

Quality Enhancement of Video Denoising Algorithm with Combined Effect of Different Views

Snehal Ambulkar¹, Priti Golar²

¹Department of Computer Engineering, St. Vincent Pallotti College of Engineering & Technology, R.T.M. Nagpur University (Maharashtra), India.)

²Department of Information Technology, St. Vincent Pallotti College of Engineering & Technology, R.T.M Nagpur University (Maharashtra), India)

Abstract: An unwanted factor in digital images and videos called as Noise is responsible for hiding the details and destroying image information. Hence denoising has great importance in the detail restoration and to improve the quality measures. A simple adaptive median filter (SAMF) is one of the most promising algorithm capable of denoising the videos to better extent, but incapable to use all the neighboring pixel information while noise processing and hence quality of the denoised video can be further improved if all neighboring pixel information is considered. To make this possible the SAMF is combined with the multi-view denoising and then a fusion strategy is used to generate enhanced denoised video sequence. The performance of the combined system is measured through the image quality matrix called as Peak Signal to Noise Ratio (PSNR) at different noise contamination level.

Keywords: simple adaptive median filter (SAMF), Multiview Fusion (MVF), Adaptive Filter, Denoising, Impulse noise, Median filter, Salt-and-pepper, Switching Median Filter.

I. Introduction

An important part of real world is achieving incredible developments in computer vision and image understanding technologies. Today images and videos collected from different sources are seems to be a great source of information for day to day applications like image acquisition and display in more sophisticated applications such as video surveillance, defense, medical diagnosis, autonomous vehicles etc.[1]. Inputs to such systems are from digital cameras, scanners, mobile cameras, webcams, etc. The quality of image depends on many factors like camera lenses and sensors, acquisition environment, channel quality while transmission, analogue to digital conversion and vice a versa, etc. Due to all these factors images and image sequences are often get corrupted by noise which leads to the distortion in image pixel values hiding the quality information.

For proper information gathering it is necessary to make image and image sequences free from noise contamination. It is achieved through effective video denoising algorithms that can remove or reduce the noise to acceptable level in various application environments such as network visual communications. They supply video signals that have better perceptual quality and help to improve the performance of the succeeding processes such as resizing, compression, segmentation, and object detection and recognition [2].

To solve these problems lots of denoising algorithms were proposed in the past era. The performance of any denoising algorithm depends on its effectiveness in separating noise from original image data. The better its differentiating capability, the better would be the performance of the algorithm. Since the literature lack in algorithm that can achieve 100% accuracy there is further scope to improve the performance of algorithm to preserve image details.

II. Literature Survey

The past years comes with the lots of image denoising algorithms in the search of better result generating algorithm which not only improves the image quality but also preserves the minute image details and helps to enhance the quality of digital image. These existing algorithms are generally classified depending upon the domain of denoising i.e. spatial or transform domain [3], subdivided into linear or non-linear filters. Instead of that for video denoising these algorithms are further categorized in three categories depending upon the scope of information gathering [4]. First category involves the algorithms which denoise the video signals on frame-by-frame basis. Each frame is considered as an individual 2D image and not having any correlation with any other frame in the video. The correlation of both intra- and inter-frame information was exploited in the second category of video denoising algorithms. Due to significant motion in video every frame has some similar information with little difference, which founds greater impact on video denoising. The third category of denoising algorithms treats video sequences as 3D volumes.

The most popular and conventional filter in spatial domain for Salt-and-pepper noise is median filter. This filter is simple and has capability of edge preservation [5]. Since this filter process both "Noisy"

and “Noise Free” pixels the resultant image results in blurring, distortion and loss of lines and corners. This is the reason for variation and improvements in median filters.

Adaptive median filter [3] with flexible filter window size was introduced to minimize the blurring effect. This filter puts the limitation on filtering window size. As it reaches to maximum window size it outputs a median value which does not provide a guarantee of whether the value is impulse or not.

For further reducing the blurring effect weight allotment concept was introduced. In Weighted median filter [6], Center weighted median filter (CWMF) [5], Recursive weighted median filter [7] some specific pixels in filtering window were assigned with more weight.

In [8][9], A new kind of adaptive weighted median filters was proposed which uses the block uniformity as standard to detect the impulse noise in the image and an efficient scalar quantization (SQ) method to partition the observation vector space which obtain the optimal weight for each block. A modified adaptive center weighted median filter (MACWM) was proposed in [10] which obtain the adjustable central weight by partitioning the observation vector space using fuzzy clustering technique. In [11], a new adaptive center-weighted hybrid mean and median filter was formulated and used within a novel optimal-size windowing framework to reduce the effects of sensor noise. A Stuck-pixel filter is another concept introduced to remove stuck-pixel noise occurred due to long-exposure of images.

The median value calculations and replacement of pixels in conventional median filtering are independent of whether pixels are noisy or not. This makes the resultant image with detail distortion. Switching median filtering (SMF) [12] overcomes this drawback by implementing decision making process in the filtering framework. SMF makes use of median filter for optimal median calculation. Since median filter uses fixed window filter size it is quite difficult to find correct value of median which is close to real pixel value in original image. Hence to overcome the drawbacks Simple adaptive median filter was proposed in the literature which is described in subsequent section.

III. Simple Adaptive Median Filter

Simple adaptive median filter (SAMF) [13] is hybrid filter developed by combining the features of Adaptive median filter and switching median filter. Adaptive median filter provides flexibility of adapting variable size of filtering window according to local noise level in an image. Whereas switching median filter speeds up the process by ignoring the noise free pixels while processing. The combination of these two features makes Simple adaptive median filter more powerful.

This filter consists of two basic steps:

- (a) Finding Noisy Pixels
- (b) Noise Removal

For impulse noise detection the two intensity levels i.e minimum and maximum are considered to be noisy whereas others are considered as noisy free. Let L be the number of intensity levels present in an image then 0 and $L - 1$ are the minimum and maximum intensity values, respectively, considered as noise contaminated pixels which may create Black and white dots in an image. As per intensity values a mask $\alpha(x, y)$ is prepared to detect whether the pixel is noisy or noise free using conditional equation:

$$\alpha(x, y) = \begin{cases} 1 & : f(x, y) = 0 \\ 1 & : f(x, y) = L - 1 \quad \dots (1) \\ 0 & : \text{Otherwise} \end{cases}$$

Where 1 indicates the “Noisy Pixel” and 0 indicates “Noise Free Pixel”.

Let $f(x, y)$ be the original image pixel value. The output image pixel value $g(x, y)$ is either having value $f(x, y)$ or $m(x, y)$ i.e. median value calculated. This is because of mask $\alpha(x, y)$ which consists of only two values 0 and 1. To determine $m(x, y)$ following set of rules are defined.

Algorithm:

For each noisy pixel location (x, y) with $\alpha(x, y) = 1$ do:

1. Initialize the filter window size $W_{M \times N}$ with eqⁿ

$$W = W_M = W_N = 2R_{\min} + 1 \quad \dots (2)$$

Where, R_{\min} is a positive integer value.

$$R_{\min} = \frac{1}{2} \sqrt{\frac{7}{(1-\eta)}} \quad \dots (3)$$

Estimation of rough noise density (η) takes a ratio of the noisy pixel to the total number of pixels in image. Value of η is always between 0 and 1 i.e. $0 \leq \eta \leq 1$.

$$\eta = \frac{K}{(MN)} \dots (4)$$

2. Compute the number of noisy pixel (K) for region defined by window size initialized in step 1. K is given as:

$$K = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \alpha(x,y) \dots (5)$$

3. If number of "Noise free pixels" (MN-K) is less than eight pixels, Then increase window size as $W = W + 2$, and return to step 2.
4. Calculate the median value of $m(x,y)$ depending upon the noise free pixels present in considered window size using adaptive median filter. This means the median intensity value of noise free pixels in given window size.
5. Update the value of $g(x,y)$ using either

$$g(x,y) = [1 - \alpha(x,y)]f(x,y) + \alpha(x,y)m(x,y)$$

OR

$$g(x,y) = \begin{cases} f(x,y) & : \alpha(x,y) = 0 \\ m(x,y) & : \text{otherwise} \end{cases} \dots (6)$$

Here calculation of the median value $m(x,y)$ depends on the noise free pixels and hence to get more appropriate value the minimum number of noise free pixel is considered as eight. Whereas the minimum filter size is defined as 5×5 and maximum filter size is defined as 21×21 .

IV. Plan Of Work

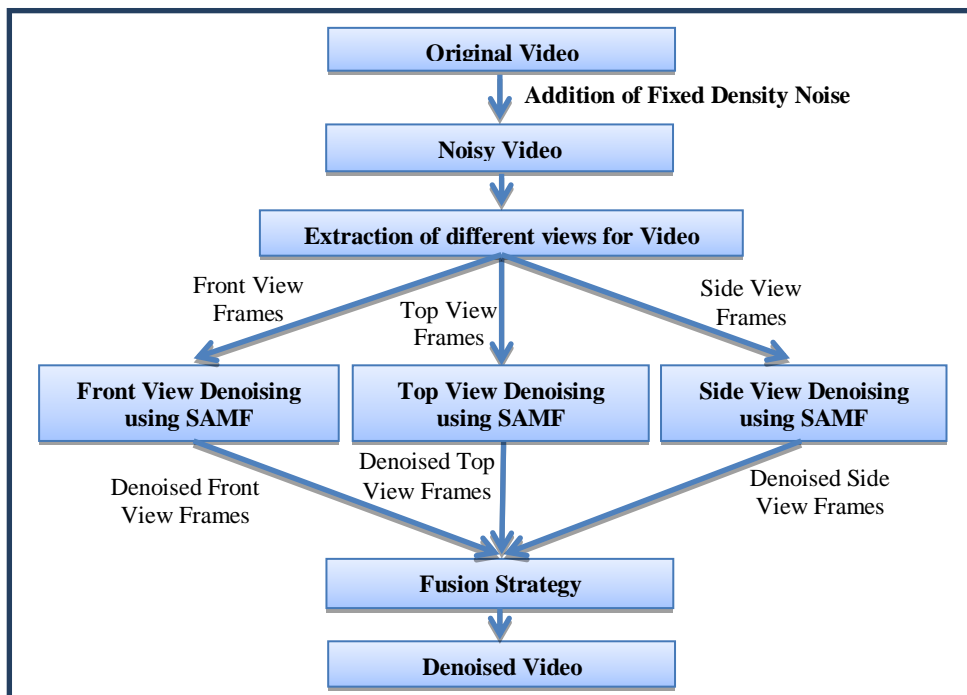


Fig.1: Concept Structure

When video contains significant motion it is very challenging and unproductive to implement simple adaptive median filter (SAMF) directly. Even if direct 3D filtering is implemented it creates complications and inefficiency. Since SAMF adopts frame by frame denoising, it is incapable to involve the information of all neighbouring pixel which results in distortion of resultant pixel value. Other filters with frame by frame strategy and 2D denoising approach also face the same drawback. So it is essential to take on a strategy which makes better consumption of neighbouring pixel information from all the three sides. To enhance the performance of SAMF, we will reshape video by making use of 2D information as 3D volume data. When these three reshaped videos are separately denoised followed by fusion may generate effect having advantages of 2D approaches using the information all the neighboring pixels from all the three dimensions for each pixel. Hence we are

going to use A Simple Adaptive Median Filter (SAMF) with multi-view fusion (MVF) approach to improve the performance of SAMF. “Fig.1” shows the concept structure of proposed system.

A digital video can be expressed as 3-D function, $f(u, v, t)$, which is discrete in both space and time, where u and v as the horizontal and vertical spatial indices, respectively, with t as time index. Normally video is played along the time axis in which at every instant i.e. at $t=t_0$ the image $f(u, v, t_0)$ represents 2D front-view of the frame. When we consider a video signal as 3-D volume data, then it can also be viewed from the side or the top. Hence the two other ways are to play video along horizontal axis and vertical axis as top view and side view. A 2D top view video is an image sequence of top-view images $f(u_i, v, t)$, for different values of i , likewise a sequence of side-view images $f(u, v_i, t)$, for different values of i , generates 2D side-view video. An example is shown in “Fig. 2 (d), (e)”, where the rarely observed side- and top-view images of a single frame from demo video are given.

Let f be an original noise-free video signal that is contaminated by additive independent fixed density salt-and-pepper noise resulting in a noisy signal $f' = f + n$. Then the three different noise contaminated views are denoised with SAMF i.e. $D(.)$ giving resultant noise free video $g_{FV}(u, v, t_0)$, $g_{TV}(u_0, v, t)$, $g_{SV}(u, v_0, t)$. Now we have to use fusion operators $F(.)$ on the different views to obtain the resultant denoised video $g(u, v, t_i)$ given as,

$$g(u, v, t_i) = F(g_{FV}(u, v, t_0), g_{TV}(u_0, v, t), g_{SV}(u, v_0, t)) \\ = F(D(f_{FV}(u, v, t_0)), D(f_{TV}(u_0, v, t)), D(f_{SV}(u, v_0, t)))$$

$D(.)$ and $F(.)$ are predetermined in this case i.e SAMF and Average wavelet coefficient fusion method respectively. Pick-signal-to-noise ratio (PSNR) [3] is one of the most popular quality measurement factor to make qualitative analysis in video processing.

V. Result Analysis

To demonstrate the proposed strategy some publicly available video sequences are used, which includes “Carphone”, “Forman”, “Traffic”, “Train”, “Rhios”, and “Warnsigns”. The sizes of the sequences are $176 \times 140 \times 100$, $176 \times 144 \times 100$, $160 \times 120 \times 90$, $360 \times 240 \times 90$, $320 \times 240 \times 105$ and $360 \times 180 \times 120$ respectively. An independent Salt-&-pepper noise is introduced to the original video sequences, where the noise density varies between 5% and 90%. All sequences are having RGB color format.

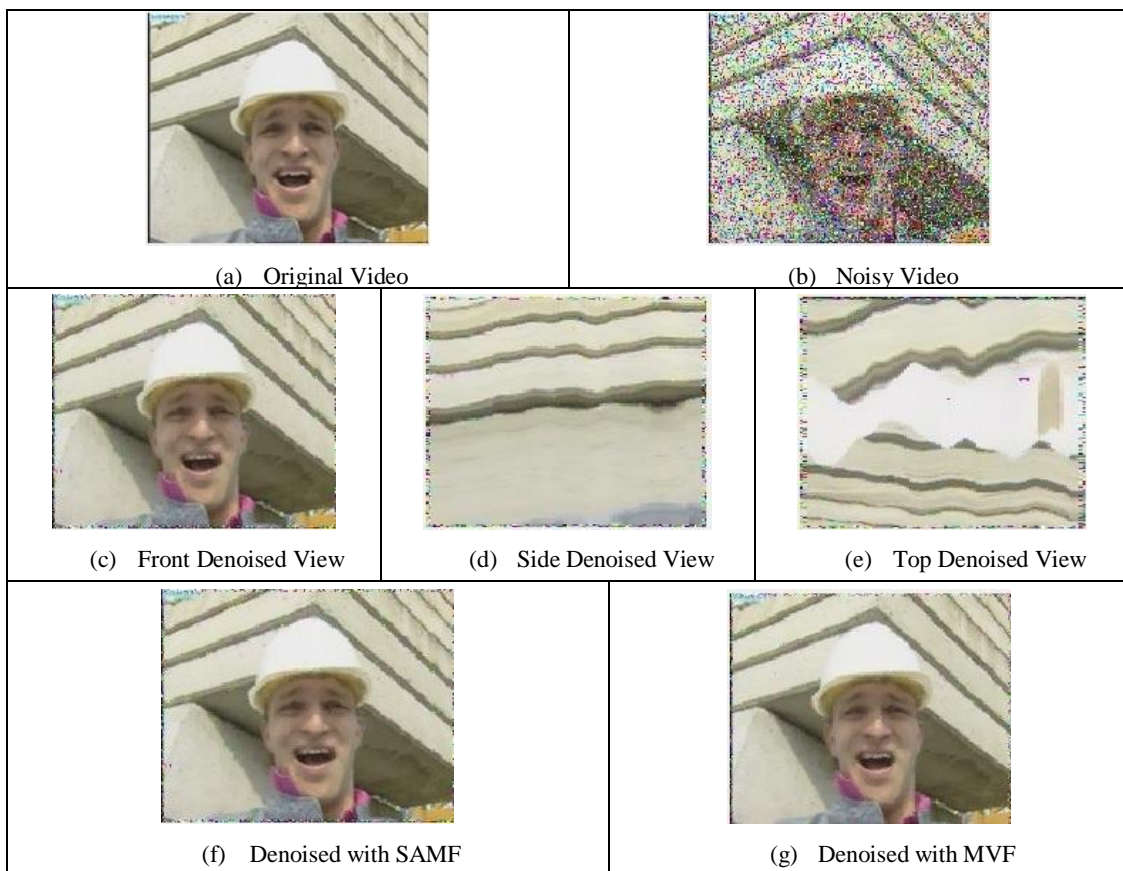


Figure 2: Visual analysis of Frame 25 from test video sequence “Forman” as a function of noise density 30%

Table 1:PSNR comparisons of Simple Adaptive Median Filter with and without MVF

Noise Density (%)		Forman		Traffic	
10%	SAMF	59.49	0.06	58.84	0.10
	with MVF	59.55		58.94	
20%	SAMF	56.48	0.08	55.84	0.17
	with MVF	56.56		56.01	
30%	SAMF	54.71	0.13	54.09	0.21
	with MVF	54.84		54.3	
40%	SAMF	53.47	0.15	52.84	0.24
	with MVF	53.62		53.08	
50%	SAMF	52.51	0.19	51.88	0.29
	with MVF	52.7		52.17	
60%	SAMF	51.72	0.28	51.09	0.38
	with MVF	52		51.47	
70%	SAMF	51.04	0.41	50.37	0.52
	with MVF	51.45		50.89	
80%	SAMF	50.32	0.71	49.59	0.82
	with MVF	51.03		50.41	
90%	SAMF	49.46	1.24	48.61	1.38
	with MVF	50.7		49.99	
Noise Density (%)		Train		Warnsigns	
10%	SAMF	64.5	0.08	63.54	0.06
	with MVF	64.58		63.6	
20%	SAMF	61.5	0.10	60.54	0.09
	with MVF	61.6		60.63	
30%	SAMF	59.74	0.13	58.73	0.17
	with MVF	59.87		58.9	
40%	SAMF	58.49	0.17	57.55	0.13
	with MVF	58.66		57.68	
50%	SAMF	57.54	0.21	56.6	0.18
	with MVF	57.75		56.78	
60%	SAMF	56.75	0.3	55.81	0.26
	with MVF	57.05		56.07	
70%	SAMF	56.05	0.45	55.1	0.39
	with MVF	56.5		55.49	
80%	SAMF	55.33	0.76	54.35	0.67
	with MVF	56.09		55.02	
90%	SAMF	54.45	1.32	53.42	1.17
	with MVF	55.77		54.59	
Noise Density (%)		Rhios		Carphones	
10%	SAMF	64.26	0.12	59.21	0.09
	with MVF	64.38		59.3	
20%	SAMF	61.26	0.15	56.21	0.15
	with MVF	61.41		56.36	
30%	SAMF	59.51	0.17	54.45	0.19
	with MVF	59.68		54.64	
40%	SAMF	58.27	0.19	53.2	0.23
	with MVF	58.46		53.43	
50%	SAMF	57.3	0.25	52.24	0.30
	with MVF	57.55		52.54	
60%	SAMF	56.51	0.32	51.47	0.37
	with MVF	56.83		51.84	
70%	SAMF	55.8	0.46	50.8	0.52
	with MVF	56.26		51.32	
80%	SAMF	55.05	0.74	50.13	0.84
	with MVF	55.79		50.97	
90%	SAMF	54.11	1.27	49.3	1.43
	with MVF	55.38		50.73	

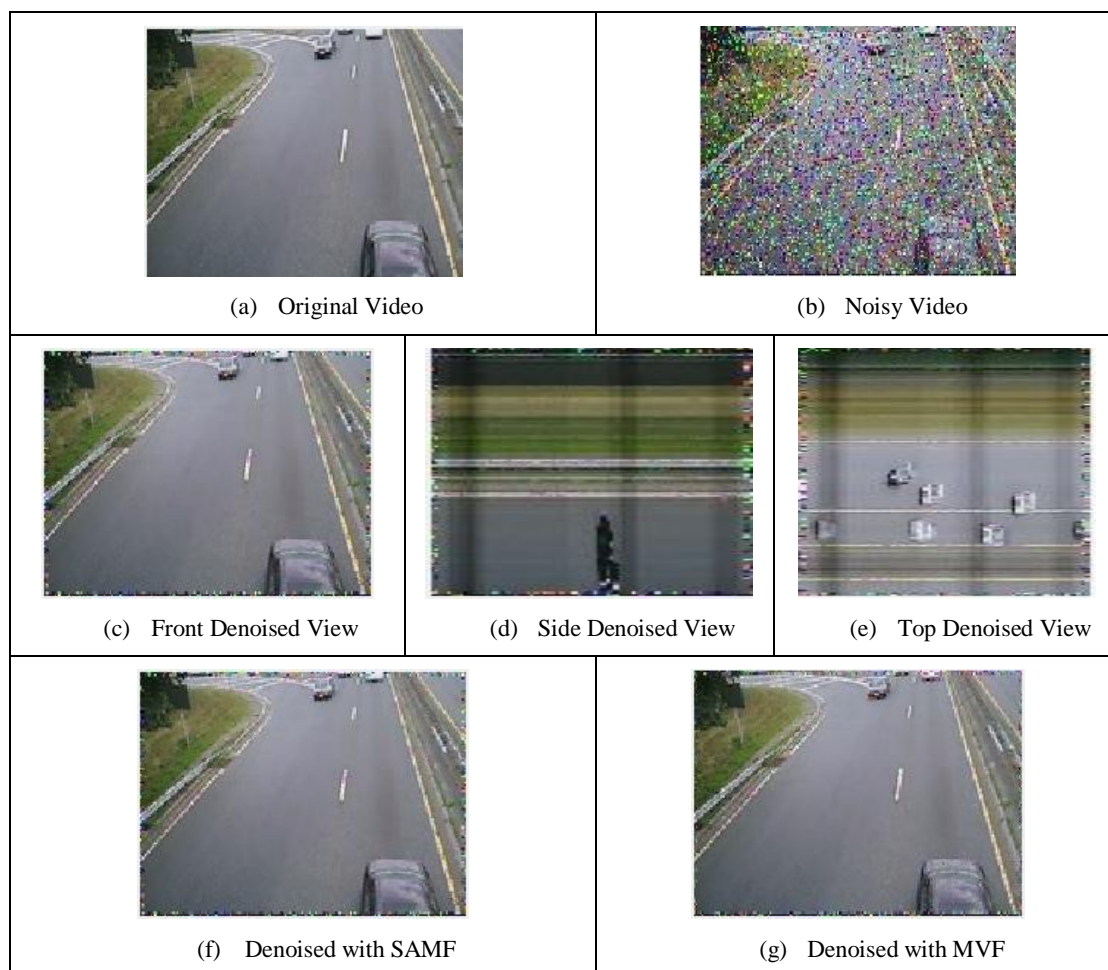


Figure 3: Visual analysis of Frame 20 from test video sequence “Traffic” as a function of noise density 20%

A quality matrix, namely PSNR, is employed to evaluate the quality of denoised video quantitatively. PSNR is the most widely used method in the literature. Since SAMF is a spatial domain filter and basically works on pixel value it is quite interesting to observe the improvement over it. Comparison of the PSNR values generated for the SAMF with and without MVF is mentioned in the “Table 1”. PSNR generation involves the individual frame value of all the three RGB components.

As per the result generated, MVF with SAMF approach consistently leads to performance gain over normal denoising using SAMF. By observing the results it is found that as the noise density in video get increases a substantial gain in PSNR also get increase. At low noise density improvement in performance of algorithm is very low but at high noise levels improvement is upto 1.43 dB in terms of PSNR over denoising by SAMF without MVF.

The time complexity of the technique totality depends on the complexity of base denoising algorithm (SAMF). Fusion of video frames at last level is not time consuming if a better fusion algorithm is selected. “Fig.2” and “Fig.3” provides visual comparisons of the denoising results for frame number 25 extracted from the “Forman” sequence at noise level 30% and frame no 20 from the “Traffic” sequence at noise level 20% correspondingly. As we observed the changes it is found that at the high noise density level the noise reduction is more with some good extent.

To demonstrate the performance improvement “fig. 4” shows graphical analysis for four video sequences. It clearly identifies the PSNR gain in SAMF due to MVF approach. For the “Forman” Sequence highest gain of 1.24 DB is detected at 90% noise density value. Whereas the maximum gain detected is of 1.48 DB for video sequence “Carphones” at 90%.

VI. Conclusion And Future Scope

For video denoising a basic 2D Simple adaptive Median Filter is enhanced using fusion of different view denoising. This experiment demonstrates significant and reliable improvement over existing Simple

adaptive Median Filter for video denoising. The MVF method is simple, easy-to-use, and computationally effective.

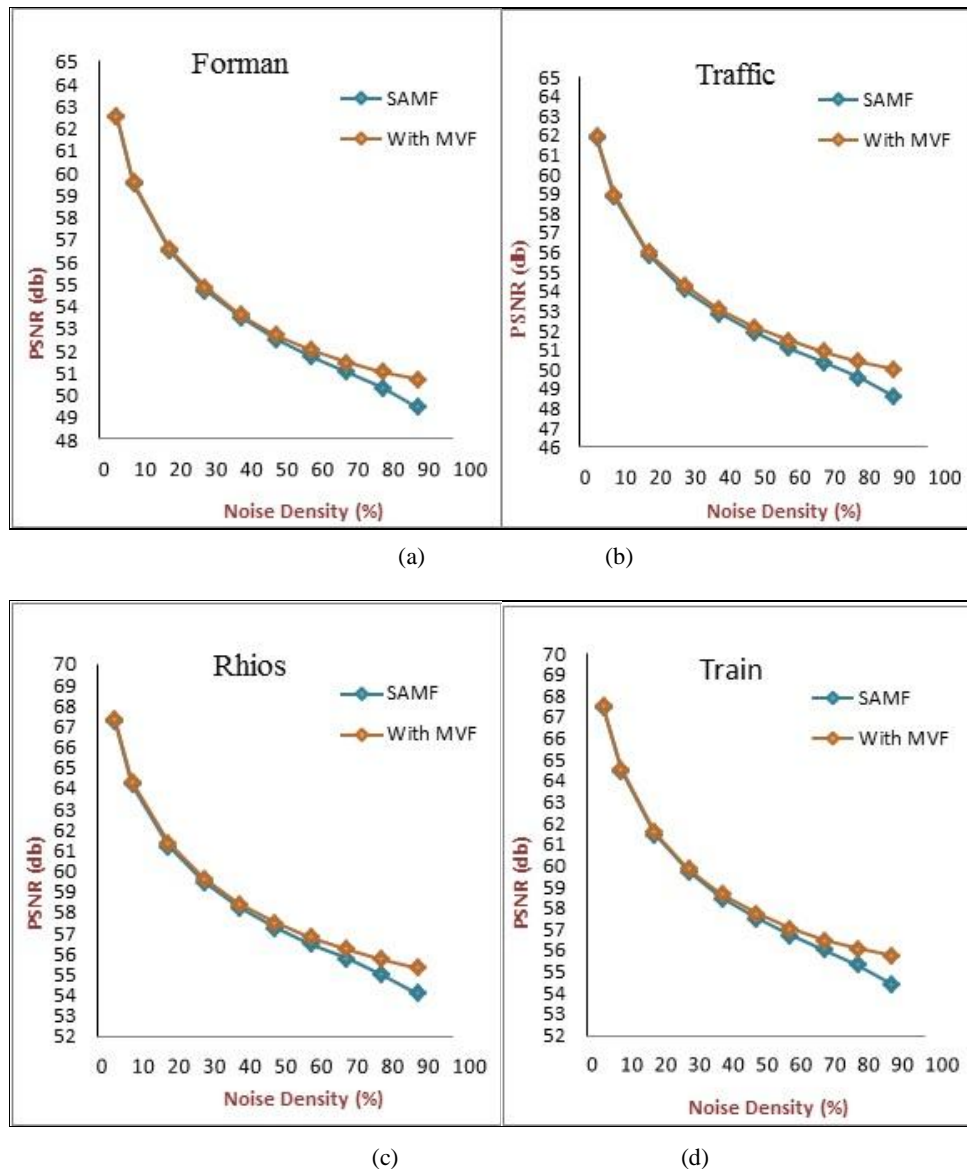


Figure 4: Graphical analysis of test video sequence as a function of noise density and PSNR (a): Forman Video Sequence, 176×144×100 (b): Traffic Video Sequence, 160×120×90 (c): Rhios Video Sequence, 320×240×105(d): Train Video Sequence, 360×240×90

The complexity of the whole method depends on the base denoising algorithm i.e. SAMF, and not on procedure we have applied. Though applying multi-view fusion to all the existing video is possible, it would not generate the better results with 3D approached algorithms. 3D algorithms works with 3D patches in which all the dependencies between neighbouring pixels from all directions are already considered, and thus applying them from different views may generate similar results that would not be significant. The experimental work accumulates all video frames involved in the denoising and fusion process which needs memory storage. In case of long videos, segmentation of video along temporal direction and then denoising of each segment independently is a better memory management. The memory requirement can be controlled by adjusting the length of the segments.

In future, better denoising results may be obtained by incorporating more advanced denoising algorithm than that of SAMF or by improving the fusion method. The current implementation fuses the denoising results using same base denoiser applied along three views. One can further attempt to use finite number of denoising algorithms to produce denoising results which will further used for fusion. But for doing so

one has to take care of the correlation between base denoising algorithms and then there will be need of the proper selection of fusion policy.

References

- [1] Vladimir Zlokolica, Aleksandra Pizurica, and Wilfried Philips, "Video denoising using multiple class averaging with Multiresolution", University of Ghent, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium, VLBV2003, LNCS 2849, PP. 172-179, 2003, Springer verlog berlin Heidelberg 2003.
- [2] Bovik, Alan C.: Handbook of Image and Video Processing (Communications, Networking and Multimedia). Academic Press, Inc., Orlando, FL, USA, (2005).
- [3] Rafael C. Gonzalez, Richard E. Woods, Book on "Digital Image Processing", third edition, Pearson Education, pp.156-157,326,332-335,369©2008.
- [4] Kai Zeng, and Zhou Wang; Polyview Fusion: A Strategy to Enhance Video-Denoising Algorithms; IEEE Transactions On Image Processing, Vol. 21, No. 4, APRIL (2012).
- [5] Sung-JeaKo, and YoogHoon Lee, "Center weighted median filters and their application to image enhancement", IEEE Trans. Circuits and System, vol. 38, no. 9, pp. 984-993, September 1991.
- [6] R. Yang, L. Yin, M. Gabbouj, J. Astola, and Y. Neuvo, "optimal weighted median filters under structural constraints", IEEE Trans. Signal Processing, vol. 43, pp. 591-604, Mar. 1995.
- [7] Arce G, Paredes J, "Recursive Weighted Median Filters Admitting Negative Weights and Their Optimization", IEEE Trans. on Signal Proc., Vol. 48, 2000, pp 768-779.
- [8] WANG Chang-you, LI Lin-lin, YANG Fu-ping, GONG Hui, "A new kind of adaptive weighted median filter algorithm", International Conference on Computer Application and System Modeling (ICCASM 2010), volume 11-667,2010, 978-1-4244-7237-6/\$26.00 C 2010 IEEE.
- [9] T.-C. Lin, P.-T. Yu, "A new adaptive center weighted median filter for suppressing impulsive noise in images", Information Sciences 177 (2007) 1073–1087.
- [10] HadiSadoghiYazdi, FaranakHomayouni, "Impulse Noise Suppression Of Image Using Adaptive median Filter", International journal of signal processing, image processing and pattern Recognition, vol. 3, No. 3, September 2010.
- [11] Tamer Rabie, "Adaptive hybrid mean and median filtering of high-ISO long-exposure sensor noise for digital photography", Journal of Electronic Imaging, volume 13(2), 264–277 (April 2004).
- [12] KavitaTewari, Manorama V. Tiwari, "Efficient Removal of Impulse Noise in Digital Images", International Journal of Scientific and Research Publications, Volume 2, Issue 10, October 2012.
- [13] Haidi Ibrahim, Nicholas SiaPik Kong, and Theam Foo Ng, "Simple adaptive median filter for the removal of impulse noise from highly corrupted images", IEEE Trans. Consumer Electronics, vol. 54, no. 4, pp. 1920-1927, November 2008.