

# CHARACTERIZATION OF CORRELATED NOISE IN VIDEO SEQUENCES AND ITS APPLICATIONS TO NOISE REMOVAL

*Nemanja Petrovic, Vladimir Zlokolica, Ljubomir Jovanov, Bart Goossens, Aleksandra Pižurica and Wilfried Philips*

Ghent University, Dept. of Telecommunications and Information Processing (TELIN), IPI,  
Sint-Pietersnieuwstraat 41, 9000 Gent, Belgium

## ABSTRACT

Video sequences in TV and surveillance systems usually contain noise which decreases the visual quality and the performance of various post-processing tasks in the video chain. Usually only white Gaussian noise is assumed within these video application. However in practice that assumption does not always hold and results in poor denoising performance of standard video enhancement algorithm.

In order to solve these problems we propose a new adaptive wavelet-based video denoising method. The method consist of a novel noise modeling scheme and the proposed noise-adaptive spatio-temporal filter. Specifically, the correlated (granulated) noise is characterized by a covariance matrix estimated from the noisy video sequence. Based on the estimated noise covariance and the presence of signal in a spatially local ares, the shrinkage factor for each wavelet band and the spatial position is determined, for spatial denoising. The spatial filtering is followed by the recursive temporal filtering in order to remove the remainder of noise.

## 1. INTRODUCTION

Video sequences are used within a number of applications such as broadcasting, video-phone, tele-conferencing systems, satellite observations or surveillance systems, auto-navigation and medical imaging. The noise can be introduced during acquisition, recording and transmission. Certain noise sources are located in the camera (acquisition noise) hardware and become amplified under bad lighting conditions. Other noise sources are due to transmission over analogue channels, e.g., satellite or terrestrial broadcasting. Furthermore noise can be introduced into the signal by video recording devices.

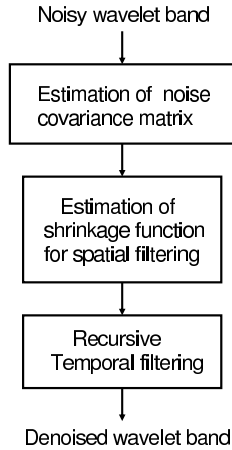
Noise in video sequence is most often modeled with stationary, additive white Gaussian noise. From theoretical point of view white Gaussian noise is a good model for noise in electronic circuitry (which is essentially random process), because according to central limit theorem each process which is the sum of infinite number of random processes converges to Gaussian distribution. Recently, there

has been a considerable amount of algorithms developed for noise level estimation of white Gaussian noise in video sequences [1, 2, 3].

A correctly estimated noise level of the white Gaussian noise is necessary for superior performance of denoising methods for video [4, 5, 6, 7] and for still images [8, 9, 10]. However, often in practice in real noisy scenarios, the noise is spatially correlated and has to be modeled in a more sophisticated manner for superior denoising results [11]. One way of modeling correlated noise is through estimation of noise covariance matrix [12, 13]. Such estimated covariance matrix can further be used for effective vector-based noise reduction [14] or facilitate easier distinguishment between the signal of interest and noise. An advanced noise modeling and adaptive denoising can not only significantly improve the quality of video but facilitate post-processing tasks in a video chain as well.

In our work we have considered video sequences acquired by surveillance and TV cameras, where the noise in the sequences was either introduced by the camera acquisition or was coming from transmission interferences (in analogue domain) via TV tuner. The camera acquisition noise is mostly white Gaussian noise of fairly low noise levels and it is dependent on the lighting conditions: in regions with poorer lighting conditions relatively higher noise level is introduced than in the well lighted ones. On the other hand, the interferences in the transmission of the TV signal through analogue channels (before it is digitally captured) are mostly due to power and channel signal interferences and due to poor cable connections; the noise there is usually spatially correlated (non-white) with particular harmonics and sometimes locally granulated.

In this paper, we propose a new method for noise modeling and denoising in the wavelet domain, using a non-decimated (redundant) transform [15]. The characterization of the correlated (granulated) noise is done by estimating a noise covariance matrix by the proposed method for noise modeling, where noise is assumed to be spatially loose stationary. The estimated noise covariance matrix is further used for adaptive spatial filtering in the wavelet do-



**Fig. 1.** General scheme for wavelet-based video denoising for each wavelet band.

main. This spatial filtering is followed by the recursive temporal filtering method proposed in [5]. This results in an advanced noise removal efficiency which is superior to the performance of other denoising algorithms which assume only white Gaussian noise and comparable to the other algorithms that take the spatial correlation of the noise into account.

The paper is organized as follows: We explain the general scheme of our video denoising algorithm in Section 2. Specifically, in Section 2.1 we describe the proposed method for noise covariance estimation and in Section 2.2 the proposed spatio-temporal denoising method is explained. Finally, in Section 3 we present experimental results and conclude the paper in Section 4.

## 2. PROPOSED SCHEME FOR VIDEO DENOISING

In this Section we describe the proposed scheme for wavelet based video denoising. First wavelet transform is applied on the noisy input sequence, and the decomposed wavelet bands are spatially filtered. Subsequently, the temporal filtering is applied on spatially filtered sequence. The general scheme for the proposed denoising method, in each wavelet band is shown in Fig. 1. The reason for first applying the proposed spatial filter to noisy sequence is to first remove the correlated noise as much as possible, i.e. to decorrelate noise. After that the temporal recursive filtering developed for Gaussian noise will be efficient and remove the rest of the noise.

The proposed method uses a non-decimated wavelet transform implemented with the à trous algorithm [16]. We apply a two-dimensional (spatial) wavelet transform to each video frame and denote *wavelet bands* of this spatial wavelet transform by  $WB = LL, LH, HL, HH$  for the low-pass

(approximation), horizontal, vertical and diagonal orientation bands, respectively. We use a subscript to denote the noisy or denoised band as follows:  $WB_{sn}$  - noisy band,  $WB_{sf}$  - spatially filtered and  $WB_{stf}$  - spatio-temporally filtered band. Additionally, we denote the spatial position as  $\mathbf{r} = (x, y)$  and frame index (time) as  $t$ . The decomposition level is denoted by a superscript  $(l)$ , where  $l = 1, \dots, N$  ( $1$  denotes the finest scale and  $N$  the coarsest).

### 2.1. Noise Estimation

Estimating the statistical parameters of noise is a common problem in noise filtering. Complexity of finding proper solution is mainly conditioned by the noise nature itself. One of the simplest noise models - white Gaussian noise is characterized with only one parameter, noise variance. On the other hand, colored Gaussian noise which exhibits spatial correlation has to be described in a more complex manner, e.g., through a covariance matrix.

Because correlated noise in general has its specific spatial structure it can be easily misinterpreted as signal of interest. Consequently, it is difficult to distinguishing between the local signal (noise-free) structure and noise itself.

In this paper, we propose a simple and efficient heuristic method for estimating noise covariance matrix. Covariance matrix is estimated in the wavelet domain, in each subband and video frame separately. We assume that local statistical properties of the signal of interest are spatially non-stationary, i.e. differ from one spatial position to another. On the other hand, we assume that noise is spatially stationary throughout the whole wavelet subband.

In the wavelet subbands there is often a considerable amount of regions where the amplitude of the wavelet coefficients, corresponding to the noise-free signal, is negligible to one coming from noise.<sup>1</sup> In our method we aim at detecting these areas and estimating the noise covariance matrix throughout these parts in the wavelet subband, by using their statistics for noise estimation.

Let  $WB_o^{(l)}$  and  $WB_n^{(l)}$  represent wavelet band at scale  $l$  for the corresponding noise-free signal and noise, respectively.<sup>2</sup> Then the wavelet coefficients of the signal corrupted by additive correlated noise can be formulated, as follows:

$$WB_{sn}(x, y, t) = WB_o(x, y, t) + WB_n(x, y, t) \quad (1)$$

Hence, the covariance matrix for wavelet band of the corrupted signal is then determined as:

$$C_{sn} = C_o + C_n. \quad (2)$$

<sup>1</sup>This is a reasonable assumption for many applications, such as TV and video surveillance systems, where there are usually enough spatially homogeneous image regions.

<sup>2</sup>For the sake of clarity, in the remainder of paper, we will not use the superscript  $l$  if not necessary.

where  $C_o$  and  $C_n$  stand for covariance matrices of noise-free signal and noise, respectively. Note that the covariance matrix of noise is the same for the whole wavelet band, while  $C_{sn}$  and  $C_o$  change for different spatial positions.

In general a covariance matrix  $C_{sn}$  can be approximated in a local neighborhood as follows:

$$C_{sn} \approx \frac{1}{K} \sum_{k=1}^K \mathbf{V}_k \bullet \mathbf{V}_k^T. \quad (3)$$

where  $\mathbf{V}_k(x, y, t)$  is a vector of noisy wavelet coefficients:

$$\mathbf{V}_k(x, y, t) = WB_n(x_k + i, y_k + j, t) - M \quad (4)$$

with  $M$  being the mean value of the wavelet coefficients in the area used for computing. Additionally,  $i, j \in \{-1, 0, 1\}$ , i.e.,  $\mathbf{V}_k$  is a vector of wavelet coefficients belonging to a local spatial window and subtracted by their mean value. The symbol  $\bullet$  used in (3) stands for the vector product and  $k$  represents the spatial positions in the wavelet band where the covariance matrix is computed. We note that the noise estimation by (3) is considered as reliable if the number  $K$  of the observations is relatively large.

In the proposed method we first compute sample covariance matrices in small neighborhoods randomly chosen across the wavelet band. The set of these matrices is denoted as  $\{C_{sn}\}$ . The number  $S$  of computed matrices is large enough to represent significant sample and the size  $G$  of the neighborhood used for each covariance matrix computation is relatively small compared to image size. These parameters are experimentally determined, as described in Section 3.

Because we assume that the covariance for noise is constant for the whole wavelet band and differs for the signal of interest, we aim at estimating the noise covariance matrix  $C_n$  by averaging subset of  $\{C_{sn}\}$  that contain most similar matrices.<sup>3</sup> The matrix similarity measure is determined using Frobenius norm of matrix differences. We calculate Frobenius norm of all possible matrix differences in a given set  $\{C_{sn}\}$ , as follows:

$$D_{ij} = \|C_{sn}^i - C_{sn}^j\| = \sum_{q=1}^L \sum_{r=1}^L |c_{sn}^i(q, r) - c_{sn}^j(q, r)|^2, \quad (5)$$

where  $L$  is dimension of covariance matrix  $C_{sn}$  and  $c_{bn}(q, r)$  are matrix elements at position  $(q, r)$ . The final estimate of noise covariance  $C_n$  is obtained by averaging matrices from  $\{C_{sn}\}$  for which  $D_{ij}$  is smaller than a given threshold  $T$ :

$$C_n \approx E(\{C^i | C^i \in \{C_{bn}\} \wedge C^j \in \{C_{bn}\} \wedge D_{ij} < T, \forall j\}). \quad (6)$$

<sup>3</sup>This is equivalent to determining which members of  $\{C_{sn}\}$  belong to image regions without signal of interest. Therefore, the expected value of the estimated covariance matrix of this subset according to (2) is  $E(C_{sn}) \approx C_n$ .

The parameter  $T$  determines the size of a subset from which the noise covariance estimation is computed and depends on the subband and scale of wavelet decomposition.

We found experimentally that the most appropriate estimation of this parameter can be determined by analyzing histogram of all possible distances  $D_{ij}$ , calculated for sample matrix set  $\{C_{sn}\}$ . Following that, the threshold  $T$  should satisfy following equation:

$$\int_0^T hist(D_{ij}) dh = p \int_0^\infty hist(D_{ij}) dh, \quad (7)$$

where  $hist(D_{ij})$  is histogram of all values  $D_{ij}$ ,  $dh$  stands for summation over histogram values and  $p$  is in range 0.5%–2%.

## 2.2. Spatio-temporal denoising

As previously described in Section 2.1, the covariance matrix of the noise  $C_n^{WB}$  is estimated for each wavelet band  $WB_n$ . This matrix is used for determining the wavelet shrinkage factor  $\gamma_{WB}(x, y, t)$  for each spatial position  $(x, y, t)$  based on the vector product of covariance matrix of noise and the corrupted signal in  $3 \times 3$  window as follows:

$$\gamma_{WB}(x, y, t) = \|C_n^{WB} \bullet \mathbf{V}_{WB}(x, y, t)\|^2 \quad (8)$$

where  $\mathbf{V}_{WB}(x, y, t)$  in (8) is a vector of noisy wavelet coefficients  $WB_n(x+i, y+j, t)$ , with  $i, j \in \{-1, 0, 1\}$ , i.e. it is a vector of wavelet coefficients belonging to a local spatial window.

The denoised wavelet coefficient is then determined as follows:

$$WB_{sf}(x, y, t) = N_1 \gamma_{WB}(x, y, t) WB_n(x, y, t) \quad (9)$$

where  $N_1$  is normalizing parameter experimentally found in the mean squared sense.

The subsequent temporal filtering is performed after the spatial filtering in the wavelet domain, using the proposed scheme for motion compensated filtering [5]. The reason for applying the temporal filter after the spatial one (which is the opposite of proposed in [5]) is mainly done because the temporal filter of [5] is sensitive to correlated noise. Specifically, the noise-robust motion estimation is not as reliable as in the case of white Gaussian noise and consequently results in reduced temporal filtering performance. By applying the spatial filter first, we aim at first decorrelating the noise and also removing it to some extent; after that an efficient temporal filtering scheme of [5] can be applied.

## 3. EXPERIMENTAL RESULTS

In the implementation the proposed spatio-temporal filter a Daubechies [17] wavelet with 8 tap filter bank was used.

**Table 1.** Relative mean squared error, Err.

	l=1	l=2	l=3	l=4
<i>HL</i>	0.0367	0.0174	0.0369	0.469
<i>LH</i>	0.0027	0.0232	0.0712	0.5369
<i>HH</i>	0.0313	0.0253	0.2051	0.6369

We use non-decimated wavelet transform, where 4 decomposition scales was used for the proposed spatial filter and 2 scales for the temporal one.

In our experiments we first evaluated the performance of the proposed noise covariance estimation method. For that we used “Lena” image with artificially added correlated noise, as shown in Fig 2(a). Specifically, the noise added was obtained by cutting particular high frequency components of the white Gaussian noise in the Fourier domain. This produces “spectrally colored” noise which can be seen in Fig 2(a) as vertical and horizontal stripes. Note that this noise although spatially correlated is still spatially stationary.

The accuracy of noise covariance estimation is determined by using relative mean squared error between the noise covariance  $C_r$  of known noise and the noise covariance  $C_n$  estimated from the corrupted image, as follows:

$$Err = \frac{\sum_k (C_r[k] - C_n[k])^2}{\sum_k C_n[k]^2} \quad (10)$$

where  $k$  in (10) stands for an index, i.e., the position of an element in the corresponding covariance matrix.

The experimental results showed better accuracy for higher resolution scales (smaller  $l$ ) than for the low resolution ones. In table 1 we show the results for the considered noise and “Lena” image. There, it can be observed that after third scale the errors become significantly large; however for the smaller scale there are acceptably small. Hence at smaller resolution scales the proposed method for noise covariance estimation can be considered as relatively good.

Next, by using the proposed noise estimation method and the proposed spatial denoising scheme (as described in Section 2.2 we show results for the processed “Lena” image with added correlated noise (Fig 2(a)) in Fig 2(b). In Fig 2 it can be noticed that the correlated noise is efficiently removed with small spatial blurring introduced. However, sometimes extreme noisy “stripes” are not completely removed; this comes from the fact that we assume spatial stationarity which is not always the case in fact. Nevertheless this happens very rarely and does not introduce serious artifacts.

Finally, we show the results for a video sequence captured from TV which was corrupted by correlated noise due to transmission interferences (in analogue domain) via TV

tuner. In Fig. 3 we show one frame of noisy sequence, denoised by the proposed spatial and spatio-temporal filter. In the same figure we show the results of the spatial filtering approach of [14], where we used non-decimated wavelet transform<sup>4</sup> with 4 scales and local window of  $3 \times 3$  and Gaussian scale mixture with 13 components (the noise covariance estimated from a clean (signal-free) area (right part of the image)). Furthermore, in Fig. 3 we also show results for the spatio-temporal method of [5], which only assumes white Gaussian noise into account.

The results show the the proposed spatial filter (Fig. 3(b)) efficiently removes correlated noise, while it fails to completely remove some impulse-like noise (which is spatially local). However, this is not a problem because the proposed temporal filter (Fig. 3(c)) applied subsequently removes efficiently the rest of the noise; some little amount of noise is still left, but less annoying. The spatial-only filter of [14] (Fig. 3(d)) showed comparable and reasonably good results as well, but with some artifacts also introduced (low frequency noise). Finally, it can be seen from Fig. 3(e) that the biggest amount of noise is left when noise is assumed to be white Gaussian.

#### 4. CONCLUSION AND FUTURE WORK

In this paper we proposed a new scheme for wavelet-based video denoising, where noise is assumed to be spatially correlated and stationary. Specifically, a novel technique for estimating noise covariance matrix in video is developed and applied to a wavelet-based shrinkage-wise denoising scheme for spatial filtering. Subsequently the temporal filtering is performed on the spatially denoised video frames. The results show superior performance of the method when only white Gaussian noise is assumed; nevertheless further improvements are still necessary for the optimal noise removal of correlated noise in video sequences.

#### 5. REFERENCES

- [1] A. Amer and E. Dubois, “Fast and reliable structure-oriented video noise estimation,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 1, pp. 113–118, Jan. 2005.
- [2] B.C. Song and K.W. Chun, “Noise power estimation for effective de-noising in a video encoder,” in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, USA, Apr. 2005, pp. 357–360.
- [3] V. Zlokolica, A. Pižurica, and W. Philips, “Noise estimation for video processing based on spatio-temporal gradients,” *IEEE Signal Processing Letters*, vol. 13, no. 6, pp. 337–340, 2006.
- [4] E. J. Balster, Y. F. Zheng, and R. L. Ewing, “Combined spatial and temporal domain wavelet shrinkage algorithm for video denoising,” *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 220–230, Feb. 2006.

<sup>4</sup>Daubechies [17] wavelet with 8 tap filter bank was used for this purpose.

(a) Noisy



(b) Denoised



**Fig. 2.** “Lena” image (a) with added correlated noise and (b) denoised image by the proposed spatial filter.

- [5] V. Zlokolica, Pižurica, and W. Philips, “Wavelet-domain video denoising based on reliability measures,” *IEEE Trans. on Circuits and Systems for Video Technology*, Sep 2006.
- [6] Fu. Jin, P. Fieguth, and L. Winger, “Wavelet video denoising with regularized multiresolution motion estimation,” *EURASIP Journal of Applied Signal Processing*, vol. ID 72705, 2006.
- [7] A. Pižurica, V. Zlokolica, and W. Philips, “Noise reduction in video sequences using wavelet-domain and temporal filtering,” in *SPIE Conference on Wavelet Applications in Industrial Processing*, Providence, RI, USA, Oct. 2003, vol. 5266.
- [8] S.G. Chang, B. Yu, and M. Vetterli, “Spatially adaptive wavelet thresholding with context modeling for image denoising,” *IEEE Trans. on Image processing*, vol. 9, no. 9, pp. 1522–1531, Sept. 2000.
- [9] L. Sendur and I.W. Selesnick, “Bivariate shrinkage functions for wavelet-based denoising exploiting interscale dependences,” *IEEE Trans. on Image processing*, vol. 50, no. 11, pp. 2744–2756, Nov. 1999.
- [10] H.A. Chipman, E. D. Kolaczyk, and R.E. McCulloch, “Adaptive bayesian wavelet shrinkage,” vol. 92, pp. 1413–1421, 1997.
- [11] A. Pižurica, W. Philips, I. Lemahieu, and M. Archeroy, “Despeckling sar images using wavelets and a new class of adaptive shrinkage estimators,” in *IEEE Conference on Image Processing*, Sept. 2001, vol. 2, pp. 233–236.
- [12] R. F. Wagner, S.W. Smith, J.M. Sandrik, and H. Lopez, “Statistics of speckle in ultrasound b-scans,” *IEEE Trans. Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156–163, May 1983.
- [13] D. Nadadur, R.M. Haralick, and D.E. Gustafson, “A bayesian framework for noise covariance estimation using the facet model,” *IEEE Trans. on Image Processing*, vol. 14, no. 11, pp. 1902–1917, Nov. 2005.
- [14] J. Portilla, V. Strella, M Wainwright, and E.P Simoncelli, “Image denoising using scale mixtures of Gaussians in the wavelet domain,” *IEEE Trans. on Image processing*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.
- [15] S. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. 14, pp. 674–693, July 1989.
- [16] S. Mallat, *A wavelet tour of signal processing (2nd ed.)*, Academic Press, Oval Road, London, third edition, 1999.
- [17] I. Daubechies, “The wavelet transform, time-frequency localization and signal analysis,” *IEEE Trans. on Information Theory*, vol. 36, no. 5, pp. 961–1005, 1990.

(a) Noisy TV sequence



(c) The proposed spatial filter



(b) The method of [14]



(d) The proposed spatio-temporal filter



(e) The spatio-temporal method of [5]



**Fig. 3.** “TV sequences from Kanaal3”: (a) noisy input, (b) denoised by the spatial filter of [14], (c) denoised by the proposed spatial filter, (d) denoised by the proposed spatio-temporal filter, denoised by spatio-temporal method of [5].