A Method for Performance Analysis of Grayscale Image Denoising Techniques Based on the Wavelet Transform

Fabrizio Russo

***Abstract*—** Wavelet transform-based filters are widely adopted for noise removal from grayscale digital images because these techniques can effectively combine cancellation of noise and preservation of image details. The aim of this paper is to provide accurate quantitative evaluations of these key filtering features without the limitations (and the errors) of current metrics. For the first time, the exact amounts of filtering distortion and unfiltered noise produced by a wavelet-based denoising filter are formally computed resorting to the filter theory only. Computer simulations are reported in the paper in order to show how residual noise and filtering distortion affect the results at the pixel level. Comparisons with current metrics are also provided.

***Keywords*—** image denoising, wavelet transform, wavelet shrinkage, image quality assessment, Gaussian noise.

I. Introduction

Denoising algorithms are of paramount importance in digital imaging because noise can significantly degrade essential information that is embedded in the image data [1]. In this framework, wavelet transform-based filters are one of the most powerful and widely adopted approaches to noise removal especially in the field of medical imaging [2-3]. Indeed, these techniques can realize an effective trade-off between noise reduction and preservation of image details. As an example, medical applications encompass (but are not limited to) digital mammography [4], ultrasound imaging [5], computed tomography (CT) [6-7] and magnetic resonance imaging (MRI) [8-9]. The basic principle of wavelet denoising consists in performing the wavelet transform of the noisy data, thresholding the wavelet coefficients (*wavelet shrinkage*) and finally inverting the wavelet transform to obtain a denoised version of the input data. Many different approaches are available in the literature in order to choose wavelet function and shrinkage method [10-17]. The choice of shrinkage method is a very important aspect in wavelet denoising because it has a direct impact on the accuracy of the result in terms of *residual noise* (RN) and *collateral distortion* (CD). The former represents the noise left by insufficient filtering, whereas the latter takes into account detail blur and artifacts produced by excessive (or wrong) smoothing. The information lost during noise cancellation is a very critical issue in many applications. Thus, quantitative evaluations of RN and CD would be of paramount importance for the validation of any new denoising technique. Several full-reference metrics have been proposed for measuring RN and CD in grayscale images. In vector approaches, a vector error is typically evaluated whose components aim at estimating the different amounts of RN and CD left after filtering [18-19]. Vector approaches overcome the drawbacks of classical scalar metrics, such as the mean squared error (MSE) and the peak signal-to-noise ratio (PSNR) that cannot distinguish between noise cancellation and detail preservation. Vector metrics also overcome the limitations of techniques that mimic the human perception [20]. As observed in [21], these metrics can produce the same result for filtered pictures having different combinations of RN and CD. However, a very compelling issue with current metrics yielding separate estimates of RN and CD is to know how accurate these estimates are [22-24].

In this paper we show how a solution to this problem can be found focusing on wavelet denoising theory. For the first time, we shall *theoretically* evaluate the *true values* of RN and CD for this class of filters, without the limitations and the inaccuracies of current scalar and vector metrics. Results of many computer simulations are reported in the paper in order to show how RN and CD depend upon the threshold of the wavelet coefficients and, for a comparison, how erroneous the available metrics are. The exact locations of filtering errors due to RN and CD will be shown too. This paper is organized as follows. Section II describes the theoretical evaluation of RN and CD, Section III presents the results of many computer simulations, and, finally, Section IV reports conclusions.

II. Computing the True Values of RN and CD

Let us deal with digitized images having L gray levels (typically L=256). Let r(i,j) and x(i,j) be the pixel luminances at location [i,j] in the reference (noise-free) and in the noisy pictures, respectively (i=1,…M; j=1,…,N). Regardless of the mechanism of noise generation, let n(i,j) be the amount of noise corruption affecting the pixel at location [i,j]:

$$n(i, j) = x(i, j) - r(i, j) \tag{1}$$

The basic approach to wavelet denoising typically involves three steps: a linear discrete wavelet transformation (DWT), a nonlinear shrinkage of the wavelet coefficients, and a linear

This work was supported by the University of Trieste, Trieste, Italy.

F. Russo is with the Department of Engineering and Architecture, University of Trieste, I-34127 Trieste, Italy (e-mail: rusfab@units.it).

inverse transformation (IDWT). Thus, let us apply a generic DWT to the set of noisy input data $\{x(i,j)\}$ and let $u_x^{(h,k)}(p,q)$ denote the wavelet coefficient at location [p,q] in the k-th subband of the h-th stage. According to (1), the noisy image $\{x(i,j)\}$ can be considered as the sum of two components:

$$\{x(i,j)\} = \{r(i,j)\} + \{n(i,j)\} \tag{2}$$

where $\{r(i,j)\}$ is the reference (noise-free) picture and $\{n(i,j)\}$ briefly denotes the noise. Since the transform is linear, we have:

$$DWT(\{x(i,j)\}) = DWT(\{r(i,j)\}) + DWT(\{n(i,j)\}) \tag{3}$$

The corresponding wavelet coefficients are given by the following relationship:

$$u_x^{(h,k)}(p,q) = u_r^{(h,k)}(p,q) + u_n^{(h,k)}(p,q) \tag{4}$$

Now, let $v_x^{(h,k)}(p,q)$ be the wavelet coefficient modified after shrinkage. Regardless of the specific shrinkage method that is adopted, we shall express such modification as follows:

$$v_x^{(h,k)}(p,q) = w^{(h,k)}(p,q)\ \ u_x^{(h,k)}(p,q) \tag{5}$$

where $w^{(h,k)}(p,q)$ is a weight ranging from zero to unity $(0 \le w^{(h,k)}(p,q) \le 1)$. When $w^{(h,k)}(p,q) = 1$, the wavelet coefficient is passed unchanged. Conversely, when $w^{(h,k)}(p,q) = 0$, the role of the coefficient is cancelled, as typically occurs in hard thresholding. The filtered image $\{f(i,j)\}$ is obtained by performing the IDWT:

$$\{f(i,j)\} = IDWT\left(\{w^{(h,k)}(p,q)\ \ u_x^{(h,k)}(p,q)\}\right) \tag{6}$$

Thus, remembering (4), we have:

$$\{f(i,j)\} = \{f_r(i,j)\} + \{f_n(i,j)\} \tag{7}$$

where:

$$\{f_n(i,j)\} = IDWT\left(\{w^{(h,k)}(p,q)\ \ u_n^{(h,k)}(p,q)\}\right) \tag{8}$$

$$\{f_r(i,j)\} = IDWT\left(\{w^{(h,k)}(p,q)\ \ u_r^{(h,k)}(p,q)\}\right) \tag{9}$$

The term $\{f_n(i,j)\}$ is the result of the filtering action that aims at reducing the noise $\{n(i,j)\}$. On the contrary, $\{f_r(i,j)\}$ represents the unwanted effect that modifies the original information $\{r(i,j)\}$.

The filtering error $E(i,j) = f(i,j) - r(i,j)$ can be expressed as follows:

$$E(i,j) = E_n(i,j) + E_r(i,j) \tag{10}$$

$$E_n(i,j) = f_n(i,j) \tag{11}$$

$$E_r(i,j) = f_r(i,j) - r(i,j) \tag{12}$$

where $E_n(i,j)$ is the (signed) error component dealing with RN and $E_r(i,j)$ is the (signed) error component that generates CD. The *resulting* error components depend on the possible compensation of $E_n(i,j)$ and $E_r(i,j)$. In order to evaluate these components, we shall perform a decomposition of the absolute error $e(i,j)$:

$$e(i,j) = |E(i,j)| = e_{RN}(i,j) + e_{CD}(i,j) \tag{13}$$

where $e_{RN}(i,j)$ and $e_{CD}(i,j)$ clearly denote the absolute error components addressing RN and CD, respectively ($e_{RN}(i,j) \ge 0, e_{CD}(i,j) \ge 0$). The computation of these components depends on the signs and amounts of $E_n(i,j)$ and $E_r(i,j)$, as follows.

1) Let us suppose that $E_{RN}(i,j)$ and $E_{CD}(i,j)$ have the same signs. Since no compensation occurs, we have: $e_{RN}(i,j) = |E_{RN}(i,j)|$, $e_{CD}(i,j) = |E_{CD}(i,j)|$.

2) Let us suppose that $E_{RN}(i,j)$ and $E_{CD}(i,j)$ have different signs and $|E_{RN}(i,j)| \ge |E_{CD}(i,j)|$. In this case, RN prevails: $e_{RN}(i,j) = e(i,j)$, $e_{CD}(i,j) = 0$.

3) Finally, let us suppose that $E_{RN}(i,j)$ and $E_{CD}(i,j)$ have different signs and $|E_{RN}(i,j)| < |E_{CD}(i,j)|$. Since CD prevails, we have: $e_{RN}(i,j) = 0$, $e_{CD}(i,j) = e(i,j)$.

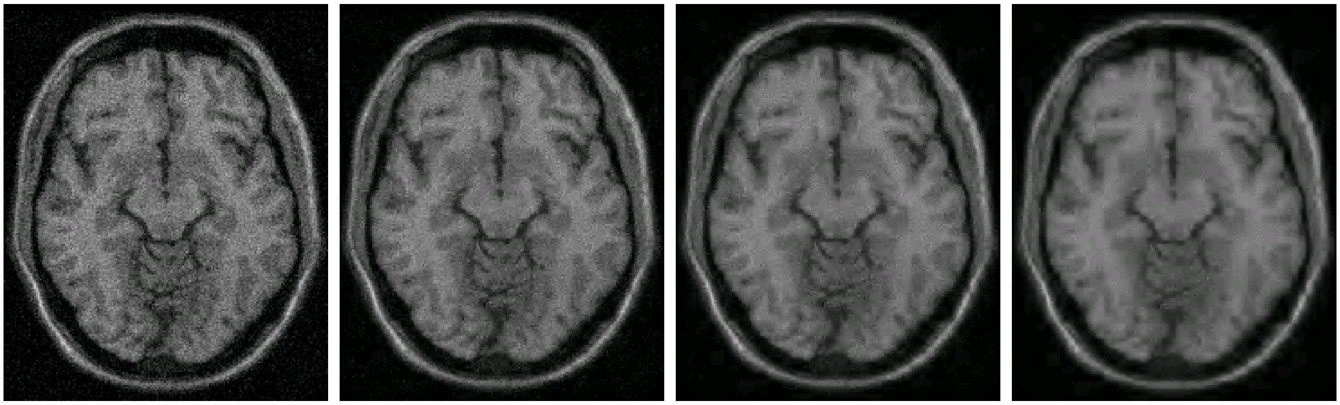
Now, the RN and CD on the entire picture can be computed in terms of *mean absolute errors* MAE_{RN} and MAE_{CD} as follows:

$$MAE = MAE_{RN} + MAE_{CD} \tag{14}$$

$$MAE_{RN} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} e_{RN}(i,j) \tag{15}$$

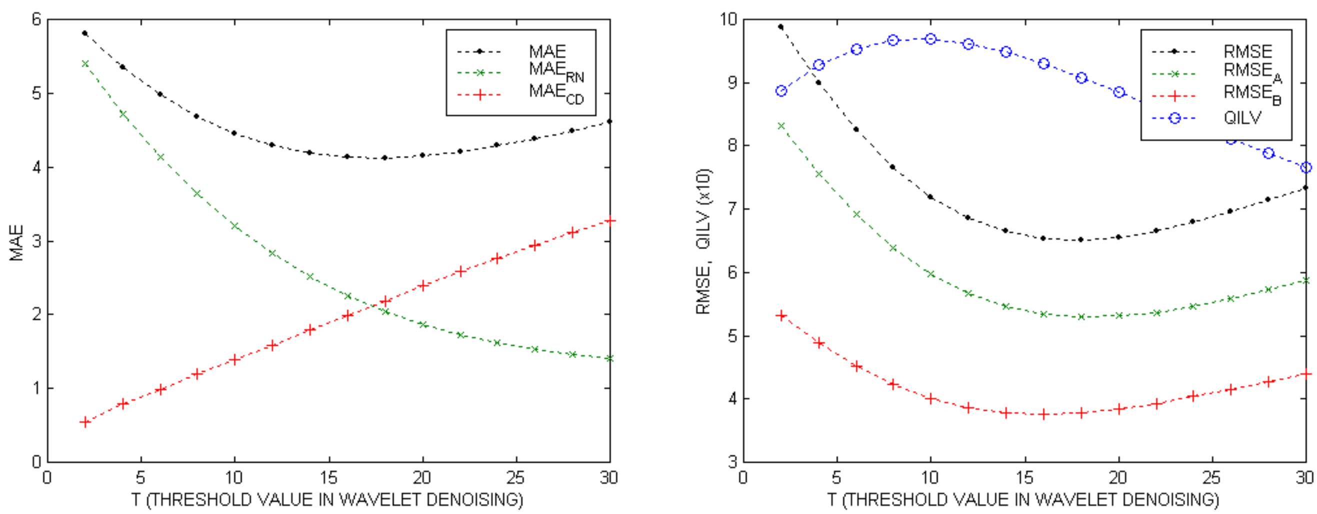
$$MAE_{CD} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} e_{CD}(i,j) \tag{16}$$

We chose the MAE instead of the mean squared error (MSE), because a two-terms RN-CD decomposition of the MSE would be erroneous, according to (13).



(a) (b) (c) (d)

Fig.1 - Simulated BrainWeb data corrupted by Gaussian noise (σ^2=225) and processed by wavelet filtering with different threshold values T: (a) T=2, (b) T=10, (c) T=18, (d) T=28.



(a) (b)

Fig.2 – Results given by the proposed approach and current metrics: (a) MAE, MAE$_{RN}$ and MAE$_{CD}$ evaluations, (b) RMSE, RMSE$_A$, RMSE$_B$ and QILV evaluations ($2 \leq T \leq 30$).

III. Results of Computer simulations

In this section we shall investigate how the nonlinear shrinkage of the wavelet coefficients affects the quality of the result in terms of RN and CD. In all the experiments, we adopted the excellent software package for forward 2-D DWT, soft thresholding and inverse 2-D DWT, available in [25-26]. The first test deals with simulated BrainWeb data [27-29]. We corrupted the original data by adding zero-mean Gaussian noise with standard deviation σ^2=225 and we performed wavelet filtering by choosing increasing values of threshold T. Some results are depicted in Figs.1a (T=2), 1b (T=10), 1c (T=18) and 1d (T=28). From visual inspection, we can easily see that residual noise and detail preservation decrease as the value of T increases. The evaluations of MAE$_{RN}$ and MAE$_{CD}$ that are achieved when the T ranges from 2 to 30 are graphically depicted in Fig.2a. As the threshold T becomes larger, the MAE$_{RN}$ correctly decreases, whereas the MAE$_{CD}$ increases, as it should be. For a comparison, we considered the Quality Index based on Local Variance (QILV) [30] that is often adopted as a measure of detail preservation in medical imaging. If we observe the QILV evaluations in Fig.2b, however, we see that this index wrongly increases for growing values of T in the interval $2 \leq T \leq 8$. We also considered for a comparison vector metrics such as the VRMSE method presented in [18]. In this approach, the RMSE$_A$ and RMSE$_B$ components yield the filtering errors in the uniform and edge regions of the filtered image aiming at estimating residual noise and collateral distortion, respectively. The incorrect behavior of these metrics, however, is apparent (Fig.2b): the RMSE$_A$ should not increase ($20 \leq T \leq 30$) and the RMSE$_B$ should not decrease ($2 \leq T \leq 16$).

We performed a second group of tests using the well-known Shepp-Logan phantom image corrupted by zero-mean Gaussian noise (σ^2=225). The results yielded by wavelet filtering with increasing values of threshold T are shown

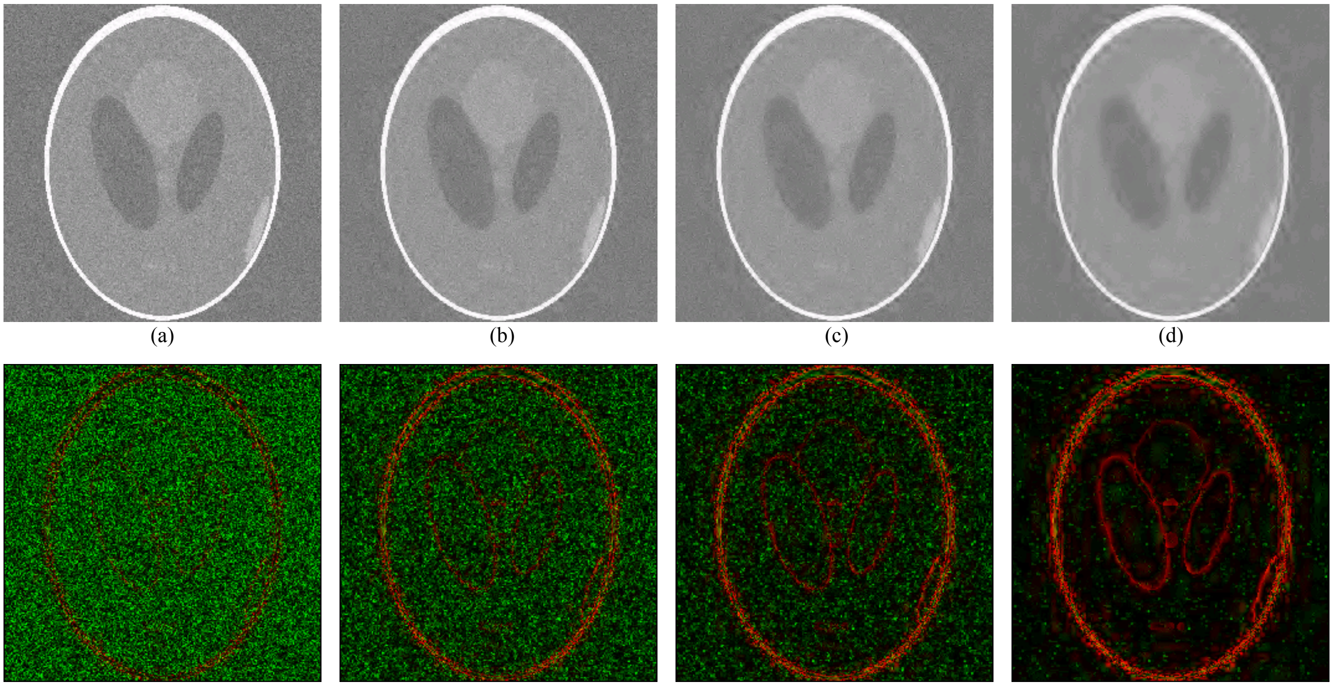
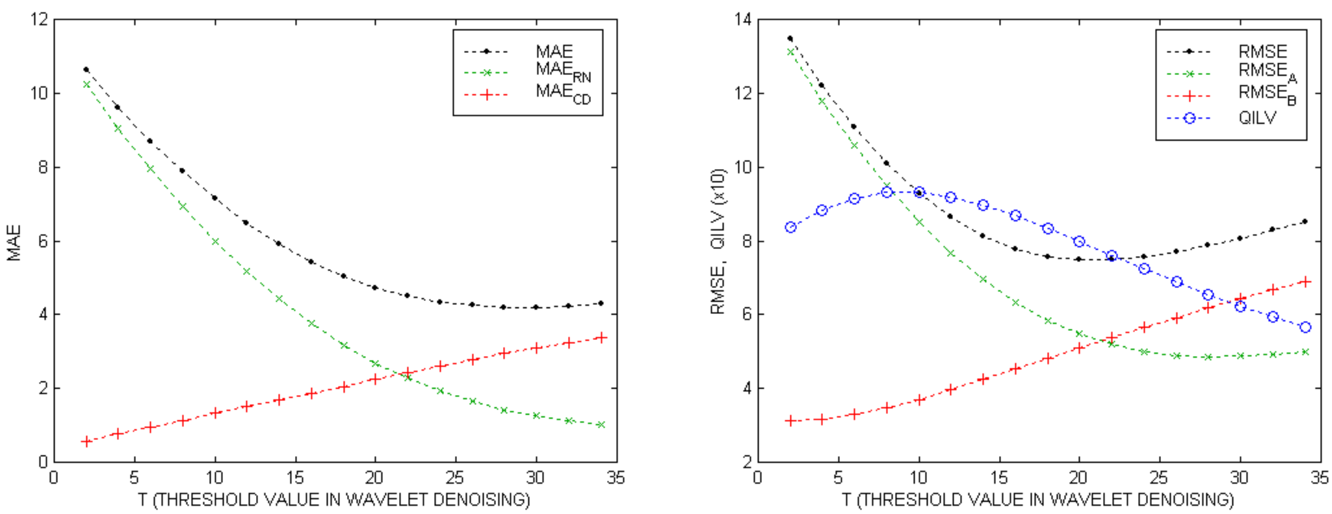


Fig.3 – Shepp-Logan phantom image corrupted by Gaussian noise (σ^2=225) and processed by wavelet filtering with different threshold values T: (a) T=6, (b) T=12, (c) T=18, (d) T=30; (e)-(f)-(g)-(h) corresponding error maps of the true error components $e_{RN}(i.j)$ (green) and $e_{CD}(i.j)$ (red).



(a) (b)

Fig.4 – Results given by the proposed approach and current metrics: (a) MAE, MAE_{RN} and MAE_{CD} evaluations, (b) RMSE, $RMSE_A$, $RMSE_B$ and QILV evaluations ($2 \le T \le 34$).

in Fig.3. The corresponding maps of absolute error components $e_{RN}(i,j)$ (green) and $e_{CD}(i,j)$ (red) are graphically depicted too. For T≤12, we easily see that large amounts of noise affect the filtered data. For T=18, some unfiltered noise is apparent, whereas the presence of collateral distortion becomes well perceivable. For T=30, a small number of pixels is still noisy, whereas the most annoying effect is represented by distortion also including typical wavelet artifacts [31]. The evaluations of MAE_{RN} and MAE_{CD} that are achieved when T ranges from 2 to 34 are graphically depicted in Fig.4a. As in the previous group of experiments, the correct behavior of MAE_{RN} and MAE_{CD} is apparent. For growing values of T, the MAE_{RN} decreases and the MAE_{CD} increases, as it should be. Conversely, the QILV becomes larger for values of T ranging from 2 to 10, and the $RMSE_A$ wrongly increases in the interval 28≤T≤34 (Fig.4b). The proposed approach can yield exact quantitative evaluations of RN and CD, whereas other methods cannot.

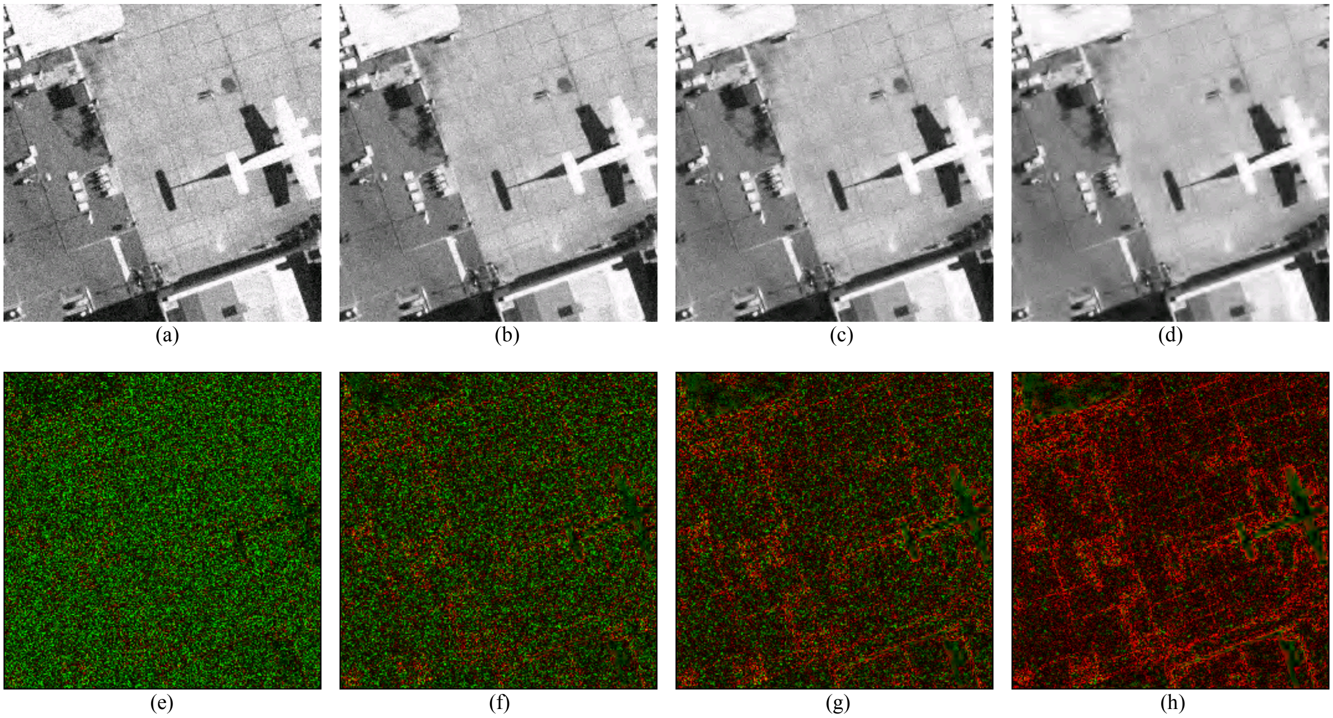


Fig.5 – “Airfield” image corrupted by Gaussian noise (σ^2=300) and processed by wavelet filtering with different threshold values T: (a) T=6, (b) T=12, (c) T=18, (d) T=30; (e)-(f)-(g)-(h) corresponding error maps of the true error components $e_{RN}(i.j)$ (green) and $e_{CD}(i.j)$ (red).

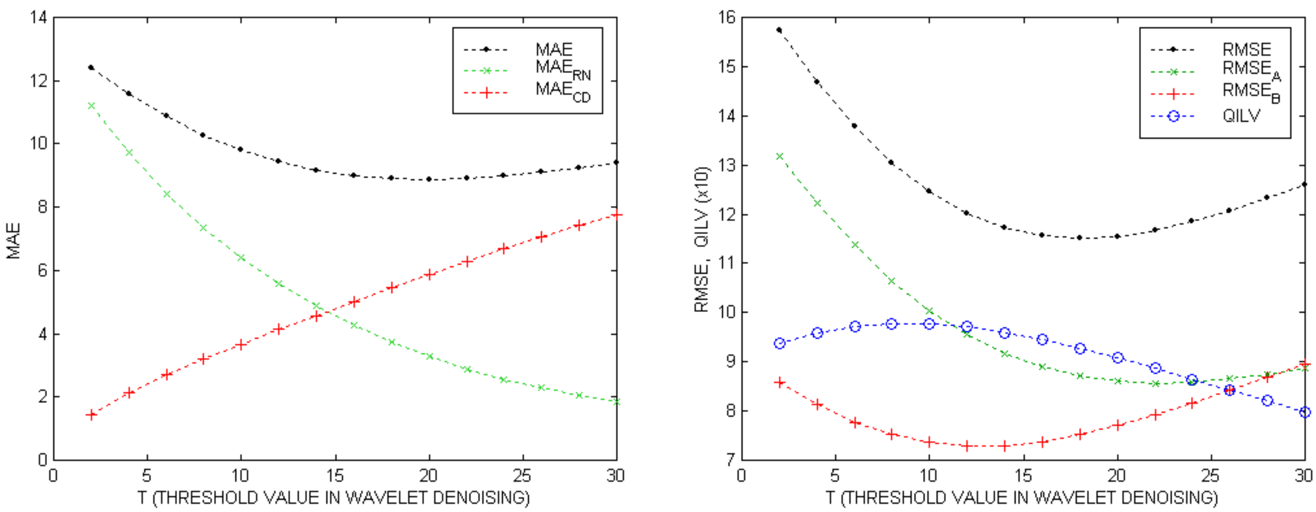


Fig.6 – Results given by the proposed approach and current metrics: (a) MAE, MAE_{RN} and MAE_{CD} evaluations, (b) RMSE, $RMSE_A$, $RMSE_B$ and QILV evaluations $(2 \leq T \leq 30)$.

We performed a third group of tests using the “Airfield” picture corrupted by zero-mean Gaussian noise with variance σ^2=300. The results given by wavelet filtering for growing values of threshold T are shown in Fig.5. The corresponding maps of absolute error components are graphically provided too. The evaluations of MAE_{RN} and MAE_{CD} that are achieved when T ranges from 2 to 30 are graphically represented in Fig.6a, whereas the values of RMSE, $RMSE_A$, $RMSE_B$ and QILV are reported in Fig.6b. As in the previous case, the incorrect behavior of competing metrics is apparent. The QILV aims at measuring the detail preservation: however, it incorrectly increases for growing amounts of smoothing ($2 \leq T \leq 8$). On the other hand, the $RMSE_A$ (measuring the unfiltered noise) wrongly increases in the interval $22 \leq T \leq 30$, and the $RMSE_B$ (measuring filtering distortion) decreases in the interval $2 \leq T \leq 15$.

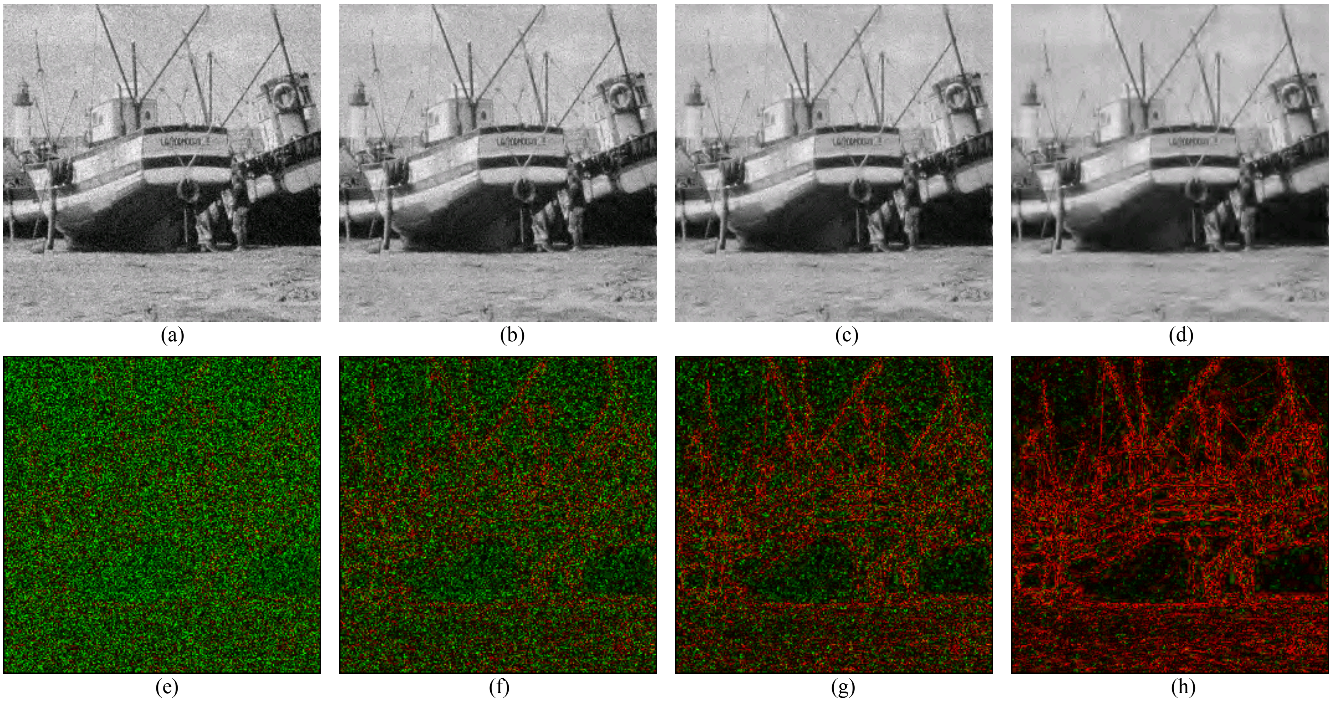


Fig.7 – “Boats” image corrupted by Gaussian noise (σ^2=350) and processed by wavelet filtering with different threshold values T: (a) T=6, (b) T=12, (c) T=18, (d) T=30; (e)-(f)-(g)-(h) corresponding error maps of the true error components $e_{RN}(i.j)$ (green) and $e_{CD}(i.j)$ (red).

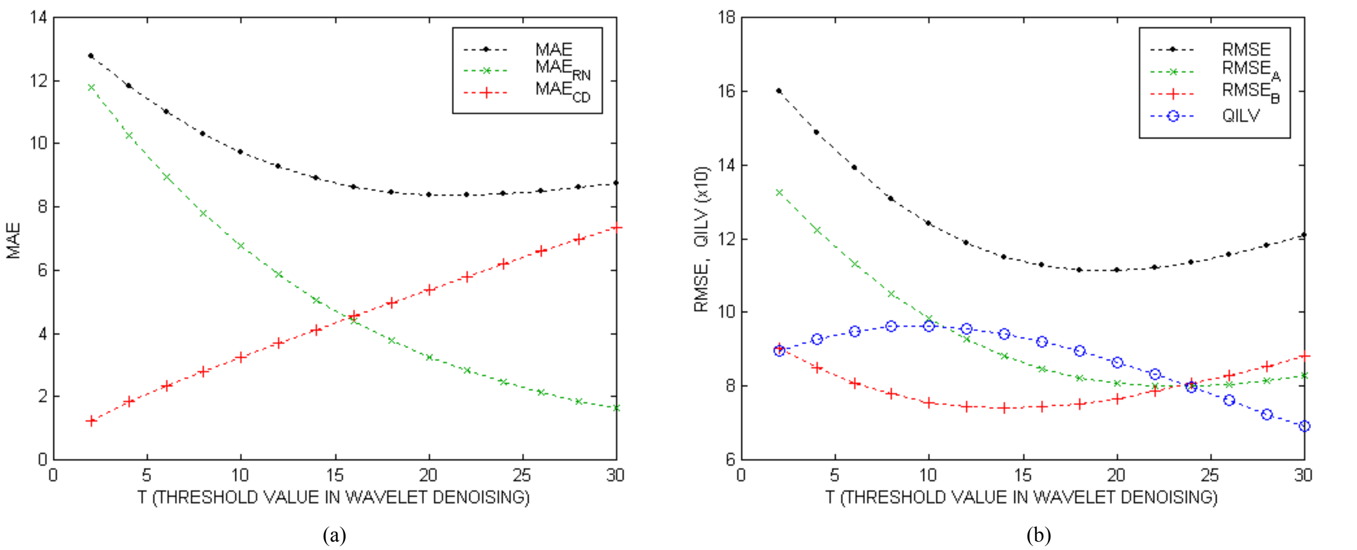


Fig.8 – Results given by the proposed approach and current metrics: (a) MAE, MAE_{RN} and MAE_{CD} evaluations, (b) RMSE, $RMSE_A$, $RMSE_B$ and QILV evaluations $(2 \leq T \leq 30)$.

We performed a fourth group of tests using the “Boats” test image corrupted by zero-mean Gaussian noise with variance σ^2=350. The results given by wavelet filtering are shown in Fig.7. The correct behavior of MAE_{RN} and MAE_{CD} is graphically depicted in Fig.8a, whereas the values of RMSE, $RMSE_A$, $RMSE_B$ and QILV are reported in Fig.8b. The limitations and inaccuracies of these metrics are apparent, as in the previous experiments.

IV. Conclusions

In this paper we have presented a new approach to the investigation of the accuracy of wavelet-based image denoising. We have shown how the formal expressions for key features such as residual noise and collateral distortion can be directly derived from wavelet denoising theory regardless of the specific choice of wavelet function and shrinkage method. Result of computer simulations have shown how exact quantitative evaluations of these important features can now be computed without being impaired by the limitations and inaccuracies of current metrics. It is expected that the availability of more accurate information about the filtering behavior could very likely lead to more powerful classes of wavelet-based filters.

REFERENCES

[1] M. Mafi, H. Martin, M. Cabrerizo, J. Andrian, A. Barreto and M. Adjouadi, "A comprehensive survey on impulse and Gaussian denoising filters for digital images", Signal Processing, vol.157, 2019, pp.236–260.

[2] A. Ouahabi, "A review of wavelet denoising in medical imaging", 2013 8th Int. Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, 2013, pp. 19-26.

[3] D. Cho, *Image Denoising Using Wavelet Transform*, VDM, 2008.

[4] A. Mencattini, M. Salmeri, R. Lojacono, M. Frigerio, and F. Caselli, "Mammographic images enhancement and denoising for breast cancer detection using dyadic wavelet processing transform", IEEE Transactiions on Instrumentation and Measurements, vol.57, n.7, 2008, pp.1422–1430.

[5] G.Andria, F Attivissimo, G. Cavone, A.M.L. Lanzolla, "Selection of wavelet functions and thresholding parameteres in ultrasound image denoising", in Proc. MeMeA 2013, Gatineau, QC, Canada, May 4-5, 2013, pp. 49-52.

[6] A.Borsdorf, R. Raupach, T Flohr, et al., "Wavelet based noise reduction in CT-images using correlation analysis", IEEE Transactions on Medical Imaging, vol.27, n.12, 2008, pp. 1685–1703.

[7] M. Diwakar, M. Kumar, "CT image denoising using NLM and correlation-based wavelet packet thresholding", IET Image Process., vol.12 n.5, 2018, pp.708-715.

[8] A. Meije Wink and J. B. T. M. Roerdink, "Denoising Functional MR Images: A Comparison of Wavelet Denoising and Gaussian Smoothing", IEEE Transactions On Medical Imaging, vol.23, n.3, 2004, pp.374-387.

[9] H. Morsheddost, D. Asemani and N. Mirahadi, "Optimization of MDL-based wavelet denoising for fMRI data analysis", Proc. IEEE 11th International Symposium on Biomedical Imaging (ISBI- 2014), 2014, pp.33-36.

[10] D.L. Donoho, "Denoising by Soft-Thresholding", IEEE Trans. On Inf. Theory, 41, 3, 1995, pp. 613–627.

[11] D.L. Donoho, I.M. Johnstone, "Minimax Estimation via Wavelet Shrinkage", Annals of Statistics 26, 1998, pp. 879-921.

[12] S. Sardy, "Minimax threshold for denoising complex signals with waveshrink," IEEE Transactions on Signal Processing, vol. 48, no. 4, Apr. 2000, pp.1023–1028.

[13] M. A. Figueiredo and R. D. Nowak, "Wavelet-based image estimation: An empirical Bayes approach using Jeffreys's noninformative prior," IEEE Transactions on Image Processing, vol. 10, no. 9, Sep. 2001, pp. 1322–1331.

[14] A. Pizurica, W. Philips, I. Lemahieu and M. Acheroy, "A Versatile Wavelet Domain Noise Filtration Technique for Medical Imaging", IEEE Transactions on Medical Imaging, vol.22, n.3, 2003, pp. 323-331.

[15] A. Dineva, A. R. Várkonyi-Kóczy, J. K. Tar, "Fuzzy expert system for automatic wavelet shrinkage procedure selection for noise suppression", Proc. INES 2014, IEEE 18th Int. Conf. on Intelligent Engineering Systems, July 3-5, Tihany, Hungary, 2014.

[16] A. Mohammad Zaki, S. Ghofrani, "Comparing Threshold Based Denoising Methods in Wavelet Transform Domain", International Journal of Circuits, Systems and Signal Processing, vol.12, 2018, pp. 65-73, ISSN 1998-4464.

[17] Z. Li, J. Wen, Z. Xiao and M. Wang,"Comparative Evaluation of Denoising of Tree Radar B-scan Images using Multi-resolution Analysis Algorithms", International Journal of Circuits, Systems and Signal Processing, Vol.13, 2019, pp. 205-212, ISSN 1998-4464.

[18] 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, n.1, 2009, pp. 14–25.

[19] 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.

[20] 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.

[21] F. Russo, "New method for performance evaluation of grayscale image denoising filters", IEEE Signal Processing Letters, vol.17, n.5, 2010, pp.417-420.

[22] F. Russo, "Study of the Accuracy of the Color Peak Signal-to-Blur Ratio (CPSBR)", International Journal of Circuits, Systems and Signal Processing, Vol.10, 2016, pp. 242-253, ISSN 1998-4464.

[23] F. Russo, "Performance Evaluation of Non-local Means (NLM) Algorithms for Grayscale Image Denoising", Proc. IEEE EECS, Bern, Switzerland, 2018, pp.204-210.

[24] F. Russo, "On the Accuracy of Denoising Algorithms in Medical Imaging: A Case Study", Proc. IEEE MeMeA 2018, Rome, Italy, 2018.

[25] http://taco.poly.edu/WaveletSoftware/

[26] A. Farras Abdelnour, I. W. Selesnick, "Design of 2-band orthogonal near-symmetric CQF", Proc. 2001 IEEE ICASSP, 2001, pp.3693-3696.

[27] http://www.bic.mni.mcgill.ca/brainweb/

[28] C.A. Cocosco, V. Kollokian, R.K.-S. Kwan and A.C. Evans, "BrainWeb: online interface to a 3D MRI simulated brain database", NeuroImage, vol.5, no.4, part 2/4, S425, 1997.

[29] R.K.-S. Kwan, A.C. Evans and G.B. Pike, "MRI simulation-based evaluation of image-processing and classification methods", IEEE Trans. on Medical Imaging, vol.18, n.11, 1999, pp.1085-97.

[30] S. Aja-Fernández, R. San Jose Estepar, C. Alberola-López and C.-F. Westin, "Image Quality Assessment based on Local Variance", Proc. EMBS Annual International Conference New York City, USA, Aug 30-Sept 3, 2006, pp.4815-4818.

[31] Y. Ding, I. W. Selesnick, "Artifact-Free Wavelet Denoising: Non-convex Sparse Regularization, Convex Optimization" IEEE Signal Processing Letters, vol.22 , n.9, 2015, pp.1364-1368.

F. Russo is currently Associate Professor of electrical and electronic measurements in the Department of Engineering and Architecture of the University of Trieste, Italy. He is author/coauthor of more than 100 papers in international journals, textbooks, and conference proceedings including the Wiley Encyclopedia of Electrical and Electronics Engineering, (J. G. Webster ed.). His research interests presently include nonlinear and fuzzy techniques for image denoising, image enhancement and image quality measurement. He was one of the organizers of the 2004 and 2005 IEEE International Workshops on Imaging Systems and Techniques. He served as Technical Program Co-Chairman of the 2006 IEEE Instrumentation and Measurement Technology Conference and as Co-Guest Editor for the Special Issue of the IEEE Transactions on instrumentation published in August 2007. He was invited speaker of the plenary session "Fuzzy Models for Low-level Computer Vision: a Comprehensive Approach" at the IEEE Symposium on Intelligent Signal Processing (WISP 2007). He has served as session chairman/co-chairman in many conferences organized by IEEE, IMEKO, WSEAS and NAUN.