ROBUST VIDEO ANALYSIS: WHICH STATISTICS TO USE ?

M. El Hassouni, H. Cherifi

LIRSA Laboratory
University of Bourgogne
Dijon, France.
{Mohamed.Elhassouni,Hocine.Cherifi}@u-bourgogne.fr

D. Aboutajdine

GSCM Laboratory
University of Mohammed V
Rabat, Morocco.
aboutaj@fsr.ac.ma

ABSTRACT

The aim of this paper is to discuss which statistics are efficient when video sequences are highly corrupted by impulsive noise. Usually the gaussian model is implicitly used for modelling the noise. Many papers in the literature use the gaussian mixture as an alternative of this later for image and video processing [1]. No a priori assumption for the noise pdf will be considered. We will use higher-order statistics to improve the processing results instead of second order ones used in the Gaussian case. Experimentations show the effectiveness of the proposed approach.

Keywords: *video denoising, spatio-temporal filters, motion estimation , higher order statistics, impulsive noise*

1. INTRODUCTION

Video sequences occur naturally in digital video broadcasting, teleconference, videophone, medical imaging, surveillance and defence. The demand for enhancement of video sequences has increased not only to improve visual quality bu also to aid subsequent processing.

These sequences are also corrupted by impulsive noise. . The noisy sequence is modelled as follows

$$g_n(i,j) = f_n(i,j) + \eta_n(i,j) \tag{1}$$

where $g_n(i,j)$ denotes the intensity of the pixel (i,j) of the n^{th} frame of the observed sequence, $f_n(i,j)$ stands for the intensity of the pixel (i,j) of the n^{th} frame of the original sequence and $\eta_n(i,j)$ denotes the zero mean additive noise. We will use for modelling this noise the following distributions:

- α-stable(α,δ): the symmetric α-stable noise. Where α and δ presents respectively, the exponential and the dispersion parameters.
- GG(α,β): The Generalized Gaussian noise. Where α et β denotes respectively the scale and the shape parameters .
- Mixed(σ,p): mixed noise (Gaussian + "salt and pepper "). Where σ and p presents respectively, the standard deviation and the pourcentage of impulses.

In general, spatio-temporal filter performance will be improved with accurate motion estimation. Motion compensation along the temporal direction is performed by HOS-based region recursive method [2]. After motion compensation, we apply a HOS-based filter to the reconstructed frames [3].

This paper is organized as follows. Section 2 presents the motion estimation step. Section 3 describes the proposed spatio-temporal filtering scheme. Experimental results are described in Section 4. Section 5 contains some conclusions.

2. MOTION ESTIMATION

To represent the motion of every region, we choose the affine linear model described by the vector

$$\Theta = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)^T \tag{2}$$

The vector of displacement $d = (di, dj)^T$ of an object will be described as follows:

$$\begin{cases} di = \theta_1 + \theta_3(i - i_g) + \theta_5(j - j_g) \\ dj = \theta_2 + \theta_4(i - i_g) + \theta_6(j - j_g) \end{cases} \tag{3}$$

Given a sequence of images $\{g_1, ..., g_N\}$, with $g_n = \{g_n(s)\}$, where $s = (i,j)^T$. The Displaced Frame Difference is denoted by $DFD(s) = g_{n+1}(s + d(s)) - g_n(s)$.

The correct displacement is found by minimizing the cost function based on the kurtosis measure of the DFD in a local area.

$$\hat{\Theta} = argmin\{E\{DFD^4(s)\} - 3E\{DFD^2(s)\}^2\} \tag{4}$$

The optimization algorithm uses the well known steepest descent gradient method and yields an algorithmic of the cumulant method.

3. SPATIO-TEMPORAL NOISE FILTERING

The intensity estimate at image location (i, j) for the n^{th} image is computed using a cube of size $(2p+1) \times (2q+1) \times (2l+1)$ and is given by :

$$\hat{f}_n(i,j) = \sum_{(p,q,l) \in S} \hat{a}(p,q,l) g^r_{n-l}(i-p, j-q) \quad (5)$$

where S is the filter support, $g^r_n(i,j)$ is the ordered observations and $\hat{a}(p,q,l)$ filter coefficient vector that minimizes the Kurtosis of the difference between the estimated and the original images given by :

$$J(s) = E\{\varepsilon^4(s)\} - 3E\{\varepsilon^2(s)\}^2 \quad (6)$$

where $s = (i, j)$ denotes the pixel location, $\varepsilon(s)$ is the estimation error at pixel s, i.e., $\varepsilon(s) = \hat{f}(s) - f(s)$. The algorithm for adjusting the filter coefficient is derived such as :

$$\hat{a}^{s+1}(p,q,l) = \hat{a}^s(p,q,l) + \mu \frac{\partial J(s)}{\partial a(p,q,l)} \quad (7)$$

Where μ is the step size controlling the convergence rate and the stability of the adaptation procedure. When the adaptive filter is to operate in non stationary environment, it is reasonable to employ a time/space varying step-size parameter $\mu(s)$. If $\mu(s)$ is chosen to be :

$$\mu(s) = \frac{1}{\|g^r_n(s)\|^2} \quad (8)$$

The normalized least Kurtosis filter is easily written by substituting the adaptive value of $\mu(s)$ in equation (6). A nonzero initialization of the parameters is necessary to start the adaptation procedure.

4. EXPERIMENTAL RESULTS

In our simulations, we use two video sequences ”Trevor White” and Caltrain. These sequences are corrupted by the three types of noise cited above. In order to have the same statistical characteristics, the parameters of noise distributions are optimized in ordre to have the same *Median Absolute Deviation* MAD value which is defined by:

$$MAD = \text{med}(|X_i - \text{med}(X)|) \quad (9)$$

Where X denotes the noise matrix. The table 1 shows the different noise realizations for the cited distributions using different parameters. We choose some of this realizations for our comparisons. First, we have the motion estimation method to different noisy sequences. Table 2 shows some of this results on both sequences with MAD value of 5. The obtained mean PSNR values of the sequence show a good performance of detection with the different types of noise.

Second, we apply the HOS-based spatio-temporal filter to motion-compensated sequences. The notation used in our figure legends and also later in summarizing our results is explained below:

- **SOS-SOS** means that we have used motion estimation method based on second order statistics (SOS) and spatio-temporal LMS L-filter.
- **HOS-SOS** means that we have used motion estimation method based on Higher order statistics (HOS) and spatio-temporal LMS L-filter.
- **SOS-HOS** means that we have used motion estimation method based on Second order statistics (SOS) and spatio-temporal LMK L-filter.
- **HOS-HOS** means that we have used motion estimation method based on Higher order statistics (HOS) and spatio-temporal LMK L-filter.

We give some results in tables 3, 4 and 5 with different values of MAD. These tables show the Signal to Noise Ratio Improvement (SNRI) values when using developed method comparing to second order statistics based methods.
Based on these results, we conclude that HOS-HOS based-method outperforms the other methods. the effectiveness of the method is classified according to the following order: (1: α-stable noise, 2: mixed noise, 3: Generalized Gaussian noise). This explains the robustness of the presented filter to highly impulsive noise effects. The figures 1 and 2 illustrates the effectiveness of the chosen method.

5. CONCLUSION

In this paper, a 3-D adaptive motion-compensated LMK L-filter for non-Gaussian noise suppression in video sequences was investigated. Higher-order statistics based method showed good performances comparing to second-order based method and for different types of noise.

6. REFERENCES

[1] J.C. Brailean, R.P. Kleihorst, S.N. Efstratiadis and A.K. Katsaggelos. Noise reduction filters for dynamic image sequences : A review. *Proc. IEEE* , vol. 83, no. 9, pp. 1236-1251, September 1995.

[2] E. Ibn-Elhaj, D. Aboutajdine, S. Pateux and L. Morin. HOS-based method of global motion estimation for noisy image sequences. *Electronics Letters*, vol. 35, no. 16, pp. 1320-1322, August 1999.

[3] M. El Hassouni, H. Cherifi and D. Aboutajdine. HOS-based image sequence noise removal. To appear in *IEEE Transactions on Image Processing* (2005).

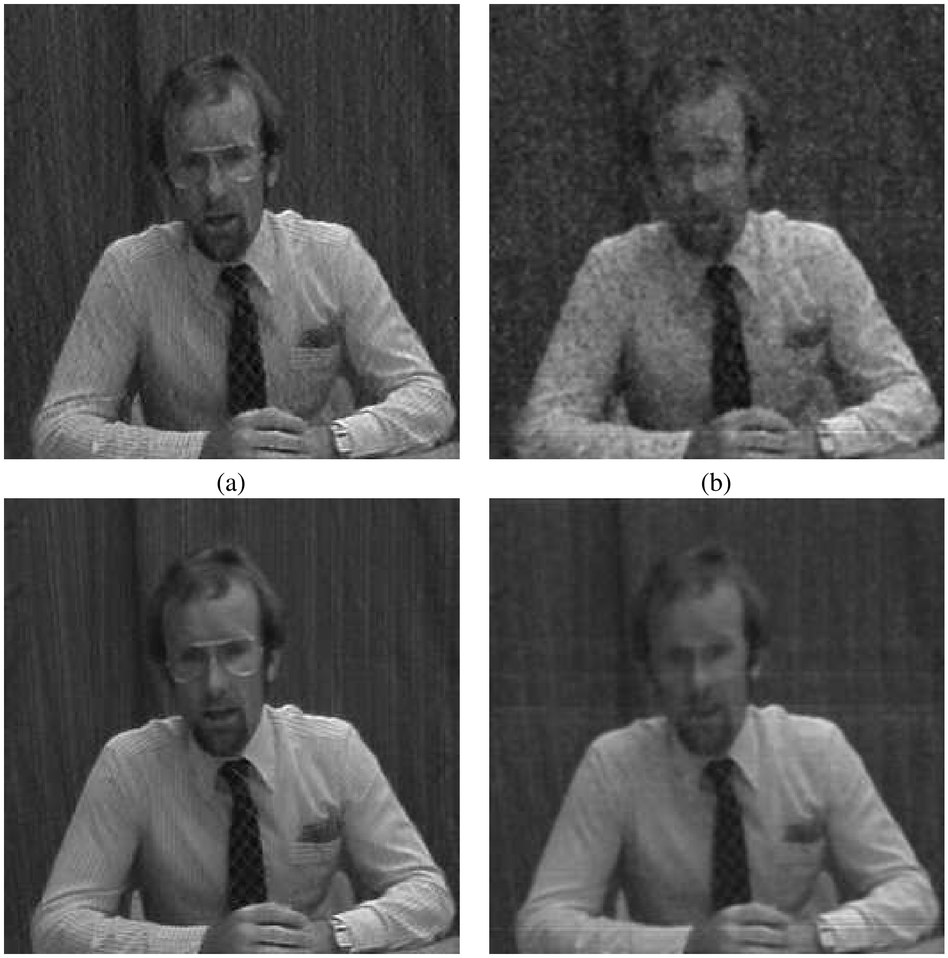


Fig. 1. The third filtered Trevor sequence. (a) α-stable(0.5,1.95) noise by SOS-SOS combinaison. (b) α-stable(0.5,1.95) noise by SOS-SSO combinaison. (c) GG(0.5,1.5) noise by SOS-SOS combinaison. (d)G(0.5,1.5) noise by SOS-SSO combinaison.

MAD=5	MAD=10	MAD=15
α-stable(0.5,1.95)	α-stable(0.5,2.8)	α-stable(0.5,3.4)
α-stable(1,5)	α-stable(1,10)	α-stable(1,15)
α-stable(1.5,12)	α-stable(1.5,33)	α-stable(1.5,61)
GG(0.5,1.5)	GG(0.5,2.1)	GG(0.5,2.6)
GG(1,10)	GG(1,20)	GG(1,30)
GG(1.5,60)	GG(1.5,170)	GG(1.5,310)
Mixte(20,5)	Mixte(20,10)	Mixte(20,30)
Mixte(15,8)	Mixte(15,25)	Mixte(15,36)
Mixte(10,13)	Mixte(10,37)	Mixte(10,45)

Table 1. Different noise realizations with different parameters

Sequences	Trevor White		Caltrain	
PSNR moyen	SSO	SOS	SSO	SOS
α-stable(0.5,1.95)	12.12	12.49	11.88	12.47
α-stable(1,5)	17.36	18.17	16.48	17.73
α-stable(1.5,12)	22.37	23.69	20.42	23.03
GG(0.5,1.5)	19.93	21.05	19.09	21.01
GG(1,10)	24.93	26.44	21.82	25.56
GG(1.5,60)	25.81	27.73	22.46	26.68
Mixte(10,13)	12.25	12.93	12.33	12.95
Mixte(15,8)	18.35	19.25	17.34	18.39
Mixte(20,5)	19.67	20.54	18.48	20.24

Table 2. The mean PSNR values for the reconstructed images for MAD=5

Sequences	Trevor White				Caltrain			
SNRI	SOS-SOS	SOS-HOS	HOS-SOS	HOS-HOS	SOS-SOS	SOS-HOS	HOS-SOS	HOS-HOS
α-stable(0.5,1.95)	-3.08	-7.60	-10.07	-12.37	-2.92	-6.46	-7.34	-8.49
α-stable(1,5)	-3.52	-4.02	-4.43	-5.19	-2.42	-2.81	-3.41	-4.09
α-stable(1.5,12)	-3.12	-3.65	-4.72	-5.28	-1.68	-2.33	-3.56	-3.94
GG(0.5,1.5)	-4.85	-7.60	-7.48	-9.06	-3.40	-6.80	-4.80	-5.00
GG(1,10)	-4.08	-4.14	-3.90	-4.41	-2.44	-3.02	-3.84	-4.12
GG(1.5,60)	-4.10	-4.31	-3.34	-4.38	-2.39	-2.82	-2.75	-3.06
Mixed(20,5)	-5.29	-5.47	-7.46	-9.59	-3.83	-5.56	-6.11	-7.59
Mixed(15,3)	-3.17	-3.56	-3.05	-3.50	-2.45	-2.65	-3.93	-4.38
Mixed(10,13)	-4.11	-4.52	-6.01	-6.25	-2.14	-2.82	-3.95	-4.71

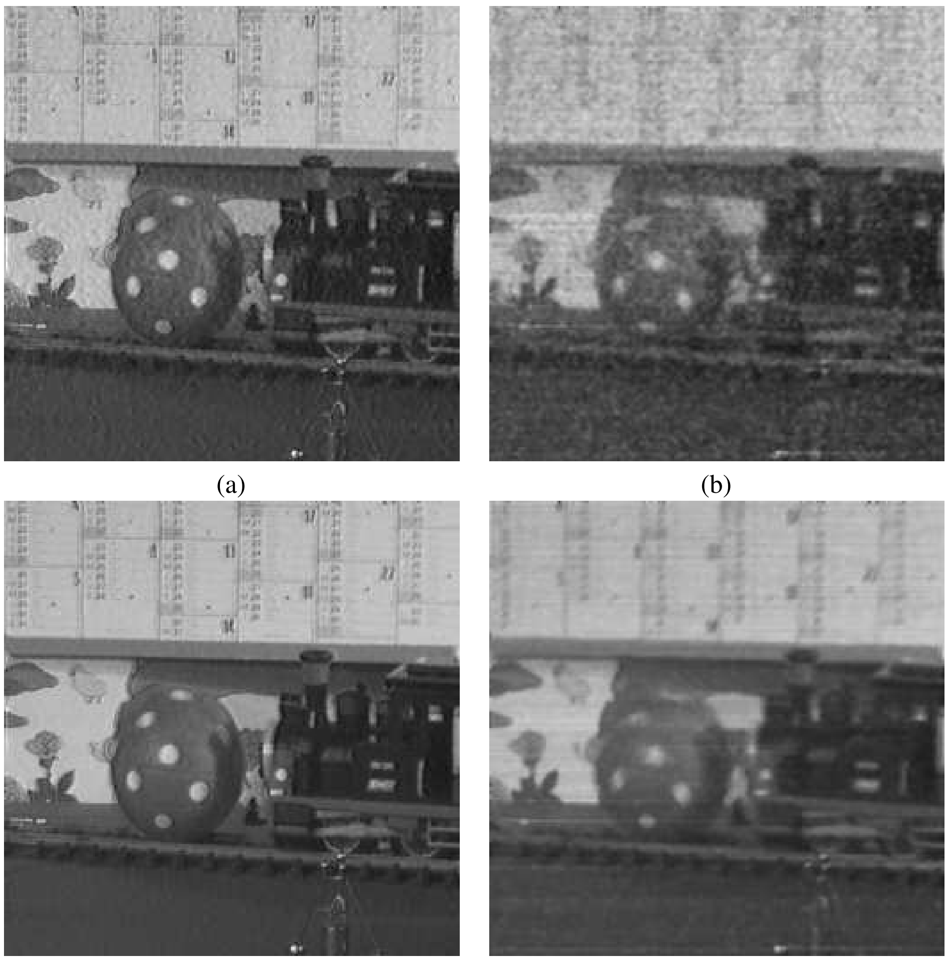
Table 3. SNRI values for MAD=5

Sequences	Trevor White				Caltrain			
SNRI moyen	SSO-SSO	SSO-SOS	SOS-SSO	SOS-SOS	SSO-SSO	SSO-SOS	SOS-SSO	SOS-SOS
α-stable(0.5,2.8)	-1.98	-2.03	-6.30	-6.41	-1.85	-2.74	-4.44	-5.19
α-stable(1,10)	-2.66	-3.46	-5.06	-5.57	-2.12	-2.60	-3.90	-3.71
α-stable(1.5,33)	-3.46	-4.01	-3.59	-4.08	-1.22	-1.35	-2.01	-2.59
GG(0.5,2.1)	-5.13	-6.12	-7.81	-9.21	-4.11	-4.57	-5.94	-6.55
GG(1,20)	-6.44	-7.48	-6.46	-7.16	-4.36	-4.62	-5.86	-6.77
GG(1.5,170)	-6.48	-7.59	-6.36	-7.06	-2.86	-3.47	-4.54	-5.18
Mixte(20,10)	-2.44	-3.52	-5.72	-6.08	-2.69	-3.17	-3.72	-4.36
Mixte(15,25)	-1.25	-2.20	-6.53	-7.42	-1.06	-2.38	-4.86	-5.55
Mixte(10,37)	-2.72	-3.39	-6.07	-6.96	-1.63	-2.52	-5.25	-6.23

Table 4. SNRI values for MAD=10

Sequences	Trevor White				Caltrain			
SNRI moyen	SSO-SSO	SSO-SOS	SOS-SSO	SOS-SOS	SSO-SSO	SSO-SOS	SOS-SSO	SOS-SOS
α-stable(0.5,3.4)	-4.77	-5.08	-6.28	-6.25	-2.64	-3.72	-4.64	-5.24
α-stable(1,15)	-2.55	-2.98	-5.74	-5.83	-3.15	-3.42	-4.22	-5.15
α-stable(1.5,61)	-3.04	-3.65	-4.50	-4.76	-2.12	-2.28	-2.66	-3.09
GG(0.5,2.6)	-5.61	-5.79	-9.17	-11.27	-6.48	-6.71	-7.32	-8.12
GG(1,30)	-6.49	-8.04	-8.11	-9.58	-4.92	-5.78	-6.55	-6.99
GG(1.5,310)	-7.59	-8.98	-7.98	-9.42	-6.13	-6.89	-7.59	-7.95
Mixte(20,30)	-3.19	-4.64	-6.46	-6.48	-3.82	-4.03	-4.95	-5.79
Mixte(15,36)	-2.24	-2.43	-6.45	-6.55	-4.12	-4.38	-5.10	-6.10
Mixte(10,45)	-1.47	-2.27	-6.78	-7.37	-3.86	-4.44	-5.41	-6.49

Table 5. SNRI values for MAD=15



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

Fig. 2. The third filtered Caltrain sequence. (a) α-stable(0.5,1.95) noise by SOS-SOS combinaison. (b) α-stable(0.5,1.95) noise by SOS-SSO combinaison. (c) GG(0.5,1.5) noise by SOS-SOS combinaison. (d)G(0.5,1.5) noise by SOS-SSO combinaison.