A Robust Video Denoising System Using Optimised Surfacelet Transform

Mohammed Khalid
Department of ECE
Sree Chitra Thirunal College of Engineering
Pappanamcode, Trivandrum, Kerala

P. Sajith Sethu
Assistant Professor, Department of ECE
Sree Chitra Thirunal College of Engineering
Pappanamcode, Trivandrum, Kerala

Sethunadh R
Scientist / Engineer
Vikram Sarabhai Space Centre
Indian Space Research Organisation
Trivandrum, Kerala, India

Abstract— **The conventional video denoising algorithms utilizes either a strenuous motion estimation step or by the frame by frame wavelet transformation without exploiting the correlation between neighboring frames. However, these schemes of video denoising results in videos with jittery edges and curves. The limitations of motion estimation based schemes are that they suffer due to aperture problems in optical flow and lighting variations. Surfacelet transform is a powerful tool for the representation of multidimensional data. Video signals can be treated as a different type of 3D signal and therefore can be processed using surfacelet transform which preserves the edge information and visual quality. By incorporating the advantages of surfacelet transformation with an advanced optimization technique we propose a novel video denoising technique which was found to be delivering overwhelming results in terms of peak signal to noise ratio(PSNR) and structural similarity(SSIM) index.**

Keywords— **Surfacelet transform , Video denoising, Threshold,NDFB,PSNR,SSIM.**

I. Introduction

The primary aim of all video denoising systems is to remove noise from a corrupted video sequence. A video is corrupted often due to the limitations of the acquisition and processing devices. Most of the conventional video denoising schemes employ the technique of motion estimation or the optical flow estimation. Motion estimation is mostly an arduous technique particularly in conditions with lighting variations. Motion estimation step is also worsened due to the aperture problem of the optical flow estimation. This limitation of motion estimation paved the way for wavelet transform based video denoising techniques. Unfortunately, those systems resulted in videos with jittery edges and curves.

Video representation with edge preservation is one of the main challenges faced by the researchers. The authors in [1]–[4] used motion based 3D transforms for video representation. Bamberger and Smith proposed a novel technique for 2D data representation called the directional filter bank [5]. In 2007, the authors in [6] realized the surfacelet transformation by integrating directional filter bank (DFB) in higher dimensions utilizing a different kind of pyramid for multiscale decomposition. In the surfacelet based approach, the video is decomposed into motion selective subbands. Hence, the motion information in video signals is accurately preserved. Another important feature of surfacelet transformation is that its directional information can be improved by introducing more decomposition levels.

Researchers developed diverse schemes for video denoising methods overcoming the limitations of the wavelet based methods. In 1999, Park proposed a 3-D velocity selective filter bank by applying two 2-D DFBs separately along two signal planes for video representation and processing [7]. The authors in [8] utilized a spatial filter operating in wavelet domain with Markov Random Field model. The authors presented different spatial and spatio-temporal filters [8]-[13] to remove noise from video sequences. At high noise levels the spatial filters cause blurring. The authors in [12] developed a pixel based approach wherein the new pixel values are obtained by utilizing weighted average of motion corrected frames. An edge preserving spatio-temporal video noise filter that combines 2D Wiener and Kalman filters has been presented in [13]. A video filtering algorithm utilizing wavelet transform is presented in [14]. A content adaptive video denoising filter [15] has been proposed in 2005 for video denoising. These methods however struggled in high noise variances and in videos with higher edge details. The proposed algorithm outperformed other existing algorithms contrasted on the basis of visual quality, PSNR and SSIM.

The paper is categorized as follows. Section II gives an overview of Surfacelet Transform. Section III introduces the surfacelet based video denoising method. In Section IV, the results and discussions of the proposed method with other existing methods are presented. Conclusions are given in the final section.

II. Surfacelet Transform

The surfacelet transform is a potent multidimensional processing tool developed by Yue Lu and Minh N Do in 2007.

Surfacelet transform was developed by integrating an N dimensional directional filter bank (NDFB) with a new multiscale pyramid [6]. The key points regarding surfacelet transform is discussed in this section.

The primary element of the surfacelet transform is the N dimensional directional filter bank(NDFB) which is actually a higher dimensional realization of directional filter bank developed by Bamberger and Smith in 1992 [5]. The frequency partition in 2D representation given in Figure 1(a) should be realized in higher dimensions. Thus in 3D volumetric data (like video signals), the ideal pass bands of the component filters are rectangular-based pyramids radiating out from the origin at different orientations and tiling the entire frequency space [6] which is shown in Figure 1(b). A three-channel undecimated filter bank is used to obtain the first level of decomposition in NDFB which decomposes the frequency spectrum of the input signal into three subbands, with their directions aligned with the ω_1, ω_2 and ω_3 axes respectively. The output of the undecimated filter bank is then decomposed sequentially by 2D filter banks IRC (1, i) operating along (n_1, n_i) planes, where i=2,3.

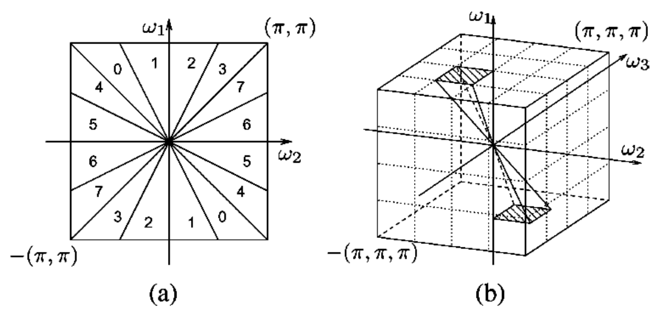


Figure 1 (a) Frequency partitioning of the directional filter bank with three levels of decomposition (b) Frequency partitioning of NDFB in 3-D

The NDFB filter bank described above has the following useful properties:

1) Directional decomposition
2) Efficient tree-structured construction
3) Ideal reconstruction
4) Small redundancy

The extension of DFB to higher dimensions was one of the major challenges faced by researchers. The researchers developed diverse methods for the same. Bamberger proposed a 3-D subband decomposition scheme implemented by applying the checkerboard filter banks separately along two orthogonal signal planes followed by a 2-D DFB decomposition on one of the planes [5]. But unfortunately such filter banks had pass bands in the shape of triangular based prisms and it didn't exhibit a single dominant direction. In 1999, Park proposed a 3-D velocity selective filter bank by applying two 2-D DFBs separately along two signal planes [8]. The filter bank proposed by Park offered the optimal frequency partition for N dimensional directional decomposition at the price of increased redundancy. An optimum solution for the quest was proposed in [6], called NDFB, which provides the optimal frequency partition at minimum redundancy.

Surfacelet transform is actually a multiscale version of NDFB. Surfacelet transform is developed by scaling the output of NDFB at different scales. Scaling the signal at different scales help in grabbing surface singularities in 3D volume. In contourlet transform, the DFB is integrated with multiscale operation by the use of a laplacian pyramid. Unlike contourlet transform [3], the surfacelet transform utilizes a different kind of pyramid for multiscale decomposition. The multiscale pyramid construction is explained in future discussion.

III. The Proposed Method

The main steps for video denoising using surfacelet transform can be summarized as:

1) Initially the original video sequences are corrupted with Gaussian noise with mean zero and variance σ^2
2) The noisy video sequence is then applied to preprocessing filter for initial stage of noise removal
3) The surfacelet transformation of the preprocessed video sequence is taken
4) Otsu thresholding optimized using particle swarm optimization is then applied to the surfacelet coefficients
5) Apply inverse surfacelet transform to denoised coefficients to get the denoised video sequences

The block diagram of the proposed method is shown in Figure 2.

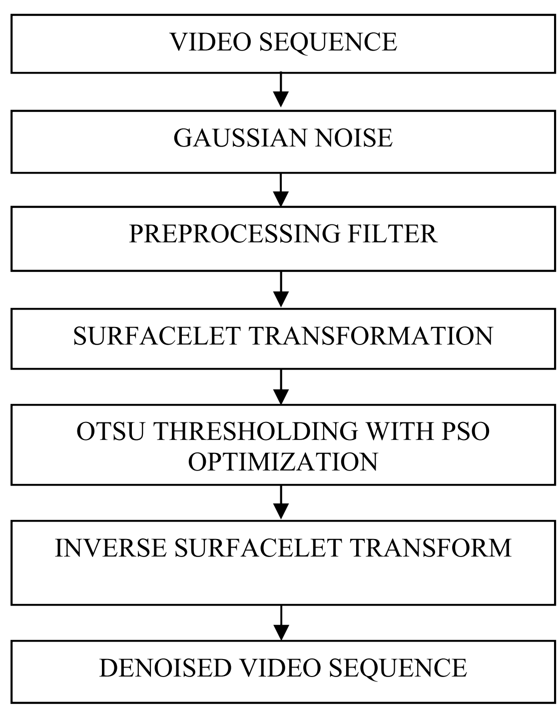


Figure 2. Block diagram of the proposed method

The block diagram is elaborated below. Additive white Gaussian noise with different variances is added to obtain noisy video sequences. Preprocessing of video sequences is done by using bilateral filtering. The noisy sequences are then transformed into surfacelet domain. Video denoising aims at the removal of noisy coefficients from the real ones, which can be accomplished by thresholding. Here, Otsu thresholding optimized by PSO is applied to the surfacelet coefficients. After thresholding, denoised surfacelet coefficients are obtained. Finally, the video data is obtained by inverse surfacelet transformation of the denoised surfacelet coefficients.

The major steps involved in the denoising algorithm are explained in the following subsections.

A. Otsu Thresholding

Let $S_K^j(x,y)$ represent the initial surfacelet coefficient in the point (x, y) in each sub-band $K \epsilon \{K_1^j\ K_2^j \dots K_k^j\}$at scale j. The aim of this paper is to obtain denoised coefficient $D_k^j(\mathrm{x},\mathrm{y})$ at the point $S_K^j(x,y)$ by adjusting the pixel values. In this paper, we use Otsu thresholding [16] to obtain the denoised surfacelet coefficients i.e.,

$$D_k^j(\mathrm{x},\mathrm{y}) = \begin{cases} S_K^j(x,y) & if\ S_K^j(x,y) > T \\ 0 & if\ S_K^j(x,y) < T \end{cases} \quad (1)$$

where T is Otsu threshold. The procedure to obtain T is summarized below [16]:

1) Compute normalized histogram
2) Set the threshold T=K and divide $f(x,y)$ ino two classes C_1 and C_2. Probability that $f(x,y)$ is in class C_1 is $P_1(K)$ and probability that $f(x,y)$ in class C_2 is $P_2(K)=1-P_1(K)$
3) Compute cumulative mean of C_1 and C_2 i.e., M1 and M_2
4) Compute Global mean, $M_G = P_1M_1 + P_2M_2$
5) Compute the in-between class variance,

$$\sigma_B^2(k) = \frac{[M_G\,P_1(k) - M(k)]^2}{P_1(k)\ \times P_2(k)} \quad (2)$$

6) Obtain Otsu Threshold, T= argmax $[\sigma_B^2(k)]$

B. Particle Swarm Optimization

Particle Swarm Optimization (PSO) [17] algorithm is a population based optimization technique. PSO algorithm is inspired from the food searching behavior of a birds' flock. PSO can be classified into two: global-best PSO and local-best PSO. For global best PSO, the entire swarm is taken into account for calculating the personal best position for each particle. On the other hand, in local best PSO, the entire swarm is divided into sub regions and the best is chosen based on the position of other swarm agents. Each solution options are represented by each swarm agents in the search space. Each swarm agent remembers the coordinates in the search space which provided the best solution provided by the swarm agent called the personal best. Global best value is the best solution obtained by any swarm agent inside the entire search space. The core concept of PSO lies in moving each swarm agents toward its p_{best} and g_{best} locations, using a random weighted acceleration (w) at each iteration. The swarm agent's velocity is updated using the equation below

$$X_i(t+1) = w.X_i(t) + c_1r_1\left(Z_{pbesti}(t) - Z_i(t)\right) + c_2r_2\left(Z_{gbesti}(t) - Z_i(t)\right) \quad (3)$$

Where, $X_i(t)$ is the velocity of the i^{th} agent at time't', c_1 and c_2 are the self confidence and swarm confidence respectively, r_1 and r_2 are uniformly distributed weighting functions

Finally, the position of each swarm agent is updated using the formula

$$Z_i(t+1) = Z_i(t) + X_i(t+1) \quad (4)$$

PSO is a metaheuristic algorithm as it requires no prior knowledge about the optimal solution. It can be implemented in wider search spaces.

IV. RESULTS

The proposed video denoising system was tested using 'Miss America' and 'Coastguard' video sequences. The denoising algorithm was implemented in MATLAB 2013 and surfacelet transformation with four levels of decomposition using Surfbox toolbox. The resolution of the video was truncated to 192 x 192 and the number of frames was limited to 192.The performance of the algorithm was validated at different noise levels and by using different thresholding schemes on the basis of PSNR and SSIM.

The peak signal to noise ratio is given by,

$$PSNR = 20\log_{10}\left(\frac{255}{MSE}\right) \quad (5)$$

where MSE is the mean square error. Given an image $f_r(i,j)$ and original image$f_o(i,j)$, then MSE is given by,

$$MSE = \sum_{i,j}\frac{[f_o(i,j) - f_r(i,j)]^2}{M\times N} \quad (6)$$

where M ×N is the video resolution.

SSIM is a metric used for benchmarking the performance of different denoising and enhancement systems. SSIM is given by

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (7)$$

where x, y are image patches from original and distorted images, c_i= $(k_i\text{L})^2$ are two variables to stabilize the division with weak denominator with k_1=0.01 and k_2=0.03 and L is the dynamic range of the frame.

A comparison of PSNR and SSIM values calculated on the two sequences under different noise levels are tabulated in Table I and Table II respectively.

TABLE I
COMPARISON OF THE PERFORMANCE OF THE PROPOSED METHOD TO OTHER METHODS IN TERMS OF PSNR (dB) FOR VIDEO SEQUENCES

Sequence	Miss America	Miss America	Coastguard	Coastguard
Techniques / σ	σ=10	σ=20	σ=10	σ=20
Wiener 2D	22.14	21.79	22.15	20.89
3D curvelet	24.56	23.18	27.25	27.56
Surfacelet	28.13	26.63	30.15	27.13
The Proposed Method	36.57	35.43	35.83	33.73

TABLE II
COMPARISON OF THE PERFORMANCE OF THE PROPOSED METHOD TO OTHER METHODS IN TERMS OF SSIM FOR VIDEO SEQUENCES

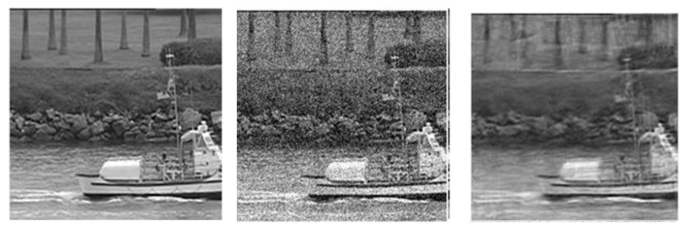
Sequence	Miss America	Miss America	Coastguard	Coastguard
Techniques /σ	σ=10	σ=20	σ=10	σ=20
Wiener 2D	0.860	0.774	0.859	0.778
3D Curvelet	0.865	0.839	0.839	0.818
Surfacelet	0.937	0.917	0.915	0.903
The Proposed Method	0.956	0.935	0.937	0.925

The original, noisy and denoised versions of the Miss America sequence are shown in Figure 3.

(a) (b) (c)

Figure 3:Denoised Miss America Sequence (a) Original sequence, (b) Noisy sequence, (c) Denoised Sequence

The original, noisy and denoised versions of the coastguard sequence are shown in Figure 4.



(a) (b) (c)

Figure 4: Denoised Coastguard sequence, (a) Original sequence, (b) Noisy sequence, (c) Denoised Sequence.

On the basis of the observation, the proposed method using Otsu threshold with PSO optimization outperforms existing video denoising algorithms in terms of PSNR and SSIM.

V. CONCLUSION

A novel framework for video denoising using surfacelet transformation with PSO optimized Otsu thresholding was explored. The denoising algorithm was tested with different standard video sequences. A significant increase in PSNR and SSIM was obtained and the denoised videos were visually appealing with well-preserved edges.

REFERENCES

[1] F. A. Mujica, J.-P. Leduc, R. Murenzi, and M. J. T.Smith, “A new motion parameter estimation algorithm based on the continuous wavelet transform,” *IEEE Trans. Image Proc.*, vol. 9, no. 5, pp. 873–888, May2000.

[2] W. Selesnick and K. Y. Li, “Video denoising using2D and 3D dual-tree complex wavelet transforms,” in *Proc. of SPIE conference on Wavelet Applications in Signal and Image Processing X*, San Diego, USA, August2003.

[3] M. N. Do and M. Vetterli, “The contourlet transform: an efficient directional multiresolution image representation,” *IEEE Trans. Image Proc.*, vol. 14, no. 12, December 2005.

[4] E. Chang and A. Zakhor, “Subband video coding based on velocity filters,” in *Proc. IEEE International Symposium on Circuits and Systems*, May 1992.

[5] R. H. Bamberger and M. J. T. Smith, “A filter bank for the directional decomposition of images: Theory and design,” *IEEE Trans. Signal Process.*, vol. 40, no. 4, Apr. 1992, pp. 882–893.

[6] Y. M. Lu and M. N. Do, “Multidimensional Directional Filter Banks and Surfacelets," in *IEEE Transactions on Image Processing*, vol. 16, no. 4, pp. 918-931, April 2007.

[7] S. Park, “New directional filter banks and their applications in image processing,” Ph.D. dissertation, Georgia Inst. Technol., Atlanta, 1999

[8] Malfait M, Roose D, “Wavelet Based Image Denoising Using a Markov Random Field a Priori Model,” *IEEE Transaction on Image Processing,* Vol. 6, No. 4, 1997, pp.549-565.

[9] Kazubek M, “Wavelet Domain Image Denoising by Thresholding and Wiener Filtering,” *IEEE Signal Processing*, Vol. 10, No. 11, 2003, pp. 324-326.

[10] Van De Ville D, Van der Weken D ,Nachtegael M, Kerre E. E., Philips W., and Lemahieu I, “Noise Reduction by Fuzzy Image Filtering,” *IEEE Transaction on Fuzzy Systems,* Vol. 11, No. 4, 2003,pp. 429-436

[11] Sadhar S. I., and Rajagopalan A. N, “Image Estimation in Film-Grain Noise,” *IEEE Signal Processing Letters*, Vol. 12, No. 3, 2005, pp.238-241.

[12] Ozkan M. K., Sezan I., and Tekalp A. M, “Adaptive Motion Compensated Filtering of Noisy Image Sequences,” *IEEE Transaction on Circuits and Systems for Video Technology*, Vol.3, No. 4, 1993, pp.277-290.

[13] Dugad R., and Ajuha N., “Video Denoising by Combining Kalman and Wiener Estimates,” *IEEE International Conference on Image Processing*, Kobe, Japan, 1999, pp. 152-161.

[14] Gupta N, Swamy M. N. S, and Plotkin E. I, “Low Complexity Video Noise Reduction in Wavelet Domain,” *IEEE 6th Workshop on Multimedia Signal Processing*, 2004, pp. 239-242.

[15] Chan T.-W, Au O. C, Chong T. S, and Chau W.S, “A Novel Content-Adaptive Video Denoising Filter,” *IEEE ICASSP*, Philadelphia, PA, USA, Vol. 2, 2005, pp. 649-652.

[16] N.Otsu, “A Threshold Selection Method from Gray-Level Histograms” *IEEE Transactions On Systems, Man, And Cybernetics*, vol.9, no.1,1979, pp.62-66.

[17] S. Das, A Abraham and A Konar, “Particle Swarm Optimization and Differential Evolution Algorithms: Technical Analysis, Applications and Hybridization Perspectives, *“Studies in Computational Intelligence (SCI)*, 2008.