MOTION ADAPTIVE SPATIO-TEMPORAL GAUSSIAN NOISE REDUCTION FILTER FOR DOUBLE-SHOT IMAGES

Shao-Yi Chien and Tse-Wei Chen

Media IC and System Lab
Graduate Institute of Electronics Engineering and Department of Electrical Engineering
National Taiwan University
BL-421, No. 1, Sec. 4, Roosevelt Rd., Taipei 106, Taiwan
sychien@cc.ee.ntu.edu.tw

ABSTRACT

The high performance of the conventional spatio-temporal image noise reduction filers comes at the cost of high computationally intensity. In this paper, a motion adaptive spatio-temporal filter is proposed for double-shot images to achieve high performance with low computing power requirement. Two images are used simultaneously to separate the static regions and dynamic regions, where different noise reduction approaches are employed. Temporal average and 2-D adaptive filter is applied for static regions, where spatio-temporal filter with motion compensation is only applied for dynamic regions. Experiments show that the proposed noise reduction filter can achieve better noise reduction performance with less computationally intensity than other previous spatio-temporal filters.

Index Terms**—** Noise reduction, double-shot images, spatio-temporal filter, motion adaptive

1. INTRODUCTION

The existence of noise signals is inevitable in image capture devices, such as digital still cameras and digital video camcoders. For semiconductor photodetectors, the temporal noise can be modeled as independent additive Gaussian noise [1] [2], which can be represented as the following equation:

$$g(x, y, t) = f(x, y, t) + n(x, y, t), \tag{1}$$

where $f(x, y, t)$ is the original pixel intensity at location (x, y) at time t, $n(x, y, t)$ is the Gaussian noise with zero mean and σ_η^2 variance, and $g(x, y, t)$ is the degraded observed pixel intensity.

To restore the image and enhance the image quality, several noise reduction filters have been proposed. They can be classified into two main categories: spatial noise reduction filters and sptio-temporal noise reduction filters. The spatial filters remove noise in spatial domain. That is, only one image is considered at a time. In order to remove noise and preserve the detailed texture in an image, the developing trend is to add some adaptation in the filters [3] [4]. From our experiments, the 2-D adaptive filter outperforms other spatial filters [5], where the filter strength is adapted according to the variance of image and the noise variance. The performance is good especially for removing noise with small variance. Note that, noise variance estimation [6] [7] is required for this filter.

If multiple images can be considered at the same time, for example, for video camcoders or digital still cameras with continuous shooting, spatio-temporal noise reduction filters can be employed. This kind of filter can usually outperform other kind of filters [1] since signals in image sequences are more stationary in temporal direction, which leads the noise signals to be easily separated from the original signals. For example, the motion-compensated spatio-temporal filter takes 2^n images as input [2]. It first decorrelates the 2^n images in temporal domain with Hadamard transform to 2-D images and then applies 2-D Wiener filter. It can be viewed as a fast 3-D Wiener filter, and it can achieve good performance. Another example of spatio-temporal filter is the selective filter based on double-shot pictures [8]. With considering two images, the high spatial frequency parts in an original image can be separated from noise, and the noise reduction filter can be only applied on the smooth regions. Although the performances of spatio-temporal noise reduction filters are better, they always require large computing power especially when motion compensating operation is involved.

For double-shot pictures, we think the spatio-temporal noise reduction filter can be done in a more efficient way with further considering motion adaptation. That is, for static regions and dynamic regions, different denoising approaches can be employed. For static regions, the signals are temporally stationary. The noise variance can be reduced with simple temporal filters, and the noise variance can be easily estimated. For dynamic regions, spatio-temporal filters can be employed, and motion compensation operation is only used in these regions to increase the temporally stationary. Therefore, in this paper, we would propose a motion adaptive spatio-temporal noise reduction filter to achieve high performance with low computational intensity.

This paper is organized as follows. In Section 2, the proposed noise reduction filter is introduced. The experimental results are then shown in Section 3. Finally, Section 4 gives the concluding remarks.

2. PROPOSED NOISE REDUCTION FILTER

The block diagram of the proposed motion adaptive spatio-temporal filter (MASTF) is shown in Fig. 1. Two images, including an input image $g(x, y, t)$ and an auxiliary image $g(x, y, t+1)$, are considered at the same time. There are five important steps in the proposed filter: *Moving Pixel Detection*, *Temporal Average Filter*, *Spatio-Temporal Noise Reduction Filter*, *Image Fusion*, and *2-D Adaptive Filter in Static Regions*. *Moving Pixel Detection* first separates static regions and dynamic regions in an image. Next, the noise variance in static regions can be reduced with *Temporal Average Filter*, and the *Spatio-Temporal Noise Reduction Filter* with motion compensation is applied in the dynamic regions. *Image Fusion* then combines both the output images as an image with low noise variance. At the last step,

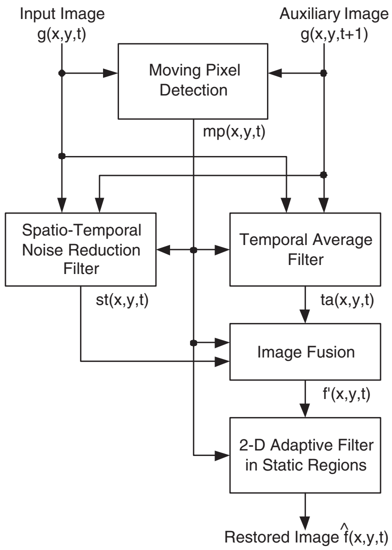


Fig. 1. Block diagram of the proposed motion adaptive noise reduction filter.

2-D Adaptive Filter in Static Regions can further reduce the noise in the static regions. The details of each operation are described as follows.

We first define the mean absolute difference (MAD) of a correlation window around a pixel (x, y) with motion vector (u, v) as

$$MAD_{(x,y,t)}(u,v) = \frac{1}{(2M+1)(2N+1)} \sum_{m=-M}^{M} \sum_{n=-N}^{N} |g(x+m, y+n, t) - g(x+m+u, y+n+v, t+1)|, \quad (2)$$

where $(2M + 1)$ and $(2N + 1)$ are the width and height of the correlation window. With $MAD_{(x,y,t)}(0, 0)$, we can separate dynamic regions and static regions and record the positions of dynamic pixels in the "moving pixel mask," $mp(x, y, t)$:

$$mp(x,y,t) = \begin{cases} 0 & \text{if } MAD_{(x,y,t)}(0,0) < mth \\ 1 & \text{else} \end{cases}, \quad (3)$$

where mth is a threshold value and is proportional to the noise standard deviation. The empirical value of mth is $1.5\sigma_\eta$. An example of the moving pixel mask is shown in Fig. 2(b), where the MPEG video sequence *Weather* is used. Note that, the noise variance σ_η^2 can be derived by noise variance estimation [6], or it can also be calculated iteratively from the frame difference of the static regions.

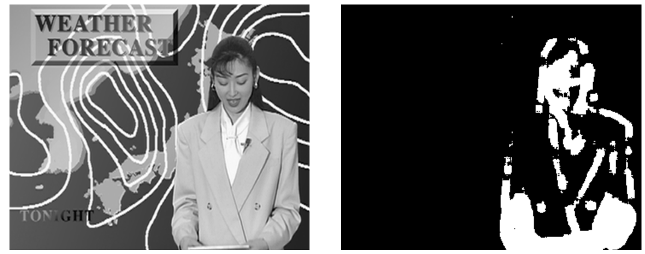
In static regions, the signals are temporally stationary. A simple temporal average filter can be applied to reduce the variance of noise signals in *Temporal Average Filter* step. That is, for $mp(x, y, t) = 0$,

$$ta(x,y,t) = \frac{1}{2}[g(x,y,t) + g(x,y,t+1)], \quad (4)$$

with which the noise variance is reduced to $\sigma_\eta^2/2$.

On the other hand, in the dynamic regions, the dense motion field is first derived as the following equation.

$$(U_{(x,y,t)}, V_{(x,y,t)}) = \arg\min_{(u,v)} MAD_{(x,y,t)}(u,v), \quad (5)$$



(a) (b)

Fig. 2. (a) Sequence *Weather*. (b) The moving pixel mask $mp(x, y, t)$.

where $(U_{(x,y,t)}, V_{(x,y,t)})$ is the motion vector of the pixel (x, y). Next, a spatio-temporal filter can be employed with motion compensation in *Spatio-Temporal Noise Reduction Filter* step. Here, a 3-D extension of a 2-D adaptive filter [5] is employed as follows.

$$st(x,y,t) = g(x,y,t) - \frac{\min\{\sigma_{tl}^2, \sigma_\eta^2\}}{\sigma_{tl}^2}[g(x,y,t) - mt(x,y,t)], \quad (6)$$

where

$$mt(x,y,t) = \frac{1}{2(2R+1)(2S+1)} \times \{\sum_{m=-R}^{R}\sum_{n=-S}^{S} g(x+m, y+n, t) + \sum_{m=-R}^{R}\sum_{n=-S}^{S} g(x+m+U, y+n+V, t+1)\}, \quad (7)$$

$$\sigma_{tl}^2 = \frac{1}{2(2R+1)(2S+1)} \times \{\sum_{m=-R}^{R}\sum_{n=-S}^{S} [g(x+m, y+n, t) - mt(x,y,t)]^2 + \sum_{m=-R}^{R}\sum_{n=-S}^{S} [g(x+m+U, y+n+V, t) - mt(x,y,t)]^2\}. \quad (8)$$

Note that, only pixels with $mp(x, y, t) = 1$ need to be considered for both motion estimation and spatio-temporal filter, which can save a lot of computation.

In *Image Fusion*, the two images, $ta(x, y, t)$ and $st(x, y, t)$, are then fused together to form a partial restored image $f'(x, y, t)$ as the following equation.

$$f'(x,y,t) = \begin{cases} ta(x,y,t) & \text{if } mp(x,y,t) = 0 \\ st(x,y,t) & \text{else} \end{cases}. \quad (9)$$

After that, in *2-D Adaptive Filter in Static Regions* step, since the noise variance of the static regions in $f'(x, y, t)$ is smaller that that of $g(x, y, t)$, 2-D adaptive filter can perform well, as shown in the following equations.

$$\hat{f}(x,y,t) = \begin{cases} f'(x,y,t) & \text{if } mp(x,y,t) = 1 \\ f'(x,y,t) - \frac{\min\{\sigma_L^2, \sigma_\eta^2/2\}}{\sigma_L^2} \times [f'(x,y,t) - m(x,y,t)] & \text{else} \end{cases}, \quad (10)$$

where

$$m(x,y,t) = \frac{1}{(2R+1)(2S+1)} \times \sum_{m=-R}^{R} \sum_{n=-S}^{S} f'(x+m, y+n, t), \quad (11)$$

$$\sigma_L^2 = \frac{1}{(2R+1)(2S+1)} \times \sum_{m=-R}^{R} \sum_{n=-S}^{S} [f'(x+m, y+n, t) - m(x,y,t)]^2. \quad (12)$$

The final output restored image is $\hat{f}(x,y,t)$.

3. EXPERIMENTAL RESULTS

Six noise reduction filters are used for comparison. They are 2-D adaptive filter (2DA) [5], detail-preserving filter (DP) [3], selective filter based on double-shot images (DS) [8], motion-compensated spatio-temporal filter with two images and four images (MC2 and MC4) [2], and the proposed motion adaptive spatio-temporal filter (MASTF), whose parameters are set as $M = N = 4$ and $R = S = 1$. Among them, 2DA and DP are spatial filters, and the other four filters are spatio-temporal filters. Note that, the same dense motion estimation algorithm is applied for MC2, MC4, and MASTF. For DS, we did not implement the motion compensation operation, and so the performance may be lower than the original filter for test sequences with dynamic scenes. Three test sequences are employed. *Sean* and *Weather* are standard test video sequences from MPEG, and *Peppers* is an image. For *Sean* and *Weather*, frame 19, 19–20, or 19–22, are used if one, two, or four images are required for each algorithm. The image *Peppers* is duplicated to several copies, and they are added noise independently as the test bench for an ideal static scene situation.

The objective evaluation results are shown in Table 1, where PSNR is used as the evaluation criterion, and sequences with different noise variances ($\sigma_\eta = 5, 10,$ and 15) are tested. From the table, it shows that in most cases, spatio-temporal filters perform better than spatial filters. It also shows that the performance of the proposed filter, MASTF, is similar to that of MC4 and is better than MC2 and DS. On the other side, Table 2 presents the runtime comparison of different spatio-temporal filters, where the test platform is a PC with Intel Pentium M 1.6 MHz CPU. It shows that the proposed filter requires much less computation power. From these experiments, it is proved that the proposed filter can achieve the similar performance of MC4 with much less computational intensity.

The subjective evaluation results with sequence *Sean* with $\sigma_\eta^2 = 100$ are shown in Fig. 3. It also shows the performance of MASTF is similar to that of MC4 and is better than those of other filters, since it can remove the noise and maintain the details of the face and tie as well. Another test image sequences *NTU Sports Center* are captured by a digital still camera, Canon EOS 350D, with double shooting at ISO value of 1600, as shown in Fig. 4. Note that, raw data format is used to avoid the influence from image signal processing and JPEG compression. With the noise variance estimation [6], we find the noise variance is only 22. Parts of the noise-reduced images are shown in Fig. 5. It also shows that compared with other filters, the proposed noise reduction filter can perform well and keep the details as well.

Table 1. PSNR Comparison of Different Noise Reduction Filters.

Sequence (PSNR (dB))	PSNR (dB) 2DA [5]	DP [3]	DS [8]	MC2 [2]	MC4 [2]	MASTF
Sean (34.79)	37.94	33.02	37.69	39.12	39.90	39.54
Sean (28.56)	32.99	32.00	32.89	34.01	35.00	35.03
Sean (24.88)	30.08	29.57	30.23	30.85	31.67	32.26
Weather (34.91)	37.43	31.00	35.61	38.82	39.73	39.05
Weather (28.59)	32.02	30.12	31.23	33.39	34.57	33.92
Weather (24.98)	28.94	28.24	28.57	30.15	31.20	31.02
Peppers (34.86)	37.19	33.92	37.51	38.37	39.60	39.04
Peppers (28.49)	33.12	32.98	33.56	33.72	34.44	34.99
Peppers (24.89)	30.65	30.29	31.21	31.02	31.52	32.70

Table 2. Runtime Comparison of Different Spatio-Temporal Filters.

Sequence (size)	Runtime (s) MC2	MC4	MASTF
Sean (352x288)	67.86	203.03	0.44
Weather (352x288)	67.72	203.06	4.08
Peppers (512x512)	177.16	531.08	0.49

4. CONCLUSION

In this paper, a motion adaptive spatio-temporal noise reduction filter is proposed for double-shot pictures. This filter separates static regions and dynamic regions first and adaptively applies different noise reduction techniques for different regions. For static regions, temporal average filter followed by 2-D adaptive filter is used. Spatio-temporal filter is employed for dynamic regions with motion compensation. Experiments show that the proposed filter can achieve the best performance with only light computation power requirement. The proposed algorithm still has the limitation to be employed only for situations with zero or near zero camera motion in the period of double-shot, for example, when digital still cameras are placed on tripods. It will be included in our future work to extend it to more general situations.

5. REFERENCES

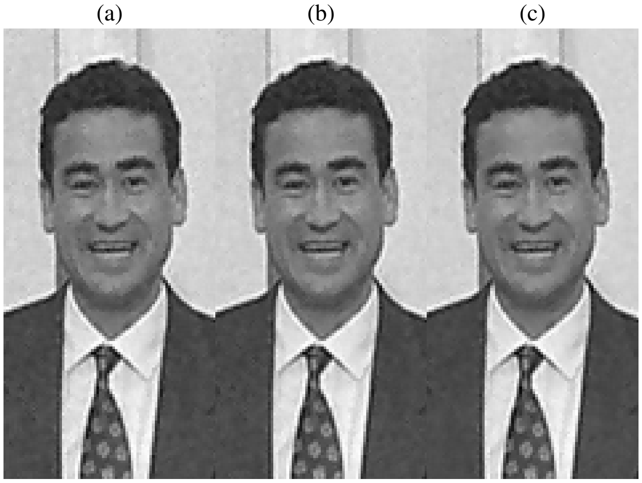
[1] J. C. Brailean, R. P. Kleihorst, S. Efstratiadis, A. K. Katsaggelos, and R. L. Lagendijk, “Noise reduction filters for dynamic image sequences: a review,” *Proceedings of the IEEE*, vol. 83, no. 9, pp. 1272 – 1292, Sept. 1995.

[2] K. J. Boo and N. K. Bose, “A motion-compensated spatio-temporal filter for image sequences with signal-dependent noise,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 3, pp. 287 – 298, June 1998.

[3] G. Ramponi, “Detail-preserving filter for noisy images,” *Electronics Letters*, vol. 31, no. 11, pp. 865 – 866, May 1995.

[4] J. N. Lin, X. Nie, and R. Unbehauen, “Two-dimensional LMS adaptive filter incorporating a local-mean estimator for image processing,” *IEEE Transactions on Circuits and Systems–II*

(a) (b) (c)



(d) (e) (f)

Fig. 3. (a) The degraded image, *Sean*, with $\sigma_\eta^2 = 100$. (b) The output image of 2DA filter. (c) The output image of DS filter. (d) The output image of MC2 filter. (e) The output image of MC4 filter. (f) The output image of the proposed filter, MASTF.

Analog and Digital Signal Processing, vol. 40, no. 7, pp. 417–428, July 1993.

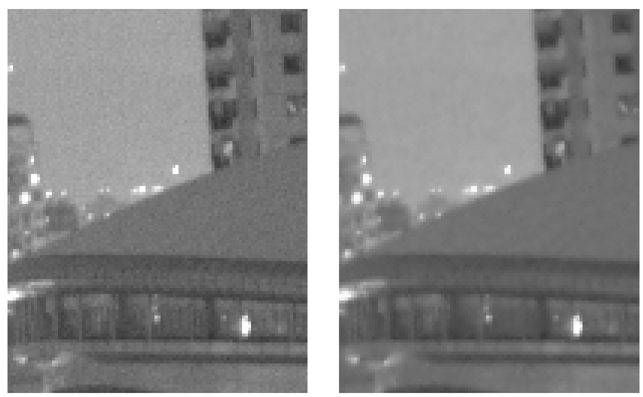
[5] J.-S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-2, no. 2, pp. 165 – 168, Mar. 1980.

[6] K. Rank, M. Lendl, and R. Unbehauen, "Estimation of image noise variance," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 146, no. 2, pp. 80–84, Apr. 1999.

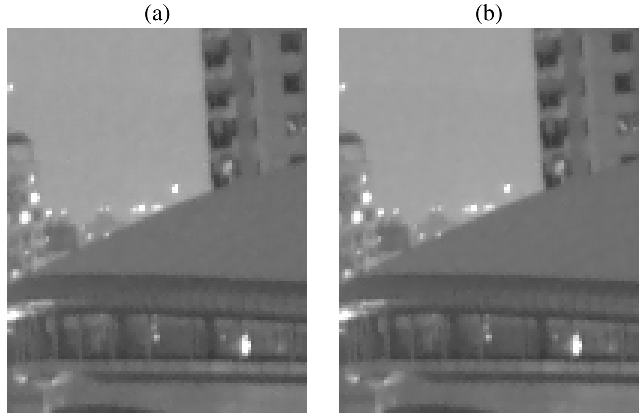
[7] A. Bosco, A. Bruna, G. Messina, and G. Spampinato, "Fast method for noise level estimation and integrated noise reduction," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 3, pp. 1028 – 1033, Aug. 2005.

[8] D. A. Florêncio, "Image de-noising by selective filtering based on double-shot pictures," in *Proceedings of 2005 International Conference on Image Processing*, 2005, pp. III–353 – III–356.

Fig. 4. Test image sequence captured with a digital still camera: *NTU Sports Center*.



(a) (b)



(c) (d)

Fig. 5. (a) The captured image. (b) The output image of DS filter. (c) The output image of MC2 filter. (d) The output image of the proposed filter, MASTF.