Abstract—Residual noise and collateral distortion are two key features for any image denoising filter. The former is the amount of noise still affecting the data after filtering, the latter represents the price to be paid in terms of detail blur. Measuring these effects is of paramount importance for the validation of a denoising algorithm. This work focuses on Non-Local Means (NLM) filtering that represents one of the most effective approaches to grayscale image denoising. The exact values of residual noise and collateral distortion are derived from NLM theory and an in-depth analysis of these features is provided for different input data and different parameter settings.

Keywords—Image denoising, image filtering, Gaussian noise, nonlinear filtering.

I. INTRODUCTION

It is known that nonlinear filters have become a powerful and widespread used resource for digital image denoising because they can distinguish between noise (to be removed) and image details (to be preserved) [1-6]. In this framework, non-local means (NLM) filters [7] represent one of the most effective and attractive approaches to grayscale image denoising. Unlike other filtering techniques 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 [8-12] or to improve the filtering performance with respect to the basic NLM algorithm [13-17]. 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 [18]. 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 [19-22]. 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 [20], 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 [23-24]. 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 [25-26].

The aim of this paper is twofold: to show a novel approach to the evaluation of RN and CD and to provide an in-depth analysis of the filtering behavior of the basic NLM algorithm. 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 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 many computer simulations, Section IV focuses on some metrological aspects and, finally, Section V 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
Let \( r(i,j) \) be the pixel luminance at location \([i,j]\) in the reference (noise-free) image \((i=1,\ldots,L_1; j=1,\ldots,L_2)\) 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 \([7]\), as follows:

\[
f(i,j) = \sum_{p=-M}^{M} \sum_{q=-M}^{M} w(i,j,p,q) c(i-p,j-q)
\]

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:

\[
0 \leq w(i,j,p,q) \leq 1
\]

\[
\sum_{p=-M}^{M} \sum_{q=-M}^{M} w(i,j,p,q) = 1
\]

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:

\[
d(i,j,p,q) = \sqrt{\sum_{u=-N}^{N} \sum_{v=-N}^{N} G_a(u,v) \left[ (i-p-u,j-q-v) - (i-u,j-v) \right]^2}
\]

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

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

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

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 \( f^-(i,j) \):

\[
f(i,j) = f^+(i,j) + f^-(i,j)
\]

\[
f^+(i,j) = \sum_{p=-M}^{M} \sum_{q=-M}^{M} w(i,j,p,q) c(i-p,j-q)
\]

\[
f^-(i,j) = \sum_{p=-M}^{M} \sum_{q=-M}^{M} w(i,j,p,q) r(i-p,j-q)
\]

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:

\[
E(i,j) = E^+(i,j) + E^-(i,j)
\]

\[
E^+(i,j) = f^+(i,j)
\]

\[
E^-(i,j) = f^-(i,j) - r(i,j)
\]

\[
E'(i,j) \text{ and } E(i,j)
\]

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:

\[
AE(i,j) = |E(i,j)| = AE_{RN}(i,j) + AE_{CD}(i,j)
\]

where \( AE_{RN}(i,j) \) and \( AE_{CD}(i,j) \) are the absolute error components representing the resulting residual noise and collateral distortion, respectively \((AE_{RN}(i,j)\geq0, AE_{CD}(i,j)\geq0)\). In order to evaluate \( AE_{RN}(i,j) \) and \( AE_{CD}(i,j) \), the following cases should be considered:

a) 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).

b) If \( E'(i,j)\neq0 \) and \( E(i,j)=0 \) then \( AE_{RN}(i,j)\neq AE_{CD}(i,j)=0 \).

c) If \( E'(i,j)=0 \) and \( E(i,j)\neq0 \) then \( AE_{RN}(i,j)=0 \) and \( AE_{CD}(i,j)\neq AE_{CD}(i,j) \).

d) If \( E'(i,j)>0 \) and \( E(i,j)>0 \) then \( AE_{RN}(i,j)=E'(i,j) \) and \( AE_{CD}(i,j)=E(i,j) \).

e) If \( E'(i,j)<0 \) and \( E(i,j)<0 \) then \( AE_{RN}(i,j)=-E'(i,j) \) and \( AE_{CD}(i,j)=-E(i,j) \).

f) 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 \( AE_{CD}(i,j)=0 \).

g) If \( E'(i,j)<0 \) and \( E(i,j)>0 \) and \( |E'(i,j)|<|E(i,j)| \) then \( AE_{RN}(i,j)=0 \) and \( AE_{CD}(i,j)=AE(i,j) \).
If $E^+(i,j)>0$ and $E^-(i,j)<0$ and $|E^+(i,j)|=|E^-(i,j)|$ then $AERN(i,j)=0$ and $AECD(i,j)=0$.

i) If $E^+(i,j)<0$ and $E^-(i,j)>0$ and $|E^+(i,j)|>|E^-(i,j)|$ then $AERN(i,j)=AE(i,j)$ and $AECD(i,j)=0$.

j) If $E^+(i,j)<0$ and $E^-(i,j)>0$ and $|E^+(i,j)|<|E^-(i,j)|$ then $AERN(i,j)=0$ and $AECD(i,j)=AE(i,j)$.

k) If $E^+(i,j)<0$ and $E^-(i,j)>0$ and $|E^+(i,j)|=|E^-(i,j)|$ then $AERN(i,j)=0$ and $AECD(i,j)=0$.

Once $AERN(i,j)$ and $AECD(i,j)$ 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:

$$MAE_{RN} = \frac{1}{L_1L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE_{RN}(i,j)$$  \hspace{1cm} (14)$$

$$MAE_{CD} = \frac{1}{L_1L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE_{CD}(i,j)$$  \hspace{1cm} (15)$$

Clearly, we have:

$$MAE = \frac{1}{L_1L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} AE(i,j) = MAE_{RN} + MAE_{CD}$$  \hspace{1cm} (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:
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”, (d) “Cat”.

\[
MSE = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \left[ AE_{RN}(i, j) + AE_{CD}(i, j) \right]^2
\]

\[
= MSE_{RN} + MSE_{CD} + MSE_{MIX}
\]  \hspace{1cm} (17)

where:

\[
MSE_{RN} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \left[ AE_{RN}(i, j) \right]^2
\]  \hspace{1cm} (18)

\[
MSE_{CD} = \frac{1}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \left[ AE_{CD}(i, j) \right]^2
\]  \hspace{1cm} (19)

\[
MSE_{MIX} = \frac{2}{L_1 L_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \left[ AE_{RN}(i, j) \cdot AE_{CD}(i, j) \right]
\]  \hspace{1cm} (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. 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 four 512×512 grayscale images: “Lena”, “Lighthouse”, “Motorbike” and “Cat” (Fig. 1). In the first group of tests, we generated four 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 $\text{MAE}_{RN}$ and $\text{MAE}_{CD}$ 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 $\text{MAE}_{RN}$ correctly decreases as the smoothing parameter $h$ becomes larger. On the contrary, $\text{MAE}_{CD}$ 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 $\text{AE}_{RN}(i,j)$ (green) and $\text{AE}_{CD}(i,j)$ (red) are also reported. We can see that, for $h<50$, the presence of unfiltered noise is apparent.
Fig. 5. MAE, MAE\textsubscript{RN} and MAE\textsubscript{CD} 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”, (d) “Cat”.

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\textsubscript{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\textsubscript{RN} and MAE\textsubscript{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\textsubscript{CD} component is generally larger than the corresponding MAE\textsubscript{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.

In the third group of tests, we investigated the dependence of MAE, MAE\textsubscript{RN} and MAE\textsubscript{CD} on the variance \(a\) of the Gaussian kernel (see eq.(4)). Fig.6(a) shows the graphical representations of MAE in dependence on parameter \(h\) (10\(\leq h \leq 150\)) and parameter \(a\) (0.5\(\leq a \leq 5\)) for the “Lena” picture corrupted by Gaussian noise with variance \(\sigma^2=200\). We can notice that the parameter \(a\) does not play a very critical role: about the same minimum value of MAE (\(\approx 4\)) can be obtained in the interval 2\(\leq a \leq 5\) (although for different values of \(h\)).
Similar considerations can be applied to the case $\sigma^2=400$ (Fig.6(b)). The corresponding graphical representations of MAE$_{RN}$ and MAE$_{CD}$ are reported in Fig.7 and Fig.8, respectively. We can see that the parameter $a$ has some influence on the values of MAE$_{RN}$ and MAE$_{CD}$. Indeed, let the variance $a$ range from 0.5 to 5 for a given value of the smoothing parameter $h$: the MAE$_{RN}$ increases whereas MAE$_{CD}$ decreases. A satisfactory choice is represented by $a=2$, adopted in the previous groups of experiments.

IV. METROLOGICAL CONSIDERATIONS

It is worth pointing out that MAE$_{RN}$ and MAE$_{CD}$, evaluated by (14-15), yield the true values of residual noise and collateral distortion during NLM filtering. Hence, these metrics are more accurate than any previous evaluation system including the human eye. On one hand, the results given by
MAER\textsubscript{N} and MAEC\textsubscript{D} are in perfect agreement with the visual inspection of the filtered data (see Figs.3-4). On the other hand, these metrics can yield a quantitative evaluation of the filtering features whereas the human perception cannot. Furthermore, the error maps provide the exact location and type of filtering errors at the pixel level. In order to clarify the situation, let us focus on the following experiment (Fig.9), where we considered the “Lighthouse” picture corrupted by Gaussian noise ($\sigma^2=200$) and processed by slightly different values of the smoothing parameter $h$ chosen in the proximity of the minimum MAE: $h=77$ (Fig.9(c)) and $h=78$ (Fig.9(d)). From visual inspection, the image data look almost identical. The measurements of residual noise and collateral distortion reveal the differences. Indeed, we obtained the following evaluations: MAER\textsubscript{N}=1.695 and MAEC\textsubscript{D}=2.621 ($h=77$), MAER\textsubscript{N}=1.640 and MAEC\textsubscript{D}=2.875 ($h=78$). The image filtered by the larger value of $h$ is really affected by more distortion and less noise, although the human eye cannot appreciate the difference neither the exact distribution of the filtering errors (Figs.9(e) and 9(f)). The analysis of these maps leads to some additional considerations about the accuracy of previous algorithms for quality assessment in image filtering. For instance, type-1 vector metrics [19] resorted to a simple idea: to compute the filtering distortion along the object borders (obtained from a map of edge gradients of the reference picture). The error maps depicted in Figs.9(c) and 9(f) show, however, that the filtering behavior is not so simple: some object contours are affected by collateral distortion, whereas other edge regions are impaired by residual noise only. Indeed, the filtering errors in the proximity of such edges change from residual noise to distortion for growing values of the smoothing parameter $h$ (this effect is also apparent in Fig.3 and Fig.4). Even if the validation of previous metrics is not the scope of this work, a study of their accuracy would greatly
benefit from the availability of specific cases, where the exact amounts of residual noise and collateral distortion are known for any image.

V. CONCLUSIONS

A new approach to quantitative evaluation of key filtering features in NLM processing 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 main smoothing parameter and the variance of Gaussian kernel. Furthermore the reported error maps have provided very accurate information about kind and spatial location of the 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


