

# Efficient and Robust Non-Local Means Denoising Methods for Biomedical Images

Matt Judson<sup>1,\*</sup>, Troy Viger<sup>2,\*\*</sup>, and Hyeona Lim<sup>2,\*\*\*</sup>

<sup>1</sup>EMC Insurance, USA

<sup>2</sup>Mississippi State University, USA

**Abstract.** Denoising is an important step to improve image quality and to increase the performance of image analysis. However, conventional partial differential equation based image denoising methods, especially total variation functional minimization techniques, do not work very well on biomedical images such as magnetic resonance images (MRI), ultrasound, and X-ray images. These images present small structures with signals barely detectable above the noise level which involve more complex noise and unclear edges. The recently developed non-local means (NLM) filtering method can treat these types of images better. The standard NLM filter uses the weighted averages of similar patches present in the image. Since the NLM filter is a non-local averaging method, it is very accurate in removing noise but has computational complexity. We develop efficient and optimized NLM based methods and their associate numerical algorithms. The new methods are still accurate enough and more efficient than the original NLM filter. Numerical results show that the new methods compare favorably to the conventional denoising methods.

## 1 Introduction

Image denoising is among the most fundamental problems in image processing with many applications. A large range of methods covering various fields of mathematics and engineering are available for denoising an image. The initial denoising models were derived from energy minimization using nonlinear partial differential equations (PDEs) [1]-[6]. The filtering based models using smoothing operators have also been used for quite a long time especially for engineering applications [7]-[11]. The most successful one among them was the non-local means method proposed by Buades, *et al.* in 2005 [11]. There have also been some efforts to extend some known PDEs and variational techniques in image processing to the non-local framework [12]-[14].

The filtering based models are mainly based on convolutions. The image is convoluted with a symmetric kernel in effort to smooth out the noise. To remove noise, the affected pixels are smoothed out using information from surrounding pixels. The choice of neighboring pixels play an important role in these models. The most basic smoothing

---

\*e-mail: matt\_judson935@yahoo.com

\*\*e-mail: tvv11@msstate.edu

\*\*\*e-mail: hlim@math.msstate.edu

filter is the Gaussian filter, where the image is locally convoluted with Gaussian kernel. It has been shown by Buades, *et al.* in [11] that the Gaussian convolution is optimal for harmonic functions, but it destroys edges and textures. The first non-local filtering method, where neighbors are chosen from the whole image as opposed to the adjacent pixels in local filters, was developed by Yaroslavsky [7]. The neighborhood filter has a big advantage of not blurring the edges much, compared to the local filters. However, the method will not give expected results when both pixel values are noisy. Moreover, it was also shown in [11] that neighborhood filters can create artificial shocks. In [11], the authors proposed the new state of art image denoising algorithm known as non-local means (NLM) algorithm. The method is non-local since each pixel is denoised using the weighted average of all the pixels in the entire image. Several numerical examples in [11] showed that the NLM method is very accurate in removing noise in the given image compared to other conventional filtering methods and even to some PDE based methods. Also, the method works very well on biomedical images which have complex noise and unclear edges. However, the main drawback of the non-local means denoising method is on computational complexity. Due to the non-locality, the method is very slow in achieving the results and hence quite impractical. Coupé *et al.* [15] reduced the computational time by developing an optimized NLM filter using a blockwise implementation and a pre-selection of the most relevant blocks.

In this paper, we develop new blockwise NLM filters based on the model developed by Coupé *et al.* [15]. Section 2 comprises of preliminary with more detailed information on the standard NLM model and the blockwise NLM method developed by Coupé *et al.* In Section 3, new methods and their numerical procedures are introduced and discussed. The numerical results in Section 4 show that the new methods outperform the conventional ones. Finally, we summarize our outcomes in Section 5.

## 2 Preliminary

In this section, we describe the two filtering models, the original NLM model [11] and Coupé *et al.*'s blockwise NLM [15].

### 2.1 Non-local means filtering method

The NLM filter [11] restores every pixel in the image by computing the weighted average of all the pixels in the image. The weight calculates the similarity between the intensity of the local neighborhoods. The method is given by the formula:

$$NL[u](\mathbf{x}) = \frac{1}{C(\mathbf{x})} \int_{\Omega} e^{-\frac{(G_a * u(\mathbf{x}+) - u(\mathbf{y}+))^2(0)}{h^2}} u(\mathbf{y}) d\mathbf{y}, \quad (1)$$

where  $\mathbf{x} \in \Omega$  and  $C(\mathbf{x}) = \int_{\Omega} e^{-\frac{(G_a * u(\mathbf{x}+) - u(\mathbf{z}+))^2(0)}{h^2}} dz$ . Here,  $G_a(\mathbf{x}) = \frac{1}{4\pi a^2} e^{-\frac{|\mathbf{x}|^2}{4a^2}}$  is a Gaussian kernel and  $h, a$  are filtering parameters. For a given discrete noisy image  $u = \{u(i) : i \in I\}$ , the estimated value  $NL[u](i)$  is computed as

$$NL[u](i) = \sum_{j \in I} w(i, j) u(j), \quad (2)$$

where the weight  $w(i, j)$  depends on the similarity of the  $i$  and  $j$  pixels and satisfies the conditions  $0 \leq w(i, j) \leq 1$  and  $\sum_j w(i, j) = 1$ . For estimating the weight between the two pixels, the authors considered the similarities of a square neighborhood  $\mathcal{N}_k$  of a fixed size at each pixel

$k$ . The similarity is measured as a decreasing function of the Gaussian-weighted Euclidean distance [11],  $\|u(\mathcal{N}_i) - u(\mathcal{N}_j)\|_{2,\alpha}^2$ . The weight function is defined as

$$w(i, j) = \frac{1}{C(i)} e^{-\frac{\|u(\mathcal{N}_i) - u(\mathcal{N}_j)\|_{2,\alpha}^2}{h^2}}, \quad (3)$$

where  $C(i)$  is the normalizing constant given by

$$C(i) = \sum_j e^{-\frac{\|u(\mathcal{N}_i) - u(\mathcal{N}_j)\|_{2,\alpha}^2}{h^2}}. \quad (4)$$

The similar neighborhoods have a larger weight which contributes more to the approximation  $NL[u](i)$ . Therefore, the method produces more accurate denoising results. For computational purposes, the authors reduced the search window of each pixel to a window of size  $21 \times 21$  pixels and the similarity square neighborhood  $\mathcal{N}_i$  to a  $7 \times 7$  block. Therefore for the image of size  $N \times N$ , the final complexity of the algorithm is in the order of  $O(49 \cdot 441 \cdot N^2)$ .

## 2.2 Blockwise non-local means method

The standard NLM method is very accurate in removing noise compared to other conventional denoising methods. However, the main drawback is on computational inefficiency due to its non-locality. To overcome the drawback, the authors of [15] suggested the use of blocks  $B_{i_k}$  of size  $(2\alpha + 1)^2$  with overlapping supports and performed NLM restoration on these blocks instead of pixels. The following is the proposed formula to restore images:

$$NL[u](B_{i_k}) = \sum_{B_j \in V_{i_k}} w(B_{i_k}, B_j) u(B_j) \quad (5)$$

with

$$w(B_{i_k}, B_j) = \frac{1}{Z_{i_k}} e^{-\frac{\|u(B_{i_k}) - u(B_j)\|_2^2}{\hat{h}^2}}, \quad (6)$$

where  $Z_{i_k}$  is a normalization constant satisfying  $\sum_{B_j \in V_{i_k}} w(B_{i_k}, B_j) = 1$  and  $\hat{h}$  is a smoothing parameter. Inspired by Mahmoudi and Sapiro [16] who proposed a method to pre-select the most similar neighborhoods only in order to avoid unnecessary weight computation, the weight  $w(B_{i_k}, B_j)$  was only considered when the ratios of mean and variance of  $u(B_{i_k})$  and  $u(B_j)$  are within sufficiently small ranges around 1. This blockwise NLM method significantly reduced the computation time for the NLM method.

## 3 New blockwise based non-local means algorithms

In this section, we develop more efficient denoising algorithms based on the blockwise NLM method developed by Coupé *et al.* [15]. The blockwise NLM method significantly reduced the computational time of the standard NLM method. The blockwise NLM filter is more practical in medical applications since the MRI and X-ray images are mostly 3D or stacks of 2D, which requires a large amount of computation to achieve desirable denoising results. Therefore, computational efficiency plays a crucial role in biomedical image denoising as long as it does not lose much accuracy. In [15], authors considered blocks overlapping with their neighboring blocks by one column or one row of pixels to ensure global continuity. As a result, they obtained multiple approximations of the restored intensity  $NL(u)(x_i)$  at the pixel  $x_i$  for up to 4 values. The final restored intensity value at the pixel  $x_i$  was then averaged

among those several approximations. Averaging intensity values at overlapping locations can increase the computational costs especially when dealing with a large size data.

We consider blocks that are still overlapping with neighboring blocks and perform a NLM-like restoration Eqs. (5)-(6). But instead of calculating the average of multiple intensity approximations, we replace the value with the most current value as we proceed to the next block. We also consider non-overlapping blocks to have just one intensity approximation at each pixel. In both cases, the denoised images are globally continuous although the new approaches may produce more checkerboard effects. Both methods do not require extra computations of averaging multiple intensities at overlapping blocks and therefore enhance the computational efficiency of the original blockwise NLM developed by Coupé *et al.* To further reduce the computational complexity, we also adopt the strategy in [15] that the Gaussian-weighted Euclidean distance  $\| \cdot \|_{2,a}^2$  is replaced by the classical Euclidean distance  $\| \cdot \|_2^2$ . We summarize the steps of new algorithms as follows:

1. Use a standard Euclidean norm for calculating the weights.
2. Perform NLM-like restoration for each block  $B_{i_k}$  Eqs. (5)-(6) in the following two ways:
  - (a) *Overlapping Block by Replacement (OBR)-NLM Method*: Replace the intensity values at the overlapping pixels with the most updated values.
  - (b) *Non-Overlapping Block (NOB)-NLM Method*: Blocks do not overlap and therefore each pixel will have only one restoration outcome.
3. Weights are calculated only when mean and variance of intensities of two blocks are similar:

$$w(B_{i_k}, B_j) = \begin{cases} \frac{1}{Z_{i_k}} e^{-\frac{\|u(B_{i_k}) - u(B_j)\|_2^2}{h^2}}, & \text{if } \mu_1 < \frac{\overline{u(B_{i_k})}}{u(B_j)} < \frac{1}{\mu_1} \text{ and } \sigma_1^2 < \frac{Var(u(B_{i_k}))}{Var(u(B_j))} < \frac{1}{\sigma_1^2} \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

for some appropriate parameters  $\mu_1$  and  $\sigma_1$ . Here,  $\overline{u(B_{i_k})}$  and  $Var(u(B_{i_k}))$  represent the mean and the variance of  $u(B_{i_k})$ , respectively.

## 4 Numerical results

In this section, we test our new algorithms, OBR-NLM and NOB-NLM, using two different images: Lenna and an MRI image for a human brain (Brain-MRI). The results are compared to standard NLM and Coupé *et al.*'s Blockwise (CB)-NLM model. The CB-NLM adopted other techniques such as *Automatic Tuning of NLM filter* and *Parallel Computation* [15], to further improve the computational efficiency of the standard NLM. For fair comparison with our new methods, we only apply the selective weight computation Eq. (7) to all blockwise NLM methods. All algorithms presented in this section are implemented in Matlab using an Intel Core i7 CPU 1.60 GHz Desktop Processor.

The peak-signal-to-noise ratio (PSNR) is defined as follows to measure the accuracy of the methods:

$$PSNR = 10 \log_{10} \left( \frac{\sum_{ij} 255^2}{\sum_{ij} (g_{ij} - u_{ij})^2} \right).$$

Here,  $g$  is a noise free clean image and  $u$  is the restored image obtained by denoising a noisy version of  $g$ . For a pre-selection of weights as described in Eq. (7), the optimal values of

**Table 1.** Comparison table for computation time and PSNR using different methods applied to Lenna and Brain-MRI images.

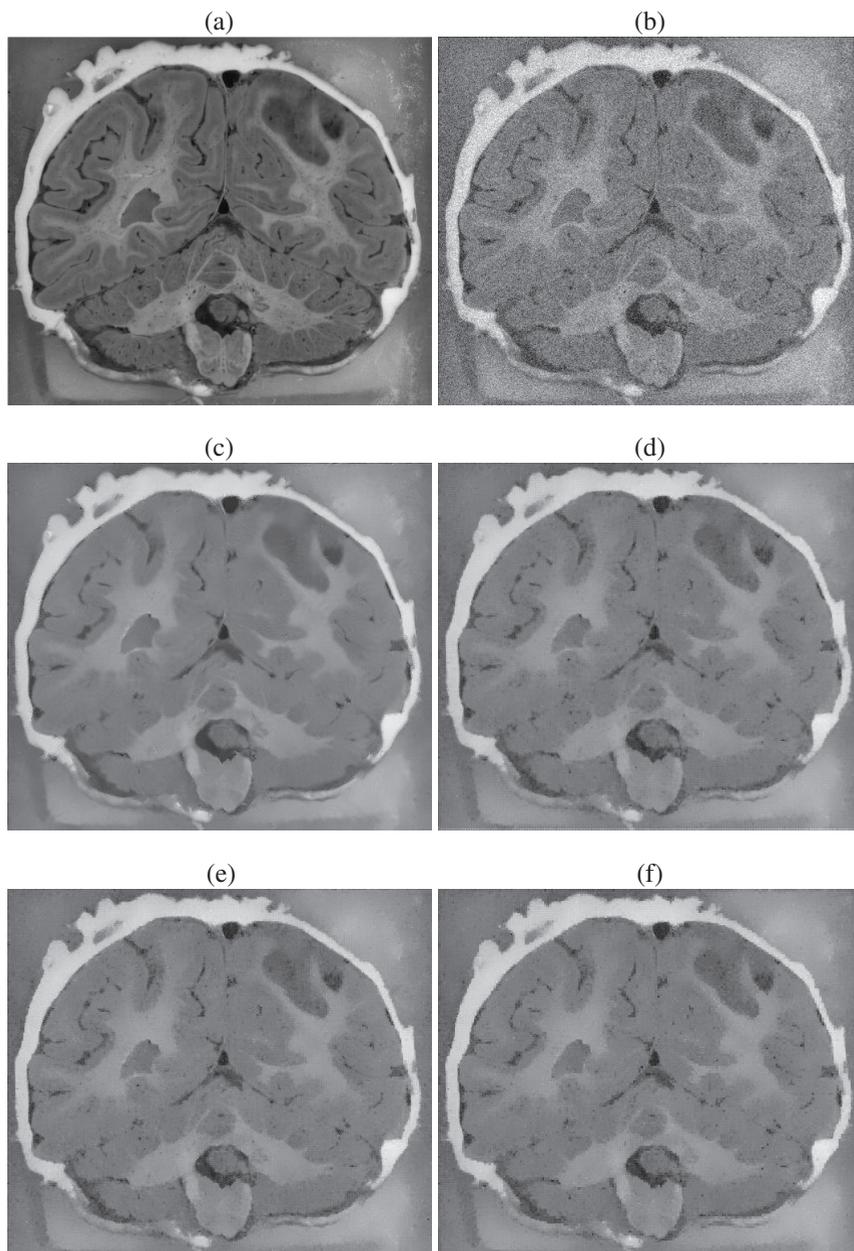
	Lenna		Brain-MRI	
	Time (s)	PSNR	Time (s)	PSNR
Noisy		24.6160		20.2441
NLM	90.3650	30.8132	366.8318	29.7496
CB-NLM	3.0015	29.4870	12.8395	28.1306
OBR-NLM	2.7780	28.6876	10.5851	27.6968
NOB-NLM	1.8350	28.4962	6.8082	27.4672



**Figure 1.** Lenna: (a) Original, (b) Noisy (with 15% of Gaussian noise), (c) NLM, (d) CB-NLM, (e) OBR-NLM, and (f) NOB-NLM. The size of image is  $256 \times 256$  pixels.

the parameters  $\mu_1$  and  $\sigma_1^2$  were chosen experimentally in [15] and they were 0.95 and 0.5, respectively. We follow their suggestions but slightly modify the value of  $\mu_1$  to be 0.8 to obtain the best results. For Lenna, methods were tested on the image where 15% of Gaussian noise was synthetically added to the clean image. For Brain-MRI, we added 25% of Gaussian noise, more noise than Lenna, to mimic the case when medical images have considerable amount of noise and artifacts. For blockwise algorithms, the smoothing parameter  $\hat{h}$  was set to 100 for Lenna and 500 for Brain-MRI. The smoothing parameter  $h$  for the standard NLM was chosen as 12 for Lenna and 15 for Brain-MRI.

As one can see in Table 1, the new algorithms OBR-NLM and NOB-NLM both outperform the standard NLM and CB-NLM in terms of computational efficiency. The visual verification in Figure 1 and Figure 2, and the PSNR values in Table 1 show that the new methods still do not lose much accuracy. One can notice that the PSNR values for Brain-MRI in Table 1 were greatly improved although the image was quite seriously contaminated by 25% of Gaussian



**Figure 2.** Brain-MRI: (a) Original, (b) Noisy (with 25% of Gaussian noise), (c) NLM, (d) CB-NLM, (e) OBR-NLM, and (f) NOB-NLM. The size of image is  $528 \times 455$  pixels.

noise. It can also be seen in Table 1 that the computation time was significantly reduced for the blockwise NLM methods compared to the standard NLM method. The Coupé *et al.*'s block NLM method (CB-NLM) is about 30 times faster than the standard NLM. It has been even further reduced by almost half using the new NOB-NLM method. This could be significant in biomedical applications since those images are often in larger sizes that require a huge amount of computation.

## 5 Conclusions

In this article, the authors have studied efficient and robust filtering based methods to deal with biomedical images, which have complex noise and unclear edges. The computational efficiency of the methods for denoising biomedical images needs to be optimal to be applied to a larger size images, such as 3D scan of MRI or 2D stacks of X-ray images. The blockwise algorithm developed by Coupé *et al.* reduced the computational burden of the standard NLM method but it was further enhanced by the new algorithms. Numerical examples show that the new methods outperform both the standard NLM and Coupé *et al.*'s blockwise NLM method on their computational efficiency without losing much accuracy.

## References

- [1] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and variational problems," *Commun. Pure Appl. Math.*, **vol. 42**, pp. 577–685, (1989)
- [2] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, **vol. 12**, no. 7, pp. 629–639, (1990)
- [3] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, **vol. 60**, pp. 259–268, (1992)
- [4] F. Catté, P.-L. Lions, J.-M. Morel, and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," *SIAM J. Numer. Anal.*, **vol. 29**, no. 1, pp. 182–193, (1992)
- [5] L. Alvarez, P.-L. Lions, and J.-M. Morel, "Image selective smoothing and edge detection by nonlinear diffusion. ii," *SIAM J. Numer. Anal.*, **vol. 29**, no. 3, pp. 845–866, (1992)
- [6] A. Marquina and S. Osher, "Explicit algorithms for a new time dependent model based on level set motion for nonlinear deblurring and noise removal," *SIAM J. Sci. Comput*, **vol. 22**, pp. 387–405, (1999)
- [7] L. Yaroslavsky. *Digital Picture Processing - An Introduction* (Springer Verlag, 1985)
- [8] S. M. Smith and J. M. Brady, "SUSAN - A new approach to low level image processing," *Int. J. Comput. Vis.*, **vol. 23**, no. 1, pp. 45–78, (1997)
- [9] J. van de Weijer and R. van den Boomgaard, "Local mode filtering," in IEEE Computer Soc. Conf. Computer Vis. Pattern Recognit., Kauai, HI, Dec. 2001, pp. 428–433.
- [10] M. Elad, "On the origin of the bilateral filter and ways to improve it," *IEEE Trans. Image Process.*, **vol. 11**, no. 10, pp. 1141–1151, (2002)
- [11] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simul.*, **vol. 4**, no. 2, pp. 490-530 (2005)
- [12] S. Kindermann, S. Osher, and P. W. Jones, "Deblurring and denoising of images by nonlocal functionals," *Multiscale Model. Simul.*, **vol. 4**, no. 4, pp. 1091–1115, (2005)
- [13] G. Gilboa and S. Osher, "Nonlocal operators with applications to image processing." *Multiscale Model. Simul.*, **vol. 7**, no. 3, pp. 1005–1028, (2008)
- [14] A. B. Misra and H. Lim, "Nonlocal speckle denoising model based on non-linear partial differential equations," *Information Systems Design and Intelligent Applications*, pp 165-176, (2015)
- [15] 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 Trans. Med Imaging*, **vol. 27**, no. 4, pp. 425-441, (2008)
- [16] M. Mahmoudi and G. Sapiro, "Fast image and video denoising via nonlocal means of similar neighborhoods," *IEEE Signal Process. Lett.*, **vol. 12**, no. 12, pp. 839–842, (2005)