Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI.

Nicolas Wiest-Daesslé, Sylvain Prima, Pierrick Coupé, Sean Patrick Morrissey, Christian Barillot

Non-local means variants for denoising of diffusion-weighted and diffusion tensor MRI

Nicolas Wiest-Daesslé, Sylvain Prima, Pierrick Coupé, Sean Patrick Morrissey, Christian Barillot

Unit/Project VisAGeS U746, INSERM - INRIA - CNRS - Univ-Rennes 1,
IRISA campus Beaulieu 35042 Rennes Cedex, France
{nwiestda,sprima,pcoupe,spmorris,cbarillo}@irisa.fr,
http://www.irisa.fr/visages

Abstract. Diffusion tensor imaging (DT-MRI) is very sensitive to corrupting noise due to the non linear relationship between the diffusion-weighted image intensities (DW-MRI) and the resulting diffusion tensor. Denoising is a crucial step to increase the quality of the estimated tensor field. This enhanced quality allows for a better quantification and a better image interpretation. The methods proposed in this paper are based on the Non-Local (NL) means algorithm. This approach uses the natural redundancy of information in images to remove the noise. We introduce three variations of the NL-means algorithms adapted to DW-MRI and to DT-MRI. Experiments were carried out on a set of 12 diffusion-weighted images (DW-MRI) of the same subject. The results show that the intensity based NL-means approaches give better results in the context of DT-MRI than other classical denoising methods, such as Gaussian Smoothing, Anisotropic Diffusion and Total Variation.

1 Introduction

Image processing procedures needed for fully automated and quantitative analysis (registration, segmentation, visualisation) require images with the best signal-to-noise ratio and the least artifacts in order to improve their performances. Most of the time, the hardware introduces artifacts during the acquisition (noise, intensity non-uniformities, geometrical deformations). Therefore, one critical issue is to remove the noise while keeping relevant image information. This is particularly true for diffusion-weighted MRI (DW-MRI) especially when they are acquired with high diffusion (b-value) coefficient. This paper focuses on denoising using variants of the non-local means (NL-means) method modified to deal with DT-MRI (NLMt) and DW-MRI, either gradient-by-gradient (NLM) or as a multi-spectral (NLMv) image. The NL-means variants are compared with the simple Gaussian Filter (GF), the Anisotropic Diffusion (AD) [13] and the Total Variation (TV) [15]. In particular the AD filter is frequently used for diffusion image denoising [4, 9] or tensor field regularisation [16].

The original publication is available at www.springerlink.com

2 Methods

2.1 The non-local means algorithm

First introduced by Buades *et al.* in [3], the NL-means algorithm is based on the natural redundancy of information in images to remove noise. In the theoretical formulation of the NL-means algorithm, the restored intensity of the voxel x_i, $NL(v)(x_i)$, is a weighted average of all voxels intensities in the image I. Let us denote:

$$NL(v)(x_i) = \sum_{x_j \in I} w(x_i, x_j) v(x_j), \tag{1}$$

where v is the intensity function and thus $v(x_j)$ is the intensity of voxel x_j and $w(x_i, x_j)$ the weight assigned to $v(x_j)$ in the restoration of $v(x_i)$. More precisely, the weight quantifies the similarity of voxels x_i and x_j under the assumptions that $w(x_i, x_j) \in [0, 1]$ and $\sum_{x_j \in I} w(x_i, x_j) = 1$.

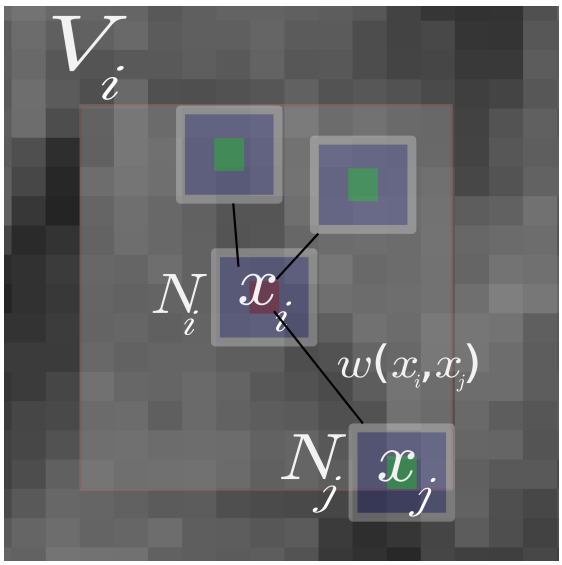

Fig. 1. Two-dimensional illustration of the NL-means principle. The restored value of voxel x_i is a weighted average of all intensities of voxels x_j in the search volume V_i. The weight $w(x_i, x_j)$ is based on the similarity of the intensities in cubic neighborhoods N_i and N_j around x_i and x_j.

In practice, voxels similar to i are only searched over a neighborhood V_i, so Eq. 1 is: $NL(v)(x_i) = \sum_{x_j \in V_i} w(x_i, x_j) v(x_j)$. For each voxel x_j in V_i, the weight $w(x_i, x_j)$ is related to the distance $d(v(N_i), v(N_j))$, N_i and N_j being neighborhoods around x_i and x_j, following:

$$w(x_i, x_j) = \frac{1}{Z(i)} e^{-\frac{d\left(v(N_i), v(N_j)\right)}{(h\hat{\sigma})^2}} \tag{2}$$

where $Z(i)$ is a normalization constant with $Z(i) = \sum_j w(x_i, x_j)$, $\hat{\sigma}$ is the estimation of the standard deviation of the noise using the pseudo-residuals

method [8] and h acts as a filtering parameter (for more details see [6] and Fig. 1). The distance d is expressed in general terms as: $d(v(N_i), v(N_j)) = \sqrt{\frac{1}{N}\sum_k^N \Delta(v(y_k), v(z_k))}$ where $N = \operatorname{card} N_i = \operatorname{card} N_j$ and y_k and z_k are the k-th voxels in the neighborhoods N_i and N_j. For a grey-level image, Δ is $\Delta(v(y_k), v(z_k)) = \|v(y_k) - v(z_k)\|^2$.

2.2 DW- and DT-MRI adaptations

This section introduces the NL-means as a method to remove noise from either the whole DW-MR dataset (with n directions, $n \geq 6$, plus the B_0 image) or the resulting DT-MR image. Three variants are proposed here, two acting on the DW-MRI and one on the DT-MRI:

1. **NLM**: each DW-MRI is denoised individually as described in Section 2.1 and the DT-MRI is estimated from these denoised DW-MRI,
2. **NLMv**: the whole set of DW-MRI is considered as a multi-spectral image, each voxel being a $(n+1)$-dimensional vector. The Δ is defined as:

$$\Delta(v(y_k), v(z_k)) = \sum_{i=1}^{n+1} \|v^i(y_k) - v^i(z_k)\|^2, \tag{3}$$

   $v^i(.)$ being the i-th component of the vector $v(.)$.
3. **NLMt**: the DT-MRI is computed from the raw DW-MRI and then denoised. The weighted average of the MRI intensities (grey levels) Eq. 1 is replaced by a Log-Euclidean weighted average [1, 12]: of the image diffusion tensors. Δ is defined as:

$$\Delta(v(y_k), v(z_k)) = \|\log(v(y_k)^{-\frac{1}{2}} v(z_k) v(y_k)^{-\frac{1}{2}})\|^2, \tag{4}$$

   $v(y_k)$ and $v(z_k)$ being the tensors at voxels y_k and z_k.

The Log-Euclidean framework could have been replaced by other methods [7, 12, 17].

2.3 Implementation details

The NLM method uses a cubic neighborhood ($\operatorname{card} N_i = 27$). For NLMv and NLMt, considering that both vectors and tensors convey enough information for denoising, $\operatorname{card} N_i$ is set to 1. In these cases, having larger neighborhoods N_i makes it difficult to find similar blocks in the search area V_i and thus limits the denoising capacities of the algorithms. The search area V_i is chosen to be identical for all the NL-means variants ($\operatorname{card} V_i = 11^3$ voxels).

2.4 Comparison measure

A comparison measure is needed to validate the denoising methods with respect to a ground truth. We define the distance between two DT-MRI as:

$$RMS = \sqrt{\frac{1}{\text{card}\,\Omega} \sum_{\text{card}\,\Omega} d\left(\widehat{I}, I_d\right)^2}, \tag{5}$$

with Ω the masked diffusion tensor image grid, $\widehat{I}$ the reference DT-MRI, I_d the denoised DT-MRI, and d a distance over tensors. The Log-Euclidean distance is selected as it is specifically designed for tensors as described in Sec. 2.2. The comparison is restricted to cerebral tissues, where the estimation of a diffusion tensor is relevant.

3 Validation and results

3.1 Dataset

In order to evaluate the performances of the different algorithms on DT-MRI multiple tests are performed on a reference data set. The reference data set is constructed by averaging multiple acquisitions of the same subject. The acquisition protocol is a single-shot spin echo EPI sequence on a Siemens 1.5T scanner, with diffusion encoding (10 directions, $b = 1000$ s/mm^2, voxels=1.875×1.875×5 mm^3, 24 slices, 24 cm FOV). The acquisition is repeated 12 times with identical slice locations and each acquisition has a run time of 8 minutes. Each diffusion-weighted acquisition is corrected for distortions [18].

Numerous methods exist for the estimation of the tensor [10, 16]. We simply choose to estimate the tensor by classical linear regression.

3.2 Leave-one-out comparison

To assess the validity of the proposed denoising methods, a leave-one-out approach is devised. For each DW acquisition I^i_{noisy} the 11 other DW-MRI are averaged gradient-by-gradient, giving $I^i_{average}$ (*cf* Fig. 2, left). A DT-MRI is estimated from $I^i_{average}$ and serves as a comparison basis. The selected image, I^i_{noisy} (or its corresponding DT-MRI), is then denoised with the 6 denoising techniques, and the resulting denoised DT-MRI is estimated. The error between this denoised image and the ground truth data built from $I^i_{average}$ is computed using the measure defined in Section 2.4. The process is then iterated, yielding 12 RMS errors, which are finally averaged to give a global RMS error. These error measures are displayed on Fig. 2 (right) for the 6 denoising methods. This leave-one-out method helps avoiding the introduction of bias.

The denoising using the NLMt techniques yields very poor results, probably due to the poor redundancy of tensor information in the image. Computing the weights for each voxel shows that on average only 8 significantly similar tensors are found, whereas for grey level or vector images the number of significantly similar blocks is generally higher than 100.

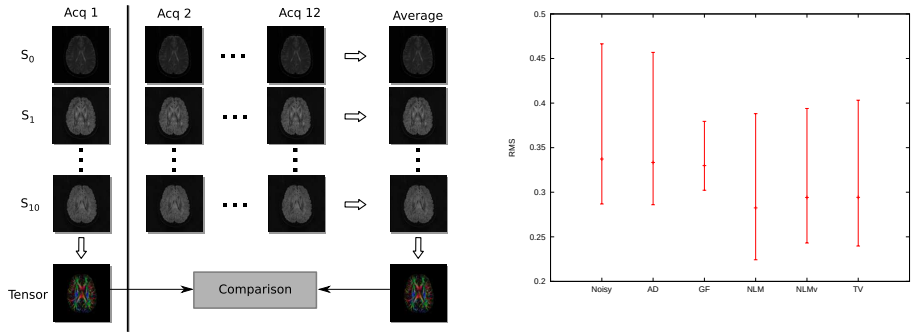

Fig. 2. Left: **Scheme of the first step of the leave-one-out validation.** An acquisition is selected (I^1_{noisy}), and the others are averaged, giving $I^1_{average}$. The DW-MRI (or DT-MRI) corresponding to I^1_{noisy} are denoised with each of the 6 algorithm; the associated DT-MRI is computed and quantitatively compared with the DT-MRI computed from $I^1_{average}$, giving an error measure ϵ_1. The process is then repeated with $i = 2, ..., 12$ and the global error measure is computed as $\frac{1}{12}\sum_i \epsilon_i$. Right: **Error plot of the RMS for each method.** The bar length indicates the min and max error over the 12 experiments; the middle mark indicates the mean value. The acronyms are as follows: GF: Gaussian Filter, AD: Anisotropic Diffusion, TV: Total Variation, NLMv: NL-means vector, NLM: NL-means gradient-by-gradient, NLMt: NL-means tensor. The NLMt method is not plotted due to poor results: average RMS is 1.2.

3.3 Comparison with different noise levels

In this section, the average of the 12 images I_{avg} is used as a reference. A new image I_n is built by adding Rician noise with different levels. In Collins *et al.* [5], the noise percentage p is related to the standard deviation of the Gaussian noise σ and the mean value ν of the brightest tissue following $p = 100\sigma/\nu$. The same idea is used here with Rician noise. The mean intensity of the CSF in the non-diffusion-weighted image (S_0) is used for the ν value. The RMS error is computed between each denoised DT-MRI and the ground truth DT-MRI computed from $I_{average}$.

Results are shown in Fig. 3. At low levels of noise (below 4%), TV and AD perform better than the NL-means filters. That could be partially explained by the fact that the estimation of the noise by pseudo-residuals used in the NL-means variants is known to be overestimated for these low levels of noise. At higher levels (in the range 5-10%), usually met in real DW-MRI, the NL-means filters outperform all the other filters, NLM being constantly better than NLMv.

3.4 Choice of the filters parameters

Each proposed method needs specific parameters for denoising. For a fair comparisons of all the methods, those parameters are selected with an optimisation procedure so that each method gives its best result for a given experiment. In

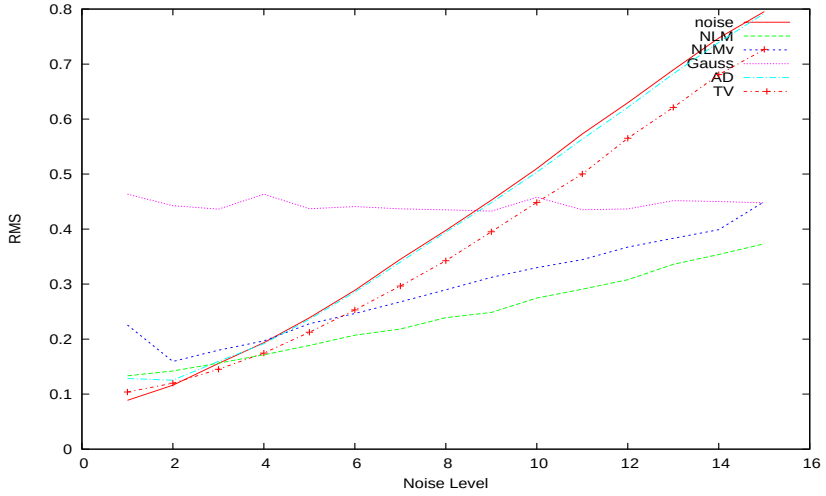

Fig. 3. Plot of different noise levels and RMS. Noise is added to the reference image. The image is then denoised and compared to the original.

practice, according to the 12 different RMS errors computed in the leave-one-out experiment, the noise level is between 6 and 7 percents. All the parameters are optimised for this level of noise added to the $I_{average}$ DW-MRI.

This optimisation is performed with the Nelder-Mead's downhill simplex algorithm [14]. The cost function for this optimiser is the measure described in Sec. 2.4. Initial guesses for the parameters are empirically chosen after a few manual tests, and are used to initialise the downhill simplex. The unknown parameters are: number of iterations and regularisation strength (TV and AD), kernel size (GF), and filtering parameter h (NL-means variants).

3.5 Visual assessment

In Figure 4, we display axial slices at the level of the ventricles for the ground truth data (DT-MRI computed from $I_{average}$), the raw DT-MRI computed from one of the acquisitions I^i_{noisy}, and the 6 denoised images. The color encodes the principal direction of diffusion (colinear to the eigenvector of the tensor with maximal eigenvalue), weighted by the fractional anisotropy [11]. The reference image has smooth color transitions but also sharp edges. The GF filtered image efficiently removes the noise but suppresses the edges and lowers the anisotropy of tensors. The standard NL-means seems to be the best filter, followed by TV, NLMv, AD and NLMt, which confirms the quantitative values in Fig. 3.

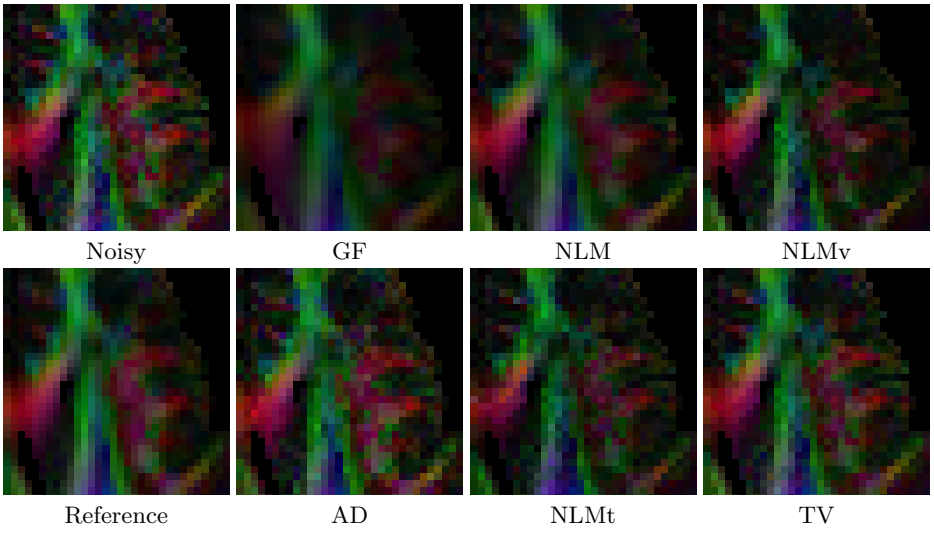

Fig. 4. Visual Comparison of the different algorithms. The color encodes the principal direction of diffusion (colinear to the eigenvector of the tensor with maximal eigenvalue), weighted by the fractional anisotropy [11].

4 Conclusion and further works

This paper presents new variants of the Non-Local (NL) means algorithm, applied to diffusion-weighted and diffusion tensor images. The validations performed on a reference dataset underline how the NL-means denoising outperforms well-established other methods, such as Anisotropic Diffusion [13] and Total Variation [15]. The results are obtained from the denoising of either the diffusion images or the diffusion tensor. Our comparison does not take into account the Rician nature of the noise, and comparison with more specific denoising methods [2] will be performed in the future. The direct denoising of the DT-MRI with our proposed NL-means variant does not achieve good performances. This relates to the fact that the number of similar tensors inside the search region is quite low (≈ 8). The lower quality of the direct denoising of DT-MRI compared to denoising on DW-MRI is in line with the literature [2].

The effect of such denoising techniques needs to be investigated in pathological cases. For instance Multiple Sclerosis clearly shows changes in diffusion coefficients (such as fractional anisotropy and mean diffusivity). The effect of denoising must be studied in lesion areas to make sure these are well preserved in terms of their diffusion characteristics. Moreover, the impact of this NL-means denoising variants on the performances of post-processing algorithms, such as segmentation and fiber tracking has to be further investigated.

References

1. V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Fast and simple calculus on tensors in the Log-Euclidean framework. In *MICCAI'05*, volume 3749, pages 115–122, October 2005.
2. S. Basu, P.x T. Fletcher, and Ross T. Whitaker. Rician noise removal in diffusion tensor mri. In *MICCAI'2006*, pages 117–125, 2006.
3. A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.
4. B. Chen and E. Hsu. Pde denoising of MR diffusion tensor imaging data. In *ISBI'04*, pages 1040–1042, 2004.
5. D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans. Design and construction of a realistic digital brain phantom. *IEEE Trans. Med. Imaging*, 17(3):463–468, 1998.
6. P. Coupé, P. Yger, and C. Barillot. Fast Non Local Means Denoising for 3D MR Images. In *MICCAI'2006*, volume 4191, pages 33–40, October 2006.
7. P. T. Fletcher and S. C. Joshi. Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors. In *CVAMIA and MMBIA 2004*, volume 3117, pages 87–98. Springer, May 2004.
8. T. Gasser, L. Sroka, and C. Steinmetz. Residual variance and residual pattern in non linear regression. *Biometrika*, 73(3):625–633, 1986.
9. J. E. Lee, M. K. Chung, and A. L. Alexander. Evaluation of anisotropic filters for diffusion tensor imaging. In *Biomedical Imaging: Macro to Nano, 2006. 3rd IEEE International Symposium on*, pages 77–78, 2006.
10. J.-F. Mangin, C. Poupon, C. Clark, D. Le Bihan, and I. Bloch. Distortion correction and robust tensor estimation for MR diffusion imaging. *Med Image Anal*, 6(3):191–198, September 2002.
11. S. Pajevic and C. Pierpaoli. Color schemes to represent the orientation of anisotropic tissues from diffusion tensor data: application to white matter fiber tract mapping in the human brain. *Magn Reson Med*, 42(3):526–540, September 1999.
12. X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.
13. P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(7):629–639, 1990.
14. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, Cambridge (UK) and New York, 2nd edition, 1992.
15. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. pages 259–268, 1992.
16. D. Tschumperlé and R. Deriche. Variational frameworks for DT-MRI estimation, regularization and visualization. In *ICCV'03*, pages 116–121, October 2003.
17. Z. Wang and B. C. Vemuri. DTI segmentation using an information theoretic tensor dissimilarity measure. *IEEE Trans Med Imaging*, 24(10):1267–1277, October 2005.
18. N. Wiest-Daesslé, S. Prima, S. P. Morrissey, and C. Barillot. Validation of a new optimisation algorithm for registration tasks in medical imaging. In *ISBI'07*, April 2007.