International Journal of Computer and Communication Technology

Volume 8 | Issue 2 Article 3

April 2017

# VIDEO DENOISING BASED ON EXPLOITING INTRASCALE AND INTERSCALE DEPENDENCY IN WAVELET DOMAIN

Karabi Devi
*Deptt. of Electronics and Communication Engineering, School of Engineering Tezpur University (A Central University), Napaam, Tezpur, Assam, India*, karabidevi20@yahoo.com

DEEPIKA HAZARIKA
*Deptt. of Electronics and Communication Engineering, School of Engineering Tezpur University (A Central University), Napaam, Tezpur, Assam, India*, deepika@tezu.ernet.in

V. K. NATH
*Deptt. of Electronics and Communication Engineering, School of Engineering Tezpur University (A Central University), Napaam, Tezpur, Assam, India*, vknath@tezu.ernet.in

# VIDEO DENOISING BASED ON EXPLOITING INTRASCALE AND INTERSCALE DEPENDENCY IN WAVELET DOMAIN

**K. DEVI[1], D. HAZARIKA[2], V. K. NATH[3]**

[1,2,3]Deptt. of Electronics and Communication Engineering, School of Engineering
Tezpur University (A Central University), Napaam, Tezpur, Assam, India
Email: karabidevi20@yahoo.com, deepika@tezu.ernet.in & vknath@tezu.ernet.in

**Abstract-** In this paper, we propose a new video denoising algorithm which uses an efficient wavelet based spatio-temporal filter. The filter first applies 2D discrete wavelet transform (DWT) in horizontal and vertical directions on an input noisy video frame and then applies 1-D discrete cosine transform (DCT) in the temporal direction in order to reduce the redundancies which exist among the wavelet coefficients in the temporal direction. We observe that the subband coefficients with large magnitudes occur in clusters in locations corresponding to the edge locations even after applying the above spatio-temporal filter. In this paper, we propose to use two different low complexity wavelet shrinkage based methods to denoise the noisy wavelet coefficients in different subbands. The first method exploits the intra-scale dependencies between the coefficients and thresholds the wavelet coefficients based on the measure of sum of squares of all wavelet coefficients within a square neighborhood window. The second method exploits the inter-scale dependencies between the coefficients at different scales in an individual slice of coefficients. After filtering the individual slices of coefficients, the denoised video frames in time domain are obtained after inverse transforms. We propose to exploit the temporal redundancies between the successive frames again in the time domain using low complexity selective recursive temporal filtering (SRTF). In the proposed video denoising scheme, since the temporal redundancy is exploited both in the time and wavelet domain, the denoising capability of the scheme is hence increased. The video denoising performance using the two proposed approaches outperform many existing well known video denoising techniques including one recent well known method which uses the similar transformation, both in terms of PSNR and visual quality. We also show that, simple soft thresholding using Donoho's threshold when used with this wavelet based spatio-temporal filter even outperforms many well known non linear based video denoising techniques.

***Keywords****- Video denoising, discrete wavelet transform, wavelet shrinkage, discrete cosine transform, Temporal filtering.*

## I. INTRODUCTION

Digital video, which is a sequence of images or frames are often corrupted by noise during acquisition or transmission process that affects further processing of the video. The noise deteriorates the quality of an image and makes the tasks like compression and segmentation very difficult. Hence it is very important to reduce the noise in video sequences without affecting the important features or details. In the video applications, the noise is well approximated by the white additive Gaussian model, which is considered here.

In the recent years, a number of video denoising methods have been proposed [2]-[26]. A review of several noise reduction filters for video sequences is given in [2]. In the literature, several non-linear techniques for video denoising have been proposed. Based on the order statistics of the pixel values in the local spatio-temporal window, the alpha trimmed mean filter [21] based on averaging of several neighbors, estimates the filtered value of the given pixel.

Many video denoising techniques employ spatio-temporal filters. The spatio-temporal filters for video denoising can be classified as fully 3D or non-separable and 2D+1D or separable. In literature, the separable filters are available in 3 different forms - spatial filtering is done before temporal filtering [14][15], temporal filtering is done before spatial filtering[16] [17] and both spatial and temporal filters are applied parallel [2][9]. In [6], the authors extended the spatio-temporal filter in the wavelet domain. In [26], Nath et al. proposed a spatio-temporal filter where 2D Lapped transform was applied in the spatial domain and 1D DCT was applied in the temporal direction. Local Wiener filtering was used to filter the noisy transform coefficients. Recently, Plotkin et al. proposed a similar transformation where 2D DWT was applied in the spatial and 1D DCT was applied in the temporal direction. The authors used hierarchically adapted threshold to filter the transform coefficients.
Some of the video denoising techniques uses motion compensation ([9], [10], [13], [16], [17] to better exploit the considerable temporal redundancy in video by temporally smoothing pixel values along the motion trajectories. But, these motion compensated filters are quite complex due to its computational complexity and takes longer processing time.

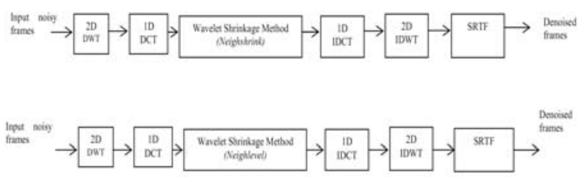


Fig.1 Block diagram of the proposed method using (a) *Neighshrink* technique (b) *Neighlevel* technique

In this paper, we have proposed two efficient video denoising methods which uses the similar transformation as proposed in [19]. The method tries to exploit the temporal correlation both in the time and wavelet domain which increases the efficiency of the proposed algorithm. In order to obtain the estimate of the noise free coefficients in an individual slice, we propose to use two different wavelet shrinkage based methods which efficiently exploits the ‘clustering’ feature of the transformed coefficients which exist strongly in the subband locally and also across the scales. The proposed method outperforms the recent well known method [19] which uses the similar transformation, in terms of PSNR.

The paper is organized as follows. The proposed video denoising scheme is presented in section II. Experimental results are given in section III and in section IV, the paper is concluded.

## II. PROPOSED VIDEO DENOISING SCHEME

Fig.1 shows the proposed video denoising framework where 2D DWT is first applied spatially on the input noisy video frames followed by a 1D DCT which is applied temporally on the wavelet coefficients. The application of 1D DCT in wavelet domain decorrelates the temporal redundancies which exist between the corresponding wavelet coefficients in the temporal direction. We observe that the transform coefficients obtained after application of 2D DWT and 1D DCT still shows highly non Gaussian statistics with sharp peak and heavy tails and also exhibits strong ‘clustering’ feature which means that the coefficients with comparatively large magnitudes in a subband tends to occur in clusters in positions corresponding to edge positions in the spatial domain.
We propose to use two different wavelet shrinkage methods which use neighboring coefficients in a local window to obtain the estimate of the noise free wavelet coefficients in an individual slice.
In wavelet domain, the noisy slice can be expressed as

$$y_k = x_k + n_k \qquad (1)$$

where $x_k$ represents wavelet coefficients of the original slice.

The method popularly known as Neighshrink [18] is used to denoise the individual slices of transform coefficients in the first proposed method. The method shrinks the transform coefficient based on the information of sum of squares of all the wavelet coefficients within a neighborhood local window. The method assumes high correlation between the neighboring coefficients in a local window. The method first considers the Donoho’s universal threshold [27] assuming an $NxN$ size signal

$$\lambda = \sqrt{2\sigma^2 \log N^2} \qquad (2)$$

In soft thresholding [27], the thresholded coefficient is obtained as

$$X_{soft} = sign(X)(|X| - \lambda) \quad \text{if} \quad |\ \ | > \lambda \qquad (3)$$
$$= \quad \text{else}$$

It is seen that the universal threshold is designed for smoothness rather than minimizing the errors. But the natural images are not sufficiently smooth. Moreover, the universal threshold reduces more the noise coefficients as $N$ tends to infinity. But, natural images are not composed of infinite number of elements. Hence, an optimal threshold $\lambda^* = \alpha\lambda$ is used in our proposed methods which minimize mean square error. The performance of the above mentioned thresholding technique based on Donoho’s threshold value is compared with that of our proposed methods in experimental results. The optimal threshold value is substituted in an expression which finally gives [18]

$$\hat{x}_k = y_k \left( 1 - M^2 \lambda^{*2} \Big/ \sum_{z_l \in N(i,j)} z_l^2 \right)_+ \qquad (4)$$

where $N(i, j)$ is an $MxM$ local window centered over the coefficient to be filtered. From the experiments we have investigated the optimal threshold by varying $\alpha$ and found it to be 0.25 which is very much less than 1.0 for the simulated videos, as shown in Fig.3.
We show in the experimental results section that simple soft thresholding with Donoho’s threshold even shows encouraging results with this efficient wavelet based spatio-temporal filter.

TABLE I PSNR comparison of different algorithms for ‘Miss America’ and ‘Hall Monitor’ sequence

| Video Seq. | ATM [21] | KNN [21] | RF [3] | JKW [9] | LMMSE [13] | AWA [13] | MCA [6] | Ref. [19] | Pr-Soft | Proposed1 (Neighshrink) | Proposed2 (Neighlevel) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 20dB | | | | | | | | | | | |
| MA | 30.37 | 29 | 29.29 | 30.65 | 30.96 | 29.01 | 31.21 | 35.53 | 35.74 | 36.42 | 36.63 |
| Ha | 27.66 | 27.29 | 27.31 | 28.68 | 27.35 | 27.75 | 28.09 | 30.38 | 31.53 | 33.08 | 33.23 |
| 25dB | | | | | | | | | | | |
| MA | 33.61 | 33.45 | 34.18 | 32.22 | 34.57 | 33.36 | 34.92 | 37.52 | 38.25 | 39.32 | 39.46 |
| Ha | 28.74 | 29.98 | 30.34 | 32.37 | 30.86 | 31.07 | 32.04 | 33.65 | 33.66 | 35.62 | 35.73 |

TABLE II PSNR comparison of different algorithms for ‘Salesman’, ‘Table Tennis’ and ‘Flower Garden’ sequence

| Video sequences | Noise ($\sigma_n$) | Noisy PSNR | RF [3] | Soft 3D [25] | SEQWT [14] | WRSTF [24] | 3D Mixed [26] | Pr-Soft | Proposed1 (Neighshrink) | Proposed2 (Neighlevel) |
|---|---|---|---|---|---|---|---|---|---|---|
| Salesman | 10 | 28.13 | 30.55 | 34.85 | 34.17 | 35.82 | 35.85 | 33.75 | 37.02 | 37.12 |
| | 15 | 24.60 | 29.50 | 33.29 | 31.84 | 33.91 | 33.62 | 32.07 | 35.07 | 35.19 |
| | 20 | 22.11 | 28.27 | 32.10 | 30.10 | 32.40 | 31.99 | 30.84 | 33.70 | 33.83 |
| Table Tennis | 10 | 28.13 | 28.39 | 31.86 | 31.17 | 32.41 | 31.35 | 29.03 | 31.64 | 31.71 |
| | 15 | 24.61 | 27.42 | 29.86 | 29.07 | 30.12 | 29.61 | 25.41 | 29.70 | 29.79 |
| | 20 | 22.12 | 26.35 | 28.58 | 27.83 | 28.68 | 28.34 | 23.03 | 28.41 | 28.52 |
| Flower Garden | 10 | 28.14 | 23.32 | 30.23 | 29.65 | 30.80 | 29.51 | 29.30 | 30.04 | 30.12 |
| | 15 | 24.61 | 22.96 | 27.71 | 27.07 | 29.19 | 27.15 | 25.60 | 27.42 | 27.51 |
| | 20 | 22.10 | 22.46 | 26.01 | 25.34 | 26.39 | 25.47 | 23.22 | 25.71 | 25.80 |

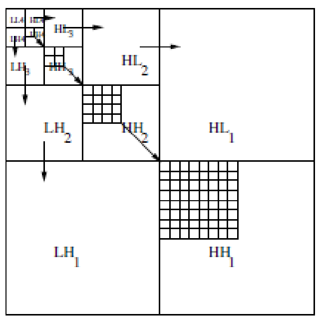


**Fig.2 Parent-child relationships in a 4-level orthogonal discrete wavelet transform**

The second proposed method exploits the inter subband correlation in an individual slice of coefficients. We assume that there exists high correlation between the wavelet coefficients lying in different scales in an individual slice. In image compression and denoising applications, this statistical correlation between wavelet coefficients at different scales is widely used and is known as parent child interdependency as shown in Fig. 2. We propose to use the 'Neighlevel' [18] method which is based on exploitation of intersubband correlation. The filtering equation based on 'Neighlevel' method [18] is given by

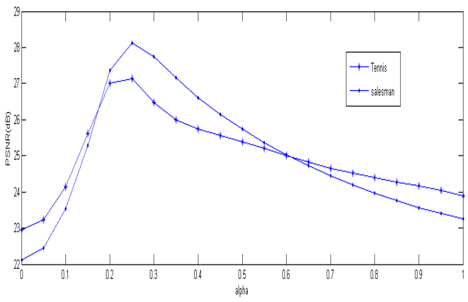


**Fig.3 PSNR performance v/s $\alpha$ in $\lambda^* = \alpha\lambda$ for two sequences.**

$$\hat{x}_k = y_k \left( 1 - \frac{(M^2+1)\lambda^{*2}}{\left( \sum_{z_l \in N(i,j)} z_l^2 + p^2 \right)} \right)_+ \qquad (5)$$

where $z_l$ denotes the coefficient to be filtered and $N(i, j)$ is an $MxM$ local window centered over the coefficient to be filtered. $p$ is the parent of the coefficient to be filtered and $\lambda^* = \alpha\lambda$ .

After denoising all the individual slices of wavelet coefficients, 1D IDCT followed by 2D IDWT is applied to obtain the denoised time domain frames. In order to exploit the temporal redundancy in time domain, we apply the simple selective recursive temporal filtering (SRTF) in the time domain [14]. The SRTF further improves the denoising performance of the proposed framework by temporal averaging using motion information. It is to be noted that SRTF uses very low complexity pixel based motion detector.

## III. EXPERIMENTAL RESULTS

For evaluation of the performance of the proposed scheme, we have compared its output peak signal to noise ratio (PSNR) values and visual quality with some well known video noise reduction algorithms. Five different standard video sequences 'Miss America', 'Hall' ,'Salesman', 'Tennis' and 'Flower Garden' contaminated with pure additive white Gaussian noise are processed with both of the proposed methods . We process the noisy input video frames in a group of 10 to keep the computational complexity low. The implementation of soft thresholding using Donoho's threshold in the wavelet (2D DWT+1D DCT) domain is referred to as Pr-Soft method.

In Table I, the resultant PSNR values are compared with that of 3D KNN filter [21], 3D alpha-trimmed mean (ATM) filter [21], 3D RF filter [3], joint Kalman-Wiener filter [9], LMMSE filter [13] , adapted weighted averaging (AWA) filter [13] ,Pr-Soft and Ref. [19] where similar transformation (2D DWT+1D DCT) is used. The PSNR was averaged over a total 60 frames. In Table II, the average PSNR values of the test sequences 'Salesman', 'Tennis' and 'Flower Garden' after being processed with our proposed methods are compared with some existing methods [3][14][24][25][26]. For average PSNR computation, we have taken 40 frames for 'Salesman', and 52 frames for both 'Tennis' and 'Flower Garden'. We used an orthogonal wavelet transform (Daubechies length-8 wavelet) with four levels of decomposition with a center square shaped window of 5X5 size and alpha value of 0.25 in both of the proposed schemes. A threshold value of 24 was chosen based on experimental study, for selective recursive temporal filtering. The experimental results in Table I shows that the proposed video denoising framework outperforms many well known video denoising schemes including Ref. [19] which uses the similar transformation, both in terms of PSNR and visual quality. This shows that the proposed filtering using Neighshrink and Neighlevel better exploits the intra-subband and inter-subband correlation respectively as compared to hierarchically adapted thresholding [19] in the wavelet (2D DWT+1D DCT) domain.

In Table II, except for WRSTF [24] and soft3D [25] methods for Tennis and Flower garden sequence, both the proposed methods outperform all other methods. It is to be noted that unlike WRSTF, the proposed schemes uses very low complexity pixel based motion detector for temporal filtering. From both Table I and Table II, it can be seen that the proposed methods based on the modified optimal threshold value gives better results than the method based on the universal soft threshold with Donoho's threshold, for all the sequences which is quite obvious. From Table I we can observe that the Pr-soft clearly outperforms the non –linear methods as well as two wavelet based methods in terms of PSNR.

In Fig. 4, PSNR per frame comparison for different techniques from frames 10 to 50 has been shown. In Fig. 5 and Fig. 6, the visual results for the 5th frame of 'Flower Garden' sequence and 9th frame of 'Table Tennis' sequence are shown respectively.

It can be observed that the proposed method provides denoised video frames that are of good visual quality and a smoother look with less distortion.

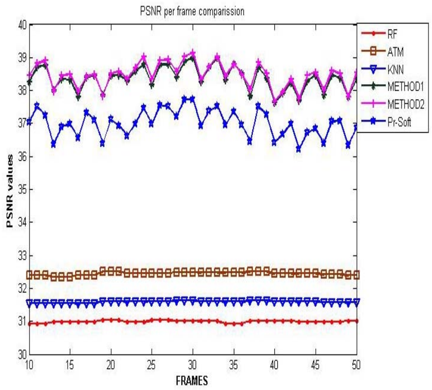


**Fig.4 PSNR per frame comparison of different algorithms for the sequence 'Miss America' with input PSNR=20dB.**

## IV. CONCLUSION

In this paper, two new video denoising techniques are presented which are based on an efficient wavelet based spatio-temporal filter. The proposed framework not only exploits the temporal correlation in wavelet domain, but also exploits it in the time domain which is one of the main features of the proposed scheme. The wavelet coefficients obtained after the application of 2D DWT in spatial domain followed by 1D DCT in temporal domain, are denoised using two wavelet shrinkage methods 'Neighshrink' and 'Neighlevel' which utilizes the intra-subband and inter-subband correlation between the coefficients. The temporal correlation in time domain is exploited using selective recursive temporal filtering. The experimental results demonstrates that the proposed filtering using 'Neighshrink' and 'Neighlevel' in the subbands of individual slice performs better than the hierarchically adapted threshold based method and several other well known methods in terms of PSNR. It is shown that even simple soft thresholding using Donoho's threshold when used with this spatio-temporal filter performs better than many non linear filter based schemes. The proposed framework reduces the noise very efficiently while preserving the true edges and other important details.

## REFERENCES

[1] V. Zlokolica, A. Pizurica, E. Vansteenkiste, W. Philips "Spatio-temporal approach for noise estimation", in proceedings of IEEE international conference on Acoustic Speech and Signal Processing, May 2006.

[2] J. C. Brailean, R. P. Kleihorst, S. Efstratidis, K. A. Katsaggeleos, and R. L. Lagendijk, "Noise reduction filters for dynamic image sequences:a review," Proc. IEEE, vol. 83, no. 9, pp. 1272–1292, Sep. 1995.

[3] F. Cocchia, S. Carrato, and G. Ramponi, "Design and real-time implementation of a 3-D rational filter for edge preserving smoothing," IEEE Trans. Consum. Electron., vol. 43, no. 4, pp. 1291–1300, Nov. 1997.

[4] W. I. Selesnick and K. Y. Li, "Video denoising using 2d and 3d dual-tree complex wavelet transforms," in Proc. SPIE Wavelet Applicat. Signal Image Process., San Diego, CA, Aug. 2003, pp. 607–618.

[5] N. Rajpoot, Z. Yao, and R. Wilson, "Adaptive wavelet restoration of video sequences," in Proc. Int. Conf. Image Process., Singapore, 2004, p. 957–960.

[6] V. Zlokolica, A. Pizurica, and W. Philips, "Video denoising using multiple class averaging with multiresolution," in Proc. Int. Workshop VLBV03, Madrid, Spain, Sep. 2003, pp. 172–179.

[7] R. Rajagopalan and M. T. Orchard, "Synthesizing processed video by filtering temporal relationships," IEEE Trans. Image Process., vol. 11, no. 1, pp. 26–36, Jan. 2002.

[8] O. A. Ojo and T. G. Kwaaitaal-Spassova, "An algorithm for integradted noise reduction and sharpness enhacement," IEEE Trans. Consum. Electron., vol. 46, no. 3, pp. 474–480, Aug. 2000.

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

[10] R. P. Kleihorst, R. L. Legendrijk, and J. Biemond, "Noise reduction of image sequences using motion compensation and signal decomposition," IEEE Trans. Image Process., vol. 4, no. 3, pp. 274–284, Mar. 1995.

[11] P. M. B. Van Roosmalen, S. J. P. Westen, R. L. Lagendrijk, and J. Biemond, "Noise reduction for image sequences using an oriented pyramid thresholding technique," in Int. Conf. Image Process., Lausanne, Switzerland, 1996, pp. 375–378.

[12] A. Pizurica, V. Zlokolica, and W. Philips, "Noise reduction in video sequences using wavelet-domain and temporal filtering," in Proc. SPIE Conf.Wavelet Applicat. Industrial Process., Providence, RI, Oct. 2003, pp. 48–59. 5.

[13] K. O. Mehmet, S. Ibrahim, and T. Murat, "Adaptive motion-compensated filtering of noisy image sequences," IEEE Trans. Circuits Syst. Video Technol., vol. 3, no. 4, pp. 277–290, Aug. 1993.

[14] A. Pizurica, V. Zlokolika, and W. Philips, "Combined wavelet domain and temporal video denoising," in Proceedings of IEEE International Conference advanced Video and Signal Based Surveillance (AVSS), Jul. 2003.

[15] Cheong, H., Tourapis, A., Llach, J., Boyce, J, "Adaptive spatiotemporal filtering for video denoising", Proceedings of IEEE International Conference on Image Processing, pp. 965968 Singapore, 2004.

[16] Zlokolica, V., Pizurica, A., Wilfried, P, " Recursive temporal denoising and motion estimation of video", Proceedings of IEEE International Conference on Image Processing, 2004, pp.1465-1468, Singapore.

[17] Kleihorst, R., Legendrijk, R., Biemond, J, " Noise reduction of image sequences using motion compensation and signal decomposition", IEEE Transactions on Image Processing, 4(3), 1995, pp.274284.

[18] Dongwook Cho, Tien D. Bui and Guangyi Chen ,"Image denoising based on wavelet shrinkage using neighbor and level dependency". International jouranal of Wavelet Multiresolution and Information Processing,Vol7,No.3,2009

[19] Nikhil Gupta, M. N. S. Swamy and Eugene I. Plotkin, "Video Noise Reduction in the Wavelet Domain using Temporal Decorrelation and Adaptive Thresholding",in proceedings of IEEE international Symposium Circuits and Systems,2006.

[20] J. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans. Signal Process., 41, 3445–3462, 1993.

[21] V. Zlokolica, W. Phillips, and D. Van de Ville, “ Robust non-linear filtering for video processing,” in Proc. Intl. Conf. on Dig. Sig. Process. (DSP2002), Santorini, Greece, July 2002, vol. 2, pp. 571–574.

[22] V. K. Nath, D. Hazarika, “Video Denoising using combined Bilateral and Temporal filtering”, National Conference on trends in machine intelligence, Tezpur, Assam, India,July, 2011.

[23] V. K. Nath, D. Hazarika, “Noise Reduction in Video using Wienerbilateral filter”, Journal of Computer Engineering,Vol. 3, No.1, 2012.

[23] V. K. Nath, D. Hazarika, “Noise Reduction in Video using Wienerbilateral filter”, Journal of Computer Engineering,Vol. 3, No.1, 2012.

[24] V. Zlokolica, A. Pizurica, and W. Philips, “Wavelet Domain Noise-Robust Motion Estimation and Noise Estimation for Video Denoising,” First International Workshop on Video Processing and

Quality Metrics for Consumer Electronics, Scotssdale, Arizona, 23-25January, 2005, USA.

[25] I. Selesnick and K. Y. Li, “Video denoising using 2D and 3D dualtree

complex wavelet transforms,” presented at the Wavelet Applications in Signal and Image Processing X (SPIE), San Diego, CA,Aug. 2003.

[26] V. K. Nath, D. Hazarika, “Video noise reduction in 3-D mixed transform domain using its efficient wavelet structure”, IEEE INDICON 2009, Gujarat.

[27] D. L. Donoho, “ Denoising by soft thresholding “, IEEE Trans. on Inf. Theory, vol. 41, 3, pp. 613-627.

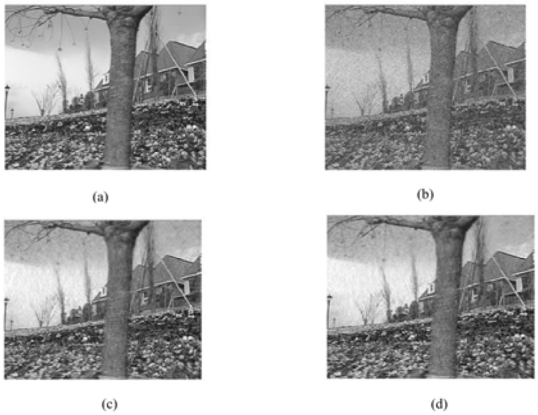


**Fig. 5 Visual Results for ‘FlowerGarden’ sequence (Frame no. 5, noisy PSNR=20 dB) (a) Original (b) Noisy (c) Proposed method1(using ‘Neighshrink’) (d) Proposed method 2(using ‘Neighlevel)**

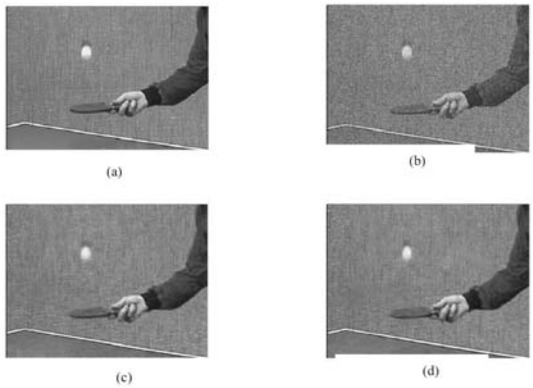


**Fig. 6 Visual Results for ‘Table tennis’ sequence (Frame no. 9, noisy PSNR=20 dB) (a) Original (b) Noisy (c) Proposed method 1using ‘Neighshrink’) (d) Proposed method 2(using ‘Neighlevel)**