Video Noise Reduction Using H.264 Multi-Frame Trajectory

[1]*Jong-Tzy Wang (王仲資)*, [2]*Kai-Wen Liang (梁凱雯)*, [2]*Ni-Na Chen (陳旎娜)*, *and* [2]*Pao-Chi Chang (張寶基)*

[1]Department of Electronic Engineering, Jinwen University of Science and Technology, Shindian, Taiwan 231
[2]Department of Communication Engineering, National Central University, Jhongli, Taiwan 320
E-mail: jtwang@jwit.edu.tw, { nnchen, kwliang, pcchang }@vaplab.ee.ncu.edu.tw

ABSTRACT

In this paper, we propose a spatio-temporal filtering scheme that utilizes motion compensation with multi-reference frames in H.264 to remove noise. It utilizes the variation of inter mode distributions to detect noise and determine the parameters for spatial filter. In the time domain, a reference pixel is selected from multi-reference frames according to the motion trajectory and MSE criterion. The proposed adaptive spatial filter and non-linear temporal filter can effectively remove Gaussian and impulse noises to improve video quality up to 0.83dB and 8.17dB in PSNR respectively. Meanwhile, it can also boost the compression efficiency by reducing the bitrate up to 84.85% in our experiments.

1. INTRODUCTION

With the development of information technology, the applications of digital images and videos increase rapidly. The demand for noise reduction to provide images and videos with high quality is growing because of increasing use of video systems in consumer, medical, and communication applications.

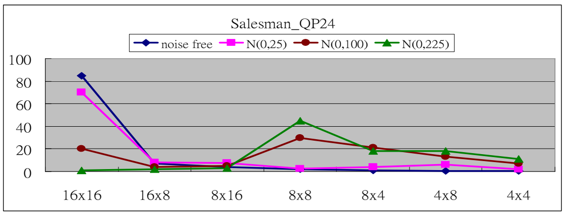
Noise reduction techniques were developed starting from either spatial filtering or temporal filtering to the current joint spatio-temporal filtering. The spatial filtering methods proposed in [1] improved the blurring problem in images but could not avoid the ghost effect without considering motion compensation in video coding. For temporal filtering, [2],[3],[4] adopted block searching method to find matched blocks between frames and then utilized the correlation between signals to perform linear filtering. Boyce method [2] executed the block matching from consecutive frames and applied temporal averaging to reduce noise. This method needs to estimate the variance of noise in advance. In addition, block matching accuracy reduces when SNR is low. Spatio-temporal filters [3],[5],[6],[7] consider the correlation of signal in both spatial and time domains such that more information could be used to reduce noise. In [7], the spatial filter using the weighted averaging pixels in single frame is adopted first, and the temporal filter proposed in [2] is then used to further reduce the remaining noise to get better performance. In our proposed algorithm, we utilize the motion information produced from motion estimation and compensation in H.264 [8] to perform block matching. The proposed method cannot only reduce noise but also improve compression efficiency at the same time.

The paper is organized as follows. Section 2 provides a brief description of the noise detection with H.264 mode variations. Section 3 introduces the proposed approach. Experimental results are presented in Section 4. Section 5 concludes this paper.

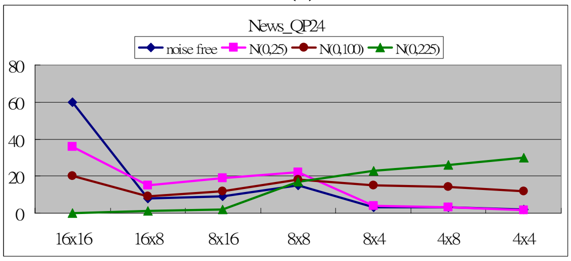
2. NOISE DETECTION USING VARIATIONS OF H.264 INTER MODE DISTRIBUTIONS

H.264 uses variable block size to perform motion estimation. In general, large block size modes, such as 16x16, 16x8, and 8x16 are used in smooth areas. In contrary, areas with more textures are predicted with small block size modes. When the light condition is poor or the camera sensor has defects, Gaussian and impulse noises will appear in the captured images. The existence of noise makes the image more complicated and the number of small block size modes increases substantially. Figure 1 shows the prediction modes versus their occurrence probability with different noise levels in two video sequences. As shown in Figure 1, the probability of small blocks increases with the noise level. We can estimate the existence and level of noise by the variation of mode distributions. Figure 2 shows the statistical results of mode distributions. We set four thresholds to determine the level of noise. The seven block sizes are classified into two modes, i.e., Mode 1 and Mode 2. Mode 1 represents the occurrence of large block sizes, i.e., 16x16, 16x8, and 8x16 while Mode 2 represents the occurrence of small block sizes, i.e., 8x8, 8x4, 4x8, and 4x4. We determine the level of noise by three conditions as follows:

1. Low noise: $T1 \le Mode1$ and $Mode2 \le T3$
2. Median noise: $T2 < Mode1 < T1$ and $T3 < Mode2 < T4$
3. High noise: $Mode1 \le T2$ and $T4 \le Mode2$



(a)



(b)

Figure 1. mode distribution with QP24 (a) Salesman (b) News

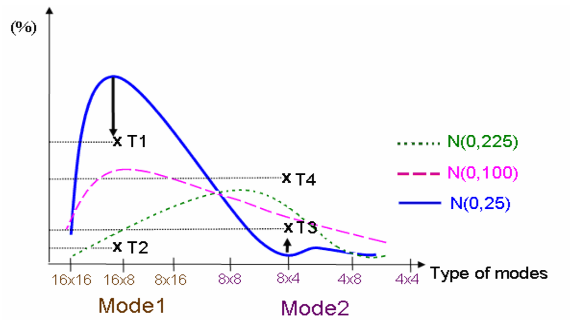


Figure 2. Variation of mode distributions

3. THE PROPOSED NOISE REDUCTION ALGORITHM

The proposed algorithm uses multi-frame tracking technique to reduce noise in video. This algorithm processes noise reduction by utilizing the characteristics of H.264 compression standard. The block diagram is shown in Figure 3. The input video is first processed by the noise detection unit based on the variation of mode distributions. The noisy video XN is then sent into the adaptive LUM spatial filter. The non-linear temporal filter is finally applied to the filtered sequence y^* to eliminate the blurring phenomenon produced from spatial filtering.

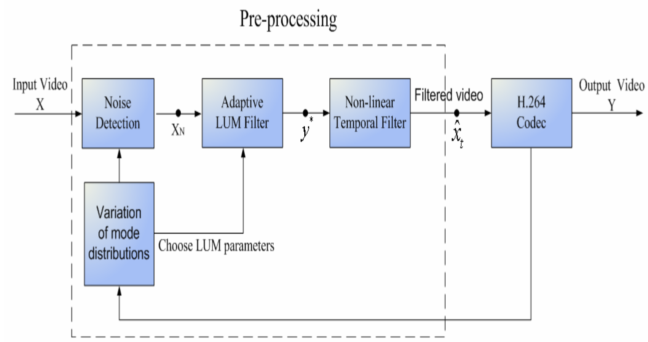


Figure 3. Block diagram of the proposed spatio-temporal filter

3.1. Adaptive spatial filter

Most spatial filters use masks to process images pixel by pixel. Among them, the median filter is popularly used because it is able to preserve edges and remove dispersed noise. However, the median filter will cause serious distortion when the noise is distributed in cluster. Hardie and Boncelet [9] proposed LUM filter with more parameters based on a special ordering. By properly using two predefined adaptive parameters, i.e., smooth parameter k and sharpening parameter l, the LUM filter has noise reduction, edge detection, and edge enhancement capabilities. Therefore, LUM filter can be designed as a low pass smoother, high pass sharpener, or noise reduction filter in different applications. The principle of the LUM filter is to reorder the input signal in the filter by its strength and uses a non-linear function to determine the output signal.

We propose effective LUM parameter sets for different noise levels based on the statistical results from various video sequences and different noise levels as shown in Table Ⅰ.

Table Ⅰ LUM PARAMETERS (a)Gaussian Noise (b)Impulse Noise

(a)

	k	l	m
Noise ≤ N(0,25)	3	5	1
N(0, 25) < Noise ≤ N(0,100)	5	13	2
N(0,100) < Noise	9	12	2

(b)

	k	l	m
Noise ≤ 1%	3	5	1
1% < Noise ≤ 10%	5	13	2
10% < Noise	9	12	2

We use the distribution of Mode 1 and Mode 2 in the previous frame to determine the noise level and select the corresponding parameters for LUM.

3.2 Non-linear temporal filter

The noise reduction performed by spatial filters usually blurs the image. A temporal filter may reduce the blurring artifact when it is properly arranged. Motion trajectory that is a key factor connecting two consecutive frames can be used in designing a temporal filter. The motion trajectory can be calculated by motion estimation and motion compensation in H.264. We utilize the motion trajectory to design the proposed non-linear temporal filter. Two assumptions are applied in the design as follows. (1) Signal has higher correlation than noise in the time domain along the motion trajectory because the occurrences of noise are regarded as random. (2) The reference frame that is closer to the current frame has higher correlation.

A non-linear temporal filter in the proposed algorithm reduces blurring artifact in the spatial domain operates as the following steps:

Step1: Select reference pixels for each pixel from motion trajectory, $\{x_t, x_{t-1}, x_{t-2}\ldots, x_{ref\le 5}\}$, as shown in Figure 4. The reference frames that are too far are discarded when the accumulated motion vector $mv = \sum_{i=0}^{ref} |mv_i|$ is large than a threshold.

Step2: Calculate the MSE of pixels between any two frames including the current frame and all reference frames.

$$D = \left\{ d_k = sum\_of\_MSE = \frac{1}{(1+ref)} \sum_{i=0}^{ref} (x_{t-i} - x_{t-k})^2 \right.$$

$$k = 0,1,2,\ldots ref \le 5; i \ne k$$

$$x_{t-k} : pixel \ \ at \ \ frame \ \ f_{t-k} \ \ along \ \ motion \ \ trajectory$$

Step3: Choose the pixel that produces the minimum D to be the reconstructed value denoted as $\hat{x}_t$.

Step4: When the same D are produced by different reference frames, we choose the pixel by the order as $x_t > x_{t-1} > x_{t-2} \ldots > x_{t-(ref\le 5)}$

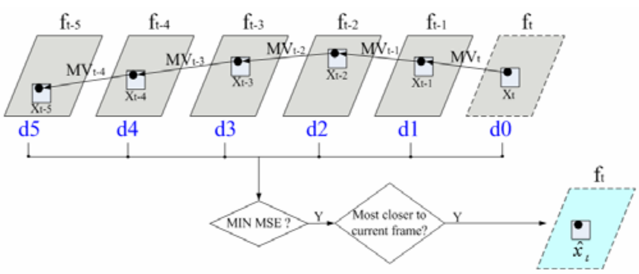


Figure 4. Noise reduction in temporal domain

When Intra mode is used, there will be no motion information to use. Therefore, neighboring points in the current and previous frame are used to reduce noise in the temporal filter as follows.

$$\hat{x}_t(x,y) = \frac{1}{8}(x_t(x,y) + x_t(x-1,y) + x_t(x,y-1) + x_t(x-1,y-1)$$
$$+ x_{t-1}(x,y) + x_{t-1}(x-1,y) + x_{t-1}(x,y-1) + x_{t-1}(x-1,y-1))$$

4. EXPERIMENTAL RESULTS

Experiments with several sets of moving sequences are performed to show the effectiveness of the proposed noise reduction algorithm, in which additive zero-mean white Gaussian noise and impulse noise are employed. The test sequences are Salesman, News, and Stefan in CIF format with 150 frames in each. The proposed algorithm is applied to all the video sequences, in which noise can be reduced gradually by utilizing the previously denoised reference frames.

As shown in Figure 5, our PSNR calculation compares the noise reduced frame and noise free frame.

$$PSNR_{IY} = 10\log\frac{255^2}{MSE} \quad MSE = \frac{\sum_{j=0}^{H-1}\sum_{i=0}^{W-1}(x(i,j) - y(i,j))^2}{H \times W}$$

where H and W are the height and width of each frame.

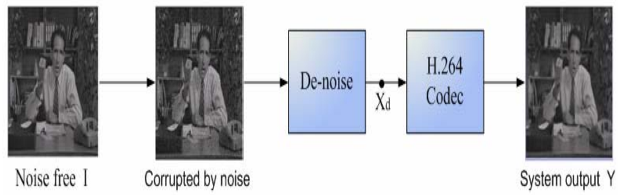
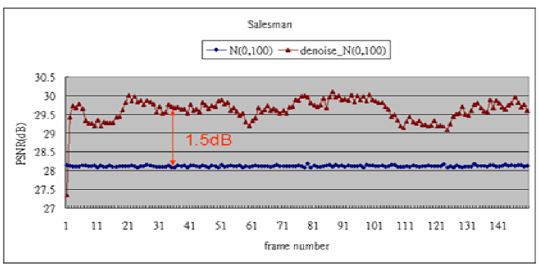


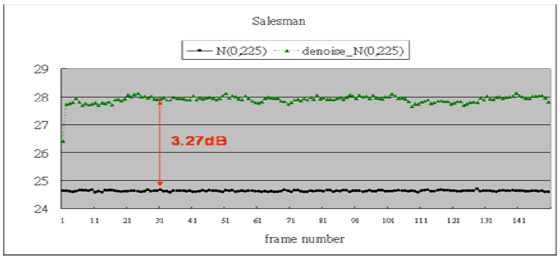
Figure 5. PSNR comparison

Figure 6 and Figure 7 show the PSNR between the original video and the video after denoising.

The proposed denoising algorithm can increase 1.5dB in PSNR for Salesman and 2.4dB for News. In Figure 6(b), for the noise N(0, 225) that the variance is large, the PSNR increment is up to 3.27dB for Salesman. The PSNR performance fluctuates with variation of motions in video sequence because of the non-linear characteristic of the filter. The experimental results of impulse noise are shown in Figure 7. It can be seen that the proposed algorithm performs even better when deals with impulse noise.



N(0,100)



N(0,225)
Figure 6. PSNR for Gaussian noise

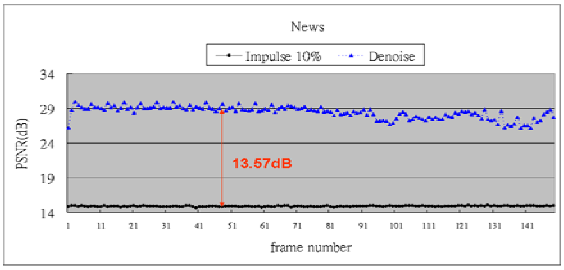
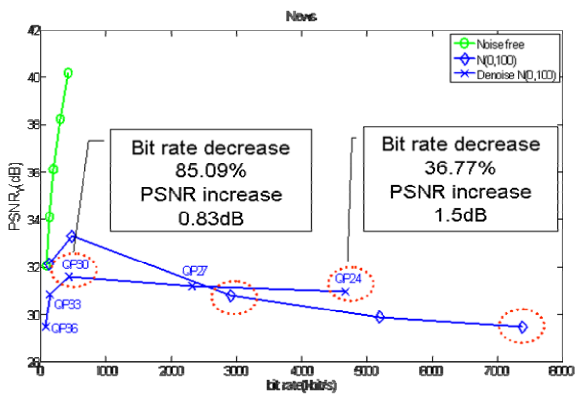


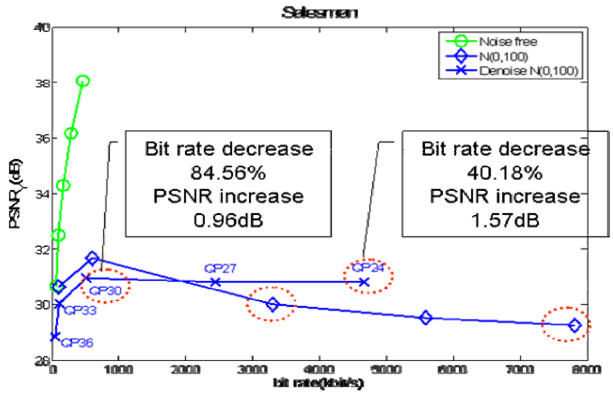
Figure 7. PSNR for impulse 10%

When video sequence suffers from severe noise, the compression efficiency will be decreased seriously. We compare the bitrates between three cases in each sequence, which are noise free, with, and without noise reduction as shown in Figure 8. These experiments are operated at five different quantization levels, QP24, QP27, QP30, QP33, and QP36. These results show that the proposed algorithm reduces the bitrate by 40.18% and increases 1.57dB in PSNR at the same time for Salesman with QP24.

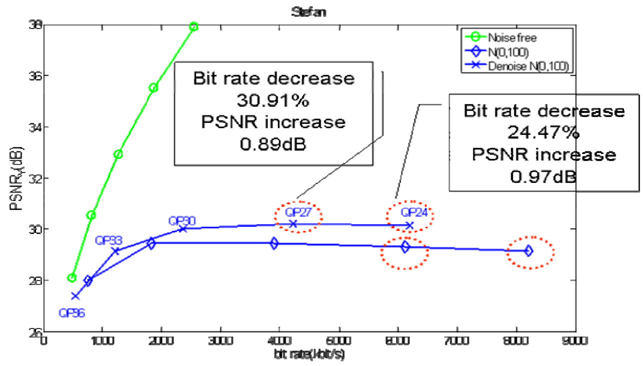
Figure 9 shows the experimental results for sequences with 10% impulse noise. The impulse noise seriously corrupts the video quality and increases the bitrate significantly. The impulse noise reduction operation can obviously help reduce bitrate more significant than Gaussian noise. For Salesman, the bitrate reduction can be achieved up to 91.2% and PSNR improvement is 15.13dB.



(a)

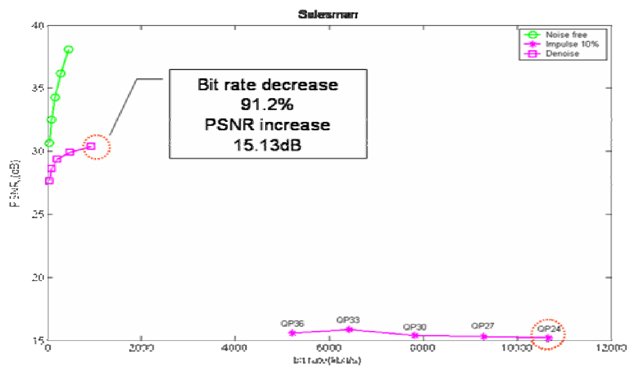


(b)

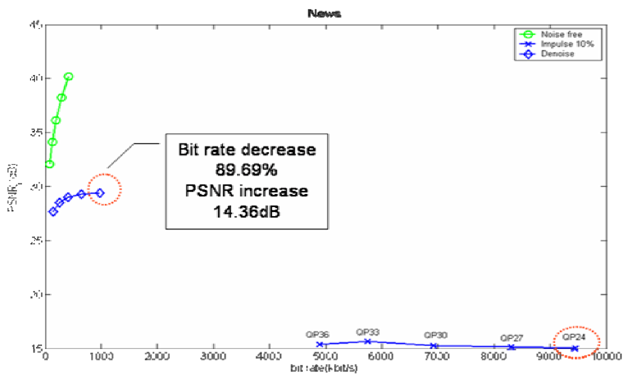


(c)

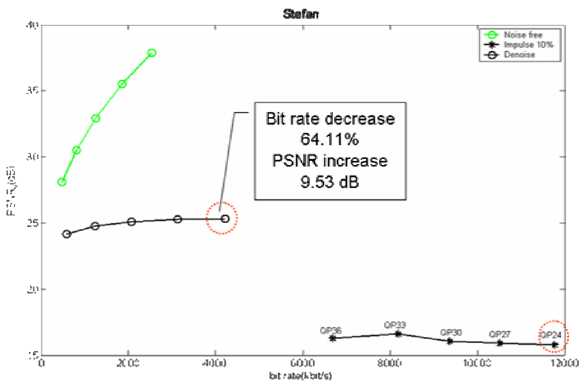
Figure 8. Birate versus PSNR comparison (a)News (b)Salesman (c)Stefan



(a)



(b)



(c)

Figure 9. Birate versus PSNR comparison
(a)Salesman (b)News (c)Stefan

5. CONCLUSION

We have presented a spatio-temporal filter utilizing motion compensation information to reduce Gaussian and impulse noises. This scheme preprocesses video sequence to benefit compression efficiency. We utilize variation of mode distributions to detect noise and determine parameters for spatial filter. In the temporal domain, the motion information is used to perform temporal filtering. The experimental results show that the proposed method can improve PSNR by at least 0.83dB for Gaussian noise and reduce 24.46% bitrate. As to impulse noise, the PSNR increases 8.17dB and the bitrate reduction is 84.85%.

REFERENCES

[1] M. Meguro, A. Taguchi and N. Hamada. ”Data-dependent weighted average filtering for image sequence restoration”, *IEICE Trans.* Vol. J82-A, no.10,pp.1623-1632, Oct 1999 (in Japanese)

[2] J.M. Boyce, ”Noise reduction of image sequences using adaptive motion compensated frame averaging ,” *IEEE Int. Conf. Acoust. ,Speech & Signal Process.*, pp.461-464, 1992

[3] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsaggelos, and R. L. Lagendijk, “Noise reduction filters for dynamic image sequences: A review,” *In Proceedings of the IEEE*, vol. 83, no. 9, pp. 1272-1292, Sept. 1995.

[4] B. C. Song; K. W. Chun, "Motion-compensated temporal pre-filtering for noise reduction ic," ICIP '04. 2004 *International Conference on Image Processing,* vol.2, no.pp. 1221- 1224, 24-27 Oct. 2004.

[5] M. Saeidi, S.A. Motamedi, A. Behrad, B. Saeidi, R. Saeidi, "Noise reduction of consecutive images using a new adaptive weighted averaging filter," *IEEE Workshop on Signal Processing Systems Design and Implementation,* 2005., pp. 455- 460, 2-4 Nov. 2005.

[6] D.H. Kim, R.H. Park, S. Yang, "Noise reduction of image sequences using H.264/AVC variable block-size motion compensation," *ICCE '06. Digest of Technical Papers. International Conference on Consumer Electronics,* .pp. 387-388, 7-11 Jan. 2006.

[7] K. Miyata, A. Taguchi, "Spatio-temporal separable data-dependent weighted average filtering for restoration of the image sequences," *in Proceedings., IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002 (ICASSP '02).* vol.4, no., pp. IV-3696-IV-3699, 2002.

[8] T. Wiegand, G. J. Sullian, G. Bjontegaard, and A. Luthra, “Overview of the H.264/AVC Video Coding Standard,” *IEEE Trans. Circuits System Video Technology*, vol. 13, pp. 560-576, July 2003.

[9] R.C. Hardie and C. G. Boncelet, “LUM Filters：A Class of Rank-Order-Based Filters for Smoothing and Sharpening,” *IEEE Trans. Signal. Processing* .vol.41, no. 3, Page(s). 1064-1076, Mar. 1993.