Research Journal of Applied Sciences, Engineering and Technology 3(10): 1182-1187, 2011
ISSN: 2040-7467

Submitted: July 20, 2011 Accepted: September 25, 2011 Published: October 20, 2011

A Video Denoising Method Based on Grouping the Similar Blocks and Surfacelet

[1]Peng Geng, [1]Li Xianbin, [2]Yu Wanhai, [1]Yan Shiqiang and [3]Liu Jianhu
[1]The Center of Educational Technology, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China
[2]Xingtai Polytechnic College, Xingtai, 054035, China
[3]Electrical and Electronics Engineering Department, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China

Abstract: In this study, a novel video denoising method combining similar block matching based on the PDS searching method, grouping these similar block strategy and Surfacelet transform is proposed. Firstly, we apply the SAD block matching criterion and PDS search algorithm which we proposed by searching all frames for blocks which are similar to the currently processed one. In the complex motion case, the method of PDS can effectively decrease the number of search points and search number. Secondly, the matched blocks are put together to stack into some new third dimension sub-video sequences and because of the similarity between them, the data in the video array exists high level of correlation. We apply the surfacelet transform to them and effectively attenuate the noise by solid threshold shrinkage of the coefficients. Finally, inversely transforming the coefficients and obtaining the denoising video according to the obtained locations in the block matching process. This algorithm is obviously better than other 3D Curvlet and 3D wavelet method in the denoising effect and the PSNR is increased about 0.6dB. In terms of visual quality, the proposed method can effectively preserve the video detail, and the trajectory of motion object is very smooth, which is especially suitable to process the video flames with plenty of large area movement object or background change.

Key words: NDDFB, PDS, SAD, Surfacelet transform, video denoising

INTRODUCTION

The traditional wavelet denoising is usually each frame denoising, without considering the correlation between each frame movement, moving objects trailing phenomenon. A new video denoising algorithm is the video signal as a special 3d signal, three-dimensional transform to regard it as a whole, the algorithm is effective to solve the moving object trailing, flashing and algorithm robustness problems. (Selesnick and Li, 2003), (Yang *et al*., 2009) and (Aquino *et al*., 2009) use the 3D double tree complex wavelet (3DTCWT) to decompose signal to six directions, has good selectivity, but when the movement in video are very intense, the video texture is very much, its directions are obviously insufficient. (Ying *et al*., 2005) and (Minhas *et al*., 2011) proposed 3D Curvelet method like the 3DTCWT method. Yue Lu's Surfacelet Transform (Lu and Do, 2007) is a new kind of 3D transformation, has many direction decomposition, anisotropy, the efficiency of the tree structure filter, completely rebuilding and low redundancy properties, which is very suitable for video processing. (Xiao *et al*., 2008) proposed the 3DCMST method that the ST coefficient matrix is divided into several parts using 3D context model in the threshold value selection according to the energy, but without considering the original video frame correlation, no motion detection and motion estimation, and so has great redundancy, especially for the high correlation between frames.

This study presents a BMG-3DST (Block-matching Grouping 3D surfacelet transform) Video denoising algorithm. Firstly, using SAD standards and PDS (Pseudo-diamond Search) search algorithm, the motion estimation, motion detection and interframe block matching are made between the video frames. Secondly, group the matched bocks from the video into some groups which are new 3D subsequence. Adopting Surfacelet transform on the new 3D subsequence and filtering the surfacelet transform coefficients using the hard threshold. Finally, inverse surfacelet transform the filtered coefficients to the 3D subsequence. And restore the 3D subsequence to original video. Experiments show that this method can achieve better visual effect and PSNR. Especially, the method is suitable for high correlation between neighboring frames of the video, such as the object image containing acuteness movement video.

Corresponding Author: Peng Geng, The Center of Educational Technology, Shijiazhuang Tiedao University, Shijiazhuang, 050043, China

BLOCK MATCHING AND GROUPING

Video sequence has the space redundancy and temporal redundancies. The motion estimation can effectively eliminate motion video frame exists between the temporal redundancies. The video compression standards such as MPEG1/24, h.261, h.263 h.264 are based on block matching motion estimation algorithm.

Block matching criterion: Block matching algorithm of matching effect depends on the search mode, matching criterion, etc. Matching criterion refers to the cost function, motion estimation for that cost function is the smallest reference image block relative to the current image block of displacement, and the displacement is sports vector. For the same image block, different matching criterion may have different motion vector, decoding the image quality will also have a difference. Commonly block matching criterion including MSE, MAD, and MSE etc. My method uses SAD as the Block matching criterion.

$$SAD(m,n) = \sum_{i=0}^{N-1}\sum_{i=0}^{N-1} f_c(i,j) - f_r(i+m, j+n) \qquad (1)$$

$f_r(i,j)$ is pixel value in the reference block. $f_c(i,j)$ is the pixels of the current block. The size of the block matching is N*N. m and n represents the search step of the block. The best motion vector is the m and n which make the smallest SAD

Block matching algorithm: In 2002, (Tao, 2002) proposed a Pseudo-Diamond Search Method to matching the blocks in different frame. This method and take full account of the details and the correlation between pixels. The basic search point is shown in Fig. 1. The concrete steps are as follows:

- **Step 1:** Searching the point of center (the hollow points) and the four points in the four points (the gray points) in Fig. 1 and calculating the value of SAD. Then comparing it with the threshold. If the SAD is less than the threshold Point, the point is a matching point and stop searching. Otherwise, enter step 2.
- **Step 2:** Making the minimum value of SAD points (sub-optimum) in step 1 as the center, searching the black points of apex of the diamond in Fig. 1 and calculating the value of SAD. Comparing it with the threshold. If the SAD is less than the threshold Point, the point is the matching point and stop searching. If the SAD of the two points in any the diamond's edge is small but larger than the threshold and similar to each other (Less than the special gray difference), then go to step 3. Otherwise, go to step 4.
- **Step 3:** Searching the other point not being searched on the diagonal based on the two points. Calculating

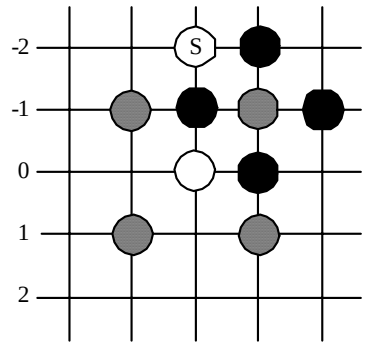


Fig. 1: Schematic diagram of the basic search point of the PDS method

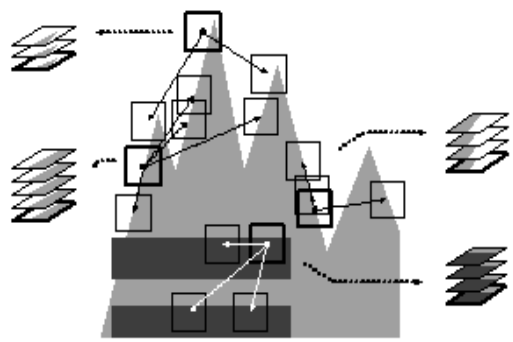


Fig. 2: Grouping only in a frame

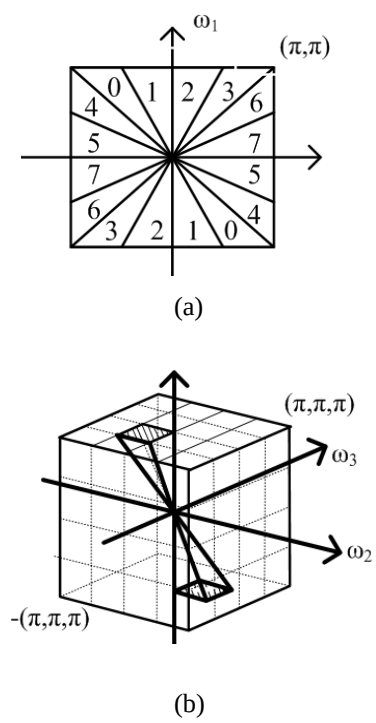


Fig. 3: Ideal multidimensional filt

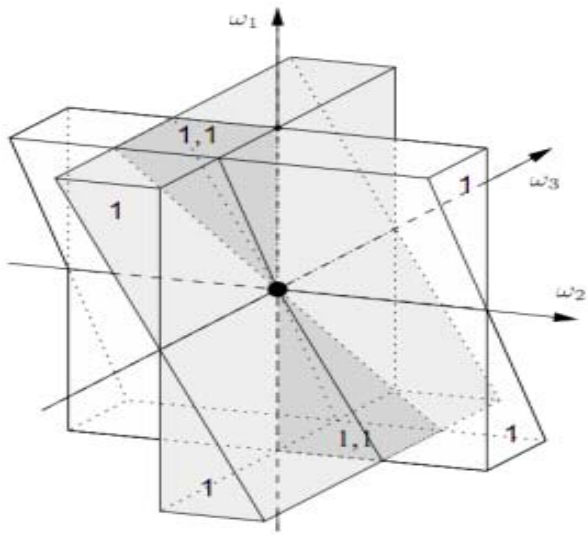


Fig. 4: 2D-DFB and 3D-DFB frequency decomposition (a) 2D-DFB Frequency decomposition(b) 3D-DFB Frequency decomposition

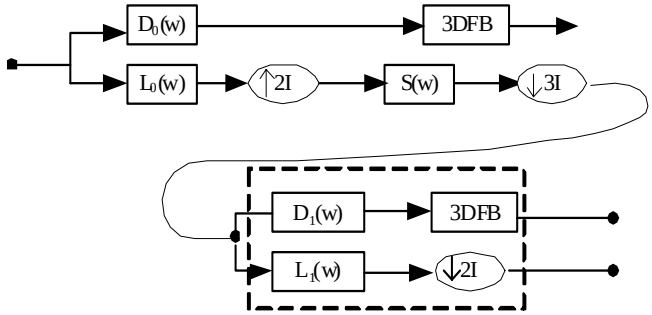


Fig. 5: The block diagrams of the proposed surfacelet transform (The forward transform 3DFB)

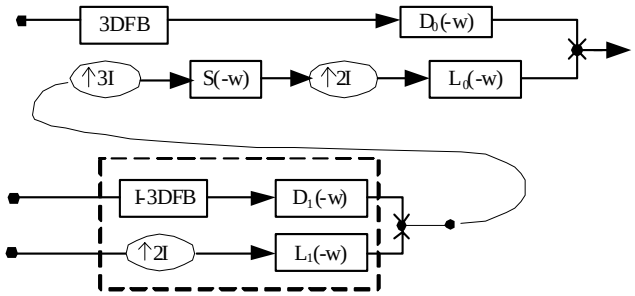
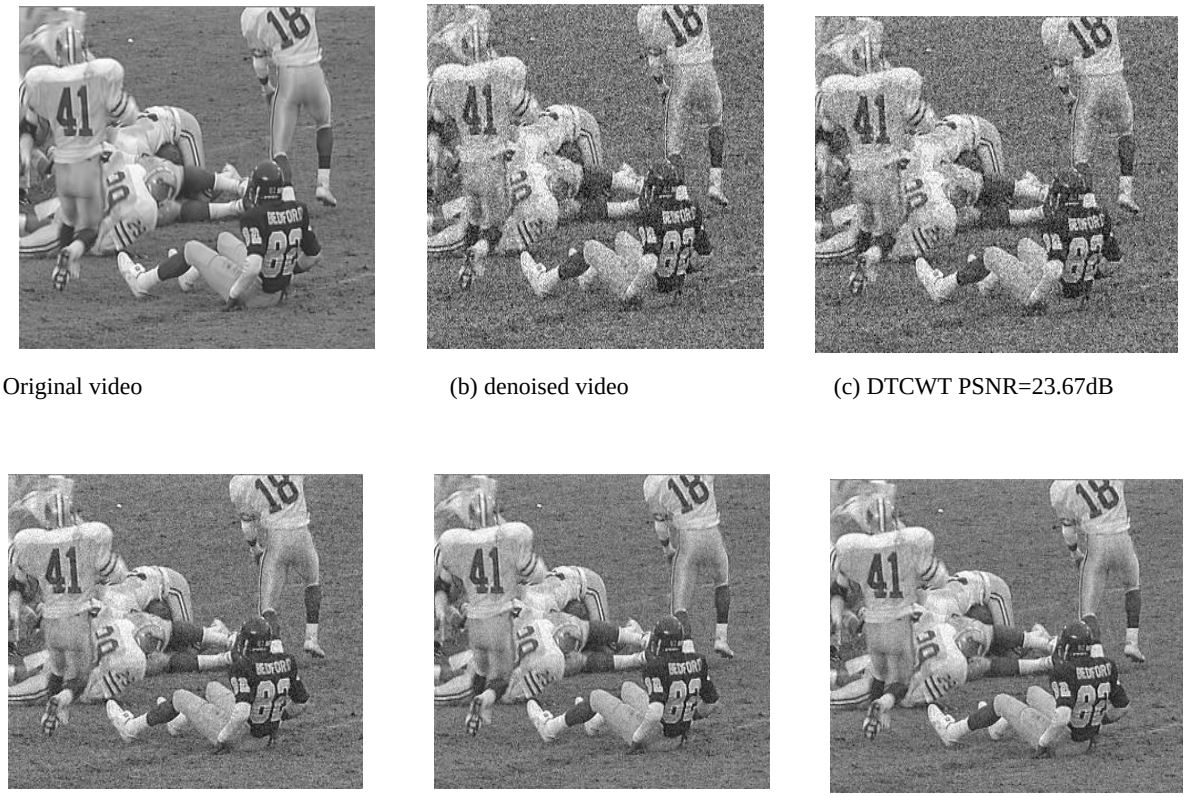


Fig. 6: The block diagram of the proposed inverse surfacelet transform (its inverse I-3DFB are attached to the highpass)

he value of SAD and comparing it with the threshold. If the SAD is less than the threshold point, the point is the optimum point and stop searching. Otherwise, enter step.

- **Step 4:** Making the previous point with the minimum SAD (sub-optimum) as the center, then go to step 1.

Grouping: According to block matching algorithm and the matching principle as above, the grouping is method finding the image blocks satisfying matching criterion from the reference frames and the rest frame in the video sequences (Fig. 2) V(x,y,t) and stacking these matching blocks into a new group stored into a new 3D matrix



(a) Original video (b) denoised video (c) DTCWT PSNR=23.67dB

(d) ST PSNR=24.98dB (e) 3DCMST PSNR=26.31dB (f) BMG-3DST SNR=27.12dB

Fig. 7: Effect of the denoised football video

G_L (i, j, k). L is the number of group. In a sense, the G_L (i, j, k) is a video sequence and suitable for 3D Surfacelet transform. Figure 3 is the Schematic diagram for block matching in a frame. Block matching grouping between frames are similar and the description method is omitted because of the limited length.

SURFACELET

ND-DFB: In 1992, Bamberger and Smith proposed firstly the DFB (Directional Filter Bank). DFB is a non redundant transformations, can achierve the complete reconstructin the singal.

The 2D-DFB frequency decomposition is as following in Fig. 3a. Lu and Do (2007) expanded the 2D-DFB to N dimension and puts forward a new ND-DFB (N Dimension Directional Filter Banks).

Figure 4 illustrates one of the key ideas of NDFB. The dark region in the Fig. 4 shows the pyramid shaped support in frequency that we want to get. We can see that the pyramid support can be obtained as an intersection of two wedge supports shown in gray regions. This observation leads to the idea of obtaining pyramid shaped supports in 3D DFB by a concatenation of two 2D DFB on appropriate dimensions. The 3D DFB frequency division is as in Fig. 3b.

In general, the ideal passbands of the NDFB in dimmensions are hypercube-based hyperpyramids out from the origin. With the following distinctive properties:

- **Construction:** The NDFB has an efficient tree structured implementation using iterated filter banks.
- **Angular resolution:** The number of directional subbands can be increased by iteratively invoking more levels of decomposition through a simple expansion rule. In general, there can be N×l ($l \geq 0$) different directional subbands in the N-D case.
- **Perfect reconstruction:** The original signal can be exactly reconstructed from its transform coefficients in the absence of noise.
- **Small redundancy:** The NDFB is N-times expansive in the N-dimensional case.

Surfacelet transform: Do and Vetterli (2005) proposed the Contourlet transform to obtain a multi-directional and multi-resolution method of sparse image in combination on the basis of the pyramidal decomposition and the 2D-DFB. Combining the NDFB with a new multiscale pyramid, Lu and Do (2007) proposes the Surfacelet transform, which can be used to efficiently capture and represent surface ike singularities in multidimensional data. This strategy is analogous to the contourlet

construction, in which the original 2 DFB is combined with a multiscale decomposition.

The multidimensional surface singular of signal can be captured by the 3D-DFB frequency decomposition. Because the 3D-DFB only processes the high frequency part of signal, therefore in the use of 3D-DFB signal, signal must be decomposed by multi-scale. A new Surfacelet transformation is constituted by Combining multi-scale decomposition and the 3D-DFB. In ST transform, multi-scale decomposition adopted a pyramid structure as shown in Fig. 5 and 6. In the new multiscalepyramid depicted in Fig. 5, the lowpass filter $L_0(w)$ in the first level is downsampled by a non-integer factor of 1.5 along each dimension. Although this fractional sampling factor makes the new pyramid slightly more redundant than the Laplacian pyramid, the added redundancy to be very useful in reducing the frequency domain aliasing of the NDFB, which is concentrated on the boundaries of the frequency $[-\pi, \pi]^N$ cell and mainly caused by the 2π periodicity of the frequency spectrum of discrete signals. Intuitively, the new multiscale pyramid achieves the task of eliminating aliasing components by only keeping the middle (alias-free) portion of the NDFB filter responses. Consequently, the constructed surfacelet are well-localized in both the spatial and frequency domain. The lowpass filters $L_i(\omega)$ (i = 0,1) in the frequency domain as:

$$L_i(\omega) = d_i . \prod_{n-1}^{N} L_i^{1D}(\omega_n) \tag{2}$$

where $d_1 = 6^{N/2}$ and $d_2 = 6^{N/2}$; $d_1 = 6^{N/2}$ is a 1-D lowpass filter along the ω_n axis with passband frequency $\omega_{p,i}$ stopband frequency $w_{s,i}$ and a smooth transition band, defined as:

$$L_i^{ID}(\omega) = 1 \quad |\omega| \le \omega_{p,i} \tag{3}$$

$$L_i^{1D}(\omega) = 1 \quad \omega_{p,i} \le |\omega| \le \pi \tag{4}$$

$$L_i^{ID}(\omega) = \frac{1}{2} + \frac{1}{2} Cos \frac{(|\omega| - \omega_{pi})}{\omega_{si} - \omega_{pi}} \text{, Other} \tag{5}$$

The main advantage of designing the filters in the frequency domain is that we can let their frequency responses to be strictly zero beyond some cutoff frequencies.

By choosing $\omega_{p.A}$, $\omega_{s,A}$ and $\omega_{s,i}$ properly, we can ensure that the aliasing introduced by the upsampling and downsampling operations will be completely cancelled, and the perfect reconstruction condition for the multiscale pyramid can be simplified as:

$$\frac{L_i(\omega)}{d_i^2} + |D_i(\omega)|^2 \equiv 1 \; i = 0,1 \tag{6}$$

By combining the multiscale pyramid with the NDFB, the surfacelet transform has the ideal passband supports as pairs of concentric cubes radiating out from the origin with different directions and scales. In the spatial domain, the surfacelet basis is localized surface patches with different normal directions and spatial locations.

Surfacelet Transform has the characteristics of directional decomposition, the more efficient tree filter banks, completely reconstruction and low redundancy. It can use different scale and different frequence direction sub-block accurately to capture the surface singularity of a 3D signal. The ST coefficient energy is very concentrated.

Algorithm: Algorithm is listed below:

- **Step 1:** Using SAD standards, PDS search algorithm to find the locations of the blocks in a frame that are similar to the currently processed one.
- **Step 2:** Form a 3D array (group) by stacking the blocks located at the obtained locations until all blocks in the video are matched
- **Step 3:** Executing the 3D Surfacelet Transform on the every noisy groups (the sub-video)
- **Step 4:** Denoising the Surfacelet Transform coefficient using hard threshold
- **Step 5:** Reconstruction the groups from the denoised Surfacelet Transform coefficients
- **Step 6:** According to the steps 1, restoring the every block matched in the every group to its initial position in the video.

RESULTS

Analysis of experimental results: Our experiment is based on Matlab 2009a and the 4GB memory, and uses the MAT formats video sequences of the size of 240×320×60 such as the Bus and Flower sequence with large background changes and the Football sequence with acute movement.

Firstly, adding the variance (30 and 50) Gaussian white noise with zero mean to the video sequences. Secondly, denoising the noisy sequences by the method such as the dual-tree complex wavelet transform (DTCWT), the Surfacelet Transform with hard threshold, the 3DCMST and our proposed method based on block-matching grouping denoising with Surfacelet Transform (BMG-3DST). In the experiment, all method is executed by decomposing the video sequence into three levels.

To investigate the effect by the decomposition layers in our algorithm, applying three layers decomposition of

Table 1: The PSNR (dB) of different method

	Flower		Football		Bus	
Variance	30	50	30	50	30	50
DTCWT	24.60	22.59	23.67	20.90	23.60	21.59
ST	25.89	23.21	24.98	21.98	24.61	21.95
3DCMST	26.48	23.92	26.31	23.5l	25.66	22.12
BM-3DST	27.12	24.55	27.12	24.20	26.21	23.78

ST. Corresponding to the number of 3D-DFB direction respectively are 64, 16, 4, shrinkage only used on the two finest levels coefficients.

Table 1 gives PSNR after denoising by the four algorithms. Fig. 7 are the comparison of the football video denoising result by the four algorithms.

From the experiment results can be obtained in the following conclusions:

- The PSNR of algorithm is obviously superior to the other three algorithms. This algorithm can improve the PSNR than other algorithms about 0.6dB.
- The visual effect of subjective evaluation is better than other algorithms. The denoising effect is apparently better when this algorithm used to deal with video containing acuteness movement.

We innovatively propose the video denoising algorithm based on PDS grouping and Surfacelet transform. This algorithm can significantly improve the denoising video PSNR.

Not only can keep well the video details, and attain very good effect for movement video especially for the strenuous movement objects and the large area changes in the background. The movement objects are very fluent and does not exist the phenomenon such as flash and ghosting in the traditional algorithm.

We will focus on the next research work such as researching the more suitable block matching algorithm, adopting the different characteristics of filter for the grouping of block matching using the block-matching ST denoising to the image.

ACKNOWLEDGMENT

This study was supported in part by National Science Foundation of China under grant 61072148 and Hebei Province University Science Research Project under grant Z2010306.

REFERENCES

Aquino, A., V. Starostenko, O. Ramirez-Cortes, 2009. Detection of microcalcifications in digital mammograms using the dual-tree complex wavelet trans form. Engine. Int. Syst., 17(1): 49-63.

Do, M.N. and M. Vetterli, 2005 Contourlet transform: An efficient directional multiresolution image representation. IEEE Trans. Image Process., 14(2).

Lu, Y. and M.N. Do, 2007. Multidimensional directional filter banks and surfacelets. IEEE Trans. Image Proc., 16(4): 919-931.

Minhas, R., A. Mohammed and Q.M. Wu, 2011. Shape from focus using fast discrete curvelet transform. Pattern Rec., 44(4): 839-853.

Selesnick, I.W. and K.Y. Li, 2003. Video Denoising using 2D and 3D Dual Tree Complex Wavelet Transforms in Proceedings of Wavelets: Applications in Signal and Image Processing X., San Diego, CA, USA.

Tao, G., 2002, A modified motion estimation algorithm--pseudo-diamond search method. J. Image Grap., 7(11): 1181-1186.

Xiao, H.Z., J.W. Yan and X.B. Qu, 2008. A novel video denoising method with 3D context model Based on Sur face let transform. Chinese J. Elect., 36(7): 1460-1464.

Yang, J.Y., W. Yao, W.L. Xu and Q.H. Dai, 2009. Image and video denoising using adaptive dual-tree discrete wavelet packets. IEEE Trans. Circuits Syst. Video Technol., 19(5): 642-655.

Ying, L.X., L. Demanet and E. Candes, 2005. 3D Discrete Curvelet Transform. Applied and Computational Mathematics, In Proceedings of the International Society for Optical Engineering SPIE 05, Caltech, Pasadena, CA.