A STATISTICAL BASED APPROACH FOR REMOVING HEAVY TAIL NOISE FROM IMAGES

M. El Hassouni and H. Cherifi
LIRSA, University of Dijon
9, avenue Alain Savary 21000 Dijon, France

Keywords: image restoration, l_p-norm filter, α-stable processes, Fractional lower order statistics.

Abstract: In this paper, we propose to use a class of filters based on fractional lower order statistics (FLOS) for still image restoration in the presence of α-stable noise. For this purpose, we present a family of 2-D finite-impulse response (FIR) adaptive filters optimized by the least mean l_p-norm (LMP) algorithm. Experiments performed on natural images prove that the proposed algorithms provide superior performance in impulsive noise environments compared to LMS and Weighted Myriad filters.

1 INTRODUCTION

Several distributions that exist can be good candidates for modelling noise signals. The most common in the literature, and especially in signal and image processing, is the Gaussian distribution. The use of the Gaussian distribution is frequently motivated by the physics of the problem, and in most cases it ensures an analytical solution. This has led to the development of numerous algorithms based on second-order statistics.

However, in many real-world problems the noise encountered is more impulsive in nature than that presented by a Gaussian distribution. Important non-Gaussian impulsive noise, found in radar and mobile communications for example, can be efficiently modelled by infinite variance processes for which the theory of SOS is not useful(Gonzales and Arce, 1997). There exists a class of distributions, called α-stable distributions that can be used to model these types of noise. A variety of FLOS-based algorithms have been proposed to filter α-stable random processes (Aydin et al., 1999; Shao and Nikias, 1993; Kidmose, 2000). This property has been used extensively in one-dimensional (1-D) signal processing, e.g. in removing some undesired effects in communication channels or synthesizing a deconvolution filter. Nonlinear least l_p-norm polynomial filters expansion is a powerful tool in signal processing (Kuruoglu et al., 1998; Kuruoglu, 2002). However, a serious problem is the increased filter complexity as compared to linear filtering. Developing this techniques for two-dimensional (2-D) signal processing (i.e. image processing) is of current research interest. Weighted myriad filters (WMyF) have been used with considerable success in robust communications and image processing. They have been derived based on the Cauchy distribution, which is a special case of α-stable distributions. WMyF are inherently more powerful than weighted median filters (WMF), which along with other order statistics filters, have been used widely in image processing, due to their edge preservation and outlier rejection properties (Gonzalez and Arce, 1996).

In this paper, we present a 2-D adaptive least l_p-norm (LMP) filter and its normalized version for image restoration under additive α-stable noise. The algorithms are based on fractional lower order statistics (FLOS). The performances of these algorithms are compared to those of the normalized least mean square-type algorithm which is developed under the Gaussian assumption.

The organization of the article is as follows. In Section 2, we present the problem formulation with some definitions of the α-stable processes. Section 3 is then devoted to the presentation of the 2-D adaptive l_p-norm filter and some related work. Finally, experimental results and some concluding remarks are given in Section 4 and Section 5, respectively.

El Hassouni M. and Cherifi H. (2006).
A STATISTICAL BASED APPROACH FOR REMOVING HEAVY TAIL NOISE FROM IMAGES.
In *Proceedings of the First International Conference on Computer Vision Theory and Applications*, pages 157-161
DOI: 10.5220/0001371901570161

2 PROBLEM FORMULATION

Let us consider that the observed image $g(i,j)$ can be expressed as a sum of noise-free image $f(i,j)$ plus (2-D) additive noise, i.e.,

$$g(i,j) = f(i,j) + \eta(i,j) \tag{1}$$

where (i,j) denotes the pixel coordinates. In image processing a neighborhood is defined around each pixel (i,j). The random noise $\eta(i,j)$ is modelled as a symmetric α-stable process ($S\alpha S$).

The stable distribution law is a direct generalization of the Gaussian distribution and in fact includes the Gaussian as a limiting case. The main difference between the non-Gaussian stable distribution and the Gaussian distribution is that the tails of the stable density are heavier than those of the Gaussian density. The characteristic of the stable distribution is one of the main reasons why the stable distribution is suitable for modeling signals and noise of impulsive nature. The α-stable distributions do not have close form probability density function except the cases $\alpha = 1$ (Cauchy distribution) and $\alpha = 2$ (Gaussian distribution). The $S\alpha S$ probability density function (pdf) is defined by means of its characteristic function

$$\phi(\omega) = \exp(\delta i\omega - \gamma|\omega|^{\alpha}) \tag{2}$$

where

- $i)$ $\alpha(0 < \alpha \leq 2)$ is the characteristic exponent, controlling the heaviness of the pdf tails,
- $ii)$ $\gamma(\gamma > 0)$ is the dispersion, which plays an analogous role to the variance, and
- $iii)$ δ is the location parameter, the symmetry axis of the pdf.

Due to the heavy tails, stable distributions do not have finite second- or higher-order moments, except the limiting case of $\alpha = 2$. More precisely, for X, an α-stable random variable with $(0 < \alpha < 2)$

$$E[|X|^p] = \infty \qquad \text{if} \qquad p \geq \alpha \tag{3}$$

However, for $0 < p \leq \alpha$, the fractional lower order moment (FLOM) is finite, i.e.,

$$E[|X|^p] < \infty \qquad \text{if} \qquad 0 \leq p < \alpha \tag{4}$$

If $\alpha = 2$, then

$$E[|X|^p] < \infty \qquad \text{for} \qquad p \geq 0 \tag{5}$$

The fractional pth-order moment of an $S\alpha S$ random variable with zero location parameter, $\delta = 0$, is given by

$$E[|X|^p] = C(p,\alpha)\gamma^{p/\alpha}, \qquad \text{for} \quad 0 < p < \alpha \tag{6}$$

where

$$C(p,\alpha) = \frac{2^{p+1}\Gamma(\frac{p+1}{2})\Gamma(-p/\alpha)}{\alpha\sqrt{\pi}\Gamma(-p/2)} \tag{7}$$

$\Gamma(.)$ denotes the gamma function.

3 2-D ADAPTIVE LEAST L_P-NORM FILTER

Our purpose is to design a filter on the pixel neighborhood (called the filter window) that aims at estimating the noise-free central image pixel value by minimizing a certain criterion. Among the several filter masks that are used in digital image processing, we rely on the square window of dimension $\Xi \times \Xi$ where Ξ is generally assumed to be an odd number, i.e., $\Xi = 2\xi + 1$.

$$\begin{bmatrix} g(i-\xi,j-\xi) & g(i-\xi,j-\xi+1) & \ldots & g(i-\xi,j+\xi) \\ \vdots & & & \vdots \\ g(i+\xi,j-\xi) & g(i+\xi,j-\xi+1) & \ldots & g(i+\xi,j+\xi) \end{bmatrix} \tag{8}$$

Let us rearrange the above $\Xi \times \Xi$ filter window in a lexicographic order (i.e., row by row) to a $N \times 1$ vector, where $N = \Xi^2$.

We assume that the filter window is sliding over image in a raster scan fashion. If K and L denote the image rows and columns respectively, a scalar running index m is defined by

$$m = (i-1)K+j \qquad 1 \leq i \leq K \qquad 1 \leq j \leq L \tag{9}$$

Consider an adaptive 2D FIR filter, the intensity estimate at pixel (i,j) is given by :

$$\hat{f}(i,j) = \sum_{k=-\xi}^{+\xi}\sum_{l=-\xi}^{+\xi} a(k,l)g(i-k,j-l) \tag{10}$$

Where $\hat{f}(i,j)$ is the filtered gray-level pixel located at (i,j); $a(k,l)$ is the filter weights around pixel (i,j); $g(i-k,j-l)$ is the corrupted pixel at location $(i-k,j-l)$.

The filter weights are estimated by minimizing a cost function $J(e(i,j))$, where $e(i,j) = \hat{f}(i,j) - f(i,j)$ is the filter estimation error at pixel (i,j).

The update equations are defined by :

$$\hat{a}_{m+1}(k,l) = \hat{a}_m(k,l) + \mu\frac{\partial J(e(i,j))}{\partial a_m(k,l)} \tag{11}$$

Where μ is the step-size parameter that controls the stability and the rate of convergence.

3.1 Related Work

The classical and the simplest adaptive type algorithm is the *least mean square* (LMS). LMS adapts the linear filter weights with every coming sample in the steepest descent direction. The cost function can be defined as :

$$J(e(i,j)) = E\{e^2(i,j)\} \tag{12}$$

the update equations of LMS filter and it's adaptive version NLMS filter are defined by:

- 2-D LMS:
$\hat{a}_{m+1}(k,l) = \hat{a}_m(k,l) + 2\mu e(i,j) g(i-k, j-l)$
- 2-D NLMS:
$\hat{a}_{m+1}(k,l) = \hat{a}_m(k,l) + 2\mu \frac{e(i,j)}{\|g(i,j)\|^2+\lambda} g(i-k, j-l)$

Where λ is an algorithm parameter, the quantity $\|g(i,j)\|^2$ is given by:

$$\|g(i,j)\|^2 = \sum_{k=-\xi}^{+\xi} \sum_{l=-\xi}^{+\xi} |g(i-k, j-l)|^2 \quad (13)$$

3.2 Proposed 2-D Adaptive Least l_p Norm Filter

The LMS algorithm has sever convergence problems for signals with more probability mass in the tails, than the Gaussian distribution. Recently filter theory for $S\alpha S$ signals has been developed, and the least Mean p-norm (LMP) algorithm has been proposed (Shao and Nikias, 1993). The objective is to minimize the dispersion of the error (p-norm cost function) which is defined by :

$$J(e(i,j)) = E\{|e(i,j)|^p\} \quad (14)$$

So, the update equation is:

- 2-D LMP :
$\hat{a}_{m+1}(k,l) = \hat{a}_m(k,l) + \mu p |e(i,j)|^{p-1} \text{sign}(e(i,j)) g(i-k, j-l)$

Although LMP is much more robust to impulsive noise than LMS, in some cases with the appearance of extremely impulsive noise it can become unstable. Motivated by the stability and increased speed of the normalized version of LMS, namely NLMS, an adaptive version for LMP, namely normalized LMP (NLMP), which can be expressed by :

- 2-D NLMP :
$\hat{a}_{m+1}(k,l) = \hat{a}_m(k,l) + \mu p \frac{|e(i,j)|^{p-1} sign(e(i,j))}{\|g(i,j)\|_p^p+\lambda} g(i-k, j-l)$

Where is a small λ which is included to avoid the division by zero, the quantity $\|g(i,j)\|_p^p$ is given by:

$$\|g(i,j)\|_p^p = \sum_{k=-\xi}^{+\xi} \sum_{l=-\xi}^{+\xi} |g(i-k, j-l)|^p \quad (15)$$

In all of these algorithms, the coefficient vector can be conveniently initialized to a good guess or to a zero vector if no prior information is available.

4 EXPERIMENTAL RESULTS

Here we present a simulation example of our application experiments demonstrating the behavior of the 2-D adaptive lp-norm filter. Due to the lack of original noise-free images in practice, the filters are trained over an image other than those to be filtered in order to represent a realistic scenario. Specially, the couple image is used in deriving the weighting coefficients, which are then used in filtering other images (Kotropoulos and Pitas, 2001). The images are corrupted by symmetrical α-stable noise with ($\alpha = 0.5; ...; 1.5$), ($\gamma = 1$) and ($\delta = 0$).

For objective quality evaluation, two criteria have been employed, namely, the fractional order SNR (FSNR(Kuruoglu et al., 1998)) defined as the logarithm of the ratio of the pth-order moments of the noise and the signal where $0 < p < \alpha$ instead of their powers, that is,

$$FSNR(dB) = 10 \log_{10} \frac{\sum_{i=1}^{K} \sum_{j=1}^{L} |f(i,j)|^p}{\sum_{i=1}^{K} \sum_{j=1}^{L} |\hat{f}(i,j) - f(i,j)|^p} \quad (16)$$

This evaluation criterion may favor the least l_p-norm filters since it is defined for the pth-order. However, we propose to use the Mean Structural SIMilarity (MSSIM) index described in (Wang et al., 2004).

$$MSSIM(X,Y) = \frac{1}{M} \sum_{j=1}^{M} SSIM(x_j, y_j) \quad (17)$$

where X and Y are the reference and the distorted image, respectively; x_j and y_j are the image contents at the jth local window; and M is the number of local windows of the image. The SSIM index is defined by:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C1)(2\sigma_{xy} + C2)}{(\mu_x^2 + \mu_y^2 + C1)(\sigma_x^2 + \sigma_y^2 + C2)} \quad (18)$$

Where the constants $C1 = (K_1 L^2)$ and $C2 = (K_2 L^2)$ are included to avoid the instability when $\mu_x^2 + \mu_y^2$ and $\sigma_x^2 + \sigma_y^2$ are very close to zero, respectively. L is the dynamic range of the pixel values (255 for 8-bit grayscale images), and K_1, K_2 are small constants $\ll 1$.

A point that needs some clarification is the choice of the scale of the p norm where the performance comparisons have to be made. So, the common way of determining p of the algorithm is to use the equality suggested by Money and al. (Money et al., 1982) that relates the kurtosis of the non-Gaussian data p of the l_p-norm minimization algorithm, which is

$$p = \frac{9}{\mathcal{K}^2} + 1 \quad (19)$$

where $\mathcal{K}$ is the kurtosis measure. Although this measure is very commonly used in literature, we have doubts concerning its efficiency, since the kurtosis is defined through the fourth and the second order moments which are not finite for α-stable distributed random variables. In our simulations, the p values have been tuned experimentally, Figure 1 shows an example of PSNR and MSSIM values obtained several values of p. It is shown that both measures give the similar optimum values of p. For example, with $\alpha = 1.5$ the optimum value of p is around 1.3, and with $\alpha = 1.9$ the optimum value of p is around 1.7.

We show in Figure 2 the results of the adaptive l_p-norm filter in suppressing impulsive noise in Lena image. Figure 2(a) shows the image corrupted by impulsive noise which is represented by Symmetrical α-stable noise (α=0.5). The output of the NLMS filter is shown is Figure 2(b). Figure 2(c) shows the filtered image by using the weighted median filter (WMF). The filtered image using the weighted myriad filter (WMyF) is shown in Figure 2(d). Figure 2(e) shows the filtered image using the 2-D adaptive LMP filter with fixed step-size. Finally, the filtered image by the normalized l_p-norm filter is shown in Figure 2(f). The filter parameters have been tuned to obtain the best visual results.

The visual quality demonstrates the superiority of using fractional lower-order statistics filtering algorithms which gives a good performance with edges and fine image details preserving.

Table 1 summarizes respectively the FSNR and the MSSIM achieved by the adaptive proposed filters with different values of α . According to these results, the filtered image using l_p-norm filter has higher FSNR and MSSIM improvement from the LMP/NLMP linear equalizers than those of the NLMS filter, weighted median filter and weighted myriad filter. The measures achieved by the normalized l_p-norm filter give good improvement in term of visual quality and signal-to-noise ratio improvement.

When a reference image is not available. We use the same methodology as described in (Kotropoulos and Pitas, 2001). We have tested the robustness of the filter coefficients that are determined at the end of a training session and are applied to filter a noisy image that has been produced by corrupting a different reference image than the one used in the training session. In Figure 3, we present a filtered house image using the coefficients determined at the end of a training session on lenna image. The same noise parameters have been used during the training session. When, regarding the quality of the filtered images and the quality evaluation measures, we can say that the proposed filter is very good for α-stable noise performance measurement.

5 CONCLUSION

In this paper, we presented a 2-D adaptive l_p-norm filter for noise suppression in images. Experimental results on natural images showed marked improvement in visual and numerical qualities when using the normalized l_p-norm algorithm adaptation. The filter is very suitable for α-stable and impulsive noise removal.

REFERENCES

G. Aydin, O. Arikan, and E. Cetin. Robust adaptive filtering algorithms for α-stable random processes. *IEEE Trans. On Circuits and Systems II*, vol. 46, N. 2, pp. 198-202, February 1999.

M. Shao, C. L. Nikias. Signal processing with fractionnal lower order moments : Stable processes and their applications. *Proc. IEEE*, vol. 81, pp. 986-1009, 1993.

J. G. Gonzalez, G. R. Arce. Weighted myriad filters : A robust filtering framework derived from alpha-stable distributions. *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, Atlanta, GA, May 1996, Vol. 5, pp. 2833-2836.

J. G. Gonzales, G.R. Arce. Zero-order statistics : a signal processing framework for very impulsive processes. *Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics*, Banff, Canada, July 1997, pp. 254-258.

P. Kidmose. Adaptive filtering for non-Gaussian noise processes. *Proceedings of International Conference on Acoustics, Speech and Signal Processing*, ICASSP2000, pp. 424-427.

Ercan E. Kuruoglu, Peter J. W. Rayner, William J. Fitzgerald. Least l_p-norm impulsive noise cancellation with polynomial filters. *Signal Processing*. v. 69, pp. 1-14. 1998.

A. Ben Hamza and H. Krim. Image denoising : A nonlinear robust statistical approach. *IEEE Trans. on Signal Processing*, Vol. 49, N0. 12, Decembre 2001.

Ercan E. Kuruoglu. Nonlinear least lp-norm filters for nonlinear autoregressive α-stable processes. Digital Signal Processing 12, 119-142 (2002).

O. Tanrikulu and A. G. Constantindes, ”Least-mean Kurtosis : A novel higher-order statistics based adaptive filtering algorithm”, *Electronics Letters*, vol. 30, no. 3, pp. 189-190, February 1994.

C. Kotropoulos and I. Pitas, Nonlinear Model-Based Image/Video Processing and Analysis, J. Wiley, 2001.

A. H. Money, J. F. Affleck-Graves, G.D.I. Barr, The linear regression model: L-p-norm estimation and the choice of p, Communications in Statist. Simulation Comput. 11 (1) (1982) 89-109.

Zhou Wang, Alan C. Bovik, Hamid R. Sheikh and Eero P. Simoncelli. Image quality assessment: from error measurement to structural similarity. To appear in IEEE transactions on image processing, vol. 13, No. 1, January 2004.

Table 1: FSNR and MSSIM achieved by different filters.

Alpha values	α=0.5		α=1.0		α=1.5	
Measure	FSNR(dB)	MSSIM	FSNR(dB)	MSSIM	FSNR(dB)	MSSIM
2-D NLMS	11.46	0.4865	14.01	0.806	13.66	0.824
2-D WMF	15.74	0.8230	15.35	0.8592	13.97	0.8678
2-D WMyF	15.87	0.8422	16.53	0.927	14.63	0.9034
2-D LMP	16.85	0.8736	16.28	0.9064	15.32	0.9126
2-D NLMP	17.30	0.8984	16.33	0.9162	15.55	0.9135

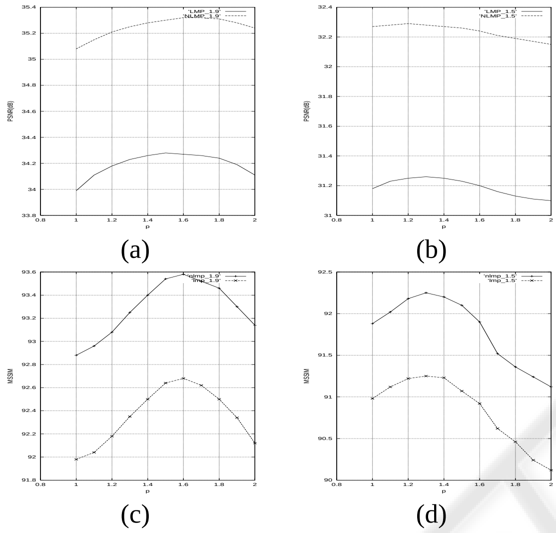

Figure 1: a) PSNR versus p for Lenna corrupted with $S\alpha S$ $\alpha = 1.9$. b) PSNR versus Lenna corrupted with $S\alpha S$ $\alpha = 1.5$. c) MSSIM versus p for Lenna corrupted with $S\alpha S$ $\alpha = 1.9$. d) MSSIM versus Lenna corrupted with $S\alpha S$ $\alpha = 1.5$.

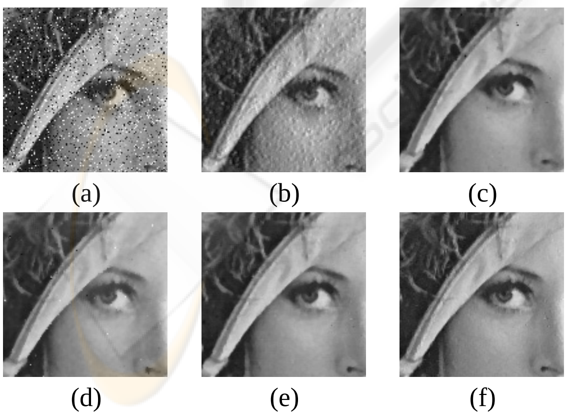

Figure 2: (a) Corrupted Lenna image ($\alpha = 0.5$) (b) Filtered image using NLMS filter (c) Weighted Median Filter (d) Filtered image using weighted myriad filter (e) filtered image using LMP filter (f) filtered image using NLMP filter.

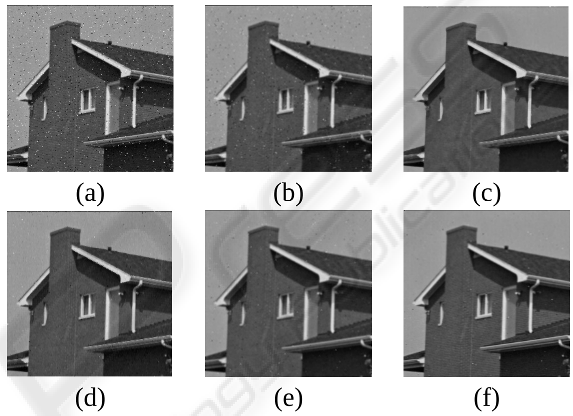

Figure 3: (a) Corrupted house image ($\alpha = 0.9$) (b) Filtered image using NLMS filter (FSNR=12.49 dB, MSSIM=0.74) (c) Weighted Median Filter (FSNR=13.94 dB, MSSIM=0.82) (d) Filtered image using weighted myriad filter (FSNR=15.50 dB, MSSIM=0.87) (e) filtered image using LMP filter (FSNR=14.84 dB, MSSIM=0.86) (f) filtered image using NLMP filter (FSNR=16.33 dB, MSSIM=0.92.)