

# Noise Reduction of Image Sequences Using Motion Compensation and Signal Decomposition

Richard P. Kleihorst, *Student Member, IEEE*, Reginald L. Lagendijk, *Member, IEEE*, and Jan Biemond, *Fellow, IEEE*

**Abstract**—In this paper, a new spatio-temporal filtering method for removing noise from image sequences is proposed. This method combines the use of motion compensation and signal decomposition to account for the effects of object motion.

Because of object motion, image sequences are temporally nonstationary, which requires the use of adaptive filters. By motion compensating the sequence prior to filtering, nonstationarities, i.e., parts of the signal that are momentarily not stationary, can be reduced significantly. However, since not all nonstationarities can be accounted for by motion, a motion-compensated signal still contains nonstationarities.

An adaptive algorithm based on order statistics is described that decomposes the motion-compensated signal into a noise-free nonstationary part and a noisy stationary part. An RLS filter is then used to filter the noise from the stationary signal. Our new method is experimentally compared with various noise filtering approaches from literature.

## I. INTRODUCTION

**I**IMAGE sequences, consisting of digitized recordings of a time-varying scene, may be corrupted by noise for various reasons, such as imperfections of the scanner, transmission, or recording medium. Noise filtering not only improves the visual quality but also increases the performance of subsequent image processing tasks, such as coding, analysis, or interpretation.

The observed noisy sequence  $g(i, j, k)$  is modeled as

$$g(i, j, k) = f(i, j, k) + n(i, j, k) \quad (1)$$

where  $f(i, j, k)$  is the original signal and  $n(i, j, k)$  is the observation noise. The indices for horizontal, vertical, and temporal directions are  $i$ ,  $j$ , and  $k$ , respectively. The noise is assumed to be independent of  $f(i, j, k)$ , spatio-temporally invariant, and uncorrelated. In this paper, no assumption is made about the specific probability density function of the noise. The purpose of filtering is to estimate  $f(i, j, k)$  from its noisy version  $g(i, j, k)$ .

Most methods for image sequence filtering operate in the temporal direction [1]–[3] to avoid introducing artifacts, such as the blurring of object boundaries in the spatial domain, and to exploit the high temporal correlation in stationary regions. This means that the sequence is regarded as a set of 1-D time series or pixel trajectories  $g(k)$  for every position  $i, j$ .

Manuscript received July 30, 1992; revised January 5, 1994. The material in this paper is based on work supported by NATO under Grant 0103/88. The associate editor coordinating the review of this paper and approving it for publication was Dr. Hsueh-Ming Hang.

The authors are with Delft University of Technology, Department of Electrical Engineering, Information Theory Group, Delft, The Netherlands. IEEE Log Number 9408202.

Movements in the scene yield temporal edges. Therefore, the temporal signal  $g(k)$  cannot be regarded as a stationary signal. Since most linear stochastic filters are designed to work on stationary signals, they blur moving objects. Nonlinear or adaptive filters can avoid these blurring effects.

A popular adaptive filter structure used for image sequence filtering is [4]

$$\hat{f}(k) = \hat{f}(k-1) + \alpha(k)[g(k) - \hat{f}(k-1)]. \quad (2)$$

Here  $\alpha(k)$  controls the amount of filtering and is typically determined by a function of  $|\hat{f}(k-1) - g(k)|$ . The signal dependency of  $\alpha(k)$  makes the filter nonlinear. More recently, Arce [5], Alp and Neuvo [6], and Kleihorst *et al.* [7] employed spatio-temporal order statistic (OS) filters, such as median filters that preserve spatio-temporal edges to deal with the nonstationary signals.

Another way to take temporal edges into account is to perform motion estimation and then filter along the motion trajectories as done by Sezan *et al.* [3], Chan and Sullivan [8], Katsaggelos *et al.* [9], [10], Dubois and Sabri [2], Dubois [11], and Efstratiadis [12]. The motion estimator has to be able to produce consistent vector fields in the presence of observation noise. In particular, in low-detail areas where the local signal-to-noise-ratio is low, the vector field must be consistent [13], [14]. For filtering image sequences, motion estimators that show good noise behavior are preferred, such as those used by Sezan *et al.* [3], Boyce [15], Kleihorst *et al.* [13], [14], Özkan *et al.* [16], and Woods and Kim [17].

Most motion estimators assume translational locally stationary motion. Since this model is only an approximation of the true motion, the compensated sequence may still exhibit temporally nonstationary parts, e.g., in occluding areas. Therefore, even if motion-compensated filtering is used, a nonlinear or adaptive filtering approach is still required.

In this paper, we propose a new method for filtering image sequences that is based on a combination of motion compensation, signal decomposition, and adaptive temporal filtering (Fig. 1). First, a motion estimator that produces consistent motion estimates is used to create a motion-compensated version of  $g(k)$  that is (temporally) more stationary than the signal  $g(k)$  itself. Next, this signal is adaptively decomposed into a location signal and a residual signal. The residual signal is finally normalized in its variance. The aim of the decomposition is to map the noise-free nonstationary part of the signal into the location signal and the noisy but stationary part into the residual. Ordered statistics are used to estimate the required local signal statistics. After noise filtering the residual

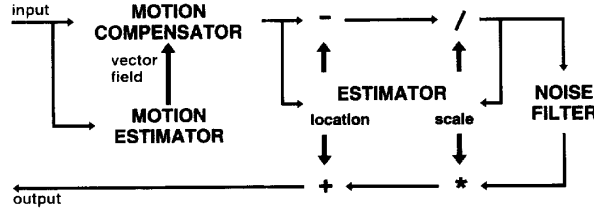


Fig. 1. System overview.

signal with an adaptive filter, the result is combined with the location signal to produce the filtered sequence.

The outline of this paper is as follows. In Section II the motion estimator used is presented. This is a recursive block-matching algorithm that produces consistent vector fields in the presence of noise. In Section III, the compensated temporal signal is modeled by a time-varying model incorporating current signal location and residual signal. Then, we derive adaptive estimators to obtain the required local statistics. In Section IV we describe the adaptive noise filter used. In Section V, the proposed method is evaluated on sequences with different noise levels, comprising Gaussian and Laplacian noise. We compare our results with several methods from literature.

## II. MOTION ESTIMATION IN NOISE

In a motion-compensating preprocessing step, temporal non-stationarities that can be described by an underlying motion model are compensated for. This increases the temporal correlation and thus improves the noise suppression achievable.

The motion model used is

$$f(i, j, k) = f(i + d_i, j + d_j, k - 1) \quad (3)$$

where  $(d_i, d_j)^t$  is the displacement vector to be estimated. Note that with (3) we assume that all motion is translational, no occlusion occurs, and no frame-to-frame intensity variations of objects occur. In addition to this we will assume that the motion can be described reasonably well by integer displacement vectors.

Motion estimation is performed on noisy observed sequences. Therefore, the motion estimator has to be insensitive to noise. Several robust motion estimation algorithms are proposed in literature. More complex methods take the presence of noise directly into account, such as the method using the generalized maximum likelihood criterion by Namazi and Lee [18], the method using cumulants by Anderson and Giannakakis [19] for Gaussian noise, and the method based on a vector coupled Gauss-Markov random field model by Brailean and Katsaggelos [20]. Methods that are tuned to give consistent vector fields are used with good results in noise filtering schemes. Among them are the algorithm by Fogel [21] used by Sezan *et al.* [3], the hierarchical block-matching algorithm by Bierling [22] used by Woods and Kim [17], and the switching block-matching algorithm used by Boyce [15].

We have used a modified 3-D recursive search block-matching algorithm from De Haan *et al.* [23]. This method ensures a consistent behavior because it limits the number of

possible displacements and uses spatial and temporal recursion. Therefore, it is more robust to noise than a standard block-matching algorithm. It was shown that the motion estimation algorithm used is robust to noise for SNR's of 10 dB or more [13], [14]. The algorithm is described briefly in the following.

The current frame is divided into nonoverlapping square blocks. For each block, in a scanning manner, the displacement relative to the previous frame is estimated. Instead of evaluating all possible displacements within a search region as a full-search block-matching algorithm does, the recursive block-matching algorithm evaluates only fine adjustments around a predicted value. For the prediction of the displacement of the current block, the estimated displacement of a neighboring block is used.

This neighboring block can be taken from the same image yielding spatial recursion or from the previous image yielding spatio-temporal recursion. Predictions from different directions can be used to produce additional candidate vectors for displacement. The spatial predictions at position  $i, j, k$  are given by the estimation results of two previous blocks from the same frame, yielding the following two candidates:

$$\begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i,j,k)} = \begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i-N,j-N,k)} + \begin{pmatrix} u_i \\ u_j \end{pmatrix} \quad (4)$$

$$\begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i,j,k)} = \begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i+N,j+N,k)} + \begin{pmatrix} u_i \\ u_j \end{pmatrix}. \quad (5)$$

Here,  $(\hat{d}_i, \hat{d}_j)^t$  are the previous estimates,  $(\hat{d}_i', \hat{d}_j')^t$  the candidate vectors;  $N$  denotes the size of the square blocks; and  $(u_i, u_j)^t$  takes on values from the following set of allowed updates:

$$\begin{pmatrix} u_i \\ u_j \end{pmatrix} \in \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\}. \quad (6)$$

Spatio-temporal recursion is introduced by using the displacement estimate of two blocks from the previous image as predictions for the displacement of the current block:

$$\begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i,j,k)} = \begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i-2N,j+2N,k-1)} \quad (7)$$

and

$$\begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i,j,k)} = \begin{pmatrix} \hat{d}_i' \\ \hat{d}_j' \end{pmatrix}_{(i+2N,j+2N,k-1)}. \quad (8)$$

The total number of candidate displacement vectors that has to be evaluated in the above estimator is 12. From these displacement vectors, the candidate that gives rise to the smallest *square error* matching the block to the previous frame is selected as the final estimate. The resultant vector field, with one integer vector per block is interpolated using a median filter to one integer vector per pixel.

In our experiments, the block size was set to  $N = 8$ , which gave the best results for several noise levels. The specific locations of the previous blocks stated for the recursion in (4), (5), (7), and (8) are found by experimental optimization using a number of video sequences [23]. They are located at offset positions from the current block to be able to notice and

track motion sooner. In effect, four directions are guarded for moving objects in the scene. Of course, the spatial recursion has to be causal. The reader can refer to [23] for a more thorough discussion of the motion estimation algorithm used.

### III. SIGNAL DECOMPOSITION WITH ESTIMATORS BASED ON ORDER STATISTICS

In general, the use of motion compensation yields a signal that is temporally more stationary. However, the signal is still not sufficiently stationary to be suitable for temporal linear filtering. This is because of the finite accuracy of the motion estimates and the incompleteness of the translational motion model underlying block-matching and most other motion estimation algorithms.

A classical way to cope with nonstationary signals is trend removal and normalization [24]. The aim is to decompose the signal in a noisy stationary and noiseless nonstationary part. Therefore, a regular noise filter can be applied to filter the stationary signal [25], [26]. Here, we focus on temporal trend estimation and scale estimation using ordered statistics.

#### A. Signal Decomposition

The nonstationary observed temporal signal is modeled as follows [27], [10]:

$$g(k) = \mu(k) + \sigma(k) \cdot y(k) \quad (9)$$

where  $\mu(k)$  is a slowly changing function known as the *ensemble location* at position  $k$  and  $\sigma(k)$  is the *ensemble scale*, which is the amplification factor of the signal  $y(k)$ . The residual signal  $y(k)$  is modeled to have a time invariant, zero mean and a time invariant unity variance. If the observation noise has zero mean, then  $\mu(k)$  is noise free.

The signal  $y(k)$  can be written using (1) and (9) as

$$y(k) = \frac{f(k) + n(k) - \mu(k)}{\sigma(k)} = \frac{f(k) - \mu(k)}{\sigma(k)} + \frac{n(k)}{\sigma(k)} = r(k) + v(k) \quad (10)$$

where all signals are now along the motion trajectory. For instance,  $f(k)$  is the original temporal signal along the motion trajectory. We assume that  $v(k)$  is a zero mean, uncorrelated, and independent signal.

In practice, signal  $y(k)$  is estimated from  $g(k)$  in the following way:

$$\hat{y}(k) = \frac{g(k) - \hat{\mu}(k)}{\hat{\sigma}(k)} \quad (11)$$

where  $\hat{\mu}(k)$  and  $\hat{\sigma}(k)$  are estimated from  $g(k)$ .

The signal  $\hat{y}(k)$ , stationary in mean and variance, can then be filtered by a regular noise filter to estimate  $r(k)$ . Finally an estimate of the original image sequence is established by combining  $\hat{r}(k)$  with the estimated location and scale

$$\hat{f}(k) = \hat{r}(k) \cdot \hat{\sigma}(k) + \hat{\mu}(k). \quad (12)$$

#### B. Derivation of the Estimators

In order to obtain an estimate of the residual signal  $y(k)$  by (11), the location  $\mu(k)$  and scale  $\sigma(k)$  have to be estimated from  $g(k)$ . Estimators of local scale and location based on the class of order statistic (OS) estimators have been successfully applied in digital signal processing [28], [29], [7]. In our application, they have the form

$$\hat{\theta}(k) = C \vec{X}_{(k)} \quad (13)$$

where  $\vec{X}_{(k)}(k) = [x_{(1)}, \dots, x_{(m)}]^t$ . The *ranks*  $x_{(r)}$  with  $x_{(1)} \leq \dots \leq x_{(m)}$  are the ordered realizations of the signal value  $g(k)$ . The subscripts  $(\cdot)$  will denote that the vector  $\vec{X}_{(k)}(k)$  is ordered. Note that (13) refers to an ensemble operation; the relation with spatio-temporal filtering is addressed in Section III-C.

Alternatively, (13) can also be written as

$$\begin{bmatrix} \hat{\mu}(k) \\ \hat{\sigma}(k) \end{bmatrix} = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,m} \\ c_{2,1} & c_{2,2} & \dots & c_{2,m} \end{bmatrix} \cdot \begin{bmatrix} x_{(1)} \\ x_{(2)} \\ \vdots \\ x_{(m)} \end{bmatrix}. \quad (14)$$

A set of optimal coefficients  $c_i$  needs to be calculated for estimating  $\mu(k)$  and  $\sigma(k)$  from  $g(k)$ . To this end, we first assume that  $g(k)$  is distributed according to

$$p(g(k)) = \frac{1}{\sigma(k)} Q\left(\frac{g(k) - \mu(k)}{\sigma(k)}\right), \quad \sigma(k) > 0. \quad (15)$$

This connects to each observation  $g(k)$  a parent distribution  $Q(\cdot)$  and a particular  $\mu(k)$  and  $\sigma(k)$ . By normalizing  $g(k)$  as

$$y(k) = \frac{g(k) - \mu(k)}{\sigma(k)} \quad (16)$$

$y(k)$  has  $Q(\cdot)$  as probability density function with zero location and unity scale. Let us for the moment assume that the parent distribution  $Q(\cdot)$  is known.

The following relations exists between the ranks  $x_{(r)}$ , ( $1 \leq r \leq m$ ) and the ranks  $y_{(r)}$  of an ordered vector  $\vec{Y}_{(k)}(k)$  with realizations of  $y(k)$  after taking expectations:

$$\begin{aligned} E\{x_{(1)}\} &= \mu(k) + \sigma(k) \cdot E\{y_{(1)}\}, \\ E\{x_{(2)}\} &= \mu(k) + \sigma(k) \cdot E\{y_{(2)}\}, \\ &\vdots \\ E\{x_{(m)}\} &= \mu(k) + \sigma(k) \cdot E\{y_{(m)}\}. \end{aligned} \quad (17)$$

The expectations of the ranks of  $\vec{Y}_{(k)}(k)$ ,  $E\{y_{(r)}\}$  depend on  $Q(\cdot)$ . If  $Q(\cdot)$  is known *a priori*, they can be calculated. Equation (17) shows that for ordered realizations of  $g(k)$ , we expect to find a set of  $m$  linear relations. From this set,  $\hat{\mu}(k)$  and  $\hat{\sigma}(k)$  can be found using generalized weighted least-squares estimation [30], as illustrated in Fig. 2 for  $m = 5$ .

With the definition of  $A$  as

$$A = \begin{bmatrix} \vec{1} & E\{\vec{Y}_{(k)}\} \end{bmatrix} = \begin{bmatrix} 1 & E\{y_{(1)}\} \\ \vdots & \vdots \\ 1 & E\{y_{(m)}\} \end{bmatrix} \quad (18)$$

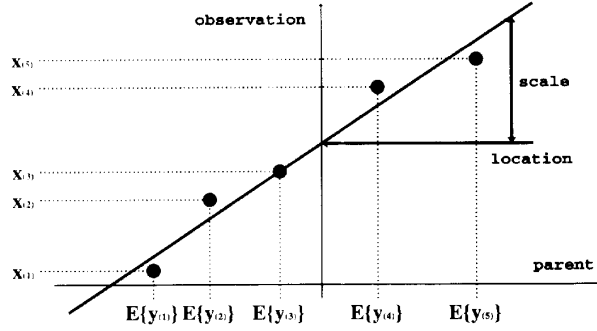


Fig. 2. There exists a linear relation between the observation vector and a vector from the parent distribution.

we rewrite (17) as follows:

$$E\{\vec{X}_0(k)\} = A\vec{\theta}(k) = [\vec{I} \quad E\{\vec{Y}_0\}] \vec{\theta}(k). \quad (19)$$

The minimal variance estimate of  $\vec{\theta}(k)$  is defined as

$$\hat{\vec{\theta}}(k) \leftarrow \min_{\vec{\theta}(k)} (\vec{X}_0(k) - A\vec{\theta}(k))^T B^{-1} (\vec{X}_0(k) - A\vec{\theta}(k)). \quad (20)$$

This inner product is weighted with the inverse variance matrix  $B$  of  $\vec{Y}_0$  for decorrelation purposes. The solution to (20) is given by

$$\hat{\vec{\theta}}(k) = (A^T B^{-1} A)^{-1} A^T B^{-1} \vec{X}_0(k) = C \vec{X}_0(k). \quad (21)$$

Equation (21) comprises the two OS estimators for location and scale. The coefficients of  $C$  can be interpreted as follows: The upper row are the weights to find  $\hat{\mu}(k)$  and the lower row are the weights to find  $\hat{\sigma}(k)$ . When the elements of the upper row  $c_{11}, c_{12}, \dots, c_{1m}$  are equal in value, the estimator of location is the sample average. This is the case for Gaussian distributed samples, when the sample average is the maximum likelihood estimator of location. For the exponential distribution, the coefficients represent the sample median that is optimal for this distribution [28], [31]. For other distributions  $Q()$ , the coefficients are optimal in a minimum variance sense.

### C. Implementation Aspects

The estimation of location and scale is performed on the motion-compensated observed signal. We make the assumption that an image sequence is locally ergodic, making it feasible to replace the ensemble statistics by the local spatio-temporal statistics. This means that the sample  $\vec{X}(k)$  is given by the values within a spatio-temporal window of size  $m$  along a motion trajectory. This window is chosen to incorporate the current pixel in combination with its six nearest neighbors. This motion-compensated window contains the pixels

$$\begin{aligned} \vec{X}(k) = & [g(i + \hat{d}_i(i, j, k), j + \hat{d}_j(i, j, k), k - 1), \\ & g(i - \hat{d}_i(i, j, k), j - \hat{d}_j(i, j, k), k + 1), \\ & g(i, j, k), g(i - 1, j, k), g(i + 1, j, k), \\ & g(i, j - 1, k), g(i, j + 1, k)]^T. \end{aligned} \quad (22)$$

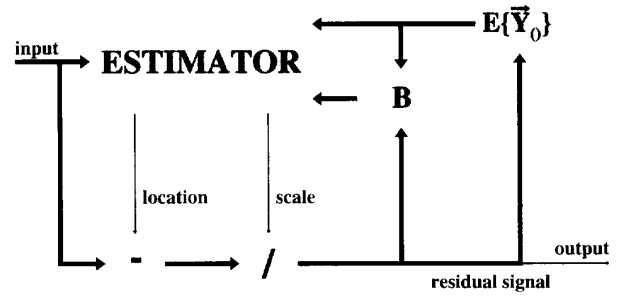


Fig. 3. Update process of  $B_{(yy)}$  and  $E\{\vec{Y}_0\}$  for the OS estimator.

which proved in a number of experiments to be a good trade-off between low estimate variance and correctness of the assumption of local ergodicity.

By extending the displacement estimates to find corresponding pixels in the future frame, we have assumed that the object velocity is constant within these frames. As described in Section II, the displacement vectors are estimated from the current frame relative to the previous frame.

So far we have assumed that the parent distribution  $Q()$  was known *a priori*. The shape and properties of this distribution are reflected by  $E\{\vec{Y}_0\}$  and the inverse variance matrix  $B^{-1}$  which determine the optimal estimators.

In general, the parent distribution is not known *a priori* and may vary spatio-temporally. However, the residual signal  $y(k)$  is distributed according to the parent distribution  $Q()$  with  $\mu(k) = 0$  and  $\sigma(k) = 1$ , which means that  $B^{-1}$  and  $E\{\vec{Y}_0\}$  can be estimated from the residual signal itself in order to describe  $Q()$ . The recursive estimation procedure used is illustrated in Fig. 3.

We have used the matrix inversion lemma [32], [33] to estimate  $B^{-1}$  directly in inverted form. A forgetting factor  $\lambda$  was used in the usual way to be able to adapt to signal variations.

The recurrent formula used to estimate  $E\{\vec{Y}_0\}(k)$  is

$$\hat{E}\{\vec{Y}_0\}(k) = \frac{\sum_{i=1}^k \lambda^{k-i} \vec{Y}_0(i)}{\sum_{i=1}^k \lambda^{k-i}}. \quad (23)$$

The forgetting factor used was set at  $\lambda = 0.97$ , which was found by experimental tuning for maximum achievement on a number of representative image sequences.

As initial estimates of  $B^{-1}$  and  $E\{\vec{Y}_0\}$ , the inverse variance matrix and the average vector of the distribution of ordered samples of observation noise are used. For more information regarding the adaptation of the estimators, the reader is referred to [7].

## IV. THE NOISE FILTER

After the motion compensation and decomposition step, the residual signal  $\hat{y}(k)$  is stationary in mean and variance. The noise  $n(k)$  is mapped into this signal according to (10). Noise filtering this residual signal and combining the result with the estimated location and scale renders an estimate of the original signal  $f(k)$ .



Fig. 4. Original and noisy (10 dB) image from the "Trevor White" sequence.

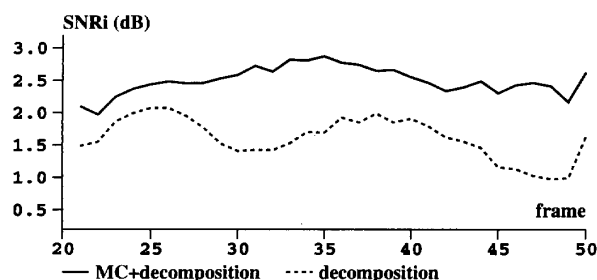


Fig. 5. Motion-compensated and nonmotion-compensated results of the proposed decomposition method for 20 dB SNR on the "Trevor White" sequence.

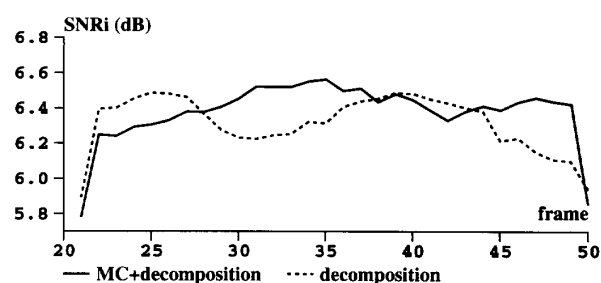


Fig. 6. Motion-compensated and nonmotion-compensated results of the proposed decomposition method for 10 dB SNR on the "Trevor White" sequence.

Because only the first two moments of the residual signal were normalized, its autocorrelation still changes. Therefore, we have used an adaptive recursive least-squares (RLS) filter to estimate  $r(k)$  from  $\hat{y}(k)$ :

$$\hat{r}(k) = \sum_{i=0}^2 w_i(k) \hat{y}(k-i). \quad (24)$$

Here  $w_i(k)$  are the components of the weight vector  $\vec{w}(k)$ , which is defined by recurrently minimizing the cumulative square error

$$\min_{\vec{w}(k)} \left\{ \sum_{i=1}^k \lambda^{k-i} (\hat{y}(i) - \hat{r}(i))^2 \right\}. \quad (25)$$

Again,  $\lambda$  serves as a forgetting factor [32] and is given the same value as used for the update of the estimators, i.e.,  $\lambda = 0.97$ . The support of the RLS filter was motivated by experimental evaluation on a number of image sequences.

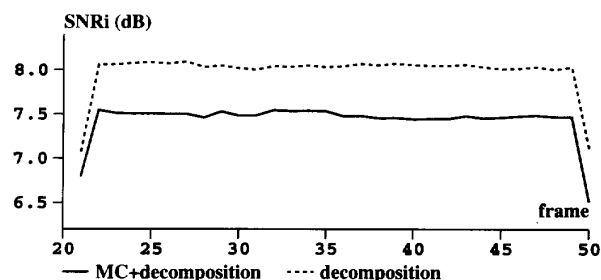


Fig. 7. Motion-compensated and nonmotion-compensated results of the proposed decomposition method for 0 dB SNR on the "Trevor White" sequence.

## V. EXPERIMENTS

In this section, the proposed algorithm, shown in Fig. 1, is evaluated on noisy image sequences. We test its performance for several noise levels comprising Gaussian and Laplacian noise.

The sequences used are parts of the luminance components

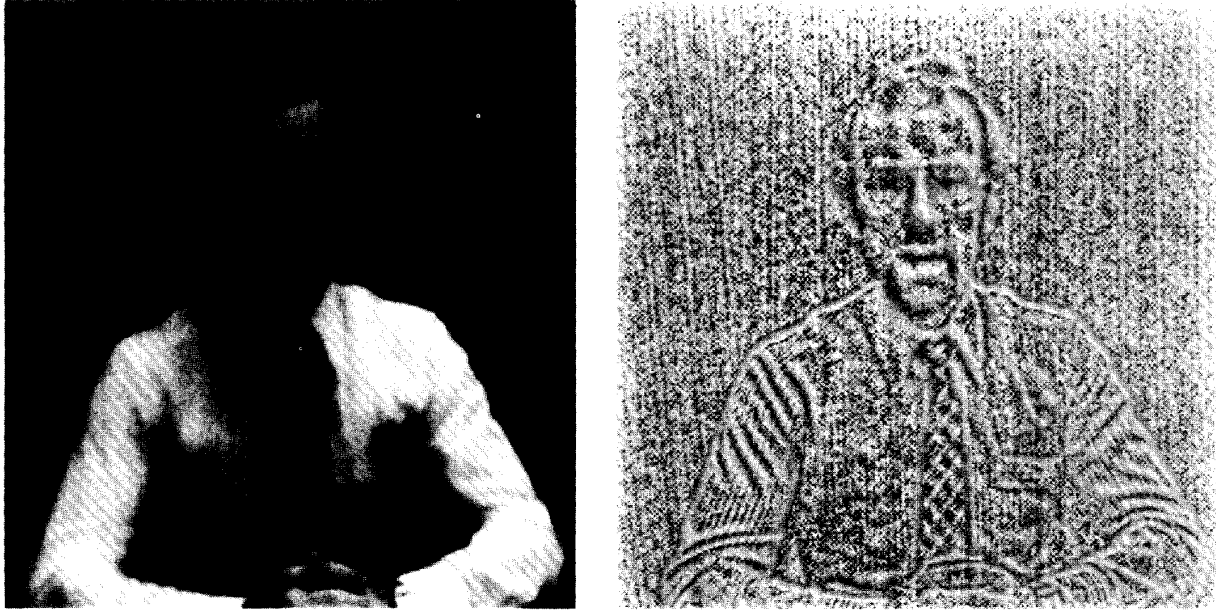


Fig. 8. Location and stretched residual signal of the “Trevor White” sequence with 10 dB noise without motion compensation.

of the “Trevor White” sequence (images 21 to 50) and the “Calendar Train” sequence (images 1 to 40). The level of corruption by the noise is specified by the signal-to-noise ratio (SNR), which is the ratio between the original signal variance  $\sigma_f^2$  and the noise variance  $\sigma_n^2$ :

$$\text{SNR} = 10 \log_{10} \left\{ \frac{\sigma_f^2}{\sigma_n^2} \right\} (\text{dB}).$$

The improvement in quality as a result of a filtering strategy is expressed as the improvement in SNR. The SNRi (dB) for image  $k$  is defined as

$$\begin{aligned} \text{SNRi}(k) \\ = 10 \log_{10} \left\{ \frac{IJ \sigma_n^2}{\sum_{i=1}^I \sum_{j=1}^J [\hat{f}(i, j, k) - f(i, j, k)]^2} \right\} \end{aligned}$$

where the scalars  $J$  and  $I$  are the number of rows and the number of columns in one image of the sequence.

First, we illustrate the effect of using motion compensation for different noise levels. In Fig. 4, a frame from the “Trevor White” sequence is shown in both original and corrupted form with Gaussian noise added to an SNR of 10 dB. For each noise level, the filtering is performed with and without motion compensation. In Figs. 5–7, the results are shown for 20, 10, and 0 dB noise levels, respectively. For high SNR’s (20 dB), motion compensation is a useful preprocessing step. For moderate levels (10 dB), the two curves are approximately identical, while for low SNR’s (0 dB) the motion estimator fails. Therefore, the overall noise filtering becomes less successful.

To illustrate the signals created by the estimator, we show the location signal  $\hat{\mu}(i, j, k)$  and the residual signal  $\hat{y}(i, j, k)$  for a single frame in Fig. 8. These frames are taken from the

experiment on 10 dB data without motion compensation. It can be seen that the location and essentially noise-free signal is very smooth. The residual signal is stretched and an offset is added for maximum visibility.

The location and residual signal from the experiment with motion compensation are shown in Fig. 9. It can be seen that the motion-compensated location signal is sharper and the residual signal contains fewer spatio-temporal edges. The residual signal is scaled with the same parameters used for the residual signal in Fig. 8.

The final results of applying the proposed algorithm without and with motion compensation are shown in Fig. 10 for an SNR of 10 dB.

To deal with the nonstationarities in the temporal signal, two approaches are proposed in literature: first, the use of nonlinear filters [4]–[6], and second, the use of motion compensation [9], [10], [3]. We compare our algorithm with some well-performing methods from these two categories. First, we compare our results with the algorithm based on 3-D median filters presented by Arce in [5].

Arce uses bidirectional multistage median filters (MMF’s), which is a method of combining the output of median subfilters. In a spatio-temporal cube, four bidirectional median filters with  $4N+1$  points are constructed. They are shown for  $N=1$  in Fig. 11. The final result for the center pixel is defined as

$$\begin{aligned} \hat{f}(i, j, k) = \text{median} \{ \max(m_1, m_2, m_3, m_4), \\ g(i, j, k), \min(m_1, m_2, m_3, m_4) \} \end{aligned}$$

where  $m_1, m_2, m_3$ , and  $m_4$  are the output of the four bidirectional median filters.

We evaluated the MMF on the “Trevor White” sequence corrupted with Gaussian noise to a level of 10 dB for  $N=1, 2$ , and 3. The SNRi curves are shown, together with the



Fig. 9. Location and stretched residual signal of the "Trevor White" sequence with 10 dB noise with motion compensation.



Fig. 10. Filtered frames with the proposed decomposition method without (left) and with (right) motion compensation for an SNR of 10 dB.

proposed motion compensated and OS decomposition result in Fig. 12. From the relatively flat appearance it can be seen that Arce's filters preserve motion very well, but they are (at least for small  $N$ ) not very effective in reducing noise. Note that Arce proposed to use  $N = 1$  [5].

The second algorithm we consider is a motion-compensated filter proposed by Katsaggelos *et al.* [10]. A pel-recursive motion estimator is used. The motion-compensated signal is filtered by a 3-D recursive filter, which is a concatenation of

three 1-D first-order recursive filters

$$\begin{aligned}\hat{f}(i, j, k) &= [1 - F_k]\rho_k \hat{f}(i, j, k-1) + F_k \hat{f}_j(i, j, k), \\ \hat{f}_j(i, j, k) &= [1 - F_j]\rho_j \hat{f}(i, j-1, k) + F_j \hat{f}_i(i, j, k), \\ \hat{f}_i(i, j, k) &= [1 - F_i]\rho_i \hat{f}(i-1, j, k) + F_i g(i, j, k).\end{aligned}$$

Here,  $F_i$ ,  $F_j$ , and  $F_k$  are the horizontal, vertical, and temporal filter gains respectively, and  $\rho_i$ ,  $\rho_j$ , and  $\rho_k$  are the horizontal, vertical, and temporal correlations. This filter is used in a nonadaptive version and an adaptive version. In the nonadaptive version, the correlations were estimated off-line and the gains

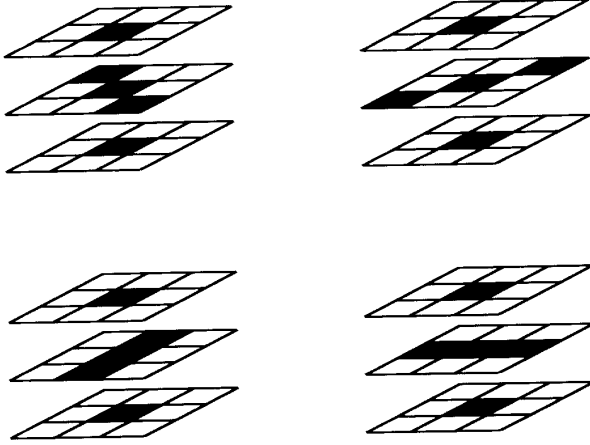
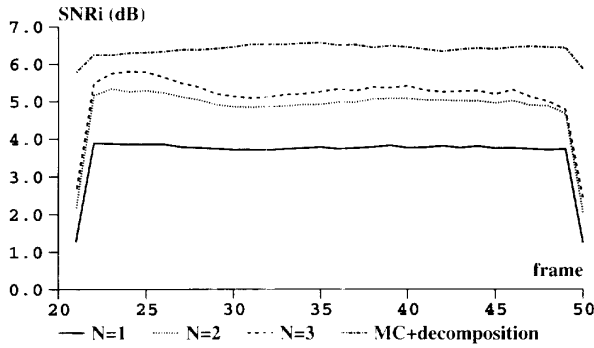
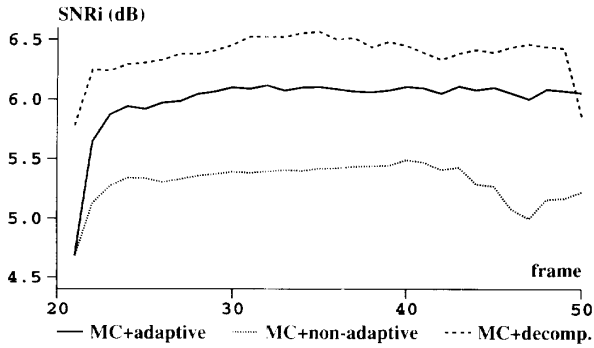
Fig. 11. Arce's bidirectional median filter supports ( $N=1$ ).Fig. 12. Comparison with the method proposed by Arce (for  $N=1, 2$ , and  $3$ ), and the decomposition method including motion compensation on the "Trevor White" sequence with an SNR of 10 dB.

Fig. 13. Comparison with the nonadaptive and adaptive method proposed by Katsaggelos and the decomposition method including motion compensation on the "Trevor White" sequence with an SNR of 10 dB.

were calculated from the data in order to be globally optimal. As this tends to blur spatio-temporal edges, an adaptive version was designed where the correlations and gains were controlled by edge detectors in the spatial directions and by the motion compensation error in the temporal direction [10].

The results are evaluated for "Trevor White" at a noise level of 10 dB (shown in Fig. 13) and at a level of 20 dB (shown in Fig. 14).

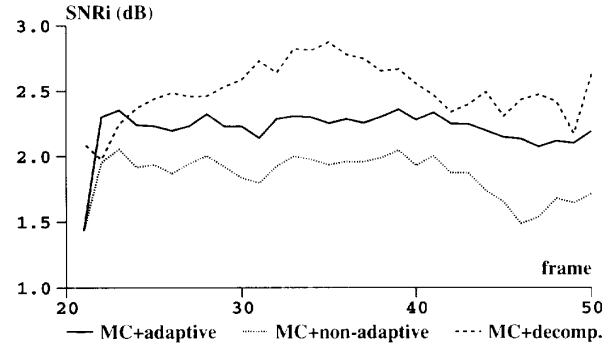


Fig. 14. Comparison with the nonadaptive and adaptive method proposed by Katsaggelos and the decomposition method including motion compensation on the "Trevor White" sequence with an SNR of 20 dB.

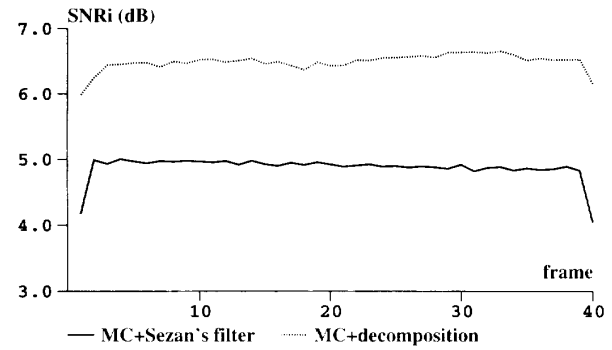
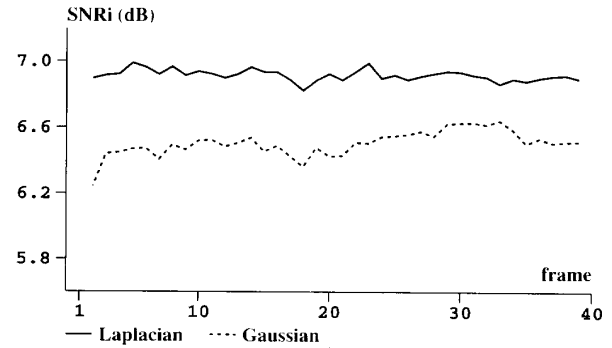
Fig. 15. Comparison with the method proposed by Sezan *et al.* and the decomposition method including motion compensation on the "Calendar Train" sequence with an SNR of 10 dB.

Fig. 16. Results of the proposed decomposition method with motion compensation for Gaussian and Laplacian noise at 10 dB SNR for the "Calendar Train" sequence.

As a third comparison, we consider the adaptive motion-compensated filter described by Sezan *et al.* [3]. Instead of using their proposed motion estimator, we have used the 3-D recursive search block-matching algorithm described in Section II. The noise filter used is composed as

$$\hat{f}(i, j, k) = \hat{\mu}(i, j, k) + \frac{\max(0, \hat{\sigma}^2(i, j, k) - \hat{\sigma}_n^2)}{\max(\hat{\sigma}^2(i, j, k), \hat{\sigma}_n^2)} \times (g(i, j, k) - \hat{\mu}(i, j, k)).$$

Here,  $\hat{\sigma}_n^2$  is the estimated noise variance. We have used our



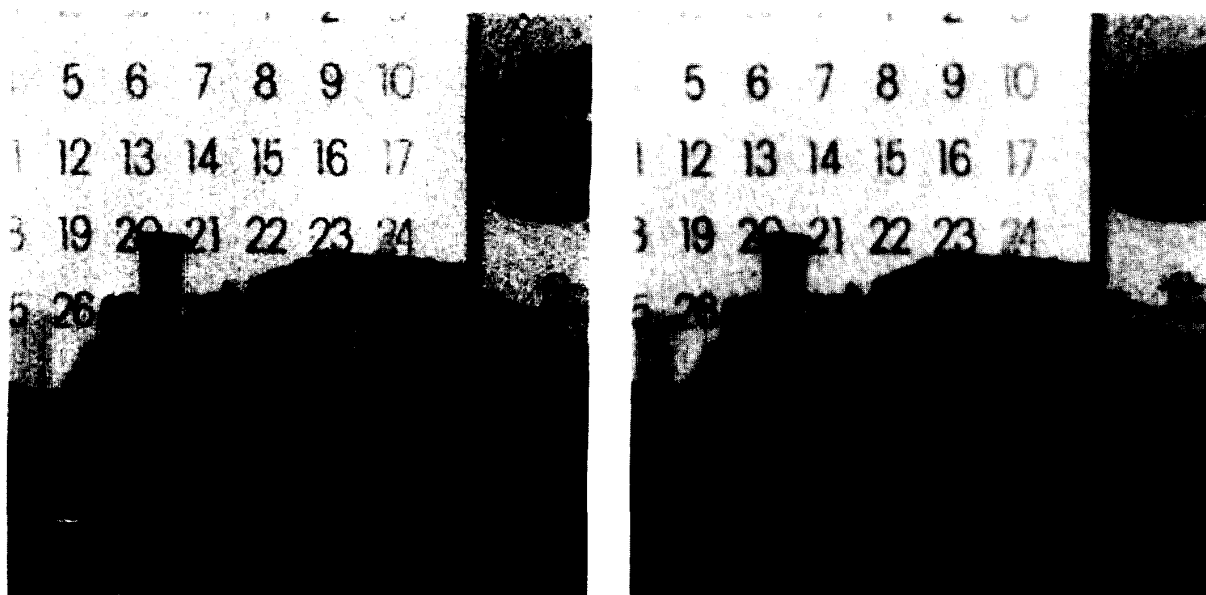


Fig. 17. Observed image of "Calendar Train" at 10 dB Gaussian noise and result image of the proposed decomposition method with motion compensation.

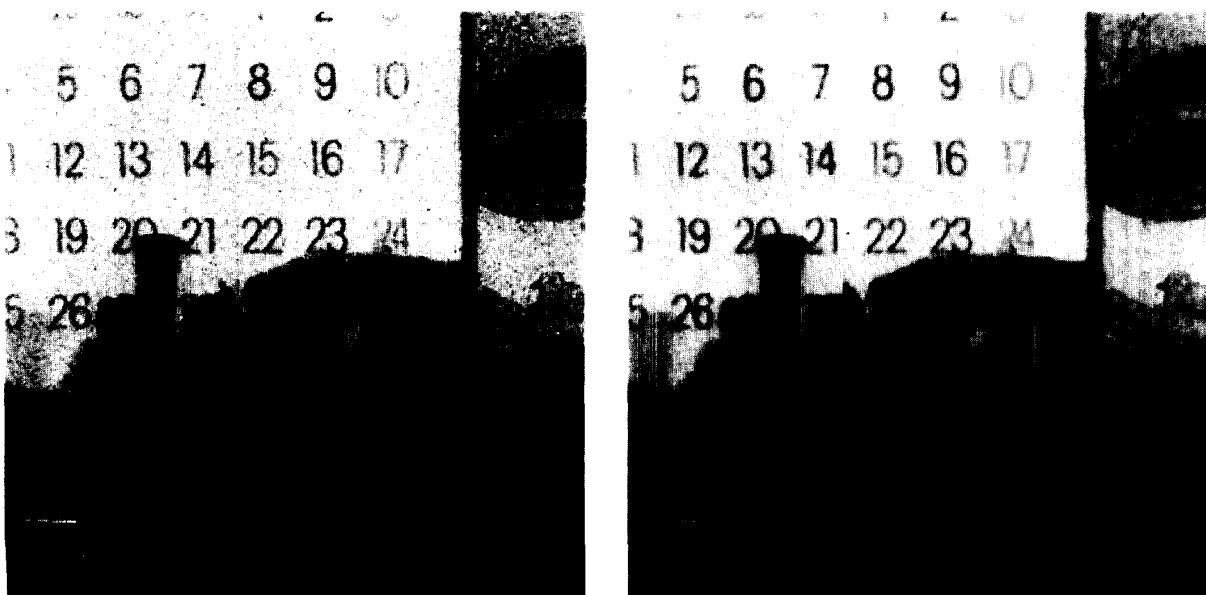


Fig. 18. Observed image of "Calendar Train" at 10 dB Laplacian noise and result image of the decomposition method with motion compensation.

adaptive OS estimator from (21) to estimate the parameters  $\hat{\mu}(i, j, k)$  and  $\hat{\sigma}(i, j, k)$  for this filter.

The SNR curves yielded by filtering the "Calendar Train" sequence at a noise level of 10 dB are shown in Fig. 15.

Finally, to illustrate that our estimator is able to decompose a signal with other noise distributions we have corrupted the "Calendar-Train" sequence with Gaussian and Laplacian noise at a level of 10 dB. The results of this experiment are presented in Fig. 16. An observed and result image of the sequence corrupted with Gaussian noise is shown in Fig. 17.

An observed and result image of the sequence corrupted with Laplacian noise is shown in Fig. 18.

From a subjective point of view regarding our method, the noise is substantially removed without affecting the sharpness. It was found that the use of motion compensation is necessary at higher SNR's to avoid annoying blurring of spatio-temporal edges. For very low SNR's, the results look better without motion compensation.

Using a motion estimator that is more robust to noise as the algorithm proposed in this paper, lower noise levels can be

tackled. A step in this direction has been the introduction of a criterion based on cumulants in the proposed recursive block-matching algorithm which enabled reliable motion estimation and filtering at 0 dB SNR [34].

The results achieved by the method from Arce [5] still contain visible noise. The nonadaptive method from Katsaggelos [10] is suffering from blurring of edges; however, the results achieved by the adaptive version are good. The results achieved by the method of Sezan *et al.* [3] are sharp and well cleaned from noise. Subjectively, the method presented in this work performs at least as good and usually better than the methods from [5], [10], and [3].

## VI. DISCUSSION

In this paper, we have proposed an algorithm for reducing noise in image sequences. It consists of a motion compensation step based on a consistent block-matching algorithm to account for the nonstationarities caused by motion and a decomposition based on an adaptive OS estimator to account for the incompleteness of the motion model. A simple temporal noise filter reduces the noise in the residual signal.

The method has been shown to outperform well-known methods from recent literature. The price that has to be paid for the improved performance is a computationally demanding determination of optimal estimators of location and scale for every pixel.

## ACKNOWLEDGMENT

The authors wish to thank Dr. G. De Haan from the Philips Research Laboratories, Eindhoven, The Netherlands, for the discussions on motion estimation. We also thank Professor Katsaggelos from Northwestern University, Evanston, IL, and Dr. Efstratiadis for their collaboration in a joint research project and invigorating discussions.

## REFERENCES

- [1] T. S. Huang and Y. P. Hsu, "Image sequence enhancement," in *Image Sequence Analysis* (T. S. Huang, Ed.), New York: Springer-Verlag, 1981, pp. 289–309, chapter 4.
- [2] E. Dubois and S. Sabri, "Noise reduction in image sequences using motion compensated temporal filtering," *IEEE Trans. Commun.*, vol. COM-32, pp. 826–831, July 1984.
- [3] M. I. Sezan, M. K. Özkan, and S. V. Fogel, "Temporally adaptive filtering of noisy image sequences using a robust motion estimation algorithm," in *Proc. ICASSP '91*, Toronto, Canada, May 14–17, 1991, vol. 4, pp. 2429–2432.
- [4] D. I. Crawford, "Spatio/temporal prefiltering for a video conference coder," in *Proc. Int. IEEE Conf. Electronic Image Processing*, York, U.K., July 1982, pp. 236–242.
- [5] G. R. Arce, "Multistage order statistic filters for image sequence processing," *IEEE Trans. Signal Processing*, vol. 39, no. 5, pp. 1147–1163, May 1991.
- [6] M. B. Alp and Y. Neuvo, "3-dimensional median filters for image sequence processing," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Canada, May 14–17, 1991, vol. 4, pp. 2917–2920.
- [7] R. P. Kleihorst, "Nonlinear filtering of image sequences using order statistics," *Chartered Designer's Report*, Delft Univ. of Technology, Delft, The Netherlands, Jan. 1992.
- [8] C. L. Chan and B. J. Sullivan, "Nonlinear model-based spatio-temporal filtering of image sequences," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Canada, May 1991, vol. 4, pp. 2989–2992.
- [9] A. K. Katsaggelos, J. N. Driessen, S. N. Efstratiadis, and R. L. Lagendijk, "Temporal motion-compensated filtering of image sequences," in *Proc. SPIE Conf. Visual Commun. and Image Processing*, Boston, MA, 1989, vol. 1199, pp. 61–70.
- [10] A. K. Katsaggelos, R. P. Kleihorst, S. N. Efstratiadis, and R. L. Lagendijk, "Adaptive image sequence noise filtering methods," in *Proc. SPIE Conf. Visual Commun. Image Processing*, Boston, MA, Nov. 10–13, 1991, vol. 1606, pp. 716–727.
- [11] E. Dubois, "Motion compensated filtering of time-varying images," *Multidimensional Syst. Signal Processing*, vol. 3, pp. 211–239, May 1992.
- [12] S. N. Efstratiadis, "Recursive motion compensated image sequence estimation," Ph.D. dissertation, Dept. of EECS, Northwestern Univ., Evanston, IL, 1991.
- [13] R. P. Kleihorst, G. de Haan, R. L. Lagendijk, and J. Biemond, "Noise filtering of image sequences with double compensation for motion," in *Proc. 13th Symp. Inform. Theory Benelux*, Enschede, The Netherlands, June 1–2, 1992, pp. 197–204.
- [14] ———, "Motion compensated noise filtering of image sequences," in *Proc. Eusipco '92*, Brussels, Belgium, Aug. 24–27, 1992, pp. 1385–1388.
- [15] J. M. Boyce, "Noise reduction of image sequences using adaptive motion compensated frame averaging," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, San Francisco, CA, 1992, pp. 461–464.
- [16] M. K. Özkan, M. I. Sezan, and A. M. Tekalp, "Motion-adaptive weighted averaging for temporal filtering of noisy image sequences," in *Image Processing Algorithms and Techniques III*, San Jose, CA, Feb. 10–13, 1992, SPIE, vol. 1657, pp. 201–212.
- [17] J. W. Woods and J. Kim, "Motion compensated spatiotemporal Kalman filter," in *Motion Analysis and Image Sequence Processing* (R. L. Lagendijk and M. I. Sezan, Eds.), Norwell, MA: Kluwer, 1993, ch. 12.
- [18] N. M. Namazi and C. H. Lee, "Nonuniform image motion estimation from noisy data," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 2, pp. 364–366, Feb. 1990.
- [19] J. M. M. Anderson and G. B. Giannakis, "Noise insensitive motion estimation using cumulants," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, Toronto, Canada, May 1991, pp. 2721–2724.
- [20] J. C. Braillean and A. K. Katsaggelos, "Displacement field estimation in noisy image sequences," in *Signal Processing VI: Theories and Applications* (J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, Eds.), North-Holland: Elsevier, 1992, vol. 3, pp. 1319–1322.
- [21] S. V. Fogel, "Estimation of velocity vector fields from time-varying image sequences," *CVGIP: Image Understanding*, vol. 53, pp. 253–287, May 1991.
- [22] M. Bierling, "Displacement estimation by hierarchical block matching," in *Proc. SPIE Conf. Visual Commun. Image Processing*, Cambridge, MA, Nov. 9–11, 1988.
- [23] G. de Haan, P. W. A. C. Biezen, H. Huijgen, and O. A. Ojo, "True-motion estimation with 3-D recursive search block matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, no. 5, pp. 368–379, Oct. 1993.
- [24] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, New York: Springer-Verlag, 1986.
- [25] F. C. Jeng and J. W. Woods, "Inhomogeneous Gaussian models for estimation and restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1305–1312, Aug. 1988.
- [26] A. D. Hillery and R. T. Chin, "Restoration of images with nonstationary mean and autocorrelation," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing*, New York, 1988, pp. 1008–1011.
- [27] R. P. Kleihorst, R. L. Lagendijk, and J. Biemond, "Nonlinear filtering of image sequences using order statistics," in *Proc. 12th Symp. Inform. Theory Benelux*, Veldhoven, The Netherlands, May 23–24, 1991, pp. 49–55.
- [28] A. C. Bovik, T. S. Huang, and D. C. Munson, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 31, pp. 1326–1337, Dec. 1983.
- [29] A. Restrepo and A. C. Bovik, "Adaptive trimmed mean filters for image restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1326–1337, Aug. 1988.
- [30] E. H. Lloyd, "Least squares estimation of location and scale parameters using order statistics," *Biometrika*, vol. 39, pp. 88–95, 1952.
- [31] H. A. David, *Order Statistics*, New York: Wiley, 1981.
- [32] S. T. Alexander, *Adaptive Signal Processing*, New York: Springer-Verlag, 1986.
- [33] K. J. Astrom and B. Wittenmark, *Adaptive Control*, New York: Addison-Wesley, 1989.

- [34] R. P. Kleihorst, R. L. Lagendijk, and J. Biemond, "Noise reduction of severely corrupted image sequences," in *IEEE Proc. Int. Conf. Acoust., Speech, Signal Processing.*, Minneapolis, MN, Apr. 27-30, 1993.



**Richard P. Kleihorst** (S'94) was born in Schiedam, The Netherlands, in 1965. He received the M.Sc. degree in electrical engineering from Delft University of Technology, The Netherlands, in 1989.

In 1989, he worked at the Philips Research Laboratories in Eindhoven, The Netherlands, on fuzzy classification techniques for character descriptions. Since 1990 he has been pursuing his Ph.D. degree, working as a Research Assistant in the Laboratory for Information Theory of Delft University of Technology. His interests currently include digital signal

processing, with emphasis on enhancement of images and image sequences. He is an active member of the International Ford Capri club, the Vintage Mustang Register, and the Austin 7 Club. His efforts have resulted in a number of technical inventions in this field, including pulsed airstream media-blasting, cathodic rust-removal methods, cathodic protection of vehicles against oxidation, and fixed-duration/variable-frequency electronic ignition.



**Reginald L. Lagendijk** (M'91) was born in Leiden, The Netherlands, in 1962. He received the M.Sc. and Ph.D. degrees in electrical engineering from Delft University of Technology in 1985 and 1990, respectively.

In 1987, he became an Assistant Professor in the Laboratory for Information Theory of Delft University of Technology. He was a Visiting Scientist in the Electronic Image Processing Laboratories of Eastman Kodak Research in Rochester, NY, in 1991. Since 1993, he has been an Associate

Professor in the Laboratory for Information Theory of Delft University of Technology. He is author of the book *Iterative Identification and Restoration of Images* (Kluwer, 1991) and coauthor of the book *Motion Analysis and Image Sequences Processing* (Kluwer, 1993). His research interests currently include multidimensional signal processing and communication theory, with emphasis on filtering, compression, and analysis of image sequences.

He is associate editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING



**Jan Biemond** (F'92) was born in De Kaag, The Netherlands, on March 27, 1947. He received the M.S. and Ph.D. degrees in electrical engineering from Delft University of Technology, Delft, The Netherlands, in 1973 and 1982, respectively.

He is currently Professor and Chairman of the Information Theory Group of the Department of Electrical Engineering at Delft University of Technology. His research interests include multidimensional signal processing, image enhancement and restoration, video compression (digital TV, stereo-

scopic TV, and HDTV), and motion estimation with applications in image coding and computer vision. He has authored and coauthored over 140 papers in these fields. In 1983 he was a Visiting Researcher at Rensselaer Polytechnic Institute, Troy, NY, and at Georgia Institute of Technology, Atlanta, GA.

Dr. Biemond is a member of the IEEE-SP Technical Committee on Image and Multidimensional Signal Processing and a member of the IEEE-CAS Technical Committee on Visual Signal Processing and Communications. He served as the General Chairman of the Fifth ASSP/EURASIP Workshop on Multidimensional Signal Processing, held at Noordwijkerhout, The Netherlands, in September 1987. Further, he is an AdCom member of the European Association for Signal Processing (EURASIP) and has been a member of the Board of Governors of the IEEE-SP Society since 1994. He is Co-Editor of the International Journal on Multidimensional Systems and Signal Processing and he serves on the Editorial Boards of Image Communication and the Journal of Visual Communication and Image Representation. He is the scientific editor of a series of books on Image Communication. He is a Distinguished Lecturer of the IEEE Signal Processing Society for 1993-1994.