Multimedia Tools and Applications (2024) 83:4065–4092
https://doi.org/10.1007/s11042-023-15339-4

Dual Non-Local Means: a two-stage information-theoretic filter for image denoising

André R. de Brito[1] · **Alexandre L. M. Levada**[1]

Received: 4 August 2022 / Revised: 5 January 2023 / Accepted: 6 April 2023 /
Published online: 20 May 2023

Abstract
Image denoise has been explored with the development of various filters used to remove or reduce random disruptions on observed data, but at the same time it preserves most of the edges and the fine details of the scene. The issue caused by the combined deterioration of the Gaussian noise succeeds the scattering through all the signal frequencies. Thus, the most effective filters for this type of noise are implemented in spatial domain. In this article, we proposed a Non-Local Means filter that combines the average of each fragment of a browser window, by using four measures – of distinct similarities – among the Gaussian densities that are estimated from the following fragments: the Kullback-Leibler divergence, the Bhattacharyya distance, the Hellinger distance and the Cauchy-Schwarz divergence. Computational experiments were done in a set of 7 images that were deteriorated by a noise of Gaussian type, considering that the data obtained show that the proposed methods are capable of producing, on average, a Peak Signal-to-Noise Ratio significantly greater than the one the combination of Total Variation, Non-Local Means, BM3D, Anisotropic Diffusion, Wiener, Wavelet e Bilateral filters does when they are applied independently.

Keywords Filtering · Denoising · Non-local means · Information theory

1 Introduction

Image denoise on digital images is a crucial process in the stage of image preprocessing on applications that envolve the areas of image processing, computational vision and pattern recognition. This stage is characterized by a simple math operation of softening or even the localization and recognition of objects. Thereby, the process of restoration of noisy images is a topic of scientific interest for great part of the research community in the areas of signal processing and computational vision. In summary, the main goal of denoising is to

✉ André R. de Brito
andrerb_1992@hotmail.com

Alexandre L. M. Levada
alexandre.levada@ufscar.br

[1] Computing Department, Federal University of São Carlos, São Carlos, SP, Brazil

minimize and/or to reduce random disruptions of a signal in an image, in order to preserve the maximum amount of relevant information subsequent stages.

In the image processing area, the concept of quadratic noise is flagged as a random variation of intensity information, which occurs during the process of acquisition, recording, processing and transmition of images [1]. In this context, certain types of noise present some characteristics about the intensity of the degraded pixels, a fact that randomly leads to black and white spots, and image blurring, or they assume random variations in their parameters both in the spatial domain and in the frequency domain, which causes and compromises the ability of visual interpretation of an image.

As a result, the methods of image filtering exist in order to reduce the effects caused by the noises and they are usually classified into two categories: (1) in the spatial domain; (2) in the frequency domain [8]. The filtering methods that work in the spatial domain operate directly over the matrix of the image intensity, through operations of convolution with a mask [27]. n the other hand, the filtering methods that operate in the frequency domain are based on the modification of the Fourier transform of the image [31].

In this research, only filtering methods that are on the spatial domain are used: Wiener [13], Total Variation [25], Anisotropic Diffusion [21], Wavelets [18], Bilateral [29], Non-Local Means (NLM) [3] and BM3D [4]. The main similarity between these image filtering methods in the spatial domain is that they normally operate through the convolution process with a mask over the matrix of pixel/image intensity or through calculations of the distance between patches and pixels.

Among the many filtering methods that were mentioned above, the Wiener filter is considered one of the classical methods in the literature of image filtering, which is applied in the process of reduction of different types of signal dependent noises [13]. This filter maps the image and its noise on random variables, in order to find an estimate between the reference image and the filtered image, in such a way that the mean square error (MSE) between them is minimized. Likewise, this filter is excellent in terms of estimating the minimum linear mean square error in the process of filtering and in the softening of image noises.

Meanwhile, the Total Variation filter is an optimized algorithm that is applied for the noise reduction in images [25]. To sum up, the approach adopted by this method solves the problem of denoising in images through a mathematical modeling in an optimized way, while aiming to reduce the signal-noise ration (SNR) of a reference image with a term defined by the magnitude of the absolute gradient of the image.

Regarding the Anisotropic Difussion, it is an adaptative and non-linear method based on partial differential equations (PDE) for an equation of heat diffusion, in which the diffusion coefficient is a function of the image gradient [21]. In short, the conception of this method is to apply the convolution of an original image with a Gaussian core in different scales (variance). In this way, the result of this convolution allows to obtain blurred images in multiple resolutions, an intra-region softening and the preservation of the picture edge.

The Bilateral filter is also an adaptative, non-line filtering method that uses a mathematical modeling to substitute the intensity of each pixel by a weighted average of its neighboring pixels [29].The results of the wrights obtained by the weighted average are defined in terms of two local differences: spatial differences calculated by the Euclidian distance, and radiometric differences between the central pixels and their neighboring ones. Consequently, this method allows to preserve the edges and to reduce the noise in uniform regions of an image.

In addition, the Wavelet filter is a filtering method that is applied to decompose and represent an orthogonal multi-resolution sign [18].The Wavelet representation is between the

spatial and the frequency domains. In the spatial domain, a transformed Wavelet is applied in order to obtain the coefficients in the Wavelets sparse domain, using a Wavelet base. In this domain, the noise slightly degrades all the coefficients, in a way that it changes zero coefficients into nonzero coefficients. Like that, according to a threshold T, lower coefficients must be defined for zero, while higher coefficients are attenuated or they remain unchanged. To sum up, the key-point is to determine the value of T.

Furthermore, the BM3D filter presents the prefiltration and post-filtration stages [4]. In the prefiltration stage, non-local techniques and the transformed Wavelet are applied to decompose the image into fragments, then they are grouped according to their similiarity. The post-filtration stage is charachterized by the usage of the Wiener filter in order to minimize the edge effects and the denoising on the image.

Great part of the theoretical background of this article is related to the Non-Local Means (NLM) filtering method. The NLM filter is a technique applied to deal with the reduction of additive noise of Gaussian type [3]. This method is based on the fact that digital images have characteristics that are repeated in the image not only in regions, but also in a global way. In this way, using the Euclidian distance as a measure of similarity, this method aims to find the estimated value of the intensity of each pixel in a certain region in the image.

In the last decades, different methodologies and mathematical models have been reported in the history of the filtering methods art to deal with issues regarding Gaussian denoising in digital images [6, 19, 20, 24]. In this context, the conception of the classic filtering method called NLM is based on the calculation of the Euclidian distance between patches, which is appropriated for images that have additive white Gaussian noise (AWGN). Therefore, an adaptative method is necessary to eliminate the different types of noise in digital images, considering the various types of intensity distribution.

Hence, the method that is proposed in this article suggests the development of a new adaptative filtering method of images that aims to reduce the Gaussian noise in a way that it individually combines the NLM filter with four different divergences of the information theory, in their variations. In order to measure the similarity between patches in the same browser window and based on Levada's work to spread the NLM filter to the Gaussian noise [16], the following divergences are used: (1) the Kullback-Leibler divergence; (2) the Bhatthacharyya distance; (3) the Hellinger distance; and (4) the Cauchy- Schwarz divergence. Like that, the capacity of the proposed method can be relevant in the processing of different image types, such as the tomographic and hyperspectral images. Besides that, this method can be extended to other types of degradation, like the speckle noise and non-Gaussian noises.

The main contributions of this research are: 1) it was proposed a new filtering method, called Dual Non-Local Means, which acts as an extension to the traditional Non-Local Means filter, and it is applied in images degraded by Gaussian noises; 2) the L1 norm of the Dual Non-Local Means filter was used and the stochastic distances of the Kullback-Leibler divergence, Bhatthacharyya distance, Hellinger distance and Cauchy-Schwarz divergence were used as a measurement of similarity; and 3) the results obtained through the 7 different images degraded by a Gaussian noise indicate that the proposed method can produce superior outcomes comparing to the following filters: Wiener, Total Variation, Anisotropic Diffusion, Wavelets, Bilateral, Non-Local Means (NLM) and BM3D. That comparison was done in a quantitative way, by applying the Peak Signal-to-Noise Ratio (PSNR) metrics for analysis of mean, minimum and maximum regarding the results that were obtained through the other filters that were compared.

This article is organized in the following way: Section 2 is meant to describe works related to this research, by presenting the relation between this study and the various important and relevant methods for the image filtering area, such as the Wiener, Bilateral, Wavelet, Total Variation, Aniostropic filters and, specially, the Non-Local Means and the BM3D filters. Section 3 describes the mathematical formulation of the stochastic distances of the Kullback-Leibler divergence, Bhatthacharyya distance, Hellinger distance and Cauchy-Schwarz divergence, when applied to the Non- Local Means filtering model. Section 4 detailes the proposed method: Non-Local Means using the KL divergence, Non-Local Means using the Bhatthacharyya coefficient, Non-Local Means using the Hellinger distance and Non- Local Means using the Cauchy-Schwarz divergence. Section 5 shows the experiments done and the results obtained in terms of os the PSNR quantitative metrics and the qualitative metrics. Finally, Section 6 shows the conclusions, final considerations and some directions for future works.

2 Related work

The literature on the image filtering methods is really long, therefore an extensive literary review assisted in the scope of this article. In this respect, this section determines the relation between this work and other important and recent methods in the area of image filtering, such as the Wiener, Bilateral, Wavelet, Anisotropic Diffusion, Total Variation, BM3D filters and, in special, the Non- Local Means filter.

In 2020, Petkova and Draganov presented, in their article, a proposal of application of the Wiener filtering method on digital images that denote an unknown level of Gaussian noise [22]. In this context, the variance of Gaussian noise found on images was caused by the distribution of intensities in homogenous areas contained in the image. Thus, the simplest method of the Wiener filter for denoising was applied when obtaining the results of the noisy images. In this way, the authors conducted an extensive analysis on the influence of the size of the filter Wiener mask regarding the variance of the noise. Consequently, the results of that influence led to the conclusion that the usage of the Wiener adaptative filter is efficient, in terms of general (PSNR) and structural (SSIM) preservation.

In 2020, Jin and Luan revealed, in their study, a new approach for the denoising in digital images, based on the Total Variation filtering method and on the weighting function [11]. The approach proposed by the authors initially analyzes the ladder effect caused by the traditional Total Variation filter. Besides that, a second analysis was done, based on the effects of the weighting function in edge regions, in flat regions and in gradient regions. Through these analyses and the information provided by the traditional method, the authors used the Total Variation filter to modify the process of image denoising in a problem of minimization of the energy function. Thereby, after the filter application, they used the weighting function to calculate the gradient magnitude and the value of the local variance of each pixel. By the end of the filter application and after the mathematical calculus, it was possible to analyze in detail the characteristics of different parts of an image. From this new approach, the authors concluded that the Total Variation filter can effectively extinguish the ladder effect of the traditional Total Variation filtering method.

In 2020, Zhang and Sun showed, in their research, a new adaptation of the BM3D algorithm for the denoising in digital images, without affecting the the intra-region and edge details [32]. In the first filtering stage, the algorithm proposed by the authors used the Anisotropic Diffusion (DA) method, along the BM3D method, to search for similar blocks

based on the vertical directions of the edge. Therefore, more concrete information could be obtained about the edges and the details of the processing effects. During the second stage, in order to calculate the function of the diffusion coefficient, a mathematic model based on the hyperbolic tangent function was introduced. In that way, the values obtained by the gradient of the neighboring of the eight directions in the image are used for the AD filter application. Through this approach, the researchers concluded that the adaptation of the AD filter, along the BM3D algorithm, could promote better results than the traditional BM3D method, in terms of PSNR and SSIM evaluation. In addition, it provided superior results regarding denoising, edge preservation and image detailing.

Similar to the methodology proposed by Zhang and Sun (2020), which seeks to adapt the BM3D method of filtering noises through another type of filter, Yahya and collaboratores presented, in their work, a model of adaptation of the BM3D method, done through an adaptative filtering technique [30]. In this context, the authors proposed to divide that method into two stages, aiming at reducing the Gaussian noise and at preserving the edge. The first part of this adaptation seeks to replace the traditional hard-thresholding technique that is on BM3D by the Total Variation method of adaptative filtering. Like that, the Total Variation filter is applied on image areas that contain slight noises, in contrast to the traditional hard-thresholding technique used in areas of high noises. This adaptation that was proposed by the authors allows a high performance regarding denoising and edge preservation. Thus, the second part of this stage uses the calculus of the adaptive weight function and the k-means clustering technique to calculate the spatial distance between a reference patch and its candidates. Consequently, using the Total Variation adaptative filter, the adaptive weight function and the k-means clustering technique, as well as through the PSNR and SSIM metrics, the authors noticed the superiority of this method in comparison to the traditional BM3D method.

In 2021, Salehi and Vahidi demonstrated, in their study, a new method of hybrid filtering, which is composed by three stages and three filters for image [26]. In this scenario, the combination of the three filtering methods used by the authors was based on the Wiener, Bilateral and Wavelet filters. In this way, the first stage of the process to denoise concerns to obtain the coefficient of variation and to apply the fuzzy c-means technique to classify the image regions. Then, the second stage consists of the combination and application of denoising filters, being them the Bilateral filter for homgenous regions, and the Wiener and Wavelet filters in regions that contain details and edges. In the third and last stage, the resulting image is evaluated through the logic fuzzy approach. Through the approach of the three named stages, the authors concluded that the combination of the three filtering methods was able to overcome other methods that exist in literature. Besides that, this method could preserve important details and edges on the image.

In 2021, Gupta and Lamba made evident, in their research, two new guidelines to be applied in the traditional Anisotropic Diffusion filtering model [23]. The proposals made by the authors include the traditional Anisotropic Diffusion filter, which is based on a new coefficient of diffusion and on a new threshold, depending on the image. The new model of diffusion coefficiente relied on the function of the tangent sigmoid, so that there was a greater speed in the rate of the function convergence. Regarding the new threshold, a mathematical calculation of weighted absolute mean deviation of the gradient of each processed image was done. So, the researchers concluded that the proposed method demonstrated a higher performance in denoising and edge preservation, besides effectively supplying the ladder effects and blurred edges in relation to the traditional anisotropic diffusion filtering method.

Recently, Kundu and collaborators presented, in their article, a new way of evaluating the value of intensity and retention of edges of NLM filter, using a genetic algorithm [15]. In this scenario, instead of applying the calculation of the weighted average to find the neighboring pixels, the authors used some techniques of genetic algorithms to choose pixels that were more relevant in the local neighborhood, through the introduction of new intensity values. By doing that, the selection of these significant pixels aided in denoising and also improved the process of filtering in terms of edge preservation and evaluation of an intensity value that is considered deprived of noise. Through this new adaptation of the NLM filter and doing an empirical analysis, the authors demonstrated that the proposed filter exceeds the traditional NLM filter.

Therefore, the methodology that is proposed in this article aims to improve the quality of the method of Gaussian noises reduction at a considerable level in comparison to the other attempts done on previous works. Besides that, the previous efforts provided several results for different variances of Gaussian, Salt and Pepper, and Poisson noises, while the method proposed in this article was applied and analyzed for additive Gaussian noise. The exhaustive experimental approach was used in order to show the efficiency of the metrics of information theory when applied on the Dual NLM method that was proposed.

3 Information theoretic distances

In the literary context, the metrics of information theory have been applied on several works, and it has obtained success in the areas of Mathematics and Statistics to quantify the similarity level between random variables. Among the various metrics that exist in the information theory, in the context of this article the Kullback-Leibler divergence, the Bhattacharyya distance, the Hellinger distance and the Cauchy-Schwarz divergence are used and analyzed.

3.1 The Kullback-Leibler divergence

The first metrics used in this work is the Kullback- Leibler (KL) divergence. The KL divergence, also known as Relative Entropy, was initially proposed on the On Information and Sufficiency article, in 1951, by Kullback and Leibler [14]. In this context, the KL method aims to calculate the divergence between two probability distributions (or relative frequencies). In this way, it is possible to represent the math equation of the KL divergence by the following expression:

$$D_{KL}(p,q) = \int d_{\mu 1}(x)\, log \frac{p(x)}{q(x)} = \int p(x)\, log \frac{p(x)}{q(x)} d\lambda(x) \tag{1}$$

In which the parameters p and q denote the discrete distribution of probabilities of a random variance X with parameter λ that is determined from the sample x.

Given the univariate Gaussian context, it is possible to compute the KL divergence as:

$$D_{KL}^{sym}(p,q) = \frac{1}{2}\left[D_{KL}(p,q) + D_{KL}(q,p)\right] = \frac{1}{4}\left[\frac{\sigma_1^2(\mu_1-\mu_2)^2}{\sigma_2^2} + \frac{\sigma_2^2(\mu_1-\mu_2)^2}{\sigma_1^2} - 2\right] \tag{2}$$

Likewise, (2) is summarized as:

$$D_{KL}^{sym}(p,q) = \frac{1}{4\sigma_1^2\sigma_2^2}\left[\left(\sigma_1^2-\sigma_2^2\right)^2 + \left(\mu_1^2-\mu_2^2\right)^2\left(\sigma_1^2-\sigma_2^2\right)\right] \tag{3}$$

In which the parameters σ and μ show the variances and non-local means of the distributions.

3.2 The Cauchy-Schwarz divergence

The second metrics used in this work is the Cauchy- Schwar divergence. Based on the probability theories and on the math theory, the metrics present on the information theory that are applied more frequently on the literature are the Kullback-Leibler divergence and the Rényi divergence [7, 9, 10, 17]. However, the Kullback-Leibler divergence and the Rényi divergence make fast and efficient calculus impossible on applications involving the classification of objects in the static recognition. In this scenario, the Cauchy-Schwarz divergence arises, which is an analytical expression, closed for a Gaussian mixture (MoG), that enables fast and efficient calculations in applications of the computational vision and classifying objects areas.

The Cauchy-Schwarz divergence for two densities of random vectors $p(x)$ and $q(x)$ is defined as:

$$\begin{aligned} D_{CS}(p,q) &= -log\left(\frac{\int p(x)q(x)dx}{\sqrt{\int p(x)^2dx \int q(x)^2dx}}\right) \\ &= \tfrac{1}{2}\, log\left(\int p(x)^2\, dx\right) + \left(\int q(x)^2\, dx\right) - log\left(\int p(x)\, q(x)\, dx\right) \end{aligned} \tag{4}$$

in which the parameters p and q represent two symmetric measures regarding the probabilities distributions, such as $0 \leq D_{CS} < \infty$,in which the result of the minimum value is obtained if and only if $p(x) = q(x)$.

It can been seen that, in the univariate Gaussian case, the Cauchy-Schwarz divergence can be computed by [28]:

$$D_{CS}(p,q) = \frac{1}{2} log\left(\frac{\left(\sigma_1^2+\sigma_2^2\right)^2}{4\sigma_1^2\sigma_2^2}\right) + \frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_1^2} \tag{5}$$

where the parameters σ and μ represent the variances and non-local means of distributions, respectively.

3.3 The Bhatthacharyya distance

The third metrics used in this work is the Bhat- tacharyya distance. Based on the process of stochastic distances, the Bhattacharyya distance was originally proposed on The Divergence and Bhattacharry Distace Measures in Signal Selection article, in 1967, by Thomas Kailath [12]. In this context, this method defines a normalized distance between two coefficients:

$$D_{BC}(p,q) = -ln\left(\sum_{x=1}^{n}\sqrt{p(x), q(x)}\right) \tag{6}$$

in which the parameters p and q represent the distributions of normalized probabilities and N the number of distribution compartments. In addition, the Bhattacharyya distance must be limited between $0 \leq D_{BC}(p,q) \leq \infty$.

Furhtermore, in the univariate Gaussian case, the Bhattacharyya distance can be computed by:

$$BC_{Bhat}(p,q) = -lnBC(p,q) \tag{7}$$

in wich $BC(p, q)$ is the Bhattacharyya coefficient, in the Gaussian case given by:

$$BC\,(p,q) = \sqrt{\frac{2\sigma_1\sigma_2}{\sigma_1^2+\sigma_2^2}}exp\left\{-\frac{1}{4}\frac{(\mu_1-\mu_2)^2}{\sigma_1^2+\sigma_2^2}\right\} \tag{8}$$

in wich parameters σ and μ denote the variances and non-local means of distributions, respectively.

3.4 The Hellinger distance

The fourth and last metrics to be applied in this work concerns the Hellinger distance. Its origin dates back to 1907, by the Germanpelo mathematician Ernst David Hellinger, and it presentes a mathematical modeling different from the Riemann integral to measure the distance between distributions of discrete probabilities. Besides its application for the distance calculation, the Hellinger method is classified for calculus that envolve metrics and divergence [5]. Like that, the Hellinger distance is defined as:

$$D_H\,(p,q) = \frac{1}{\sqrt{2}}\left[\sum_{x\in\Omega}\left(\sqrt{p\,(x)}-\sqrt{q\,(x)}\right)^2\right]^{\frac{1}{2}} \tag{9}$$

in which the parameters p and q q represent the probabilities distributions of a countable Ω space. In this case, the Hellinger distance is restricted by $0 \leq D_H\,(p,q) \leq 1$. Consequently, when the result of the Hellinger distance corresponds to 0, it means that there was no divergence; on the other hand, if it corresponds to 1, the distributions of probability do not share a common support..

In the univariate Gaussian case, the Hellinger distance can be computed by:

$$D_{Hel}\,(p,q) = 1 - BC\,(p,q) \tag{10}$$

in which the function $BC(p, q)$ denotes a Bhattacharyya distance coefficient, given by (8).

4 The proposed method

In this section, the standard filtering method is presented and compared to the method proposed in this study, in order to describe in detail the functioning of each mathematical variance. To sum up, the idea of the Non-Local Means method is to deal with Gaussian noises by the calculation of the weighted median instead of the definition of a weighted average.

Considering an additive Gaussian noise, uncorrelated and independent of the signal, the mathematical model used to describe the process of filtering is given by the following equation:

$$y_i = x_i + n_i \tag{11}$$

in which y_i denotes the result of the noisy pixel, x_i,is associated with the noise-free pixel and n_i is an operator of the additive Gaussian noise. In the traditional approach $n_i\ N(0, \sigma^2)$ and $x_i\ N(\mu_0, \sigma_0^2)$ are defined. It should be emphasized that the noise is not correlated, that is, $E[n_i n_j] = \sigma_0^2\delta_{i,j}$, whereby, $\delta_{i,j} = 1$ if $i = j$ and $\delta_{i,j} = 0$ if $i \neq j$. In this way, the goal is to recuperate xi since x_i from y_i, given p_i the ith patch.

4.1 Non-Local Means

The Non-Local Means traditional filtering method (NLM) was proposed in 2005, by Buades, Coll and Morel, as a method applied to deal with the denoising of additive Gaussian noises in images [3]]. In this scenario, this filter aims to scan the whole image searching for similar pixels by using the concept of similiarity measurement between patches, as shown in Fig. 1 [16].

Thereby, the Non-Local Means filtering method when applied to the noise is contextualized in such a way that given the noisy $y = y_i \mid \epsilon I$, the great value for the noise-free pixel x_i, denoted by $NL[x_j]$ is computed as a weighted average of all the pixels in image expressed by:

$$NLM\,[x]\,(i) = \sum_{j \in f} w\,(i,\,j)\,x_j \tag{12}$$

in which $w(i, j)$ represents the weights assigned to the similarity between pixels i and j, meeting the condition $0 \leq w(i, j) \leq 1$ and $\sum_j w(i, j) = 1$. In this way, the similarity between the i and j pixels has as analogy the similarity of vectors of intensities in the levels of gray $x(N_i)$ and $x(N_j)$. Therefore, the N_k parameter represents a patch with a central k pixel. Then (12) of the traditional NLM filter, that is expresses by the $w\,(i,\,j)$ weights, is defined as:

$$w\,(i,\,j) = \frac{1}{Z\,(i)} exp\left\{\frac{-d_E(i,j)}{h^2}\right\} = \frac{1}{Z\,(i)} exp\left\{\frac{-\left\|x(N_i)-x(N_j)\right\|^2_{2,\tau}}{h^2}\right\} \tag{13}$$

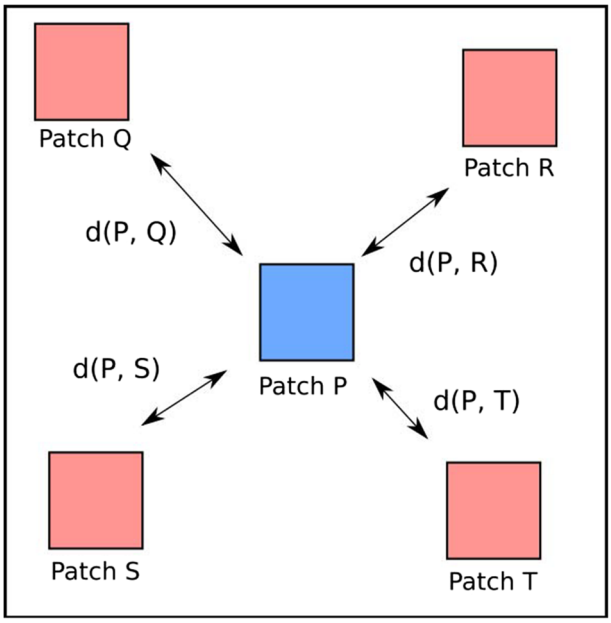

Fig. 1 Process performed by the Non-Local Means filter employing the concept of measuring similarity between image patches using weighted averaging

in which the value $\tau > 1$, h^2 is a parameter that controls the level of softening, and the constant $Z(i)$ refers to a function of normalization that is given by:

$$Z(i) = \sum_j e^{-\frac{\left\|x(N_i)-x(N_j)\right\|^2_{2,\tau}}{h^2}} \tag{14}$$

Hence, the calculation of the sum expressed by (14) do not envolve all the image pixels, but only those that belong to a browser window of $t \times t$ size, as a way to reduce the computational cost and to make the method viable.

4.2 Dual Non-Local Means

In na attempt to overcome the main limitations of the Non-Local Means traditional filtering method, the Dual Non-Local Means method is proposed in this article. The Dual Non-Local Means filter consists of a method for the reduction of Gaussian noise in two stages, which incorporates measures of non-Euclidian similarity, based on the information theory, to measure the distance between the patches. Works from the literature that consider other types of noise, such as the Poisson one, were successfully developed [2]. However, improvements for the original NLM filter are not much explored in the literature, in the case of the Gaussian noise.

In this context, the conception of the Dual Non-Local Means method considers for each $i = 1, 2, ..., n$ the e mean and the variance in each pixel for the definition of the parametric vector $\vec{theta} = (\mu_i, \sigma_i^2)$. While the variance on i is locally estimated by using all the pixels insiude the ith patch, the means are estimated in a non-local way, directly applied to the standard NLM filter. So, by using the Euclidian NLM exit, that is represented by the non-local estimates of parameters μ_i to compute the parametric version of great weights, is defined $w(i, j)$ as:

$$w(i, j) = exp\left\{\frac{-d_E(N_i, N_j)}{h^2}\right\} \tag{15}$$

in which $d_p(N_i, N_j)$ representes the parametric measure based on the information theory.

Therefore, the idea of the Dual Non-Local Means method is to substitute the L1 norm by the function of Kullback-Leibler divergence, the Bhattacharyya distance, the Hellinger distance and the Cauchy-Schwarz divergence. In this way, the Dual Non-Local Means actually defines a double filtering process, in which the first stage, based on Euclidian NLM, is responsible for the estimation of the model parameters, while the second stage is responsible for the computation of measures of parametric similarity. Because of that, this techinique is called Dual Non-Local Means.

4.2.1 Dual Non-Local Means KL

In the context of applied mathematics for the control of similarity between random variances and for the improvement of multivariate data analysis processes, the usage of measures of information theory is proposed, which is called KL divergence or relative entropy, as alternatives to benefit data grouping and filtering. To sum up, the Dual Non-Local Means KL filter works in the following way:

- From a noisy image, the standard (Euclidian) NLM filter is applied to estimate the μ_i in a non-local way, for $i = 1, 2, ..., n$;

- For each pixel x_i, two local variances σ_i^2 are estimated inside patch N_k of $f \times f$ size, defined as:

$$\sigma_i^2 = \frac{1}{f^2} \sum_{j \epsilon p_i} (x_i - \mu_i)^2 \tag{16}$$

- KL divergence is calculated between the central $\hat{N}_i$,by using (3);
- Calculate the weight $w\,(i,\,j)$ as:

$$w\,(i,\,j) = \frac{1}{Z\,(i)} exp \left\{ \frac{-d_{KL}(p,q)}{h^2} \right\} \tag{17}$$

- Noise-free pixel x_i is evaluated as:

$$NLM\,[x]\,(i) = \sum_{j \in f} w\,(i,\,j)\,m_j \tag{18}$$

4.2.2 Dual Non-Local Means Cauchy-Schwarz

Analogous to the Dual Non-Local Means KL method, the Dual Non-Local Means CS filter is used in the following way:

- From a noisy image, the standard (Euclidian) NLM filter is applied to estimate theμ_i means in a non-local way, for $i = 1, 2, ..., n$
- For each x_i pixel, we estimate two local variance σ_i^2 within the patch N_k of size $f \times f$, defined by:

$$\sigma_i^2 = \frac{1}{f^2} \sum_{j \epsilon p_i} (x_i - \mu_i)^2 \tag{19}$$

- CS divergence is calculated between the central $\hat{N}_i$ by using (5);
- Calculate the weight $w\,(i,\,j)$ as:

$$w\,(i,\,j) = \frac{1}{Z\,(i)} exp \left\{ \frac{-d_{CS}(p,q)}{h^2} \right\} \tag{20}$$

- Noise-free pixel x_i is evaluated as:

$$NLM\,[x]\,(i) = \sum_{j \in f} w\,(i,\,j)\,m_j \tag{21}$$

4.2.3 Dual Non-Local Means Bhattacharyya

Similar to the filtering process of the Dual Non-Local Means KL method, the Dual Non-Local Means Bhattacharyya filter is applied in the following way:

- From a noisy image, the standard (Euclidian) NLM filter is applied to estimate the μ_i means in a non-local way, for $i = 1, 2, ..., n$
- For each x_i pixel, two local variance σ_i^2 are estimated inside patch N_k of $f \times f$ size, defined as:

$$\sigma_i^2 = \frac{1}{f^2} \sum_{j \epsilon p_i} (x_i - \mu_i)^2 \tag{22}$$

- Bhattacharyya distance is calculated between the central patch $\hat{N}_i$ and $\hat{N}_j$, using the (8);

- Calculate the weight $w\,(i,\,j)$ as:

$$w\,(i,\,j) = \frac{1}{Z\,(i)} exp\left\{\frac{-d_{BC}(p,q)}{h^2}\right\} \tag{23}$$

- Noise-free pixel x_i is evaluated as:

$$NLM\,[x]\,(i) = \sum_{j \in f} w\,(i,\,j)\,m_j \tag{24}$$

4.2.4 Dual Non-Local Means Hellinger

When compared to the Dual Non-Local Means KL method, the Dual Non-Local Means Hellinger filter is used in the following way:

- From a noisy image, the standard (Euclidian) NLM filter is applied to estimate the μ_i means in a non-local way, for $i = 1, 2, ..., n$;
- For each x_i pixel, two local variance σ_i^2 are estimate inside patch N_k of $f \times f$ size, defined as:

$$\sigma_i^2 = \frac{1}{f^2} \sum_{j \epsilon p_i} (x_i - \mu_i)^2 \tag{25}$$

- Hellinger distance is calculated between the central patch $\hat{N}_i$ by using the (10);
- Calculate the weight $w\,(i,\,j)$ as:

$$w\,(i,\,j) = \frac{1}{Z\,(i)} exp\left\{\frac{-d_H(p,q)}{h^2}\right\} \tag{26}$$

- Noise-free pixel x_i is evaluated as:

$$NLM\,[x]\,(i) = \sum_{j \in f} w\,(i,\,j)\,m_j \tag{27}$$

5 Experiments and results

In order to test and evaluate the performance of the Dual NLM filtering method when applied to the process of denoising in digital images, a group of 7 types of images, with sizes of 512 x 512 pixels, 8 bits and in shades of grey were used. The images that were used refer to: Airplane, Barbara, Camera, Car, House, Lena and Peppers. The set of images used were taken from the dataset: https://sipi.usc.edu/database/. All the images taken from USC Image Database are intended for research purposes.

In this scenario, the performance of the Dual NLM filter was compared to other filters that exist in the literature, such as: the usual Wiener filter, NLM, the Bilateral one, Total Variation, Wavelet, Anisotropic Diffusion and BM3D.

In order to compare the different methods, quantitative metrics (PSNR) were selected to evaluate the maximum peak in the signal-to-noise ratio between a reference image and its filtered image. The higher the resultant value of this index, the better is the result of the applied filter. All the images used in this article were degraded by additive Gaussian noise, with a variance of $\sigma_n^2 = 10$ of image pixels, which were randomly selected.

Table 1 presents the results of the evalutation done with the PSNR metrics for 7 types of different images, considering a type of mathematical model of the information theory for each column.

Table 1 PSNR's obtained after filtering with the parametric NLM filter (Kullback-Leibler, Bhatthacharyya, Hellinger and Cauchy-Schwarz) for images corrupted by Gaussian noise with $\sigma = 10$

Image	Dual NLM KL	NLM Pattern	Dual NLM Bhattacharyya	NLM Pattern	Dual NLM CS	NLM Pattern	Dual NLM Hellinger	NLM Pattern
Airplane	32,0844	32,1147	31,2266	30,4348	**32,1393**	31,8127	30,8756	30,4330
Airplane	31,8908	31,5071	31,2485	31,0526	**32,0395**	31,5161	31,1704	30,2292
Airplane	**31,9330**	31,7151	31,6718	30,7826	31,9116	31,6685	31,5132	30,8737
Barbara	32,9981	32,8209	32,7934	30,7440	**33,0523**	32,7391	32,6114	30,6365
Barbara	32,9058	31,7102	32,7085	31,8039	**32,9923**	31,7702	32,5891	31,2905
Barbara	32,9442	32,3621	32,6674	32,1129	**33,0133**	32,4392	32,6196	32,0643
Camera	32,8568	**32,9608**	31,8689	31,0983	32,9464	32,8968	31,6510	31,1699
Camera	32,6702	32,4009	32,0136	31,1396	**32,8642**	32,4612	32,1806	31,3401
Camera	32,6046	32,5907	32,3543	31,6692	**32,7421**	32,5864	32,4814	31,9388
Car	**31,4145**	30,4801	31,1243	30,2227	31,3523	30,3956	31,0602	31,1699
Car	**31,5073**	30,9908	31,2452	30,6201	31,3067	30,7715	30,9149	30,4667
Car	**31,3631**	30,9573	31,1335	30,1947	31,2420	30,9400	31,1618	30,2750
House	34,7831	34,5490	34,2851	32,8635	**34,7834**	34,4527	34,1532	33,0249
House	34,7418	33,9662	34,2738	33,6994	**34,8563**	34,1595	34,0750	31,5590
House	**34,8378**	34,4572	34,1893	32,6717	34,6574	34,3192	34,2635	32,5386
Lena	**33,3640**	31,7605	33,1867	31,2676	33,2871	32,7381	32,9576	31,5141
Lena	**33,4148**	32,7093	32,9334	32,4004	33,4001	32,3561	32,8861	31,8240
Lena	33,2812	32,9471	33,1102	31,2587	**33,5353**	33,0187	33,0532	31,2978
Peppers	33,3935	33,2182	33,1270	31,2097	**33,3970**	33,1317	33,1623	31,2310
Peppers	**33,4561**	32,1744	33,1463	31,7917	33,4020	32,2288	32,9756	31,6075
Peppers	33,4144	32,8588	33,1963	32,5958	**33,4276**	32,8173	33,1844	32,5901

Table 1 (continued)

Image	Dual NLM KL	NLM Pattern	Dual NLM Bhattacharyya	NLM Pattern	Dual NLM CS	NLM Pattern	Dual NLM Hellinger	NLM Pattern
Average	32,9457	32,4405	32,5478	31,5064	**32,9690**	32,4390	32,4543	31,3845
Median	32,9442	32,4009	32,7085	31,2587	**33,0133**	32,4612	32,6114	31,2978
Minimum	**31,3631**	30,4801	31,1243	30,1947	31,242	30,3956	30,8756	30,2292
Maximum	34,8378	34,549	34,2851	33,6994	**34,8563**	34,4527	34,2635	33,0249
Std.Dev.	1,033761468	1,07493735	1,036529439	0,940214	1,048707266	**1,083043**	1,065175	0,768234

When analyzing the results presented on Table 1, the Dual-NLM Cauchy-Schwarz filtering method showed the greatest final average when compared to the other filtering methods. However, there were also situation in which the Dual-NLM Kullback-Leibler filtering method allowed the obtention of satisfactory results.

Table 2 presents the results of evaluation done with the PSNR metrics for 7 types of different images and 3 different values of entrance as parameter, applied to the calculation of KL divergence.

Through the analysis of the results show on Table 2,the Dual-NLM KL filtering method presented, in great part of the images, the best result to reduce Graussian noises. However, the standard NLM filtering method also made it possible to obtain satisfactory results.

Furthermore, Table 3 show the results of evaluation done with the PSNR metrics for 7 types of different images and 3 different entrance values as parameter, applied to the calculation of the Bhattacharrya distance.

Table 2 PSNR obtained after filtering with the standard NLM filter and parametric NLM Kullback-Leibler filter for images corrupted by Gaussian noise with $\sigma = 10$

Using KL distance calculation

Image	List	t	f	h	Dual NLM KL	NLM Pattern
Airplane	50	3	1	1,2	32,0844	**32,1147**
Airplane	60	3	1	1,1	**31,8908**	31,5071
Airplane	70	3	2	1,7	**31,9330**	31,7151
Barbara	50	3	1	1,3	**32,9981**	32,8209
Barbara	60	3	2	1,9	**32,9058**	31,7102
Barbara	70	3	2	1,8	**32,9442**	32,3621
Camera	50	3	1	1,1	32,8568	**32,9608**
Camera	60	3	2	1,6	**32,6702**	32,4009
Camera	70	3	2	1,6	**32,6046**	32,5907
Car	50	4	2	1,9	**31,4145**	30,4801
Car	60	3	2	1,8	**31,5073**	30,9908
Car	70	2	2	1,9	**31,3631**	30,9573
House	50	4	1	1,2	**34,7831**	34,5490
House	60	4	2	1,8	**34,7418**	33,9662
House	70	4	2	1,7	**34,8378**	34,4572
Lena	50	4	2	1,9	**33,3640**	31,7605
Lena	60	4	2	1,8	**33,4148**	32,7093
Lena	70	3	2	1,7	**33,2812**	32,9471
Peppers	50	3	1	1,3	**33,3935**	33,2182
Peppers	60	3	2	2	**33,4561**	32,1744
Peppers	70	3	2	1,8	**33,4144**	32,8588
Average	60,0000	3,2381	1,7143	1,6238	**32,9457**	32,4405
Median	60,0000	3,0000	2,0000	1,7000	**32,9442**	32,4009
Minimum	50	2	1	1,1	**31,3631**	30,4801
Maximum	70	4	2	2	**34,8378**	34,549
Std.Dev.	8,3666	0,538958	0,46291	0,294796	1,033761468	**1,07493735**

Table 3 PSNR obtained after filtering with the standard NLM filter and parametric NLM Bhattacharrya filter for images corrupted by Gaussian noise with $\sigma = 10$

Using the Bhattacharrya Distance Calculation						
Image	List	t	f	h	Dual NLM Bhattacharyya	NLM Pattern
Airplane	50	2	2	0,8	**31,2266**	30,4348
Airplane	60	2	2	0,8	**31,2485**	31,0526
Airplane	70	4	3	0,8	**31,6718**	30,7826
Barbara	50	3	2	0,8	**32,7934**	30,7440
Barbara	60	3	2	0,8	**32,7085**	31,8039
Barbara	70	2	2	0,8	**32,6674**	32,1129
Camera	50	2	2	0,8	**31,8689**	31,0983
Camera	60	4	3	0,8	**32,0136**	31,1396
Camera	70	4	3	0,8	**32,3543**	31,6692
Car	50	3	2	0,8	**31,1243**	30,2227
Car	60	2	2	0,8	**31,2452**	30,6201
Car	70	4	3	0,9	**31,1335**	30,1947
House	50	4	2	0,8	**34,2851**	32,8635
House	60	3	2	0,8	**34,2738**	33,6994
House	70	4	3	0,8	**34,1893**	32,6717
Lena	50	3	2	0,8	**33,1867**	31,2676
Lena	60	3	2	0,8	**32,9334**	32,4004
Lena	70	4	3	0,9	**33,1102**	31,2587
Peppers	50	3	2	0,8	**33,1270**	31,2097
Peppers	60	2	2	0,8	**33,1463**	31,7917
Peppers	70	2	2	0,8	**33,1963**	32,5958
Average	60,0000	3,0000	2,2857	0,8095	**32,5478**	31,5064
Median	60,0000	3,0000	2,0000	0,8000	**32,7085**	31,2587
Minimum	50	2	2	0,8	**31,1243**	30,1947
Maximum	70	4	3	0,9	**34,2851**	33,6994
Std.Dev.	8,3666	0,83666	0,46291	0,030079	**1,036529439**	0,940214

By analyzing the results presented on Table 3,the Dual-NLM Bhattacharrya filtering method showed the best result for the reduction of Gaussian noise.

Table 4 provides the results of evaluation done with the PSNR metrics for 7 types of different images and 3 different entrance values as parameter, applied to the calculation of Cauchy- Schwarz divergence.

Analyzing the results presented on Table 4, the Dual-NLM de Cauchy-Schwarz filtering method had the greatest result to reduce Gaussian noises.

Table 5 shows the results of evaluation done with PSNR metrics for 7 types of different images and 3 different entrance values as parameter, applied to the calculation of the Hellinger distance.

Table 4 PSNR obtained after filtering with the standard NLM filter and parametric Cauchy-Schwarz NLM for images corrupted by Gaussian noise with $\sigma = 10$

Using Cauchy-Schwarz Distance Calculation						
Image	List	t	f	h	Dual NLM CS	NLM Pattern
Airplane	50	4	1	0,9	**32,1393**	31,8127
Airplane	60	3	1	0,9	**32,0395**	31,5161
Airplane	70	3	2	1,6	**31,9116**	31,6685
Barbara	50	3	1	1,1	**33,0523**	32,7391
Barbara	60	3	2	1,8	**32,9923**	31,7702
Barbara	70	3	2	1,6	**33,0133**	32,4392
Camera	50	4	1	0,8	**32,9464**	32,8968
Camera	60	3	2	1,6	**32,8642**	32,4612
Camera	70	3	2	1,3	**32,7421**	32,5864
Car	50	4	2	1,8	**31,3523**	30,3956
Car	60	3	2	1,7	**31,3067**	30,7715
Car	70	2	2	1,7	**31,2420**	30,9400
House	50	4	1	1	**34,7834**	34,4527
House	60	4	1	0,9	**34,8563**	34,1595
House	70	4	2	1,6	**34,6574**	34,3192
Lena	50	4	1	0,9	**33,2871**	32,7381
Lena	60	3	2	1,7	**33,4001**	32,3561
Lena	70	3	2	1,6	**33,5353**	33,0187
Peppers	50	3	1	1,1	**33,3970**	33,1317
Peppers	60	3	2	1,8	**33,4020**	32,2288
Peppers	70	3	2	1,6	**33,4276**	32,8173
Average	60,0000	3,2857	1,6190	1,3810	**32,9690**	32,4390
Median	60,0000	3,0000	2,0000	1,6000	**33,0133**	32,4612
Minimum	50	2	1	0,8	**31,242**	30,3956
Maximum	70	4	2	1,8	**34,8563**	34,4527
Std.Dev.	8,3666	0,560612	0,497613	0,366905	1,048707266	**1,083043**

Analyzing the results presented on Table 5, the Dual-NLM Hellinger filtering method showed the best result for the reduction of Gaussian noises when compared to the traditional NLM method.

Table 6 shows the results of evaluation done with the PSNR metrics for the 7 images. Besides that, the results presented were considered for each filter: BM3D, Bilateral, Wiener, Wavelet, Anisotropic Diffusion and Total Variation.

Analyzing the results presented on Table 6, the BM3D and Wiener filtering methods shows the greatest results for reduction of Gaussian noises in comparison to the other methods.

To illustrate the difference between those methods, Figs. 2, 3, 4, 5, 6, 7 and 8, show the visual results for the Airplane, Barbara, Camera, Car, House, Lena and Peppers images.

Table 5 PSNR obtained after filtering with the standard NLM filter and parametric Hellinger NLM filter for images corrupted by Gaussian noise with $\sigma = 10$

Using the Hellinger Distance Calculation						
Image	List	t	f	h	Dual NLM Hellinger	NLM Pattern
Airplane	50	4	1	0,9	**32,1393**	31,8127
Airplane	60	3	1	0,9	**32,0395**	31,5161
Airplane	70	3	2	1,6	**31,9116**	31,6685
Barbara	50	3	1	1,1	**33,0523**	32,7391
Barbara	60	3	2	1,8	**32,9923**	31,7702
Barbara	70	3	2	1,6	**33,0133**	32,4392
Camera	50	4	1	0,8	**32,9464**	32,8968
Camera	60	3	2	1,6	**32,8642**	32,4612
Camera	70	3	2	1,3	**32,7421**	32,5864
Car	50	4	2	1,8	**31,3523**	30,3956
Car	60	3	2	1,7	**31,3067**	30,7715
Car	70	2	2	1,7	**31,2420**	30,9400
House	50	4	1	1	**34,7834**	34,4527
House	60	4	1	0,9	**34,8563**	34,1595
House	70	4	2	1,6	**34,6574**	34,3192
Lena	50	4	1	0,9	**33,2871**	32,7381
Lena	60	3	2	1,7	**33,4001**	32,3561
Lena	70	3	2	1,6	**33,5353**	33,0187
Peppers	50	3	1	1,1	**33,3970**	33,1317
Peppers	60	3	2	1,8	**33,4020**	32,2288
Peppers	70	3	2	1,6	**33,4276**	32,8173
Average	60,0000	3,2857	1,6190	1,3810	**32,9690**	32,4390
Median	60,0000	3,0000	2,0000	1,6000	**33,0133**	32,4612
Minimum	50	2	1	0,8	**31,242**	30,3956
Maximum	70	4	2	1,8	**34,8563**	34,4527
Std.Dev.	8,3666	0,560612	0,497613	0,366905	1,048707266	**1,083043**

By doing that, it is possible to observe that there is a significant different regarding the level of residual noise in the images that were filtered by the Dual NLM Cauchy-Schwar method. Furthermore, the variances of the as Cauchy- Schwar and Kullback-Leibler non-local divergences offer a greater relation between denoising and edge preservation.

From the results obtained on Tables 1 and 6,with the evaluation of the filtering methods, it is possible to observe that both filters, Cauchy-Schwarz and KL, respectively, presented a high value of percentage regarding the evaluation of PSNR. In this context, in great part of the results, the usage of the Dual NLM of Cauchy-Schwarz filter proved to be satisfactory for the application on images that were degraded by Gaussian noise. However,

Table 6 BM3D, Bilateral, Wiener, Wavelet, Anisotropic Diffusion and Total Variation filters, for images corrupted by Gaussian noise with $\sigma = 10$

Image	BM3D	Bilateral	Wiener	Wavelet	Anis. Dif.[1]	T. Var.[2]
Airplane	**31,5459**	23,5558	31,1822	30,2621	30,5749	30,1123
Barbara	**32,7356**	22,3183	31,2905	29,9925	31,1407	30,5626
Camera	28,7898	22,7394	**31,9659**	29,1573	31,1109	30,7823
Car	**30,8849**	22,6322	30,4537	29,6981	30,0906	29,2298
House	**34,0587**	23,6012	32,2004	30,3949	32,0037	32,5141
Lena	**33,0817**	24,0036	31,2352	30,5116	31,5479	31,2257
Peppers	**32,9906**	23,9805	31,1313	30,6854	32,0584	32,1650
Average	**32,0125**	23,2616	31,3513	30,1003	31,2182	30,9417
Median	**32,7356**	23,5558	31,2352	30,2621	31,1407	30,7823
Minimum	28,7898	22,3183	**30,4537**	29,1573	30,0906	29,2298
Maximum	**34,0587**	24,0036	32,2004	30,6854	32,0584	32,5141
Std.Dev.	**1,764685**	0,686527	0,577142442	0,530269	0,722782189	1,14289

[1] Anisotropic Diffusion filter

[2] Total Variation filter

the variances found regarding the PSNR evaluation of the Dual NLM KL filter were significant, what allowed the achieving results that also showed to be satisfactory regarding the filtering methods of traditional NLM and the other filtering methods that were applied in this work.

6 Conclusions and final remarks

The process of denoising in images degraded by Graussian noises is a challenge task in the computational vision area, since the recent methods of image filtering are based on functions of spatial domain and frequency domain are not efficient. The filtering methods that are based on spatial domain are usually the best option to solve issues that deal with the impulsive Gaussian noise. Given this, in this article, a Dual Non-Local Means filter was presented, which combines the characteristics of classification order, non-local strategies and mathematical models that are seen on information theory.

In this scenario, the Dual Non-Local Means filtering method can be considered a philosophy of the NLM filter to solve problems of images that are degraded by impulsive noises. The variances of the mathematical metrics based on the concept of the information theory unify two types of distinct behaviors but essential to deal with the Gaussian noise. In this case, the behaviors of the proposed method have the Non-Local Means and the Dual Non-Local Means filters as approaches. In this way, several computational experiments were done during the course of this work, with multiple digital images degraded by Gaussian noise, which showed that the proposed method can generate, on average, significantly better outcomes in terms of PSNR when compared to the continuous application of the standard NLM, Total Variation, BM3D, Anisotropic Diffusion, Wiener, Wavelet and Bilateral filters.

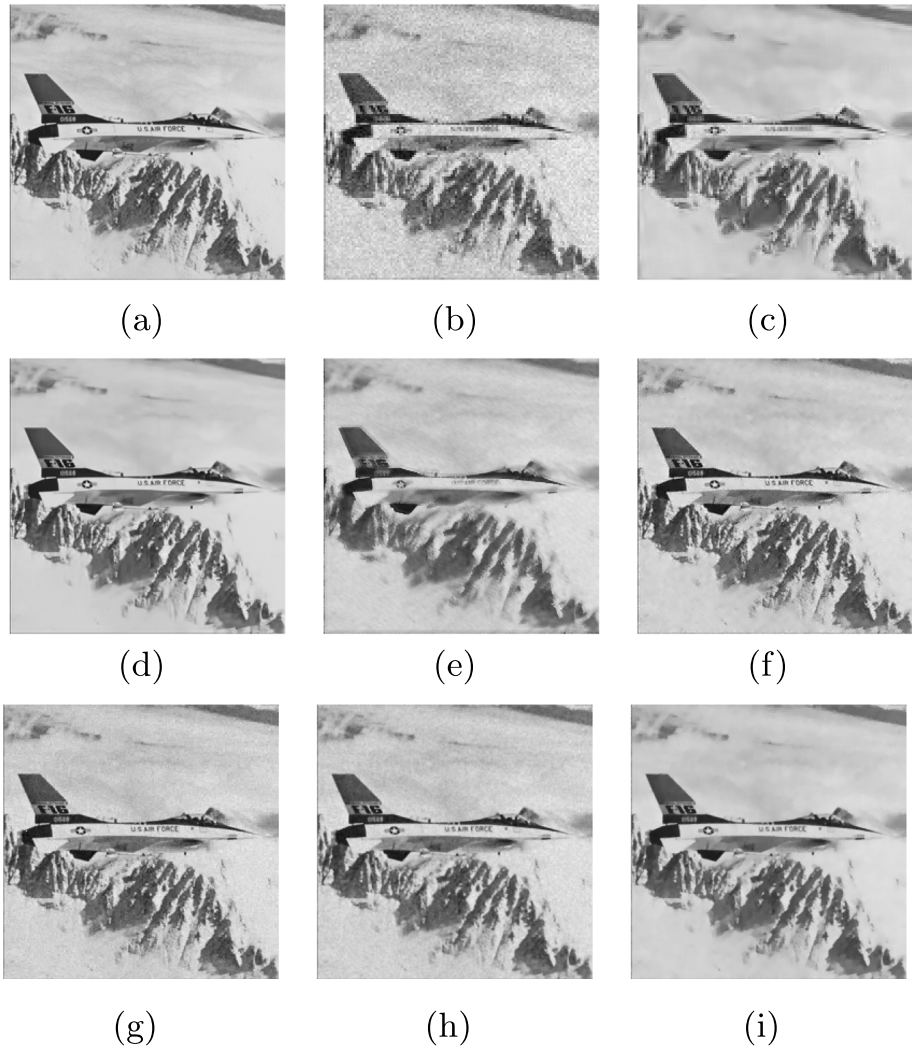

Fig. 2 Example image result (a), noisy image (b), Dual NLM Cauchy-Schwarz filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

Finally, future works can include the usage of different families of entropy, such as Renyi's and Sharma-Mittal's entropies. Methods that are applied to solve problems of dimensionality reduction, for example, PCA, can be used to better understand a more compact and significant representation for patches inside the browser window. Besides that, methods like the Parametric PCA, the ISOMAP and the Laplacian Eigemaps can be applied before the calculation of the Euclidian distances as a way of asymptotically guarantee the greatest similarity measures.

Fig. 3 Example image result (a), noisy image (b), Dual NLM Cauchy-Schwarz filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic diffusion filter (h) and Total Variation filter (i)

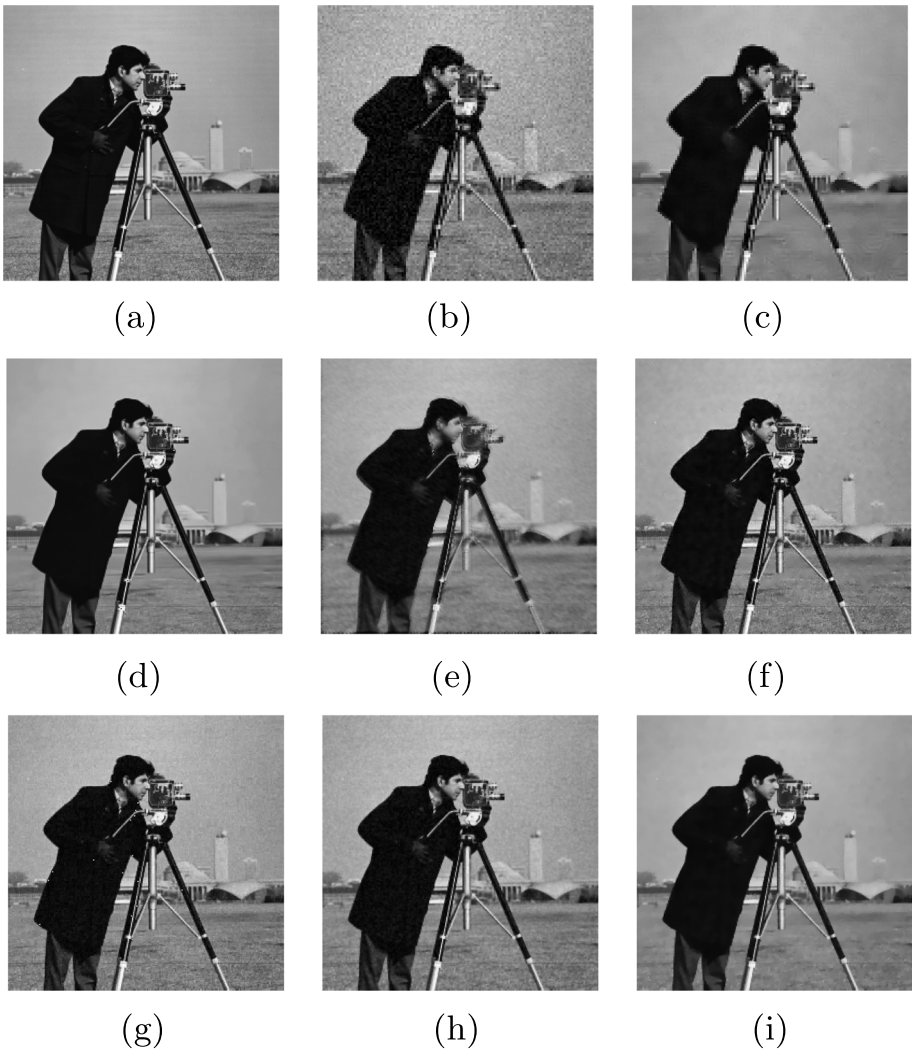

Fig. 4 Example image result (a), noisy image (b), Dual NLM Cauchy-Schwarz filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

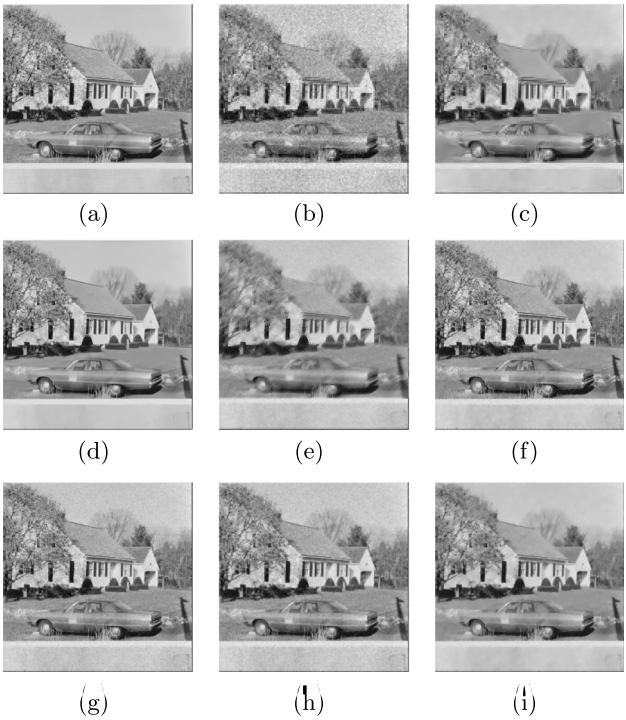

Fig. 5 Example image result (a), noisy image (b), Dual NLM Kullback-Leibler filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

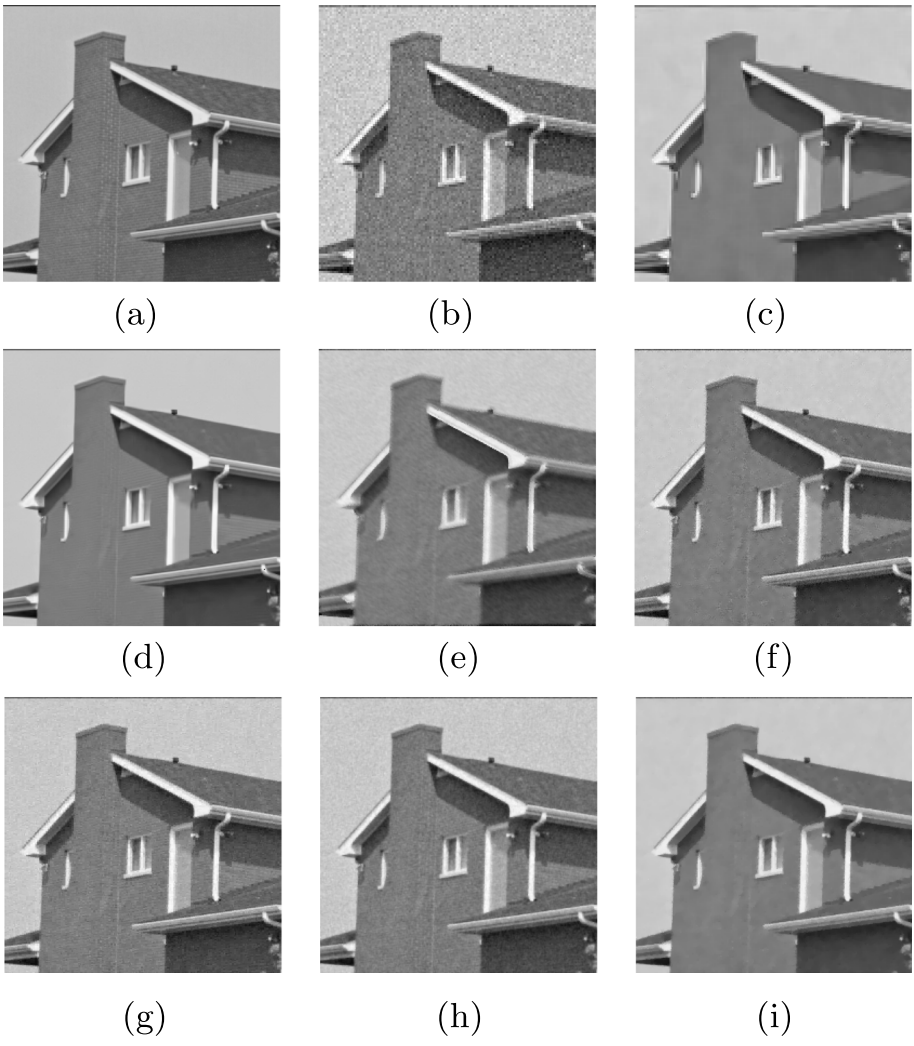

Fig. 6 Example image result (a), noisy image (b), Dual NLM Kullback-Leibler filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

Fig. 7 Example image result (a), noisy image (b), Dual NLM Kullback-Leibler filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

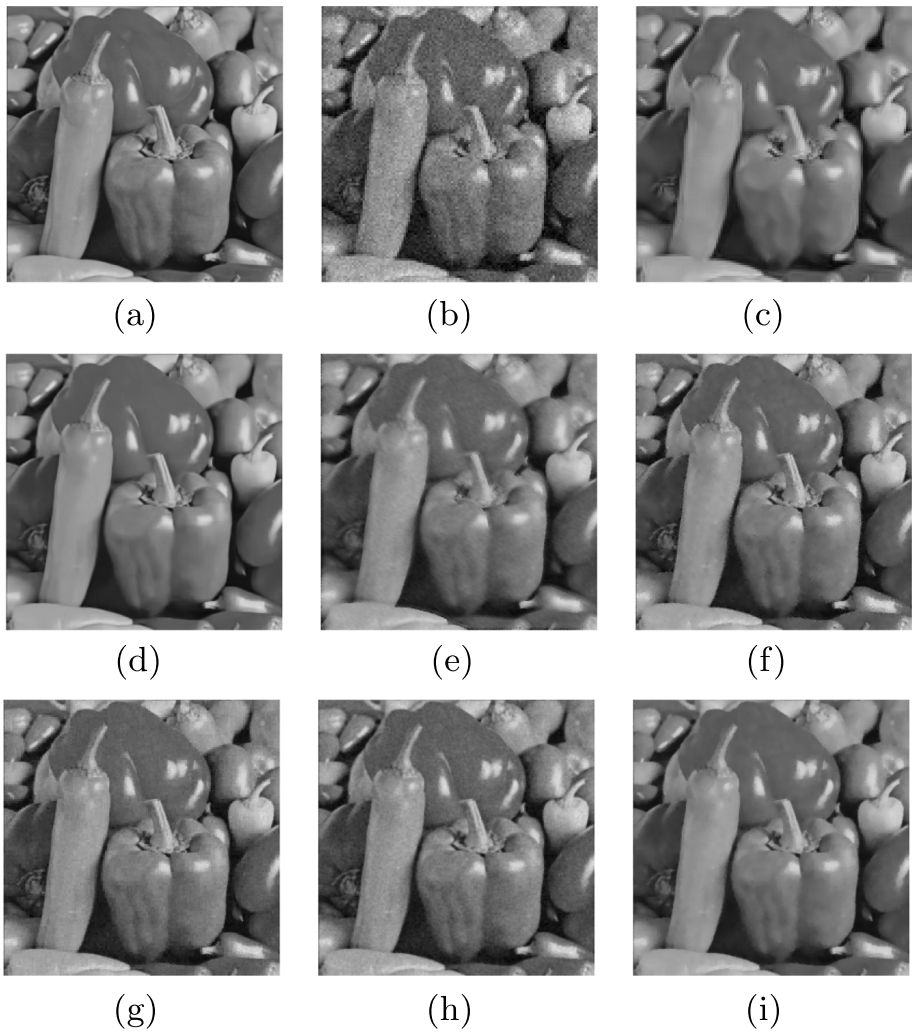

Fig. 8 Example image result (a), noisy image (b), Dual NLM Kullback-Leibler filter (c), BM3D filter (d), Bilateral filter (e), Wiener filter (f), Wavelet filter (g), Anisotropic Diffusion filter (h) and Total Variation filter (i)

Acknowledgments This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Data Availability Statement Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of Interests The authors declare that they have no conflict of interest.

References

1. Amer A, Dubois E (2005) Fast and reliable structure-oriented video noise estimation. IEEE Trans Circuits Syst Video Technol. https://doi.org/10.1109/TCSVT.2004.837017
2. Bindilatti AA, Mascarenhas NDA (2013) A nonlocal poisson denoising algorithm based on stochastic distances. IEEE Signal Process Lett. https://doi.org/10.1109/LSP.2013.2277111
3. Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. In: Proceedings - 2005 IEEE computer society conference on computer vision and pattern recognition, CVPR 2005. https://doi.org/10.1109/CVPR.2005.38
4. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans Image Process. https://doi.org/10.1109/TIP.2007.901238
5. Diaconis P, Zabell SL (1982) Updating subjective probability. J Am Stat Assoc. https://doi.org/10.1080/01621459.1982.10477893
6. Hasan M, El-Sakka MR (2018) Improved BM3D image denoising using SSIM-optimized Wiener filter. Eurasip Journal on Image and Video Processing 2018(1). https://doi.org/10.1186/s13640-018-0264-z
7. Hoang HG, Vo BN, Vo BT, Mahler R (2014) The Cauchy-Schwarz divergence for poisson point processes. In: IEEE Workshop on statistical signal processing proceedings. https://doi.org/10.1109/SSP.2014.6884620
8. Jain P, Tyagi V (2016) A survey of edge-preserving image denoising methods. Inf Syst Front. https://doi.org/10.1007/s10796-014-9527-0
9. Jenssen R, Erdogmus D, Hild KE, Principe JC, Eltoft T (2005) Optimizing the Cauchy-Schwarz PDF distance for information theoretic, non-parametric clustering. In: Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). https://doi.org/10.1007/11585978_3
10. Jenssen R, Principe JC, Erdogmus D, Eltoft T (2006) The Cauchy-Schwarz divergence and Parzen windowing: connections to graph theory and Mercer kernels. J Franklin Inst. https://doi.org/10.1016/j.jfranklin.2006.03.018
11. Jin C, Luan N (2020) An image denoising iterative approach based on total variation and weighting function. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-020-08871-0
12. Kailath T (1967) The divergence and Bhattacharyya distance measures in signal selection. IEEE Trans Commun Technol. https://doi.org/10.1109/TCOM.1967.1089532
13. Kuan DT, Sawchuk AA, Strand TC, Chavel P (1985) Adaptive noise smoothing filter for images with signal-dependent noise. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/TPAMI.1985.4767641
14. Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22(1):79–86
15. Kundu R, Chakrabarti A, Lenka P (2022) A novel technique for image denoising using non-local means and genetic algorithm. National Academy Science Letters. https://doi.org/10.1007/s40009-021-01052-z
16. Levada ALM (2021) Non-local medians filter for joint Gaussian and impulsive image denoising. In: Proceedings - 2021 34th SIBGRAPI conference on graphics, patterns and images, SIBGRAPI 2021, pp 152–159. https://doi.org/10.1109/SIBGRAPI54419.2021.00029
17. Mahler RPS (1998) Global posterior densities for sensor management. In: Masten MK, Stockum LA (eds) Acquisition, tracking, and pointing XII, vol 3365. SPIE, pp 252–263, https://doi.org/10.1117/12.317518. International Society for Optics and Photonics
18. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/34.192463
19. Milanfar P (2013) A tour of modern image filtering. IEEE Signal Process Mag 30(1):106–123. https://doi.org/10.1109/MSP.2011.2179329
20. Patidar P, Gupta M, Nagawat AK (2010) Image denoising by various filters for different noise image de-noising by various filters for different noise Sumit Srivastava. Int J Comput Applic 9(4):975–8887
21. Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. IEEE Trans Pattern Anal Mach Intell. https://doi.org/10.1109/34.56205
22. Petkova L, Draganov I (2020) Noise adaptive Wiener filtering of images. In: 2020 55th International scientific conference on information, communication and energy systems and technologies, ICEST 2020 - proceedings. https://doi.org/10.1109/ICEST49890.2020.9232887
23. Riya, Gupta B, Lamba SS (2021) An efficient anisotropic diffusion model for image denoising with edge preservation. Comput Math Appl. https://doi.org/10.1016/j.camwa.2021.03.029
24. Rohit V, Ali J (2013) A comparative study of various types of image noise and efficient noise removal techniques. Int J Adv Res Comput Sci Softw Eng 3(10):2277–128

25. Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. Physica D. https://doi.org/10.1016/0167-2789(92)90242-F
26. Salehi H, Vahidi J (2021) A novel hybrid filter for image despeckling based on improved adaptive wiener filter, bilateral filter and wavelet filter. International Journal of Image and Graphics. https://doi.org/10.1142/S0219467821500364
27. Shao L, Yan R, Li X, Liu Y (2014) From heuristic optimization to dictionary learning: a review and comprehensive comparison of image denoising algorithms. IEEE Transactions on Cybernetics. https://doi.org/10.1109/TCYB.2013.2278548
28. Spurek P, Pałka W (2016) Clustering of Gaussian distributions. In: Proceedings of the international joint conference on neural networks. https://doi.org/10.1109/IJCNN.2016.7727627
29. Tomasi C, Manduchi R (1998) Bilateral filtering for gray and color images. In: Proceedings of the IEEE international conference on computer vision. https://doi.org/10.1109/iccv.1998.710815
30. Yahya AA, Tan J, Su B, Hu M, Wang Y, Liu K, Hadi AN (2020) BM3D image denoising algorithm based on an adaptive filtering. Multimedia Tools and Applications. https://doi.org/10.1007/s11042-020-08815-8
31. Yue H, Sun X, Yang J, Wu F (2014) CID: combined image denoising in spatial and frequency domains using web images. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition. https://doi.org/10.1109/CVPR.2014.375
32. Zhang Y, Sun J (2020) An improved BM3D algorithm based on anisotropic diffusion equation. Math Biosci Eng. https://doi.org/10.3934/mbe.2020269