VIDEO DENOISING WITH ADAPTIVE TEMPORAL AVERAGING

V. B. Surya Prasath[*]

Department of Electrical Engineering and Computer Science, University of Cincinnati, OH 45221 USA
Department of Biomedical Informatics, University of Cincinnati, OH 45267 USA
Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, OH 45229 USA

ARTICLE INFO	Abstract:
	Recently the proliferation of digital videos has increased exponentially due to availability of consumer cameras . Despite the improvement in the sensor technologies, one of the fundamental problems is that of noise affecting the video scenes. Recently, adaptive, pixel-wise, temporal averaging methods can advocate in denoising videos. In this work, we adapt the edge maps of frames within temporal averaging to guide the denoising away from the edges. This allows the filtering to remove noise in intermediate flat regions while respecting boundaries of objects better. The experimental results indicate that we can obtain improved video denoising results in comparison to other filtering methods.

1 Introduction

Videos are prevalent now due to the internet penetration in various parts of the world and the availability of video shooting mode in millions of digital cameras and smartphones. Although hardware improvements paved the way for obtaining higher quality videos, noise is still a problem and it requires careful consideration. Removing digital noise in videos is still an active area research and various filters available in the image processing literature can be modified, and adapted to the video domain [1-7]. Bartovčak and Vrankić [2] recently proposed an pixel-wise adaptive temporal averaging (ATA) based video denoising method. Temporal coherence of the signal is established using neighborhood signal points for averaging intervals. Simple thresholding based decision is used along with mean values to obtain denoised results. In this work, we propose to improve the ATA method by using edge maps for denoising videos more efficiently. Though the original ATA does denoising using the information from neighboring frames, it ignores any edge information available in the current frame under consideration. For this purpose we have computed an edge map and introduce a decision criteria incorporated into the original ATA approach. The experimental results are given on various noisy video sequences to show the advantage of our improved method compared to the original and related denoising filters from the literature.

The rest of this paper is organized as follows. Section 2 introduces the temporal averaging video denoising method along with our improvement. Section 3 presents experimental results, and Section 4 concludes the paper.

[*] Corresponding author.
E-mail address: surya.iit@gmail.com.

2   The proposed improved video denoising method

2.1 Adaptive temporal averaging (ATA) approach

We follow [2] in notations and basic concepts for adaptive temporal averaging (ATA) method. Let $f[k]$ be the noisy signal that is a representation of time-dependent pixel intensity from a monochromatic video. The values taken by $f[k]$ are between 0 and 255 and have a finite number of samples, $k = 1, 2, 3, …, m$. The standard additive Gaussian noise model is assumed here, then the noisy signal is obtained by

$$f[k] = s[k] + n[k] \tag{1}$$

where $s[k]$ is the latent signal and $n[k]$ is Gaussian noise with zero mean and variance σ^2. One can adapt various image filtering techniques here to obtain a clean signal, however without utilizing the temporal coherence these filters are limited in terms of noise removal efficiency. The basic idea proposed in Bartovčak and Vrankić [2] is that of utilizing temporal coherence of the signal to establish averaging intervals. They proposed the following mean value based estimation

$$e[k] = \frac{f[l_k] + … + f[k] + … + f[r_k]}{r_k - l_k + 1} \tag{2}$$

where

$$r_k \geq k \geq l_k \tag{3}$$

are the right and left border of the averaging interval. Two thresholds were utilized on both sides of the pixel (say index k, sample) under consideration to determine the right and left borders of the averaging intervals. First threshold (denoted by Threshold A), is used to check the immediate samples (neighborhoods) on the right side ($k+1$) and left side ($k-1$). If the absolute difference is greater than the Threshold A, the border of the averaging interval is found. Moreover, further neighborhoods on the right side ($k+2, k+3,…$), and left side ($…, k-3, k-2$) are added and checked against the second threshold (denoted by Threshold B). This cumulative difference of the pixel under consideration (k) is

greater than the Threshold B. The border of the averaging interval is found.

The mean values between these right and left borders are then computed to take as the denoised version of the sample. Fig. 1 shows the ATA method in detail on a 1-D signal.
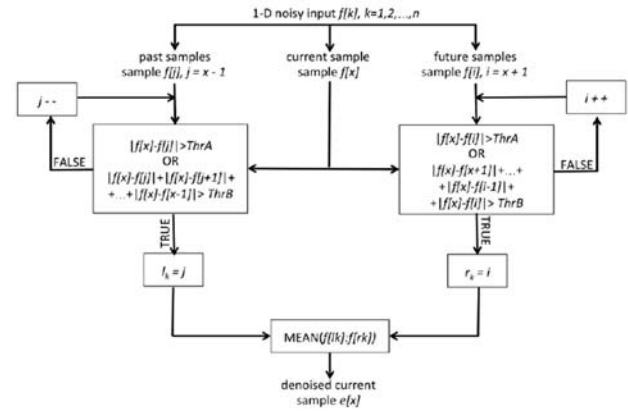

Figure 1. Flowchart of the original adaptive temporal averaging method, adapted from [2].

2.2 Proposed improvement

Note that the immediate pixels based approach employed by the ATA method [2] is equivalent to finding gradient jumps ($f[k]$-$f[k+1]$ or $f[k]$-$f[k-1]$ are first order approximations of first derivative). Similarly, the cumulative differences are higher order approximations of the derivative. Thus, the jumps in the temporal direction are captured effectively, thereby leading to good denoising results with simple average values of neighbor samples. However, the edge information within each of the frames is not utilized.

To remedy this here we consider the Canny edge detector output (a binary image) at every video frame and incorporate another decision criteria based on whether the sample under consideration is an edge pixel or not. The averaging interval is on either side (right or left) and it is stopped if it is an edge pixel. This additional conditional criterion helps to avoid the edge blurring due to the usage of the mean values. The canny edge detector [4] is used here, as it is well known to obtain robust edge detection results under various additive Gaussian noise levels. Fig. 2 shows the improved ATA (IATA) method based on the same 1-D signal notations shown in Fig. 1.
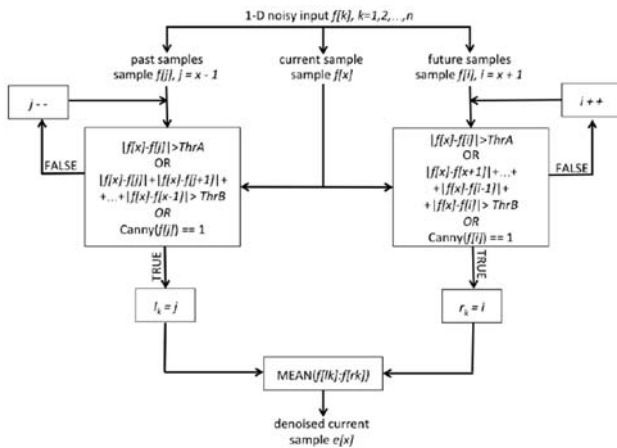
Figure 2. Flowchart of our proposed improved adaptive temporal averaging (IATA) method that incorporates edge information.

3 The experimental results and discussion

The parameters are fixed following the guidelines provided in [2]. In [2] the authors chose the Threshold A = 5σ and the Threshold B = 10σ as empirical optimal values. We utilize the default Canny edge detector parameters available in MATLAB in-built function *edge(f,'Canny')*. We fix these parameters in all the experimental results reported here.

3.1 1-D signals

We used the two 1-D signals to compare the original ATA method with our proposed modification. Two signals with different characteristics were used:

- **Signal A** – quasi-rectangular signal; features areas of constant values and big jumps in signal values at two locations (one high and one low).
- **Signal B** – quasi-sine wave; continuous smooth change of signal values with no abrupt jumps.

These two 1-D signals were corrupted by adding Gaussian noise of standard deviation $\sigma = 10$, and 20. Figs. 3 and 4 show the comparison of 1-D signals denoising under different noise levels against the original ATA method [2]. As it can be seen, the proposed method performs better with edge preservation in the Signal A case with almost identical to the noise-free original signal in low noise level $\sigma = 10$, and better than then ATA method result in higher noise level $\sigma = 20$.

In the smoother Signal B case, our proposed IATA outperforms the ATA method in both low and higher noise levels. Though ATA method is able to achieve

substantial noise removal, it can obtain spurious smooth bumps due to the averaging mean values used to noisy jumps. Our method in contrast uses true edge jump locations thereby alleviating the smoothing effects of the mean filtering on true edge signal locations and smoothens out at other noisy signal locations. Table 1 shows the mean squared error (MSE) values indicating quantitatively that we have obtained better noise removal compared to the original ATA method under various noise levels.

Table 1. MSE values of 1-D signals denoised using our proposed IATA method in comparison with the ATA [3].

Noise σ		10	15	20
MSE values				
Signal A	ATA	2.372	3.720	21.292
	Proposed	1.784	3.112	20.501
Signal B	ATA	9.632	10.224	13.765
	Proposed	9.001	9.793	13.185

Note that the MSE values (as well as the signals shown in Figs. 3 and 4) slightly differ from the ones reported in [2] since we used different random noise initializations (seeds) in generating the Gaussian noise profiles using MATLAB.

3.2 Image sequences

Next, we test the denoising methods on multiple standard (gray scale) image sequences of various scenes, namely "Salesman", "Miss America", "Tennis" under different noise levels $\sigma = 10$, 15, and 20. The original ATA method utilizes the mean values of the neighboring signal points. However, in corresponding denoised videos we can observe edge blurring due to the lack of the edge information utilized in the averaging interval computation and the mean filter employed. In contrast, our proposed method has obtained better edge preservation and noise removal. The usage of extra condition utilazing the edge pixels stopped the mean filtering from smoothing out the edges.

We compare the denoising performance of our proposed approach with state-of-the-art video denoising algorithms SEQWT [4], IFSM [5], ST-GSM [7], along with the original ATA [2] method. As a baseline, we use the still GSM [6], a static image GSM denoising method applied on a frame-by-frame bases.
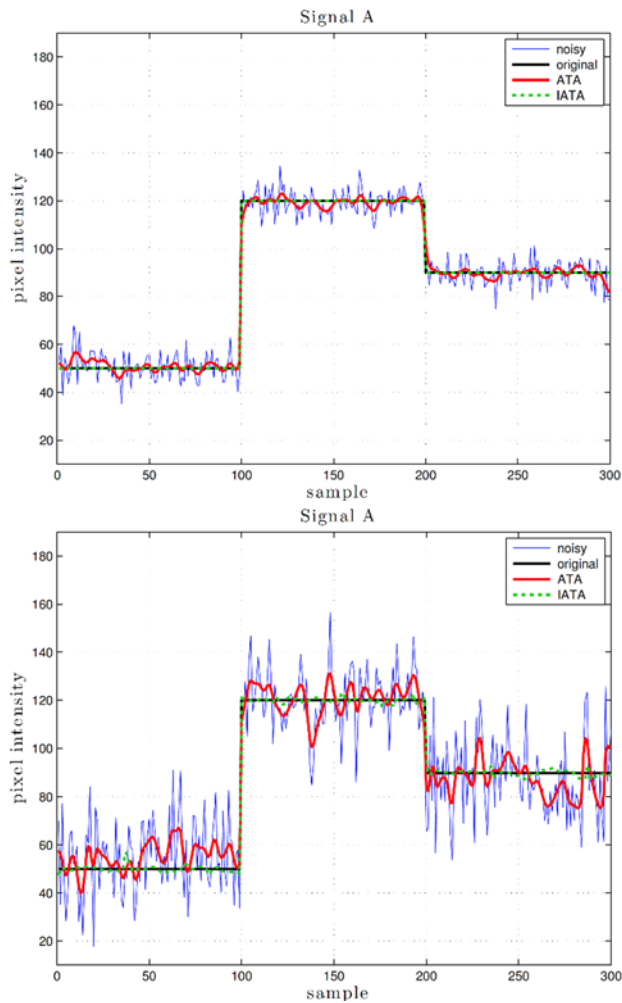
Figure 3. Denoising 1-D Signal A with original ATA [3] and proposed IATA method for noise levels σ = 10 (top), and 20 (bottom).

Figure 4. Denoising 1-D Signal B with original ATA [3] and proposed IATA method for noise levels σ = 10 (top), and 20 (bottom).

To quantitatively compare the denoising performance we utilize two common error metrics form the image processing literature, namely, the peak signal to noise ratio (PSNR), and structural similarity (SSIM). Table 2 shows the PSNR (in decibels - dB) and SSIM (in range [0 1]) values for different denoising methods on different image sequences.

As it can be seen, our proposed method obtained an overall improvement over the original ATA method [2] and performs closer to the state-of-the-art ST-GSM [7] method. It is also apparent from the mean SSIM (MSSIM) values that our method outperforms the other compared denoising methods though still less than the values obtained with more sophisticated and state of the art method of ST-GSM [7].
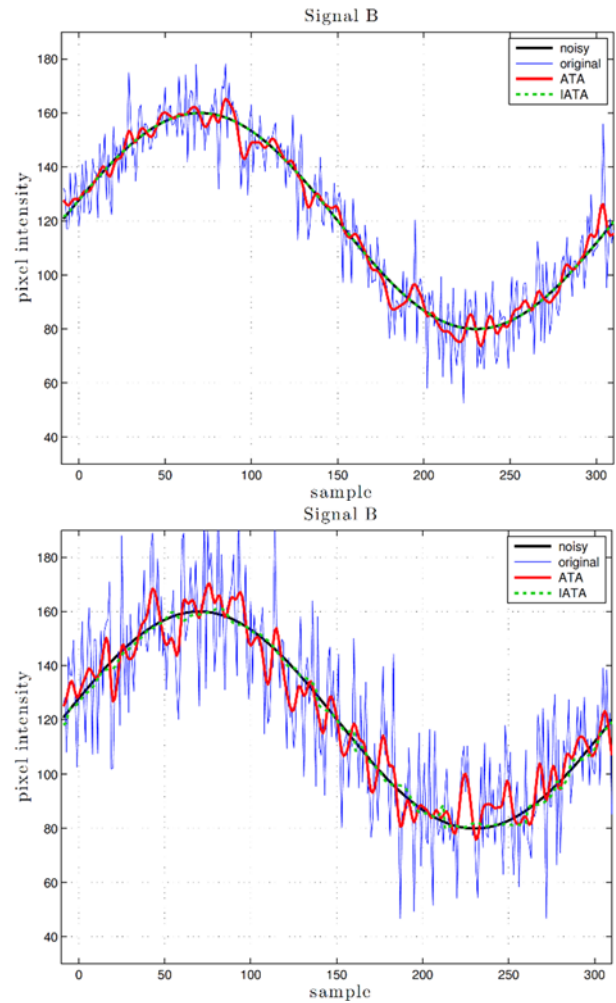
4   Conclusion

Video denoising with improved adaptive pixel-wise temporal averaging is considered in this work. By utilizing the edge pixels estimation, we have improved the temporal averaging method for obtaining better denoising without introducing any artifacts.

The extension to color video sequences require careful modeling of color changes across the image frames and defines an interesting avenue to be explored [9]. Further, improvements can be obtained by using other rule based methods [3] along with the temporal averaging approach.

Table 2. PSNR and MSSIM values of video sequences denoised using the proposed IATA method and comparison to other methods.

Videos	"Salesman"			"Miss America"			"Tennis"		
Noise σ	10	15	20	10	15	20	10	15	20
PSNR (dB)									
Noisy	28.16	24.72	22.32	28.15	24.62	22.29	28.22	24.72	22.25
SEQWT [4]	32.86	30.59	29.09	NA	NA	NA	31.19	29.14	27.59
IFSM [5]	34.22	31.85	30.22	37.52	35.41	33.86	32.41	30.10	28.56
still GSM [6]	33.80	31.73	30.28	38.52	37.14	36.14	31.82	29.87	28.65
ST-GSM [7]	38.04	36.03	34.61	40.57	39.40	38.50	34.05	31.97	30.59
ATA [2]	35.64	33.78	32.51	37.16	35.77	34.30	32.95	30.55	28.72
Proposed	36.23	34.12	32.85	38.48	36.01	35.16	33.33	30.81	29.11
MSSIM									
Noisy	0.718	0.574	0.467	0.493	0.321	0.226	0.719	0.573	0.466
SEQWT [4]	0.900	0.846	0.796	NA	NA	NA	0.842	0.772	0.716
IFSM [5]	0.904	0.851	0.801	0.904	0.857	0.812	0.855	0.776	0.709
still GSM [6]	0.909	0.865	0.825	0.936	0.922	0.913	0.831	0.758	0.711
ST-GSM [7]	0.960	0.941	0.923	0.952	0.943	0.936	0.894	0.841	0.797
ATA [2]	0.941	0.920	0.899	0.910	0.884	0.861	0.870	0.813	0.756
Proposed	0.950	0.931	0.891	0.934	0.905	0.908	0.852	0.824	0.769

References

[1] Aubert, G., Kornprobst, P.: *Mathematical Problems in Image Processing: Partial Differential Equation and Calculus of Variations*, Springer-Verlag, New York, 2006.

[2] Bartovcak, D., Vrankic, M.: *Video denoising based on adaptive temporal averaging*, Engineering Review, 32 (2012), 2, 64-69.

[3] Lerga, J., Sucic, V., Grbac, E.: *An adaptive method for video denoising based on the ICI rule*, Engineering Review, 32 (2012), 1, 33-40.

[4] Zlokolica, V., Pizurica, A., Philips, W.: *Combined wavelet and temporal video denoising*, Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, Miami, Florida USA, 2003, 334-341.

[5] Rahman, S. M. M., Ahmad, M. O., Swamy, M. N. S.: *Video denoising based on inter-frame statistical modeling of wavelet coefficients*, IEEE Transactions on Circuits and Systems for Video Technology, 17 (2007), 2, 187-198.

[6] Portilla, J., Strela, V., Wainwright, M. J., Simoncelli, E. P.: *Image denoising using scale mixtures of Gaussians in the wavelet domain*, IEEE Transactions on Image Processing, 12 (2003), 11, 1338-1351.

[7] Varghese, G., Wang, Z.: *Video denoising using a spatiotemporal statistical modeling of wavelet coefficients*, Proceedings of the IEEE Acoustics, Speech and Signal Processing International Conference, Las Vegas, Nevada USA, 2008, 1257-1260.

[8] Canny, J. F.: *A computational approach to edge detection*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8 (1986), 6, 679-698.

[9] Moreno, J. C., Prasath, V. B. S., Neves, J. C.: *Color image processing by vectorial total variation with gradient channels coupling*. Inverse Problems and Imaging, 10 (2016), 2, 461-497.