Spatio-Temporal Video Denoising by Block-Based Motion detection

Seema Mishra[1], Preety D Swami[2]
[1]*PG Student, Department of Electronics and communication Engineering, S.A.T.I. Vidisha, INDIA (M.P.)*
[2]*Associate Professor, Department of Electronics and Instrumentation Engineering, S.A.T.I. Vidisha, INDIA (M.P.)*

***Abstract* – This paper proposes a new video denoising technique where spatially adaptive noise filtering in wavelet (transform) domain is combined with temporal filtering in signal domain. AWGN is being considered which behaves as Gaussian random variable. In this paper, spatial filtering of individual frames is done in the wavelet domain, and the filtering between the frames is done by recursive temporal filter. Spatial filtering is done by taking wavelet transform of individual frames and then modifying the wavelet coefficients by spatially adaptive bayesian wavelet shrinkage method. The denoising artifacts and residual noise differ from frame to frame which produces unpleasant visual effect. Hence filtering in time domain is essential. Temporal filtering is based on a simple block based motion detector and on selective recursive time averaging of frames. This technique outperforms sequential spatio-temporal filters, 2-D spatial filters and 3-D (spatio-temporal) in terms of visual quality as well as quantitative (PSNR) performance measures**.

***Keywords:* Motion detection, Recursive temporal filtering, Spatial adaptive Bayesian shrinkage, Video denoising, Wavelet transform.**

I. INTRODUCTION

Video signals are often distorted by noise during acquisition and transmission [1]. There are certain noise sources, located in camera hardware which becomes prevalent under bad lighting condition. Noise reduction is highly desirable in many applications, e.g. for improving visual quality in video surveillance, television, teleconferencing and medical imaging; for video coding and as a preprocessing step for improving the accuracy of subsequent processes like object detection, feature extraction etc. [2].

Video denoising algorithms can be classified into two categories: spatial domain and transform domain.
Weighted averaging within local 2-D or 3-D windows is usually done in spatial domain denoising, where the weights may be either fixed or adapted based on local image characteristics. A review of spatial domain filtering methods is given in [3]. In transform domain, the noisy signal is first decorrelated using a linear transform (e.g., wavelet transform) and then transform coefficients of original signal are recovered (e.g., by hard/soft thresholding [4] or Bayesian estimation [5]), followed by inverse transform that produces the signal back into spatial domain. There is high degree of correlation between adjacent frames which provides additional information for estimating the original signal. However, the process becomes complicated due to the presence of motion between frames. Motion information or temporal correlation may be incorporated into the algorithms by motion estimation/compensation techniques [6] or simple pixel based motion detection and performing some special operation in case of detected motion [7]. One of the common approach is Spatio-temporal denoising which exploits both spatial and temporal correlation in video sequences to reduce noise. 2-D spatial filter and 1-D temporal filter are applied either sequentially [7] or separately [8]. Pizurica et.al. [7] proposed a technique where 2-D wavelet denoising is followed by selective, recursive time averaging. Temporal filter reduces the denoising artifacts of 2-D wavelet filter and residual noise present between the frames.

In this paper, sequential wavelet domain and temporal filtering approach is used. Spatial filtering of individual frames is done in the wavelet domain, which is combined with filtering in time domain.

Spatial filtering is done by taking wavelet transform of individual frames and then modifying the wavelet coefficients by spatially adaptive bayesian wavelet shrinkage method, followed by an inverse wavelet transform. The denoising artifacts and residual noise differ from to frame which degrades the visual quality. Hence temporal filtering is combined with wavelet domain denoising. Temporal filtering is based on a simple block based motion detection and selective recursive time averaging of spatially filtered frames. The recursive filter resets at the positions where motion is detected to avoid edge blurring.

The rest of this paper is organized as follows. In Section 2, the general theory of discrete wavelet transform and redundant wavelet transform is reviewed. Section 3 presents a wavelet based denoising technique based on Bayesian shrinkage. Section 4 addresses temporal filtering. In Section 5, the proposed technique is elaborated. Section 6 presents the results and comparison. The concluding remarks are in Section 7.

II. WAVELET TRANSFORM

In this section we briefly review the wavelet decomposition and its use in noise filtering. A comprehensive review of wavelets can be seen in [9]-11].

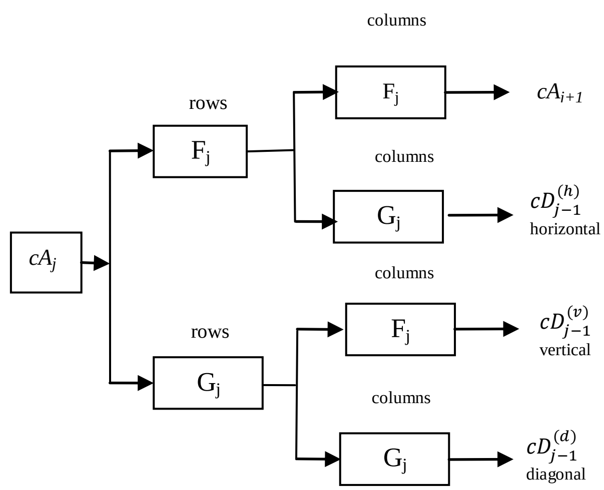
A. Discrete Wavelet Transform

The Discrete wavelet transform (DWT) can be seen as a filter bank algorithm that is iterated on low pass output [10]. In the discrete wavelet transform, an image signal is passed through an analysis filter bank followed by decimation operation. The analysis filter bank consists of low-pass and high-pass filter at each level of decomposition. When the signal passes through these filters, it splits into bands. The low pass filtering produces an approximation of the signal, while the high pass filtering reveals the details that are expressed by wavelet coefficients. For reconstruction, the approximation and detail coefficient are upsampled and then filtered with a low pass and high pass filter followed by summation of the output.

The above described Discrete wavelet transform is critically sampled (non-redundant) and it is well known that the noise suppression improves when it is implemented in redundant representation.

B. Redundant wavelet transform

The redundant wavelet transform (RWT) is a wavelet transform algorithm designed to overcome the lack of translation-invariance of the discrete wavelet transform which is achieved by removing the downsamplers and upsamplers in the DWT and upsampling the filter coefficients by a factor of $2^{(j-1)}$ in the j^{th} level of the algorithm. The output of each level of RWT contains the same number of samples as the input. Hence, for a decomposition of N levels there is a redundancy of N in the wavelet coefficients. Both, the approximation and detail coefficients at level 1 are of size *N*, which is the signal length. The common step *j* convolves the approximation coefficients at level *j*–1, with up sampled versions of the appropriate original filters, to generate the approximation and detail coefficients at level *j* [12]. The redundant wavelet transform at first level (*j=1*) is given in Fig.1.



where rows

$\boxed{X}$ Convolve with filter X the rows of input signal

columns

$\boxed{X}$ Convolve with filter X the columns of input signal

Fig. 1. 2D – Redundant wavelet transform at the first level

III. NOISE FILTERING IN WAVELET DOMAIN

Here the wavelet based denoising technique that is considered in our method is briefly reviewed.

A. The marginal statistics of image wavelet coefficients

The histograms of wavelet coefficient in each subband for natural noise-free images are usually long-tailed and sharply peaked at zero. These are commonly modelled by generalized Laplacian(also called generalized Gaussian) density [10]

$$p(y) = \frac{\lambda\nu}{2\Gamma\left(\frac{1}{\nu}\right)} \exp(-\lambda|y|^{\nu}), \qquad \lambda, \nu > 0, \quad (1)$$

where $\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$ is the Gamma function. The shape parameter ν for natural images is typically $\nu \in [0,1]$. The kurtosis and the variance of a generalized Laplacian signal are [4]

$$\kappa_y = \frac{\Gamma\left(\frac{1}{\nu}\right)\Gamma\left(\frac{5}{\nu}\right)}{\Gamma^2\left(\frac{3}{\nu}\right)}, \qquad \sigma_y^2 = \frac{\Gamma\left(\frac{3}{\nu}\right)}{\lambda^2\,\Gamma\left(\frac{1}{\nu}\right)} \qquad (2)$$

The model parameters ν and λ, in case of additive white Gaussian noise are estimated from the noisy coefficient histogram using the following equations [4,13]

$$\frac{\Gamma\left(\frac{1}{\nu}\right)\Gamma\left(\frac{5}{\nu}\right)}{\Gamma^2\left(\frac{3}{\nu}\right)} = \frac{m_4 + 3\sigma^4 - 6\sigma^2\sigma_\omega^2}{(\sigma_\omega^2 - \sigma^2)^2},$$

$$\lambda = \left((\sigma_\omega^2 - \sigma^2)\frac{\Gamma\left(\frac{1}{\nu}\right)}{\Gamma\left(\frac{1}{\nu}\right)}\right)^{-1/2}, \qquad (3)$$

where σ is the standard deviation of noise, σ_ω^2 and $m_{4,\omega}$ are variance and the fourth moment of the noisy histogram, respectively. The scale parameter is then estimated as [4]

$$\lambda = [0.5(\sigma_\omega^2 - \sigma^2)]^{-1/2} \qquad (4)$$

B. Denoising by wavelet shrinkage

Let us assume that each input video frame **f = [**f_1 **…fn]** is contaminated with additive white Gaussian noise of zero mean and variance σ^2. Since the wavelet transform has linear property, the noise remains additive in the transform domain as well

$$\omega_i = y_i + \epsilon_i, \qquad i = 1, \ldots, n \quad (5)$$

where y_i are unknown noise-free wavelet coefficients and ϵ_i are noise contributions. In case of orthogonal wavelet transform, ϵ_i are independent identically distributed normal random variables such that $\epsilon_i \sim N(0, \sigma^2)$. If the input standard deviation σ is not known, one can easily estimate it as the median absolute deviation of the highest- frequency subband [14].

In spite of the type of employed discrete wavelet transform(e.g., non-redundant or non-decimated), noise reduction is generally done by wavelet shrinkage: the magnitude of each coefficient is decreased by a given amount depending on the noise level and depending on how likely it is that a given coefficient represents an actual discontinuity. Thresholding [4], [14] is common shrinkage approach, which puts the wavelet coefficients with "small" magnitudes to zero while keeping the remaining ones unchanged ("hard-thresholding") or shrinking in magnitude the remaining ones ("soft-thresholding"). Thresholding with uniform threshold in each subband is attractive due to its simplicity. But, the performance is inadequate and the denoising quality is often not satisfactory. Apart from this, many spatially adaptive wavelet shrinkage methods [13], [15], [16], [17] have been developed over recent years.

In this paper, spatially adaptive Bayesian wavelet shrinkage is considered which makes use of prior distributions of noise-free coefficients.

C. Spatially adaptive Bayesian shrinkage approach

In this approach, each wavelet coefficient is shrinked according to probability that it represents a "signal of interest" [7], [15]. The imperative coefficient whose magnitude is above a certain threshold is defined as the signal of interest. Three parameters are considered to estimate how probable it is that a coefficient represents a "signal of interest" based on [7], [15]

- the wavelet coefficient value ϵ_l
- a local spatial activity indicator (LSAI), η_l used as the locally averaged coefficient magnitude in a small window $\delta(l)$: $z_l = \sum_{k\in\delta(l)} |\omega_k|$ and
- the global subband statistic, ρ.

The following shrinkage estimator [7] is given by considering two hypotheses H_1 – "signal of interest present: |y>T" and H_0 – "signal of interest absent: |y| < T"

$$\hat{y}_l = \frac{\rho\epsilon_l\eta_l}{1+\rho\epsilon_l\eta_l}\omega_l \qquad (6)$$

where

$$\rho = \frac{P(H_1)}{P(H_0)}, \quad \epsilon_l = \frac{p(\omega_l/H_1)}{p(\omega_l/H_0)} \text{ and } \quad \eta_l = \frac{p(z_l/H_1)}{p(z_l/H_0)}$$

Here $p(\omega_l/H_1)$ and $p(\omega_l/H_0)$ represent the conditional probability density functions of the noisy coefficients given the presence and absence of a signal of interest. Correspondingly, $p(z_l/H_1)$ and $p(z_l/H_0)$ denote the conditional probability density functions of the local spatial activity indicator. The required ratios ρ, ϵ_l, η_l are directly estimated from the observed image coefficients.

IV. TEMPORAL FILTERING

Temporal filtering is an approach of exploiting temporal correlation to reduce noise in a video sequence. A video sequence contains not only spatial correlation but also temporal correlation between consecutive frames. Temporal video denoising methods can remove the artifacts caused by spatial methods by tracking object motions through frames and thus make certain temporal consistency.

There are two types of temporal filtering techniques: the techniques [7] where motion is detected first and then some special operation is performed in case of detected motion. The motion detector evaluates the absolute difference between the pixel value from the current frame and previous frame, where the considered pixels have the same spatial positions. If the absolute difference is above a predefined threshold then the temporal filtering is switched off otherwise recursive time averaging of frames is done.

Motion estimation/compensation techniques [6] attempt to better exploit the considerable temporal redundancy in video by temporally smoothing pixel values over their estimated motion trajectories. Generally, it is often impossible or impractical to determine the exact temporal correspondences between consecutive frames for all pixels due to the absence of information in the video content (e.g., occlusion), imperfections of the motion estimates. When motion estimation fails, motion-compensated temporal filtering techniques can produce disturbing artifact.

In the proposed sequential spatio-temporal filtering, motion detection and selective recursive temporal filtering is performed over spatially denoised frames. The motion adaptive temporal filtering is benefited from the use of high-quality spatial denoising [7]. The inter-frame differences due to remaining noise and artifacts after spatial wavelet denoising are relatively small compared to the actual inter-frame differences produced by motion. Therefore, simple block based motion detection technique is considered. The recursive motion adaptive filtering is given in Fig. 2

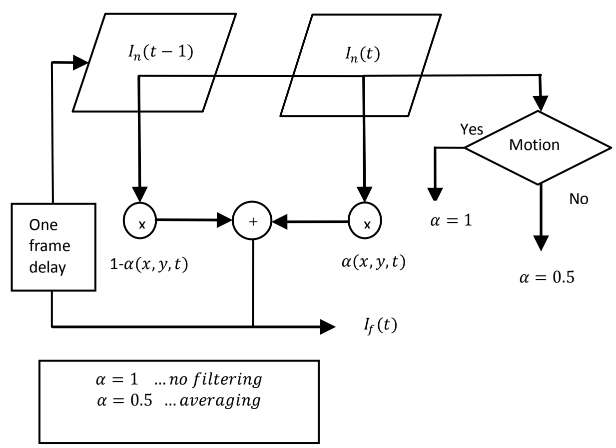


Fig. 2. Recursive motion adaptive filtering

V. PROPOSED VIDEO DENOISING ALGORITHM

In this section the proposed video denoising method is discussed. The original video frames g, of size $m \times n$ are contaminated by additive white Gaussian noise with zero mean and standard deviation σ. The denoising algorithm is performed in two steps; Step 1: Spatial filtering of individual frames in wavelet domain, Step 2: Temporal filtering by block based motion detection and recursive time averaging of the spatially filtered frames. The block diagram of the proposed method is given in Fig. 3.

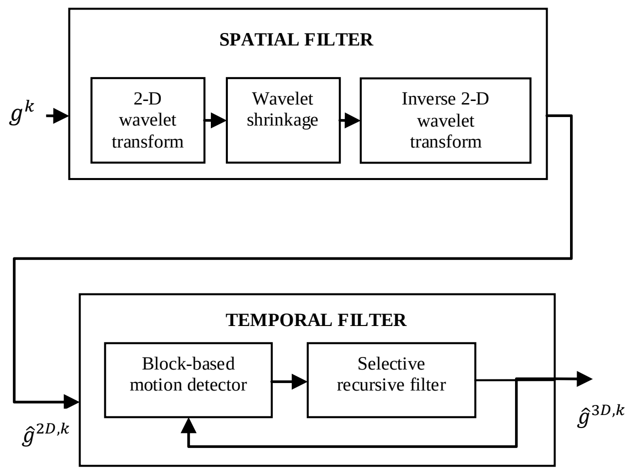


Fig. 3. Block diagram of proposed video denoising algorithm

Step 1: In the first step, the frames of size $m \times n$ are converted to images of size 512×512. The redundant wavelet transform is applied to each frame. The frames are decomposed to four levels. The wavelet used is symmlet [9], [10] with four vanishing moments. The wavelet coefficients obtained at each level of decomposition are shrinked by adaptive bayesian wavelet shrinkage method [7]. The modified coefficients are reconstructed by taking inverse wavelet transform. The reconstructed images are converted back to frames of original sizes.

Step 2: It is well known in video denoising literature that spatial denoising produces disturbing artifacts and unpleasant visual quality [3]. This is due to the fact that residual noise and annoying artifacts differ from frame to frame causing unpleasant "flickering" effect. In the proposed video denoising algorithm (Fig. 3), a temporal filter reduces the residual noise and artifacts produced by the 2-D wavelet domain filter. In the second step, temporal filtering is based on a simple block based motion detection and recursive time averaging of spatially filtered frames. The recursive filtering is switched off at those positions where motion is detected. A brief description follows.

Let $\mathbf{g^k}$ denote the k-th frame of a noise-free video sequence and $\mathbf{d^k} = \mathbf{g^k} + \mathbf{n^k}$ the corresponding noisy frame, where $\mathbf{n^k}$ is the noise field. And let the k-th 2-D denoised frame is denoted by

$$\hat{\boldsymbol{g}}^{2D,k} = [\hat{g}_1^{2D,k}, \dots, \hat{g}_L^{2D,k}] \qquad (7)$$

Each denoised frame is divided into blocks of size 4×4. Firstly, we calculate the MAD between the pixels in the corresponding blocks in the current and the previous frames

$$D_{i,j}^k = \frac{1}{N^2} \sum_{m=1}^{N} \sum_{n=1}^{N} \left| g_{m,n}^{k,i,j} - g_{m,n}^{k-1,i,j} \right| \qquad (8)$$

where k is the frame number, i, j, are the spatial coordinates of a block, m, n are the coordinates of a pixel inside the block, and N is the block size. In the filtering step, we determine whether motion exists in each block by comparing the absolute block difference with a threshold T.

The motion field $\boldsymbol{m}_{i,j}^{k,}$ of the k-th frame with respect to the previous frame is defined as follows

- $m_{i,j}^k = 0$ if there is no (considerable) motion at the block position i, j from the frame k-1 to the frame k, meaning that $g_{i,j}^k \approx g_{i,j}^{k-1}$.

- $m_{i,j}^k = 1$ if there is motion at the block position i, j from the frame k-1 to the frame k, meaning that $g_{i,j}^k$ and $g_{i,j}^{k-1}$ differ significantly.

This motion field is estimated from the denoised frames as

$$\widehat{m}_{i,j}^{k} = \begin{cases} 0, & if \ \left|\hat{g}_{i,j}^{2D,k} - \hat{g}_{i,j}^{3D,k-1}\right| < T \\ 1, & otherwise \end{cases} \quad (9)$$

here T is a threshold that is selected as T = σ. Recursive time averaging is applied at the spatial positions where no motion was detected, yielding the final 3D filtered block:

$$\hat{g}_{i,j}^{3D,k} = \begin{cases} \alpha\hat{g}_{i,j}^{2D,k} + (1-\alpha)\hat{g}_{i,j}^{3D,k-1}, & m_{i,j}^{k} = 0, \\ \widehat{g}_{i,j}^{2D,k}, & otherwise, \end{cases} \quad (10)$$

where $0 \leq \alpha \leq 1$. This recursive filter accumulates and averages the pixel intensities of a given block from all the previous frames if the motion was not present at that block. The detection of a motion resets the filter.

VI. RESULTS AND COMPARISON

The performance of the proposed denoising algorithm is tested on four different gray-scale videos: "Tennis," "Miss America," "Foreman," and "Salesman." These test videos have been corrupted with Gaussian noise of the following standard deviation values: σ = 10, 15, 20, 25, 30. The spatial filtering part is implemented with a non-decimated wavelet transform with 4 decomposition levels. The wavelet used is 'sym4' for decomposition instead of 'sym8' used by the authors in [7]. The prior assumed to model the noise-free wavelet coefficients is generalized Laplacian of (1). The optimal value of the threshold in terms of the mean squared error is T = σ. For most test video frame images and noise levels, the window size is 5x5. The block based motion adaptive recursive temporal filter involves two parameters: the motion threshold T and the weighting parameter α for the recursive filtering. In our experiments, we set the weighting parameter α to a constant value α = 0.6 and the optimal value of threshold is selected as T = σ. Quantative quality evaluations of the denoising results were employed by two objective criteria, namely the Peak Signal to Noise Ratio (PSNR) and the Structural Similarity (SSIM) index [18]. Specifically, PSNR is defined as

$$\text{PSNR} = 20\log_{10}\left(\frac{255}{RMSE}\right) \quad (11)$$

Where RMSE is root mean squared error between the noise-free and denoised frame. SSIM is calculated within local windows using

$$\text{SSIM}(\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (12)$$

where **x** and **y** are the image patches obtained from the local window from the original and contaminated images. σ_x^2, μ_x, and σ_{xy} are the variance, mean, and cross- correlation computed within the local window, respectively. SSIM is a better indicator of perceived image quality. The PSNR and SSIM results for the denoised video sequences are calculated as the frame average of the full sequence. Table 1 shows the experimental results in terms of PSNR and SSIM. Table 2 shows the comparison of denoising results of proposed method for three noise levels (σ = 10, 15, 20) with Wiener3D filtering [19], Wavelet2D filtering, and Sequential wavelet domain temporal filtering SEQWT [7] in terms of PSNR and SSIM. The denoising results for Wiener3D were obtained from the processed sequences available in [19], and the other results were computed by ourselves.

It is very clear from the comparative table that only for higher noise deviations (σ = 15, 20), PSNR of denoised frames Miss America sequence are low as compared with Wiener3D filtering method while for low noise deviations, it is the highest among all the methods under comparison. Also, SSIM of denoised frames Miss America sequence for low noise deviation (σ = 10, 15) are low as compared with Wiener filtering method while for high noise deviation, it is the highest among all the methods under comparison. PSNR and SSIM of denoised frames of all other video sequences under observation of proposed method for all noise deviations is high as compared to all the methods under comparison. Nevertheless, the proposed method yielded an improved visual quality of all the tested sequences. Denoised frame extracted from Miss America, Salesman, Foreman, and Tennis sequences by using different methods (Wavelet2D filtering, SEQWT, and proposed method), together with a noisy version of the same frame are given in Fig. 4, Fig. 5, Fig. 6 and Fig. 7 respectively. The quantitative performance of the proposed method on the parts of Miss America, Salesman, and Foreman and Tennis sequences in comparison with SEQWT filter is given in Fig. 8.

Fig. 4. Test frame images (a) 50th frame of Miss America (b) 18th frame of Salesman (c) 7th frame of Foreman (d) 30th frame of tennis.

TABLE I
NUMERICAL RESULTS ON THE TEST VIDEO FRAMES AVERAGED OVER 50 FRAMES

Noise deviation (σ)	Threshold parameter(T)	Salesman		Miss America		Foreman		Tennis	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
σ = 10	T = 10	32.72	0.911	36.97	0.897	33.06	0.902	29.51	0.801
σ = 15	T = 15	31.75	0.880	35.24	0.859	32.34	0.853	28.06	0.722
σ = 20	T = 20	30.38	0.842	33.74	0.814	30.77	0.809	27.17	0.670
σ = 25	T = 20	29.27	0.804	32.42	0.767	29.06	0.764	26.39	0.621
σ = 30	T = 30	28.33	0.765	31.22	0.716	28.61	0.722	25.71	0.578

Table II.
COMPARISON BETWEEN THE WIENER3D FILTERING METHOD [19], WAVELET2D FILTERING METHOD, SEQUENTIAL WAVELET DOMAIN TEMPORAL FILTERING SEQWT [7], AND PROPOSED METHOD IN TERMS OF PSNR AND SSIM AVERAGED OVER 50 FRAMES

Video sequence	**Salesman**			**Foreman**			**Miss America**			**Tennis**		
Noise dev.(σ)	**10**	**15**	**20**	**10**	**15**	**20**	**10**	**15**	**20**	**10**	**15**	**20**
						PSNR (dB)						
Wiener3D [19]	29.59	29.30	28.88	29.54	29.26	28.87	36.95	**35.60**	**34.06**	22.88	22.81	22.71
Wavelet2D	31.85	30.34	28.73	32.46	31.21	29.38	35.20	33.02	31.18	29.29	27.55	26.33
SEQWT [7]	32.33	31.06	29.66	33.03	31.73	30.12	36.30	34.60	33.04	29.30	27.71	26.66
Proposed	**32.72**	**31.75**	**30.38**	**33.06**	**32.34**	**30.77**	**36.97**	35.24	33.74	**29.51**	**28.06**	**27.17**
						SSIM						
Wiener3D [19]	0.839	0.818	0.786	0.865	0.839	0.808	**0.907**	**0.868**	0.808	0.577	0.560	0.539
Wavelet2D	0.888	0.837	0.781	0.872	0.799	0.731	0.850	0.785	0.714	0.798	0.705	0.635
SEQWT [7]	0.901	0.863	0.821	0.883	0.824	0.771	0.882	0.842	0.792	0.799	0.713	0.650
Proposed	**0.911**	**0.880**	**0.842**	**0.902**	**0.853**	**0.809**	0.897	0.859	**0.814**	**0.801**	**0.722**	**0.670**

Fig. 5. Results for the 50^{th} frame of the denoised "Miss America" sequence. (a) Noisy image frame with σ = 15; (b) Wavelet2D filtering, PSNR=32.98dB; (c) SEQWT filter, PSNR=34.57dB; (d) Proposed method, PSNR=35. 26dB.

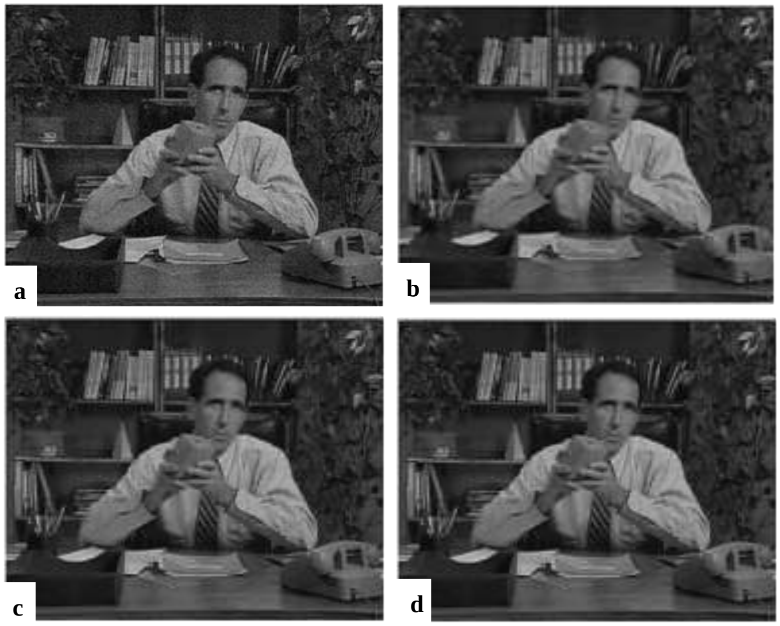


Fig. 6. Results for the 18th frame of the denoised "Salesman" sequence. (a) Noisy image frame with σ = 10; (b) Wavelet2D filterimg , PSNR=31.94dB; (c) SEQWT filter, PSNR=32.39dB; (d) Proposed method, PSNR=32.90dB.

Fig. 7. Results for the 7th frame of the denoised "Foreman" sequence. (a) Noisy image frame with σ = 15; (b) Wavelet2D filtering, PSNR=31.30dB; (c) SEQWT filter, PSNR=31.98dB; (d) Proposed method, PSNR=33.01dB.

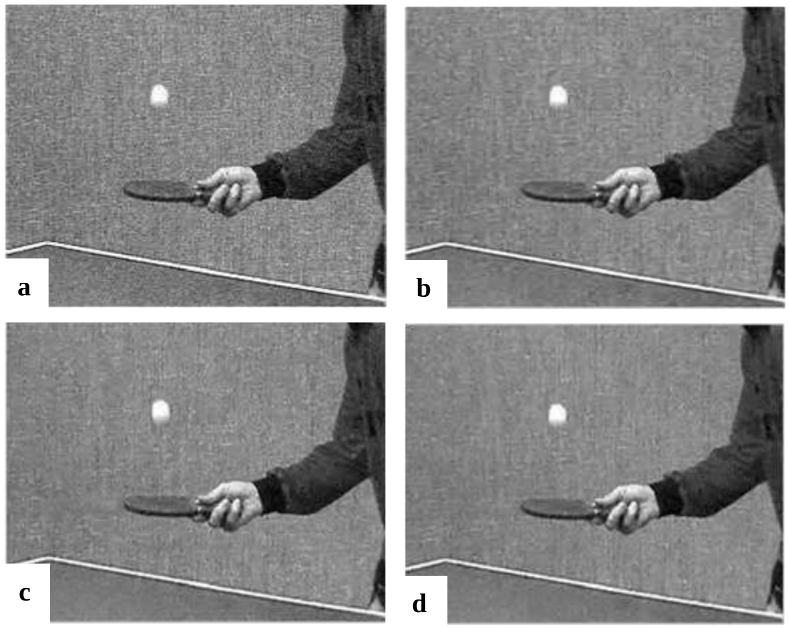
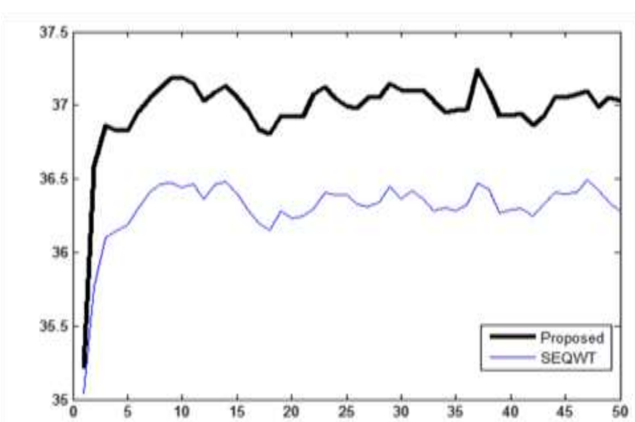


Fig. 8. Results for the 30th frame of the denoised "Tennis" sequence. (a) Noisy image frame with σ = 15; (b) Wavelet2D filtering, PSNR=27.88dB; (c) SEQWT filter, PSNR=27.82dB; (d) Proposed method, PSNR=27.95dB.

MISS AMERICA sequence, σ = 10

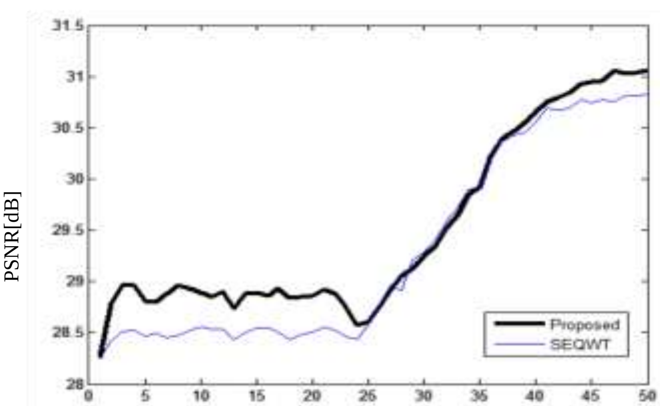
PSNR [dB]



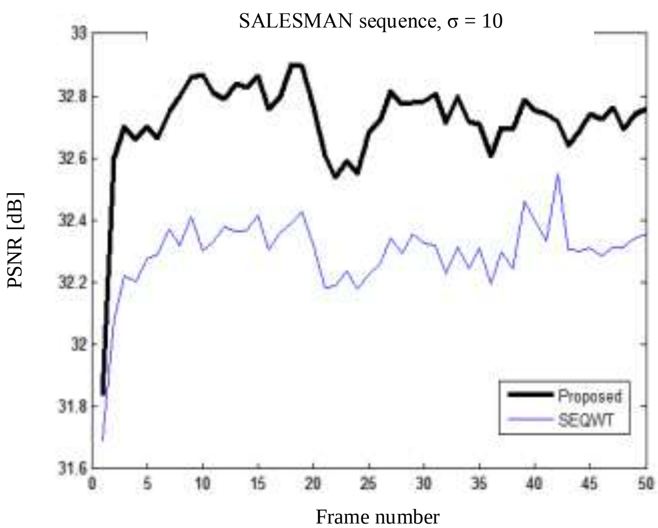
Frame number

TENNIS sequence, σ = 10

PSNR[dB]



Frame number



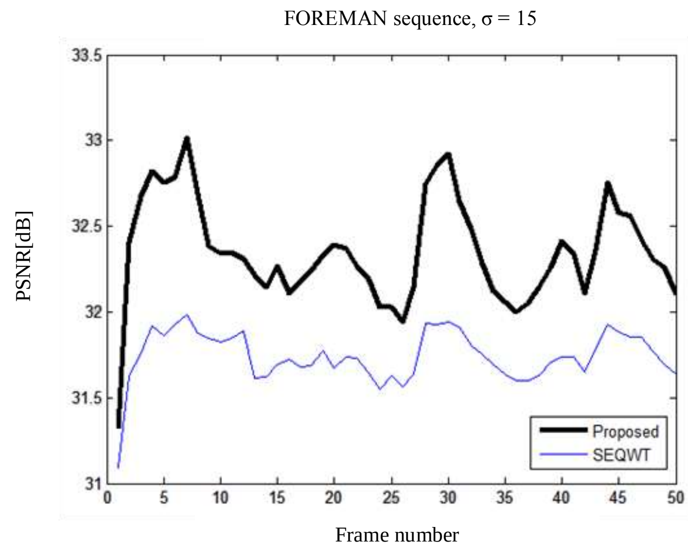


Fig. 9. The quantitative performance of the proposed filter on the parts of the Miss America, Salesman, Tennis, sequences corrupted by additive white Gaussian noise of standard deviation σ = 10 and Foreman sequence corrupted by additive white Gaussian noise of standard deviation σ = 15 in comparison with SEQWT filter [7].

VII. CONCLUSION

In this paper, a video denoising algorithm is proposed where spatially adaptive noise filtering in wavelet (transform) domain is combined with temporal filtering in time domain. In the first stage, spatial filtering is done by taking wavelet transform of individual frames and then modifying the wavelet coefficients by spatially adaptive bayesian wavelet shrinkage method. The 2D adaptive wavelet domain filtering produces unpleasant “flickering” artifacts due to lack of filtering in time. Hence temporal filtering is combined with 2D wavelet filtering to improve the results of the 2D wavelet filtering. Temporal filtering is based on a simple block based motion detector and on selective recursive time averaging of frames. The experimental results show that this combination of the 2D wavelet domain and temporal filtering outperforms 3D (spatio-temporal) and sequential (2D-spatial+1D-temporal) methods for video denoising in terms of both PSNR and SSIM evaluations.

REFERENCES

[1]. P. Karunakaran, S.Venkatraman, I.Hameem Shanavas , and T.Kapilachander, “Denoising of noisy pixels in video by neighborhood correlation filtering algorithm,” *I.J. Image, Graphics and Signal Proc.*, vol. 4, no. 7, pp. 61-67, Jul. 2012.

[2]. V. Zlokolica, A. Pizurica , and W. Philips, “Wavelet-domain video denoising based on reliability measures,” *IEEE Trans. on Circuits and Systems for Video Technology* , vol. 16, no. 8, pp. 993-1007, Aug. 2006.

[3]. J. C. Brailean, R. P. Kleihorst, S. Efstratidis, K. A Katsaggeleos, and R. L. Lagendijk, “Noise reduction filters for dynamic image sequences: A review,” *IEEE Trans. on Image Process.*, vol. 83, no. 9, pp. 1272–1292, Sep. 1995.

[4]. D. L. Donoho, “De-noising by soft- thresholding,” *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 613–627, May 2005.

[5]. E. P. Simoncelli and E. H. Adelson, “Noise removal via Bayesian wavelet coring,” in *Proc. IEEE Int. Conf. Image Proc., Lausanne, Switzerland*, vol. 1, pp. 379–382, Sep.1996.

[6]. V. Zlokolica, “Advanced nonlinear methods for video denoising”, Ph.D. Thesis, Ghent University, 2006.

[7]. A. Pizurica, V. Zlokolica, and W. Philips, "Combined wavelet domain and temporal video denoising," in *Proc. IEEE Conf. Adv. Video Signal- Based Surveillance, Miami, Fla, USA* vol. 1, pp. 334–341, Jul. 2003.

[8]. R. Dugad and N. Ahuja, "Video denoising by combining Kalman and Wiener estimates," in *Proc. IEEE Int. Conf. Image Process, Kobe, Japan*, vol. 4, pp. 156–159, Oct. 1999.

[9]. I. Daubechies, Ten *Lectures on Wavelets*, Philadelphia: SIAM, 1992.

[10]. S. Mallat, *A wavelet tour of signal processing*, Academic Press, London, 1998.

[11]. M. Vetterli and J. Kovacevi´c, *Wavelets and Subband Coding*, Prentice-Hall 1995.

[12]. K. Kannan, S. Arumuga Perumal, and K. Arulmozhi, "Optimal decomposition level of discrete, stationary and dual tree complex wavelet transform for pixel based fusion of multi-focused images, " *Serbian Journal of Electrical Engineering* , vol. 7, no. 1, 81-93, May 2010.

[13]. S.G. Chang, B. Yu, and M. Vetterli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," *IEEE Trans. Image Proc*., vol. 9, no. 9, pp. 1522–1531, Sep. 2000

[14]. D.L. Donoho and I.M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, Aug. 1994.

[15]. A. Pizurica, W. Philips, I. Lemahieu, and M. Acheroy, "A joint inter- and intrascale statistical model for wavelet based bayesian image denoising," *IEEE Trans. Image Proc*., vol. 11, no. 5, pp. 545–557, May 2002.

[16]. A. Pizurica and W. Philips, "Estimating the probability of the presence of a signal of interest in multiresolution single and multiband image denoising," *IEEE Trans. Image Process., vol. 15, no. 3, pp. 654–665,* Mar. 2006.

[17]. M.K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin, "Low-complexity image denoising based on statistical modeling of wavelet coefficients," *IEEE Signal Proc. Lett.*, vol. 6, no. 12, pp. 300–303, Dec.1999.

[18]. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[19]. G. Varghese and Zhou Wang, "Video denoising based on a spatiotemporal Gaussian scale mixture model," IEEE *Trans on Circuits and Systems for Video Technology*, vol. 20, no. 7, Jul. 2010.