-of-squares images. The noisy intensity in this case was generated as

$$y = \sqrt{\sum_{c=1}^{C} [(a_c + n_{1c})^2 + (n_{2c})^2]},$$

where $a_c = a$ is the noise-free intensity for each coil, n_{1c} and n_{2c} are white Gaussian noise with the same standard deviation as in the Rician distributed data. A total of four coils were simulated ($C = 4$) and uniform sensitivity maps were assumed for each coil (noting that the sensitivity map does not affect noise distribution). The proposed method was applied to denoise the simulated series with the noncentral χ likelihood model (Eq. [11]) with $L = 8$ and the same w_q assignment as in the Rician noise case. The MNLM algorithm was also

applied to the squared image series and the output intensities were shifted by $2C\sigma^2$, considering that the squared intensity follows a noncentral chi-square distribution with expectation $E(y^2) = E(a^2) + 2C\sigma^2$.

Experimental data

In-vivo human brain diffusion imaging data from healthy volunteers was used to evaluate the performance of the proposed method in practical scenarios. The first data set was obtained from a multi-*b*-value diffusion experiment. The data was acquired on a 3T Siemens scanner (Trio, Siemens Healthcare USA) equipped with a 32-channel receiver headcoil, using a double spin-echo EPI sequence with partial Fourier sampling (5/8 of the *k*-space in phase-encoding direction). The imaging parameters were: TR/TE = 6300/98ms, matrix size = 128 × 128 with a FOV of 256 × 256mm^2 and 35 slices with 2mm thickness. The DW image series was acquired at multiple *b*-values (b = 1000, 2000, 3000s/mm^2) in 30 diffusion encoding directions for each nonzero *b*-value. Partial Fourier reconstruction (48) was applied to each coil to obtain the original noisy images. The proposed method with Rician likelihood model was then applied to denoise the reconstructed image series coil by coil with $L = 16$. λ was chosen based on visual inspection of the denoised images to balance between noise reduction and oversmoothing (a similar procedure was applied for all the in-vivo data). Sum-of-squares was used to combine the denoised images from all coils for diffusion tensor estimation.

The second data set was from a DTI experiment on a different 3T Siemens scanner (Skyra, Siemens Healthcare USA) equipped with a standard 16-channel receiver headcoil, using a single spin-echo EPI sequence. The data acquisition parameters were: TR/TE = 7800/91ms, 2 × 2 mm^2 in-plane resolution and 2mm slice thickness. Diffusion encodings along 64 directions with b = 2000s/mm^2 and a $b = 0$ image were acquired. Parallel imaging with an acceleration factor of 2 was used and GRAPPA reconstruction (49) was performed. The proposed method with noncentral χ likelihood model was applied to denoise the sum-of-squares image series, with $L = 16$ and λ chosen by visual inspection.

The third data set was again acquired on the 3T Siemens Skyra using the same sequence as above. Sixty-four diffusion directions with b = 2000s/mm^2 were acquired covering a full sphere in the diffusion *q*-space (50, 51). In addition, the diffusion directions in the left hemisphere (in *q*-space) are symmetric to the diffusion directions in the right hemisphere. Again, GRAPPA reconstruction with an acceleration factor of 2 was applied. The diffusion tensors estimated from either hemisphere before and after denoising were compared. Theoretically, the tensors estimated from either hemisphere in *q*-space should be the same. However, the presence of noise alters this symmetry. Therefore, this comparison provides a direct quantitative measure of noise reduction. A similar approach was used in (4).

Results

Representative results from both simulated and in-vivo data are shown to illustrate the effectiveness of the proposed method.

Simulation Results

Figure 1 shows some representative DW images (for the same diffusion encoding direction), calculated FA maps and the corresponding error maps for the four alternative schemes (described in the first paragraph of the Methods section). As can be seen, effective noise reduction due to edge constraints can be observed in both columns 2 and 3. However, the Gaussian noise assumption introduced noticeable bias (see the error image and FA error map in column 2). The results in column 4 have the smallest errors in both the DW images and FA maps, illustrating the advantages of integrating the data model, the low-rank model and the edge constraints.

Figure 2 compares representative image frames, FA maps and color-coded FA maps from the noisy images, the MNLM algorithm and the proposed method, for the simulated Rician distributed data. Zoom-in images corresponding to a specific brain region are also shown. As can be seen, the proposed method (column 4) not only significantly reduces the noise level but also helps recover important diffusion features originally contaminated by noise. The denoising results from the MNLM algorithm (column 3) look good visually but oversmoothing is present in several regions. The difference between MNLM and the proposed method can be seen more clearly in the zoom-in regions. Figure 3 compares the quantification results from different methods. Note that FA errors of the proposed method (column 3) are more concentrated around zero (i.e., lower bias and variance) and less dependent on the true FA values, indicating the effectiveness of the proposed method in reducing noise while maintaining the anisotropy of the diffusion tensors.

Figure 4 compares the denoising results for the simulated noncentral χ distributed data (sum-of-squares images). Figure 5 shows the corresponding comparison of the FA estimation errors. The image quality and quantification accuracy of the proposed method are again superior to those of the alternative methods.

Figure 6 shows a more comprehensive quantitative comparison between different cases under different noise levels. The peak signal-to-noise ratio (PSNR), computed as

$$\mathrm{PSNR} = 10\log_{10}\frac{M}{\|a_q - \hat{a}_q\|_2^2}, \quad [19]$$

is shown for representative frames from each method. a_q and $\hat{A}_q$ in the equation correspond to the noise-free image and the noisy/denoised images, respectively. PSNR is commonly used for measuring the approximation accuracy of a noisy/reconstructed image to the original image. Higher PSNR (higher approximation accuracy) usually indicates better visual quality. However, better visual quality does not necessarily lead to higher parameter estimation accuracy in quantitative diffusion imaging. Therefore, we also computed the root-mean-squared-error (RMSE) of the estimated FA values, defined as

$$\text{FA}-\text{RMSE}=\sqrt{\frac{\sum_{m=1}^{M_1}(FA_m-\hat{FA}_m)^2}{M_1}}, \quad [20]$$

for evaluating quantification accuracy. *FA* and $\hat{FA}$ in the equation are FA maps estimated from the noise-free images and images from different methods, respectively, and M_1 is the number of voxels extracted from the non-background regions for computing the RMSE. It can be observed from the curves that the proposed method significantly improves the image quality while providing more accurate parameter estimation. The MNLM algorithm provides very good PSNR results but the quantification accuracy is lower compared to the proposed method. The difference is even more obvious for the noncentral χ distributed data. Also note that the proposed method can handle very high noise levels.

To examine the effects of L (the low-rank model order) and λ (the regularization parameter) on the denoising performance, simulations were run with different combinations of L and λ for the Rician noise case. Figure 7 compares the PSNR and FA-RMSE for different parameter combinations. As can be seen, although lower ranks yield better PSNR, the diffusion parameter estimation accuracy degrades significantly if L is too small (< 6), as expected. If L is too high, the denoising performance degrades in terms of both PSNR and FA-RMSE, because the rank constraint was not effective in this case. The proposed method is in general robust to a reasonable range of λ values. Furthermore, we observe that, to achieve a higher PSNR (better visual image quality), using edge constraints only (full rank, $L = Q$) will tend to favor larger regularization parameters, which can degrade the parameter estimation accuracy. Using the joint constraints provides a better trade-off for both metrics. This feature allows us to choose a proper λ by visual inspection of the denoised images.

Experimental Results

Figure 8 shows the denoising results of the proposed method for one slice of the in-vivo Rician distributed data. Representative DW images, FA maps and color-coded FA maps corresponding to the three different *b*-values are shown. Although the gold standard is not available, qualitative improvement is evident as shown by the denoised images and the parameter maps. Before denoising, several diffusion structures (as identified by the red arrows in the DW images) were corrupted by noise, especially for higher *b*-values. After denoising by the proposed method, these structures became better defined. Furthermore, the proposed method significantly reduces noise contamination in the diffusion parameters. The fiber bundles appeared much more uniform after denoising while the boundaries remained sharp (as identified by the red arrows in the RGB images). The improvement is consistent across different *b*-values.

Denoising results from the in-vivo noncentral χ distributed data are shown in Figure 9. One slice of the DW images and the corresponding FA maps are shown, which illustrate effective noise removal and recovery of diffusion parameters achieved by the proposed method for practical sum-of-squares reconstructions.

Figure 10 compares the histograms of Log-Euclidean distances (52) between tensors estimated from the right hemisphere and the left hemisphere of q-space, before and after denoising. The average Log-Euclidean distances are also calculated. It is well known that diffusion data are symmetric in q-space. This symmetry provides another mean to validate our method on experimental data without a gold standard. As can be seen, the symmetry between tensors estimated from each of the two hemispheres is significantly improved after denoising. This result quantitatively demonstrates the effectiveness of the proposed method in reducing the uncertainty introduced by signal-dependent noise in the original magnitude data.

Discussion and Conclusions

We have demonstrated effective denoising of DW magnitude image series by combining noisy magnitude signal modeling with rank and edge constraints. The proposed method has two desirable advantages. First, the proposed method introduces a unified mathematical framework that integrates low-rank modeling and a spatial prior to maximize the noise reduction effect while preserving important diffusion information. Second, the capability to incorporate both Rician and noncentral χ distributions into the MAP estimation framework allows for effective denoising of magnitude images from both single-coil and multiple-coil acquisitions using an image prior. Note that magnitude images from a single coil follow Rician distribution (5, 8, 9), while the noisy reconstructions from multiple coils depend on the reconstruction scheme used (11). For example, GRAPPA reconstructions with sum-of-squares combination can be assumed to follow noncentral χ distribution, while SENSE reconstructions can be assumed to follow Rician distribution. For multiple-coil acquisitions, if reconstructions from individual coils were available, denoising them individually before combination would be better because (1) the Rician distribution is simpler, and (2) correlation between coils could complicate the distribution of the combined data. However, in many practical situations, we do not have reconstructions from each individual coil. Therefore, the capability of denoising with a more complicated distribution (e.g., noncentral χ distributions) increases the flexibility and applicability of the proposed method.

As with most denoising methods, the performance of the proposed method relies on an accurate estimation of the noise variance σ^2. The estimation method used in this paper is simple and provides highly accurate results when images contain reasonably large background regions (with noise only). If the background regions contain artifacts (e.g., due to aliasing or motion artifacts), the estimation of σ^2 is less accurate. For these more general scenarios, a variety of robust noise estimation methods have been proposed. They include simultaneous background detection and noise estimation (53, 54), robust noise estimation for images without background regions (55–57), and estimation of spatially varying noise (28, 58, 59). Proper integration of the proposed method with these noise estimation methods may be desirable for handling more general data acquisition and reconstruction schemes. A thorough investigation of this issue is application-specific and beyond the scope of this paper.

Theoretically, the rank L of the image series A is determined by the number of distinct types of diffusion properties in the tissues (whether it is described by a tensor or higher-order

models). Practically, L can be small without introducing significant low-rank approximation error due to the spatial-diffusion correlation, which seems to be strongly supported by our results. In general, the optimized rank choice will be application dependent. For DTI, a lower rank is reasonable since every tensor lies in a six-dimensional subspace. But for high angular resolution diffusion imaging, a higher rank is necessary to avoid oversmoothing complex fiber structures.

Choosing a proper regularization parameter λ is also important to the proposed method. In practical denoising scenarios, calculating the optimal λ is very challenging with the absence of a gold standard. The discrepancy principle and generalized cross-validation (60, 61) are useful for choosing regularization parameters in least-squares problems with Tikhonov regularization, but they are not directly applicable in our case due to the nonlinearity of the problem and non-Gaussian noise model. In this paper, we simply chose λ by visual inspection and it has provided reasonably good results empirically. However, as indicated by the results in Fig. 7, λ can be "optimized" with respect to a specific data acquisition scheme and application scenario using a "reference" data set and then be translated to practical measurements, but this issue needs to be further investigated. Furthermore, establishing an objective metric for choosing λ is worth pursuing in future research.

In summary, a new method for denoising a sequence of diffusion-weighted magnitude images has been described. The proposed method formulates the denoising problem within an MAP estimation framework, incorporating noisy magnitude signal modeling, low-rank approximation of the image series, and a special edge prior. An efficient algorithm is used to solve the associated optimization problem. Representative results from simulated and experimental data demonstrate that the proposed method substantially improves both the image quality and the diffusion parameter estimation accuracy. In addition, significant FA estimation error reduction is achieved for a wide range of anisotropy levels, which is especially important for comparing diffusion properties of different tissues or conducting longitudinal study of tissue changes for a certain disease. We expect the proposed method to prove useful for reducing noise-induced uncertainty or improving spatial resolution for quantitative diffusion MRI with magnitude data. Furthermore, although the proposed method was developed for DW images, it could also be used in different applications where a sequence of magnitude images is acquired and analyzed, e.g., functional MRI and parametric mapping. Such extensions could be investigated in the future.

Acknowledgments

This work was supported in part by NIH-P41-EB015904, NIH-P41-EB001977 and NIH-1RO1-EB013695.

Appendix

Gradient of Equation [16] for the U sub-problem

Defining $C_q(U)=L_q(U)+\lambda\|\mathrm{WB}(U\hat{V}^{(k)})_q\|_2^2$, where

$$L_q(U)=\sum_{m=1}^{M}\frac{\left(U\hat{V}^{(k)}\right)_{mq}^2}{2\sigma^2}-\log I_0\left(\frac{y_{mq}\left(U\hat{V}^{(k)}\right)_{mq}}{\sigma^2}\right), \quad [21]$$

we can express the cost function in Eq. [16] as $f(U)=\sum_{q=1}^{Q}C_q(U)$. Then the gradient of f(U) can be written as

$$\begin{aligned}\nabla f(U)&=\sum_{q=1}^{Q}\nabla C_q(U)\\&=\sum_{q=1}^{Q}\left[\nabla L_q(U)+2\lambda(\mathrm{WB})^T\mathrm{WBU}\hat{v}_q^{(k)}(\hat{v}_q^{(k)})^T\right].\end{aligned} \quad [22]$$

Where $\left[\nabla L_q(U)\right]_{ij}=\partial L_q(U)/\partial U_{ij}$ and $\hat{v}_q^{(k)}$ is the q^{th} column of $\mathrm{V}^{(\hat{k})}$. For brevity, let us denote

$$a_{mq}=\frac{y_{mq}(\mathrm{UV}^{(k)})_{mq}}{\sigma^2}.$$

Then, the directional derivative $[\nabla L_q(\mathrm{U})]_{ij}$ is given as

$$\begin{aligned}\frac{\partial L_q(U)}{\partial U_{ij}}&=\sum_{m=1}^{M}\frac{1}{2\sigma^2}\frac{\partial\left(\sum_{l=1}^{L}U_{ml}\hat{V}_{lq}^{(k)}\right)^2}{\partial U_{ij}}-\sum_{m=1}^{M}\frac{I_1(a_{mq})}{I_0(a_{mq})}\frac{\partial y_{mq}\left(U\hat{V}^{(k)}\right)_{mq}}{\sigma^2\partial U_{ij}}\\&=\frac{\sum_{l=1}^{L}U_{il}\hat{V}_{lq}^{(k)}}{\sigma^2}\hat{V}_{jq}^{(k)}-\frac{I_1(a_{iq})}{\sigma^2 I_0(a_{iq})}y_{iq}\hat{V}_{jq}^{(k)}\\&=\left(\left(U\hat{V}^{(k)}\right)_{iq}-\frac{I_1(a_{iq})}{I_0(a_{iq})}y_{iq}\right)\frac{\hat{V}_{jq}^{(k)}}{\sigma^2}.\end{aligned}$$

Using this equation, the gradient $\nabla L_q(\mathrm{U})\in R^{M\times L}$ can be expressed in a matrix-vector multiplication form as

$$\nabla L_q(U)=\frac{1}{\sigma^2}\left(\left(U\hat{V}^{(k)}\right)_q-\frac{I_1(a_q)}{I_0(a_q)}\odot y_q\right)\left(\hat{V}_q^{(k)}\right)^T, \quad [23]$$

where $a_q=\dfrac{y_q\odot\left(U\hat{v}_q^{(k)}\right)}{\sigma^2}$ is a column vector, $\odot$ denotes element-wise multiplication and $I_{0/1}(\cdot)$ stands for element-wise evaluation of modified Bessel function of the first kind. Therefore, the gradient can be computed efficiently using this closed-form expression. Note that the variable being updated is a matrix, thus we are taking derivative w.r.t. each element in this matrix. However, in the actual implementation, the variable and the gradient can be vectorized to fit into an optimization routine.

Gradient of Equation [17] for the V sub-problem

Defining $g(\mathrm{v}_q)$ as the cost function, we have

$$g(v_q)=L(v_q)+\lambda R(v_q), \quad [24]$$

where

$$L(v_q)=\sum_{m=1}^{M}\frac{\left(\hat{U}^{(k)}v_q\right)_m^2}{2\sigma^2}-\log I_0\left(\frac{y_{mq}\left(\hat{U}^{(k)}v_q\right)_m}{\sigma^2}\right),$$

and $R(v_q)=\lambda\|\mathrm{WB}\hat{U}^{(k)}v_q\|_2^2$. Following a similar strategy for deriving the gradient of $f(\mathrm{U})$ in the previous section, we write the directional derivative of $L(\mathrm{v}_q)$ as follows:

$$\begin{aligned}\frac{\partial L(v_q)}{\partial v_{q,l}}&=\sum_{m=1}^{M}\frac{\left(\hat{U}^{(k)}v_q\right)_m\hat{U}^{(k)}_{m,l}}{\sigma^2}-\frac{I_1(a_{mq})}{I_0(a_{mq})}\frac{y_{mq}\hat{U}^{(k)}_{m,l}}{\sigma^2}\\&=\sum_{m=1}^{M}\left(\left(\hat{U}^{(k)}v_q\right)_m-\frac{I_1(a_{mq})}{I_0(a_{mq})}y_{mq}\right)\frac{\hat{U}^{(k)}_{m,l}}{\sigma^2},\end{aligned} \quad [25]$$

where a_{mq} is defined the same way as in the previous section. According to Eq. [25], to compute the derivative of $L(\mathrm{v}_q)$ w.r.t. the lth element, we first compute a length-M column vector:

$$b_q=\frac{1}{\sigma^2}\left(\left(\hat{U}^{(k)}v_q\right)-\frac{I_1(a_q)}{I_0(a_q)}\odot y_q\right), \quad [26]$$

Then a new column vector is obtained by element-wise multiplication of b_q and the lth column of $\hat{\mathrm{U}}^{(k)}$, and hence $\partial L(\mathrm{v}_q)/\partial \mathrm{v}_{q,l}$ is obtained by summing all the elements in this vector. Finally, the gradient of $g(\mathrm{v}_q)$ can be written as

$$\nabla g(v_q)=\nabla L(v_q)+2\lambda\left(\hat{U}^{(k)}\right)^T(\mathrm{WB})^T\mathrm{WB}\hat{U}^{(k)}v_q. \quad [27]$$

Gradients for the non-central χ distribution

The gradient derivation for non-central χ likelihood follows the same procedures except the extra log function term and the order difference in the modified Bessel function of the first kind, $I_{C-1}(\cdot)$ as shown in Eq. [11]. The gradient w.r.t. the additional log function is straightforward to derive and is not shown here. As for the derivative for $I_{C-1}(\cdot)$ with an arbitrary order $C-1$, we use the following identity

$$I'_{C-1}(\cdot)=\frac{1}{2}(I_{C-2}(\cdot)+I_C(\cdot)). \quad [28]$$

Therefore, the terms in Eq. [23] and Eq. [26] can be modified for the non-central χ distribution as

$$\nabla L_q(U)=\left(\frac{C-1}{\left(U\hat{V}^{(k)}\right)_q}+\frac{\left(U\hat{V}^{(k)}\right)_q}{\sigma^2}-\frac{I_{C-2}(a_q)+I_C(a_q)}{2\sigma^2 I_{C-1}(a_q)}\odot y_q\right)\left(\hat{V}_q^{(k)}\right)^T. \quad [29]$$

and as

$$b_q=\frac{C-1}{\hat{U}^{(k)}v_q}+\frac{1}{\sigma^2}\left((\hat{U}^{(k)}v_q)-\frac{I_{C-2}(a_q)+I_C(a_q)}{2I_{C-1}(a_q)}\odot y_q\right). \quad [30]$$

References

1. Pierpaoli C, Basser PJ. Toward a quantitative assessment of diffusion anisotropy. Magn Reson Med. 1996; 36:893–906. [PubMed: 8946355]
2. Jones DK, Basser PJ. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. Magn Reson Med. 2004; 52:979–993. [PubMed: 15508154]
3. Wang Z, Vemuri BC, Chen Y, Mareci TH. A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex DWI. IEEE Trans Med Imag. 2004; 23:930–939.
4. Haldar JP, Wedeen VJ, Nezamzadeh M, Dai G, Weiner MW, Schuff N, Liang ZP. Improved diffusion imaging through SNR-enhancing joint reconstruction. Magn Reson Med. 2012 In press.
5. Gudbjartsson H, Patz S. The Rician distribution of noisy MRI data. Magn Reson Med. 1995; 34:910–914. [PubMed: 8598820]
6. Rice SO. Mathematical analysis of random noise. Bell System Technical Journal. 1944:23–24.
7. Koay CG, Basser PJ. Analytically exact correction scheme for signal extraction from noisy magnitude MR signals. J Magn Reson. 2006; 179:317–322. [PubMed: 16488635]
8. Henkelman RM. Measurement of signal intensities in the presence of noise in MR images. Med Phys. 1985; 12:232–233. [PubMed: 4000083]
9. Bernstein MA, Thomasson DM, Perman WH. Improved detectability in low signal-to-noise ratio magnetic resonance images by means of a phase-corrected real reconstruction. Med Phys. 1989; 16:813–817. [PubMed: 2811764]
10. Constantinides CD, Atalar E, McVeigh ER. Signal-to-noise measurements in magnitude images from NMR phased arrays. Magn Reson Med. 1997; 38:852–857. [PubMed: 9358462]
11. Dietrich O, Raya JG, Reeder SB, Ingrisch M, Reiser MF, Schoenberg SO. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. Magn Reson Imag. 2008; 26:754–762.
12. Aja-Fernandez S, Tristan-Vega A. Influence of noise correlation in multiple-coil statistical models with sum of squares reconstruction. Magn Reson Med. 2012; 67:580–585. [PubMed: 21656560]
13. Andersson JLR. Maximum a posteriori estimation of diffusion tensor parameters using a Rician noise model: Why, how and but. NeuroImage. 2008; 42:1340–1356. [PubMed: 18602480]

14. Descoteaux M, Angelino E, Fitzgibbons S, Deriche R. Apparent diffusion coefficients from high angular resolution diffusion imaging: Estimation and applications. Magn Reson Med. 2006; 56:395–410. [PubMed: 16802316]
15. Nowak RD. Wavelet-based Rician noise removal for magnetic resonance imaging. IEEE Trans Image Process. 1999; 8:1408–1419. [PubMed: 18267412]
16. Aja-Fernandez S, Niethammer M, Kubicki M, Shenton ME, Westin CF. Restoration of DWI data using a Rician LMMSE estimator. IEEE Trans Med Imag. 2008; 27:1389–1403.
17. Koay CG, Ozarslan E, Basser PJ. A signal transformational framework for breaking the noise floor and its applications in MRI. J Magn Reson. 2009; 197:108–119. [PubMed: 19138540]
18. Foi A. Noise estimation and removal in MR imaging: the variance-stabilization approach. Proc IEEE Int Symp on Biomed Imag. 2011:1809–1814.
19. Luisier F, Wolfe PJ. Chi-square unbiased risk estimate for denoising magnitude MR images. Proc IEEE ICIP. 2011:1561–1564.
20. Basu S, Fletcher T, Whitaker R. Rician noise removal in diffusion tensor MRI. Med Image Comput Comput Assist Interv. 2006; 9:117–125. [PubMed: 17354881]
21. McGraw T, Vemuri B, Ozarslan E, Chen Y, Mareci T. Variational denoising of diffusion weighted MRI. Inverse Problems and Imaging. 2009; 3:625–648.
22. Ding Z, Gore JC, Anderson AW. Reduction of noise in diffusion tensor images using anisotropic smoothing. Magn Reson Med. 2005; 53:485–490. [PubMed: 15678537]
23. Van Hecke W, Leemans A, De Backer S, Jeurissen B, Parizel PM, Sijbers J. Comparing isotropic and anisotropic smoothing for voxel-based DTI analyses: A simulation study. Human Brain Mapping. 2010; 31:98–114. [PubMed: 19593775]
24. Walker-Samuel S, Orton M, Boult JKR, Robinson SP. Improving apparent diffusion coefficient estimates and elucidating tumor heterogeneity using Bayesian adaptive smoothing. Magn Reson Med. 2011; 65:438–447. [PubMed: 21264934]
25. Patel V, Shi Y, Thompson PM, Toga AW. K-SVD for HARDI denoising. Proc IEEE Int Symp on Biomed Imag. 2011:1805–1808.
26. Wiest-Daessle N, Prima S, Coupe P, Morrissey SP, Barillot C. Rician noise removal by nonlocal means filtering for low signal-to-noise ratio MRI: Applications to DT-MRI. Proc MICCAI. 2008:171–179.
27. Descoteaux M, Wiest-Daessle N, Prima S, Barillot C, Deriche R. Impact of Rician adapted non-local means filtering on HARDI. Proc MICCAI. 2008:122–130.
28. Manjon JV, Coupe P, Marti-Bonmati L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. J Magn Reson Imag. 2010; 31:192–203.
29. Manjon JV, Thacker NA, Lull JJ, Garcia-Marti G, Marti-Bonmati L, Robles M. Multicomponent MR image denoising. Int J Biomed Imaging. 2009:2009.
30. Liang ZP. Spatiotemporal imaging with partially separable functions. Proc IEEE Int Symp on Biomed Imag. 2007:988–991.
31. Haldar JP, Liang ZP. Spatiotemporal imaging with partially separable functions: A matrix recovery approach. Proc IEEE Int Symp on Biomed Imag. 2010:716–719.
32. Zhao B, Haldar JP, Christodoulou AG, Liang ZP. Image reconstruction from highly under-sampled (k, t)-space data with joint partial separability and sparsity constraints. IEEE Trans Med Imag. 2012; 31:1809–1820.
33. Lingala SG, Hu Y, DiBella E, Jacob M. Accelerated dynamic MRI exploiting sparsity and low-rank structure: k-t SLR. IEEE Trans Med Imag. 2011; 30:1042–1054.
34. Manjon JV, Coupe P, Concha L, Buades A, Collins L, Robles M. DWI denoising using overcomplete local PCA decomposition. Proc Int Soc Magn Reson Med. 2012:3580.
35. Nguyen HM, Peng X, Do MN, Liang ZP. Spatiotemporal denoising of MR spectroscopic imaging data by low-rank approximations. Proc IEEE Int Symp on Biomed Imag. 2011:857–860.
36. Lam F, Babacan SD, Schuff N, Liang ZP. Denoising diffusion-weighted MR images using low rank structure and edge constraints. Proc Int Soc Magn Reson Med. 2012:4308.

37. Haldar JP, Hernando D. Rank-constrained solutions to linear matrix equations using powerfactorization. IEEE Signal Process Lett. 2009; 16:584–587. [PubMed: 22389578]
38. Basser PJ, Jones DK. Diffusion-tensor MRI: Theory, experimental design and data analysis – a technical review. NMR Biomed. 2002; 15:456–467. [PubMed: 12489095]
39. Yablonskiy DA, Sukstanskii AL. Theoretical models of the diffusion weighted MR signal. NMR in Biomedicine. 2010; 23:661–681. [PubMed: 20886562]
40. Santis SD, Gabrielli A, Palombo M, Maraviglia B, Capuani S. Non-Gaussian diffusion imaging: A brief practical review. Magn Reson Imag. 2011; 29:1410–1416.
41. Tournier J, Mori S, Leemans A. Diffusion tensor imaging and beyond. Magn Reson Med. 2011; 65:1532–1556. [PubMed: 21469191]
42. Haldar, JP. PhD thesis, University of Illinois at Urbana-Champaign. 2011. Constrained imaging: denoising and sparse sampling.
43. Nikolava M, NG MK. Analysis of half-quadratic minimization methods for signal and image recovery. SIAM Journal on Scientific Computing. 2005; 27:937–966.
44. Huber, PJ. Robust Statistics. John Wiley & Sons; 1981.
45. Zhao B, Haldar JP, Brinegar C, Liang ZP. Low rank matrix recovery for real-time cardiac MRI. Proc IEEE Int Symp on Biomed Imag. 2010:996–999.
46. Nocedal, J.; Wright, SJ. Numerical Optimization. Springer; 2006.
47. Koay CG, Chang LC, Carew JD, Pierpaoli C, Basser PJ. A unifying theoretical and algorithmic framework for least squares methods of estimation in diffusion tensor imaging. J Magn Reson. 2006; 182:115–125. [PubMed: 16828568]
48. Haldar JP, Sakaie K, Liang ZP. Resolution and noise properties of linear phase-constrained partial Fourier reconstruction. Proc Int Soc Magn Reson Med. 2009:2862.
49. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, Kiefer B, Haase A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magn Reson Med. 2002; 47:1202–1210. [PubMed: 12111967]
50. Wedeen VJ, Hagmann P, Tseng WYI, Reese TG, Weisskoff RM. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magn Reson Med. 2005; 54:1377–1386. [PubMed: 16247738]
51. Tuch DS. Q-ball imaging. Magn Reson Med. 2004; 52:1358–1372. [PubMed: 15562495]
52. Arsigny V, Fillard P, Pennec X, Ayache N. Log-Euclidean metrics for fast and simple calculus on diffusion tensors. Magn Reson Med. 2006; 56:411–421. [PubMed: 16788917]
53. Chang LC, Rohde GK, Pierpaoli C. An automatic method for estimating noise-induced signal variance in magnitude-reconstructed magnetic resonance images. Proc of SPIE. 2005:1136–1142.
54. Koay CG, Ozarslan E, Pierpaoli C. Probabilistic identification and estimation of noise (PIESNO): A self-consistent approach and its applications in MRI. J Magn Reson. 2009; 199:94–103. [PubMed: 19346143]
55. Sijbers J, Poot DHJ, Dekker AD, Pintjens W. Automatic estimation of the noise variance from the histogram of a magnetic resonance image. Phys Med and Biol. 2007; 52:1335–1348. [PubMed: 17301458]
56. Coupe P, Manjon JV, Gedamu E, Arnold D, Robles M, Collins DL. Robust Rician noise estimation for MR images. Med Image Anal. 2010; 14:483–493. [PubMed: 20417148]
57. Rajan J, Poot D, Juntu J, Sijbers J. Noise measurement from magnitude MRI using local estimates of variance and skewness. Phys Med and Biol. 2010; 55:N441–N445. [PubMed: 20679694]
58. Landman BA, Bazin PL, Smith SA, Prince JL. Robust estimation of spatially variable noise fields. Magn Reson Med. 2009; 62:500–509. [PubMed: 19526510]
59. Maximov II, Farrher E, Grinberg F, Shah NJ. Spatially variable Rician noise in magnetic resonance imaging. Med Image Anal. 2012; 16:536–548. [PubMed: 22209560]
60. Craven P, Wahba G. Smoothing noisy data with spline functions. Numerische Mathematik. 1978; 31:377–403.
61. Vogel, CR. Computational Methods for Inverse Problems. Philadelphia, PA: SIAM; 2002.

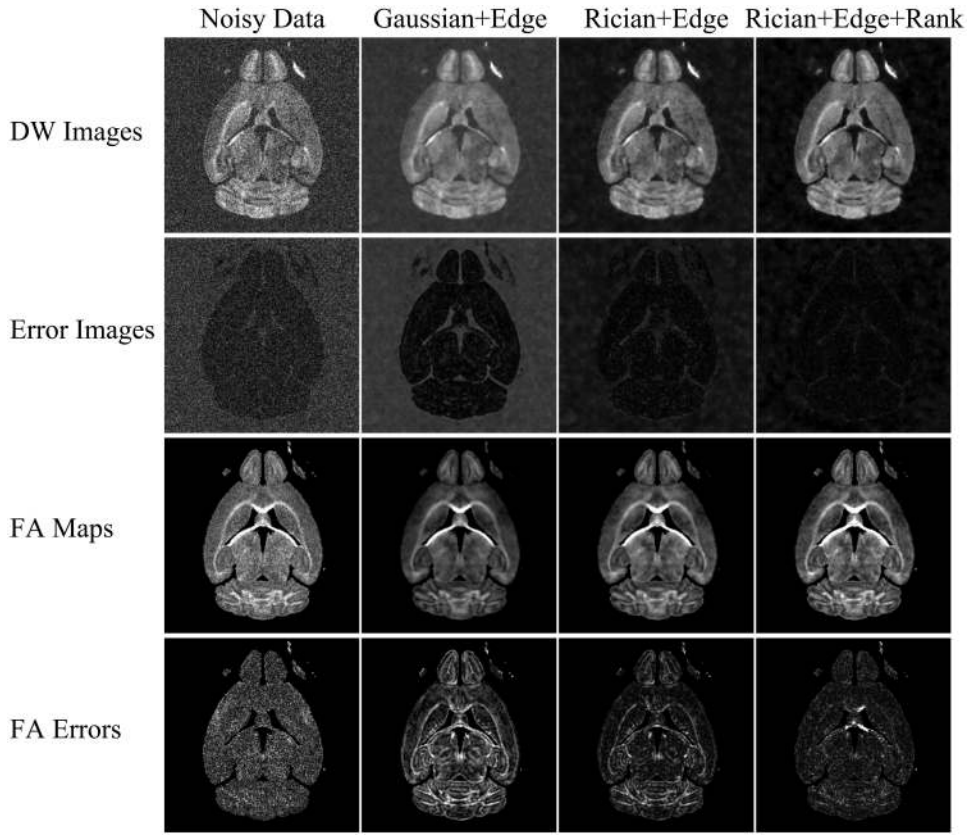

Figure 1.

Comparison of DW images (row 1), error images (row 2), FA maps (row 3) and FA error maps (row 4) for four different processing schemes. They correspond to (1) no denoising (column 1, Noisy Data); (2) denoising with Gaussian noise model and edge constraints (column 2, Gaussian+Edge); (3) denoising with Rician signal model and edge constraints (column 3, Rician+Edge); (4) denoising with Rician signal model and joint low-rank and edge constraints (column 4, Rician+Edge+Rank).

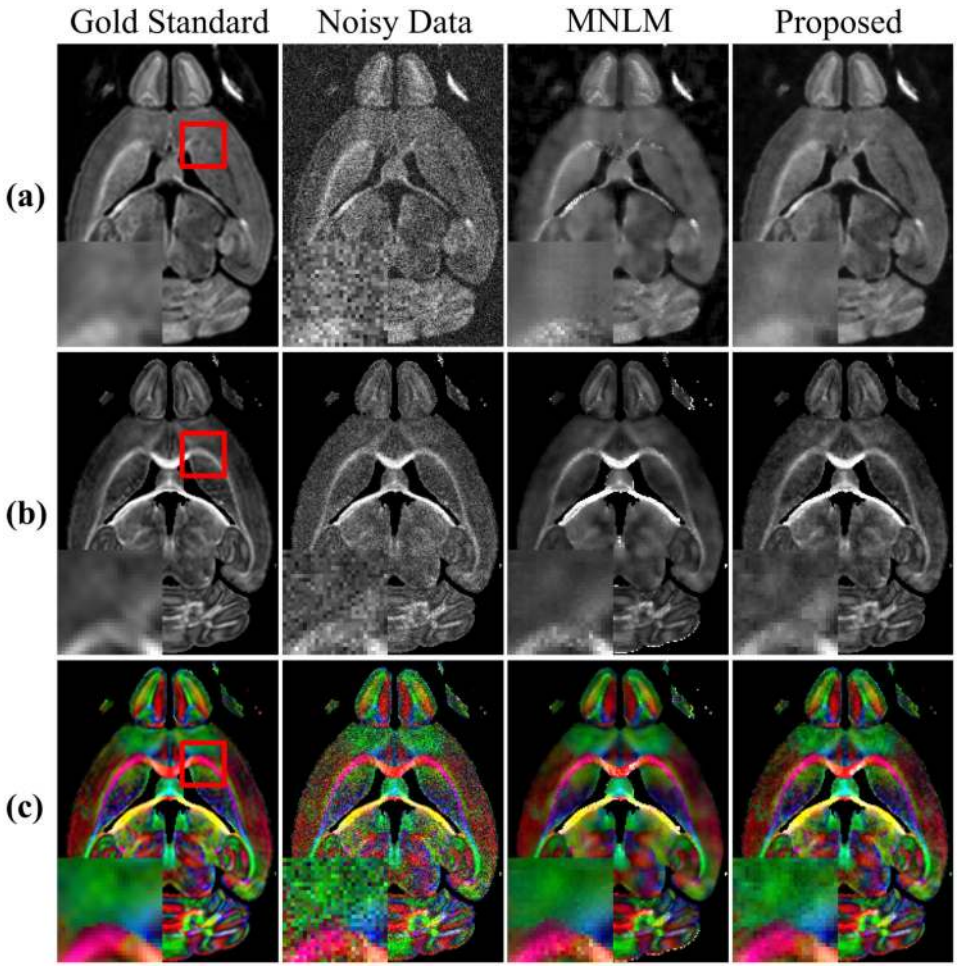

Figure 2.

Denoising results for simulated Rician distributed data (b = 2000). The three rows correspond to (Top) representative DW images, (Center) FA maps and (Bottom) color-coded FA maps from the noise-free images (column 1, Gold Standard), Rician noise corrupted images (column 2, Noisy Data), images denoised by the MNLM algorithm (column 3, MNLM) and images denoised by the proposed method (column 4). The color is based on the orientation of the primary eigenvectors of the estimated diffusion tensors: red for left-right, green for anterior-posterior and blue for superior-inferior.

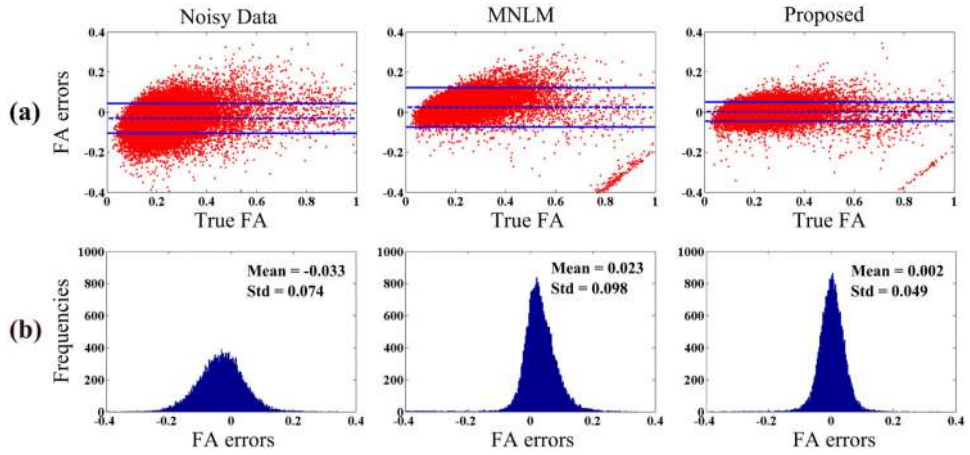

Figure 3.

Quantitative comparison of FA estimation for simulated Rician distributed data: (Top) plots of FA errors versus the true FA values and (Bottom) histograms of FA errors. The bias (dash line) and the standard deviation (solid line) are also shown. Three cases are compared, including FA errors for noisy data (column 1), FA errors for denoised data from the MNLM algorithm (column 2) and FA errors for denoised data from the proposed method (column 3).

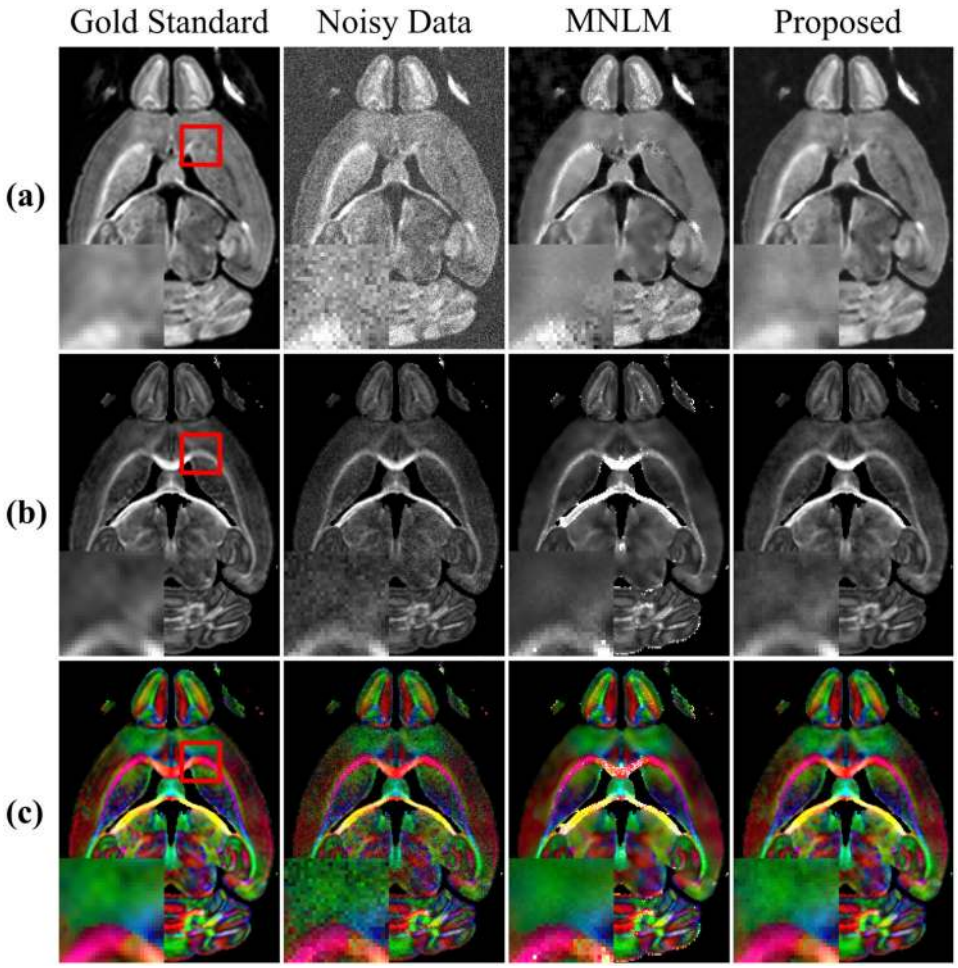

Figure 4.

Denoising results for simulated noncentral χ distributed data (sum-of-squares combined images). The three rows correspond to (Top) representative DW images, (Center) FA maps and (Bottom) color-coded FA maps from the noise-free images (column 1), noncentral χ distributed noisy images (column 2), images denoised by the MNLM algorithm (column 3) and images denoised by the proposed method (column 4). The direction information is encoded in the same way as in Fig. 2.

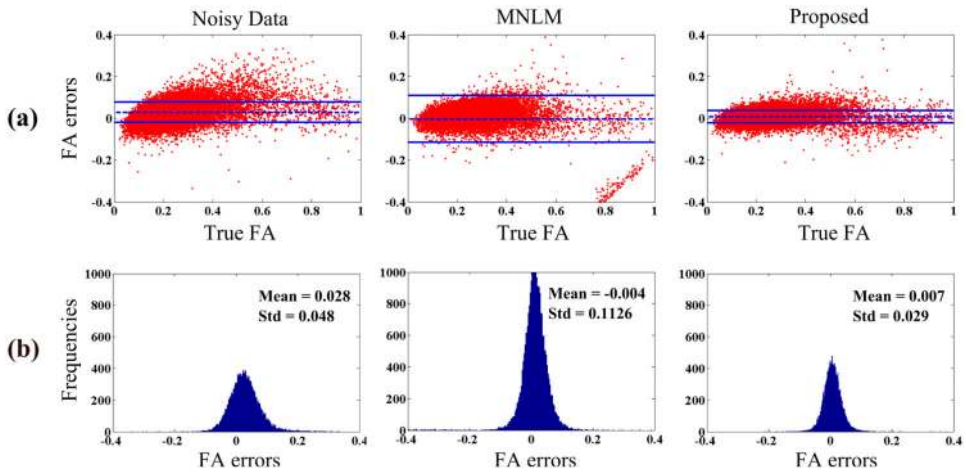

Figure 5.

Quantitative comparison of FA estimation for simulated noncentral χ distributed data: (Top) plots of FA errors versus the true FA values and (Bottom) histograms of FA errors. Three cases are compared, including FA errors for noisy data (column 1), FA errors for denoised data from the MNLM algorithm (column 2) and FA errors for denoised data from the proposed method (column 3).

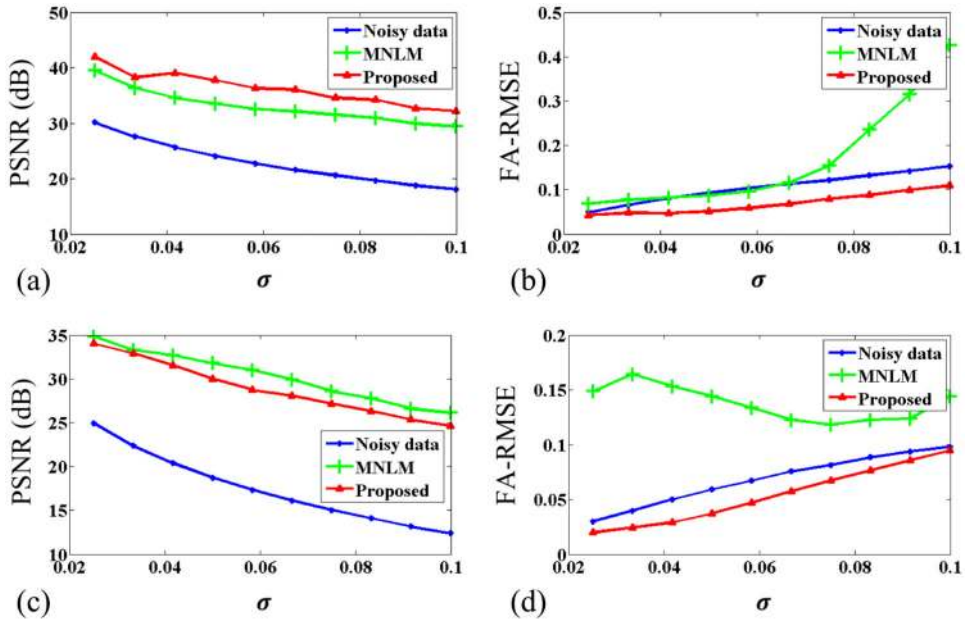

Figure 6.

Quantitative comparison of different methods under different noise levels for Rician distributed data ((a)-(b)) and noncentral χ distributed data ((c)-(d)): (a and c) PSNRs for one representative image frame; and (b and d) RMSEs for FA estimation. The x-axis represents different noise standard deviations (σ).

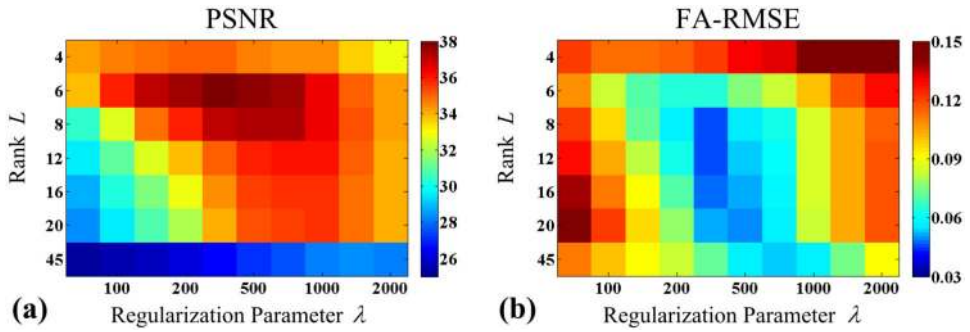

Figure 7.

Comparison of denoising performance w.r.t. different combinations of the low-rank model order L and the regularization parameter λ. The image on the left (a) shows the PSNR and the image on the right (b) shows the FA-RMSE in different cases. The x-axis shows the range for λ and the y-axis shows the range for L.

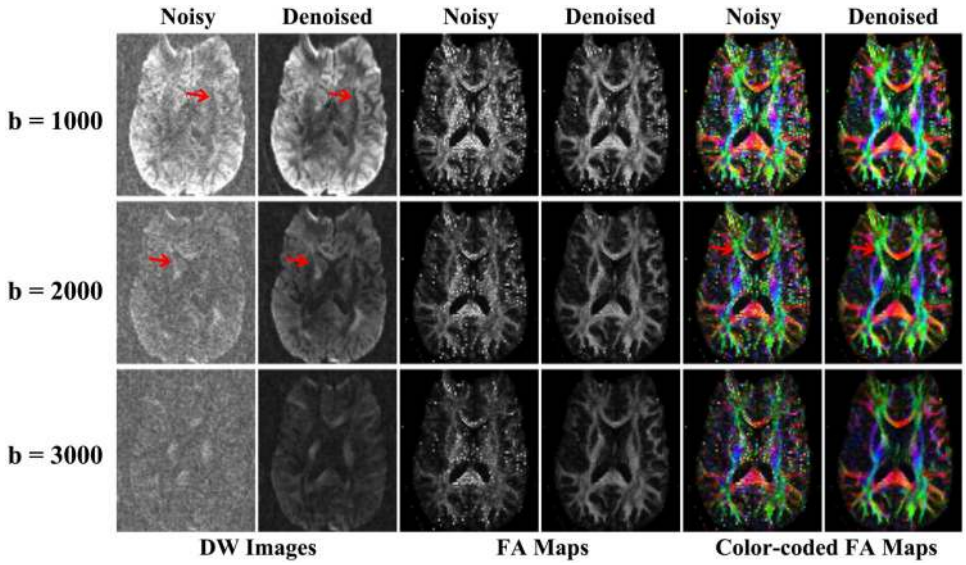

Figure 8.

Results from coil-by-coil denoising the experimental data (with Rician likelihood). The three rows correspond to images for different *b*-values. Representative image frames (column 1 and 2), calculated FA maps (column 3 and 4) and color-coded FA maps (column 5 and 6) are shown. The colors stand for: red for left-right, green for anterior-posterior and blue for superior-inferior.

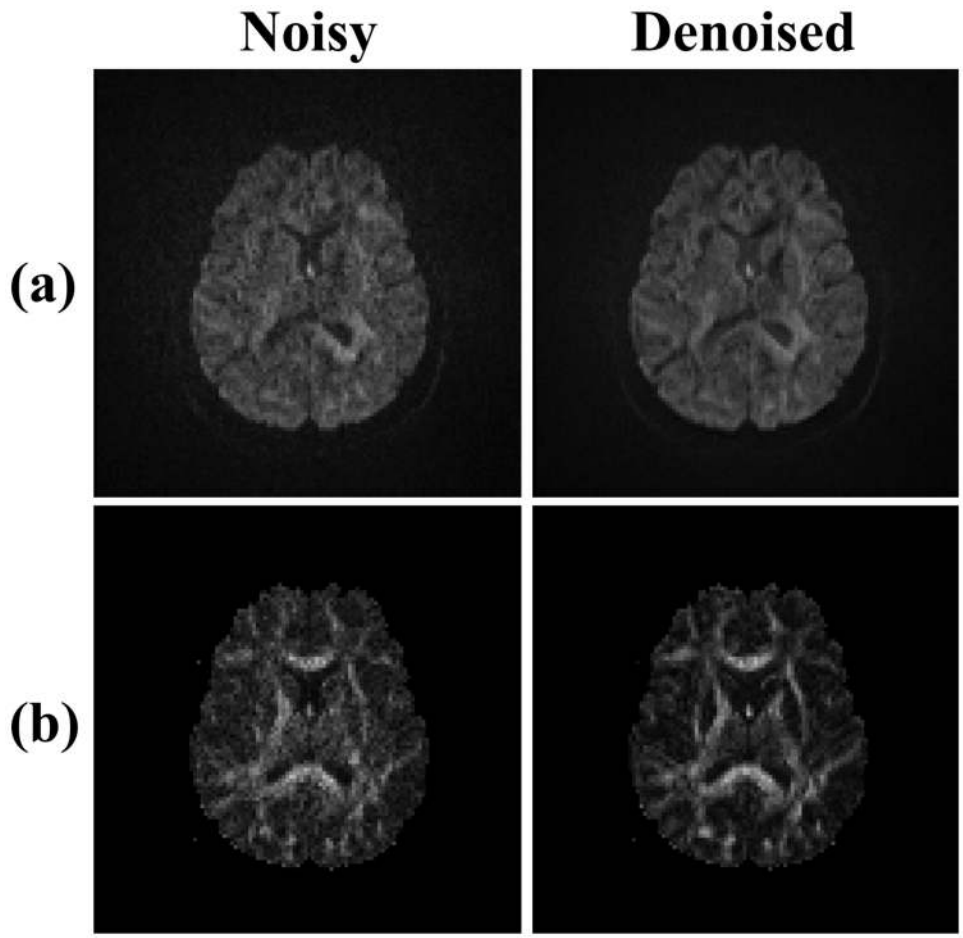

Figure 9.

Results from denoising the experimental sum-of-squares images (with noncentral χ likelihood). The images for one slice are shown and arranged as: (Top) DW images from the noisy and denoised data; and (Bottom) Estimated FA maps from the noisy and denoised data.

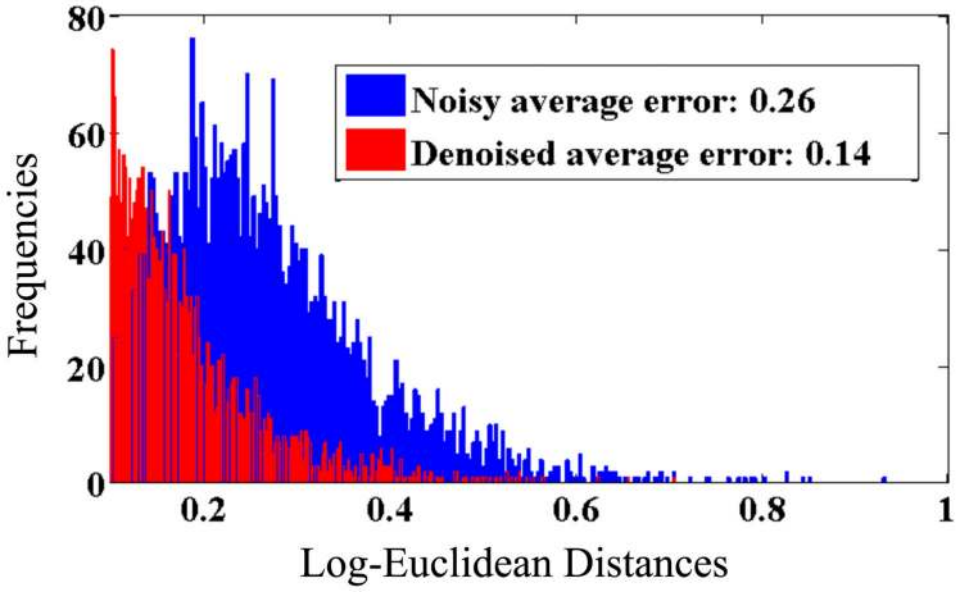

Figure 10.

Comparison of histograms for Log-Euclidean distances between tensors estimated from diffusion signals in two hemispheres of the diffusion space (*q*-space), before and after denoising. The average Log-Euclidean errors for both cases are also shown.