HOMOGENEITY-BASED DIRECTIONAL SIGMA FILTERING OF VIDEO NOISE

Mohammed Ghazal, Aishy Amer and Ali Ghrayeb

Electrical and Computer Engineering, Concordia University
Montréal, Québec, Canada
{moha_mo,amer,aghrayeb}@ece.concordia.ca

ABSTRACT

This paper proposes a real-time method for the reduction of white Gaussian video noise. The method achieves a maximum gain of 4.8 dB and is capable of preserving image content. It adapts window size, weighting and behavior to both image content and noise level in order to optimize the filtering. It starts by detecting the intensity-homogeneous direction from 8 different candidates. A variant of the Sigma filter is then applied directionally. The filtering is performed along homogeneous areas and not across edges. For noisy images, the filtering is increased automatically by using the two most homogeneous directions with a larger kernel size. The proposed filter achieves better image preservation by turning off gradually for less noisy images. It works well for both highly noisy and good-quality images without the introduction of speed or hardware implementation challenges.

1. INTRODUCTION

Video processing applications continuously increase in volume and complexity. As new video services, such as in mobile devices, emerge, the need for efficient noise reduction techniques becomes more apparent. In general, a noise filter should be both fast and able to work well with variable noise levels. This paper proposes a spatial noise filter that meets these requirements.

A number of spatial adaptive filters have been proposed over the years. The method in [1] introduced a spatial filter that is based on the $2\sigma_\eta$ criterion defined by the Chebyshev Inequality. This filter is known as the Sigma Filter (SF) and is widely used as a benchmark for testing spatial adaptive filters against. This filter selects which pixels to include in or exclude from the averaging process based on the noise level. The filter defined in [2] and [3] proposed a Recursive Sigma Filter (RSF) that adds a number of modifications to the Sigma Filter. These modifications include changing the shape of the kernel, making the filtering recursive and using a more sophisticated weighting function. With these modifications, the RSF yields a maximum gain of 3.5 dB. To adapt to the input noise level, the RSF use the noise estimation method in [3] which we have also used to adapt the SF.

A classical approach to spatial noise filtering is the noise-adaptive Wiener Filter (WF) [4]. It produces a relatively high gain for noisy images. However, it does not handle clean images well despite being noise level adaptive. It also introduces some blurring and is computationally more complex than other methods.

The proposed approach is based on [5] where an adaptive averaging Homogeneity-based Filter (HF) is applied to the intensity most homogeneous direction. The authors in [5] introduced an effective criteria to measure homogeneity.

In this paper, we improve upon the technique proposed in [5] with 1) increased gain by using a noise-adaptive kernel size and using the two most homogeneous directions and 2) increased preservation of image content by applying the $2\sigma_\eta$ criterion directionally.

The remainder of the paper is as follows. Section 2 presents the proposed approach theoretically and gives an interpretation of its good performance. Objective simulation results are presented and discussed in Section 3. Finally, Section 4 concludes the paper.

2. PROPOSED APPROACH

Video signals are spatially correlated in nature. Using low-pass filtering, spatially uncorrelated noise can be reduced while preserving the original signal if the frame is unstructured. Structured frames with fine details like edges and corners gets blurred with such filtering [5].

The principle idea of the proposed approach is not to filter similarly everywhere in the image, but to adapt the filtering in a number of ways to the frame and noise characteristics. To tailor a filter that will preserve the

This work was supported, in part, by the *Fonds de la recherche sur la nature et les technologies du Québec* (NATEQ).

structure, we detect image structure in the area surrounding the processed pixel. The main steps of the proposed method are:

1. Detect the two most homogeneous directions from eight candidates.

2. Adapt the filter window size, weighting and behavior to the input image noise as follows:

   (a) If $PSNR_\eta \leq t_\eta$, use the two most homogeneous directions with a $W = 5$ kernel where t_η is a threshold (see later, e.g., see Fig. 4).

   (b) If $PSNR_\eta > t_\eta$, use the most homogeneous directions with a $W = 3$ kernel.

   (c) Adapt pixel weights to the noise level.

The detection of image structure is done using eight intensity-homogeneity analyzers as in [5] (Fig. 1). These analyzers work as directional Laplacian operators with the coefficients $\{-1, -1, .., W-1, .., -1, -1\}$, where W is a positive odd integer. The output of applying these operators to the pixels with close intensity values is close to 0. A kernel size of $W = 3$ was used in [5].

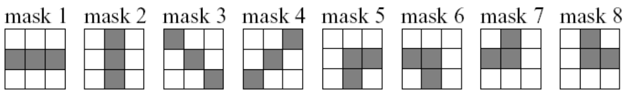

Fig. 1. Homogeneous-Intensity Analyzer Masks [5].

For highly noisy images, however, using a kernel of $W = 5$ achieves theoretically higher gain than with $W = 3$. On the other side, using a kernel of $W = 5$ introduces blurring in good-quality images. This is evident in Fig. 4 where HF[5] with $W = 3$ is compared to its $W = 5$ variant. To combine the benefits of both kernel sizes, we propose to adapt the window size to the noise level. The noise level is estimated using the approach in [6]. Fig. 4 suggests how the selection of the window size can be determined automatically by defining a threshold to decide at what noise level the switch from a $W = 5$ to $W = 3$ kernel is made. The point the 3×1 and 5×1 kernel sizes intersect corresponds to a noise level of 28 dB which represents the used threshold.

To find the two most homogeneous direction, define σ_η to be the estimated noise standard deviation and ϵ_i the output of applying $mask_i$ of size $W = 3$ to the currently processed pixel $I_{(n,m)}$ with spatial coordinates n and m where $i = 1, .., 8$ is the mask index. Depending on which i gives the minimum ϵ_i, the input-output relationship for the filter in [5] is given by Eq. 1.

$$I^o_{(n,m)} = \frac{w_{\sigma_\eta} I_{(n,m)} + \sum\limits_{a,b \in \{-1,0,1\}} f(i) I_{(n+a,m+b)}}{w_{\sigma_\eta} + 2} \quad (1)$$

where $f(i)$ is a 1 or 0 based on the direction detected to be the most homogeneous and w_{σ_η} is the central pixel weight equal to $r \cdot \sigma_\eta$ where r is a positive constant < 1. For example, if $mask_1$, i.e., the horizontal direction, was detected to be the most homogeneous, Eq. 1 reduces to Eq. 2

$$I^o_{(n,m)} = \frac{w_{\sigma_\eta} I_{(n,m)} + I_{(n,m-1)} + I_{(n,m+1)}}{w_{\sigma_\eta} + 2} \quad (2)$$

As can be seen from Eq. 1, $f(i)$ depends only on the homogeneous direction. This means that the $W-1$ neighboring pixels are always included in the averaging with weights equal to 1. In the proposed method $f(i)$ is changed to $f(i, \sigma_\eta)$ that is noise level adaptive as well. The input-output relationship of the proposed method becomes Eq. 3

$$I^o_{(n,m)} = \frac{w_{\sigma_\eta} I_{(n,m)} + \sum\limits_{a,b \in \{-1,0,1\}} f(i, \sigma_\eta) I_{(n+a,m+b)}}{w_{\sigma_\eta} + \sum\limits_{a,b \in \{-1,0,1\}} f(i, \sigma_\eta)} \quad (3)$$

where $f(i, \sigma_\eta)$ is given by Eq. 4

$$f_{(i,\sigma_\eta)} = \begin{cases} f(i) & : I_{(k,l)} - 2\sigma_\eta \leq I_{(n,m)} \leq I_{(k,l)} + 2\sigma_\eta \\ 0 & : otherwise \end{cases} \quad (4)$$

with $I_{(k,l)}$, the neighboring pixel. The new weighting function in Eq. 4 uses the $2\sigma_\eta$ probability defined in [1] as being the probability of a random variable, in our case the neighboring pixels, being $2\sigma_\eta$'s away from the current pixel. Instead of applying the adaptive averaging filter directionally as proposed in [5], the proposed method applies the sigma filter directionally.

As can be seen from Eq. 4, $\sum_{a,b \in \{-1,0,1\}} I_{(n+a,m+b)}$ defines a 3×3 neighborhood around the center pixel $I_{(n,m)}$. In [5], $f(i)$ will act as a way to select a subset from the 3×3 population by choosing pixels that make up a homogeneous direction. $f(i, \sigma_\eta)$ will refine the selection by creating a subset of $f(i)$ that includes only pixels satisfying the $2\sigma_\eta$ criterion. Fig. 2 shows the change in the shape of the kernel at different noise level ranges between [5] and the proposed method. In Fig. 2(a,b), all pixels in the homogeneous direction are assumed to belong to one population and are included in the averaging process. Fig 2(c,d) illustrates how the proposed method adapts kernel size and shape to frame and noise characteristics. To achieve higher gain

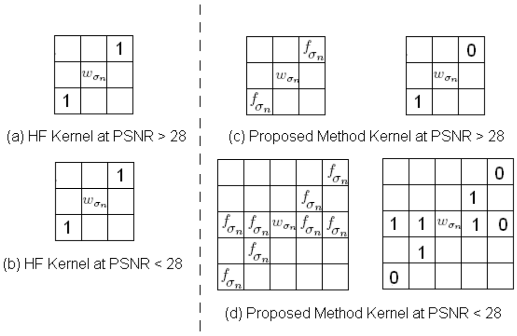

Fig. 2. Fixed HF[5] versus proposed variable kernel shape and direction at different noise levels.

for noisy frames, i.e., noise level below $t_\eta = 28$, the two most homogeneous directions are used. The higher gain is due to the increased averaging. The justification of using the second most homogeneous direction is to increase the number of samples used in the averaging process, thus increasing the probability of the mean giving a closer approximation to the noise free pixel value. To combat blurring, the sigma criterion will ensure that any pixel that is outside of the population defined by the two homogeneous direction gets excluded using the noise-adaptive $2\sigma_\eta$ rejection criterion.

The main reason the proposed method performed better than the SF is the fact that two criteria are used in the proposed method as opposed to one in the SF. The proposed method imposes an extra constraint that pixels have to be spatially grouped in a direction rather than scattered as in the SF. The proposed methods ensures that pixels that randomly satisfy the 2σ probability without actually belonging to the same population as the centered pixels gets excluded. To ensure that the new ideas do not decrease the real-time speed of the filter in [5] due to the variable kernel size, the method is designed to work for $W = 5$ with the two outermost pixels given extra binary weights of 1 or 0 based on the determined kernel size. This way those pixels can be turned off completely, thus effectively reducing the kernel size to $W = 3$ without much complexity.

3. RESULTS

To validate the proposed approach, 8 images (Fig 3(a-h)) and 4 video sequences (Fig. 3(i-l)) were used in simulation. The images were corrupted by noise levels ranging from 20-40 dB PSNR in steps of 5 dB and the video sequences with levels 20-40 in steps of 10 dB. Fig. 4 shows how the threshold of $t_\eta = 28$ was

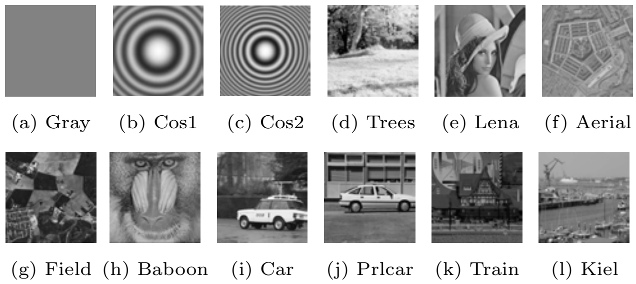

Fig. 3. Images and Video Sequences Used in Simulation.

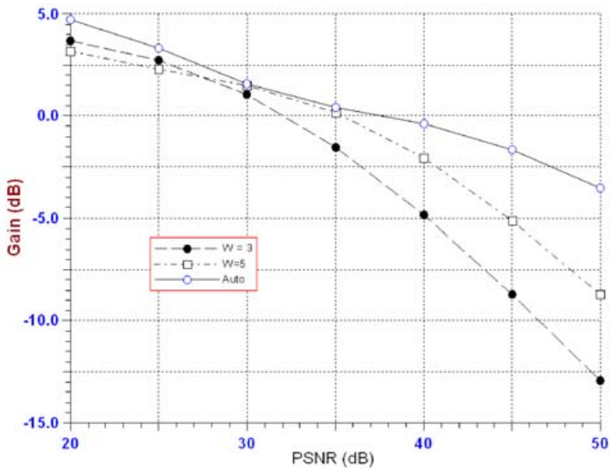

Fig. 4. Average gain at different noise levels with $W = 3$ and $W = 5$ variants of HF[5] versus window adaptation by the proposed filter.

determined at the intersection of the fixed $W = 3$ and $W = 5$ curves. A comparison of the gain achieved by the proposed and reference methods at different noise levels is shown in Fig. 5 (for the test images) and Fig. 6 (for the test video sequences). While the WF does achieve higher gain for 20 dB PSNR, it does not work as well for other noise levels despite being noise level adaptive. The RSF[2] achieves better gain than the standard SF. The proposed method outperforms both the SF and RSF methods. Overall it also outperform the WF in preserving clean images while still giving good gain for noisier images.

Table 1 shows the average time needed by each method to process a 512×512 frame when implemented using C++ under an Intel(R) Xeon(TM) CPU 2.40GHz machine running Linux. As can be seen, the proposed method remains suitable for real-time video processing compared to [5] despite the added modifications. It is also faster than the SF, RSF, and WF methods, e.g.,

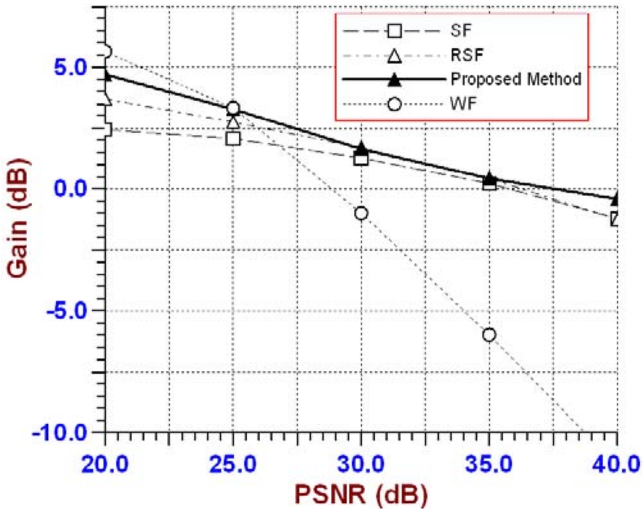

Fig. 5. Average gain achieved by proposed method, RSF[2], SF[1] and WF[4] when applied to Images.

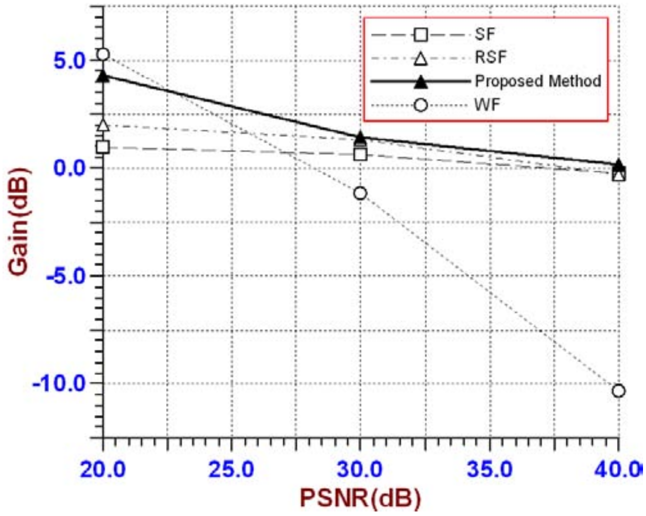

Fig. 6. Average gain achieved by by proposed method, RSF[2], SF[1] and WF[4] when applied to Videos.

it is 2.4 times faster than WF.

Table 1. Time Complexity Comparison.

Algorithm	(seconds/512x512 frame)
HF[5]	0.13
Proposed Method	0.16
SF[1]	0.20
RSF[2]	0.25
WF[4]	0.37

4. CONCLUSION

This paper proposed a new method for spatial noise reduction of white Gaussian noise that is both fast and good at preserving image contents. The method uses the sigma filter for directional smoothing along homogeneous areas and not across edges. The proposed method achieves a maximum gain of 4.8 dB by increasing the filtering through expansion of kernel size and using the two homogeneous directions. So it adapts its behavior to the image and noise content. Simulations show that the proposed method achieves higher noise reduction gain that the four reference methods. It is also more suitable for real-time video applications.

5. REFERENCES

[1] J. Lee, "Digital image smoothing and the sigma filter," *Computer Vision, Graphics and Image Processing*, vol. 24, pp. 225–269, 1983.

[2] T. Kwaaitaal-Spassova G. de Haan and M. Larragy, "Television noise reduction ic," *,IEEE Transactions on Consumer Electronic*, vol. 44, pp. 143–154, 1998.

[3] T. Kwaaitaal-Spassova G. de Haan and O. Ojo, "Automatic 2-d and 3-d noise filtering for high-quality television receivers," *International Workshop on Sigma Processing and HDTV*, vol. VI, pp. 221–230, 1996.

[4] L. Jae, *Two-Dimensional Signal and Image Processing*, pp. 536–540, Prentice Hall, Englewood Cliffs, NJ, 1990.

[5] K. Jostschulte and A. Amer, "A new cascaded spatio-temporal noise reduction scheme for interlaced video," *Int. Conf. on Image Processing (ICIP),*, vol. 2, pp. 293–294, Oct 1998.

[6] A. Amer and E. Dubois, "Fast and reliable structure-oriented video noise estimation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, Janurary 2005.