Pixel Domain Spatio-temporal Denoising for Archive Videos

M. Kemal Güllü, Oğuzhan Urhan, and Sarp Ertürk

Kocaeli University Laboratory of Image and Signal processing (KULIS),
Electronics and Telecom. Eng. Dept., University of Kocaeli, 41040, Kocaeli, Turkey
{kemalg, urhano, sertur}@kou.edu.tr

Abstract. A new pixel domain spatio-temporal video noise filter for archive film restoration has been proposed in this paper. The proposed filtering method takes motion changes and spatial information into account. Firstly, temporal filtering is carried out considering temporal changes adaptively. Afterwards, interpolation between degraded and temporally filtered images is carried out to preserve edge information using local standard deviation values. With respect to pixel domain techniques proposed in the literature, the proposed method gives better results for various test videos and particularly provides superior results for archive film.

1 Introduction

Restoration and storage of archive film materials are important for the transfer of cultural heritage to next generations. Archive films include many types of defects like flicker, blotch, scratch, and noise that are caused by storage medium conditions, playing and copying of films etc. Noise is a typically encountered degradation in archive films. After removing strong blotches that cause enormous temporal discontinuity in archive video, transparent dust effects and film-grain noise remain. These remained defects should be suppressed from archive film for improved visual quality.

Spatial (2-D) and spatio-temporal (3-D) filters [1-6] have been proposed in the literature to remove video noise. Spatial filters take only spatial information into account and as an effect can cause spatial blurring at high noise levels. This blurring effect can be reduced using both temporal and spatial information and the filtering performance can be improved at low noise levels also in this way.

A wavelet domain spatial filter whose coefficients are manipulated using a Markov Random Field (MRF) image model has been proposed in [1]. In [2], a Wiener filter is utilized in the wavelet domain in order to remove image noise. A fuzzy logic based image noise filter that takes directional deviations into account has been proposed in [3]. In [4], a recursive estimator structure has been proposed to estimate the clean image from the film-grain noisy image. Noise is considered to be related to exposure time in the form of non-Gaussian and multiplicative structure in [4].

A pixel based spatio-temporal adaptive filter that calculates new pixel values adaptively using the weighted mean of pixels over motion compensated frames has been proposed in [5]. In [6], an edge preserving spatio-temporal video noise filter that

A. Levi et al. (Eds.): ISCIS 2006, LNCS 4263, pp. 493 – 502, 2006.

combines 2D Wiener and Kalman filters has been presented. A non-linear video noise filter which calculates new pixel values using a 3D window has been proposed in [7]. This method arranges pixels in the form of a 3D window according to their difference with respect to related pixel values and averages the pixels in the window after weighting them according to their sorting order. This method gives good results in case of no- or slow local motion, but deforms image regions in cases of abrupt local motion. Video denoising using 2D and 3D dual-tree complex wavelet transforms has been proposed in [8]. In the case of local motion, the 3D filtering performance of the method is highly reduced. In order to increase the 3D filtering performance of the method proposed in [8], 2D wavelet based filtering and temporal mean filtering that uses pixel based motion detection has been proposed in [9]. A wavelet transform based video filtering technique that uses spatial and temporal redundancy has been proposed in [10]. In [11], a content adaptive video denoising filter has been proposed recently. This method filters both impulsive and non-impulsive noise but the filtering performance is highly reduced in case of Gaussian noise with high variance.

In this work, a new pixel based spatio-temporal video noise filter that takes motion changes and spatial standard deviations into account is proposed. The main objective is to suppress noise in archive video, and it is shown that the proposed method provides a successful visual quality for archive videos.

2 The Proposed Filtering Method

2.1 Noise Model

Noise can be defined as an unwanted component of the video and generally occurs as Gaussian noise, film-grain noise and quantization noise in archive video. In general, noise can be grouped as: additive (impulse, Gaussian noise) and multiplicative (film grain noise) noise. Image independent noise is described by an additive noise model as in (1).

$$I_n(x,y) = I_o(x,y) + \eta(x,y) \tag{1}$$

where I_o denotes the original image, I_n is the noisy image. The noise η is often considered as zero-mean Gaussian distribution and described by its variance σ_n^2 ($\eta(0,\sigma_n^2)$). In this work the noise for each pixel is handled as zero-mean Gaussian, independent and identically distributed (i.i.d.) (Additive White Gaussian Noise-AWGN).

2.2 Temporal Filtering Considering Local Motion

Only spatial noise reduction techniques give limited filtering performance and can produce frustrating artifacts [14]. Taking the advantage of the i.i.d. property of the noise, lost or degraded image information can typically be attained from previous and/or subsequent image frames for the current pixel $I_n(x,y,t)$ more accurately. Therefore filtering performance is increased for video sequences.

Similar to video de-noising methods in the literature, a 3D window (of size $ws \times ws \times ws$) around the (x, y, t) position has been constructed for each pixel (See Fig. 1) to obtain temporal information to the noise suppression process in our method.

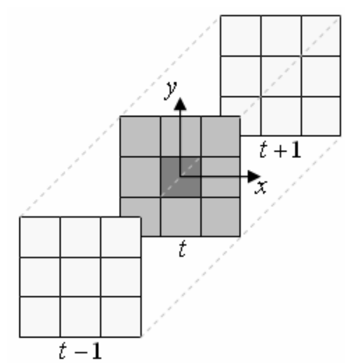


Fig. 1. The 3D window used for temporal video denoising

Suppose that preceding and subsequent frames are the same as the current frame and AWGN (Additive White Gaussian Noise) is added to each frame independently. Basically, linear average of the pixels in this 3-D window can give good results for less detailed regions in this assumption. While filtering real sequences, global and local motion effects should be taken into account to reduce temporal filtering artifacts such as blurring. Global motion effect can be recompensed using global motion compensation. However, image regions that include local motion in the preceding and/or succeeding frames that are non-matching parts of the 3-D window should not taken into account or local motion compensation should be done while filtering.

In the proposed method, two temporal contribution thresholds are calculated for each pixel in the current frame to determine which pixels from preceding and succeeding frames are to be used in the 3-D window while filtering. For this purpose, initially absolute image frame differences are thresholded using T_i and the Euclidean distance is measured over the binary difference image. Then, obtained distance images are smoothened using SOR (Successive Over-Relaxion) to avoid crisp changes [13]. Smoothened distance images are subsequently converted to temporal threshold images using a linear function $f_t(.)$ as given in (2). The proposed adaptive temporal threshold selection process is shown in Fig. 2.

$$T(x, y) = f_t(E(x, y)) = \begin{cases} \left[\dfrac{0.5 - E(x, y)}{0.5}\right] \times T_s , & E(x, y) \leq 0.5 \\ 0 , & otherwise \end{cases} \tag{2}$$

here $E(x, y)$ is Euclidean distance value for a given (x, y) pixel and T_s is the temporal deviation threshold.

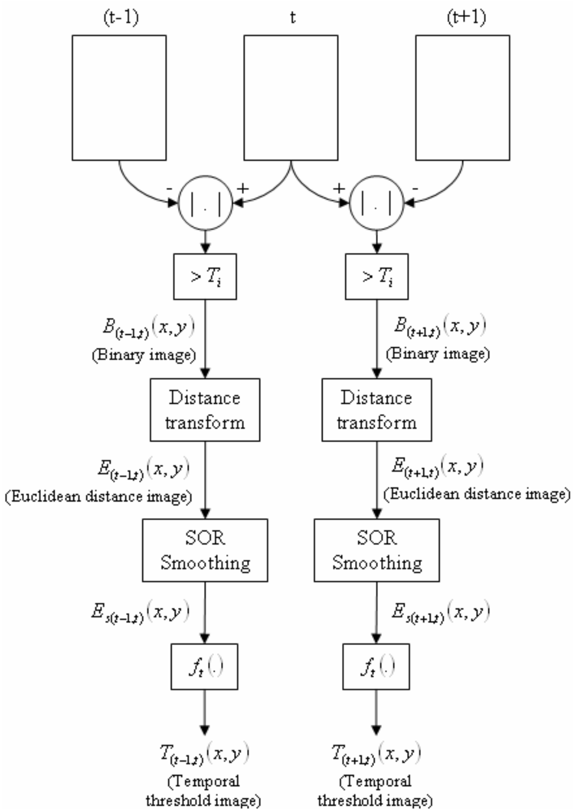


Fig. 2. Proposed adaptive temporal threshold selection stage

A simple SOR smoothing approach is utilized, which can be formulated as

$$E^{i+1}(x,y)=E^{i}(x,y)-\lambda\left(4E^{i}(x,y)-E^{i}(x-1,y)-E^{i}(x+1,y)-E^{i}(x,y-1)-E^{i}(x,y+1)\right)/4 \tag{3}$$

where i is the iteration order, and λ determines the smoothness of the operation. In this work, the maximum smoothness value is selected ($\lambda=1$). Temporal threshold images for frame #1160 of Mount archive video are given in Fig. 3.

After temporal threshold selection, local mean calculation of the $ws \times ws$ square window of t 'th frame is carried out as given (4).

$$\mu(x,y,t)=\frac{1}{ws\times ws}\sum_{m=1}^{ws}\sum_{n=1}^{ws} I_n\left(x+m-\frac{ws+1}{2},\, y+n-\frac{ws+1}{2},\, t\right) \tag{4}$$

$\mu(x,y,t)$ is used to decide pixels which can be used for temporal filtering from the previous and subsequent frames in the 3D window in case of local motion (denoted as P_b and P_f). This operation is carried out as given in (5).

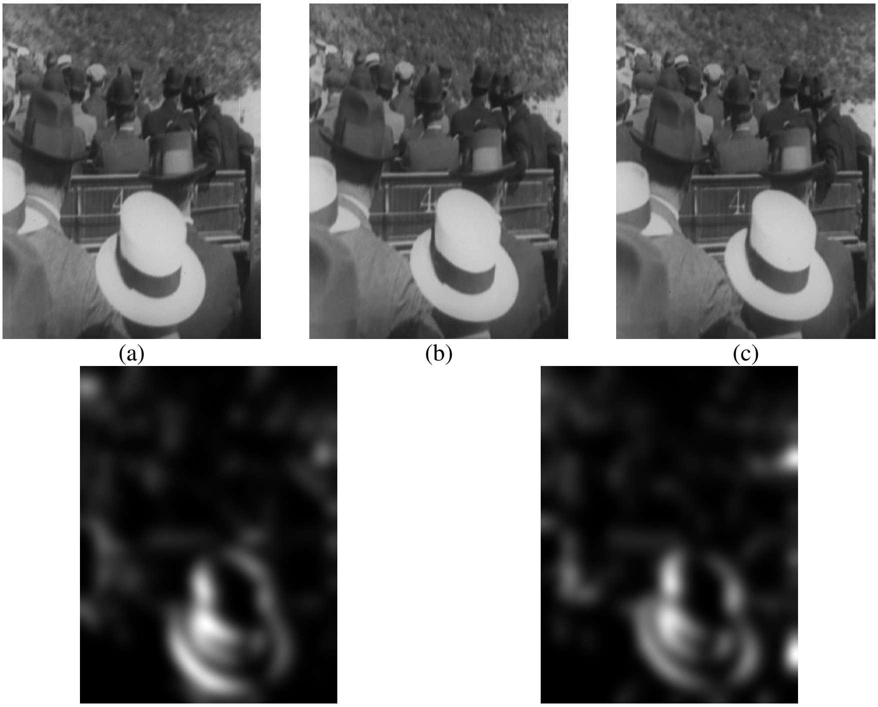


Fig. 3. Original frames a) #1159, b) #1160, c) #1161 of Mount archive video, d) $T_{(t-1,t)}$ and, e) $T_{(t+1,t)}$ threshold images

$$
\begin{aligned}
&P_b = \left\{\forall I_n(x,y) \middle| \left| I_n(x+m, y+n, t-1) - \mu(x,y,t)\right| < T_{(t-1,t)}(x,y)\right\} \\
&P_f = \left\{\forall I_n(x,y) \middle| \left| I_n(x+m, y+n, t+1) - \mu(x,y,t)\right| < T_{(t+1,t)}(x,y)\right\}, \\
&m, n = -1,0,1
\end{aligned} \tag{5}
$$

$$
P_u = P_b \cup P_f
$$

The temporal filtering result $\mu_u(x,y,t)$, which will replace each pixel value in the following stage is obtained as the average value of all pixels in P_u as shown in (6).

$$
\mu_u(x,y,t) = \frac{1}{NUP} \sum_{m=1}^{NUP} P_u(m) \tag{6}
$$

Here, NUP is the total number of used pixel in the temporal filtering process for a given (x,y,t) pixel.

2.3 Interpolation of Temporally Filtered and Noisy Image Data

Preservation of edge information is the main problem in image and video denoising techniques. In this work, interpolation is carried out between temporally filtered and degraded images taking the block based standard deviation into account to retain edges successfully. The final filtered image frame I_o' is constructed as given in (7).

$$I_o'(x,y,t)=\begin{cases} \mu_u(x,y,t) & , \quad \sigma_u(x,y,t)<T_a(t) \\ \dfrac{I_n(x,y,t)\times\dfrac{\sigma_u(x,y,t)}{T_a(t)}+\mu_u(x,y,t)}{\dfrac{\sigma_u(x,y,t)}{T_a(t)}+1} & , \quad \text{otherwise} \end{cases} \tag{7}$$

In (7), $\sigma_u(x,y,t)$ shows the standard deviation of P_u for each pixel. The standard deviation value of P_u for each pixel is calculated as given in (8).

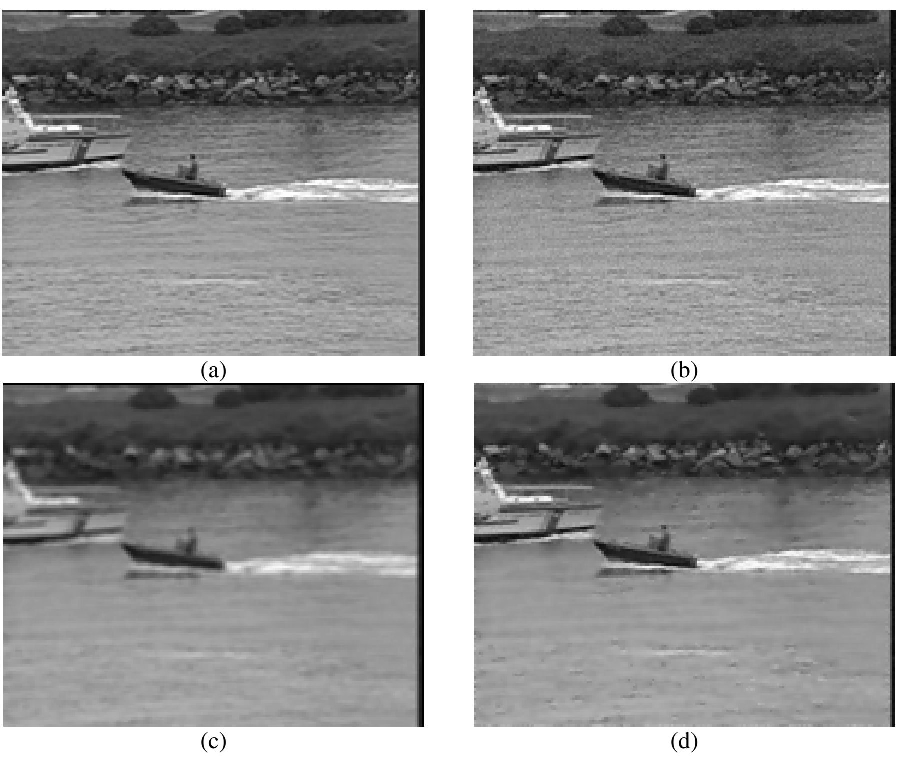


Fig. 4. Coastguard sequence a) original frame #25, b) $\sigma=5$ AWGN added frame #25, c) temporal filtering result, and d) output of the proposed filter

$$\sigma_u(x,y,t)=\sqrt{\frac{1}{NUP}\sum_{m=1}^{NUP}\left[P_u(m)-\mu_u(x,y,t)\right]^2} \tag{8}$$

T_a is an adaptive threshold that is determined for the entire image using (9).

$$T_a(t)=\frac{1}{w\times h}\sum_{m=1}^{w}\sum_{n=1}^{h}\sigma_u(m,n,t) \tag{9}$$

If $\sigma_P(x,y,t)<T_\alpha(t)$ less detail around the pixel region is detected and $I_o^{'}(x,y,t)$ is kept as $\mu_u(x,y,t)$. Interpolation is carried out if $\sigma_P(x,y,t)\geq T_\alpha(t)$ to preserve the spatial details in the image frame. In Fig. 4, the temporally filtered image frame and result of spatial interpolation are given for frame #25 of the Coastguard sequence with $\sigma=5$ AWGN added.

In this figure smoothed details are successfully retained overall but on the sea area some details can not preserved effectively.

3 Experimental Results

The standard deviation value of noise in archive videos is generally low. Therefore, the proposed method is examined for $\sigma=2$, $\sigma=5$ and $\sigma=10$ values of AWGN, artificially introduced into the first 100 frames of commonly used test sequences. To compare the objective performance of the proposed method against several pixel domain methods, Zlokolica's [7] and Chan's [11] pixel domain methods as well as the Wiener filter [12], and the Peak Signal to Noise Ratio (PSNR) measure is used. The PSNR is calculated as given in (10).

$$MSE=\frac{1}{w\times h}\sum_{x=1}^{w}\sum_{y=1}^{h}\left[I_o^{'}(x,y)-I_o(x,y)\right]^2$$
$$PSNR=20\log_{10}\frac{255}{\sqrt{MSE}} \tag{10}$$

Table 1. Average PSNR (dB) values of compared methods for various test sequences (AWGN with $\sigma=2$)

Sequence	Wiener [12]	Chan [11]	Zlokolica [7]	**Proposed**
Akiyo	39.01	39.69	38.20	42.97
Coastguard	30.01	35.76	30.23	32.32
Foreman	34.30	37.66	35.18	38.78
Hall Monitor	35.07	38.89	33.86	39.48
Mother	37.00	38.77	37.79	41.49
Salesman	33.21	36.94	35.49	36.22
Silent	33.65	36.77	33.87	38.00
Average	34.60	37.78	34.95	**38.47**

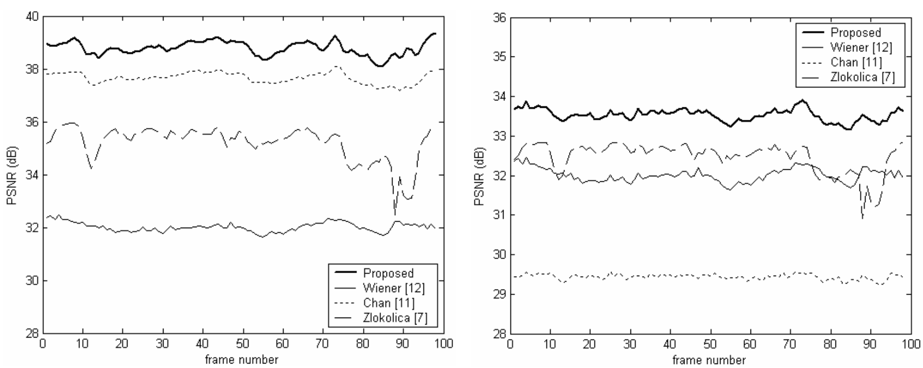
T_i and T_s thresholds are experimentally set to 20 in the proposed method. It is seen from Table 1, Table 2 and Table 3 that the proposed method gives the highest average PSNR values for all standard deviation values of the noise. PSNR graphics are given for AWGN added Foreman sequence ($\sigma = 2$, $\sigma = 5$) in Fig. 5 to evaluate

Table 2. Average PSNR (dB) values of compared methods for various test sequences (AWGN with $\sigma = 5$)

Sequence	Wiener [12]	Chan [11]	Zlokolica [7]	**Proposed**
Akiyo	37.38	36.31	36.83	39.57
Coastguard	29.74	33.74	29.97	31.67
Foreman	33.67	35.08	34.45	36.78
Hall Monitor	34.26	35.86	33.39	37.36
Mother	35.97	35.90	36.57	38.53
Salesman	32.60	34.72	34.63	34.80
Silent	33.18	34.53	33.38	36.22
Average	33.83	35.16	34.17	**36.42**

Table 3. Average PSNR (dB) values of compared methods for various test sequences (AWGN with $\sigma = 10$)

Sequence	Wiener [12]	Chan [11]	Zlokolica [7]	**Proposed**
Akiyo	34.23	29.77	33.92	35.07
Coastguard	28.91	29.00	29.11	30.25
Foreman	32.01	29.43	32.46	33.53
Hall Monitor	32.27	29.75	31.91	34.09
Mother	33.60	29.71	33.73	34.36
Salesman	31.02	29.38	32.48	32.20
Silent	31.86	29.28	31.71	33.17
Average	31.99	29.47	32.19	**33.24**



(a) (b)

Fig. 5. PSNR graphics for Foreman sequence AWGN added with a) $\sigma = 2$, b) $\sigma = 10$

the characteristics of the methods in case of different standard deviation values of the AWGN and local motion.

It is clearly seen from Fig. 5-a that the proposed method gives better results than the compared methods for the case of AWGN with $\sigma = 2$. The Wiener filter gives poor results compared to the others. PSNR results of Zlokolica's method are on the average lower by 3.5dB mainly because this method distorts the shape of image regions in case of local motion. Fig. 5-b presents the performance results of the compared techniques for the case of AWGN with $\sigma = 10$. It is obviously seen from this figure that the proposed method outperforms compared methods. Performances of the Wiener filter and Zlokalica's method increase with higher standard deviation values of the AWGN but Zlocalica's method is still influenced by local motion. Chan's method gives poor results for higher standard deviation values of AWGN compared to the other methods.

4 Conclusion

A new pixel domain spatio-temporal video noise filter for archive video restoration has been proposed in this paper. The proposed filtering method takes motion changes and spatial information into account. Initially, temporal filtering is carried out considering temporal changes for this purpose. Then, interpolation is utilized between degraded and temporally filtered images to preserve edge information taking standard deviation values into account. Filtering results are compared with several pixel domain filtering methods and the results show that the proposed method outperforms compared pixel domain filtering techniques in terms of PSNR and provides successful visual results for archive film.

Acknowledgement

This work was supported by the Turkish Scientific and Technical Research Council (Tübitak), under grant EEEAG/103E007.

References

1. Malfait M., and Roose D.: Wavelet-Based Image Denoising Using a Markov Random Field a Priori Model. IEEE Trans. on Image Processing, Vol. 6, No. 4, (1997) 549-565
2. Kazubek M.: Wavelet Domain Image Denoising by Thresholding and Wiener Filtering. IEEE Signal Processing Letters, Vol. 10, No. 11, (2003) 324-326
3. Van De Ville D., Nachtegael M., Van der Weken D., Kerre E. E., Philips W., and Lemahieu I.: Noise Reduction by Fuzzy Image Filtering. IEEE Trans. on Fuzzy Systems, Vol. 11, No. 4, (2003) 429-436
4. Sadhar S. I., and Rajagopalan A. N.: Image Estimation in Film-Grain Noise. IEEE Signal Processing Letters, Vol. 12, No. 3, (2005) 238-241
5. Ozkan M. K., Sezan I., and Tekalp A. M.: Adaptive Motion Compensated Filtering of Noisy Image Sequences. IEEE Trans. on Circuits and Systems for Video Technology, Vol. 3, No. 4, (1993) 277-290

6. Dugad R., and Ajuha N.:Video Denoising by Combining Kalman and Wiener Estimates. IEEE Int. Conf. on Image Processing, Kobe, Japan, 4, (1999) 152-161
7. Zlokolica V., Philips W., and Van De Ville D.: A New Non-Linear Filter for Video Processing. 3rd IEEE Benelux Signal Processing Symposium (SPS-2002), Leuven, Belgium, (2002) 13-16
8. Selesnick I. W., and Li K. Y.: Video Denoising Using 2d and 3d Dual-Tree Complex Wavelet Transforms. In Wavelet Applications in Signal and Image Processing (SPIE 5207), San Diego, 5207, (2003) 607-618
9. Pizurica A., Zlokolica V., and Philips W.: Combined Wavelet Domain and Temporal Video Denoising. IEEE Conf. on Advanced Video and Signal Based Surveillance (AVSS'03), (2003) 334-341
10. Gupta N., Swamy M. N. S., and Plotkin E. I.: Low-Complexity Video Noise Reduction in Wavelet Domain, IEEE 6th Workshop on Multimedia Signal Processing, (2004) 239-242
11. Chan T.-W., Au O. C., Chong T. S., and Chau W.-S.: A Novel Content-Adaptive Video Denoising Filter. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP), Philadelphia, PA, USA, Vol. 2, (2005) 649-652
12. Lim J. S.: Two-Dimensional Signal and Image Processing. Englewood Cliffs, NJ: Prentice Hall, (1990) 536-540
13. Press W. H., Teukolsky S. A., Vetterling W. T., and Flannery B. P.: Numerical Recipes in C, 2nd edition New York: Cambridge Univ. Press, (1992)
14. Brailean J., Kleihorst R., Efstratiadis S., Katsaggalegos A., and Lagendijk A.: Noise Reduction Filters for Dynamic Image Sequences: A Review, Pcoc. Of IEEE 83 (9), (1995) 1272-1292