We proposed a novel method of video noise reduction based on the spatial Wiener filter and the temporal filter. In the proposed spatial Wiener filter, both the amount of noise and the size of the mask are taken into consideration. The proposed model has a great capacity to be adaptive in each area in accordance with the amount of noise. In the proposed model, the motion detector is applied to control the noise removal process in accordance with the area's information (i.e., static or movable). More accurately, more noise removal is done in the areas that are potentially still areas and less removal in the areas that are potentially motion areas. The proposed model achieves a maximum gain of 7.6 dB and capacity of conserving the significant image features (e.g., edges). The experimental results demonstrate that the new approach is more efficient than reference methods in terms of noise removal and edges preservation.

1. Introduction

Noise introduced during acquisition of images, broadcasting over analog channels, and encoding or decoding often corrupted the video sequences, which leads to significant degradation of image quality. Hence, highlighting the importance of the role of noise reduction methods of video sequences is needed [1].

Noise reduction is a useful tool to enhance perceptual quality and increase compression effectiveness, in addition to pattern recognition processes [2].

There are many video noise reduction algorithms existing beforehand in the literature. These algorithms can be classified into three categories: first category implements in a spatial domain [3], second category implements in a temporal domain [4, 5], while the third category implements in the combination of a spatial and temporal domain [6, 7].

As an example of spatial filter, Wiener filter is considered as a classical approach for spatial noise filtering [8]. This filter has the capacity to achieve high gain in noise removal. However, it can cause serious damage to the edge of the image during the process of the noise removing, especially in noise-free areas.

To improve the traditional Wiener filter, the authors in [9] proposed a new approach based on multidirectional Wiener filter. The main idea of this filter is to conserve essential structures of the image by choosing only the homogeneous directions for filtering.

On other hand, Yan and Yanfeng in [10] have proposed a noise reduction method based on temporal filtering. The basic idea of this method is to remove the noise in continuous frames. The proposed filter has succeeded in reducing the noise in real video sequence. Nevertheless, this filter suffers from dragging effects on moving objects [11].

Rakhshanfar and Amer [12] proposed temporal video denoising filter based on temporal data blocks. The authors succeeded in reducing the noise and minimizing the blocking artifacts. However, this method has failed to prevent blur edges.

Frames in video sequence are temporally associated. Wherefore in video noise reduction algorithms, the temporal filter should be used in the motion areas of the frames to
diminish the noise extent practicable. However, temporal filter cannot be utilized alone since it may cause blurring in the motion areas. On the other hand, utilizing the spatial filter separately often causes spatial blurring. For that the spatial filter must be used in combination with temporal filter [7].

The authors in [13, 14] have proposed spatiotemporal adaptive filtering to avoid the drawbacks of all of spatial filters and temporal filters.

Zuo et al. [15] introduced a new video denoising method based on the exploitation of the strong spatiotemporal correlations of neighboring frames. In this denoising method, the authors first applied the motion estimation on the previously denoised frames and the current noisy frame. Then, Kalman filter and bilateral filter both were applied on the current noisy frame. Finally, to get a satisfying result, they weighted the denoised frame from Kalman filtering and bilateral filtering.

Maggioni et al. [16] proposed a new framework for the denoising of the videos which were corrupted by random and fixed-pattern noise. This video denoising approach is based on motion-compensated 3D spatiotemporal volumes. To sparsify the data in 3D spatiotemporal transform domain, the authors have leveraged both the spatial and temporal correlations within each volume. After that the adaptive 3D threshold array is used to shrink the coefficients of 3D volume spectrum.

Hong-zhi et al. [17] have addressed the problem of noise reduction in the video sequences. The proposed method which is based on spatial-temporal combination has the ability to discriminate the static regions from the motion regions of video frames. In this method, the temporal bilateral Kalman filtering is performed on the static regions while the spatial bilateral adaptive nonlocal means filtering is performed on the motion regions.

Esche et al. [18] proposed a new adaptive loop filter that only requires small overhead on slice level. In this filter to reconstruct the individual motion trajectory of every pixel in a frame at both encoder and decoder, the temporal information conveyed in the bit stream is used. The temporal information is exploited to perform pixelwise adaptive motion-compensated temporal filtering.

Cong et al. [19] proposed new surveillance video denoising based on hierarchical motion estimation. The main idea of this method lies in tracking matching blocks and filter along the motion trajectory. In this method, the hierarchical motion estimation proceeds from large blocks to small blocks while the corresponding motion vector field is from coarse to fine.

Wang et al. [20] proposed using the depth and texture information jointly in a spatial-temporal domain to develop the spatial-temporal depth filter. By depending on the similarity of pixel vectors, the authors selected the reference pixels of a to-be-filtered pixel in the spatial-temporal domain. Then, just the most relevant pixels are selected to be identified among reference pixels. Finally, the median filter has been selected among the identified pixels to obtain the result for the to-be-filtered pixel.

Lin et al. [21] addressed the effect of the noise on the depth pixels in the image. The authors tried to overcome this problem by using the painting method to remove the noise from object-removed images. Moreover, the holes in the depth image are filled in order to enhance the quality of this image.

Maggioni et al. [22] proposed VBM4D based on the spatiotemporal redundancy characterizing natural video sequences, which represents the state of the art in video denosing. In this filter, the paradigm of nonlocal grouping and collaborative filtering is implemented. The tracking blocks along the trajectories are used to construct the 3D spatiotemporal volumes and the mutually similar volumes are grouped together by stacking them along the 4th D.

Yahya et al. [23] proposed a video denoising method based on total variation (TV) and temporal filtering. In order to minimize the blurring of the edges in this algorithm, both the previously denoised frame and the current noisy frame are filtered by TV algorithm. Then, the TV’s output has been filtered by the temporal filter to get more improvement. Finally, the motion detector and recursive time average are applied for more noise suppression.

Most of the existing spatial-temporal methods have succeeded in removing noise but, unfortunately, they often degrade image quality by staining the free-noise areas.

To avoid this drawback, we propose a motion based on spatial-temporal filtering which takes the areas quality into consideration. In the proposed model, the recursive time averaging will be applied in the areas where the motion has not been detected. This proposed model is able to be adapted in each area depending on the information of the area. More precisely, our spatial-temporal recursive filter is going to play a good role in making noise removal more at the still areas and less at the motion areas. In this way, we can avoid blurring edges.

Our proposed model as illustrated in Figures 1 and 2 passes through three stages; in the first stage, we use the spatial Wiener filter to filter the previous and current degraded frames. The proposed Wiener adaptive filter has the ability to adapt and change in each area in accordance with the amount of noise. In the proposed spatial Wiener filter, the size of the mask is taken into consideration. More accurately, the mask with large size is applied in the areas with high noise levels, while the mask with small size is utilized in areas with low noise levels. In the second stage, we enhance the results of the spatial filter with temporal filter. At the last stage, we use the motion detector and recursive time averaging to improve the temporal filter’s output.

The remainder of this paper is organized as follows. In Section 2, we briefly describe the spatial Wiener filter. The proposed model is described in Section 3. The experimental results are presented in Section 4. Some concluding remarks are outlined in Section 5.

2. Spatial (Wiener) Filter

Suppose that \( y(i, j) \) is the input image corrupted with a white Gaussian noise \( n(i, j) \), with zero mean and variance \( \sigma_n^2(i, j) \). The observed noisy image \( y(i, j) \) is known as the sum of
the original image \(x(i, j)\) and some noise \(n(i, j)\); that is, \(y(i, j) = x(i, j) + n(i, j)\).

The main purpose of the noise reduction algorithms is to restore the image \(R(i, j)\) from the degraded image \(y(i, j)\) of the original image \(x(i, j)\). The most efficient algorithm is that one which has the ability to yield image \(R(i, j)\) so as to be as close as possible to the original image \(x(i, j)\). Wiener filter is based on this principle.

The Wiener filter formulation can be written as

\[
W(i, j) = \frac{\sigma^2_x(i, j)}{\sigma^2_x(i, j) + \sigma^2_y(i, j)} \cdot y(i, j),
\]

(1)

where \(\sigma^2_x(i, j)\) is the variance of the noise over the input image (noisy image) \(y(i, j)\), [24].

3. Proposed Model

As illustrated in Figures 1 and 2, in this section, we propose a new video noise reduction algorithm based on combination of a spatial Wiener and temporal filters and then improve the combination’s output by subjecting it to the test of motion detector.

Spatial Wiener filter is the most common filter in the field of spatial noise filtering. This filter has the ability to reduce noise effectively. But unfortunately, it can cause blurring especially in the areas with low noise levels.

To avoid this blurring, we propose increasing the filtering action in the areas with high noise levels and decreasing it in the areas with low noise levels.

The regulation of the filtering action is carried out by applying the mask size of \(5 \times 5\) in the areas which contain high levels of noise and low image features. On the other hand, we apply the mask size of \(3 \times 3\) in the areas which contain less noise and many image features.

In the areas that contain high levels of noise, the mask size of \(5 \times 5\) will outperform the mask size of \(3 \times 3\) in terms of noise removal. However, applying the mask size of \(5 \times 5\) often leads to blurring especially in the areas that contain less levels of noise. To combine the advantages of both masks, we propose using the threshold \(T1\) to determine which mask should be applied in accordance with the noise level.

More accurately, we apply the mask size of \(5 \times 5\) in the areas with high noise levels in order to reduce the highest amount of noise; however, applying the mask size of \(3 \times 3\) in the areas with low noise levels will play an important role in the conservation of the image edges.

For controlling the filtering action, we use the following function:

\[
f_n(i, j) = \begin{cases} 
\mu > T1 & \text{Mask size of } 5 \times 5, \\
\text{Otherwise} & \text{Mask size of } 3 \times 3,
\end{cases}
\]

(2)

where \(\mu\) is the amount of noise and \(T1\) is an optional test threshold.
Now let $f_n$ denote the $n$th current frame filtered by spatial Wiener filter. The temporally filtered $n$th frame $T_n$ can be expressed as

$$T_n(i, j) = \rho \cdot f_{n-1}(i, j) + (1 - \rho) \cdot f_n(i, j),$$

where

$$\rho(i, j) = \frac{\sigma_n^2(i, j)}{\sigma_n^2(i, j) + \sigma_r^2(i, j) + \sigma(i, j) \cdot \sigma_n(i, j)}$$

and $f_{n-1}$ is the previous frame filtered with spatial Wiener filter.

For further improvement of the temporally filtered frame $T_n(i, j)$, we select the following motion field to control the process of removing noise according to the areas information (static or movable) [25]:

$$M(i, j) = \begin{cases} 0 & |f_n(i, j) - T_n(i, j)| < T, \\ 1 & \text{otherwise}, \end{cases}$$

where $T$ is an optional test threshold.

From the above motion threshold we can observe the following:

(i) In case of $M(i, j) = 0$, the changes over $f_n(i, j)$ and $T_n(i, j)$ at the spatial position $(i, j)$ will be nearly zero; in other words, $f_n \approx T_n$. 

Figure 3: Visual comparison of different algorithms for frame 5 of the Miss America sequence. From left to right and from top to bottom: original frame, noisy frame ($\sigma_n = 15$), result of [26] algorithm, result of [27] algorithm, result of [28] algorithm, and result of the proposed algorithm.
In case of $M(i, j) = 1$, the changes over $f_n(i, j)$ and $T_n(i, j)$ at the spatial position $(i, j)$ will be significant. So the noise removal should be stopped in this case; otherwise, the significant information of the image in this area will be removed which will lead to a blurry image.

In case of $M(i, j) = 0$, we enforce recursive time averaging by taking the weighted average of $f_n$ and $T_n$ as follows:

$$ R(i, j) = \begin{cases} 
\beta(i, j) \cdot f_n(i, j) + (1 - \beta(i, j)) \cdot T_n(i, j), & M(i, j) = 0; \\
T_n(i, j), & M(i, j) = 1.
\end{cases} \tag{6} $$

Here the weighting factor $\beta(i, j)$ is

$$ \beta(i, j) = \frac{\sigma_n^2(i, j)}{\sigma_n^2(i, j) + \sigma_r^2(i, j) + \sqrt{\sigma_n(i, j) \cdot \sigma_r(i, j)}}. \tag{7} $$

where $\sigma_n^2(i, j)$ is the noise variance and $\sigma_r^2(i, j)$ is the residue variance.

### 4. Experimental Results

In this section, we compare the proposed approach with the algorithms in [26–28] in terms of the visual quality of denoising image and Peak Signal-to-Noise Ratio (PSNR) according to (8).

For PSNR, we use the following formula:

$$ \text{PSNR} = 10 \log_{10} \frac{255^2 \cdot N_x \cdot N_y}{\sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [D(i, j) - O(i, j)]^2}, \tag{8} $$

where $N_x$ and $N_y$ are the numbers of pixels horizontally and vertically, respectively, and $D(i, j)$ and $O(i, j)$ are the denoised frame and original frame, respectively.

To assess the superiority of our new model, we take the common video sequences, that is, Miss America, Salesman, Flower Garden, and Foreman debased by different types of noise, and filter them by the algorithms in [26–28] and the new proposed algorithm.

The experimental results are shown in Figures 3–6 where these figures illustrate: original frame, noisy frame, result of the algorithm in [26], result of the algorithm in [27], result of the algorithm in [28], and result of the new algorithm, respectively.

In our experiments, the original frames in Figures 5 and 6 are corrupted by Speckle noise, while in Figures 3 and 4 they were degraded by the additive white Gaussian noise (AWGN).

From Figures 3–6, we can see that all of the algorithms in [26–28] leave serious noise without removing, especially in the video sequence which is filtered by the algorithm in [26], unlike in the proposed model, where the noise is almost removed, and at the same time the edges are preserved.

The quantitative results of the four models are shown in Tables 1–4. From Tables 1 and 2, we can observe that the proposed model achieves 7.6 dB gain better than that of the algorithm in [26], 4.48 dB gain better than that of the algorithm in [27], and 4.61 gain better than that of the algorithm in [28], while in Tables 3 and 4 the proposed model achieves 4.92 dB gain better than that of the algorithm in [26], 2.13 dB gain better than that of the algorithm in [27], and 3.96 gain better than that of the algorithm in [28].

From Figures 7 and 9, we can see that the algorithm in [26] achieves higher PSNR than that in [27] in low noise level, while in Figures 8 and 10 the algorithm in [27] outperforms that of [26] in all noise levels in terms of PSNR. Nevertheless,
Figures 7–10 show that the proposed algorithm outperforms all of the algorithm in [26], the algorithm in [27], and the algorithm in [28] in terms of PSNR.

Experimental results demonstrate superiority of the new approach in terms of edge preservation and noise suppression, due to its ability to control the amount of noise removal according to areas information (static or movable).

As shown in Figures 1 and 2, the algorithm proposed in this paper can be carried out as follows.

Step 1. Filter out the previous frame by (2) as follows:

(a) If $\mu > T_1$, filter out the previous frame by the mask size of $5 \times 5$.

(b) If $\mu < T_1$, filter out the previous frame by the mask size of $3 \times 3$.

Step 2. Filter out the current frame by (2) as follows:

(a) If $\mu > T_1$, filter out the current frame by the mask size of $5 \times 5$.

(b) If $\mu < T_1$, filter out the current frame by the mask size of $3 \times 3$.

Step 3. Subject the outputs of (2) to (3).

Step 4. Apply the motion detector (5) to the output of temporal filter (3) in order to discriminate between static and motion areas.
5. Conclusion

This paper presented a novel model of video noise reduction based on spatial Wiener filter and temporal filter. The proposed algorithm has high ability to remove the noise very efficiently and at the same time maintain the important image features. The application of the proposed spatial Wiener filter in this paper is based on the amount of noise in each area, where the mask with large size is applied in the areas with high noise levels, while the mask with small size is applied in the areas with low noise levels. In the proposed temporal filter, the motion detector is applied to control the noise removal process in accordance with the areas quality, where the static areas have been heavily filtered to remove greater amount of noise unlike in the motion areas which are subject to less filtering in order to maintain the features of the image. Numerical experiments with four different video sequences and various levels of Speckle noise and white Gaussian noise show that our proposed model has achieved higher noise removal gain as compared with the algorithms in [26–28]. To emphasize the superiority of the proposed algorithm, we use a Peak Signal-to-Noise Ratio (PSNR) as a quantitative measurement, while the visual quality is used as a qualitative measurement.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.
Figure 6: Visual comparison of different algorithms for frame 5 of Foreman sequence. From left to right and from top to bottom: original frame, noisy frame (noise variance = 0.04), result of [26] algorithm, result of [27] algorithm, result of [28] algorithm, and result of the proposed algorithm.

Figure 7: PSNR (dB) graph of [26–28] and new algorithms for various Gaussian noises for Miss America video sequence.
Figure 8: PSNR (dB) graph of [26–28] and new algorithms for various Gaussian noises for Salesman video sequence.

Figure 9: PSNR (dB) graph of [26–28] and new algorithms for various Speckle noises for Miss America video sequence.

Figure 10: PSNR (dB) graph of [26–28] and new algorithms for various Speckle noises for Salesman video sequence.
Acknowledgments

The project is supported by the NSFC-Guangdong Joint Foundation (Key Project) under Grant no. U1135003 and the National Natural Science Foundation of China under Grant nos. 61472466 and 61072227.

References


