KYBERNETIKA — VOLUME 35 (1999), NUMBER 6, PAGES 777-792

ON THE OPTIMALITY OF A NEW CLASS OF 2D RECURSIVE FILTERS[1]

LEOPOLDO JETTO

The purpose of this paper is to prove the minimum variance property of a new class of 2D, recursive, finite-dimensional filters. The filtering algorithms are derived from general basic assumptions underlying the stochastic modelling of an image as a 2D gaussian random field. An appealing feature of the proposed algorithms is that the image pixels are estimated one at a time; this makes it possible to save computation time and memory requirement with respect to the filtering procedures based on strip processing. Experimental results show the effectiveness of the new filtering schemes.

1. INTRODUCTION

The considerable attention that the image restoration problem has been receiving in the literature has motivated the interest for extending optimal one-dimensional (1D) filtering procedures to two-dimensional (2D) data fields. In particular most authors investigated the applicability of Kalman filtering techniques to the restoration of images corrupted by additive noise.

In [3, 11, 14, 21, 22, 23, 24, 26, 29, 32] a 2D image is transformed into 1D scalar or vector stochastic process using a line-by-line scan or a vector scanning scheme. Other approaches that use a 2D model can be found in [1, 9, 10, 12, 15, 28]. All these papers are based on the common assumption that the image is the realization of a wide sense stationary random field. By this simplifying hypothesis an image model suitable to a state-space representation can be derived; nevertheless the corresponding space-invariant filters are insensitive to abrupt changes in the image signal and give restored images with reduced contrast and blurred edges. Actually, a real image is composed of an ensemble of several different regions and, in general, no correlation among them may be assumed. Thus the stationarity assumption may fit for the statistics of each single region, but not for the whole image; consequently blurring and oversmoothing phenomena occur at edge locations.

Adaptive space-variant filters based on identification-estimation algorithms have been proposed in [2, 13, 16, 17, 18, 30, 31, 33]. These methods allow the parameters describing the image model to vary inside the image itself according to the local

[1] This work was supported by Ministero dell'Università e della Ricerca Scientifica e Tecnologica.

statistics. The specific problem of reducing the numerical complexity involved in the adaptive parameter estimation procedures for a 2D image model is considered in [34].

All the above mentioned papers are based on a description of the image in terms of its statistical properties. The self-tuning methods attempt to draw this information starting from noisy data. Their main drawbacks are the computational cost, that in many cases may be unacceptable, and/or the increased structural complexity of the algorithm. Moreover it seems difficult to obtain a fast switching of the image model parameters in correspondence of a sudden change in the image statistics, such as at edge points. The other methods assume that the information on the image model is available *a priori* or that it can be obtained from the noisy free image or by a sample of similar pictures. In many practical situations it is unrealistic to assume that these data are available.

Based on the results of [5] and [8], the method proposed in [7] starts from a completely different point of view. The image modelling proposed in [7] is based on the following assumptions.

Smoothness assumption. The image is modelled as the union of open disjoint subregions whose interior is regular enough to be well described by a 2D surface of class $C^{\bar{n}}$.

Stochastic assumption. All the derivatives of order $\bar{n}+1$ of the 2D signal are modelled by means of zero-mean independent Gaussian random fields.

Inhomogeneity assumption. The random fields representing the image process relative to different subregions are independent.

As shown in [7], these hypotheses allow one to construct a space-variant image model where the problem of image parameter identification is greatly simplified and where the presence of image edges is intrinsically taken into account, so that edge oversmoothing is automatically avoided.

The space-variant Kalman filter derived from the previous assumptions is implemented in [7] by partitioning the image into parallel strips according to the procedure described in [32]. Strip filtering may be computationally attractive because it does not require the definition of a state vector of dimension equal to that resulting from considering the whole image and allows one to embed a 2D filtering problem into a 1D algorithm, moreover it does not have some of the undesirable nonstationary characteristics of line scanning. Nevertheless, strip filtering suffers from some inconveniences. First, filtered estimates of each strip are obtained neglecting the information carried by the pixels lying on the other strips; second, the computational cost may still be high. For example, the state vector defined in [5, 7, 8] associated with each pixel is composed of the image signal and its partial spatial derivatives up to the order $\bar{n}$, hence it is composed of $N=(\bar{n}+1)(\bar{n}+2)/2$ elements. Therefore, if the image is partitioned into strips of width L, the state vector involved in the strip Kalman filtering has dimensions LN. Taking into account that only the L_R

middle points are retained as final estimates to avoid edge effects, this implies a number of computation per points $\mathcal{O}((LN)^3/L_R)$ [32], that in some cases may be unacceptable. This computational burden can not be reduced exploiting the simplifying assumption of a spatial invariant structure of the signal process, as in [32], because the image model proposed in [7] is space-varying.

In this paper, two point-to-point recursive filtering algorithms are proposed and their optimality (in the minimum variance sense) is proved. Both algorithms are finite dimensional and hence exactly implementable. The first one is a causal filter given by an optimal combination of a 1D Kalman predictor and of a 1D Kalman filter, the second one is a semicausal filter obtained by optimally combining a 1D Kalman predictor with a fixed-interval smoother. No partition of the image into strips is needed; image pixels are estimated one at a time so that the dimension of the state vector is N and the number of computation per pixel is $\mathcal{O}(N^3)$ with a reduction of a factor $\mathcal{O}(L^3/L_R)$ with respect to the strip filtering procedure. The overall saving of computation time derives not only from the filter equations but even from the computation needed to construct the image model. In fact, to compute the dynamical matrix of the image model, the method described in [7] requires, at each iteration, the inversion of a $(NL) \times (NL)$ space-varying three-band matrix; this inversion is avoided here.

Other 2D recursive filtering algorithms paralleling the 1D Kalman filter and not requiring strip processing have been proposed in [6, 9, 15, 25]. These filters are based on the quarter-plane system, first introduced in [9], and their non optimality was proved in [27] and [20]. As shown in [4] there is no optimal finite-dimensional causal filter for the quarter-plane system, while a finite dimensional approximation to the optimal half-plane filter has been presented in [1]. Moreover, the edge problem is not considered in the above papers.

The paper is structured in the following way. Some preliminaries are stated in Section 2; the discrete state-space realization of the image is obtained in Section 3; the recursive filtering algorithms are given in Section 4 and numerical results are reported in Section 5.

2. PRELIMINARIES

Let $x(r,s)$ be the value of the original monochromatic image at spatial coordinates (r,s), where the continuous variables r and s denote the vertical and horizontal position respectively. For simplicity, but without loss of generality, it is assumed that $(r,s) \in [0,1]^2$. Because of the smoothness assumption it is possible to define a state vector composed of the signal and its partial derivatives with respect to r and s

$$X(r,s) = \left[\frac{\partial^n x(r,s)}{\partial r^{n-\alpha}\partial s^{\alpha}},\ n = 0,1,\ldots,\bar{n};\ \alpha = 0,1,\ldots,n\right]^T. \tag{2.1}$$

If $\bar{n}$ is the maximum order of derivatives taken into account, the dimension of $X(r,s)$ is $N = (\bar{n}+1)(\bar{n}+2)/2$. If $r = r(u) = r_0 + \gamma u$, $s = s(u) = s_0 + \beta u$, the following equation can be written

$$\dot{X}(r(u),\ .s(u)) = \left(\frac{\partial}{\partial r}X(r(u),\ s(u))\right)\gamma + \left(\frac{\partial}{\partial s}X(r(u),\ s(u))\right)\beta, \tag{2.2}$$

the dot denoting the derivative with respect to u. Moreover, by direct computation

$$\frac{\partial}{\partial r}X(r(u),\ s(u)) = A_r X(r(u),\ s(u)) + BW_r(r(u),\ s(u)), \tag{2.3}$$

$$\frac{\partial}{\partial s}X(r(u),\ s(u)) = A_s X(r(u),\ s(u)) + BW_s(r(u),\ s(u)), \tag{2.4}$$

where:

- $$A_s = \begin{bmatrix} A'_s \\ 0 \end{bmatrix}, \quad A_r = \begin{bmatrix} A'_r \\ 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

  matrices A_r and A_s have dimensions $N \times N$, blocks A'_r and A'_s have dimensions $(N-(\bar{n}+1)) \times N$ and are composed of 0 and 1 elements in a suitable position, matrix B has dimensions $N \times (\bar{n}+1)$, the dimension of the null block and of the identity matrix are $(N-(\bar{n}+1)) \times (\bar{n}+1)$ and $(\bar{n}+1) \times (\bar{n}+1)$ respectively.

- vectors $W_r(r(u),\ s(u))$ and $W_s(r(u),\ s(u))$ have dimension $\bar{n}+1$ and are given by

$$W_r(r(u),\ s(u)) = \left[\frac{\partial^{\bar{n}+1}x(r,\ s)}{\partial r^{\bar{n}-\alpha+1}\partial s^{\alpha}},\ \alpha = 0, 1, \ldots, \bar{n}\right]^T,$$

$$W_s(r(u),\ s(u)) = \left[\frac{\partial^{\bar{n}+1}x(r,\ s)}{\partial r^{\bar{n}-\alpha}\partial s^{\alpha+1}},\ \alpha = 0, 1, \ldots, \bar{n}\right]^T.$$

Each row of blocks A'_r and A'_s contains only one element equal to 1 whose position follows from the order of the elements of $X(r,\ s)$ as stated by (2.1). For example consider A'_r and denote by $X_i(r,\ s, t)$, $i = 1\ldots, N$, a generic element of $X(r,\ s)$. By definition of state vector one has that if $1 \le i \le (N - (\bar{n}+1))$, then $\frac{\partial}{\partial r}X_i(r,\ s) \triangleq X_l(r,\ s)$ is still an element of $X(r,\ s)$ and by (2.3), $X_l(r,\ s)$ is given by the product of the ith row of A'_r with $X(r,\ s)$. This means that the 1 element on the ith row of A'_r lies on the lth column. Fully analogous considerations hold for A'_s. The following Lemma holds.

Lemma 1. Matrices A_r and A_s commute.

Proof. By the way A_r and A_s are defined, one has that both A_rA_s and A_sA_r are composed of 0 and 1 elements and that each non null row of A_rA_s and A_sA_r contains only one element equal to 1. If the ith row, $i = 1, \ldots, N$, of A_rA_s is not null, then the product of the ith row of A_rA_s with $X(r,\ s, t)$ gives the lth component $X_l(r,\ s, t)$, $3 < l < N$, of $X(r,\ s)$ with $X_l(r,\ s) = \frac{\partial}{\partial r}\left(\frac{\partial}{\partial s}X_i(r,\ s)\right)$. If the value of row index i is so high that $\frac{\partial}{\partial r}(\frac{\partial}{\partial s}X_i(r,\ s)) \notin X(r,\ s)$, then the ith row of A_rA_s is null.

Analogously, if the ith row, $i = 1, \ldots, N$, of $A_s A_r$ is not null, then the product of this row with $X(r,\, s, t)$ gives $X_{l_1}(r,\, s, t)$, $3 < l_1 < N$, with $X_{l_1}(r,\, s) = \frac{\partial}{\partial s}(\frac{\partial}{\partial r} X_i(r,\, s))$. If the value of row index i is so high that $\frac{\partial}{\partial s}(\frac{\partial}{\partial r} X_i(r,\, s)) \notin X(r,\, s)$, then the ith row of $A_s A_r$ is null. As $\frac{\partial}{\partial r}\left(\frac{\partial}{\partial s} X_i(r,\, s)\right) = \frac{\partial}{\partial s}\left(\frac{\partial}{\partial r} X_i(r,\, s)\right)$, it follows $l = l_1$. This means that for each i, either the ith rows of $A_r A_s$ and $A_s A_r$ are both null, or both contain the 1 element on the same column; hence $A_r A_s = A_s A_r$. □

Using (2.3) and (2.4), equation (2.2) can be written in the following form

$$\dot{X}(r(u),\, s(u)) = (\gamma A + \beta A')\, X(r(u),\, s(u)) + B[\gamma W_r(r(u),\, s(u)) + \beta W_s(r(u),\, s(u))]. \tag{2.5}$$

Lemma 1 implies $e^{A_r} e^{A_s} = e^{(A_r + A_s)}$, so that formal integration of (2.5) with respect to u between u_0 and u_1 results in the following relation between the state vector evaluated at two generic points $(r_0 + \zeta u_0, s_0 + \theta u_0)$ and $(r_0 + \zeta u_1, s_0 + \theta u_1)$

$$X(r_0 + \zeta u_1, s_0 + \theta u_1) = e^{(\zeta A_r + \theta A_s)(u_1 - u_0)} X(r_0 + \zeta u_0, s_0 + \theta u_0) \tag{2.6}$$
$$+ \int_{u_0}^{u_1} e^{(\zeta A_r + \theta A_s)(u_1 - \tau)} B[\zeta W_r(r_0 + \zeta\tau, s_0 + \theta\tau) + \theta W_s(r_0 + \zeta\tau, s_0 + \beta\tau)]\, \mathrm{d}\tau.$$

By the stochastic assumption, the integral term in (2.7) is intended as a stochastic Wiener integral.

3. DISCRETE STATE SPACE REALIZATION

Denote by (i, j) the pixel of the sampled image with vertical coordinate $i\Delta_r$ and horizontal coordinate $j\Delta_s$, where Δ_r and Δ_s denote the distance between two adjacent pixels on a same column or on a same row respectively. If the image is sampled with an equal number $\bar{m}$ of pixels on each row and on each column, the normalized values of Δ_r and Δ_s are both equal to $1/(\bar{m} - 1)$. The true value of the sampled image at pixel (i, j) is denoted by $x_{i,j}$ and the state vector evaluated at the same pixel is denoted by $X_{i,j}$.

Provided that the pixels (i, j), $(i - 1, j)$ and $(i, j - 1)$ are not separated by edges, the relations between $X_{i,j}$, $X_{i-1,j}$ and $X_{i,j-1}$ can be obtained from equation (2.7) putting $u_0 = 0$, $u_1 = 1$ and with a suitable choice of γ and β.

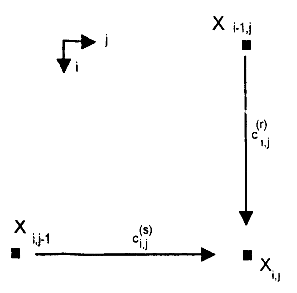

Fig. 1. Spatial structure of the dependence scheme.

The following equations are obtained

$$X_{i,j} = H_s X_{i,j-1} + W_{i,j}^{(s)}, \quad (3.1)$$

$$X_{i,j} = H_r X_{i-1,j} + W_{i,j}^{(r)}, \quad (3.2)$$

where: $H_s = e^{(A'\Delta_s)}$, $H_r = e^{(A\Delta_r)}$, and

$$W_{i,j}^{(s)} = \int_0^1 e^{A'\Delta_s(1-\tau)}\Delta_s B W_s(i\Delta_r, (j-1)\,\Delta_s + \Delta_s\tau)\,\mathrm{d}\tau, \quad (3.3)$$

$$W_{i,j}^{(r)} = \int_0^1 e^{A\Delta_r(1-\tau)}\Delta_r B W_r((i-1)\,\Delta_r + \tau\Delta_r, j\Delta_s)\,\mathrm{d}\tau. \quad (3.4)$$

From the stochastic assumption and using (3.3) and (3.4), it can be shown that $W_{i,j}^{(s)}$ and $W_{i,j}^{(r)}$ are zero mean white gaussian random fields with the following properties

$$\mathcal{E}[W_{i,j}^{(s)}W_{l,m}^{(s)^T}] = \delta_{i,l}\delta_{j,m}Q_s, \quad \mathcal{E}[W_{i,j}^{(r)}W_{l,m}^{(r)^T}] = \delta_{i,l}\delta_{j,m}Q_r, \quad \mathcal{E}[W_{i,j}^{(s)}W_{i,j}^{(r)^T}] = 0, \quad (3.5)$$

with

$$Q_s = \int_0^1 e^{(A'\Delta_s(1-\tau))}B\Psi_s B^T e^{(A'^T\Delta_s(1-\tau))}\,\mathrm{d}\tau,$$

$$Q_r = \int_0^1 e^{(A\Delta_r(1-\tau))}B\Psi_r B^T e^{(A^T\Delta_r(1-\tau))}\,\mathrm{d}\tau,$$

where Ψ_r and Ψ_s are diagonal matrices such that

$$\mathcal{E}[W_s(r,s)W_s^T(\bar{r}\bar{s})] = \Psi_s\delta(\|(r,s)-(\bar{r},\bar{s})\|),$$

$$\mathcal{E}[W_r(r,s)W_r^T(\bar{r}\bar{s})] = \Psi_r\delta(\|(r,s)-(\bar{r},\bar{s})\|),$$

where $\mathcal{E}[\cdot]$ denotes the expected value, $\delta(\cdot)$ is the Dirac delta $\Psi_s(t)$, $\Psi_r(t)$ and $\Psi_t(t)$ are diagonal matrices that can be estimated as functions of the image spectrum [8]. Hence $\{W_{i,j}^{(s)}\}$ and $\{W_{i,j}^{(s)}\}$ are mutually uncorrelated 2D white noise sequences.

If an edge occurs between pixels (i,j) and $(i-1,j)$ and/or between pixels (i,j) and $(i,j-1)$, equation (2.5) can not be integrated along the corresponding horizontal and/or vertical direction because, as a consequence of the inhomogeneity assumption, no relation exists between $X_{i,j}$ and $X_{i-1,j}$ and/or between $X_{i,j}$ and$X_{i,j-1}$. Equations (3.1) and (3.2) are then modified as

$$X_{i,j} = H_s(c_{i,j}^{(s)}X_{i,j-1} + (1-c_{i,j}^{(s)})\,X_{i,j-1}^{0(s)}) + c_{i,j}^{(s)}W_{i,j}^{(s)} + (1-c_{i,j}^{(s)})W_{i,j}^{0(s)}, \quad (3.6)$$

$$X_{i,j} = H_r(c_{i,j}^{(r)}X_{i-1,j} + (1-c_{i,j}^{(r)})\,X_{i-1,j}^{0(r)}) + c_{i,j}^{(r)}W_{i,j}^{(r)} + (1-c_{i,j}^{(r)})W_{i,j}^{0(r)}, \quad (3.7)$$

where $c_{i,j}^{(s)}$ and $c_{i,j}^{(r)}$ are coefficients that may be zero or one. According to the inhomogeneity assumption, $c_{i,j}^{(s)}$ is one if pixels (i,j) and $(i,j-1)$ are not separated by an edge and zero otherwise, analogously $c_{i,j}^{(r)}$ is one if pixels (i,j) and $(i-1,j)$ are not separated by an edge and zero otherwise; $X_{i,j-1}^{0(s)}$ and $W_{i,j}^{0(s)}$ are the initial state and the initial value of $\{W_{i,j}^{(s)}\}$ respectively, corresponding to each edge crossed during the horizontal scanning along a line and $X_{i-1,j}^{0(r)}$ and $W_{i,j}^{0(r)}$ are the initial state and the initial value of $\{W_{i,j}^{(r)}\}$ respectively, corresponding to each edge crossed during the vertical scanning along a column.

Equations (3.6) and (3.7) work like equations (3.1) and (3.2) respectively, as long as $c_{i,j}^{(s)}$ and $c_{i,j}^{(r)}$ are one, namely inside each smooth subregion. If for some pixel (i,j) one has $c_{i,j}^{(s)} = 0$ (and/or $c_{i,j}^{(r)} = 0$) this means that an edge occurs between pixels (i,j) and $(i,j-1)$ (and/or between pixels (i,j) and $(i-1,j)$), namely that a transition occurs between two contiguous smooth subregions; hence, in the light of the inhomogeneity assumption, no relation exists between $X_{i,j}$ and $X_{i,j-1}$ (and/or between $X_{i,j}$ and $X_{i-1,j}$), then a state resetting is performed in equations (3.6), (and/or (3.7)) by expressing $X_{i,j}$ as a function of the initial conditions $X_{i,j-1}^{0(s)}$, $W_{i,j}^{0(s)}$ (and/or $X_{i-1,j}^{0(r)}$, $W_{i,j}^{0(r)}$) relative to the stochastic process describing the image inside the new smooth subregion.

Taking into account that the only observed component of the state vector is the image signal, the following measure equation can be associated with equations (3.6) and (3.7)

$$y_{i,j} = CX_{i,j} + v_{i,j}, \tag{3.8}$$

where C is the $1 \times N$ row vector $[1, 0, \ldots, 0]$ and $v_{i,j}$ is a discrete white gaussian noise $\sim \mathcal{N}(0, \sigma_v^2)$ uncorrelated both with $\{W_{i,j}^{(s)}\}$ and $\{W_{i,j}^{(r)}\}$. System composed of equation (3.6), (3.7) and (3.8) constitutes the 2D discrete space-variant state-space representation of the image.

4. THE RECURSIVE FILTERING ALGORITHMS

4.1. The causal filter

Given pixel (i,j) on the ith row, let $Y_{i,j}$ and $Y_{i,j}^*$ be the sets of all the observations $y_{i,l}$ relative to pixels lying on the ith row which are not separated from pixel (i,j) by an edge and with the additional requirement $l \le j$ as for $Y_{i,j}$. Denoting by $\mathcal{E}(./.)$ the conditional expectation, let $\mathcal{E}(X_{i,j}/Y_{i,j-1}) = \hat{X}_{i,j}^{(s)-}$ and $\mathcal{E}(X_{i,j}/Y_{i,j}) = \hat{X}_{i,j}^{(s)+}$, $1 \le j \le \bar{m}$, be the predicted and filtered estimates respectively of $X_{i,j}$, $1 \le j \le \bar{m}$, obtained by applying, on the ith row, the 1D Kalman filter to the image 1D submodel given by equations (3.6) and (3.8); moreover denote by $P_{i,j}^{(s)-}$ and $P_{i,j}^{(s)+}$ the error covariance matrices of $\hat{X}_{i,j}^{(s)-}$ and $\hat{X}_{i,j}^{(s)+}$ respectively. According to (3.6), (3.8) one has

$$\hat{X}_{i,j}^{(s)+} = \hat{X}_{i,j}^{(s)-} + K_{i,j}^{(s)}(y_{i,j} - C\hat{X}_{i,j}^{(s)-}), \tag{4.1}$$

$$\hat{X}_{i,j}^{(s)-} = H_s(c_{i,j}^{(s)}\hat{X}_{i,j-1}^{(s)+} + (1 - c_{i,j}^{(s)})\hat{X}_{i,j-1}^{0(s)}), \tag{4.2}$$

$$K_{i,j}^{(s)} = P_{i,j}^{(s)-}C^T(CP_{i,j}^{(s)-}C^T + \sigma_v^2)^{-1}, \tag{4.3}$$

$$P_{i,j}^{(s)-} = H_s(P_{i,j-1}^{(s)+} + (1 - c_{i,j}^{(s)})\, P_{i,j-1}^{0(s)})\, H_s^T + c_{i,j}^{(s)} Q_s + (1 - c_{i,j}^{(s)}) Q_{i,j}^{0(s)}, \quad (4.4)$$

$$P_{i,j}^{(s)+} = (I - K_{i,j}^{(s)} C)\, P_{i,j}^{(s)-}, \quad (4.5)$$

where $\hat{X}_{i,j-1}^{0(s)}$ is the estimate of $X_{i,j-1}^{0(s)}$, $P_{i,j}^{0(s)}$ is the error covariance matrix of $\hat{X}_{i,j-1}^{0(s)}$, and $Q_{i,j}^{0(s)}$ is the covariance matrix of $W_{i,j}^{0(s)}$.

Assuming that $\hat{X}_{i-1,j}$ and $P_{i-1,j}$, $1 \le j \le \bar{m}$, have already been computed, their optimal previsions to the ith row can be obtained by the prediction equations of the 1D Kalman filter deriving from (3.7), obtaining

$$\hat{X}_{i,j}^{(r)-} = H_r(c_{i,j}^{(r)} \hat{X}_{i-1,j} + (1 - c_{i,j}^{(r)}) \hat{X}_{i-1,j}^{0(r)}), \quad (4.6)$$

$$P_{i,j}^{(r)-} = H_r(c_{i,j}^{(r)} P_{i-1,j} + (1 - c_{i,j}^{(r)})\, P_{i-1,j}^{0(r)})\, H_r^T + c_{i,j}^{(r)} Q_r + (1 - c_{i,j}^{(r)})\, Q_{i,j}^{0(r)}, \quad (4.7)$$

where $\hat{X}_{i-1,j}^{0(r)}$ is the estimate of $X_{i-1,j}^{0(r)}$, $P_{i-1,j}^{0(r)}$ is the error covariance matrix of $\hat{X}_{i-1,j}^{0(r)}$, and $Q_{i,j}^{0(r)}$ is the covariance matrix of $W_{i,j}^{0(r)}$. The final 2D estimates $\hat{X}_{i,j}$ and $P_{i,j}$ can be obtained through an optimal combination of $\hat{X}_{i,j}^{(s)+}$, $P_{i,j}^{(s)+}$, $\hat{X}_{i,j}^{(r)-}$ and $P_{i,j}^{(r)-}$. To this purpose let $(i, j_{i,\ell}), j_{i,\ell} = j_{i,1}, \ldots, j_{i,\bar{\ell}}$, with $j_{i,1} = 0$, $i = 1, \ldots, \bar{m}$, be the coordinates of discontinuity points on the ith row, namely the coordinates of those pixels for which $\hat{X}_{i,j}^{0(s)}$ is defined and denote by $\hat{E}_{i,j_{i,\ell}}^{0(s)} = X_{i,j_{i,\ell}}^{0(s)} - \hat{X}_{i,j_{i,\ell}}^{0(s)}$ the corresponding initial estimation error. The following Assumption is made.

Assumption 4.1. Each initial estimation error $\hat{E}_{i,j_{i,\ell}}^{0(s)}$, $j_{i,\ell} = j_{i,1}, \ldots, j_{i,\bar{\ell}}$, is uncorrelated with $\hat{E}_{i-1,j}$ and with $W_{i,j}^{(r)}$, $j = 1, \ldots, \bar{m}$.

This assumption is not restrictive, it simply means that at each edge point the available estimate $\hat{X}_{i,j_\ell}^{0(s)}$ of the initial value $X_{i,j_\ell}^{0(s)}$ is independent of the way the final estimates on the previous row have been computed and of the stochastic terms of state equation (3.7). Assumption 4.1 allows us to show that each estimation error $\hat{E}_{i,j}^{(s)+} = X_{i,j} - \hat{X}_{i,j}^{(s)+}$ is uncorrelated with $\hat{E}_{i,j}^{(r)-} =: X_{i,j} - \hat{X}_{i,j}^{(r)-}$, $j = 1, \ldots, \bar{m}$. This can be proved in the following way. For simplicity, but without any loss of generality, reference is made to a single smooth subregion starting at the first pixel of the generic ith row, hence in the following it is assumed that $c_{i,j}^{(s)} = c_{i,j}^{(r)} = 1$.

From (3.6), (3.7), (3.8), (4.1) and (4.6) it follows

$$\hat{E}_{i,j}^{(r)-} = H_r \hat{E}_{i-1,j} + W_{i,j}^{(r)}, \quad (4.8)$$

$$\hat{E}_{i,j}^{(s)+} = L_{i,j}^{(s)} \hat{E}_{i,j-1}^{(s)+} + M_{i,j}^{(s)} W_{i,j}^{(s)} - K_{i,j}^{(s)} v_{i,j}, \quad (4.9)$$

where

$$H_s - K_{i,j}^{(s)} C H_s \triangleq L_{i,j}^{(s)}, \quad I - K_{i,j}^{(s)} C \triangleq M_{i,j}^{(s)}.$$

Equation (4.9) gives

$$\hat{E}_{i,j}^{(s)+} = \Phi_{i(j,0)}^{(s)} \hat{E}_{i,0}^{0(s)} + \sum_{k=1}^{j} \left(\Psi_{i(j,k)}^{(s)} W_{i,k}^{(s)} - \Theta_{i(j,k)}^{(s)} v_{i,k} \right), \quad (4.10)$$

where:

$$\Phi_{i_{(j,0)}} = \begin{cases} I & \text{if } j=0, \\ L_{i,j}^{(s)} L_{i,j-1}^{(s)} \cdots L_{i,1}^{(s)} & \text{if } j>0, \end{cases}$$

$$\Phi_{i_{(j,k)}}^{(s)} M_{i,k}^{(s)} \triangleq \Psi_{i_{(j,k)}}^{(s)}, \qquad \Phi_{i_{(j,k)}}^{(s)} K_{i,k}^{(s)} \triangleq \Theta_{i_{(j,k)}}^{(s)}.$$

By Assumption 4.1 one has $\mathcal{E}[\hat{E}_{i,0}^{0(s)}(\hat{E}_{i,j}^{(r)-})^T] = 0$, moreover $W_{i,k}^{(s)}$ is in the "future" with respect to $\hat{E}_{i-1,j}$ so that $\hat{E}_{i-1,j}$ is a linear function of quantities independent of $W_{i,k}^{(s)}$, hence, taking also into account (3.5) one has $\mathcal{E}[W_{i,k}^{(s)}(\hat{E}_{i,j}^{(r)-})^T] = 0$. Analogously $v_{i,k}$ is in the "future" with respect to $\hat{E}_{i-1,j}$ and is uncorrelated with $W_{i,j}^{(r)}$ by assumption, so that $\mathcal{E}[v_{i,k}(\hat{E}_{i,j}^{(r)-})^T] = 0$. It follows $\mathcal{E}[\hat{E}_{i,j}^{(s)+}(\hat{E}_{i,j}^{(r)-})^T] = 0$.
This implies that the estimates $\hat{X}_{i,j}^{(s)+}$ and $\hat{X}_{i,j}^{(r)-}$ and the relative error covariance matrices $P_{i,j}^{(s)+}$ and $P_{i,j}^{(r)-}$ respectively, can be optimally combined to obtain $\hat{X}_{i,j}$ and $P_{i,j}$ according to

$$\hat{X}_{i,j} = P_{i,j}[(P_{i,j}^{(r)-})^{-1}\hat{X}_{i,j}^{(r)-} + (P_{i,j}^{(s)+})^{-1}\hat{X}_{i,j}^{(s)+}], \tag{4.11}$$

$$P_{i,j} = [(P_{i,j}^{(r)-})^{-1} + (P_{i,j}^{(s)+})^{-1}]^{-1}. \tag{4.12}$$

The proposed causal filtering algorithm is given by equations (4.1) – (4.7), (4.11), (4.12) applied to each row; it is referred to as the Causal Space–Variant Filter (CSVF).

4.2. The semicausal filter

Denote by $\hat{X}_{i,j}^{(s)*}$ the smoothed estimates $\mathcal{E}(X_{i,j}/Y_{i,j}^*)$ on the generic ith row and by $P_{i,j}^{(s)*}$ the relative error covariance matrices corresponding to the image 1D submodel (3.6), (3.8). The semicausal filter is obtained by replacing the filtered estimates $\hat{X}_{i,j}^{(s)+}$ with $\hat{X}_{i,j}^{(s)*}$ and $P_{i,j}^{(s)+}$ with $P_{i,j}^{(s)*}$. The smoothed estimates $\hat{X}_{i,j}^{(s)*}$ and the relative error covariance matrices $P_{i,j}^{(s)*}$, $1 \le j \le \bar{m}$, can be obtained through the 1D fixed-interval smoother equations associated with equations (4.1) – (4.5),

$$\hat{X}_{i,j}^{(s)*} = \hat{X}_{i,j}^{(s)+} + A_{i,j}(\hat{X}_{i,j+1}^{(s)*} - \hat{X}_{i,j+1}^{(s)-}), \tag{4.13}$$

$$P_{i,j}^{(s)*} = P_{i,j}^{(s)+} + A_{i,j}(P_{i,j+1}^{(s)*} - P_{i,j+1}^{(s)-})A_{i,j}^T, \tag{4.14}$$

$$A_{i,j} = c_{i,j+1}^{(s)} P_{i,j}^{(s)+} H_s^T (P_{i,j+1}^{(s)-})^{-1}. \tag{4.15}$$

Denote by $\hat{X}_{i,j}^*$ the final 2D semicausal estimates obtained by also exploiting equation (3.7) and by $P_{i,j}^*$ the relative error covariance matrices and, as for the CSVF, assume that $\hat{X}_{i-1,j}^*$ and $P_{i-1,j}^*$, $1 \le j \le \bar{m}$ have already been computed equations (4.6) and (4.7) are then replaced by

$$\hat{X}_{i,j}^{(r^*)-} = H_r(c_{i,j}^{(r)}\hat{X}_{i-1,j}^* + (1-c_{i,j}^{(r)})\hat{X}_{i-1,j}^{0(r)}), \tag{4.16}$$

$$P_{i,j}^{(r^*)-} = H_r(c_{i,j}^{(r)}P_{i-1,j}^* + (1-c_{i,j}^{(r)})P_{i-1,j}^{0(r)})\, H_r^T + c_{i,j}^{(r)}Q_r + (1-c_{i,j}^{(r)})Q_{i,j}^{0(r)}. \tag{4.17}$$

Analogously to the CSVF, the final semicausal estimates $\hat{X}^*_{i,j}$ and $P^*_{i,j}$ can be obtained through an optimal combination of $\hat{X}^{(r*)-}_{i,j}$ and $P^{(r*)-}_{i,j}$ with $\hat{X}^{(s)*}_{i,j}$ and $P^{(s)*}_{i,j}$. To this purpose, defining the final semicausal estimation error as $\hat{E}^*_{i,j} = X_{i,j} - \hat{X}^*_{i,j}$, Assumption 4.1 is reformulated as:

Assumption 4.2. Each initial estimation error $\hat{E}^{0(s)}_{i,j_{i,\ell}}$, $j_{i,\ell} = j_{i,1}, \dots, j_{i,\bar{\ell}}$, is uncorrelated with $\hat{E}^*_{i-1,j}$ and with $W^{(r)}_{i,j}$ (or $W^{0(r)}_{i,j}$), $j = 1, \dots, \bar{m}$.

Assumption 4.2 allows us to show that also $\hat{E}^{(s)*}_{i,j} = X_{i,j} - \hat{X}^{(s)*}_{i,j}$ and $\hat{E}^{(r*)-}_{i,j} = X_{i,j} - \hat{X}^{(r*)-}_{i,j}$, $j = 1, \dots, \bar{m}$, are uncorrelated. For simplicity, but without any loss of generality, reference is again made to a single smooth subregion starting at the first pixel of the generic ith row.
From (3.7) and (4.16) it readily follows

$$\hat{E}^{(r^*)-}_{i,j} = H_r \hat{E}^*_{i-1,j} + W^{(r)}_{i,j}. \tag{4.18}$$

To show that $\hat{E}^{(r*)-}_{i,j}$ is uncorrelated with $\hat{E}^{(s)*}_{i,j}$, consider equation (4.13) and replace $\hat{X}^{(s)*}_{i,j+1}$ with its expression deriving from (4.13) written for $j+1$; in the equation so obtained repeat the same procedure for $\hat{X}^{(s)*}_{i,j+2}$ and so on till $j = \bar{m}$. Taking into account that (4.13) is started with $\hat{X}^{(s)*}_{i,\bar{m}} = \hat{X}^{(s)+}_{i,\bar{m}}$, one has

$$\hat{X}^{(s)*}_{i,j} = \hat{X}^{(s)+}_{i,j} + \sum_{\ell=0}^{\bar{m}-j-1} (A_{i,j} A_{i,j+1} \cdots A_{i,j+\ell}) \cdot [\hat{X}^{(s)+}_{i,j+\ell+1} - \hat{X}^{(s)-}_{i,j+\ell+1}]. \tag{4.19}$$

Using (4.1), (4.2) and (4.10), it is found that

$$\begin{aligned} \hat{X}^{(s)+}_{i,j+\ell+1} - \hat{X}^{(s)-}_{i,j+\ell+1} &= \Gamma^{(s)}_{i_{(j+\ell+1,0)}} \hat{E}^{0(s)}_{i,0} + \sum_{k=1}^{j+\ell} (\Gamma^{(s)}_{i_{(j+\ell+1,k)}} W^{(s)}_{i,k} - \Pi^{(s)}_{i_{(j+\ell+1,k)}} v_{i,k}) \\ &\quad + K^{(s)}_{i,j+\ell+1} (C W^{(s)}_{i,j+\ell+1} + v_{i,j+\ell+1}), \end{aligned} \tag{4.20}$$

where:

$$\begin{aligned} K^{(s)}_{i,j+\ell+1} C H_s \Phi^{(s)}_{i_{(j+\ell,0)}} &\triangleq \Gamma^{(s)}_{i_{(j+\ell+1,0)}}, \\ K^{(s)}_{i,j+\ell+1} C H_s \Psi^{(s)}_{i_{(j+\ell,r)}} &\triangleq \Upsilon^{(s)}_{i_{(j+\ell+1,r)}}, \\ K^{(s)}_{i,j+\ell+1} C H_s \Theta^{(s)}_{i_{(j+\ell,r)}} &\triangleq \Pi^{(s)}_{i_{(j+\ell+1,r)}}. \end{aligned}$$

From (4.19) and (4.20), the following expression for $\hat{E}^{(s)*}_{i,j}$ is found

$$\hat{E}^{(s)*}_{i,j} = \hat{E}^{(s)+}_{i,j} - \sum_{\ell=0}^{\bar{m}-j-1} \mathcal{A}_{i_{(j,\ell)}} \Gamma^{(s)}_{i_{(j+\ell+1,0)}} \hat{E}^{0(s)}_{i,0}$$

$$
\begin{aligned}
&- \sum_{\ell=0}^{\bar{m}-j-1} \sum_{k=1}^{j+\ell} \mathcal{A}_{i_{(j,\ell)}} (\Upsilon^{(s)}_{i_{(j+\ell+1,k)}} W^{(s)}_{i,k} - \Pi^{(s)}_{i_{(j+\ell+1,r)}} v_{i,r}) \\
&- \sum_{\ell=0}^{\bar{m}-j-1} \mathcal{A}_{i_{(j,\ell)}} [K^{(s)}_{i,j+\ell+1} (C W^{(s)}_{i,j+\ell+1} + v_{i,j+\ell+1})],
\end{aligned} \tag{4.21}
$$

where $A_{i,j} A_{i,j+1} \cdots A_{i,j+\ell} \stackrel{\Delta}{=} \mathcal{A}_{i_{(j,\ell)}}$.
Replacing $\hat{E}^{(s)+}_{i,j}$ into (4.21) with its expression given by (4.10), using Assumption 4.2 and arguing as for the CSVF, it is straightforward to see that $\mathcal{E}[\hat{E}^{(s)+}_{i,j} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$. Assumption 4.2 also implies $\mathcal{E}[\hat{E}^{0(s)}_{i,j} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$. Moreover $W^{(s)}_{i,k}$ and $W^{(s)}_{i,j+\ell+1}$ are in the "future" with respect to $\hat{E}^*_{i-1,j}$, so that $\hat{E}^*_{i-1,j}$ is a linear function of quantities independent of $W^{(s)}_{i,k}$ and $W^{(s)}_{i,j+\ell+1}$, hence, taking also into account (3.5) one has $\mathcal{E}[W^{(s)}_{i,k} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$, $\mathcal{E}[W^{(s)}_{i,j+\ell+1} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$. Analogously $v_{i,k}$ is in the "future" with respect to $\hat{E}^*_{i-1,j}$ and is uncorrelated with $W^{(r)}_{i,j}$ and $W^{(s)}_{i,j+\ell+1}$ by assumption, so that $\mathcal{E}[v_{i,k} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$. It follows $\mathcal{E}[\hat{E}^{(s)*}_{i,j} (\hat{E}^{(r^*)-}_{i,j})^T] = 0$.

Hence, the final semicausal estimates $\hat{X}^*_{i,j}$ and $P^*_{i,j}$ can be obtained with formulas analogous to (4.11), (4.12):

$$
\hat{X}^*_{i,j} = P^*_{i,j} [(P^{(r^*)-}_{i,j})^{-1} \hat{X}^{(r^*)-}_{i,j} + (P^{(s)*}_{i,j})^{-1} \hat{X}^{(s)*}_{i,j}], \tag{4.22}
$$

$$
P^*_{i,j} = [(P^{(r^*)-}_{i,j})^{-1} + (P^{(s)*}_{i,j})^{-1}]^{-1}. \tag{4.23}
$$

The proposed semicausal filtering algorithm is given by equations (4.1) – (4.5), (4.13) – (4.17), (4.22), (4.23) applied to each row; it is referred to as the SemiCausal Space–Variant Filter (SCSVF).

Foregoing calculations show that the state estimates provided by the CSVF and SCSVF are really minimum variance estimates, namely $\hat{X}_{i,j} = \mathcal{E}(X_{i,j}/\boldsymbol{Y}_{i,j})$ and $\hat{X}^*_{i,j} = \mathcal{E}(X_{i,j}/\boldsymbol{Y}^*_{i,j})$, where $\boldsymbol{Y}_{i,j}$ and $\boldsymbol{Y}^*_{i,j}$ are suitably defined linear observation spaces that can be determined from the filter equations as follows.

Consider at first $\boldsymbol{Y}_{i,j}$. The CSVF is such that, at each pixel (i,j), the estimate $\hat{X}_{i,j}$ may get information on previous rows only from $\hat{X}_{i-1,j}$ through equations (4.6), (4.7). In turn, $\hat{X}_{i-1,j}$ may get information on previous rows only through $\hat{X}_{i-2,j}$ and so on backward, till an image edge is met. Further, on any row, each estimate $\hat{X}^{(s)+}_{i,j}$ is computed exploiting the information carried by all the pixels (i,k) with $k \leq j$ which are not separated from pixel (i,j) by an edge. Therefore, it follows that for each pixel (i,j), the associated $\boldsymbol{Y}_{i,j}$ is generated by the observations relative to the ensemble of segments with end-pixels $[(i-l, j-k_l), (i-l, j)]$, $l = 0, 1, \ldots, \bar{l}_{i,j}$, for some k_l and $\bar{l}_{i,j} \geq 0$, where each segment is composed of pixels not separated from pixel $(i-l, j)$ by an image edge. The value $\bar{l}_{i,j}$ is the largest value such that no edge is met scanning the jth column from pixel (i,j) to pixel $(i-\bar{l}_{i,j}, j)$. An example of pixels corresponding to observations generating $\boldsymbol{Y}_{i,j}$ is reported in Figure 2a. In the case of images coinciding with a unique smooth subregion, $\boldsymbol{Y}_{i,j}$ is the linear space spanned by all the observations $y_{l,m}$ with $1 \leq l \leq i, 1 \leq m \leq j$. Analogously, for each pixel (i,j), the associated $Y^*_{i,j}$ is generated by the observations relative to the ensemble of segments with end-pixels $[(i-l, j-k_l), (i-l, j+m_l)], l = 0, 1, \ldots, \bar{l}_{i,j}$, for some k_l, m_l and $\bar{l}_{i,j} \geq 0$, where each segment is composed of pixels not separated from pixel $(i-l, j)$ by an image edge. The value $\bar{l}_{i,j}$ is the largest value such that no edge is met scanning the jth column from pixel (i,j) to pixel $(i-\bar{l}_{i,j}, j)$. Figure 2b shows the ensemble of pixels corresponding to observations generating $Y^*_{i,j}$, with reference to the same situation of Figure 2a. In the case of images coinciding with a unique smooth subregion, $Y^*_{i,j}$ is the linear space spanned by all the observations $y_{l,m}$ with $1 \leq l \leq i, 1 \leq m \leq \bar{m}$.

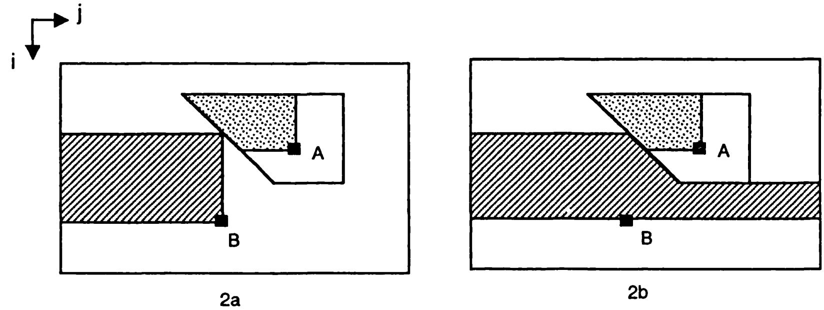

Fig. 2. Ensembles of pixels corresponding to observations concurring to determine the state estimate at pixels A (dotted area) and B (dashed area) for the CSVF (2a) and for the SCSVF (2b).

5. NUMERICAL RESULTS

The 256 × 256 pixels eight-bit image shown in Figure 3 has been used to test the performance of the proposed CSVF and SCSVF. The original has been corrupted by zero-mean white gaussian noise with a variance such that the SNR (signal variance/ noise variance) resulted to be 8. The noisy picture is reported in Figure 4. The procedure for detecting image edges, thus determining the coefficients $c^{(r)}_{i,j}$, and $c^{(s)}_{i,j}$ associated with each pixel, has been implemented using the gradient method, as described in [7]. In correspondence of each pixel for which $c^{(r)}_{i,j}$ and/or $c^{(s)}_{i,j} = 0$, the following $\hat{X}^{0(r)}_{i-1,j}$, $\hat{X}^{0(s)}_{i,j-1}$, $P^{0(r)}_{i-1,j}$, $P^{0(s)}_{i,j-1}$, $Q^{0(r)}_{i,j}$ and $Q^{0(s)}_{i,j}$ have been assumed:

$$\hat{X}^{0(r)}_{i-1,j} = \hat{X}^{0(s)}_{i,j-1} = [y_{i,j} \quad \underbrace{0 \cdots 0}_{N-1 \text{ elements}} \quad]^T,$$

$$P^{0(r)}_{i-1,j} = P^{0(s)}_{i,j-1} = \begin{bmatrix} P_0 & 0 & 0 & \cdots & 0 \\ 0 & P_1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & \cdot & \cdot & \cdots & P_{\bar{n}} \end{bmatrix},$$

$$P_k = \begin{bmatrix} \bar{m}^{2k} & 0 & 0 & \cdots & 0 \\ 0 & \bar{m}^{2k} & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ 0 & 0 & \cdots & 0 & \bar{m}^{2k} \end{bmatrix} \sigma_v^2, \quad k = 0, \ldots, \bar{n}, \quad Q_{i,j}^{0(r)} = Q_r, \; Q_{i,j}^{0(s)} = Q_s,$$

where each block P_k, $k = 0, \ldots, \bar{n}$, of matrices $P_{i-1,j}^{0(r)}$ and $P_{i,j-1}^{0(s)}$ has dimensions $(k+1) \times (k+1)$ and represents the initial error covariance matrix of the $(k+1)$ partial derivatives of order k.

The image has been processed with a model order corresponding to the choice $\bar{n} = 1$. The numerical simulations have been implemented on an Alpha AXP 3500 under Open VMS 1.5 Operating System. The Signal-to-Noise Ratio Improvement $\mathrm{SNRI_{db}}$ introduced by the filters has been estimated as

$$\mathrm{SNRI_{db}} = 10 \log_{10} \frac{\sum_i \sum_j (y_{i,j} - x_{i,j})^2}{\sum_i \sum_j (\hat{x}_{i,j} - x_{i,j})^2}$$

where $\hat{x}_{i,j}$ is the first component of the estimated state vector ($\hat{X}_{i,j}$ or $\hat{X}_{i,j}^*$), namely the estimate of the true image signal $x_{i,j}$.

Filtered images obtained with the CSVF and with the SCSVF are reported in Figures 5 and 6 respectively, the values of the SNRI and CPU times are reported in Table 1. The first term in the column of CPU times in Tables 1–4 refers to the filtering procedure, the second term is the time elapsed for the preliminary processing of image edges.

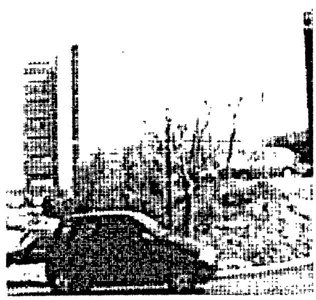

Fig. 3. Original image.

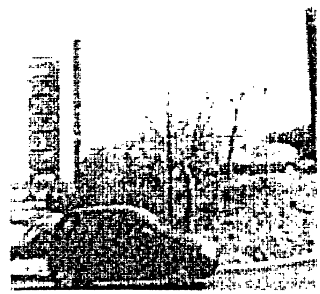

Fig. 4. Noisy image.

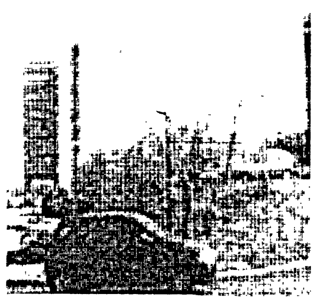

Fig. 5. CSVF restored image.

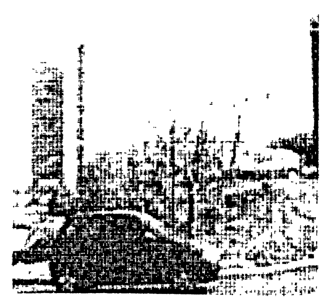

Fig. 6. SCSVF restored image.

Table 1. SNRI and CPU Times.

CSVF		SCSV	
$SNRI_{db}$	CPU Time	$SNRI_{db}$	CPU Time
4.64	45.20 s + 2.16 s	5.30	11 m 12.74 s + 2.16 s

Table 2. $SNRI_{db}$ and CPU Times.

CSVF		SCSV	
$SNRI_{db}$	CPU Time	$SNRI_{db}$	CPU Time
4.64	45.20s + 2.16s	5.30	11 m 12.74s + 2.16s

Filtered images appear to be very satisfactory. Figures 5 and 6 reveal an effective reduction of the observation noise and a good preservation of image edges. Also the computation time results to be modest for both the algorithms.

6. CONCLUSIONS

In this paper it has been shown that, starting from general basic modelling assumptions, two finite-dimensional, minimum variance 2D filters can be defined. The key points implying optimality are the following. First, the 2D state space representation of the image is given by two dynamical equations with independent white noise inputs. This allows us to define two 1D optimal estimates of the same pixel characterized by independent estimation errors. Hence the 1D estimates can be optimally combined according to (4.11), (4.12), or (4.22), (4.23), to yield the final 2D minimum variance estimate. The space variant behaviour of the proposed filters is obtained by including a structural information about edge locations in the image model. In this way, the filter transitions in correspondence of edge locations are not the result of heuristic procedures, but are justified on a theoretical basis in that they are strictly related to the image model.

(Received December 11, 1998.)

REFERENCES

[1] S. Attasi: Modeling and recursive estimation for double indexed sequences. In: System Identification: Advances and Case Studies (R. K. Mehra and D. G. Lainiotis, eds.), Academic Press, New York 1976.

[2] M. A. Azimi-Sadjadi and S. Bannour: Two-dimensional recursive parameter identification for adaptive Kalman filtering. IEEE Trans. Circuits and Systems *CAS-38* (1991), 1077–1081.

[3] M. R. Azimi-Sadjadi and K. Khorasani: Reduced order strip Kalman filtering using singular perturbation method. IEEE Trans. Circuits and Systems *CAS-37* (1990), 284–290.

[4] P. E. Barry, R. Gran and C. R. Waters: Two-dimensional filtering – a state space approach. In: Proc. of Conference Decision and Control 1976, pp. 613–618.

[5] M. A. Bedini and L. Jetto: Realization and performance evaluation of a class of image models for recursive restoration problems. Internat. J. Systems Sci. *22* (1991), 2499–2519.

[6] J. Biemond and J. J. Gerbrands: Comparison of some two-dimensional recursive point-to-point estimators based on a DPCM image model. IEEE Trans. Systems Man Cybernet. *SMC-10* (1980), 929–936.

[7] A. De Santis, A. Germani and L. Jetto: Space-variant recursive restoration of noisy images. IEEE Trans. Circuits and Systems *CAS-41* (1994), 249-261.

[8] A. Germani and L. Jetto: Image modeling and restoration: a new approach. Circuits Syst. Sign. Process. 7 (1988), 427–457.

[9] A. Habibi: Two-dimensional bayesian estimate of images. Proc. IEEE *60* (1972), 878–883.

[10] A. K. Jain: Partial differential equations and finite-difference methods in image processing. Part I: Image representation. J. Optim. Theory Appl. *23* (1977), 65-91.

[11] A. K. Jain and E. Angel: Image restoration, modeling and reduction of dimensionality. IEEE Trans. Comput. *C-23* (1974), 470-476.

[12] A. K. Jain and J. R. Jain: Partial differential equations and finite difference methods in image processing. Part II: Image restoration. IEEE Trans. Automat. Control *AC-23* (1978), 817–833.

[13] T. Katayama: Restoration of noisy images using a two-dimensional linear model. IEEE Trans. Systems Man Cybernet. *SMC-9* (1979), 711–717.

[14] T. Katayama: Estimation of images modeled by a two-dimensional separable autoregressive process. IEEE Trans. Automat. Control *AC-26* (1980), 1199-1201.

[15] T. Katayama and M. Kosaka: Recursive filtering algorithm for a two-dimensional system. IEEE Trans. Automat. Control *AC-24* (1979), 130–132.

[16] H. Kaufman, J. W. Woods, S. Dravida and A. M. Tekalp: Estimation and identification of two-dimensional images. IEEE Trans. Automat. Control *AC-28* (1983), 745–756.

[17] H. R. Keshavan and M. D. Srinath: Sequential estimation technique for enhancement of noisy images. IEEE Trans. Comput. *C-26* (1977), 971–987.

[18] H. R. Keshavan and M. D. Srinath: Enhancement of noisy images using an interpolative model in two dimensions. IEEE Trans. Systems Man Cybernet. *SMC-8* (1978), 247–259.

[19] P. B. Liebelt: An Introduction to Optimal Estimation. Addison-Wesley, Reading, MA 1967.

[20] M. S. Murphy: Comments on 'Recursive filtering algorithm for a two-dimensional system'. IEEE Trans. Automat. Control *AC-25* (1980), 336–338.

[21] M. S. Murphy and L. M. Silverman: Image model representation and line-by-line recursive restoration. IEEE Trans. Automat. Control *AC-23* (1978), 809-816.

[22] N. E. Nahi: Role of recursive estimation in statistical image enhancement. Proc. IEEE *60* (1972), 872-877.

[23] N. E. Nahi and T. Assefi: Bayesian recursive image estimation. IEEE Trans. Comput. *C-21* (1972), 734–738.

[24] N. E. Nahi and C. A. Franco: Recursive image enhancement-vector processing. IEEE Trans. Comm. *Com-21* (1973), 305-311.

[25] D. P. Panda and A. C. Kak: Recursive Filtering of Pictures. Tech. Rep. TR-EE-76, School of Electrical Engineering, Purdue University, Lafayette, Ind., 1976; also in: A. Rosenfield and A. C. Kak: Digital Picture Processing. Chapter 7. Academic Press, New York 1976.

[26] S. R. Powell and L. M. Silverman: Modelling of two-dimensional covariance function with application to image enhancement. IEEE Trans. Automat. Control *AC-19* (1974), 8-13.

[27] M. G. Strintzis: Comments on 'Two-dimensional Bayesian estimate of images'. Proc. IEEE *64* (1976), 1255-1257.

[28] B. R. Suresh and B. A. Shenoi: The state-space realization of a certain class of two-dimensional systems with applications to image restoration. Computer Graphics and Image Processing *11* (1979), 101–110.

[29] B. R. Suresh and B. A. Shenoi: New results in two-dimensional Kalman filtering with applications to image restoration. IEEE Trans. Circuits and Systems *CAS-28* (1981), 307–319.

[30] P. E. Wellstead and J. R. Caldas Pinto: Self tuning filters and predictors for two-dimensional systems. Part I: Algorithms. Internat. J. Control *42* (1985), 479–496.

[31] P. E. Wellstead and J. R. Caldas Pinto: Self tuning filters and predictors for two-dimensional systems. Part II: Smoothing applications. Internat. J. Control *42* (1985), 479–496.

[32] J. W. Woods and C. H. Radewan: Kalman filtering in two dimensions. IEEE Trans. Inform. Theory *IT-23* (1977), 809–816.

[33] Y. H. Yum and S. B. Park: Optimum recursive filtering of noisy two-dimensional data with sequential parameter identification. IEEE Trans. Pattern Anal. Mach. Intell. *PAMI-5* (1983), 337–344.

[34] C. T. Zou, E. I. Plotkin and M. N. S. Swamy: 2-D fast Kalman algorithms for adaptive estimation of nonhomogeneous gaussian Markov random field model. IEEE Trans. Circuits and Systems *41* (1994), 678–692.

Prof. Dr. Leopoldo Jetto, Dipartimento di Elettronica e Automatica, Università di Ancona, Ancona. Italy.
e-mail: L.Jetto@ee.unian.it