SPATIO-TEMPORAL SEPARABLE DATA-DEPENDENT WEIGHTED AVERAGE FILTERING FOR RESTORATION OF THE IMAGE SEQUENCES

Kouji Miyata           Akira Taguchi

Department of Electrical and Electronic Engineering
Musashi Institute of Technology
Setagaya-ku, Tokyo 158-8557, Japan

ABSTRACT

The restoration of the image sequences corrupted by the additive noise is important for getting the high quality image and the high compression ratio of the image sequences. In this paper, we propose a novel restoration method for the image sequences corrupted by the Gaussian noise. Conventional methods use the spatio-temporal filtering after the motion compensation. However, the accuracy of the motion compensation on degraded image is not satisfied. Therefore, in the proposed method, we divide spatio-temporal filter into spatial filter and temporal filter in order to achieve the high noise reduction. The spatial filter is regarded as the pre-filter for the motion compensation. We also study the suitable filtering scheme after the motion compensation.

1. INTRODUCTION

Recently, we have the remarkable developments in multimedia products. Especially, image sequences are useful for communication. However, image sequences, which are obtained by the analog equipment, are degraded by adaptive Gaussian noise. The restoration of degraded image sequences is necessary for not only improving the image quality but also pre-processing for the image compression [1].

Since spatial filtering can preserve the motion of images sequences, spatial filtering can also use for restoring the image sequences corrupted by additive noise. However, this method can’t remove additive noise sufficiently and causes the flicker. Therefore, spatio-temporal filtering is used for restoring the image sequences. In order to preserve the motion of the image sequences, the spatio-temporal filter is processed after the motion compensation. If the accuracy of the compensation is insufficient, we can’t obtain the high quality image sequences. In the conventional methods [1], in order to improve the filtering performance, researchers mainly studied about the motion estimator, which is robust for additive noise. The Boyce’s motion estimator is one of robust estimators for additive noise. However, we apply the Boyce’s motion estimator for noisy image sequences, the accuracy is not satisfied, especially, in the low signal to noise ratio condition.

In this paper, we propose a novel restoration method. In order to overcome the defect of the conventional methods, we divide the spatio-temporal filtering into spatial filtering and the temporal filtering. The spatial filter regards as the pre-filter for the motion compensation. After the motion compensation, we apply the temporal filter. It is clear that the accuracy of the motion compensation of the proposed method is improved because the Gaussian noise is reduced on the output of the spatial filtering. And the proposed method is almost same computational effort as the conventional methods. Moreover we also study the suitable filtering scheme after the motion compensation.

2. SPATIO-TEMPORAL SEPARABLE DATA-DEPENDENT WEIGHTED AVERAGE FILTER

The proposed method has the three steps for restoration.

Step 1: Spatial filtering

The first step is the spatial filtering process, which is the pre-processing for the motion compensation (i.e., Step 2). It is necessary for the spatial filtering to preserve edges and details of each frame of the input image sequences. The data-dependent weighted average (DDWA)[4],[5] filter is chosen for this task. The DDWA is defined by

\[
\hat{x}(i, j, k) = \sum_{p=-P}^{P} \sum_{q=-Q}^{Q} w(p, q) \cdot x(i + p, j + q, k)
\]

(1)

\[w(p, q) = W_T \cdot K(i, j, k) \cdot E_{p,q} \cdot D_{p,q} + 1\]

(2)
where $W_T$ is constant and around 200 is suitable[5].

$K(i, j, k)$ is the local statistics value derived by

$$K(i, j, k) = \sigma^2(i, j, k)/\left(\sigma^2(i, j, k) + \sigma_n^2\right)$$

(3)

$\sigma^2(i, j, k)$ is the estimated variance of original image which is given by

$$\sigma^2(i, j, k) = \max[\text{Var}(i, j, k) - \sigma_n^2, 0]$$

(4)

$\text{Var}(i, j, k)$ and $\sigma_n^2$ are the local variance in the filter window and the variance of the additive noise, respectively.

$K(i, j, k) \approx 1$ means the filter window is located at the edge or detail region, then, $w(p, q)$ is determined by $E_{p,q}$ and $D_{p,q}$ for preserving the edges and details. On the other hand, $K(i, j, k) \approx 0$ means that the filter window is located at the flat region, then, all $w(p, q)$ set 1 (i.e., mean filter). It is well known that the mean filter is the optimal filter for reducing the Gaussian noise [6].

$E_{p,q}$ is the difference information between the current pixel and the neighborhood pixels which is given by

$$E_{p,q} = \begin{cases} 1 & \text{if} \ |x(i, j, k) - x(i+p, j+q, k)| < \mu \cdot \sigma_n \\ 0 & \text{otherwise} \end{cases}$$

(5)

If $x(i, j, k)$ and $x(i+p, j+q, k)$ are assumed to be belonged to the same flat region, then $E_{p,q}$ is set as 1.

And, the suitable value of $\mu$ is 2 or 3.

$D_{p,q}$ is the distance information.

$$D_{p,q} = 1 - (p^2 + q^2) / (P^2 + Q^2)$$

(6)

This information is the function which has the normalized distance between $x(i, j, k)$ and $x(i+p, j+q, k)$. Generally, image signal is non-stationary. $D_{p,q}$ is lager, the relationship between $x(i, j, k)$ and $x(i+p, j+q, k)$ is weaker. We can consider this property for filtering by using the distance information.

Step 2 : Motion compensation

In the second step, we estimate and compensate the motion on the output of the first step using the boyce’s motion estimator [3]. In order to improve the accuracy of the motion estimation, we introduced spatial filtering in step 1, however, the output of the pre-filter is not reduced the noise sufficiently. Therefore, we use Boyce’s estimator. This estimator is robust for the influence of Gaussian noise.

At the first, the index of the similarity between the current frame and the reference frame is defined as

$$MAD = \sum_{n=N}^{N} \sum_{m=-N}^{N} |\hat{x}(i+n, j+n, k) - \hat{x}(i+n + d_{ij}^w(i, j), j+n + d_{ij}^w(i, j), k+l)|$$

(7)

In this equation, we assume that the macro block of the estimation is the $(2N+1) \times (2N+1)$ square region.

$d=[d_{ij}^l(i,j), d_{ij}^r(i,j)]$ is the estimated motion vector, which minimizes the index (7). In addition, when the condition as $MAD_0 < \beta \cdot MAD_{\text{min}}$ or $MAD_0 < \rho \cdot MAD_{\text{noise}}$ (8) is satisfied, we assume that the image in the macro block is stopped. $MAD_0$ is calculated by Eq.(7) with $[d_{ij}^l(i,j), d_{ij}^r(i,j)]^T = [0, 0]^T$. $MAD_{\text{noise}}$ is the value which is caused by the only Gaussian noise and derived by Ref.[5].

Step 3 : Temporal filtering

In the third step, basically, we restore the image sequences by using the temporal filtering. We study an appropriate filtering after the motion compensation, in this paper.

The output of the third step is given by

$$y(i, j, k) = \frac{\sum_{l=1}^{L} \sum_{j=1}^{P} \sum_{q=1}^{Q} w_f(p, q, l) \cdot \hat{x}(s + p, t + q, k + l)}{\sum_{l=1}^{L} \sum_{j=1}^{P} \sum_{q=1}^{Q} w_f(p, q, l)}$$

(9)

where $s = i + p + d_{ij}^l(i, j), t = j + q + d_{ij}^r(i, j), \hat{K}(i, j, k)$ is the spatial local statistics which is derived by using the output of the first step same as Eq.(3). Switching of our method is controlled by the value of $\hat{K}(i, j, k)$. If the $\hat{K}(i, j, k)$ is less than $\alpha$, we assume the filter window is located at the flat region. Thus, we can use the spatio-temporal filter for restoration, since the filtering does not degrade image signal. On the other hand, in order to preserve edge and motion, we use the temporal filter, when $\hat{K}(i, j, k)$ is larger than $\alpha$.

The weight of two filters as $w_f(p, q, l)$ and $w_e(l)$ are given by

$$w_f(p, q, l) = W_T \cdot \hat{K}(i, j, k) \cdot E_{p,q,l} \cdot D_{p,q,l} + 1$$

(10)

$$w_e(l) = W_T \cdot E_{p,q,l} + 1$$

(11)
where $\tilde{K}(i, j, k)$ is spatio-temporal local statistics. $E_{p,q,l}$ and $D_{p,q,l}$ are given by same idea of Eq.(5) as

$$E_{p,q,l} = \begin{cases} 1 \text{ if } |x(i, j, k) - x(s, t, k + l)| < \hat{\sigma}_n \, \sigma(n, m, k) \\ 0 \text{ otherwise} \end{cases}$$

$$D_{p,q,l} = 1 - \sqrt{p^2 + q^2 + \hat{\sigma}_n^2} / \sqrt{P^2 + Q^2 + \hat{E}}$$

where $\hat{\sigma}_n$ is the noise variance of the output image sequences of first step.

3. THE RESTORATION PERFORMANCE

Two remarkable points of the proposed method are as follows:

(I) Motion estimation and compensation are performed by the noise reduced image.

(II) Introducing the switching filter after the motion compensation.

These points can be realized by dividing the spatio-temporal filtering into the spatial filtering and the temporal filtering.

In this section, we show the effectiveness of the two remarkable points of the proposed method through the comparing with several conventional methods.

3.1. Accuracy of the restoration

We compare the accuracy of the several restoration methods as

a) 2D-DDWA [4]

b) Video-DDWA [5]

c) Motion compensation using degraded image sequences + 3D-DDWA [5]

d) Motion compensation using the output of 2D-DDWA filter + 3D-DDWA

e) Proposed-1
   (only spatio-temporal filter (the upper equation of Eq.(8)) is used in Step 3)

f) Proposed-2
   (only temporal filter (the under equation of Eq.(8)) is used in Step 3)

g) Proposed-3 (Switching filter is used in Step 3)

c) and d) are same methods expect for the motion estimation method. The proposed method is the same concept of d). We can see the effectiveness of the switching filter in Step 3, comparing with (e), (f) and (g).

We evaluate 7 methods by using the index $MSE-R$ which is given by

$$MSE - R = \frac{\sum \sum |y(n, m, k) - s(n, m, k)|^2}{\sum \sum |x(n, m, k) - s(n, m, k)|^2}$$

where $s(n, m, k)$ is original signal. Two image sequences shown in Fig.1 are used for simulation. These two images are added to zero mean Gaussian noise with variance 900.

In Fig.2, we show the restoration performance in the case of $\sigma_n^2 = 900$. Comparing with the results c) and d), the pre-filtering for the motion compensation which is the concept of the proposed method is effective for restoration of images. Furthermore, comparing with (d) and (g), we can understand that the proposed separable filtering scheme is suitable for restoration. From e), f) and g), we can
Fig.3 Research of the threshold

understand the effectiveness of the switching filter in step 3. In the comparison with e) and f), we can observe the difference between the two kind of filtering in Eq.(9). We think that this difference is caused by the characteristic of each image. The SALESMAN has a lot of detail regions, for example background region, on the other hand, the CLEA has monotonous image sequences. The performance of the proposed filter (i.e., Proposed-3) is same as the proposed-1 and proposed-2 in the case of the CLEA and SALESMAN, respectively. These results mean that the switching of the Step 3 of the proposed method is well worked.

We should study how to set the threshold $\alpha$ of Eq.(9). From Fig.3, we should set $0.0 \leq \alpha \leq 0.4$, then in this paper we set $\alpha = 0.0$. This value dose not fit for the CLEA, however, the proposed method shows higher performance than any other conventional methods for CLEA (see Fig.2 (B)).

At the last, we evaluate visual quantity. In Fig.4 (A)(C) are the conventional methods. We can see from Fig.4 (A), spatial filtering is not effective for noise reduction. We emphasize that the proposed filter can preserve the motion and remove noise sufficiently.

4. CONCLUSION

In this paper, we propose a restoration method for the image sequences, which were corrupted by the Gaussian noise. In the conventional methods, which used the motion estimator, the filtering performance was affected by the noise. The accuracy of the motion estimation is low, because we estimate the motion by using the corrupted image sequences. Therefore we proposed a spatio-temporal separable filtering, each filtering is data-dependent type. Before the motion estimation, spatial filtering is performed. This spatial filter is the pre-filter for the motion estimation, which can realize the robust motion compensation. After the motion compensation, we introduce the switching filtering scheme, which realize the preserving the motion and the noise reduction at the same time. High quality results are obtained by the proposed method. We show the high performance of the proposed method through many examples.

REFERENCES