Image Estimation Using Fast Modified Reduced Update Kalman Filter

Wen-Rong Wu, *Member, IEEE*, and Amlan Kundu, *Member, IEEE*

Abstract—**In this paper, we have proposed some modifications of the reduced update Kalman filter (RUKF) as applied to filtering of images corrupted by additive noise. First, we have reduced the computational complexity by reducing the state dimensionality. By doing so, it is shown in the paper that the computational requirement is reduced by an order of magnitude while the loss of performance is only marginal. Next, the RUKF is modified using the score function based approach to accommodate the non-Gaussian noise. The image is modeled as a nonstationary mean and stationary variance autoregressive Gaussian process. It is shown in the paper that the stationary variance assumption is reasonable if the nonstationary mean is computed by an edge and detail preserving efficient estimator of local nonstationary mean. Such an estimator called HMSMD filter is also described in the paper. Finally, detailed experimental results are provided which indicate the success of the new filtering scheme.**

I. Introduction

THE problem of optimal estimation of images corrupted by additive noise has been considered by a number of researchers [1]–[10], [18]–[21]. For optimal estimation of images, an image model is the primary requirement. Hunt and Cannon [4] proposed a nonstationary mean Gaussian image model. They assumed that an image f can be decomposed into a nonstationary statistical mean component $E(f)$, and a stationary residual component $f_0 = f - E(f)$. The nonstationary statistical mean component describes the gross structure of an image and the residual component describes the detail variation of the image. If we replace the nonstationary statistical (ensemble) mean $E(f)$ by a local average (spatial) that is calculated over an $M \times M$ window, and subtract the local mean from the original image, we have the residual image f_0. In general, the residual image is still a correlated nonstationary process, but the shape of the histogram is more Gaussian in nature.

Nonstationarity of images has been considered by Rajala and de Figueiredo [5] and Ingle and Woods [6]. In [5], a piecewise stationary image model is assumed, and different Kalman filters are used in different segments of the image. Ingle and Woods used reduced update Kalman filtering [6] and a composite image model which assumes that an image is composed of many different stationary components. Nonstationarity of the images has been considered by Anderson and Netravali [7] and Abramatic and Silverman [8] using the masking effect of the human vision system. The masking effect derives from the fact that noise is perceived by a human eye differently depending upon whether it occurs close to or far from an edge. In [8], the resulting optimal filter is shown to be nonlinear. But none of these approaches are designed for non-Gaussian noise though the non-Gaussian noise consideration also leads to an optimal nonlinear filter.

In order to consider the nonstationarity of the image, and not to complicate the computation too much, Kuan *et al.* [9] assumed that f_0 is a nonstationary white Gaussian process. More specifically, f_0 is statistically uncorrelated, and is characterized by its nonstationary variance. This nonstationary mean and nonstationary variance (NMNV) image model has been used by Kuan *et al.* [9] to restore images corrupted by additive noise. Their algorithm is similar to the algorithm used by Lee [10].

It has been observed in [9] that to realize the NMNV image model, the residual image f_0 should be obtained from the original image f by means of an intelligent filter. The purpose of the intelligent filter is to provide edge preserving smoothing of the original image f such that the residual image exhibits less correlation at the edges. It is our observation that if the intelligent filter is edge preserving as well as detail preserving, and if it provides sufficient smoothing over the flat regions of the image, then the white variance assumption is reasonably satisfied for some of the images, but not for all images. The role of intelligent filter, also, becomes very critical.

The nonstationary mean, nonstationary white variance image model is simple and rather inexpensive computationally, but it fails to take care of the correlation structure of the residual image. On the other hand, a nonstationary mean subtracted residual image can be modeled as an AR Gauss–Markov random field with zero mean and space-variant variance. Such an image model is theoretically more appealing though computationally more expensive. Here, the incorporation of reduced update Kalman filtering [11] can reduce the complexity substantially.

In this paper, we have modeled the residual image using the state space model given by Woods *et al.* in their significant works as described in [11], [15]. We also assume

Manuscript received September 2, 1989; revised October 15, 1990. This work was supported in part by the National Science Foundation under Grant MIP-8908082.

W.-R. Wu is with the Microelectronics and Information Science/Technology Center, National Chiao Tung University, Hsinchu, Taiwan, Republic of China.

A. Kundu is with the Department of Electrical Engineering, State University of New York at Buffalo, Amherst, NY 14260.

IEEE Log Number 9106030.

that the covariance and the variance of the residual signal are stationary. The stationary variance assumption for the residual signal is the only significant difference from the image model proposed in [15]. In [15], a simple averaging filter is used to estimate the nonstationary mean. Such a filter cannot preserve the edges or the thin-line details. As a result, the variance of the residual signal near the edges or the thin-line signal tends to be higher compared to that over the flat regions of the image. Consequently, the residual signal exhibits markedly different variances over different parts of the image. On the other hand, if the nonstationary mean component is computed by an edge and detail preserving filter, it is reasonable to think that the variance of the residual signal will not change much over the image. It is to be noted that the computation of nonstationary variance from the noisy image is not very reliable, particularly when the data contain impulse noise. On the other hand, the computation of one stationary variance is more reliable. With these considerations, the proposed model is expected to be as effective as the one proposed in [15]. One of the contributions of this paper is to propose an efficient edge and detail preserving filter to justify the assumptions on the image model.

The filtering algorithm used in [15] is the well-known reduced update Kalman filtering (RUKF) [11] scheme. Although the RUKF has successfully applied the framework of Kalman filtering to image processing, the results can still be improved. Also, the optimal filter derived in [11] is linear in form, i.e., this filter is optimal in the class of linear filters. If the residual image is assumed to be Gaussian in nature, the linear filters are globally optimal only when the additive noise is also Gaussian. But the additive noise is not always Gaussian. The occurrence of impulse noise, which can appear due to transmission error, is a prime example. The other frequent non-Gaussian disturbances are Laplacian noise, rectangular noise, etc. Because of non-Gaussian noise, the conventional RUKF has to be modified using the score function of the observation prediction density as has been described by Wu and Kundu [12].

A further consideration is the computational complexity of the filtering scheme. To reduce the computational complexity, we have incorporated the idea of strip processing with the RUKF. The idea of strip processing was considered by Nahi [21]. In this approach, the image signal is transformed into a one-dimensional signal by raster scanning before processing. This results in a time variant correlation function. An approximation method is then used to obtain a time invariant correlation function. In order to achieve a better approximation, he then divided the horizontal scan lines into nonoverlapping strips. Subsequently, a time invariant correlation function is obtained for each strip. Once the correlatio functions are obtained, state space signal models can be derived. A time variant Kalman filter, which changes its parameters between successive strips, is then applied. A different type of strip processing was proposed in [11] by Woods and Radewan. The idea is to divide an image into some overlapping strips and each strip is processed independently by a vector Kalman filter. While Nahi's strip processing technique is used for handling the nonstationary scalar correlation functions, Woods' method is used to reduce the computational complexity. Thus, the objective of our strip processing is similar to that of Woods.

In brief, we have developed some modification schemes which can significantly improve the performance of the original RUKF. These modifications are mainly threefold. 1) Reduce the computation furthermore. 2) Accommodate the non-Gaussian noise. 3) Reduce edge blurring.

Of these three items, the first two items deal with the modification of the original RUKF; and the third item is connected with the computation of edge and detail preserving nonstationary local mean. It is also shown in the paper that these modifications significantly improve the estimation results.

This paper is organized as follows: Section II presents the basic idea of RUKF. In Section III, the modification of RUKF, for reduced computational complexity and non-Gaussian observation noise, is presented. In Section IV, a new filter, called HMSMD filter, is proposed for the computation of the nonstationary local mean. Section V describes the experimental results. The concluding remarks are made in Section VI.

II. Reduced Update Kalman Filter

Kalman filtering is a state space approach and a dynamic model of the system is essential. In order to construct the dynamic model, a state has to be defined. For an image, there is no unique way to define a state. In [11], Woods and Radewan proposed one way to define the state which utilizes the concept of a nonsymmetric half-plane model. Consider an $N \times N$ image. Assume that the image is processed in a raster scan fashion. If we take the current scan point and define it as the "present," the points that have been scanned will be the "past" and the remaining points will be the "future" (see Fig. 1(a)). Let an individual point in the image be represented as $s(k, l)$ where k is the horizontal coordinate and l is the vertical coordinate. Applying the concept of two-dimensional spectrum factorization, the $M \times M$ order signal model can be obtained as follows:

$$s(k, l) = \sum_{R} c_{ij} s(k - i, l - j) + w(k, l) \qquad (1)$$

where $R = \{M \geq i \geq 0, M \geq j \geq 0) \cup \{0 \geq i > -M, M \geq j > 0\}$ and $w(k, l)$ is a random process assumed to be zero-mean Gaussian. This signal model is called the nonsymmetric half-plane (NSHP) model. We now can reconstruct a state based on this model. Let

$$\begin{aligned} x(k, l) = [&s(k, l), s(k - 1, l), \cdots, s(1, l); \\ &s(N, l - 1), \cdots, s(1, l - 1); \cdots; \\ &s(N, l - M), \cdots, s(k - M, l - M)]^t. \end{aligned} \qquad (2)$$

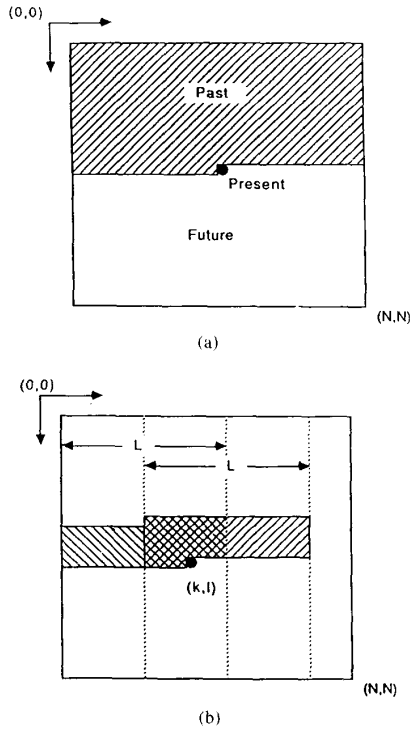


Fig. 1. (a) Nonsymmetric half-plane signal model. (b) Stripped nonsymmetric half-plane signal model.

Also, let us assume that the image is degraded by an additive noise, i.e.,

$$z(k, l) = s(k, l) + v(k, l) \tag{3}$$

where $z(k, l)$ is the observed pixel value of the image, and $v(k, l)$ is a zero mean white noise process independent of $w(k, l)$. Combining (1)–(3), a recursive dynamic model of the image is established:

$$x(k, l) = Fx(k - 1, l) + Gw(k, l) \tag{4}$$

$$z(k, l) = Hx(k, l) + v(k, l) \tag{5}$$

where

$$F = \begin{pmatrix} c_{10} & c_{20} & \cdots & c_{MM} \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 \end{pmatrix} \tag{6}$$

$$G = (1, 0, \cdots, 0)^t, \quad \text{and} \quad H = (1, 0, \cdots, 0). \tag{7), (8}$$

Once the dynamic model is obtained, the standard Kalman filter equation can be directly applied. It should be stated that the Kalman filter, in this case, is the optimal filter in the class of all linear filters. Let $Q = GG^t\sigma_w^2$. The filtering equations can be written as

$$\hat{x}(k, l) = \bar{x}(k, l) + M(k, l)H^t[HM(k, l)H^t + \sigma_v^2]^{-1} \cdot [z(k, l) - H\bar{x}(k, l)] \tag{9}$$

$$P(k, l) = M(k, l) - M(k, l)H^t[HM(k, l)H^t + \sigma_v^2]^{-1} \cdot HM(k, l) \tag{10}$$

$$\bar{x}(k + 1, l) = F\hat{x}(k, l) \tag{11}$$

$$M(k + 1, l) = FP(k, l)F^t + Q \tag{12}$$

where σ_w^2 and σ_v^2 are the variance of $w(k, l)$ and $v(k, l)$, respectively. The direct computation seems straightforward. However, if we take a look at the dimension of the state, we find that it is of the order $O(MN)$. Because N is usually a large number, the memory and the computational requirement for the Kalman filter is tremendous. Hence, with this kind of formulation, direct implementation is infeasible. Observing (9), we find that the filtering process consists of two parts, namely, prediction and update. From (11), we know that the prediction part just consists of state shifting and is computationally straightforward. However, the update part involves the matrix multiplications and additions which are computationally intensive. The main concept of the RUKF is that the bulk of the computation can be reduced by reducing the update process. We can only update the states which is in the region around the current processing point. Since the states outside this region are much less correlated with the current point, this approximation will only affect the optimality slightly. But, the computational savings is significant. For convenience, the update region is taken as the support of the $M \times M$ NSHP model. In this case, there are only $2M^2 + 2M$ points to be updated. We now introduce a vector corresponding to the state in the update region:

$$x_1(k, l) = [s(k, l), \cdots, s(k - M, l); \; s(k + M, l - 1), \cdots, s(k - M, l - 1); \cdots; \; s(k + M, l - M), \cdots, s(k - M, l - M)]^t. \tag{13}$$

Let the remaining elements of the state be $x_2(k, l)$. Ordering the original state, we can partition the original state into $x_1(k, l)$ and $x_2(k, l)$:

$$x(k, l) = [x_1(k, l)^t, x_2(k, l)^t]^t. \tag{14}$$

The dynamic model can then be written as

$$x(k + 1, l) = Cx(k, l) + Gw(k, l) \tag{15}$$

$$z(k, l) = Hx(k, l) + v(k, l) \tag{16}$$

where G and H are the same as defined before and C is the state transition matrix which is obtained by rearranging F according to the new states $x_1(k, l)$ and $x_2(k, l)$. Applying the same partitioning scheme, H, C, M_k, and P_k can also be partitioned accordingly:

$$H = (H_1, H_2), \; H_2 = O \tag{17}$$

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix} \tag{18}$$

$$M(k, l) = \begin{pmatrix} M_{11}(k, l) & M_{12}(k, l) \\ M_{21}(k, l) & M_{22}(k, l) \end{pmatrix} \tag{19}$$

$$P(k, l) = \begin{pmatrix} P_{11}(k, l) & P_{12}(k, l) \\ P_{21}(k, l) & P_{22}(k, l) \end{pmatrix}. \tag{20}$$

The RUKF can now be described as

$$K_1(k, l) = M_{11}(k, l)H_1^t(H_1M_{11}(k, l)H_1^t + \sigma_v^2)^{-1} \tag{21}$$

$$\hat{x}_1(k, l) = \bar{x}_1(k, l) + K_1(k, l)[z(k, l) - H_1\bar{x}_1(k, l)] \tag{22}$$

$$P_{11}(k, l) = [I - K_1(k, l)H_1]M_{11}(k, l) \tag{23}$$

$$P_{12}(k, l) = [I - K_1(k, l)H_1]M_{12}(k, l) \tag{24}$$

$$P_{21}(k, l) = P_{12}(k, l) \tag{25}$$

$$M(k + 1, l) = CP(k, l)C^t + Q_k \tag{26}$$

$$\bar{x}_1(k + 1, l) = C_{11}\hat{x}_1(k, l) + C_{12}x_2(k, l). \tag{27}$$

Although the RUKF has greatly reduced the complexity of the standard Kalman filter, it still needs an intensive computation to compute the error covariance matrix. This has been explicitly recognized in [15]. In addition, the optimality for the RUKF can only be ascertained under Gaussian noise. In the next section, we have proposed some modifications which can overcome these problems.

III. Fast Modified RUKF

In this section, we have described the modifications of RUKF in order to reduce the computational complexity furthermore, and to accommodate the non-Gaussian observation noise.

A. Reduction of Computational Complexity

The basic idea of RUKF is to reduce the update process. Although this algorithm does reduce the computational cost dramatically, the computational cost is still very high. This can be clearly understood by looking at (23)–(26) which are used to compute the Kalman gain. The reduction of computational complexity of RUKF has been considered in [15] where the steady state Kalman gain is used for different quantized level of the signal variance. Here, we propose another approach for reduction in computational complexity by reducing the state dimensionality. Note that (1) is obtained by the two-dimensional spectrum factorization. This results in a state of order $O(MN)$, which is dependent on the image size N. This situation can be avoided by redefining the state. Consider an $N \times N$ image. We first partition the image into strips (they are overlapped, see Fig. 1(b)). Let the width of a strip be L where $L > M$. Then, we redefine the state in the first strip as

$$\begin{aligned} x(k, l) = [&s(k, l), s(k - 1, l), \cdots, s(1, l); \\ &s(L, l - 1), \cdots, s(1, l - 1); \cdots; \\ &s(L, l - M), \cdots, s(k - M, l - M)]^t. \end{aligned} \tag{28}$$

The dimension of states now has the order of $O(ML)$ which is independent of the image size. After one line in a strip is processed, the filter goes to the next strip and starts all over again. Between two consecutive strips, there is a minimum overlap of $2M + 1$ pixels (Fig. 1(b)). So, the state in strip number 2 can be written as

$$\begin{aligned} x(k, l) = [&s(k, l), s(k - 1, l), \cdots, s(L - 2M - 1, l); \\ &s(2L - 2M - 1, l - 1), \cdots, \\ &s(L - 2M - 1, l - 1); \cdots; \\ &s(2L - 2M - 1, l - M), \cdots, \\ &s(k - M, l - M)]. \end{aligned} \tag{29}$$

We call this the strip RUKF. One key problem of this approach is: How to decide the initial state and the initial covariance matrix in each strip? Since the state inside the $M \times M$ support ($x_1(k, l)$) remains the same, the final $\hat{x}_1(k, l)$ and M_{11} of the previous strip can be directly used as the initial x_1 and P_{11} of the current strip. But, initial x_2, P_{12}, P_{21}, and P_{22} cannot be obtained. Here, we can use an approximation to overcome this problem. We assume that x_2 of the current strip is uncorrelated with the processed pixels, i.e., observations, in the previous strip. In other words, the observations of the previous strip do not affect the estimation of the initial x_2 of the current strip. Consequently, we can use the corresponding filtered value of the pixels in the current strip as the initial x_2. Also, we can use the same initial P_{12}, P_{21}, and P_{22} as the previous strip, i.e., fixed initial covariance matrices. Since the most important information is carried by x_1 and M_{11} and the correlation of the image decays fast, this approach is expected to work well. Specially, when the local mean of the image is extracted, the uncorrelated assumption, as made before, becomes more tenable. Simulations show that this scheme can perform almost as well as RUKF. Using sixteen strips in an 128×128 image, the complexity, as measured by elapsed CPU time, is reduced by a factor of 100 and the loss in gain is less than 0.1 dB. Although the direct comparison is not possible, the loss in gain for using the steady state Kalman gain for different values of the signal variance is around 0.5 dB. It should also be noted that for non-Gaussian noise, the update term in the Kalman filter is no longer linear. As a result, the conventional meaning of steady state Kalman gain does not hold.

The order of the total number of states for the fast RUKF is $O(ML)$. The order of the number of states for the update region is $O(M^2)$. Thus, the order of computation is $O(M^3L)$. The order of computation of the original

RUKF is $O(M^3N)$ [11] which is dependent on the image size N. Hence, the computational savings of fast RUKF is of order $O(N/L)$. For a constant L, the bigger the image size, the more computations we can save. The memory requirement is dominated by the storage of the covariance matrix. For the RUKF, this is $O(M^2N^2)$. For the fast RUKF, this storage is $O(M^2L^2)$. A memory saving of $O(N^2/L^2)$ is achieved.

B. Non-Gaussian Observation Noise

Next, we consider the case of non-Gaussian observation noise. We have mentioned before, the Kalman filter is an optimal filter only for the Gaussian noise. For the RUKF, the same argument holds. Note that in (16), the observation $z(k, l)$ is a scalar, so is the noise $v(k, l)$. The H matrix is $(1, 0, \cdots, 0)$. Due to this special structure, the observation prediction density is still univariate. Hence, we can apply the one-dimensional score function filtering scheme developed in [12]–[14] to accommodate the non-Gaussian observation noise.

The score function based scheme was first proposed by Masreliez [14]. Consider a linear system depicted as follows:

$$x_{k+1} = \phi_k x_k + w_k \tag{30}$$

$$z_k = H_k x_k + v_k. \tag{31}$$

Let Z^{k-1} stand for $(z_1, z_2, \cdots, z_{k-1})$, w_k be Gaussian with $E\{w(k)w(j)\} = Q_k\,\delta_{kj}$, and v_k be non-Gaussian. Assuming that $f(x_k|Z^{k-1})$ is a Gaussian density with mean $\bar{x}_k$ and covariance M_k, and $f(z_k|Z^{k-1})$ twice differentiable, Masreliez has shown that the minimum variance estimation of the state $\hat{x}_k = E(x_k|Z^k)$ and its covariance $P_k = E\{(\hat{x}_k - x_k)(\hat{x}_k - x_k)^t|Z^k\}$ can be recursively calculated as follows:

$$\hat{x}_k = \bar{x}_k + M_k H_k^t g_k(z_k) \tag{32}$$

$$P_k = M_k - M_k H_k^t G_k(z_k) H_k M_k \tag{33}$$

$$M_{k+1} = \phi_k P_k \phi_k^t + Q_k \tag{34}$$

where $g_k(\cdot)$ is called the score function and $G_k(\cdot)$ is the derivative of $g_k(\cdot)$. They are defined as

$$g_{k,i}(z_k) = -\left[\frac{\partial f(z_k|Z^{k-1})}{\partial(z_k)_i}\right][f(z_k|Z^{k-1})]^{-1} \tag{35}$$

$$G_{k,ij}(z_k) = \frac{\partial\{g_{k,i}(z_k)\}}{\partial(z_k)_j} \tag{36}$$

where $g_{k,i}$ is the ith component of g_k and $G_{k,ij}$ is the ijth component of G_k. It can be shown that the score function, which is operating on the residue signal $z_k - H_k\bar{x}_k$, will deemphasize the large residue if the noise distribution is long tailed and emphasize the large residue if the noise distribution is short tailed. This is intuitively appealing. In the special case of Gaussian observation noise, the score function becomes a linear function of the innovation. As a result, the score function approach will be reduced to standard Kalman filtering. This can be easily proved in the following way. Let v_k be Gaussian and its variance be R_k. Denote a Gaussian density by $N(a, b)$ where a is the mean and b is the variance. Then

$$f(x_k|Z^{k-1}) \sim N(\bar{x}_k, M_k) \tag{37}$$

$$f(z_k|Z^{k-1}) \sim N(H_k\bar{x}_k, H_kM_kH^t + R_k) \tag{38}$$

$$g_k(z_k) = (H_kM_kH^t + R_k)^{-1}(z_k - H_k\bar{x}_k) \tag{39}$$

$$G_k(z_k) = (H_kM_kH^t + R_k)^{-1}. \tag{40}$$

Substituting (39), (40) into (32)–(34), the standard Kalman filter equations are obtained.

Masreliez's algorithm is directly applicable to the RUKF problem. Let Z^{k-1} stand for data up to the point $(k-1, l)$, $g(\cdot)$ for the score function of $f(z(k, l)|Z^{k-1})$, and $G(\cdot)$ be the derivative of $g(\cdot)$. Now, the RUKF can be rewritten as

$$\hat{x}_1(k, l) = \bar{x}_1(k, l) + M_{11}(k, l)H_1^t g(z(k, l)) \tag{41}$$

$$P_{11}(k, l) = M_{11}(k, l) - M_{11}(k, l)H_1^t G(z(k, l)) \cdot H_1M_{11}(k, l) \tag{42}$$

$$P_{12}(k, l) = M_{12}(k, l) - M_{12}(k, l)H_1^t G(z(k, l)) \cdot H_1M_{12}(k, l) \tag{43}$$

$$P_{21}(k, l) = P_{12}(k, l) \tag{44}$$

$$M(k+1, l) = CP(k, l)C^t + Q \tag{45}$$

$$\bar{x}_1(k+1, l) = C_{11}\hat{x}_1(k, l) + C_{12}x_2(k, l). \tag{46}$$

Although the RUKF is applied to two-dimensional image signal, from the above discussion, we know that the score function is still one dimensional. Consequently, the computation turns out to be simple and straightforward.

C. Computation of the Score Function

The main thrust of the computation of the score function based scheme is the estimation of $f(z_k|Z^{k-1})$ which is obtained by a convolution of $f(H_kx_k|Z^{k-1})$ and $f(v_k)$. The convolution is very difficult to implement except for very simple cases. In [12], we have developed an efficient approximation scheme for the computation of the score function of a distribution in the scalar case. The method employs an adaptive normal expansion to expand the score function and truncates the higher order terms in the expanded series. Consequently, the score function can be approximated by a few moments of the distribution. It is shown in [13] that the approximation scheme works well for medium-tailed to long-tailed distributions.

Consider a random variable z with distribution $f(z)$ and the moment generating function $M(T)$. Let $K(T) = \ln(M(T))$. The scheme uses the concept of the conjugate density and the saddle point to approximate the score function at a given point z in the following manner [13], [22].

1) Replace the original $f(z)$ by the conjugate density $f_z(\cdot)$:

$$f_z(s) = c_z e^{a_z s} f(z + s) \tag{47}$$

where c_z and a_z are chosen such that $f_z(\cdot)$ is a probability density with expectation 0. It has been shown [13], [22] that if a_z is chosen as the saddle point of $M(T)e^{-Tz}$ and $1/c_z$ as $M(a_z)$, $f_z(\cdot)$ is the conjugate density. The saddle point is defined as the solution of the equation

$$K'(T) - z = 0. \tag{48}$$

2) Calculate the second, third, and fourth moment of $f_z(\cdot)$:

$$\sigma_{2,z} = \int_{-\infty}^{+\infty} s^2 f_z(s)\, ds, \tag{49}$$

$$\sigma_{3,z} = \int_{-\infty}^{+\infty} s^3 f_z(s)\, ds, \tag{50}$$

$$\sigma_{4,z} = \int_{-\infty}^{+\infty} s^4 f_z(s)\, ds. \tag{51}$$

If the saddle point is found, $\sigma_{2,z}$, $\sigma_{3,z}$, and $\sigma_{4,z}$ can be easily calculated as follows:

$$\sigma_{2,z} = K''(a_z) \tag{52}$$

$$\sigma_{3,z} = K'''(a_z) \tag{53}$$

$$\sigma_{4,z} = K''''(a_z). \tag{54}$$

3) Using the normal expansion for the conjugate density, it is shown [13] that the score function of the original distribution and its derivative can be approximated by

$$g(z) \approx a_z + \frac{\sigma_{3,z}}{2\sigma_{2,z}^2} \tag{55}$$

$$G(z) \approx \frac{1}{\sigma_{2,z}}\left(1 + \frac{\sigma_{4,z}}{2\sigma_{2,z}^2} - \frac{\sigma_{3,z}^2}{\sigma_{2,z}^3}\right). \tag{56}$$

For the present problem, we want to approximate the score function and its derivative for the density $f(z(k, l)\,|\,Z^{k-1})$. This density can be obtained by convolving the density $f(H_1 x_1(k, l)\,|\,Z^{k-1})$ and the observation noise density. Since $H_1 = (1, 0, \cdots)$, $f(H_1 x_1(k, l)\,|\,Z^{k-1})$ is just the density of the first state element in $x_1(k, l)$. The density $f(x_1(k, l)\,|\,Z^{k-1})$ is assumed to be Gaussian [13], [14]. Subsequently, the variance of the first state element in $x_1(k, l)$, given the data up to the point $(k - 1, l)$, is just the first element of M_{11}; and the mean is given by the first element of $\bar{x}_1(k, l)$. To evaluate the score function, one needs to know the moment generating function (MGF) of $f(z(k, l)\,|\,Z^{k-1})$. This MGF can be obtained by the multiplication of the MGF of the noise density ($M_r(T)$) and the MGF of the $f(H_1 x_1(k, l)\,|\,Z^{k-1})$, i.e., $M_x(T)$. If no closed form of the MGF is available, we assume that some approximate form, say in terms of rational Chebyshev approximation [13], is available. Now, let

$$z = z(k, l) \tag{57}$$

$$M(T) = M_r(T) M_x(T) \tag{58}$$

and

$$K(T) = \ln\,(M(T)). \tag{59}$$

Substituting (57)–(59) into (48), and (52)–(56), the score function and its derivative for $f(z(k, l)\,|\,Z^{k-1})$ are obtained.

IV. Computations of Nonstationary Mean

Because of the stationarity assumption of the image model, the Kalman filter can smear the edges. Understanding the problem, Jeng and Woods [15] proposed an approach which can reduce this effect. The main idea is to subtract a nonstationary mean component to convert the nonstationary signal into a residual signal. To do so, the local mean of an image is subtracted before filtering and the residual image is then filtered. After filtering, the local mean is added back to form the output. In [15], the averaging filter is used to estimate the local mean. Although this approach does show some improvement, the strong edges and the details are not well preserved by the averaging filter. As a result, the Jeng and Woods model has to incorporate a nonstationary variance in the image model. Since our image model assumes a stationary variance, it is absolutely essential to estimate the mean by an edge and detail preserving filter which also has a high smoothing efficiency over the flat regions of the image. For this reason, we propose a new filter which can be efficiently estimate the local mean such that the edges and the details are well preserved. This filter combines the merits of the multistage median (MSM) [16] filter and the Hodges–Lehman D filter [17]. We call the new filter hybrid-MSM-D filter, or HMSMD filter. It has been shown that the MSM filter has very good detail preserving ability, but its smoothing efficiency is poor. On the contrary, the D filter has very good smoothing efficiency, but its detail preserving ability is poor. Combining these two filters, we are able to obtain an efficient edge and detail-preserving filter. The definitions of the D and the MSM are described in the following.

A. Definition: Multistage Median (MSM) Filter

Let $s(.,.)$ be a discrete image sequence. Consider the samples inside a $(2L + 1) \times (2L + 1)$ window which is centered at (i, j). Define four subwindows as

$$W_1(i, j) = \{s(i + k, j):\ -L \le k \le L\} \tag{60}$$

$$W_2(i, j) = \{s(i + k, j + k):\ -L \le k \le L\} \tag{61}$$

$$W_3(i, j) = \{s(i, j + k):\ -L \le k \le L\} \tag{62}$$

$$W_4(i, j) = \{s(i + k, j - k):\ -L \le k \le L\}. \tag{63}$$

Let

$$z_l(i, j) = \text{median}\,[s(.,.) \in w_l(i, j)] \qquad 1 \le l \le 4 \tag{64}$$

$$y_{\max}(i, j) = \max_{1 \le l \le 4} [z_l(i, j] \tag{65}$$

$$y_{\min}(i, j) = \min_{1 \le l \le 4} [z_l(i, j)]. \tag{66}$$

The output of the MSM at (i, j) is defined as

$$y(i, j) = \text{median}\,[y_{\max}(i, j), y_{\min}(i, j), s(i, j)]. \tag{67}$$

B. Definition: D Filter

Consider an window of size $M \times M$ which scans the image in a raster scan sense. Let $M^2 = n$. The n pixel values inside the window are next sorted in the ascending order of magnitude as $y_{(1)}, \cdots, y_{(n)}$. Thus, $y_{(1)}$ is the minimum order statistic, and $y_{(n)}$ is the maximum order statistic, respectively. The D filter is given by

$$z(i) = (0.5 \cdot y_{(i)} + 0.5 \cdot y_{(n-i+1)}) \tag{68}$$

$$\hat{X}_0 = \text{median}_i\,(z(i), i = 1, \cdots, (n-1)/2 + 1) \quad \text{for } n \text{ odd} \tag{69a}$$

$$\hat{X}_0 = \text{median}_i\,(z(i), i = 1, \cdots, n/2) \quad \text{for } n \text{ even.} \tag{69b}$$

We now combine the MSM filter with the D filter in the manner described below:

C. Hybrid-MSM-D Filter

Consider an window of size $M \times M$ which scans the image in a raster scan sense. Let $M^2 = n$. Let X_0 be the filtered value of all the data inside the window using the MSM filter. Define R as the range $(X_0 - q, X_0 + q)$ where q is a preselected threshold. Next, all the pixel values inside the window and falling in this range R are selected. Let their number be m. These m pixel values are next sorted in the ascending order of magnitude as $y_{(1)}, \cdots, y_{(m)}$. Thus, $y_{(1)}$ is the minimum order statistic, and $y_{(m)}$ is the maximum order statistic, respectively. The HMSMD filter is given by

$$z(i) = (0.5 \cdot y_{(i)} + 0.5 \cdot y_{(m-i+1)}) \tag{70}$$

$$\hat{X}_0 = \text{median}_i\,(z(i), i = 1, \cdots, (m-1)/2 + 1) \quad \text{for } m \text{ odd} \tag{71a}$$

$$\hat{X}_0 = \text{median}_i\,(z(i), i = 1, \cdots, m/2) \quad \text{for } m \text{ even.} \tag{71b}$$

The value of q is directly related to the noise statistics. In a quasi-constant region, we wish to include all the data in the window to estimate the signal value. Hence, in the noise filtering problem, a reasonable choice for q is in between ($2\sigma_n$ to $3\sigma_n$) where σ_n is the noise standard deviation. For the local mean estimation problem, we have to take the signal variance into account also. Thus, the choice for q should be in the range of ($2\sqrt{\sigma_n^2 + \sigma_s^2}$ to $3\sqrt{\sigma_n^2 + \sigma_s^2}$) where σ_n^2 is the signal variance, σ_s^2 can be obtained from the residual signal of the least square fitting of the image (to find the NSHP coefficients).

V. Experimental Results

To summarize, the overall filtering scheme proposed in this paper is described as follows:

1) Use the HMSMD filter to estimate the mean.
2) Subtract the mean component to obtain the residual image.
3) Filter the residual image using the fast modified RUKF.
4) Restore the nonstationary mean.

Simulations are carried out to test the performance of the modified RUKF. An 128×128 real image is chosen. First we filter the noisy image using the HMSMD filter, and then use the filtered image to identify the coefficients of the NSHP model. The order of NSHP is set to be 2×2 as it is found that a higher order NSHP model is not necessary. The least square fitting method is used to compute the NHSP model coefficients. For all strip RUKF experiments, the width of strip overlap is 5, i.e., $2M + 1$ with $M = 2$. The signal-to-noise ratio and the mean-square error are chosen as the performance index. They are defined as follows:

$$\text{SNR} = 10 \log \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} s(i, j)^2}{\sum_{k=1}^{N} \sum_{l=1}^{N} [r(k, l) - s(k, l)]^2} \tag{72}$$

$$\text{MSE} = \frac{1}{N^2} \sum_{k=1}^{N} \sum_{l=1}^{N} [r(k, l) - s(k, l)]^2 \tag{73}$$

where $s(k, l)$ is the original uncorrupted image, $r(k, l)$ is the filtered image, and N is the size of the image. The simulation is implemented on VAX 8800 and the elapsed CPU time is chosen as the computational complexity measure. Three comparisons are made.

1) Comparison of the Original RUKF and the Strip RUKF: A Gaussian white noise with variance of 225 is added to the original image. The HMSMD filter is used to estimate the local means. The threshold for HMSMD is set at 45. The window for computing the local mean is set to be 5×5. The SNR of the corrupted image is 17.94, and the MSE is 220.1. Table I shows the elapsed CPU time for RUKF with different number of strips. Note that when one strip is used, the strip RUKF reduces to the original RUKF.

2) Comparison of the Averaging and the HMSMD Local Mean Estimator: The same noisy image is used to test the performance of the RUKF under the new local mean estimator. Sixteen strips are used in the RUKF and the threshold value for HMSMD is chosen as 45. Table II shows the SNR and MSE comparison. Figs. 2(a)–(d) show the original, the noisy, and the filtered images.

3) Comparison of the RUKF and the Score Function Based RUKF: Sixteen strips are used in the RUKF, and HMSMD is used to estimate the local mean. One percent impulse noise is added to the image already corrupted by zero-mean additive Gaussian noise with variance 100. The

TABLE I
COMPARISON OF RUFF AND STRIP RUKF CPU TIME: (HOURS : MINUTES : SECONDS)

No. Strip	SNR	MSE	CPU Time
1	22.93	69.87	02 : 40 : 45.96
4	22.89	70.49	00 : 08 : 50.83
16	22.88	70.58	00 : 01 : 13.35

TABLE II
COMPARISON OF DIFFERENT LOCAL MEAN ESTIMATORS

Filters	SNR	MSE
HMSMD	22.88	70.58
Averaging	21.94	87.52

impulse noise is modeled as a long-tailed Laplacian noise with mean zero and the parameter 30. The gray scale of the image is 0-255. The SNR of the noisy image is 16.54, and the MSE is 304.2. Table III shows the SNR and MSE comparison. Figs. 3(a)-(c) show the noisy, and the filtered images.

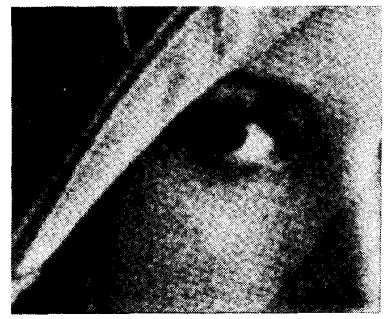
We next generate two synthetic images of size 50×50 using a DPCM image model. Let i denote the vertical coordinate and j denote the horizontal coordinate. The image model can be described as follows:

$$s(i, j) = \rho_h s(i, j - 1) + \rho_v s(i - 1, j) - \rho_h \rho_v s(i - 1, j - 1) + w(i, j) \quad (74)$$

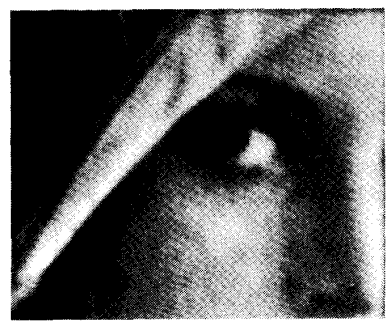
where ρ_h is the horizontal correlation coefficient, ρ_v is the vertical correlation coefficient, and $w(i, j)$ is a white Gaussian noise. In our experiment, we use $\rho_h = \rho_v = 0.9$ and $w(i, j) \sim N(0, 0.8)$. Two images are generated. One is a quarter disk and the other one is a thin ring. The means of the quarter disk and the thin ring are 180 and 150, respectively. The mean of the background is 100. To generate the quarter disk, first a 2-D signal of constant value 80 only encompassing the disk position is created. Next a square background of image size 50×50 with mean 100 is created using the DPCM model of (74). The two images are added together to create the quarter disk image. Similarly, the thin ring image is created. The parameter estimation and the strip overlapping for modified RUKF are the same as the previous experiments. A Gaussian plus impulse (1%) mixture type noise is added to the images. The variance of the Gaussian noise is 50. The original RUKF filter and the modified RUKF filter are applied to these images. The original RUKF uses the averaging filter as the local estimator of the mean. Figs. 4(a)-(d) and 5(a)-(d) show the graphic display of the results. Tables IV and V summarize the SNR and MSE comparison results.

It can be seen from the simulations that the modified RUKF greatly improves the performance of the original RUKF. For 128×128 images, the computation cost of the strip RUKF (16 strips) is less than 1% of the original RUKF. As discussed in Section III, even larger computational savings can be achieved with larger images. The SNR gain is about 1 dB using the new HMSMD local

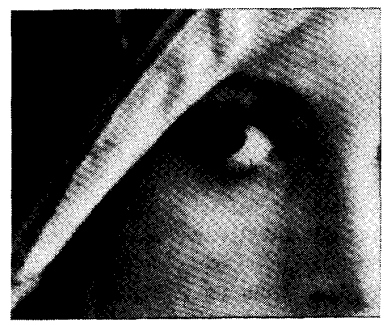
(a)



(b)



(c)

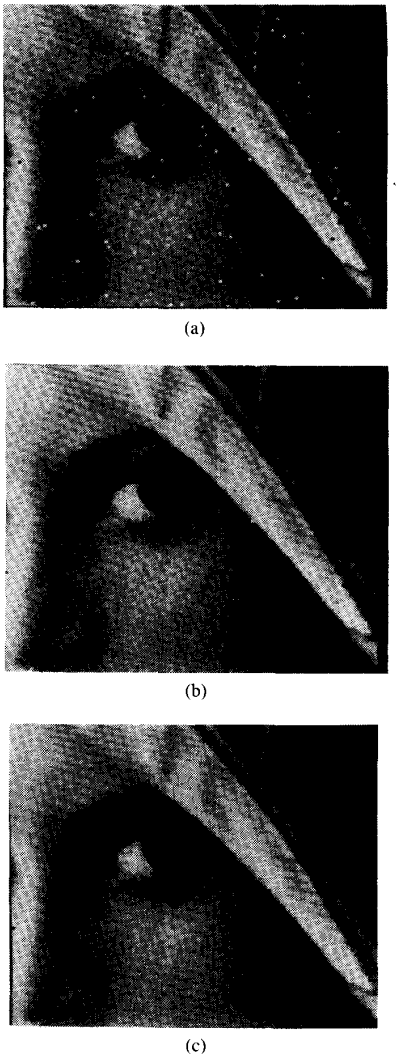


(d)

Fig. 2. (a) The original image. (b) The corrupted image: Gaussian additive noise. (c) RUKF filtered image: averaging local mean estimator. (d) RUKF filtered image: HMSMD local mean estimator.

TABLE III
COMPARISON OF RUKF AND SCORE FUNCTION BASED RUKF

Filters	SNR	MSE	CPU Time
RUKF	23.02	68.37	00:01:08.47
SRUKF	24.33	50.54	00:01:11.37



(a) (b) (c)

Fig. 3. (a) The corrupted image: Gaussian + impulse noise. (b) RUKF filtered image. (c) Score function based RUKF filtered image.

mean estimator. Also, the original RUKF cannot remove the impulse noise. Scattered white and dark points can be easily seen after filtering. However, the score function based RUKF efficiently removes both the Gaussian and the impulse noise. The score function based filtering is most appropriate when the additive noise deviates markedly from the Gaussian form. As a result, this approach is expected to work very well for many mixture type noise. In our experiments, the gain in SNR due to score function based approach with 1% impulse noise is around 1.3 dB.

The overall SNR gain using the modified RUKF, as described in this paper, is around 8 dB for natural images with input SNR of 16 dB or so. For synthetic images with input SNR of 16–17 dB or so, the gain improvement is around 12 dB. These figures compare very well with the existing schemes in Gaussian environment. For non-Gaussian observation noise, however, the new scheme outperforms the existing schemes.

The experiments reported so far use the RUKF only once. The coefficients of the NSHP can be reestimated after the filtering is done once. The new coefficients could be used to do the filtering again. This process can be iterated until no further change is found in the coefficient values. However, experiments show that if the q value is chosen properly, the iterated procedure does not improve the filtering performance in any significant manner. There could even be slight loss of SNR performance in such iterated procedure.

Although the RUKF is approximated by the strip RUKF, there is no artifact in the filtered images because the strips are not processed independently. Many parameters are passed from the previous strip to the next strip. In addition, the image is processed by a raster scan instead of strip by strip. If the correlation of the image is small outside the update region, the strip RUKF is essentially the same as the RUKF.

In our image model, we have assumed that the residual signal has a stationary variance if the nonstationary mean is estimated by an edge and detail preserving filter with good smoothing efficiency. Figs. 6(a) and (b) show the residual signal of Figs. 4(b) and 5(b), respectively. Looking at Figs. 6(a) and (b), it is obvious that the HMSMD filter does a good job in estiamting the local mean. Both Figs. 6(a) and (b) give the impression of stationary variance. Also, the impulses are exactly replicated in the residual signal as they are part of the original signal. The stationary variance does not need to take care of these impulses as the score function based scheme can reject these impulses. Figs. 6(c) and (d) show the residual signal of Figs. 4(b) and 5(b) without the impulses. These figures vindicate that the stationary variance assumption, in our case, is reasonable.

A. Comparison with Order Statistics Filter

It is well known that the MSM filter preserves the edges and the details very well. The HMSMD filter incorporates the superior smoothing property of D filter with the detail preserving ability of the MSM filter. With the proper choice of q, this filter can perform at the same level as any other order statistics filter. In the following, we have described some experiments to compare the performance of the HMSMD filter and the modified RUKF filter. A white Gaussian noise with mean 0 and variance 225 is added to two images, namely, "girl" and "bridge." The SNR and MSE of the noisy images are 17.94/220.1 and

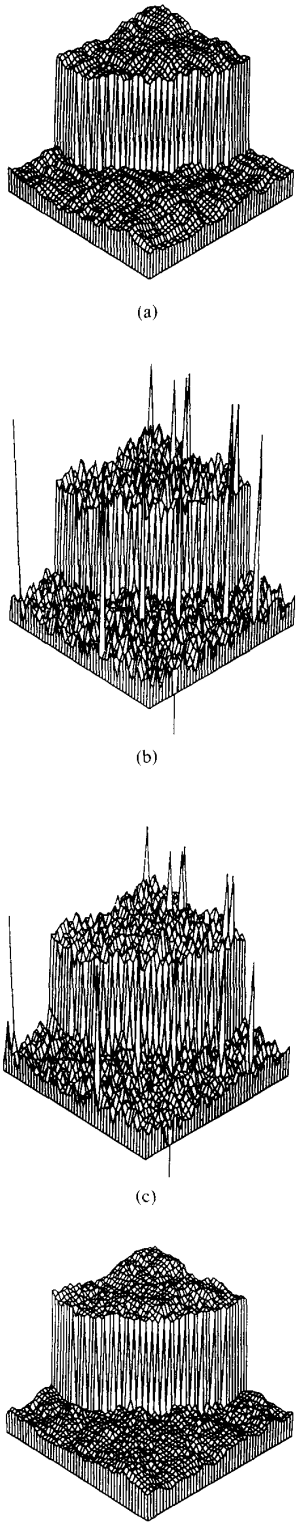


Fig. 4.(a) A synthetic image containing a sharp edge. Size 50 × 50. (b) Noise corrupted image of Fig. 4(a). The additive noise is Gaussian + impulse. (c) RUKF filtered image. (d) Fast modified RUKF filtered image.

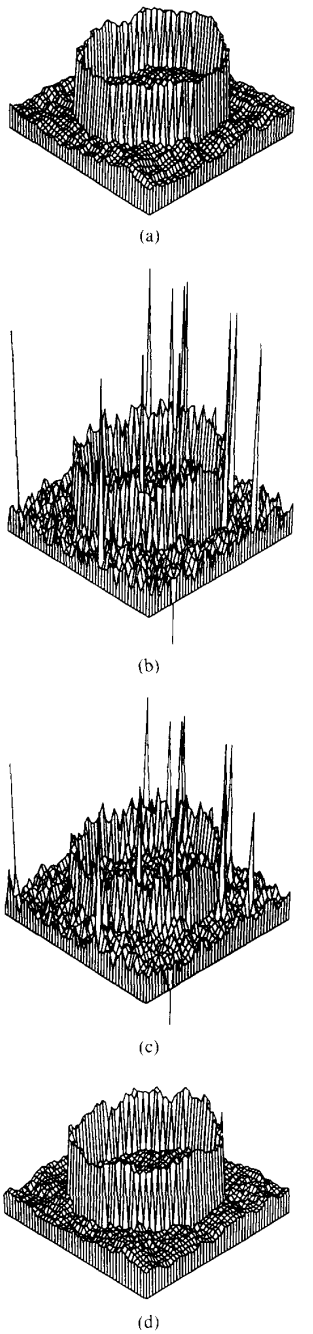


Fig. 5. (a) A synthetic image containing a sharp thin line signal. Size 50 × 50. (b) Noise corrupted image of Fig. 5(a). The additive noise is Gaussian + impulse. (c) RUKF filtered image. (d) Fast modified RUKF filtered image.

21.82/219.1, respectively. When 1% impulse noise is added to Gaussian noise corrupted images, the SNR and MSE are 15.11/422 and 18.71/448.9, respectively. The q value is selected as 35 for the "girl" image and 30 for the "bridge" image. The window size is 3 × 3 for the HMSMD. Within the constraints of the given experi-

TABLE IV
COMPARISON OF RUKF AND MRUKF FOR THE QUARTER DISK IMAGE

Q-Disk	SNR	MSE
Noisy	19.55	204.2
RUKF	22.05	114.8
MRUKF	34.42	6.665

TABLE V
COMPARISON OF RUKF AND MRUKF FOR THE THIN RING IMAGE

Thin-Ring	SNR	MSE
Noisy	16.89	222.3
RUKF	19.47	122.7
MRUKF	31.26	8.144

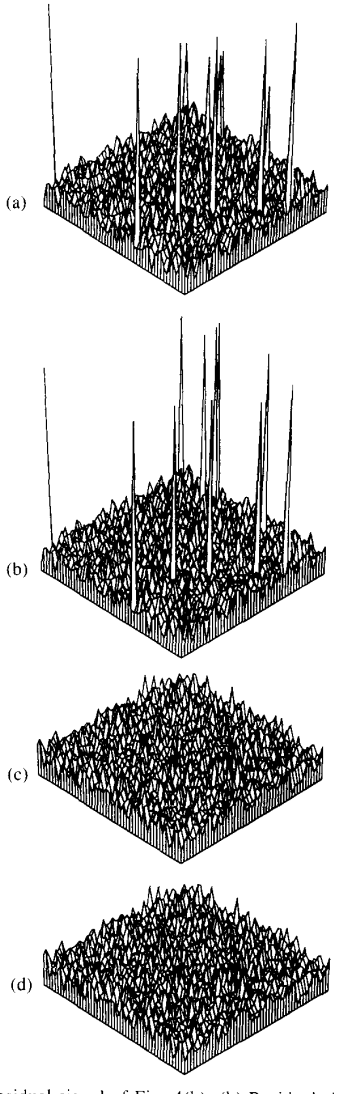


Fig. 6. (a) Residual signal of Fig. 4(b). (b) Residual signal of Fig. 5(b). (c) Residual signal of Fig. 4(b) without the impulses. (d) Residual signal of Fig. 5(b) without the impulses.

TABLE VI
COMPARISON OF HMSMD AND MRUKF FOR "GIRL" IMAGE

Filters	SNR/MSE	SNR/MSE
HMSMD	22.15/83.61	21.93/87.81
MRUKF	22.88/70.58	22.66/74.30

TABLE VII
COMPARISON OF HMSMD AND MRUKF FOR "BRIDGE" IMAGE

Filters	SNR/MSE	SNR/MSE
HMSMD	23.81/128.9	23.58/146.4
MRUKF	24.53/117.5	24.17/127.8

Fig. 7. HMSMD filtered image of Fig. 3(a).

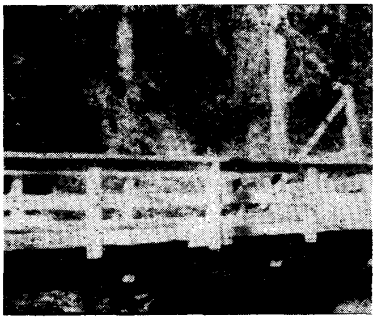


Fig. 8. Noise corrupted image "bridge" as used in experiments.

ments, these parameters give the best or nearly the best smoothing as well as good detail preservation. The parameter of modified RUKF are selected as described before. Tables VI and VII summarize the results. The right-hand column of the tables are the SNR/MSE for the filtered images when the noise also includes the impulse noise.

Fig. 7 shows the HMSMD filtered image of "girl." Fig. 8 shows the noise corrupted "bridge" image as used in our experiments. From the tables, we know that the SNR improvement of modified RUKF over HMSMD is around 0.7 dB. Also, the RUKF filtered image is visually more appealing. In the quasi-constant region, the modi-

fied RUKF provides more smoothing and, in regions of large spatial variation, it gives more edge and detail preservation. Another important aspect of modified RUKF is that it can be extended to image restoration [15].

It is possible to improve the performance of the modified RUKF furthermore by relaxing the stationary variance assumption. The idea is to segment the image into regions and then use the NSHP model with identified coefficients to estimate the AR driving noise variance in each region. When the filter is running, the corresponding variance is used for each region. This scheme is similar to the one described in [15]. This is a research topic for further investigation.

VI. Conclusions

In this paper, we have proposed some modifications of RUKF as applied to the estimation of noisy image. These modifications are in relation to the reduction of complexity, and the incorporation of non-Gaussian noise. We have also shown that by using an edge and detail preserving efficient local estimator of the nonstationary mean, a simpler image model can be used to achieve a comparable performance with respect to the existing schemes. The overall filtering scheme proposed in this paper has a gain improvement performance comparable to the existing schemes in Gaussian environment. Moreover, as the new scheme is at least suboptimal in non-Gaussian environment, the new scheme outperforms the existing schemes in non-Gaussian environment. Although a general conclusion cannot be made, the new filtering scheme also has better performance than some selected order statistics filters.

Following the outlines given in [15], the new scheme can be applied to pictures degraded by linear blur with non-Gaussian additive noise. This problem is under investigation.

References

[1] H. C. Andrews and H. R. Hunt, *Digital Image Restoration.* Englewood Cliffs, NJ: Prentice-Hall, 1977.

[2] A. K. Jain, *Fundamental of Digital Image Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1989.

[3] W. K. Pratt, "Generalized Wiener filter computation techniques," *IEEE Trans. Comput.*, vol. C-21, pp. 636-641, July 1972.

[4] B. R. Hunt and T. M. Cannon, "Nonstationary assumptions for Gaussian models of images," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 876-881, 1976.

[5] S. A. Rajala and R. J. P. de Figueiredo, "Adaptive nonlinear image restoration by a modified Kalman filtering approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1033-1042, Oct. 1981.

[6] V. K. Ingle and J. W. Woods, "Multiple model recursive estimation of images," in *Proc. ICASSP '76* (Washington, DC), 1976, pp. 642-645.

[7] C. L. Anderson and A. N. Netravali, "Image restoration based on subjective criterion," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-6, pp. 845-853, Dec. 1976.

[8] J. F. Abramatic and L. M. Silverman, "Nonlinear restoration of noisy images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-4, no. 2, pp. 141-149, 1982.

[9] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel, "Adaptive noise smoothing filter for images with signal dependent noise," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-7, no. 2, pp. 165-177, 1985.

[10] J. S. Lee, "Digital image enhancement and noise filtering by use of local statistics," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-2, no. 2, pp. 165-168, 1980.

[11] J. W. Woods and C. H. Radewan, "Kalman filtering in two dimensions," *IEEE Trans. Inform. Theory*, vol. IT-23, no. 4, pp. 473-482, 1977.

[12] W. R. Wu and A. Kundu, "Kalman filtering in non-Gaussian environment using efficient score function approximation," in *Proc. ISCAS* (Oregon), 1989, pp. 413-416.

[13] W. R. Wu, "Kalman filtering in non-Gaussian environment using efficient score function approximation," Ph.D. dissertation, SUNY, Buffalo, Aug. 1989.

[14] C. J. Masreliez, "Approximate non-Gaussian filtering with linear state and observation relations," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 107-110, 1975.

[15] F. C. Jeng and J. H. Woods, "Inhomogeneous Gaussian image model for estimation and restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 1305-1312, Aug. 1988.

[16] G. R. Arce and R. E. Froster, "Detail-preserving ranked order based filters for image processing," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 1, pp. 83-98, 1989.

[17] A. Kundu and W. R. Wu, "Double-window Hodges Lehman (D) filter and hybrid D-median filter for robust-image smoothing," *IEEE Trans. Acoust., Speech, Signal Processing*, pp. 1293-1298, Aug. 1989.

[18] H. T. Trussel and B. R. Hunt, "Sectioned models for image restoration," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, pp. 157-164, 1978.

[19] C. W. Helstrom, "Image restoration by method of least squares," *J. Opt. Soc. Amer.*, vol. 57, no. 3, pp. 297-303, 1967.

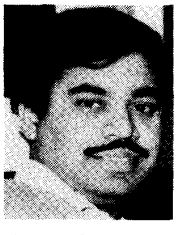
[20] A. Habibi, "Two-dimensional Bayesian estimate of images," *Proc. IEEE*, vol. 60, no. 7, pp. 878-883, 1972.

[21] N. E. Nahi, "Role of the recursive estimation in statistical image enhancement," *Proc. IEEE*, vol. 60, no. 7, pp. 872-877, 1972.

[22] H. E. Daniels, "Saddlepoint approximation in statistics," *Ann. Math. Stat.*, vol. 25, pp. 631-650, 1954.

Wen-Rong Wu (S'87-M'89) was born in Taiwan, Republic of China, in 1958. He received the B.S. degree in mechanical engineering from Tatung Institute of Technonlogy, Taiwan, in 1980, the M.S. degree in mechanical and electrical engineering, and the Ph.D. degree in electrical engineering from the State University of New York at Buffalo in 1985, 1986, and 1989, respectively.

Since August 1989 he has been a faculty member in the Department of Communication Engineering at National Chiao Tung University, Taiwan. His research interests include detection and estimation theory, digital signal processing, and image processing.



Amlan Kundu (S'83-M'84) was born in Calcutta, India, on April 8, 1955. He received the B.Tech. degree in electronics from Calcutta University in 1979, and the M.S. and Ph.D. degrees from the University of California, Santa Barbara, in 1982 and 1985, respectively.

He was with the Department of Electrical Engineering at the State University of New York at Buffalo. Currently, he is affiliated with the Center for Document Analysis and Recognition (CEDAR) SUNY at Buffalo. His research interests include various aspects of image processing, including biomedical image processing, image and signal filtering in non-Gaussian environments, pattern recognition, and VLSI implementation of signal processing algorithms. He has written about 40 scientific papers in these fields.