

# Performance Evaluation of Non-local Means (NLM) Algorithms for Grayscale Image Denoising

Fabrizio Russo

*Department of Engineering and Architecture  
University of Trieste  
Trieste, Italy  
rusfab@univ.trieste.it*

**Abstract**— Non-local Means (NLM) algorithms are state of the art nonlinear techniques for image denoising. Although computationally intensive, NLM methods have gained growing attention due to their interesting combination of noise removal and detail preservation. This paper aims at providing a quantitative evaluation of these key filtering features without resorting to existing metrics. For the first time, the true values of filtering distortion and residual noise are formally obtained from NLM theory. Many computer simulations dealing with test images corrupted by Gaussian noise are reported in the paper in order to study how the filtering errors depend upon different parameter settings. Error maps are also provided in order to show the exact location and nature of filtering effects at the pixel level.

**Keywords**— *image denoising, image filtering, image analysis, nonlinear filters, Gaussian noise.*

## I. INTRODUCTION

Nonlinear filters have become a powerful and widespread used resource for digital image denoising due to their ability to distinguish between noise (to be removed) and image details (to be preserved) [1]. In this framework, non-local means (NLM) filters [2] represent one of the most effective and attractive approaches to noise cancellation in grayscale digital images. Unlike other filtering technique that exploit only the local information in the neighborhood of the pixel to be processed, NLM methods take into account the spatial correlation in the whole image. As a result, NLM filters show superior performance over other denoising operators and are adopted in a growing number of research and application fields. Many different implementations have been proposed in order to increase the computational efficiency [3-7] or to improve the filtering performance with respect to the basic NLM algorithm [8-12]. Clearly, applications areas where the information lost during noise removal is a very critical issue (such as medical and forensic imaging) require a knowledge of the filtering behavior that goes beyond subjective appearance [13]. A quantitative evaluation of the filtering features, however, is not a trivial task. Any filtered image is affected by two main kinds of errors: residual noise (RN) due to insufficient filtering and collateral distortion (CD) caused

by excessive (or wrong) filtering that corrupts the information embedded in the image data. In order to estimate these errors, a variety of full-reference metrics have been recently proposed [14-17]. Many of them belong to the class of vector metrics. In this approach a vector error is computed whose components estimate the amounts of residual noise and collateral distortion produced by the filtering. Vector techniques overcome the limitations of classical scalar metrics, such as the mean squared error (MSE) and the peak signal-to-noise ratio (PSNR) that cannot discriminate between noise cancellation and detail preservation. Vector metrics also overcome the limitations of scalar methods that try to mimic the human perception: as shown in [15], these techniques can be insensitive to different mixtures of residual noise and detail blur and, in any case, they yield a subjective evaluation of the filtering action [18-19]. However, a common problem with metrics for RN and CD evaluation is their validation. Indeed, the accuracy of these methods has been assessed only in particular cases, where the true values of RN and CD are known and can be used for a comparison [20-21].

In this work we analyze the filtering behavior of NLM algorithms by adopting a different approach. Under the hypothesis of additive noise, we aim at theoretically evaluating the filtering effects for this class of filters and for any image. The error analysis presented in this paper overcomes the limitations of current metrics and, for the first time, reveals the exact amounts of RN and CD occurring in NLM filtering. Many computer simulations are reported in order to show how the filtering features depend upon the various parameter settings and the different amounts of noise corruption. This paper is organized as follows. Section II describes the method for the theoretical evaluation of the filtering errors, Section III shows the results of computer simulations, and, finally, Section IV reports conclusions.

## II. EXACT COMPUTATION OF RN AND CD

In this section, we briefly review the fundamentals of the NLM approach. Then, we show how formal expressions for RN and CD can be obtained from the NLM theory.

### A. The Basic NLM Algorithm

Let us deal with digitized images having  $Q$  gray levels (typically  $Q=256$ ). Let  $r(i,j)$  be the pixel luminance at location  $[i,j]$  in the reference (noise-free) image ( $i=1,\dots,L_1; j=1,\dots,L_2$ ) and let  $c(i,j)=r(i,j)+e(i,j)$  be the pixel luminance at the same location in the noisy picture, where  $e(i,j)$  represents the amount of noise corruption. Formally, the output  $f(i,j)$  of a NLM filter operating into a  $(2M+1)\times(2M+1)$  neighborhood (research window) is yielded by a nonlinear weighted average [2], as follows:

$$f(i,j) = \sum_{p=-M}^M \sum_{q=-M}^M w(i,j,p,q) c(i-p,j-q) \quad (1)$$

where  $w(i,j,p,q)$  is a weight that depends upon the similarity between the corrupted pixels  $c(i,j)$  and  $c(i-p,j-q)$ . The set of weights satisfies the following conditions:

$$0 \leq w(i,j,p,q) \leq 1 \quad (2)$$

$$\sum_{p=-M}^M \sum_{q=-M}^M w(i,j,p,q) = 1 \quad (3)$$

Unlike other approaches, the similarity between the pixels  $c(i,j)$  and  $c(i-p,j-q)$  is determined by the similarity of their neighborhoods. Let  $d(i,j,p,q)$  be the weighted Euclidean distance (in terms of gray level intensities) of two  $(2N+1)\times(2N+1)$  square neighborhoods (*comparison windows*) centered on  $c(i,j)$  and  $c(i-p,j-q)$ , respectively:

$$d(i,j,p,q) = \sqrt{\sum_{u=-N}^N \sum_{v=-N}^N G_a(u,v) [c(i-p-u,j-q-v) - c(i-u,j-v)]^2} \quad (4)$$

where  $G_a(u,v)$  is a Gaussian kernel with standard deviation  $a$  ( $a>0$ ). Thus, the weights are defined as follows:

$$w(i,j,p,q) = \frac{1}{Z(i,j)} e^{-\frac{[d(i,j,p,q)]^2}{h^2}} \quad (5)$$

$$Z(i,j) = \sum_{p=-M}^M \sum_{q=-M}^M e^{-\frac{[d(i,j,p,q)]^2}{h^2}} \quad (6)$$

where  $Z(i,j)$  is a normalizing term (according to (3)) and  $h$  is the main parameter that controls the smoothing.

### B. Theoretical Evaluation of RN and CD

Once the mathematical relationships defining NLM filtering are available, the formal expressions for residual noise and collateral distortion can be easily obtained. Remembering that

$c(i,j)=r(i,j)+e(i,j)$ , we can rewrite eq.(1) in order to highlight the presence of two different filtering actions  $f^+(i,j)$  and  $f^-(i,j)$ :

$$f(i,j) = f^+(i,j) + f^-(i,j) \quad (7)$$

$$f^+(i,j) = \sum_{p=-M}^M \sum_{q=-M}^M w(i,j,p,q) e(i-p,j-q) \quad (8)$$

$$f^-(i,j) = \sum_{p=-M}^M \sum_{q=-M}^M w(i,j,p,q) r(i-p,j-q) \quad (9)$$

The term  $f^+(i,j)$  denotes the desired filtering (*positive action*) that aims at reducing the noise  $e(i,j)$ , i.e., the initial error affecting  $r(i,j)$ . Conversely,  $f^-(i,j)$  represents the unwanted filtering (*negative action*) that corrupts the original information constituted by  $r(i,j)$ . According to (7-9), the filtering error  $E(i,j)=f(i,j)-r(i,j)$  can be expressed as follows:

$$E(i,j) = E^+(i,j) + E^-(i,j) \quad (10)$$

$$E^+(i,j) = f^+(i,j) \quad (11)$$

$$E^-(i,j) = f^-(i,j) - r(i,j) \quad (12)$$

$E^+(i,j)$  and  $E^-(i,j)$  are responsible for the generation of residual noise and collateral distortion, respectively. The actual errors, however, depends on the possible compensation of these actions. In order to evaluate the resulting effects, we shall decompose the absolute error  $AE(i,j)$  as follows:

$$AE(i,j) = |E(i,j)| = AE_{RN}(i,j) + AE_{CD}(i,j) \quad (13)$$

where  $AE_{RN}(i,j)$  and  $AE_{CD}(i,j)$  are the absolute error components representing the resulting residual noise and collateral distortion, respectively ( $AE_{RN}(i,j) \geq 0$ ,  $AE_{CD}(i,j) \geq 0$ ). In order to evaluate  $AE_{RN}(i,j)$  and  $AE_{CD}(i,j)$ , the following cases should be considered.

- If  $E^+(i,j)=0$  and  $E^-(i,j)=0$  then  $AE_{RN}(i,j)=0$  and  $AE_{CD}(i,j)=0$  (no error occurs).
- If  $E^+(i,j) \neq 0$  and  $E^-(i,j)=0$  then  $AE_{RN}(i,j)=AE(i,j)$  and  $AE_{CD}(i,j)=0$ .
- If  $E^+(i,j)=0$  and  $E^-(i,j) \neq 0$  then  $AE_{RN}(i,j)=0$  and  $AE_{CD}(i,j)=AE(i,j)$ .
- If  $E^+(i,j)>0$  and  $E^-(i,j)>0$  then  $AE_{RN}(i,j)=E^+(i,j)$  and  $AE_{CD}(i,j)=E^-(i,j)$ .
- If  $E^+(i,j)<0$  and  $E^-(i,j)<0$  then  $AE_{RN}(i,j)=-E^+(i,j)$  and  $AE_{CD}(i,j)=-E^-(i,j)$ .
- If  $E^+(i,j)>0$  and  $E^-(i,j)<0$  and  $|E^+(i,j)| > |E^-(i,j)|$  then  $AE_{RN}(i,j)=AE(i,j)$  and  $AE_{CD}(i,j)=0$ .
- If  $E^+(i,j)>0$  and  $E^-(i,j)<0$  and  $|E^+(i,j)| < |E^-(i,j)|$  then  $AE_{RN}(i,j)=0$  and  $AE_{CD}(i,j)=AE(i,j)$ .

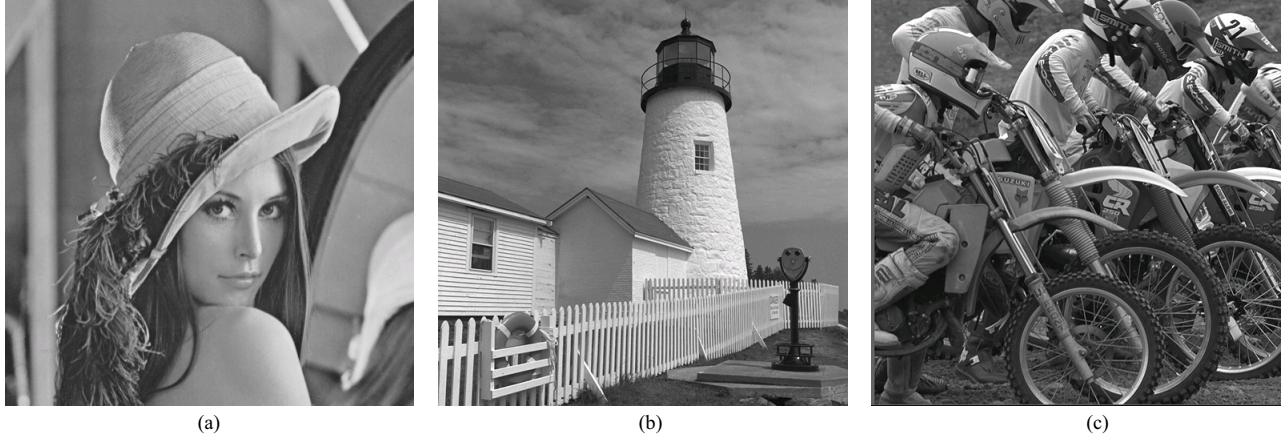


Fig.1. Test images: (a) “Lena”, (b) “Lighthouse”, (c) “Motorbike”.

- h) If  $E^+(i,j) > 0$  and  $E^-(i,j) < 0$  and  $|E^+(i,j)| = |E^-(i,j)|$  then  $AE_{RN}(i,j) = 0$  and  $AE_{CD}(i,j) = 0$ .
- i) If  $E^+(i,j) < 0$  and  $E^-(i,j) > 0$  and  $|E^+(i,j)| > |E^-(i,j)|$  then  $AE_{RN}(i,j) = AE(i,j)$  and  $AE_{CD}(i,j) = 0$ .
- j) If  $E^+(i,j) < 0$  and  $E^-(i,j) > 0$  and  $|E^+(i,j)| < |E^-(i,j)|$  then  $AE_{RN}(i,j) = 0$  and  $AE_{CD}(i,j) = AE(i,j)$ .
- k) If  $E^+(i,j) < 0$  and  $E^-(i,j) > 0$  and  $|E^+(i,j)| = |E^-(i,j)|$  then  $AE_{RN}(i,j) = 0$  and  $AE_{CD}(i,j) = 0$ .

Once  $AE_{RN}(i,j)$  and  $AE_{CD}(i,j)$  are known for each pixel, we can evaluate the residual noise and collateral distortion on the entire image in terms of *mean absolute errors*  $MAE_{RN}$  and  $MAE_{CD}$ , as follows:

$$MAE_{RN} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE_{RN}(i,j) \quad (14)$$

$$MAE_{CD} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE_{CD}(i,j) \quad (15)$$

Clearly, we have:

$$MAE = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE(i,j) = MAE_{RN} + MAE_{CD} \quad (16)$$

It is worth pointing out that the correct evaluation of residual noise and collateral distortion should be performed in terms of MAE instead of MSE. A two-terms decomposition of the MSE is conceptually inaccurate. Indeed, from eq.(13) we have:

$$\begin{aligned} MSE &= \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} [AE_{RN}(i,j) + AE_{CD}(i,j)]^2 \\ &= MSE_{RN} + MSE_{CD} + MSE_{MIX} \end{aligned} \quad (17)$$

where:

$$MSE_{RN} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} [AE_{RN}(i,j)]^2 \quad (18)$$

$$MSE_{CD} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} [AE_{CD}(i,j)]^2 \quad (19)$$

$$MSE_{MIX} = \frac{2}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} [AE_{RN}(i,j) AE_{CD}(i,j)] \quad (20)$$

Unless  $MSE_{MIX} = 0$ , the MSE cannot be decomposed into two components respectively addressing residual noise and collateral distortion only. Conversely, the MAE can. Its decomposition into  $MAE_{RN}$  and  $MAE_{CD}$  is the correct way to separate (and measure) these key filtering features. In the next section we shall adopt these metrics for an in-depth analysis of the filtering behavior of the NLM method.

### III. PERFORMANCE EVALUATION: RESULTS OF COMPUTER SIMULATIONS

We performed many computer simulations in order to study how  $MAE_{RN}$  and  $MAE_{CD}$  depend upon different parameter settings of NLM filtering. In these experiments, we considered three 512×512 grayscale images: “Lena”, “Lighthouse”, and “Motorbike” (Fig.1). In the first group of tests, we generated three noisy pictures by adding zero-mean Gaussian noise with variance  $\sigma^2 = 200$ . We set  $M=7$  and  $N=3$  because these choices are commonly adopted in the literature. We also chose  $a=2$  as standard deviation of the Gaussian kernel. The values of  $MAE_{RN}$  and  $MAE_{CD}$  that are obtained when the main parameter  $h$  ranges from 10 to 150 are graphically depicted in Fig.2. The values of the overall MAE are also reported for reference. We can see that the  $MAE_{RN}$  correctly decreases as the smoothing parameter  $h$  becomes larger.

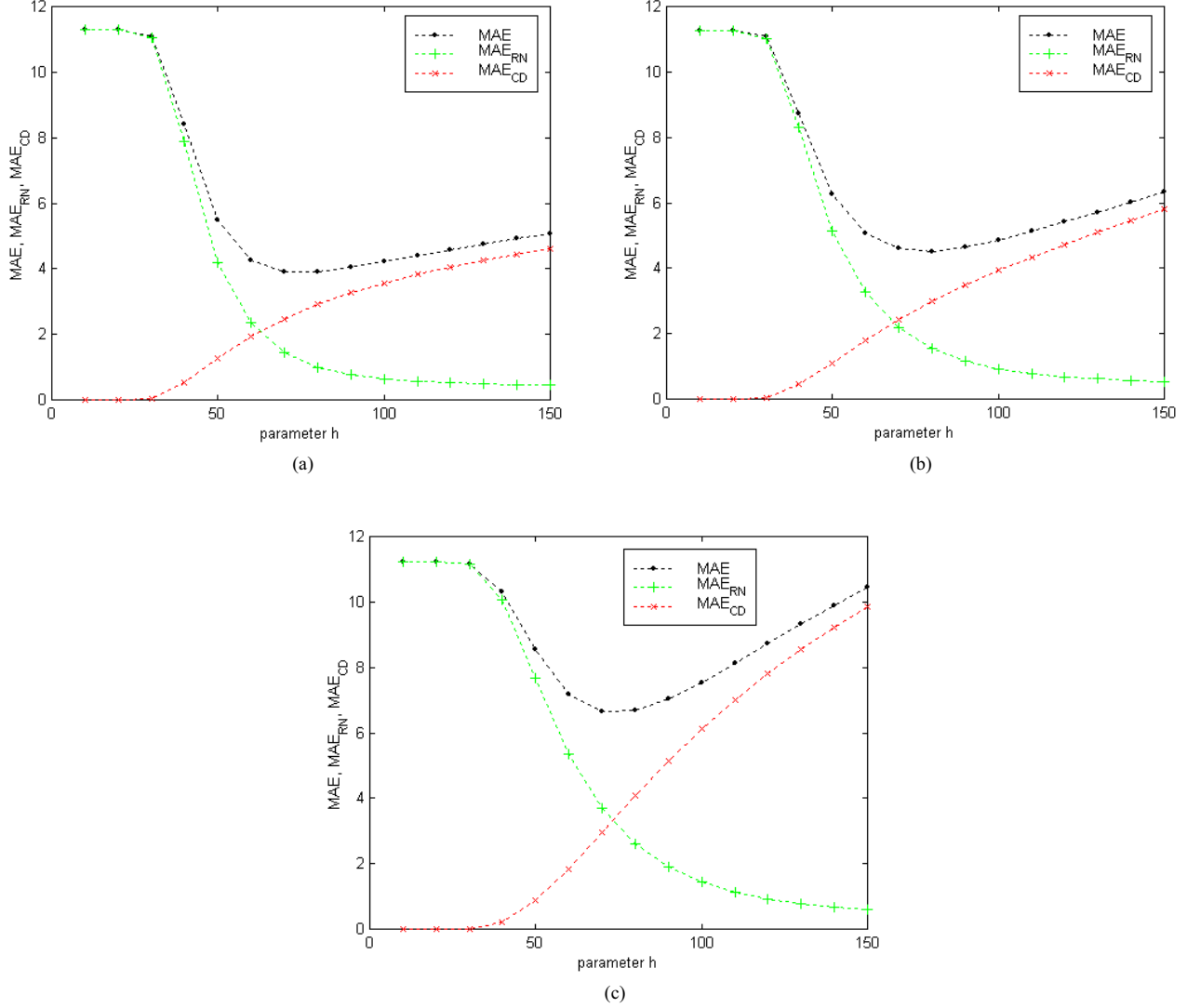


Fig.2. MAE,  $MAE_{RN}$  and  $MAE_{CD}$  evaluations for test pictures corrupted by Gaussian noise with variance  $\sigma^2=200$  and processed by NLM filtering ( $M=7$ ,  $N=3$ ,  $a=2$ ,  $10 \leq h \leq 150$ ): (a) “Lena”, (b) “Lighthouse”, (c) “Motorbike”.

On the contrary,  $MAE_{CD}$  increases because a stronger noise cancellation produces a larger collateral distortion. Samples of the processed data are shown in Fig.3 (“Lena”) and Fig.4 (“Motorbike”). Graphical representations of the absolute errors  $AE_{RN}(i,j)$  (green) and  $AE_{CD}(i,j)$  (red) are also reported. We can see that, for  $h < 50$ , the presence of unfiltered noise is apparent. For  $h=50$ , some residual noise is still well perceivable. For  $h=70$ , few pixels are still noisy and their exact locations are revealed by the error map. The price to be paid is a limited amount of collateral distortion that mainly affects small details of the image. For  $h=90$ , almost all the noise has been removed and, as a counterpart, the image contours are significantly blurred. For  $h > 90$ , the distortion becomes very annoying and heavily corrupts the image edges. Indeed, the graphical

representation of  $AE_{CD}(i,j)$  basically becomes a map of lost contours.

In the second group of tests, we increased the amount of Gaussian noise by setting  $\sigma^2=400$ . The corresponding values of MAE,  $MAE_{RN}$  and  $MAE_{CD}$  are shown in Fig.5. With respect to the previous cases, the minimum MAE is reached for larger values of the parameter  $h$ . We can also observe that, when the minimum MAE occurs, the  $MAE_{CD}$  component is generally larger than the corresponding  $MAE_{RN}$ . Until now, visual inspection was the typical (and limited) source of information about residual noise and collateral distortion. Now, the *exact quantitative* evaluations of these important features can be obtained using the proposed method.



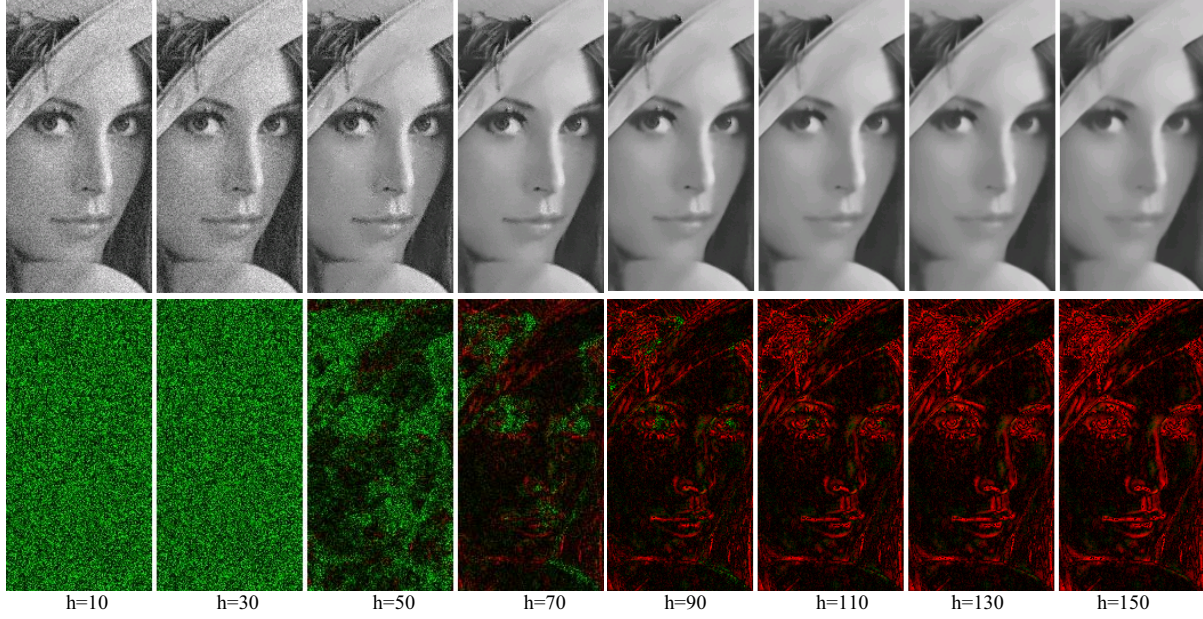


Fig.3. Portions of the “Lena” image corrupted by Gaussian noise ( $\sigma^2=200$ ) and filtered with increasing values of the parameter  $h$ . The corresponding errors maps  $AE_{RN}(i,j)$  (green) and  $AE_{CD}(i,j)$  (red) are also reported.

#### IV. CONCLUSIONS

A new approach to performance evaluation of NLM filters has been presented. Instead of resorting to existing metrics, the formal expressions for residual noise and collateral distortion have been directly derived from NLM theory. As a result, the

true values of  $MAE_{RN}$  and  $MAE_{CD}$  are now available for this class of filters and for any image corrupted by additive noise. Computer simulations have shown how residual noise and collateral distortion actually depend on the parameter settings. Furthermore the reported error maps have provided very accurate information about kind and spatial location of the

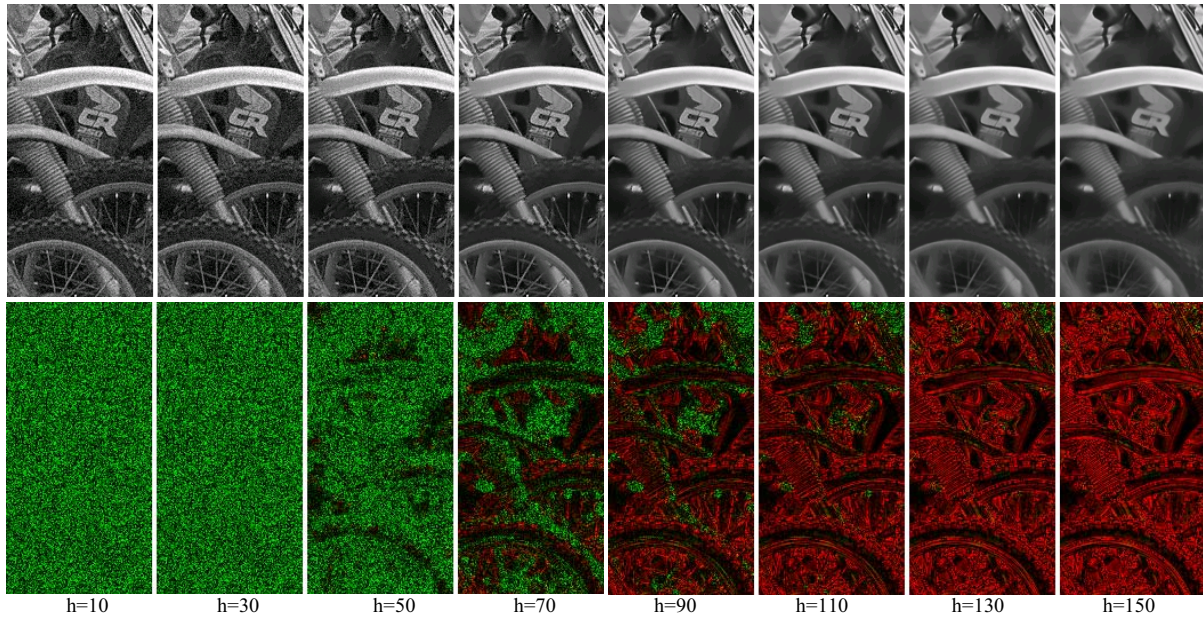


Fig.4. Portions of the “Motorbike” image corrupted by Gaussian noise ( $\sigma^2=200$ ) and filtered with increasing values of the parameter  $h$ . The corresponding errors maps  $AE_{RN}(i,j)$  (green) and  $AE_{CD}(i,j)$  (red) are also reported.

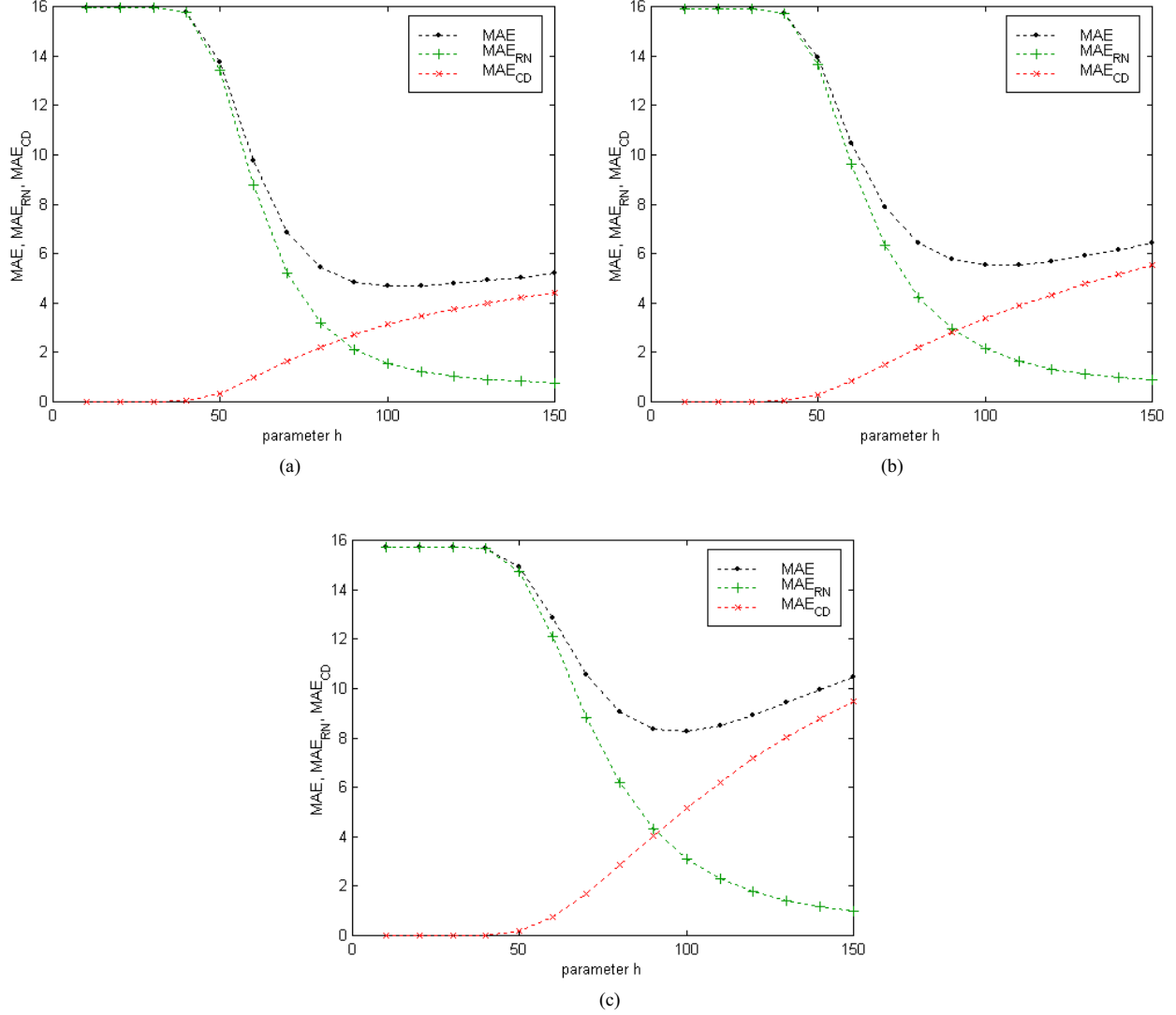


Fig.5. MAE, MAE<sub>RN</sub> and MAE<sub>CD</sub> evaluations for test pictures corrupted by Gaussian noise with variance  $\sigma^2=400$  and processed by NLM filtering ( $M=7$ ,  $N=3$ ,  $a=2$ ,  $10 \leq h \leq 150$ ): (a) “Lena”, (b) “Lighthouse”, (c) “Motorbike”.

filtering errors. Very likely, this information could suggest where the smoothing action needs to be corrected and then possibly lead to more effective generations of NLM methods.

#### REFERENCES

- [1] L. Shao, R. Yan, X. Li, and Y. Liu, “From heuristic optimisation to dictionary learning: a review and comprehensive comparison of image denoising algorithms,” *IEEE Trans. Cybern.*, vol. 44, no. 7, Jul. 2014, pp.1001–1013.
- [2] A. Buades, B. Coll, and J. M. Morel, “A review of image denoising algorithms, with a new one”, *Multiscale Model. Simul.*, vol.4, n.2, 2005, pp.490-530.
- [3] M. Mahmoudi and G. Sapiro, “Fast image and video denoising via nonlocal means of similar neighborhoods”, *IEEE Signal Process. Lett.*, Vol.12, No.1, December 2005, pp.839-842.
- [4] J. Wang, Y. Guo, Y. Ying, Y. Liu, and Q. Peng, “Fast non-local algorithm for image denoising”, 2006 IEEE International Conference on Image Processing, ICIP 2006, Atlanta, Georgia, USA, October 8-11, 2006, pp.1429-1432.
- [5] 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. on Medical Imaging*, vol.27, n.4, 2008, pp.425-441.
- [6] S. H. Chan, T. Zickler and Y. M. Lu, “Fast non-local filtering by random sampling: it works, especially for large images”, 2013 IEEE Int. Conf. on Acoustic, Speech and Sig. Process. (ICASSP 2013), Vancouver, BC, Canada, May 26-31, 2013, pp.1603-1607.
- [7] H. Bhujle, S. Chaudhuri, “Novel speed-up strategies for non-local means denoising with patch and edge patch based dictionaries”, *IEEE Trans. Image Process*, vol.23, n.1, 2014, pp. 356–365.
- [8] H. Zhong, C. Yang, and X. Zhang, “A new weight for nonlocal means denoising using method noise”, *IEEE Signal Process. Lett.*, vol. 19, n.8, Aug. 2012, pp.535–538.

- [9] J. Salmon, "On two parameters for denoising with non-local means," *IEEE Signal Process. Lett.*, vol.17, no.3, Mar. 2010, pp.269–272.
- [10] Lu Lu, Weiqi Jin, Xia Wang, "Non-local means image denoising with a soft threshold", *IEEE Signal Process. Lett.*, vol.22, n.7, 2015, pp.833-837.
- [11] V. May, Y. Keller, N. Sharon, and Y. Shkolnisky, "An Algorithm for Improving Non-Local Means Operators via Low-Rank Approximation", *IEEE Trans. Image Process.*, vol.25, n.3, 2016, pp.1340-1353.
- [12] V. Fedorov and C. Ballester, "Affine non-local means image denoising", *IEEE Trans. Image Process.*, vol. 26, n.5, 2017, pp.2137-2148.
- [13] N.A. Thacker, J.V. Manjon and P.A. Bromiley, "A Statistical Interpretation of Non-Local Means", 5-th IET International Conference on Visual Information Engineering, VIE 2008, Xi'an, China, July 29-August 1, 2008, pp.250-255.
- [14] A. De Angelis, A. Moschitta, F. Russo and P. Carbone, "A vector approach for image quality assessment and some metrological considerations", *IEEE Trans. on Instrum. and Meas.*, Vol.58, No.1, January 2009, pp.14-25.
- [15] F. Russo, "New method for performance evaluation of grayscale image denoising filters", *IEEE Signal Process. Lett.*, vol.17, n.5, 2010, pp.417-420.
- [16] F. Russo, "Validation of Denoising Algorithms for Medical Imaging", in "Advances in Biomedical Sensing, Measurements, Instrumentation and Systems", S.C. Mukhopadhyay and A. Lay-Ekuakille (Eds.), Springer-Verlag, 2010, pp. 93-105.
- [17] F. Russo, "New Vector Method for Quality Assessment in Image Denoising", 3rd International Conference on Circuits, Systems, Communications, Computers and Applications (CSCCA '14), Florence, Italy, November 22-24, 2014, pp.197-206.
- [18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", *IEEE Transactions on Image Processing*, vol.13, n.4, 2004, pp.600-612.
- [19] S. Winkler and P. Mohandas, "The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics", *IEEE Transactions on Broadcasting*, Vol.54, n.3, 2008, pp.660– 668.
- [20] F. Russo, "On the Accuracy of Vector Metrics for Quality Assessment in Image Filtering", in *Proc. 20th 2014 IMEKO TC4 International Symposium*, Benevento, Italy, September 15-17, 2014, pp.863-868.
- [21] F. Russo, "New Method for Measuring the Detail Preservation of Noise Removal Techniques in Digital Images", *WSEAS Transactions on Signal Processing*, vol.11, 2015, pp. 317-327.