Journal of Mathematical Imaging and Vision 17: 15–25, 2002

A Filter To Remove Gaussian Noise by Clustering the Gray Scale

ZAKARIA ATAE-ALLAH & JOSÉ MARTÍNEZ AROZA*
Dept. Matemática Aplicada, Universidad de Granada, Spain, Facultad de Ciencias, Avda. Fuentenueva s/n, 18071 Granada, Spain
email: jmaroza@ugr.es

Abstract. An algorithm to suppress Gaussian noise is presented, based on clustering (grouping) gray levels. The histogram of a window sliding across the image is divided into clusters, and the algorithm outputs the mean level of the group containing the central pixel of the window. This filter restores well the majority of noisy pixels, leaving only few of them very deviated, that can be finally restored with a common filter for impulsive noise, such as a median filter. In this paper the clustering filter *CF* is described, analysed and compared with other similar filters.

Keywords: image restoration, gaussian noise, histogram clustering

1. Introduction

Digital images suffer during acquisition or transmission diverse types of distortions, standing out for its incidence and interest the additive Gaussian noise, which appears usually when transmitting images by means of communication channels like a vidicom or a radar. Also, in the field of medical imaging, magnetic resonance blurred images are considered as corrupted by Gaussian noise. Mostly, noise is characterised by statistical parameters; in particular, Gaussian noise is typified by a Gaussian distribution function. Given an original image I, the image X resulting from I by adding Gaussian noise is given by

$$X_i = I_i + \epsilon_i,$$

where the subscript i runs over all locations (pixels) in I, and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$ are independent and identically distributed random variables.

Many algorithms—called filters—are used to minimize the effect of this noise. A filter can be viewed as an operator F that transforms an image X into another image Y more adequate for interpretation or later processing, $Y = F(X)$.

In the sequel, the following notations are used.

Original Image $I = (I_i)$; it represents the real image in an ideal form, with no distortion; for each i, I_i stands for the gray level of the pixel located at position i in the image I.

Input Image $X = (X_i)$; this is the image coming from the original, but altered by some kind of noise; in practice, this is the only image at hand, and it is the input of the filter.

Output Image $Y = (Y_i)$; this is the image resulting from the filtering process; if restoring is desired, then it should be as close as possible to the original I.

To evaluate the quality of a filter is not a trivial task. It is difficult to find a good method besides the mere visual inspection. Nevertheless, there are some measures used by numerous authors. The most popular (and used in this paper) are the following.

- The mean absolute error *MAE*, defined as

$$MAE = \frac{\sum_i (I_i - Y_i)}{S}$$

- The mean square error *MSE*, given by

* To whom correspondence should be addressed.

$$MSE = \frac{\sum_i (I_i - Y_i)^2}{S}$$

where S is the size of the image in pixels.

There exist plenty of filters intended to elliminate noise in images. They can be classified into two groups: linear and non linear filters.

Linear Filters

A filter F is said to be linear if for all $a, b \in \mathbb{R}$ it verifies the property

$$F(aX + bX') = aF(X) + bF(X').$$

The most simple linear filter is the mean filter. The standard mean filter with a window V of odd size $2N+1$ is defined as [4]

$$Y_i = \frac{1}{2N+1} \sum_{r \in V} X_{i+r}.$$

One of the properties of the mean filter is its great efficiency to elliminate impulsive noise; but when dealing with Gaussian noise it does not show good results, since it tends to blur edges and lose details. Linear filters in general show a weakness in restoring images with Gaussian noise, and are seldom used to elliminate this kind of noise.

Non-linear Filters

Filters not having the linearity property above are said to be non-linear [12]. Many of them have proven to be adequate for Gaussian noise; among them, the classical median filter [3] stands out because of its simplicity. The median filter is defined by

$$Y_i = \mathrm{med}\{X_{i+r} : r \in V\}$$

where med means the median of the set. This filter is a robust smoother and its results for Gaussian noise are reasonably good.

Non-linear filters specialized in Gaussian noise ellimination have appeared in recent papers [6,10,14]. One is the sigma filter [10], described as follows.

Suppose that the standard deviation σ of Gaussian noise is known, then center a window V on each position i in the image X and define

$$E_i = \{j \in V_i : |X_i - X_j| \leq 2\sigma\},$$

where V_i is the set of pixels belonging to the window centered in pixel i. Then the output of the filter is

$$Y_i = \frac{\sum_j L(X_i - X_j) X_j}{\sum_j L(X_i - X_j)},$$

where

$$L(X_i - X_j) = \begin{cases} 1 & \text{if } j \in E_i, \\ 0 & \text{if not.} \end{cases}$$

Based on the same idea, but with a continuous approach, the non-linear gaussian filter [6,14], is given by

$$Y_i = \frac{\sum_j L(X_i - X_j) \exp(-(X_i - X_j)^2/(2g^2) X_j}{\sum_j L(X_i - X_j) \exp(-(X_i - X_j)^2/(2g^2))}.$$

where g is a parameter to be determined depending on the intensity σ of noise. The aim of this filter is to reduce noise and to preserve edges and details at the same time. It can be shown that this filter gives good results in regions containing edges, but when dealing with homogeneous regions it shows a certain weakness and produces blotches.

The filter presented in this paper is based on two main concepts: grouping and separation. Scattered gray levels of a homogeneous region should be joined into a unique color[1] as close as possible to the original, while gray levels of different regions should remain separated. This procedure is known as clustering.

Cluster Analysis

Cluster analysis is a major task within statistical data analysis, that tries to evaluate the interaction between elements of a set of data and to organize them into a certain number of natural and homogenous groups

[1] Although this paper deals only with gray level images, for brevity the term ‘color’ will be used sometimes instead of ‘gray level’.

(clusters), where each element shows a high degree of similarity with other elements in the same cluster, and differs as much as possible with respect to elements in other clusters [2,13].

Cluster analysis gives the researcher prior information about the data at hand. Its results can be used to make hypotheses about the data, to classify new data, to check the homogeneity of the set of data, or even to compress the data for transmission or storage [5,11].

The most important requirement to make a good cluster analysis is to stablish an adequate criterion to obtain the subsets. More precisely, we have to define what mathematical properties of the elements of the set should be used, and in what manner the clusters should be identified.

There exist several methods for clustering [8], the most remarkable being of a hierarchical type. This group of methods impose to the data a hierarchical structure which consists in a sequence of groupings. Hierarchical methods can be classified into two groups.

Agglomerative Techniques. They begin with one cluster for each element, and the grouping sequence evolves by fusion until a stopping criterion is reached.

Division Techniques. The procedure begins with a unique cluster containing all elements, and the sequence evolves by succesive division until a stopping criterion is reached.

In both cases, new clusters are formed by means of relocation of old clusters, based on some measure of similarity or difference.

Clustering techniques are widely utilized in pattern recognition and image analysis. Many approaches to image segmentation carry out clustering of image pixels, based on their similarity or attributes such as gray level, color, texture, or gradient [7].

In the present paper, clustering methods are utilized for grouping local histograms, with a technique based on searching the best separation point between gray levels.

2. Clustering Gray Levels by Local Separation

In most cases, dividing strategies with binary data are used and the procedure utilized to divide into subgroups is based on the identification of a feature that maximizes the difference between the resulting groups.

2.1. Basic Definitions

Let $P = \{f_0, f_1, f_2, \ldots, f_K\}$ be a histogram of a region with a scale of gray levels $\mathcal{K} = \{0, 1, \ldots, K\}$. By convention, this scale will be broadened up to $\mathbb{Z}$, by enlarging P with null frequencies.

Definition 1. A gray level k is said to be proper in P iff $f_k > 0$.

Definition 2. The range $r(P)$ is defined as the maximum (chromatic) distance between proper gray levels in P,

$$r(P) = \max_{f_i>0, f_j>0} |j - i|.$$

Definition 3. The local mean $\mu_{r,s}(P)$ of P is defined as

$$\mu_{r,s}(P) = \frac{\sum_{r \le i \le s} i f_i}{\sum_{r \le i \le s} f_i} \qquad \forall s \ge r.$$

Definition 4. The separation measure $\delta_i^d(P)$ of P at level i is defined as

$$\delta_i^d(P) = \begin{cases} |\mu_{i+1,i+d}(P) - \mu_{i-d+1,i}(P)| & \text{if both exist} \\ 2d - 1 \end{cases}$$

otherwise, where $d \in \mathbb{R}_+$ is an adjustable parameter called scattering limit.

One of the immediate properties of this separation measure is that

$$1 \le \delta_i^d(P) \le 2d - 1 \quad \forall i \in \mathbb{Z}, \quad \forall d \in \mathbb{R}_+.$$

2.2. The CLOSE Algorithm

Color scattering can be represented by clustering a histogram. This will be applied in real images in order to reduce the effect of the gaussian noise. The filter presented in this paper is based on a clustering of the local histogram. A simple description of the clustering algorithm called *CLOSE* (Clustering by Local Separation) follows.

- Let P be a local histogram (supposed broadened to an infinite scale of gray levels by means of null frequencies).
- If $r(P) \leq d$, then the algorithm ends with no separation.
- Let $J = \{j \ : \ \delta_j^d(P) = \max_i \delta_i^d\}$ and $j_0 = \min J$. If $J = \emptyset$ or $\delta_{j_0}^d < d$, then the algorithm ends with no separation. Otherwise, P is divided into two parts, $\{f_0, \ldots, f_{j_0}\}$ and $\{f_{j_0+1}, \ldots, f_K\}$ and the algorithm is recursively applied to each part (see Fig. 1).

This algorithm outputs the histogram P divided into a series of subhistograms called *clusters*. Each cluster corresponds to a cohesionated set of gray levels.

The scattering of the original level can be more or less severe, depending on diverse factors. Hence, a parameter to regulate the clustering is needed, this being the scattering limit d. This parameter fulfills perfectly the requirements, and it also gives to the *CLOSE* method a depth level to cut the tree that represents the hierarchical structure of every dividing method used to obtain the proper number of groups.

2.3. Breakdown Probability

Although histogram clustering can be an adequate tool to process images with scattered colors, it is necessary to assess the success of the algorithm in assigning separate clusters to distinct colors. In order to evalute this aspect, we use the so-called breakdown probability. It is the probability that the clustering algorithm produces incorrect clusters depending on the used parameters. Since this study is mathematically complex, firstly the particular and simple case and thereinafter the general case will be studied.

2.3.1. Breakdown Probability in a Window with two Pixels.

Let I be a homogeneous image with a unique level k corrupted with gaussian noise $\mathcal{N}(0, \sigma^2)$, and let V be a window with two pixels arbitrary located in the image. The gray level of the pixels is represented by two independent random variables X_1 and X_2 such that:

$$X_1 \sim \mathcal{N}(k, \sigma^2) \qquad X_2 \sim \mathcal{N}(k, \sigma^2).$$

Let P be the observed histogram of the window V. The objective is to calculate the probability of the fact that the *CLOSE* algorithm gives from P more than one cluster depending on the d/σ ratio.

Theorem 1. *The probability Q_d of more than one cluster in the histogram of the window V (breakdown probability) is*

$$Q_d(X_1, X_2) = 1 - \frac{1}{\sqrt{\pi}\sigma} \int_0^d \exp\left(-\frac{x^2}{4\sigma^2}\right) dx$$

Proof: The probability density functions of the pixels 1 and 2 respectively are given by

$$f(x_1) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_1 - k)^2}{2\sigma^2}\right),$$
$$f(x_2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_2 - k)^2}{2\sigma^2}\right).$$

The independence of the two random variables results from the degradation process suffered by the image.

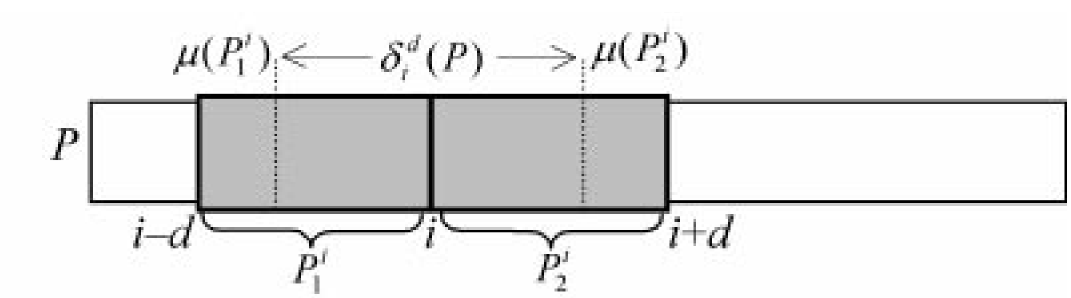


Figure 1. An illustration of the clustering process.

The clustering algorithm acts on the histogram of the window V, and there will be an unique group if the difference between means is less than the scattering limit. By supposing $X_2 \geq X_1$ without lose of generality, the breakdown probability is

$$Q_d(X_1, X_2) = P(X_2 - X_1 \geq d \mid X_2 \geq X_1) = \frac{P(0 \leq X_2 - X_1 \leq d)}{P(X_2 \geq X_1)}.$$

Due to the properties of the linear combinations of normally distributed random variables, we have

$$X_2 - X_1 \sim \mathcal{N}(0, 2\sigma^2)$$

and therefore,

$$P(X_2 \geq X_1) = \frac{1}{2}, \quad \text{and} \quad P(0 \leq X_2 - X_1 \leq d) = \frac{1}{2\sqrt{\pi}\sigma} \int_0^d \exp\left(-\frac{x^2}{4\sigma^2}\right) dx,$$

that gives the desired result. Figure 2 represents the breakdown probability of a two-pixel window depending on the d/σ ratio.

2.3.2. Breakdown Probability of a Window with any size. Now let V be a window of size n; there are n random variables $X_1, X_2, \ldots, X_n$ independent and identically distributed, each one with a normal distribution with mean k and variance σ^2. The joined distribution has the probability density function

$$f(x_1, x_2, \ldots, x_n) = \frac{1}{(\sqrt{2\pi}\sigma)^n} \exp\left(\frac{-\sum_{i=1}^n (x_i - k)^2}{2\sigma^2}\right).$$

Theorem 2. *Under the described conditions, the breakdown probability Q_d is given by*

$$Q_d(X_1, X_2, \ldots, X_n) = 1 - \frac{n}{(\sqrt{2\pi}\sigma)^n} \int_{-\infty}^{+\infty} \int_{x_1}^{+\infty} \cdots \int_{x_{n-2}}^{+\infty} \int_{x_{n-1}}^{M} \exp\left(-\sum_{i=1}^{n} \frac{(x_i - k)^2}{2\sigma^2}\right) dx_n dx_{n-1} \cdots dx_1,$$

where

$$M = \min_{r=1,2,\ldots,n-1} \left\{ (n-r)\left(d + \sum_{i=1}^{r} \frac{X_i}{r}\right) - \sum_{i=r+1}^{n-1} X_i \right\}.$$

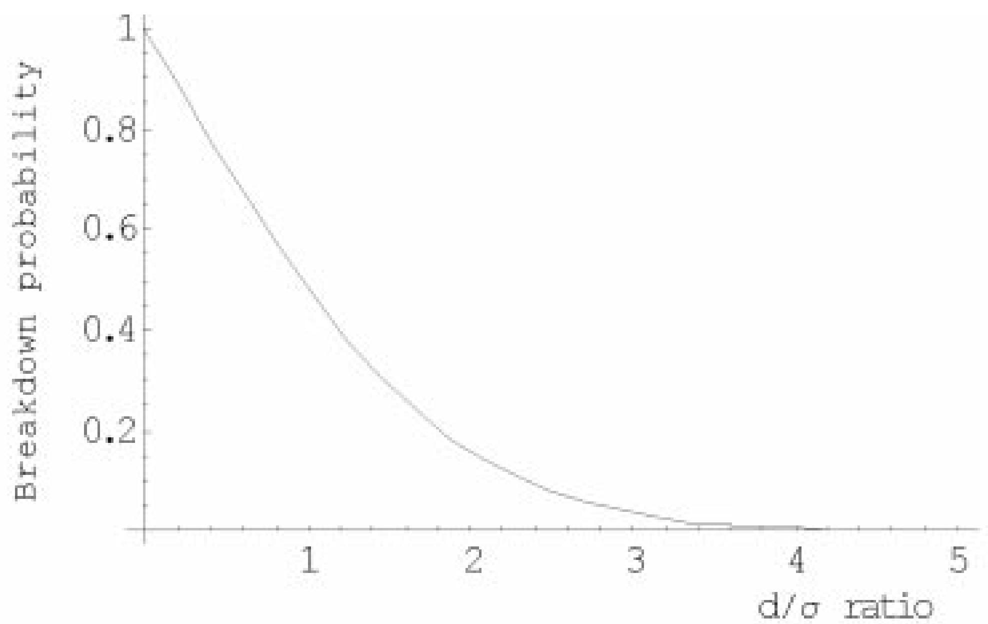


Figure 2. Breakdown probability of a window with two pixels.

Proof: The histogram of the window V has a unique cluster if there is no cut point, this is, that every difference between local means is less than the scattering limit d. For this, the following conditions must be satisfied (by supposing without losing generality that $X_1 \leq X_2 \leq \ldots \leq X_n$)

$$\frac{\sum_{i=r+1}^{n} X_i}{n-r} - \frac{\sum_{i=1}^{r} X_i}{r} \quad < \quad d \quad \forall r = 1, 2, \ldots, n-1,$$

$$\text{that is} \quad X_n \quad < \quad (n-r)\left(d + \sum_{i=1}^{r} \frac{X_i}{r}\right) - \sum_{i=r+1}^{n-1} X_i;$$

which means that X_n must verify

$$X_n < \min_{r=1,2,\ldots,n-1} \left\{ (n-r)\left(d + \sum_{i=1}^{r} \frac{X_i}{r}\right) - \sum_{i=r+1}^{n-1} X_i \right\} = M;$$

and finally an integration on the set

$$B = \{(X_1, X_2, \ldots, X_n) \in \mathbb{R} \ : -\infty < X_1 < +\infty, \\ X_1 < X_2 < +\infty, \ldots, X_{n-1} < X_n < M\}$$

allows to obtain the desired result.

Figure 3 shows the breakdown probability for several typical window sizes, depending on the d/σ ratio.

Given the little handiness of an expression with multiple integrals that can be nested up to 225 levels of depth for a 15 × 15 window, numerical values have been estimated by means of the Monte-Carlo method, using 1.000.000 of samples in all cases.

From the breakdown probabilities shown in Fig. 3, we can deduce that the larger the window, the more failures appear, and that there exists a point in each graph that can be considered as the optimal for the d/σ ratio.

Finally, the value of the scattering limit depends on the size of the used window and the intensity of the noise, and this value should be chosen so that the breakdown probability is as low as possible.

3. The Clustering Filter *CF*

In this section we present a new method to remove gaussian noise based on the clustering of the gray levels of local histograms.

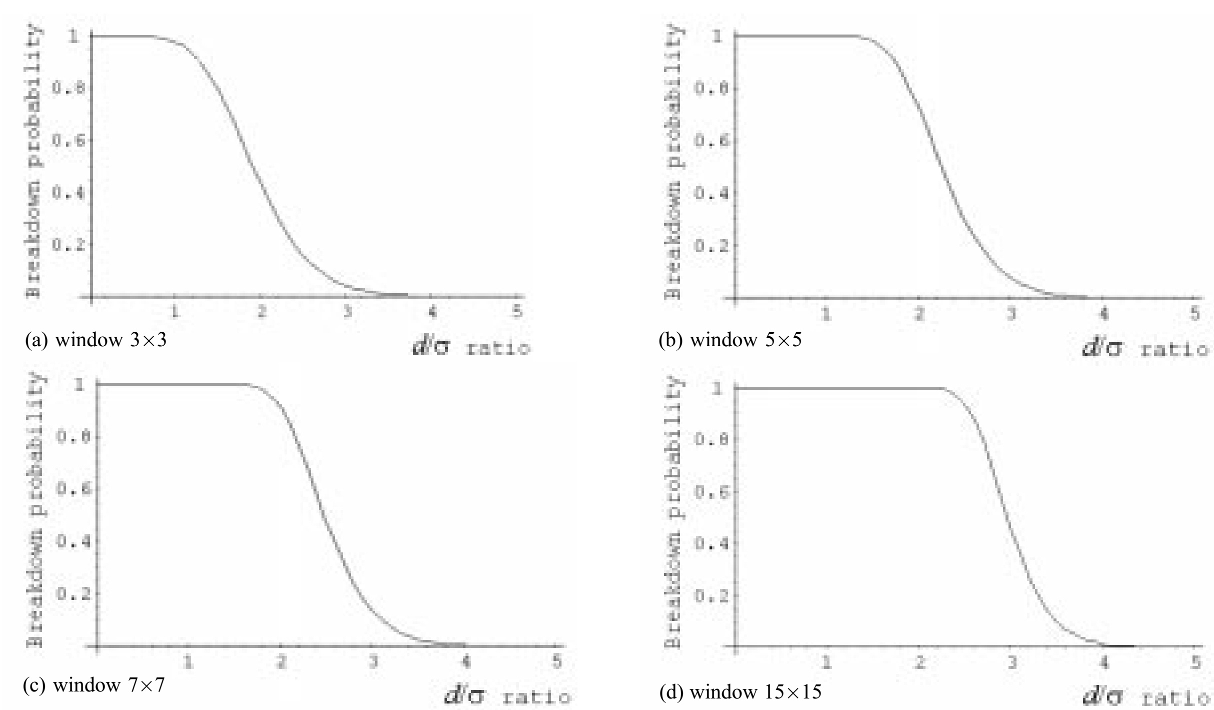


Figure 3. Breakdown probabilities obtained using the Monte-Carlo algorithm.

3.1. Definition of the CF Filter

The cluster filter CF_d is defined as follows.

Let i be a pixel in the source image, and let V_i be a sliding window centered on pixel i. The histogram of V_i is clustered by means of the *CLOSE* algorithm and the index set E_i is defined as

$$E_i = \{ j \in V_i \ : \ j \text{ belongs to the same cluster than } i \text{ in the clustered histogram.} \}$$

Then the output of the filter is

$$Y_i = \frac{\sum_j L(X_i - X_j) X_j}{\sum_j L(X_i - X_j)}$$

where

$$L(X_i - X_j) = \begin{cases} 1 & \text{if } j \in E_i \\ 0 & \text{otherwise.} \end{cases}$$

In other words, on each pixel i is centered a window V_i whose histogram is divided into clusters, and the output of the filter is the mean level of the cluster containing the central pixel. This procedure is illustrated in Fig. 4.

Since all this context is developed in a discrete domain, the final estimation of the mean of every cluster must be done by rounding off the computed average level:

$$R(Y_i) = E(Y_i + 0.5)$$

where E stands for the integer part.

This filter has several properties that back it. We can emphasize in particular its invariance for both resolution and intensity.

3.2. Properties

Definition 5. The normalized histogram of $P = \{f_0, \dots, f_k\}$ is

$$\bar{P} = \left(\frac{f_0}{n}, \frac{f_2}{n}, \dots, \frac{f_k}{n} \right), \quad \text{where} \quad n = \sum_{i=0}^{k} f_i.$$

It is evident that the relationship between histograms defined as

$$P_1 \cong P_2 \Longleftrightarrow \bar{P}_1 = \bar{P}_2$$

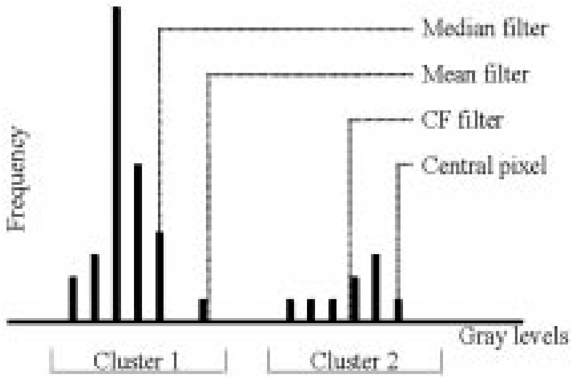


Figure 4. An illustration of the *CF* filter. The local histogram of the sliding window has been divided into two clusters. The gray level of the central pixel of the window lies in Cluster 2, so the output of the *CF* filter is the mean of Cluster 2. Both median and mean filters give outputs in the range of Cluster 1.

is an equivalence relationship. If $P_1 \cong P_2$ we will say that they are equivalent.

Clustering based on mean differences gives the same result for equivalent histograms.

Lemma 1. *If $P_1 \cong P_2$ and the same scattering limit d is used, then both give the same clustering.*

Proof. If $P_1 = \{f_0, \dots, f_K\}$ and $P_2 = \{g_0, \dots, g_K\}$ are two equivalent histograms, then the corresponding local means are equal,

$$\mu_{r,s}(P_1) = \mu_{r,s}(P_2) \qquad \forall r \geq s,$$

since

$$\mu_{r,s}(P_1) = \frac{\sum_{r \leq i \leq s} i f_i}{\sum_{r \leq i \leq s} f_i} = \frac{\sum_{r \leq i \leq s} i g_i \frac{\sum_{i=0}^{K} f_i}{\sum_{i=0}^{K} g_i}}{\sum_{r \leq i \leq s} g_i \frac{\sum_{i=0}^{K} f_i}{\sum_{i=0}^{K} g_i}} = \mu_{r,s}(P_2),$$

and therefore

$$\begin{aligned} \delta_i^d(P_1) &= |\mu_{i+1,i+d}(P_1) - \mu_{i-d+1,i}(P_1)| \\ &= |\mu_{i+1,i+d}(P_2) - \mu_{i-d+1,i}(P_2)| \\ &= \delta_i^d(P_2), \end{aligned}$$

and also if $\mu_{r,s}(P_1)$ does not exist then neither $\mu_{r,s}(P_2)$ exists.

Of course, $r(P_1) = r(P_2)$, and because of the equality for the separation measure we deduce that both histograms have the same cut points, hence the same clusters.

Theorem 3. (*invariance with respect to the resolution*) *The cluster filter CF_d is invariant with respect to the resolution. If an image is resampled under a different resolution with no gain or loss of details, then the filter CF_d gives the same output.*

Proof: A resampling of any window in the image with no gain or loss of details will give equivalent histograms, and then CF_d will give the same output.

Intensity changes are one of the possible perturbations that an image can undergo during its acquisition. The effect on the histogram is a displacement of frequencies. This phenomenon does not affect to edges and other details. Therefore, a clustering algorithm should be invariant against bright changes.

The translated histogram $P+h$ of $P=\{f_0,\ldots,f_K\}$ by means of a translation $h \in \mathbb{Z}$ is defined as $P+h=\{g_i\}$, where $g_i = f_{i-h} \quad \forall i \in \mathbb{Z}$.

As a property of the *CLOSE* algorithm we have that cut points in $P+h$ are the corresponding translated of that of P.

This is a trivial result, if we take into account that

$$\delta_i^d(P+h) = \delta_{i-h}^d(P) \qquad \forall i, d \in \mathbb{Z}.$$

Theorem 4. (*invariance with respect to intensity*) *Let X be an image. Then*

$$CF_d(X+h) = CF_d(X) + h \qquad \forall h \in \mathbb{Z}.$$

Other derived properties are:

$$d \le 1 \Longrightarrow CF_d(X_i) = X_i \quad \forall i \quad \text{(identity filter).}$$

$$d \ge K \Longrightarrow CF_d(X_i) = \frac{1}{n}\Sigma_{r\in V} X_{i+r} \quad \forall i \quad \text{(mean filter).}$$

These two extremes of the parameter d provide information about the behavior of the filter: when $d \le 1$ it is the identity filter, and when d is greater than the range of the scale it is the mean filter. Intermediate values of d, depending on the intensity of noise and the window size, will eventually enable us to obtain optimal outputs.

4. Experimental Results

4.1. Choosing Parameters

In Table 1, theoretical breakdown probabilities depending on the ratio d/σ and on the size of the window are exposed.

From this one can derive an appropriate size of window and an appropriate value of the scattering limit d depending on the intensity σ of the noise. Values providing a breakdown probability less than 5% can be considered as acceptable. For example, if a 5×5 window is advisable to use (due to nature of the image and resolution), then a ratio $d/\sigma \simeq 3.2$ will give a reasonably low breakdown probability.

4.2. Comparison

In this section, some simulation results applying the *CF* filter on images corrupted by gaussian noise are presented.

The images under consideration are made of 256×256 pixels. The mean absolute error *MAE* and the mean square error *MSE* between the original image and the filtered image are evaluated to compare the quality of the filter with other present filters in the literature.

The image chosen for comparison has very scattered local histograms. The background is a shade

Table 1. Breakdown probabilities of several sets of parameters.

	3×3	5×5	7×7	9×9	15×15
$d/\sigma = 2.6$	0.12558	0.24018	0.39274	0.54936	0.88267
$d/\sigma = 2.8$	0.07487	0.14213	0.23983	0.35467	0.69215
$d/\sigma = 3$	0.04202	0.07977	0.13676	0.20754	0.46286
$d/\sigma = 3.2$	0.02327	0.04190	0.07327	0.11204	0.27163
$d/\sigma = 3.4$	0.01233	0.02142	0.03674	0.05709	0.14472

going from black to white, and there are little details such as letters one pixel wide, corners of the star, and other meaningful details easy to lose when filtered.

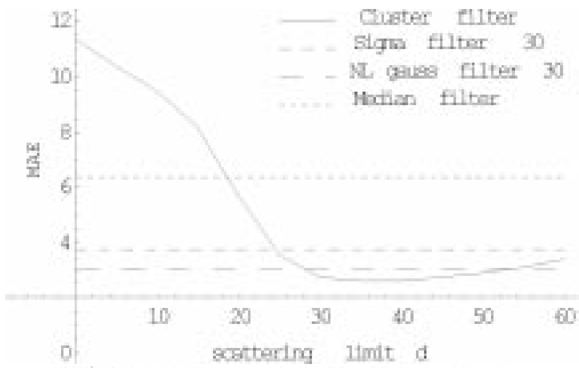
As it has been commented *MAE* and *MSE* are relative measures of quality of a filter. In our case, a minimum value of *MAE* or *MSE* does not mean that the corresponding image is visually better. The experiments have been carried out with three noise intensities: $\sigma = 15, \sigma = 20, \sigma = 25$, and the values of the parameters that gave the best results in each case were selected.

Figure 5 shows the original image, and the same image corrupted with gaussian noise of mean 0 and standard deviations $\sigma = 15$, 20 and 25 respectively.

4.2.1. Gaussian Noise with $\boldsymbol{\sigma = 15}$. The graphs in Fig. 6 show for the *CF* filter the variation of the *MAE* (left) and the variation of the *MSE* (right) with respect to the scattering limit *d*, by using a 5×5 window, and a comparison with the median, the sigma and the non-linear gaussian filters.

In the two graphs it is clear that the *MAE* and *MSE* for the *CF* are lower than for the other filters for a wide range of scattering limits; this range corresponds to optimal values of *d*. The optimal value given in paragraph 4.1 from the breakdown probability lies within this range.

On the other hand the non-linear gaussian filter in the two cases lies near to the *CF* filter, on both the *MAE* and the *MSE*; but visually (see Fig. 7) *CF* filter is a little better. Note that some degradation appears in the middle of the image; this happens when local histograms do not have separate clusters, that is, all colors in the sliding window are close each other.



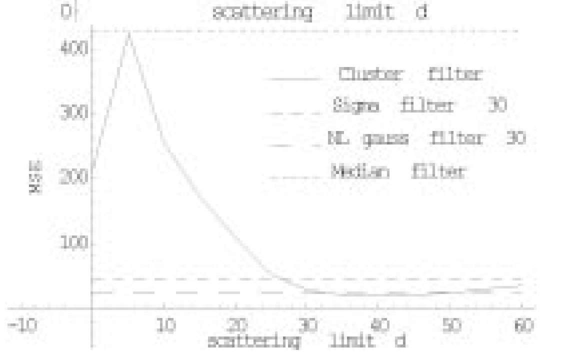
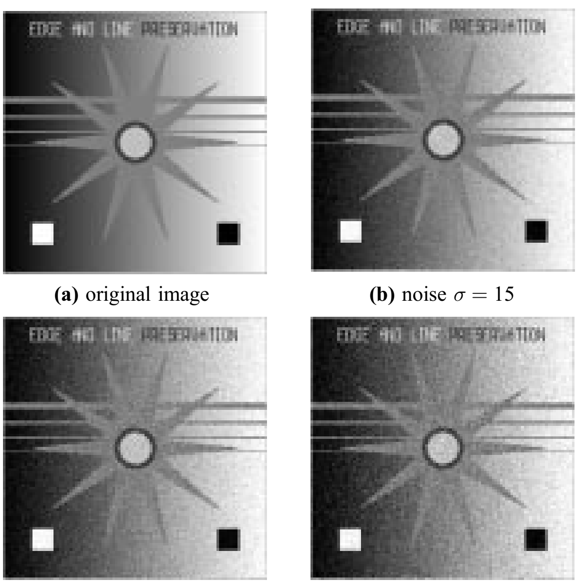


Figure 6. Evolution of *MAE* and *MSE* in function of the scattering limit for $\sigma = 15$.



(a) original image **(b)** noise $\sigma = 15$

(c) Noise $\sigma = 20$ **(d)** Noise $\sigma = 25$

Figure 5. The test images.

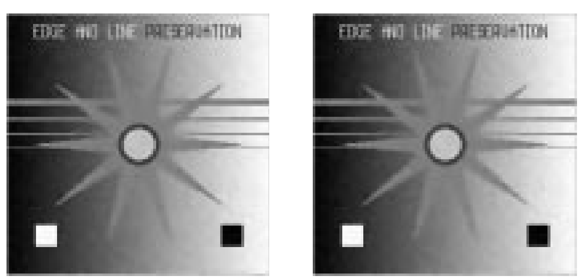


Figure 7. The image 5 (b) filtered with CF_{45} (left) and NL_{30} (right) with a 5×5 window.

4.2.2. Gaussian Noise with $\sigma = 20$ The graphs of Fig. 8 provide the same results that in the previous section, but in this situation the image is corrupted with gaussian noise of $\sigma = 20$ and a 7×7 window is used.

Together with output of CF_{60}, the result for the sigma filter is shown for visual comparison (see Fig. 9).

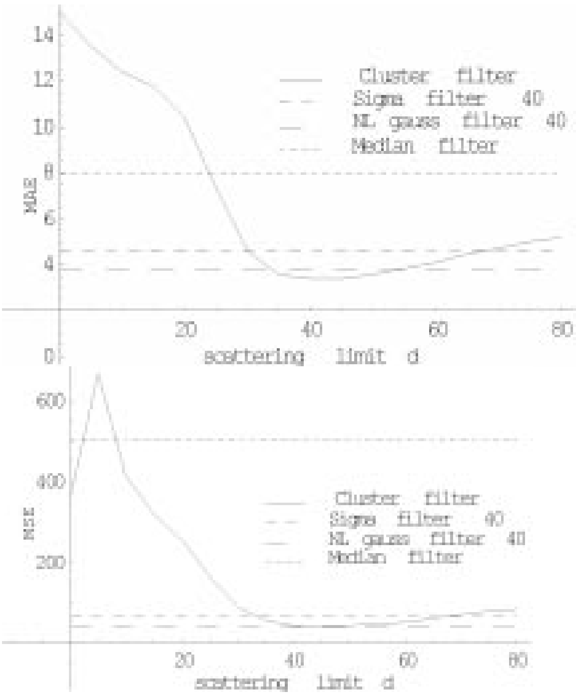


Figure 8. Evolution of *MAE* and *MSE* with respect to the scattering limit for $\sigma = 20$.

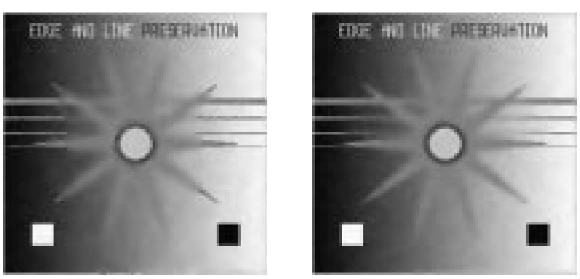


Figure 9. The image 5 (c) filtered with CF_{60} (left) and SF_{40} (right) with a 7×7 window.

4.2.3. Gaussian Noise with $\sigma = 25$. The graphs of the Fig. 10 provide the variation of *MAE* and *MSE* for the *CF* filter depending on the scattering limit, using a 5×5 window, and a comparison with the median filter, the sigma filter and the non-linear gaussian filter.

It is clear that *MAE* and *MSE* for the *CF* filter are lower than that of the other filters, and also we emphasize the presence of the stability zone for *MAE* and *MSE* with respect to the scattering limit.

On the other hand, a visual comparison with the median filter (see Fig. 11) shows that *CF* filter is much better than the classic median filter *MF* in removing gaussian noise.

The *CF* filter reduces the dispersion of the original color in images and also emphasizes the different colors that have been disturbed. This provides good results in removing the gaussian noise.

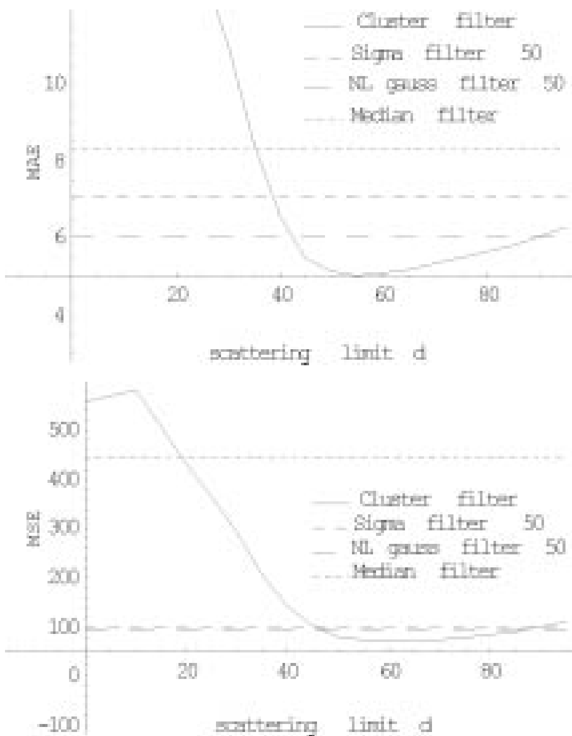


Figure 10. Evolution of *MAE* and *MSE* with respect to the scattering limit for $\sigma = 25$.

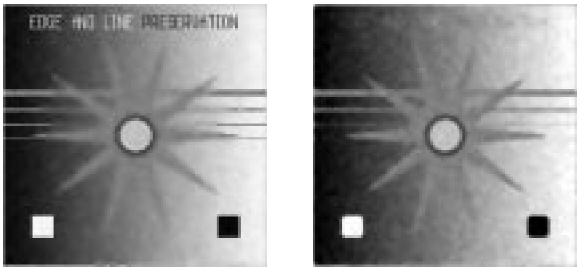


Figure 11. The image 5(d) filtered with CF_{75} (left) and *MF* (right) with a 5×5 window.

5. Conclusions

Local histograms with a large gray scale tend to be very scattered, causing difficulties to process the corresponding images.

In this work the dispersion of the histograms has been reduced, by means of a clustering algorithm based on seeking the best cut points to separate colors into clusters.

This technique is used to filter images corrupted with gaussian noise by centering on each pixel a window whose histogram is clustered, and the output of the filter is the mean level of the cluster in which the central pixel is found.

The idea has provided some very good results in comparison with the existing filters in the literature.

References

1. L. Alvarez, J. Esclarín, "Image Quantization by Nonlinear Smoothing", *SPIE*, Vol. 2567, No. 18, pp. 182-192, 1995.
2. M.R. Anderberg, R.K. Blashfield, *"Cluster Analysis"*, Sage Pub. Inc., Beverly Hills, 1984.
3. D.R.K. Brownrigg, "Weighted Median Filters", *Commun. Assoc. Comput. Machinery*, Vol. 27, pp. 807-818, August 1984.
4. H.A. David, *"Order Statistics"*, John Wiley, 1981.
5. R.C. Dubes, A.K. Jain, "Clustering Techniques: The User's Dilemma", *Pattern Recognition*, Vol. 8, pp. 247-260, 1976.
6. F. Godtliebsen, E. Spjøtvoll, "Comparison of Statistical Methods in MR Imaging", *Int. J. Imaging Systems and Technology*, Vol. 3, pp. 33-39, 1991.
7. R.M. Haralick, L.G. Shapiro, "Image Segmentation Techniques", *Computer Vision, Graphics, and Image Processing*, Vol. 29, pp. 100-132, 1985.
8. A.K. Jain, R.C. Dubes, *"Algorithms for Clustering Data"*, Prentice-Hall, Englewood Cliffs, NJ, 1988.
9. B. Jäne, *"Digital Image Processing"*, Springer-Verlag, New York, 1995.
10. J.S. Lee, "Digital Image Smoothing and the Sigma Filter", *Computer Vision, Graphics, and Image Processing*, Vol. 24, pp. 255-269, 1983.
11. R.S. Michlaski, R.F. Stepp, "Automated Construction of Classification: Conceptual Clustering Versus Numerical Taxonomy", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 5, pp. 396-410, 1983.
12. I. Pitas, A.N. Venetsanopoulos, *"Nonlinear Digital Filters: Principles and Applications"*, Kluwer Academic Publisher, 1990.
13. S. Theodoridis, K. Koutroumbas, *"Pattern Recognition"*, Academic Press, New York, 1999.
14. J. Weule, "Iteration Nichtlinearer Gauß-Filter in der Bildverarbeitung", Ph.D. Thesis, Heinrich-Heine-Universität Düsseldorf, 1994.

Zakaria Atae-Allah was born in Tanger, Morocco, in 1974. He received his B.Sc. degree in Mathematics in 1996 from the University of Teouan, Morocco. He is a secondary-school professor at Hassan-II School, Zomi, Morocco. His area of interest includes mathematical aspects of image processing.

José Martínez Aroza received his B.Sc. degree in Mathematics in 1979 from the University of Granada, Spain. From 1979 to 1987 he was a professor at the University of Almería, Spain. He received the Ph.D. degree in Applied Mathematics in 1990 from the University of Granada, Spain. He is associate professor at the Department of Applied Mathematics, University of Granada, Spain. His research interests lie on the field of digital image processing, move specifically on noise filtering and segmentation of homogeneous as well as textured regions.