MOTION COMPENSATED ENHANCEMENT OF NOISY IMAGE SEQUENCES

Dimitrios S. Kalivas and Alexander A. Sawchuk

University of Southern California
Signal and Image Processing Institute
Los Angeles, CA 90089-0272

ABSTRACT

A very important problem in image sequence processing is noise suppression and image enhancement. In this paper we present a motion compensated image sequence enhancement algorithm. First, a combined segmentation and motion estimation algorithm is employed. Then a temporal or a spatiotemporal low-pass filter is applied. We present mean and median filters as low-pass filters. The temporal filtering is performed over the motion path of each pixel, which is provided by the motion estimation algorithm. The spatial filtering does not blur the boundaries of the moving objects because the boundary locations are provided by the segmentation algorithm. The performance of the combined algorithm is examined using computer generated and real image sequences corrupted by additive white gaussian noise.

1. INTRODUCTION

Image sequence analysis is an area of image processing and computer vision which has attracted the interest of many researchers during recent years. Time sampling of a time-varying scene produces a sequence of 2-D images. A very important problem in image sequence processing is noise suppression and image enhancement. Although the enhancement of single images has been extensively studied, there has been only a limited study of image sequence enhancement [1].

The intensity of the pixels is spatially and temporally correlated. The temporal correlation is usually stronger. A temporal low-pass filter can be used to suppress the noise. This filter reduces the variance of the noise but it also blurs the moving parts of the scene. Although a temporal low-pass filter is usually the best candidate, because of memory constraints, a spatiotemporal filter is sometimes more desirable. Such a filter will also generally cause degradation by blurring the edges.

In this paper, we present a motion compensated algorithm for the enhancement of noisy image sequences. First, a combined segmentation and motion estimation algorithm is employed [4, 5, 3]. This algorithm segments the images into moving and stationary components and estimates the motion of the moving parts. Then a temporal low-pass filter is applied. For a stationary pixel the filtering is performed over a set of pixels with the same space coordinates. For a moving pixel the filtering is performed over its motion path, which is provided by the motion estimation part, so that the moving parts are not blurred. We present mean and median filters as low-pass filters. We examine the performance of mean filters for different SNR (Signal-to-Noise Ratio) and different kinds of motion. Mean filtering is more effective in the case of white gaussian noise and median filtering is more effective in the case of salt-and-pepper noise and burst noise. When using spatiotemporal filters, the spatial filtering is not performed over the boundaries because their positions are provided by the segmentation part, therefore, the edges are not blurred.

2. MODELS

2.1 Image Sequence Model

Let $f(i,j,k)$ be the original sequence of 2-D images. The first two arguments i and j are the space coordinates and take values $1,2,\ldots,N$. The third argument k is the time coordinate and takes integer values $1,2,\ldots$. The observed image sequence $g(i,j,k)$ is given by

$$g(i,j,k) = f(i,j,k) + n(i,j,k) \tag{1}$$

where $n(i,j,k)$ is a zero-mean gaussian white noise process.

Since we are interested in extracting the moving object, we express $f(i,j,k)$ in terms of the object and background as follows

$$f(i,j,k) = t(i,j,k)\lambda(i,j,k) + b(i,j,k)[1-\lambda(i,j,k)] \tag{2}$$

where $t(i,j,k)$ and $b(i,j,k)$ are two statistically independent stochastic processes corresponding to the object and the background respectively. The function $\lambda(i,j,k)$ is called the object indicator function and is defined as follows

$$\lambda(i,j,k) = \begin{cases} 1 & \text{if the point } (i,j,k) \text{ is a object point} \\ 0 & \text{otherwise} \end{cases}$$

This model is a generalization of the model introduced by Nahi and Lopez-Mora [6].

This model can be also used in a multi-object environment. In this case (2) will be modified in the following way

$$f(i,j,k) = \sum_{l=1}^{M} t_l(i,j,k)\lambda_l(i,j,k,) + b(i,j,k)[1 - \sum_{l=1}^{M} \lambda_l(i,j,k)] \tag{3}$$

where $t_l(i,j,k)$ $(l = 1,2,\ldots,M)$ and $b(i,j,k)$ are independent stochastic processes. M is the number of the objects and $t_l(i,j,k)$ is the process corresponding to the l_{th} object.

CH2847-2/90/0000-2121 $1.00 © 1990 IEEE

2.2 2-D Motion Model

We describe the motion of an object pixel in the time interval $[k, k+1]$ by the linear affine model

$$\vec{r'} = C(k)\vec{r} + \vec{s} \tag{4}$$

where $\vec{r}$ is the pixel position at time k, $\vec{r'}$ is its position at time $(k+1)$,

$$C(k) = \begin{pmatrix} c_{11}(k) & c_{12}(k) \\ c_{21}(k) & c_{22}(k) \end{pmatrix},$$

$$\vec{s}(k) = \begin{pmatrix} s_1(k) \\ s_2(k) \end{pmatrix},$$

and $c_{11}(k), c_{12}(k), c_{21}(k), c_{22}(k), s_1(k)$ and $s_2(k)$ are the motion parameters. This model describes exactly the linear motion of the object pixels and gives a very good approximation in the case of a nonlinear motion.

3. IMAGE SEQUENCE ENHANCEMENT WITHOUT MOTION COMPENSATION

There are three kinds of smoothing filters which can be used for the enhancement of a noisy image sequence: spatial filters, temporal filters, and spatiotemporal filters. A spatial filter requires much less memory capacity than the temporal filters but the time correlation of the intensity values of the pixels are much stronger than the spatial correlation. A spatial filter blurs the edges of objects and a temporal filter without motion compensation blurs the moving objects. If there are no constraints on memory capacity, a temporal filter is the best candidate. If there are some constraints in memory capacity, a spatiotemporal filter is the best candidate by necessity. In any case, a spatial filter is the worst choice.

The most commonly used smoothing filters are the mean and median filters. We suppose that the smoothing filter operates on M pixels. Then for white gaussian noise mean filtering reduces the noise variance by a factor M and median filtering reduces the noise variance by a factor of $2M/\pi$ [2]. However, for salt-and-pepper noise and burst noise the median filtering behaves better. Another interesting point concerning spatial filtering is that mean filtering blurs the image edges more than median filtering.

3.1 Temporal Filtering

A temporal filter operates on a time neighborhood of a pixel and is defined by

$$\hat{f}(i,j,k) = \mathcal{T}\{g(i,j,n); n \in T\} \tag{5}$$

where T is a suitably chosen time window.

Temporal Mean Filtering

We assume that T is a time window centered at k and having length equal to $2M+1$. Then the mean filter is given by

$$\hat{f}(i,j,k) = \frac{1}{2M+1} \sum_{n=k-M}^{k+M} g(i,j,n). \tag{6}$$

Temporal Median Filtering

The median filter is given by

$$\hat{f}(i,j,k) = \mathcal{MEDIAN}\{g(i,j,k); n \in [k-M, k+M]\}. \tag{7}$$

3.2 Spatiotemporal Filtering

A spatiotemporal filter operates on a spatiotemporal neighborhood of a pixel and is defined by

$$\hat{f}(i,j,k) = \mathcal{ST}\{g(l,m,n); (l,m) \in W, n \in T\} \tag{8}$$

where W and T are suitably chosen space and time windows respectively.

Spatiotemporal Mean Filtering

We assume that T is a time window of length $2M+1$ centered at k and W a spatial window containing L pixels centered at (i,j). The mean filter is given by

$$\hat{f}(i,j,k) = \frac{1}{(2M+1)L} \sum_{n=k-M}^{k+M} \sum_{(l,m)\in W} g(l,m,n). \tag{9}$$

Spatiotemporal Median Filter

The median filter is given by

$$\hat{f}(i,j,k) = \mathcal{MEDIAN}\{g(l,m,n); (l,m) \in W, n \in T\}. \tag{10}$$

4. EDGE PRESERVING MOTION COMPENSATED IMAGE SEQUENCE ENHANCEMENT

Temporal filtering always blurs the moving objects and spatial filtering always blurs the edges of the objects. These two image degradations can be avoided or at least significantly reduced if the motion of the moving objects and the positions of their boundaries are known. This hypothesis indicates the significance of our combined segmentation and 2-D motion estimation method in an edge preserving motion compensated image sequence enhancement algorithm.

We suppose that we want to enhance the noisy image sequence $g(i,j,k)$. Applying our combined segmentation and 2-D motion estimation algorithm we get the estimated object indicator function $\hat{\lambda}(i,j,k)$ and the estimated motion parameters $\hat{c}_{11}(k)$, $\hat{c}_{12}(k)$, $\hat{c}_{21}(k)$, $\hat{c}_{22}(k)$, $\hat{s}_1(k)$, and $\hat{s}_2(k)$.

4.1 Temporal Filtering

A pixel (i,j,k) is considered as an object pixel if

$$\hat{\lambda}(i,j,k) = 1$$

and as a background pixel if

$$\hat{\lambda}(i,j,k) = 0.$$

An object pixel, which was at position (i,j) at time k, will be or was at position (x_n, y_n) at time n. A background pixel, which was at position (i,j) at time k, will be or was at the same position (i,j) at any time n except if it will be or was occluded by a moving object.

The spatial coordinates x_n and y_n of an object pixel can be found using the estimated motion parameters. For values of n greater than k or equal to k, x_n and y_n are given by the recursive equation

$$\begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} \hat{c}_{11}(n) & \hat{c}_{12}(n) \\ \hat{c}_{21}(n) & \hat{c}_{22}(n) \end{pmatrix} \begin{pmatrix} x_n \\ y_n \end{pmatrix} + \begin{pmatrix} \hat{s}_1(n) \\ \hat{s}_2(n) \end{pmatrix} \tag{11}$$

where

$$\begin{pmatrix} x_k \\ y_k \end{pmatrix} = \begin{pmatrix} i \\ j \end{pmatrix}.$$

For values of n less than k or equal to k, x_n and y_n are given by the recursive equation

$$\begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} \hat{c}_{11}(n-1) & \hat{c}_{12}(n-1) \\ \hat{c}_{21}(n-1) & \hat{c}_{22}(n-1) \end{pmatrix} \begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} + \begin{pmatrix} \hat{s}_1(n-1) \\ \hat{s}_2(n-1) \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} x_{n-1} \\ y_{n-1} \end{pmatrix} = \frac{1}{D} \begin{pmatrix} \hat{c}_{22}(n-1) & -\hat{c}_{12}(n-1) \\ -\hat{c}_{21}(n-1) & \hat{c}_{11}(n-1) \end{pmatrix} \begin{pmatrix} x_n - \hat{s}_1(n-1) \\ y_n - \hat{s}_2(n-1) \end{pmatrix}$$

where

$$D = \hat{c}_{11}(n-1)\hat{c}_{22}(n-1) - \hat{c}_{12}(n-1)\hat{c}_{21}(n-1).$$

The temporal filter is defined by

$$\hat{f}(i,j,k) = \mathcal{T}_{MC}\left\{g(x_n, y_n, n); \hat{\lambda}(x_n, y_n, n); \hat{C}(n), \hat{\vec{s}}(n); n \in T\right\} \tag{12}$$

where the subscript MC means Motion Compensated. The spatial coordinates x_n and y_n are generally not integers. Then $g(x_n, y_n, n)$ and $\hat{\lambda}(x_n, y_n, n)$ can be approximated by linear interpolation or can be set equal to the corresponding value of the closest pixel.

Temporal Mean Filtering

We assume that T is a time window centered at k and having length equal to $2M+1$. Then the mean filter is given by

$$\hat{f}(i,j,k) = \frac{1}{2M+1} \sum_{n=k-M}^{k+M} g(x_n, y_n, n). \tag{13}$$

Temporal Median Filtering

The temporal median filter is given by

$$\hat{f}(i,j,k) = \mathcal{MEDIAN}\left\{g(x_n, y_n, n); n \in [k-M, k+M]\right\}. \tag{14}$$

4.2 Spatiotemporal Filtering

A motion compensated spatiotemporal filter is defined by

$$\begin{aligned} \hat{f}(i,j,k) &= \mathcal{ST}_{MC}\{g(z_n, v_n, n); \hat{\lambda}(x_n, y_n, n); \hat{C}(n), \hat{\vec{s}}(n); n \in T; \\ &\quad (z_n, v_n) \in W_n\} \end{aligned} \tag{15}$$

where the subscript MC means Motion Compensated, T is a time window containing k and W_n is a spatial window containing (x_n, y_n).

Spatiotemporal Mean Filtering

We assume that T has length $2M+1$ and is centered at k. W_n is spatial window centered at (i_n, j_n) which is the closest pixel to (x_n, y_n). The spatiotemporal mean filter is given by

$$\hat{f}(i,j,k) = \frac{1}{L} \sum_{n=k-M}^{n=k+M} \sum_{(l,m)\in W_n} g(l,m,n)\lambda(l,m,n) \tag{16}$$

where

$$L = \sum_{n=k-M}^{n=k+M} \sum_{(l,m)\in W_n} \lambda(l,m,n).$$

Spatiotemporal Median Filtering

The spatiotemporal median filter is given by

$$\begin{aligned} \hat{f}(i,j,k) &= \mathcal{MEDIAN}\{g(z_n, v_n, n); n \in [k-M, k+M]; \\ &\quad (z_n, v_n) \in W_n\} \end{aligned} \tag{17}$$

5. SIMULATION RESULTS

We examined the performance of the smoothing temporal and spatiotemporal mean filters for different SNR and different kinds of motion using computer generated and real image sequences. Each sequence has 31 frames. Each frame has (256 × 256) pixels. The images were corrupted by additive white gaussian noise. We used mean filters because they reduce the noise variance more than the corresponding median filters. In each sequence we smoothed the $16th$ (middle) image. For convenience, the results for each sequence are shown in one figure.

5.1 Computer Generated Image Sequences

We used two different image sequences. The first sequence n05s1 ($SNR = 0.5$) contains an object which moves in a very complicated way (Figure 1). In order to examine the performance of the smoothing filters in recovering the interior details of moving objects we used a second sequence n05s6 ($SNR = 0.5$). This sequence contains an object which has a translation motion and printed alphanumeric symbols in its interior (Figure 2).

5.2 Real Images

Additional image data was obtained by digitizing a scene consisting of a moving object on a background, and then adding noise to produce data with various *SNRs*. We call this sequence smrs4. The performance of the smoothing filters was examined for $SNR = 0.5$ (Figure 3).

5.3 Discussion

We can compare temporal smoothing with motion compensation to temporal filtering without motion compensation by observing the processed images in the following four figures (top-right: temporal filtering without motion compensation and bottom-left: temporal filtering with motion compensation). In both cases the noise suppression is the same but the moving object is well recovered in the case of motion compensated temporal enhancement and it is totally lost in the case of temporal enhancement without motion compensation.

In the case of motion compensated smoothing we can compare the temporal filtering to spatiotemporal filtering by observing the processed images in the following four figures (bottom-left: temporal filtering and bottom-right: spatiotemporal filtering). The noise suppression is more in the case of the spatiotemporal filtering. Neither filter blurs the boundary details, but the spatiotemporal filter blurs the interior details. Therefore, there is a trade-off between noise suppression and blur of the interior details.

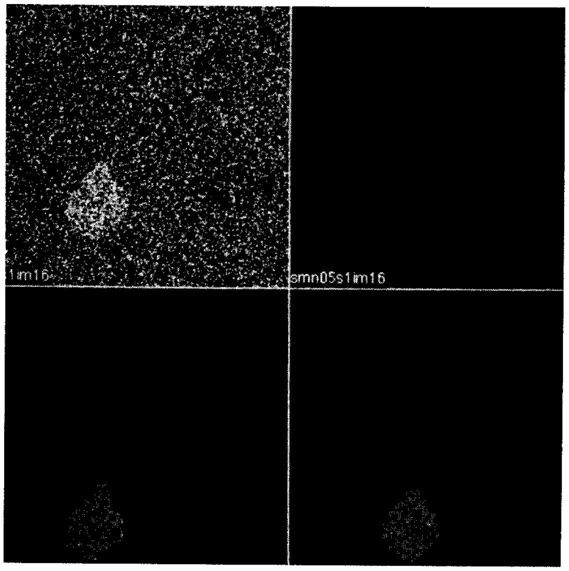

Figure 1: Noise enhancement of sequence n05s1 (SNR = 0.5). Top-left: noisy frame, top-right: simple temporal mean filtering, bottom-left: motion compensated temporal mean filtering, and bottom-right: motion compensated spatiotemporal mean filtering.

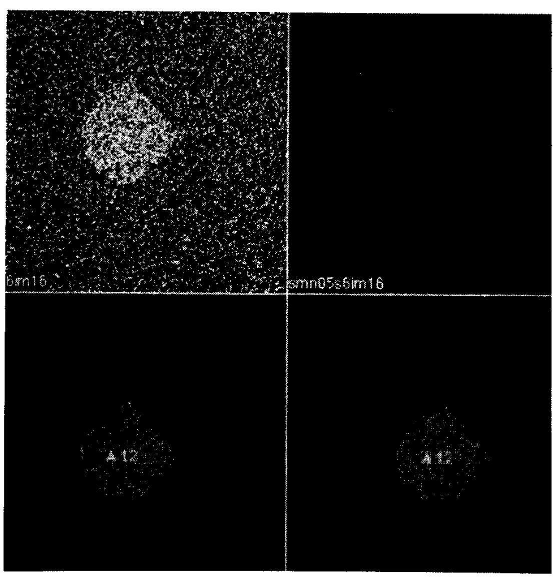

Figure 2: Noise enhancement of sequence n05s6 (SNR = 0.5). Top-left: noisy frame, top-right: simple temporal mean filtering, bottom-left: motion compensated temporal mean filtering, and bottom-right: motion compensated spatiotemporal mean filtering.

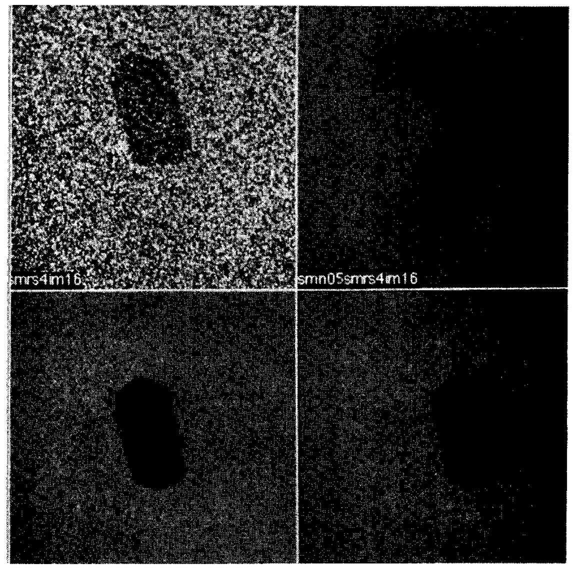

Figure 3: Noise enhancement of sequence n05smrs4 (SNR = 0.5).

6. CONCLUSIONS

In this paper we presented a motion compensated enhancement algorithm which performs very well in a very noisy environment. It does not blur the moving parts, and in the case of spatiotemporal filtering, it does not blur the edges. We presented both mean and median smoothing filters. We examined the performance of mean filters for different SNR (Signal-to-Noise Ratio) and different kinds of motion. Mean filtering is more effective in the case of white gaussian noise and median filtering is more effective in the case of salt-and-pepper noise and burst noise. This algorithm can be used for image sequence quality improvement in industrial, commercial and medical applications as well as a preprocessing step for other operations such as recognition and tracking of moving objects.

References

[1] T. S. Huang, *Image Sequence Analysis*, ch. 4, Springler-Verlag, 1981.

[2] T. S. Huang, *Two-Dimensional Digital Signal Processing II*, Springer-Verlag, 1981.

[3] D. S. Kalivas, *Segmentation and Motion Estimation of Noisy Image Sequences*, Ph.D. dissertation, University of Southern California, December 1989.

[4] D. S. Kalivas and A. A. Sawchuk, "Object Boundary Estimation in Noisy Images," *Proc. CISS 1989, The Johns Hopkins University*, March 1989.

[5] D. S. Kalivas, A. A. Sawchuk, and R. Chellappa, "Segmentation and 2-D Motion Estimation of Noisy Image Sequences," *Proc. ICASSP*, April 1988.

[6] N. E. Nahi and S. Lopez-Mora, "Estimation-Detection of Object Boundaries in Noise Images," *IEEE Trans. on Automatic Control*, Vol. Ac-23, pp. 834–845, October 1978.