Adaptive Spatio-Temporal Filtering with Motion Estimation for Mixed Noise Removal and Contrast Enhancement in Video Sequence

S. Madhura and K. Suresh

Abstract Naturally available noises in the videos are complex but fortunately they can be broadly classified as Gaussian and Impulse noises. Most of the available models for noise removal emphasize on any one kind of noise removal thus an optimum model of mixed noise removal is still a challenge. This paper describes about removal of video flickering and artifacts due to sensor motion, unprofessional recording behaviors, device defects, poor lighting conditions and high dynamic exposure. The adaptive spatio-temporal filter gives excellent result for mixed (Gaussian and Impulse) noise removal. Dense optical flow is introduced to reduce the motion blur and enhance the video. The analysis of PSNR and SSIM values were compared with existed method like Non-local Means and BM3D approach and results are tabulated. The Histogram graph gives the better intensity distribution in frames thus the proposed method even works good for low illumination or night vision surveillance videos.

Keywords Video enhancement · Noise reduction · Optical flow · Adaptive filters

1 Introduction

Video recording has become a trend in the existing rapidly developed digital advancement aided by availability of high focal length and high definition cameras. In this scenario videos are not only limited to professional application but have stepped into surveillance and customized video making since the capturing videos has become easy. Hence the artifacts such as blur, blocking, contrast distortion and

S. Madhura (✉)
Dayanandasagar College of Engineering, Bangalore, India
e-mail: madhu4tulip@gmail.com

K. Suresh
SDMIT, Ujire, Karnataka, India
e-mail: ksece1@gmail.com

S.C. Satapathy et al. (eds.), *Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications*, Advances in Intelligent Systems and Computing 516, DOI 10.1007/978-981-10-3156-4_52

ill-illumination would occur due to movement of the recording device, unprofessional recording behaviors, device defects, poor lighting conditions, high dynamic exposure, transmission loss and so forth. Even though most of the noises are random and complex they can be broadly classified as Impulse noise and Gaussian noise [1, 2]. During video acquisition additive Gaussian noise would get introduced and generally assumes zero mean distribution whereas impulse noise would assume uniform distribution and is produced due to transmission errors. It is still a challenge to develop a common model that deal equally well with different kinds of distortions. Most of the video filtering techniques defined in recent works mainly focuses on removal of impulse noise only [3].

For instance, order statistic filters work very well in removal of impulse noise which distinguishes fine features and image pixel but this method fails for shot noise or Gaussian noise [4–6]. Gaussian noise removal is better with bilateral filter [7], wavelet transforms [8] and anisotropic diffusion [9]. However some of the most effective Gaussian noise removal algorithm like Non-Local Means [10–12] and BM3D [13, 14] cannot remove impulse noise contamination. BM3D algorithm is DCT based and is generally used for Additive Gaussian White Noise (AGWN) removal. This algorithm is generally used for noise removal in still images but now it is extended to videos.

Along with the noisy frames if it is associated with blur due to camera movement or scene motion then there should be another algorithm which can normalize the frames and produce a desirable quality image. The motion blur of the image with the low temporal resolution can be directly measured by the images with high temporal resolution [8, 15]. Motivated by the above problems a novel algorithm has been developed which enhances the video which was subjected various kinds of noises and also which contain motion blur. For noise removal Adaptive Spatio-Temporal Filters would be used and for removal of blur motion estimation from optical flow would be considered. Moreover to make the algorithm faster and robust modified neighborhood relation would be introduced in the temporal domain.

2 Proposed Method

The proposed algorithm is schematically represented in Fig. 1. The entire process can be modeled into two parts: video denoising and deblurring. Firstly frames are extracted from the input video which is distorted by mixed [Gaussian and Impulse] noise and associated with motion blur. The spatially filtered frame would undergo forward and backward temporal filtering. Temporal filtering would be done more exclusively then spatial filtering [16].

Secondly for every three frames the first frame is tracked using the optical flow algorithm and hence their location in third frame would be known. With the assumption that the location of second frame is midway between first and third, the motion vectors of each pixel of first frame are divided by two which provides the

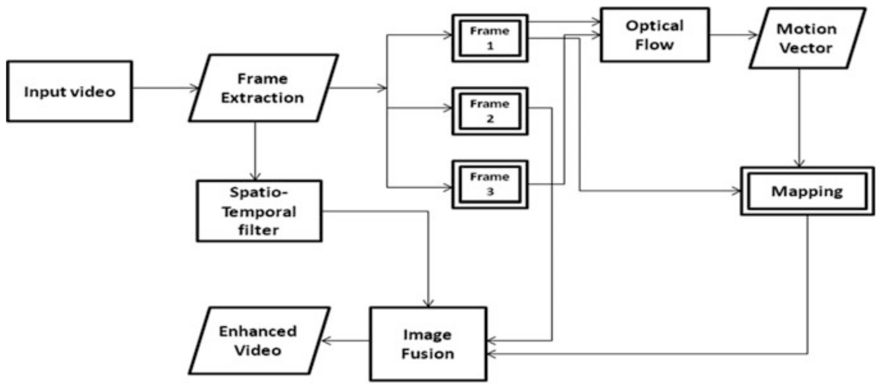

Fig. 1 Schematic of proposed algorithm

magnitude of the motion vector from first frame to second frame. However the motion vector directions remain the same as shown in Eq. (1).

$$I_1(X_1, Y_1) = I_2(X_2, Y_2) \tag{1}$$

where $I_1(X_1, Y_1)$ and $I_2(X_2, Y_2)$ are the intensities of first and second frame respectively. Finally image fusion would be simple logarithmic averaging operation between the frame reconstructed from optical flow and second original frame.

3 Adaptive Spatio-Temporal Filtering

The Adaptive Spatio-Temporal filter depends on the following points:

1. Number of pixels that it wants to combine and whether these pixels are in the area of the frame with motion.
2. The parameters for local illumination are chosen based on transition between spatial-only and temporal-only bilateral filtering.

To remove the spatial blur, blocking and temporal fluctuations within a frame sequence each frame is treated with spatial and temporal filter by proposed statistical spatio-temporal filter. Each frame estimation within the same frame is obtained by their neighboring pixels as shown in Eq. (2) which is a recursive updating function.

$$\theta_c^{(\omega+1)} = \theta_c^{(\omega)} - g\left(\theta_c^{(\omega)}\right) \tag{2}$$

where $\theta_c^{(\omega)}$ is the ω th estimation iterate of θ_c and $g(.)$ is a gradient function given by:

$$g\left(\theta_c^{(\omega)}\right) = \alpha^{\omega} \sum\nolimits_{m \in M_c} \psi\left(\theta_c^{(\omega)} - \theta_m\right) \tag{3}$$

where m is the pixel index of current element, M_c is a set of neighboring pixels, $\psi(.)$ is the influence function representing the sensitivity to the difference $(\theta_c - \theta_m)$ and α is a step size parameter.

The blocking and ringing effect are reduced by processing the above iterative method. The temporal fluctuations are suppressed by making changes in Eq. (3) as follows:

$$g\left(\theta_c^{(\omega)}\right) = \alpha^{\omega}\left(\sum\nolimits_{p \in p_c} \psi\left(\theta_c^{\omega} - \theta_p\right) + \sum\nolimits_{p \in Q_c} \psi\left(\theta_c^{(\omega)} - \theta_q\right)\right) \tag{4}$$

where p and q are pixel indexes, P_c is the immediate previous frame and Q_c is the next frame sets of neighboring pixels $\left(\theta_c^{(\omega)} - \theta_p\right)$ are the forward temporal estimation and $\left(\theta_c^{(\omega)} - \theta_q\right)$ are the backward temporal estimation.

4 Motion Estimation by Optical Flow

Optical Flow has gained lot of momentum towards much computer vision study and significant amount of research is being carried out on the same. In proposed method Lucas-Kanade approach is used.

- The pixels within the small window possess same motion.
- The offset vector from first order approximation of Taylor expansion since the motion is very small.

Consider a window slice 8 × 8, each one with p = 8^2 pixels. An overconstrained system with p equations and two variables are formed using local constraint movement.

$$\begin{array}{c} I_{X1}u + I_{y1}v + I_{t1} = 0 \\ I_{X2}u + I_{y2}v + I_{t2} = 0 \\ | \\ I_{Xp}u + I_{yp}v + I_{tp} = 0 \end{array} \tag{5}$$

Equation (5) can be solved by using Least Mean Square (LMS) method for estimating optical flow vector. The estimated optical flow for each 8 × 8 window corresponds to the optical flow vector of all the pixels in the related window.

5 Results and Discussion

Various video sequences contaminated with mixed noise and blur were analyzed. Block-Matching and 3D filtering-BM3D [14] and two dimensional Non-local Means Denoising methods were taken as reference filters. Initially our proposed algorithm was tested on an image contaminated by additive Gaussian noise and impulse noise (Fig. 2).

Objective qualitative measure such as Peak Signal to Noise Ratio (PSNR), Structural Similarity (SSIM) and Improved Minimum Mean Squared Error (IMMSE) are shown in Table 1. PSNR is calculated for RGB component while SSIM and IMMSE would be calculated only for Luminance component.

Test video that was corrupted by various noise scenarios and motion blur was tested and compared with different filters with reference to PSNR and SSIM values. Figure 3 shows the results of Non-Local Mean filter, BM3D filter and our proposed filter results for nature video sequence. The objective qualitative measures are depicted in Table 2.

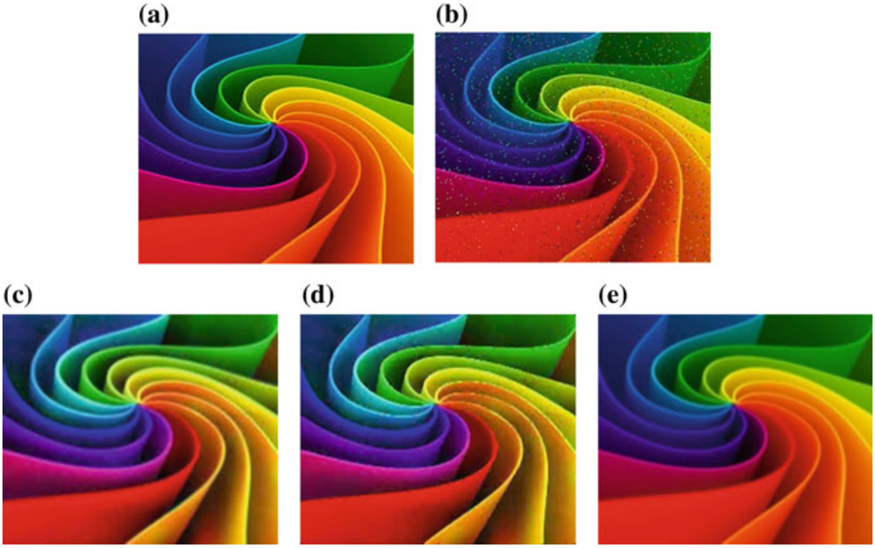

Fig. 2 Sample of **a** input image, **b** image corrupted by mixed noise, **c** non-local mean filter, **d** BM3D filter and **e** adaptive spatio-temporal filter

Table 1 Comparative qualitative measures of various methods for color image

Filters	PSNR (dB)	SSIM	IMMSE
Non-local mean	28.10	0.67	2.112
BM3D	26.07	0.71	2.330
Proposed adaptive spatio-temporal filter	33.48	0.88	2.664

Table 2 Comparative qualitative measures of various methods for nature video sequence

Filters	PSNR (dB)	SSIM	IMMSE
Non-local mean	23.50	0.23	2.912
BM3D	24.42	0.19	2.830
Adaptive spatio-temporal filter	29.48	0.816	2.564

Fig. 3 Video sequence of Nature: **a** input image, **b** image corrupted by mixed noise, **c** non-local mean filter, **d** BM3D filter and **e** adaptive spatio-temporal filter

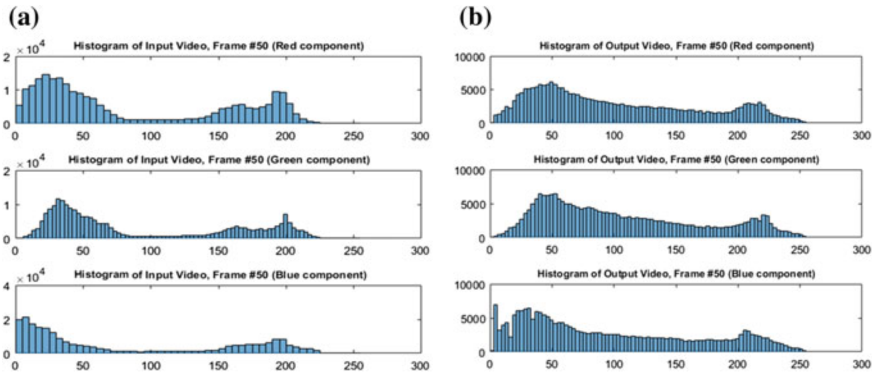

Fig. 4 Histogram graph of RGB component **a** before and **b** after execution of 50th frame of nature video sequence

The corresponding histogram of the RGB component before and after the execution is as shown in the Fig. 4. It is noted that the BM3D gives better results for Gaussian noise but the computational complexities of this filter gives limited application and also the filter requires the knowledge of noise model for execution.

Our proposed method even though it gives poor result for Gaussian noise it is simpler and gives good result for mixed noise. Even the Non-Local Mean filter does poor filtering of mixed noise.

It can be noted that the BM3D gives better results for Gaussian noise but the computational complexities of this filter gives limited application and also the filter requires the knowledge of noise model for execution. Our proposed method even though it gives poor result for Gaussian noise it is simpler and gives good result for mixed noise. Even the Non-Local Mean filter does poor filtering of mixed noise.

6 Conclusion

A universal noise removal system and video enhancement system has been presented which uses optical flow to reduce the motion blur. The adaptive Spatio-temporal filter switches from spatial to temporal based on the noise intensity which makes the model more robust. The approach in this paper removes the various scenarios of noises in video sequences and it even removes the motion artifacts caused due to undesirable camera movement. It is observed from the objective qualitative analysis of PSNR and SSIM that this proposed method gives desired result and is computationally simpler. However the execution time of this method can be reduced further if this algorithm is implemented on DSP processor with real time implementation.

References

1. P. Kisilev, S. Schein, Real-time video enhancement for high quality video conferencing. IEEE Trans. Commun. **31**(4), 532–540 (2010)
2. R. Garnett, T. Huegerich, C. Chui, W. He, A universal noise removal algorithm with an impulse detector. IEEE Trans. Image Process. **14**(11), 1747–1754 (2005)
3. M. Szczepanski, Fast spatio-temporal digital paths video filter. Real-Time Image Process. (2016). doi:10.1007/s11554-016-0561-7. Springer, Berlin, Heidelberg
4. J. Astola, P. Haavisto, Y. Neuovo, Vector median filters. IEEE Proc. **78**, 678–689 (1990)
5. B. Smolka, Peer group switching filter for impulse noise reduction incolor images. Pattern Recogn. Lett. **31**(6), 484–495 (2010). doi:10.1016/j.patrec.2009.09.012
6. B. Smolka, K. Plataniotis, A. Chydzinski, M. Szczepanski, Self-adaptive algorithm of impulsive noise reduction in color images. Patt. Recogn. **35**(8), 1771–1784 (2002)
7. C. Tomasi, R. Manduchi, Bilateral filtering for gray and color images, in *Proceedings of the 6th IEEE International Conference on Computer Vision (ICCV'98)*, Bombay, India, January 1998, pp. 839–846
8. M. Ben-Ezra, S. Nayar, Motion deblurring using hybrid imaging, in *Proceedings of CVPR 2003* (2003), pp. I-657–664
9. P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), 629–639 (1990)

10. K. Buades Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007), doi:10.1109/TIP.2007.901238
11. K. Radlak, B. Smolka, B, Trimmed non-local means technique for mixed noise removal in color images, in *2013 IEEE International Symposium on Multimedia (ISM)* (2013), pp. 405–406
12. A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in *Proceedings of CVPR 2005* (2005), pp. II-60–65
13. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Process. **16**(8), 2080–2095 (2007), doi:10.1109/TIP.2007.901238
14. M. Maggioni, V. Katkovnik, K. Egiazarian, A. Foi, Nonlocal transform-domain filter for volumetric data denoising and reconstruction. IEEE Trans. Image Process. **22**(1), 119–133 (2013). doi:10.1109/TIP.2012.2210725
15. Y.W. Tai, H. Du, M. Brown, S. Lin, Image/video deblurring using a hybrid camera. Proc. CVPR **2008**, 1–8 (2008)
16. C. Wang, L.-F. Sun, B. Yang, Y.-M. Liu, S-Q Yang, Video enhancement using adaptive spatio-temporal connective filter and piecewise mapping. EURASIP J. Adv. Signal Process. 13 pages (2008). Article ID 165792, doi:10.1155/2008/165792
17. J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain. IEEE Trans. Image Process. **12**(11), 1338–1351 (2003)