Adaptive Multiresolution Non-Local Means Filter for 3D MR Image Denoising

Pierrick Coupé, José V. Manjón, Montserrat Robles, Louis D. Collins

Adaptive Multiresolution Non-Local Means Filter for 3D MR Image Denoising

Pierrick Coupé[1], José V. Manjón[2],

Montserrat Robles[2] and D. Louis Collins[1]

[1] McConnell Brain Imaging Centre, Montréal Neurological Institute, McGill University 3801, University Street, Montréal, Canada H3A 2B4 bDivision of Neurosurgery, McGill University, Montréal, Canada

[2] Biomedical Informatics Group (IBIME), ITACA Institute, Polytechnic University of Valencia, Camino de Vera, s/n. 46022 Valencia, Spain

Abstract

In this paper, an adaptive multiresolution version of the Blockwise Non-Local (NL-) means filter is presented for 3D Magnetic Resonance (MR) images. Based on an adaptive soft wavelet coefficient mixing, the proposed filter implicitly adapts the amount of denoising according to the spatial and frequency information contained in the image. Two versions of the filter are described for Gaussian and Rician noise. Quantitative validation was carried out on Brainweb datasets by using several quality metrics. The results show that the proposed multiresolution filter obtained competitive performance compared to recently proposed Rician NL-means filters. Finally, qualitative experiments on anatomical and Diffusion-Weighted MR images show that the proposed filter efficiently removes noise while preserving fine structures in classical and very noisy cases. The impact of the proposed denoising method on fiber tracking is also presented on a HARDI dataset.

I. Introduction

In magnetic resonance (MR) imaging, denoising is an important preprocessing step used to improve image analysis. Since the noise directly impacts the accuracy of automatic approaches such as segmentation and registration, it is important to improve the Signal-Noise-Ratio (SNR) of images used during quantitative analysis. This is usually

done by using a denoising method. However, the challenge is the remove of the noise component while preservation all the image features. In the past, many approaches have been proposed to address the difficult problem of MR image denoising [1]–[9].

Recently, in 2D natural image denoising, a new filter has been introduced by Buades *et al.* [10]: the Non-Local (NL-) means filter. Based on the natural redundancy of the images, this filter uses comparison of image patches in order to achieve the denoising. Due to its simplicity and high denoising performance, this filter is one of the best state-of-the-art methods. Recently, the NL-means or similar patch-based approaches are being adapted to MR denoising [6], [9], [11], [12].

This paper focuses on a new adaptive multiresolution version of the NL-means filter. As shown in [13], there exists no set of parameters optimal for all the regions of a single image. According to the nature of the underlying structures (edge, texture, constant area...), the filtering parameters should be adapted. The main contribution of this paper is to introduce a new Adaptive Soft wavelet Coefficient Mixing (ASCM) procedure to implicitly optimize the denoising over all the space-frequency resolutions of the image. The main idea is to combine, in the wavelet transform domain, denoised images obtained with different sets of filtering parameters using the noisy input image.

II. Methods

A. The Non Local Means filter

In the classical formulation of the NL-means filter [10], the restored intensity $\hat{u}(x_i)$ of the voxel x_i, is a weighted average of the voxels intensities $u(x_j)$ in the "search volume" V_i of radius M:

$$\hat{u}(x_i) = \sum_{x_j \in V_i} w(x_i, x_j) u(x_j) \tag{1}$$

where $w(x_i, x_j)$ is the weight assigned to intensity value $u(x_j)$ to restore the intensity $u(x_i)$ at voxel x_i. More precisely, the weight evaluates the similarity between the intensity of the patches N_i and N_j centered on voxels x_i and x_j, such that $w(x_i, x_j) \in [0, 1]$ and $\sum_{x_j \in V_i} w(x_i, x_j) = 1$.

For each voxel x_j in V_i, the computation of the weight is based on the squared Euclidean distance between intensity patches $\mathbf{u}(N_j)$ and $\mathbf{u}(N_i)$, defined as:

$$w(x_i, x_j) = \frac{1}{Z_i} e^{-\frac{\|\mathbf{u}(N_i) - \mathbf{u}(N_j)\|_2^2}{h^2}} \tag{2}$$

where Z_i is a normalization constant ensuring that $\sum_j w(x_i, x_j) = 1$, and h acts as a smoothing parameter. The main limitation of the standard NL-means filter is its computational burden, especially in 3D. Therefore, we used the blockwise approach and the patch preselection strategy proposed in [6] to drastically reduce the computational time. The blockwise approach consists in denoising the entire patch at the same time. By this way, we obtained restored values for all the elements of the patch:

$$\hat{u}(N_i) = \sum_{N_j \in V_i} w(x_i, x_j) u(N_j), \tag{3}$$

Since the patches overlap, several estimations are obtained for each voxel. The final restored intensity is thus obtained by averaging these different estimations. More details can be found in [6].

B. Adaptation to Rician noise

The raw data of an MR image are complex values corrupted by white additive Gaussian noise with the same variance in real and imaginary parts [14]. In magnitude MR images, the distribution of the noise is transformed into a Rician distribution [14], [15]:

$$p(m) = \frac{m}{\sigma^2} \exp(-\frac{m^2 + A^2}{2\sigma^2}) I_0(\frac{Am}{\sigma^2}). \tag{4}$$

where σ is the standard deviation of Gaussian noise in the complex domain, A is the amplitude of the signal without noise, m is the value in the magnitude image and I_0 is the zeroth order modified Bessel function. The second-order moment of a Rician distribution can be written as:

$$\mathbb{E}(m^2) = A^2 + 2\sigma^2 \tag{5}$$

Recently, the nonlocal means filter has been adapted to Rician noise using Conventional Approach (CA) [11], [16] and Maximum Likelihood Estimators (MLE) [12].

In [12], authors proposed a Non-local Maximum Likelihood (NLML) method considering that the N most similar pixels to the current pixel (in the patch distance sense) comes from the same Rician distribution. Knowing the level of noise, an optimization procedure is used to maximize the Rician-based $\log$ likelihood function:

$$logL = \sum_{i=1}^{N} \log\left(\frac{m_i}{\sigma^2}\right) - \sum_{i=1}^{N} \frac{m_i^2 + A^2}{2\sigma^2} + \sum_{i=1}^{N} \log I_0\left(\frac{Am_i}{\sigma^2}\right) \tag{6}$$

and the MLE is obtained as:

$$\hat{A}_{MLE} = \arg\max_{A}(\log L) \tag{7}$$

The main limitation of the extention to 3D of these methods is the required computational load. Moreover, based on a minimization procedure, these approaches can converge to local minimum and require extra parameters.

In [11], [16], authors have proposed Rician-adapted versions of the original NL-means filter using the properties of the second-order moment of Rician distribution called Conventional Approach (CA). The even order moments of a Rician law are usually used due to their simplicity. Indeed, by working on the squared signal, the bias intensity can be removed by subtraction as follows:

$$\hat{A}_{CA} = \sqrt{\max\left(\widehat{\mathbb{E}(m^2)} - 2\sigma^2, 0\right)} \tag{8}$$

In denoising context, $\widehat{\mathbb{E}(m^2)}$ is usually estimated as the restored value of the squared noisy signal. The CA does not require extra computational time nor minimization procedure. Nevertheless, this method can artificially clamp background values to zero for very high levels of noise. Finally, the CA is more biased than the MLE when a small number of samples is used [15].

As we will show in our experiments, the NL-means framework is better suited to CA approach since it involves a large number of samples during estimation. Therefore, we chosen to use the CA proposed in [16] with the blockwise approach defined in [6] as follows:

$$\hat{u}(N_i) = \sqrt{max\left(\left(\sum_{N_j \in V_i} w(x_i, x_j) u(N_j)^2\right) - 2\sigma^2, 0\right)} \tag{9}$$

C. Multiresolution Framework

1) Hard Wavelet Subbands Mixing: As for all the denoising filters, the choice of the filtering parameters is crucial in NL-means-based restoration. Lukin [17] proposed a multiresolution framework to implicitly adapt the filtering parameters over the different space-frequency resolutions by combining in the transform domain the results obtained with different sets of filtering parameters. We have adapted this framework to NL-means-based restoration of 3D MR images in [7]. This Hard Subband Mixing (HSM) can be described as follows:

1) Denoise the original noisy MR image I_n using two sets of filtering parameters: S_u adapted to the image features preservation (i.e. small patches and/or low h parameter) and S_o, adapted to the noise components removal (i.e. large patches and/or high h parameter). This yields two denoised images: $\hat{I}_u$ with efficient feature preservation (slightly undersmoothed) and $\hat{I}_o$ with efficient noise removal (slightly oversmoothed).
2) Decompose $\hat{I}_u$ and $\hat{I}_o$ into subbands with 3D discrete Wavelet Transform (DWT) at first level decomposition. By using the usual notation: LLL denotes the low subband whereas LHH, LHL, LLH, HLL, HLH, HHL and HHH denote the high subbands
3) Select the LLL, LLH, LHL and HLL subbands of $\hat{I}_u$ and the LHH, HHL, HLH and HHH subbands of $\hat{I}_o$.
4) Reconstruct the final image $\hat{I}$ by an inverse 3D DWT from the combination of the selected subbands.

This procedure can be written as:

$$\begin{aligned} \hat{c}_k &= \hat{c}_{k,u} \\ \hat{d}_k &= \hat{d}_{k,u} \text{ for LLH, LHL and HLL} \\ \hat{d}_k &= \hat{d}_{k,o} \text{ for HHL, HLH, LHH and HHH} \end{aligned} \tag{10}$$

where $\hat{c}_k$, $\hat{c}_{k,u}$ are the scaling coefficients (i.e. coefficients in LLL low band) of $\hat{I}$ and $\hat{I}_u$ decompositions, and $\hat{d}_k$, $\hat{d}_{k,u}$ and $\hat{d}_{k,o}$ are the wavelet coefficients in detailed subbands of $\hat{I}$, $\hat{I}_u$ and $\hat{I}_o$ decompositions (more details can be found in [7]). This HSM procedure improves the denoising performances of the NL-means filter by selecting the optimal denoised subbands along the image frequencies. Nevertheless, the HSM is not adapted according to noise level, nor to spatial information contained in subbands (i.e. the underlying structure). Moreover, with HSM, the subband selection is empirically optimized instead of being driven by the input signal.

2) Adaptive Soft Wavelet Coefficient Mixing: In this paper, we propose to automatically adapt the mixing procedure according to the noise level and the spatial information contained in the detailed subbands. To achieve that, we use a similar strategy than in wavelet denoising thresholding method: detect wavelet coefficients corresponding to noise from coefficients corresponding to data by comparison with a threshold. Usually, the coefficients below this threshold are considered as noise and the others correspond to data information [18]. In our case, the distance between coefficients of noisy input image and the threshold is used to drive the mixing of the coefficients of the denoised images.

First, we decompose I_n, $\hat{I}_u$ and $\hat{I}_o$ into subbands with a 3D DWT. For each subband of I_n, the signed distances between absolute values of noisy wavelet coefficients $d_{k,n}$ obtained from the original noisy image I_n and a threshold are computed.

Then, this signed distance is used to perform a soft mixing between denoised wavelet coefficients, $\hat{d}_{k,u}$ and $\hat{d}_{k,o}$, obtained from the two denoised images (see Fig. 1). When signed distance is superior to zero, the corresponding noisy wavelet coefficient is mainly considered as data, thus denoised wavelet coefficient from $\hat{I}_u$ should be privileged. Inversely, when signed distance is inferior to zero, the corresponding noisy wavelet coefficient is mainly considered as noise, thus denoised wavelet coefficient from $\hat{I}_o$ should be privileged. In our approach, the weight assigned to each denoised wavelet coefficient is estimated using a weighting function as described below.

3) Adaptive Threshold: In the wavelet literature several thresholds have been proposed to perform denoising [18], [19]. In our approach the *BayesSkrink* threshold [19] has been chosen. This threshold $\hat{T}_b$ can be computed for a wavelet coefficients subband b as:

$$\hat{T}_b(\sigma) = \frac{\sigma^2}{\hat{\sigma_X}} \tag{11}$$

where

$$\hat{\sigma}_X = \sqrt{max(\hat{\sigma}_b^2 - \sigma^2, 0)} \tag{12}$$

where $\hat{\sigma}_b^2$ is the variance of the subband b.

4) Soft Coefficient Mixing: It has been shown that the soft thresholding produces less artifacts and better visual results than hard thresholding [19]. Thus, we propose to use soft mixing instead of hard mixing such as the final coefficient is a weighted average of the coefficients from the two denoised images (see Fig. 1).

$$\begin{aligned} \hat{c}_k &= \hat{c}_{k,u} \\ \hat{d}_k &= \phi(d_k, \hat{T}_b(\sigma))\hat{d}_{k,u} + (1 - \phi(d_k, \hat{T}_b(\sigma)))\hat{d}_{k,o} \end{aligned} \tag{13}$$

where

$$\phi(d_k, \hat{T}_b(\sigma)) = \frac{1}{1 + e^{(-\lambda(|d_k| - \hat{T}_b(\sigma)))}} \tag{14}$$

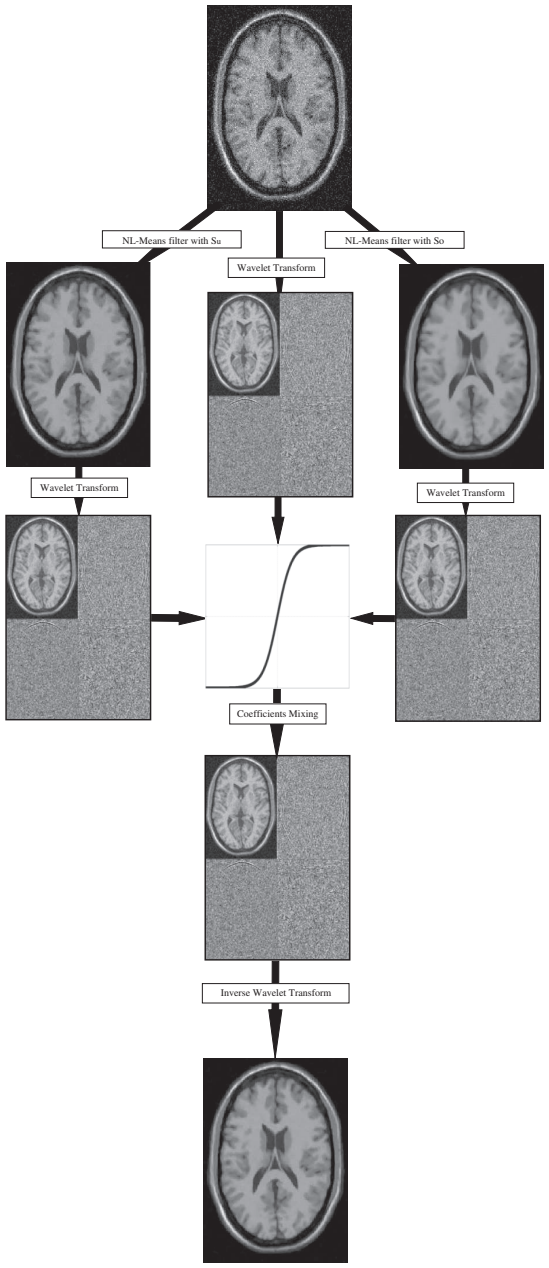

Fig. 1. **Adaptive Soft Wavelet Coefficient Mixing Procedure**. First, the wavelet transform of the original image is used to compute the threshold $\hat{T}_b$ for each subband b. Each noisy wavelet coefficient from original image is then compared to $\hat{T}_b$ in order to drive the soft mixing between the denoised wavelet coefficients. Finally, an invert wavelet transform is used to obtain final image.

The weight $\phi(d_k, \hat{T}_b(\sigma))$ assigned to each denoised wavelet coefficient, $\hat{d}_{k,u}$ and $\hat{d}_{k,o}$, depends on the noisy wavelet coefficient d_k (data driven mixing) and on the threshold $\hat{T}_b(\sigma)$, which is adapted to the frequency (subband variance) and to the noise level σ. The parameter λ controls the sharpness of the mixing and has been empirically set to 0.01 for all the experiments.

III. MATERIALS

A. Phantom Database

In order to evaluate the performance of the different variants of the NL-means filter on 3D MR images, tests were performed on the BrainWeb database [20]. T1-w phantom MR images with different levels of noise were simulated to validate the performance of the denoising. For each image, ten levels of noise from 3% up to 21% have been simulated for Gaussian and Rician distributions.

The Rician noise was built from white Gaussian noise in the complex domain:

- $I_r(x_j) = I(x_j) + \eta_1(x_j)$, $\eta_1(x_j) \backsim \mathcal{N}(0, \sigma)$
- $I_i(x_j) = \eta_2(x_j)$, $\eta_2(x_j) \backsim \mathcal{N}(0, \sigma)$

where I is the noise free "ground truth" from Brainweb, I_r is the real component and I_i is the imaginary, and σ is the standard deviation of the added white Gaussian noise. Then, the noisy image is computed as:

$$I_n(x_j) = \sqrt{I_r(x_j)^2 + I_i(x_j)^2} \tag{15}$$

B. Filters compared

During the experiment, different implementations of the NL-means filter were compared:

- **RNLM**: voxelwise implementation of the Rician adapted NL-means filter proposed in [16].
- **ONLM/ORNLM**: Optimized blockwise implementation of the NL-means filter proposed in [6][1] without (ONLM) and with (ORNLM) the Rician adaptation described in [16] .
- **ONLM/ORNLM with HSM**: ONLM or ORNLM filters with hard subbands mixing as proposed in [7].
- **ONLM/ORNLM with ASCM**: the proposed ONLM or ORNLM filters with adaptive soft coefficients mixing.
- **NLML**: 3D extention of the NonLocal-Maximum Likelihood approach proposed in [12].

The selected filtering parameters for the different implementations were:

- For the **RNLM** filter: $h^2 = \sigma^2$, $\alpha = 1$, and $M = 5$ as proposed in [16].
- For the **ONLM/ORNLM** filters: $h^2 = \sigma^2$, $\alpha = 1$, $n = 2$, $M = 5$, $\mu_1 = 0.95$ and $\sigma_1^2 = 0.5$.

[1]This filter can be freely tested at: http://www.irisa.fr/visages/benchmarks

- For the **ONLM/ORNLM with HSM/ASCM** filters as proposed in [7]: $n = 2$, $S_u = (\alpha_u, M_w, h^2) = (1, 3, \sigma^2)$ for $\hat{I}_u$ and $S_o = (\alpha_o, M_w, h^2) = (2, 3, \sigma^2)$ for $\hat{I}_o$. For both we set $\mu_1 = 0.95$ and $\sigma_1^2 = 0.5$.
- For the 3D extention of **NLML** filter: $\alpha = 1$, $N = 25$ and $M = 4$. As in [12], the number of pixels/voxels N used to maximize the likelihood function was set to 25.

As usually done in nonlocal means litterature, we use a semi-local paradigm instead of fully nonlocal paradigm by restricting the size of search volume. This enables to obtain a practical computational time but also to obtain near-optimal or optimal results excepted for image with repetitive textures [13].

C. Quality metrics

The evaluation of denoising performance depends on the balance between noise removal and structure preservation. Each quality metric tends to privilege the first or second aspect, making it difficult to compare denoising filters. Nevertheless, an optimal denoising should improve the both at the same time. We used six different metrics to evaluate the noise removal and the structure preservation performance of the compared filters.

1) The **Signal-to-Noise Ratio** (SNR):

$$SNR = 10log_{10}\left(\frac{\sum_{x_i\in\Omega} I(x_i)^2}{\sum_{x_i\in\Omega}(I(x_i) - \hat{I}(x_i))^2}\right) \tag{16}$$

where Ω is the image grid. This metric is based on the noisefree "ground truth" image I and the denoised image $\hat{I}$. Higher SNR values correspond to good denoising results.

2) The **MSE on gradient**:

$$\text{MSE on gradient} = \frac{1}{|\Omega|}\sum_{x_i\in\Omega}(\Delta I(x_i) - \Delta\hat{I}(x_i))^2 \tag{17}$$

where the gradient modulus image is computed as: $\Delta I = \sqrt{\Delta_x I^2 + \Delta_y I^2 + \Delta_z I^2}$. This metric is based on the mean L_2-norm between the first derivate of the images.

3) The **Correlation Coefficient** [21]:

$$CoC = \frac{\sum_{x_i\in\Omega}(I(x_i) - \bar{I})(\hat{I}(x_i) - \bar{\hat{I}})}{\sqrt{\sum_{x_i\in\Omega}(I(x_i) - \bar{I})^2 \sum_{x_i\in\Omega}(\hat{I}(x_i) - \bar{\hat{I}})^2}} \tag{18}$$

where $\bar{I}$ and $\bar{\hat{I}}$ are the means of the noisefree and the denoised image.

4) The **Edge Preservation Index** [21]:

$$EPI = \frac{\sum_{x_i \in \Omega}(\nabla I(x_i) - \overline{\nabla I})(\nabla \hat{I}(x_i) - \overline{\nabla \hat{I}})}{\sqrt{\sum_{x_i \in \Omega}(\nabla I(x_i) - \overline{\nabla I})^2 \sum_{x_i \in \Omega}(\hat{\nabla} I(x_i) - \overline{\nabla \hat{I}})^2}} \tag{19}$$

where the Laplacian operator ∇ is obtained with the standard $3 \times 3 \times 3$ voxels approximation. The EPI is computed as the CoC between the second derivatives of the images.

5) The **Universal Quality Index** [22]:

$$UQI = \frac{\sigma_{I\hat{I}}}{\sigma_I \sigma_{\hat{I}}} . \frac{2\bar{I}\bar{\hat{I}}}{(\bar{I})^2 + (\bar{\hat{I}})^2} . \frac{2\sigma_I \sigma_{\hat{I}}}{\sigma_I^2 + \sigma_{\hat{I}}^2} \tag{20}$$

where $\sigma_{I\hat{I}}$ corresponds to the covariance between the noisefree and the denoised image.

6) The **Mean Structural SIMilarity** (MSSIM) [22]:

$$SSIM = \frac{(2\bar{I}\bar{\hat{I}} + C_1)(2\sigma_{I\hat{I}} + C_2)}{((\bar{I})^2 + (\bar{\hat{I}})^2 + C_1)(\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2)} \tag{21}$$

The constants C_1 and C_2 are used to stabilize the UQI when the means or the variances of the compared images are close to zero. During our experiment, the SSIM is locally computed within a $5 \times 5 \times 5$ voxels Gaussian kernel. Finally, the MSSIM is estimated as the mean value of the local SSIM values.

IV. Results

A. Comparison of CA and MLE

As in [15], we propose to compare CA and MLE on a constant signal = 100 corrupted with several Rician noise levels. As in [12], a MATLAB Nelder-Mead simplex method has been for optimization procedure during MLE estimation. Each experiment was repeated 10,000 times and the mean estimation is used as the result. In [15] the size of the sample was $N = 16$ to be close to multiple acquisition number or neighborhood size used in classical denoising filters. Since NL-Means-based approaches use much larger numbers of samples (e.g. $11^3 = 1331$ with a search area similar to [6]), we have studied the impact of the size of the sample on CA and MLE with $N = 10$, $N = 100$ and $N = 1000$. The results presented in Fig. 2 show that for small sample size (i.e. $N = 10$), the MLE

obtained slightly better results than CA (around 1% at 100% noise). These results are similar to those provided by Sijbers in [15]. However, when N increases, the difference between CA and MLE decreases. At $N = 1000$, CA and MLE provide similar estimation accuracy, while the computational time required by MLE is 1500 times higher than for CA. This computational time explained why the NL-means methods based on MLE use a limited number of samples (in [12], $N = 25$ for NLML) and has been proposed only for 2D images. The presented experiment clearly demonstrates that it is better to use a CA with a high number of samples ($N > 1000$ in case of 3D NLM) than a MLE on a limited number of samples ($N = 25$ in case of NLML).

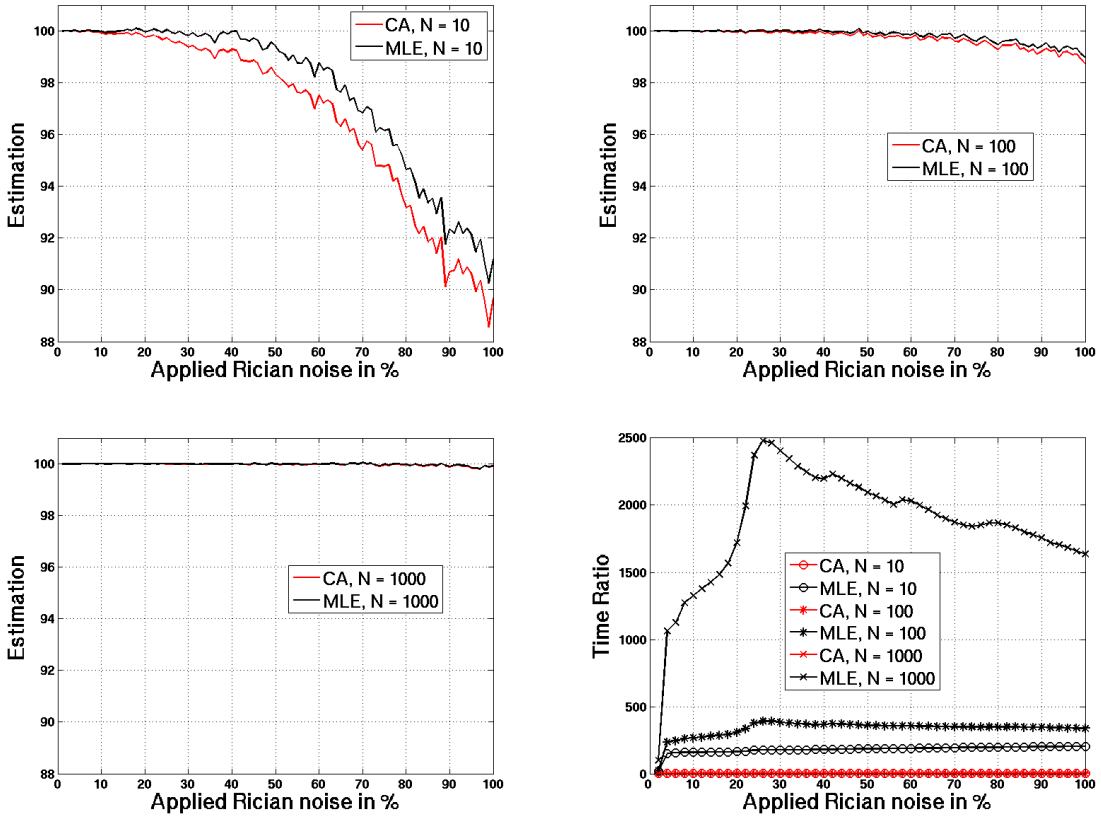

Fig. 2. Comparison between the conventional and ML estimators.

B. Impact of the ASCM on Gaussian Noise

We first studied the impact of the ASCM on Gaussian noise on a simple case (i.e. without the Rician bias intensity problem). Fig. 3 shows the results obtained for the different quality metrics. For all the metrics (except EPI) and for all the noise levels, the ASCM-based filter showed the best performances.

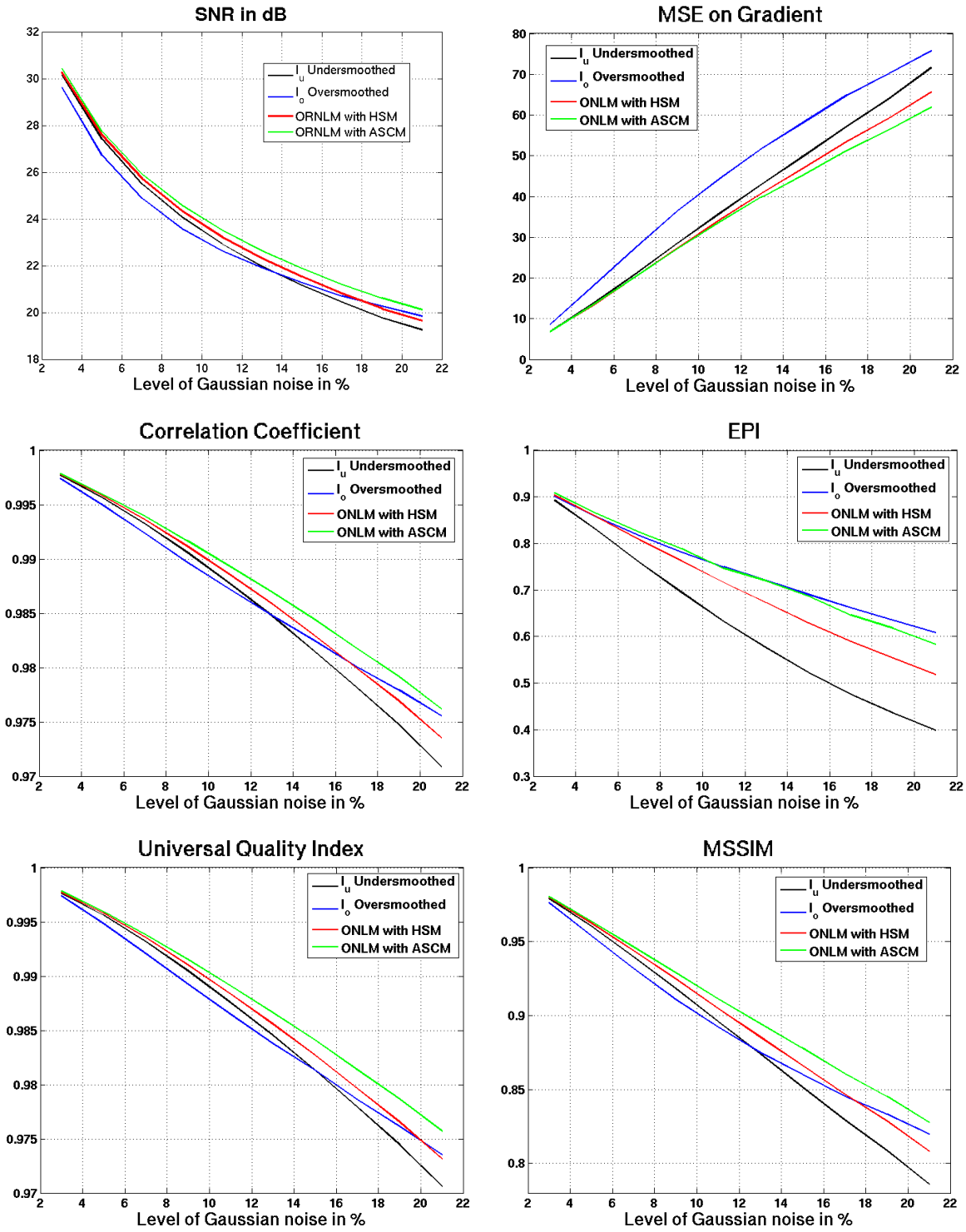

Fig. 3. Impact of the ASCM on Gaussian noise. Results obtained with the different implementations of the ONLM filter on T1-w phantom.

C. Comparison of Rician NL-means filters

Figure 4 presents the results obtained for the Rician adapted NL-means filters . For all the metrics and all the noise levels, the proposed ORNLM filter with ASCM obtained the best results. The ORNLM filter and the ORNLM

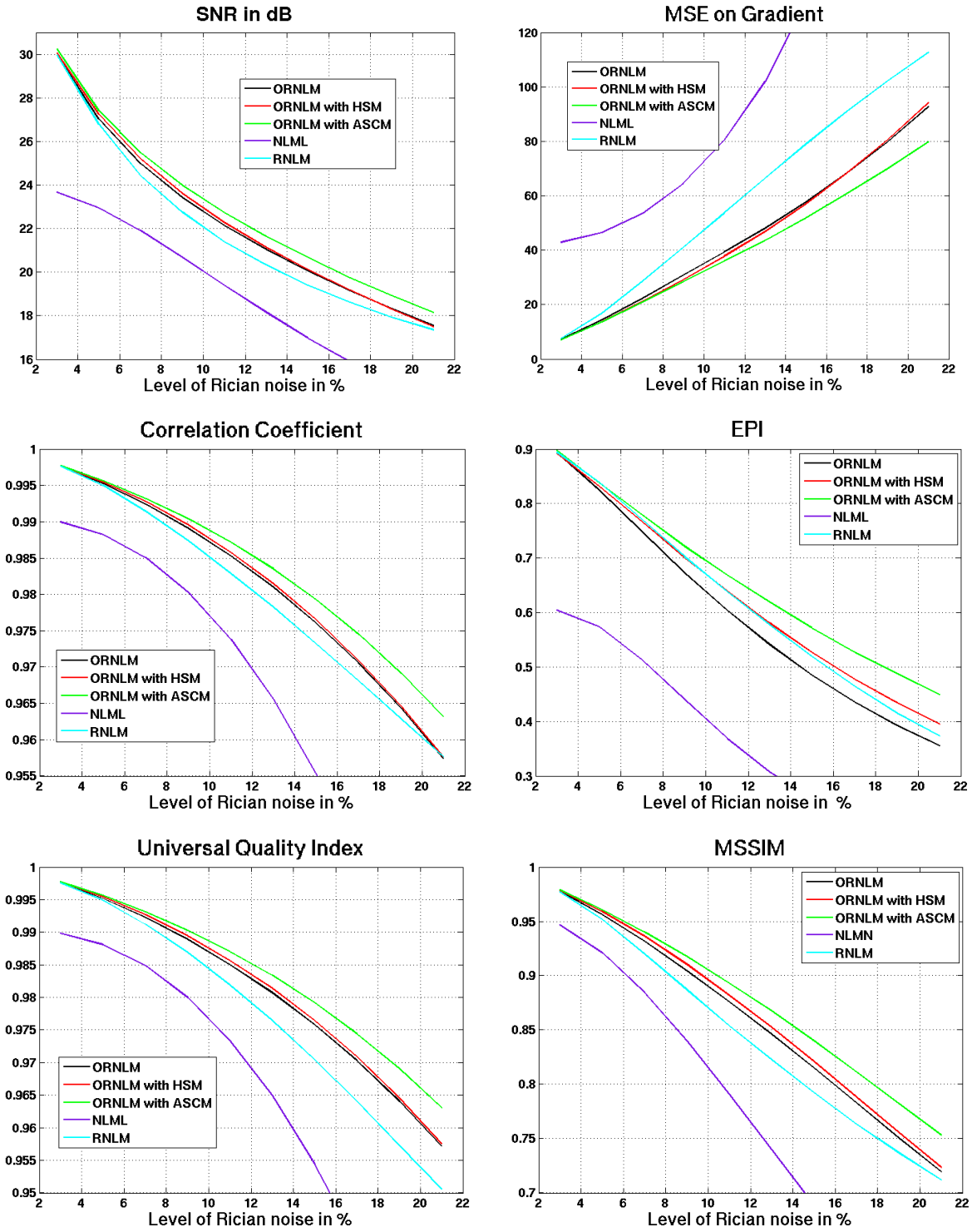

Fig. 4. Comparison of the Rician-adapted NL-means filters on Rician noise.

filter with HSM obtained similar results. The empirical subband mixing proposed for Gaussian noise in [7] did not improve results on Rician noise, while, without any tuning, the ASCM automatically optimizes the denoising for both Gaussian and Rician noises.

Figure 5 shows the denoising results at 17% Rician noise. The voxelwise implementation of Rician NL-means filter (i.e. RNLM filter) obtained worse results than ORNLM-based filters due to soft structure degradation (see Fig. 5). As shown in [6], the preselection strategy better preserves contrast and edges. The 3D extention of NLML approach proposed in [12] tended to leave more residual noise (see Fig. 5), thus adversely impacting the quality metrics.

Table I presents the computational time required by the different filters used during our comparison.

TABLE I

COMPARISON OF DIFFERENT IMPLEMENTATIONS OF RICIAN NL-MEANS FILTERS IN TERMS OF COMPUTATIONAL TIME USING ON A SINGLE CORE OF A QUADCORE PENTIUM AT 2.4 GHZ

Filter	Computational time (in s)
ORNLM	229
ORNLM with HSM	331
ORNLM with ASCM	340
NLML	65450
RNLM	4971

V. EXPERIMENTS ON CLINICAL DATASETS

A. T1-weighted anatomical MR image

The proposed method has been applied to a T1-w MR image of $170 \times 256 \times 256$ voxels acquired on a 1.5T Philips Gyroscan (Best, the Netherlands). Figure 6 shows the denoised image and the absolute residual image. Visually, the ORNLM with ASCM preserves the anatomical structures in the image while efficiently removing the noise.

B. DW-MRI with high b-value

As assessed by quantitative results on phantom dataset, the proposed multiresolution approach is particularly efficient in the case of high levels of noise. To emphasize the improvement brought by using ASCM, the ORNLM filter with ASCM has been applied on a Diffusion-Weighted MRI (DW-MRI) dataset acquired with high b-value (i.e., these images have a very high level of noise). MRI data were acquired on a Siemens 3T Trio MR scanner (Siemens

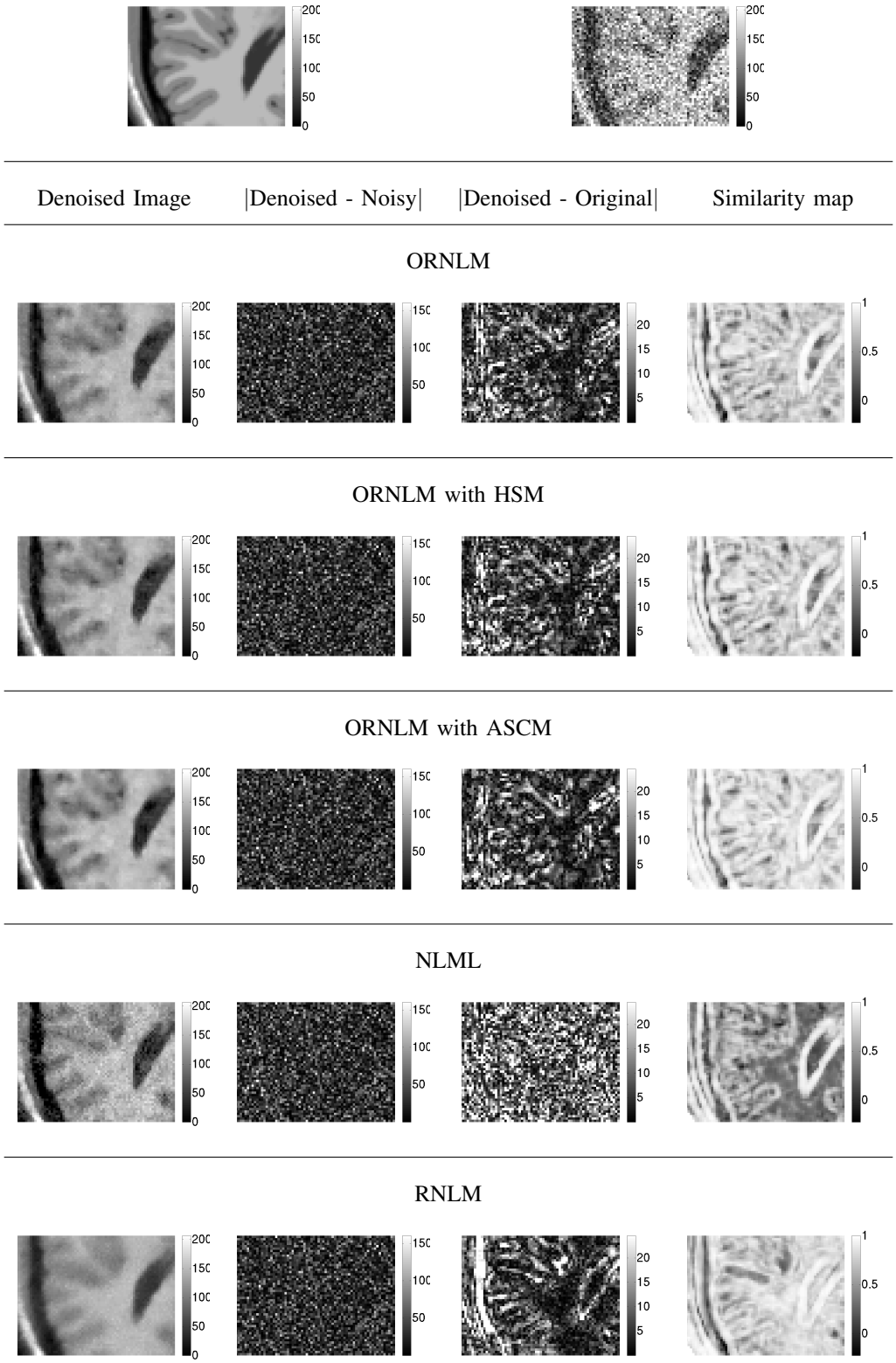

Fig. 5. Comparison of the Rician-adapted NL-means filters. Results obtained on T1-w phantom corrupted by 17% Rician noise. The images and the maps have the same scale for all the Rician filters compared.

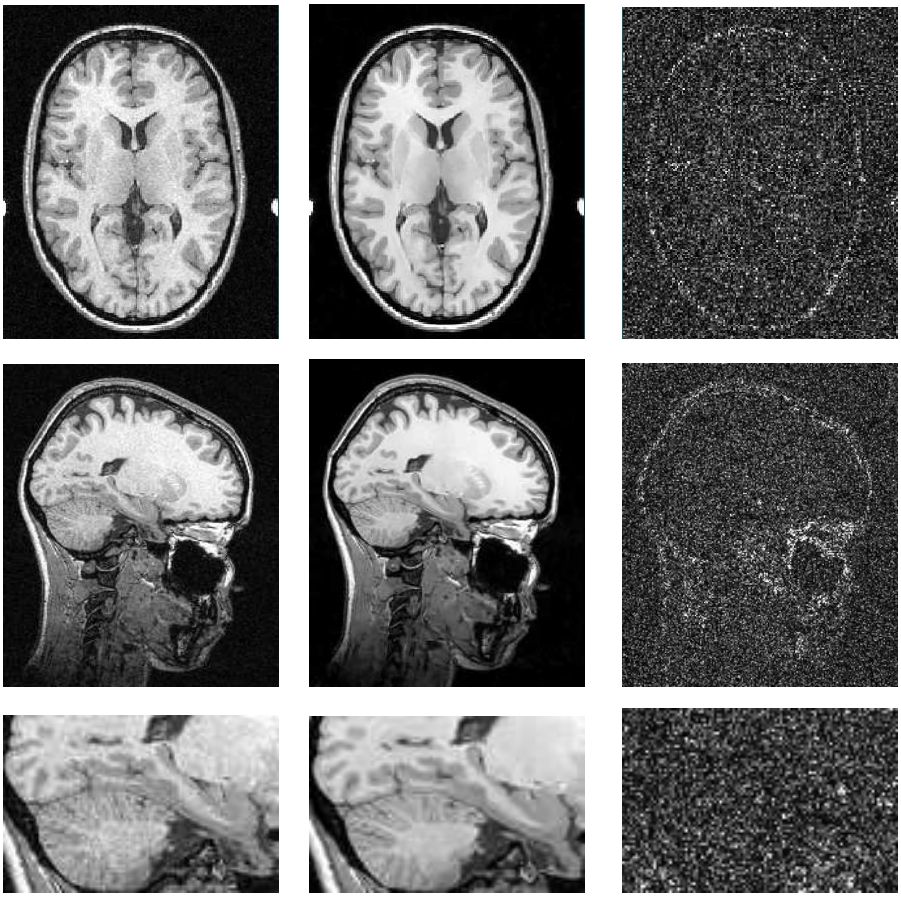

Fig. 6. Left: Original T1-w image ($SNR = 12.55dB$ estimated with [23]). Middle: Denoised image obtained with the ORNLM with ASCM. Right: Absolute residual image (i.e. absolute values of the removed noise).

Medical Systems, Erlangen, Germany) using an 8-channel phased-array head coil. The dataset was acquired with 99 diffusion encoding directions and 10 b=0 images, 2mm isotropic voxel size, 63 slices, b=3000 s/mm^2, TE=121ms and TR=11.1s. As proposed in [5], [16], each b=0 image and each gradient direction has been denoised separately.

Figure 7 shows the results obtained on one of the gradient directions. The corresponding FA map is presented in Fig. 8. The FA maps computed on the denoised DW-MRI contained less noise and less artifacts than the FA map computed on the original data, especially on the frontal region (see green box in Fig. 8).

Finally, we conclude with an example of fiber tracking to show the qualitative performance of our denoising filter on a DTI application using the same data as the previous example. Streamline tractography was obtained by using the method described in [24]. The central slice of the corpus Callosum was used as seed region for the tractography (i.e. all the voxels in this region were used for the initialization).

Figure 9 shows the tractography obtained with the noisy data and the DW-MRI denoised with the proposed filter. The fibres tracked on denoised image were longer and more dense.

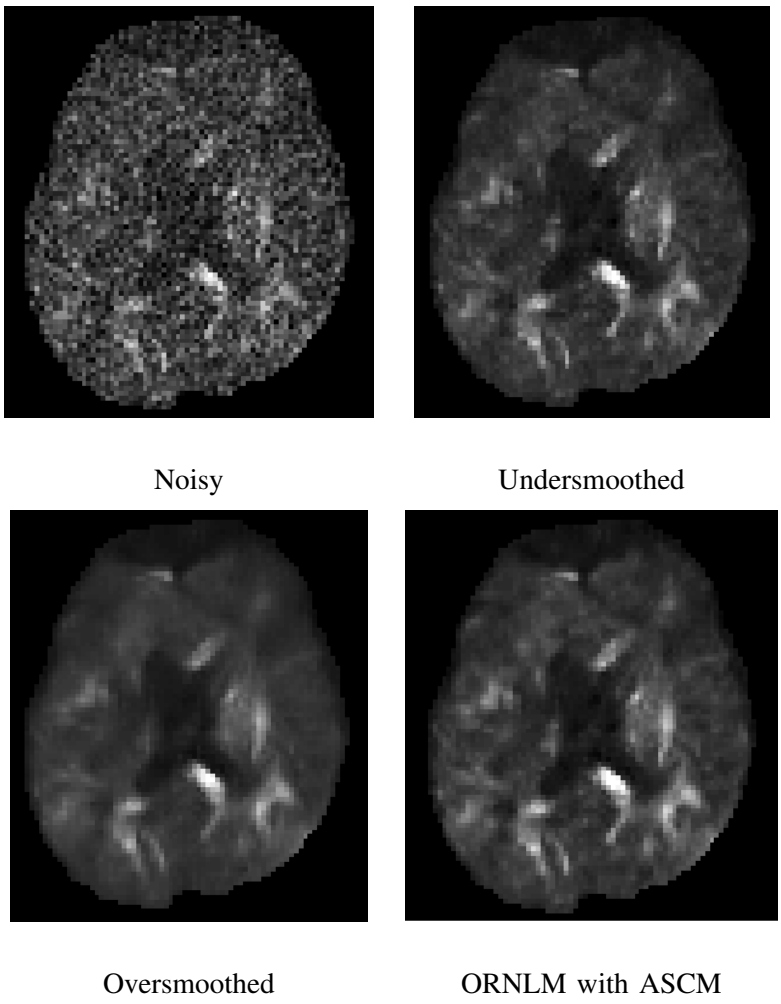

Fig. 7. Gradient direction of the DWI dataset with high b-value . **Top**. Left: Noisy DW-MRI. Right: Undersmoothed I_u DW-MRI ($SNR = 7.12dB$ estimated with [23]). **Bottom**. Left: Oversmoothed I_o DW-MRI. Right: DW-MRI denoised with ORNLM filter with ASCM.

VI. Conclusion

This paper presents a new multiresolution NL-means filter for 3D MR image denoising. Based on an adaptive soft wavelet coefficient mixing driven by spatial and frequency information, the proposed filter improves the denoising performances of the NL-means filter. In fact, compared to the ONLM filter [6], by combining two denoised images with different patch sizes in the transform domain, the proposed ONLM with ASCM filter improves denoising along all the spatial-frequency resolutions of the image. While in [7], the hard mixing was performed empirically at the subband level, in this paper we proposed a new framework enabling the soft mixing at the coefficient level of

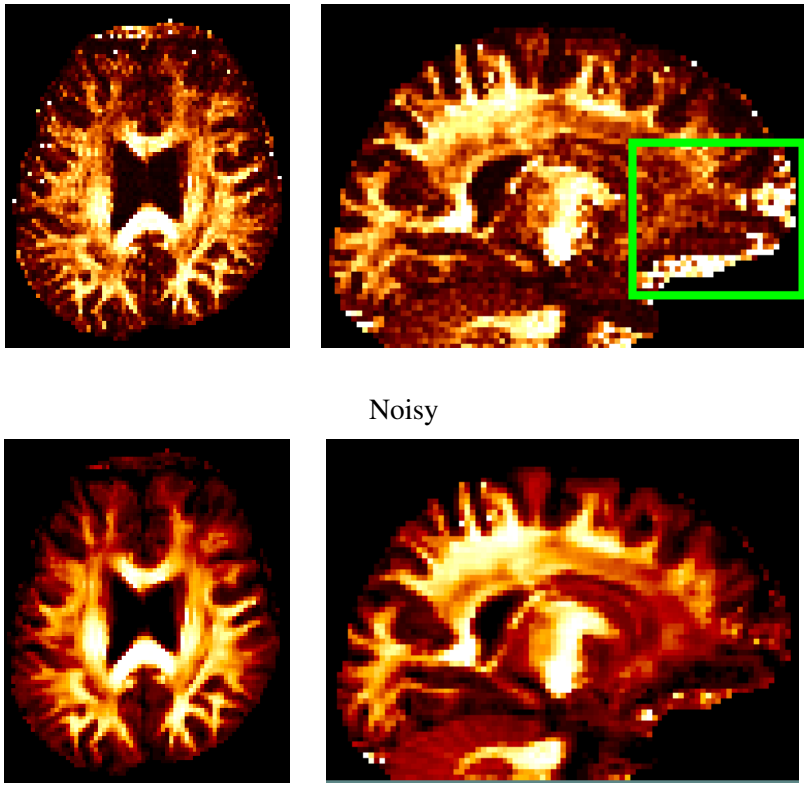

Fig. 8. Fractional Anisotropy (FA) map computed on the DW-MRI dataset acquired with high b-value. **From Top to Bottom**: FA map computed on noisy DW-MRI and on denoised DW-MRI with ORNLM filter with ASCM.

two denoised images, where the input noisy image is used to automatically drive the mixing. This new framework demonstrated better denoising capability using 6 quality metrics on Gaussian and Rician noise.

To deal with intensity bias introduced by Rician noise, Conventional Approach (CA) has been used within our NL-means framework as in [16]. Experiment on bias estimation indicates that using CA for Rician noise removal within a 3D NL-Means framework appears to be the most efficient and practical. The denoising results on the BrainWeb MR phantom shows that the proposed Adaptive Multiresolution NL-means filter outperforms the optimized implementations presented in [6], [7] on Gaussian noise. Moreover, the Rician version of the proposed filter obtained competitive results compared to recently proposed Rician NL-means filters [12], [16] in terms of denoising performance and computational time. Finally, experiments on real data emphasized the potential of the proposed filter on images with very high levels of noise such as DW-MRI with high b-values.

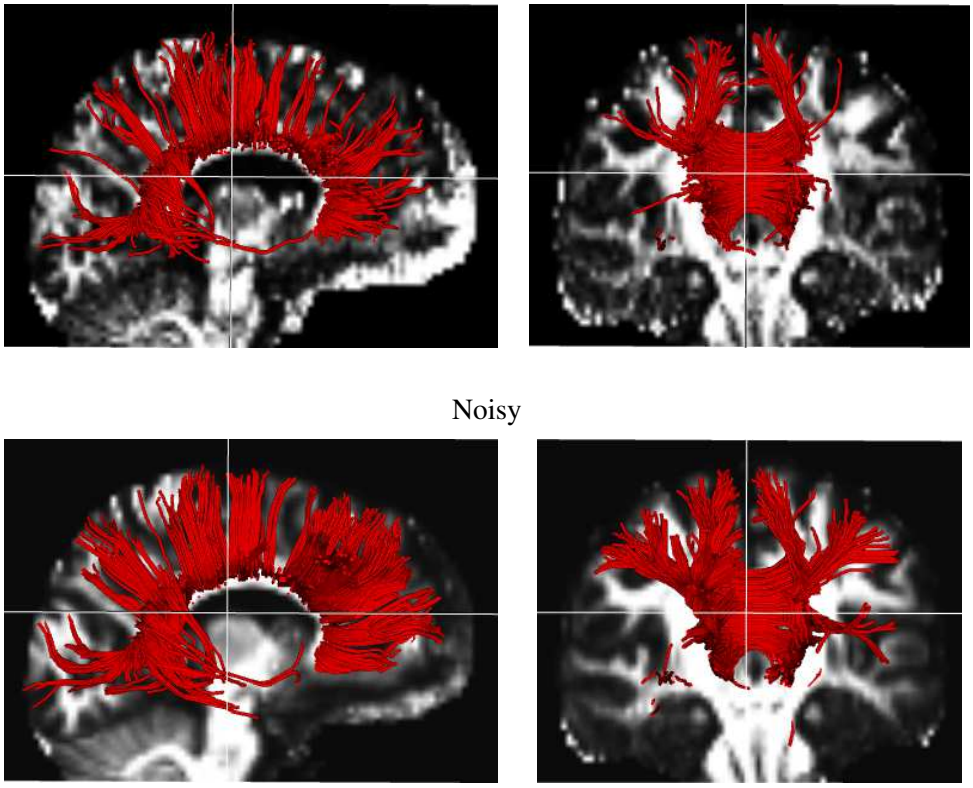

Fig. 9. Fibres seeded from the central slice of the Corpus Callosum.

ACKNOWLEDGEMENTS

We would like to thank Ilana Leppert and Prof. Bruce Pike for providing DW-MRI data and help in result analysis. We also want to thank Nicolas Guizard and Dr Vladimir Fonov for their help during tractography experiment. We thank Dr Luis Concha for his fruitful discussions about DTI processing. This work has been partially supported by the canadian grant Industry Cda (CECR)-Gevas-OE016 and by the Spanish Health Institute Carlos III through the RETICS Combiomed, RD07/0067/2001.

REFERENCES

[1] G. Gerig, R. Kikinis, O. Kübler, and F. Jolesz, "Nonlinear anisotropic filtering of MRI data," *IEEE Transactions on Medical Imaging*, vol. 11, pp. 221–232, June 1992.

[2] R. Nowak, "Wavelet-based rician noise removal for magnetic resonance imaging," *IEEE Transactions on Image Processing*, vol. 8, pp. 1408–1419, October 1999.

[3] A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile wavelet domain noise filtration technique for medical imaging," *IEEE Trans. Medical Imaging*, vol. 22, pp. 323–331, 2003.

[4] M. E. Alexander, R. Baumgartner, A. R. Summers, C. Windischberger, M. Klarhoefer, E. Moser, and R. L. Somorjai, "A wavelet-based method for improving signal-to-noise ratio and contrast in MR images.," *Magn Reson Imaging*, vol. 18, pp. 169–180, February 2000.

[5] S. Aja-Fernandez, M. Niethammer, M. Kubicki, M. E. Shenton, and C. F. Westin, "Restoration of dwi data using a rician lmmse estimator," *Medical Imaging, IEEE Transactions on*, vol. 27, no. 10, pp. 1389–1403, 2008.

[6] P. Coupé, P. Yger, S. Prima, P. Hellier, C. Kervrann, and C. Barillot, "An Optimized Blockwise NonLocal Means Denoising Filter for 3-D Magnetic Resonance Images," *IEEE Transactions on Medical Imaging*, vol. 27, pp. 425–441, April 2008.

[7] P. Coupé, P. Hellier, S. Prima, C. Kervrann, and C. Barillot, "3D wavelet subbands mixing for image denoising," *Journal of Biomedical Imaging*, vol. 2008, no. 3, pp. 1–11, 2008.

[8] K. Krissian and S. Aja-Fernández, "Noise-driven anisotropic diffusion filtering of mri.," *IEEE transactions on image processing*, vol. 18, pp. 2265–2274, October 2009.

[9] S. P. Awate and R. T. Whitaker, "Feature-preserving MRI denoising: A nonparametric empirical bayes approach," *Medical Imaging, IEEE Transactions on*, vol. 26, no. 9, pp. 1242–1255, 2007.

[10] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Modeling and Simulation*, vol. 4, no. 2, pp. 490–530, 2005.

[11] J. V. Manjón, J. Carbonell-Caballero, J. J. Lull, G. García-Martí, L. Martí-Bonmatí, and M. Robles, "MRI denoising using Non-Local Means," *Medical Image Analysis*, vol. 4, no. 12, pp. 514–523, 2008.

[12] L. He and I. R. Greenshields, "A Nonlocal Maximum Likelihood Estimation Method for Rician Noise Reduction in MR Images," *IEEE TMI*, vol. 28, no. 2, pp. 165–172, 2009.

[13] T. Brox, O. Kleinschmidt, and D. Cremers, "Efficient nonlocal means for denoising of textural patterns," *Image Processing, IEEE Transactions on*, vol. 17, no. 7, pp. 1083–1092, 2008.

[14] H. Gudbjartsson and S. Patz, "The Rician distribution of noisy MRI data," *Magnetic Resonance in Medicine*, vol. 34, pp. 910–914, 1995.

[15] J. Sijbers, A. den Dekker, J. V. Audekerke, M. Verhoye, and D. V. Dyck, "Estimation of the noise in magnitude MR images," *Magnetic Resonance Imaging*, vol. 16, no. 1, pp. 87–90, 1998.

[16] N. Wiest-Daesslé, S. Prima, P. Coupé, S. Morrissey, and C. Barillot, "Rician noise removal by non-local means filtering for low signal-to-noise ratio mri: Applications to DT-MRI," in *MICCAI'2008*, (New York, USA), September 2008.

[17] A. Lukin, "A multiresolution approach for improving quality of image denoising algorithms," in *ICASSP'06*, (Toulouse, France), pp. 857–860, May 2006.

[18] D. Donoho and I. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[19] S. G. Chang, B. Yu, and M. Vetterli, "Adaptive wavelet thresholding for image denoising and compression," *IEEE Transactions on Image Processing*, vol. 9, no. 9, pp. 1532–1546, 2000.

[20] D. Collins, A. Zijdenbos, V. Kollokian, J. Sled, N. Kabani, C. Holmes, and A. Evans, "Design and construction of a realistic digital brain phantom.," *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 463–468, 1998.

[21] F. Sattar, L. Floreby, G. Salomonsson, and B. Lovstrom, "Image enhancement based on a nonlinear multiscale method.," *IEEE transactions on image processing*, vol. 6, no. 6, pp. 888–895, 1997.

[22] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, pp. 600–612, April 2004.

[23] P. Coupé, J. V. Manjón, E. Gedamu, D. Arnold, M. Robles, and D. L. Collins, "Robust Rician noise estimation for MR images," *Medical Image Analysis*, vol. 14, no. 4, pp. 483 – 493, 2010.

[24] J. Campbell, P. Savadjiev, K. Siddiqi, and G. Pike, "Validation and regularization in diffusion MRI tractography," in *ISBI'06*, pp. 351–354, 2006.