

Multichannel L Filters Based on Reduced Ordering

Nikos Nikolaidis, *Student Member, IEEE*, and Ioannis Pitas, *Senior Member, IEEE*

Abstract—Nonlinear multichannel signal processing is an emerging research topic with numerous applications. In this paper we use the so-called reduced ordering (R-ordering) principle to introduce a new family of L filters for vector-valued observations. The coefficients of the proposed filters can be deduced so that the filters are optimal with respect to the output mean squared error. Expressions for the unconstrained, unbiased and location invariant optimal filter coefficients are derived. The calculation of moments of the R-ordered vectors that are involved in these expressions is also discussed. Experiments with noisy two-channel vector fields and noisy color images are presented in order to demonstrate the superiority of the proposed filters over other multichannel filters.

I. INTRODUCTION

MULTICHANNEL signals appear in a variety of important signal processing applications. In multispectral satellite imaging, multispectral scanners are used to record earth surface reflectance at various spectral bands, providing valuable data for military surveillance, mineral exploration, archaeology, etc. Color images can be considered as multichannel signals where each pixel is represented by the triplet of the three primary colors red, green, and blue. Motion vector field processing is crucial for image sequence analysis and compression. Multichannel signals are also involved in multiple antenna transmission systems and AM continuous wave (CW) laser range sensors.

Despite the broad range of applications, the field of multichannel signal processing has experienced a large growth rather recently. Although in most cases the various channels exhibit some degree of correlation, the method that is most commonly used to process multichannel signals is to treat each channel independently, by means of one of the various existing single-channel techniques. Obviously, this approach is not the most appropriate because it does not exploit at all the inter-channel correlation. Furthermore, as it was pointed out in [1], the use of single-channel techniques can lead to unnatural effects on the output signal. Finally, the detection of possible outliers among input data is very difficult in componentwise processing because these outliers may not be clearly distinguishable in all components. The Karhunen-Loève transform can be used alternatively in order to decorrelate the channels prior to the application of the single-channel techniques. However, multichannel techniques

that take into account channel correlation seem to be the most natural way to process multichannel signals [2].

Single-channel nonlinear filtering techniques based on data ordering have been very successful, especially in the area of digital image processing. This success is mainly due to the fact that these techniques have excellent edge preservation properties and, at the same time, are robust to outliers. The single-channel L filters [3], whose output is a linear combination of the ordered input samples, constitute a very important filter class that includes several other filters (e.g., median, α -trimmed mean, arithmetic mean) as special cases. The most significant feature of L filters is the existence of closed formulas that give the optimal (with respect to the output MSE) filter coefficients for a specific additive noise distribution. Unfortunately, L filters as well as other filters that are based on data ordering cannot be readily extended to the multichannel case due to the fact that no globally agreeable ordering scheme exists for multichannel observations. Several subordering principles (marginal, conditional, partial, reduced ordering) have been proposed in the literature [4]–[6]. A number of efforts to utilize these ordering schemes for the introduction of nonlinear multichannel filters have been presented lately. The use of marginal ordering in multichannel signal processing is discussed in [6]. Multichannel L filters based on marginal ordering are proposed in [7]. Reduced ordering has been used in [8] to introduce the R_E , R_M filters. The multichannel modified trimmed mean filter [9] and the multichannel k -nearest neighbor (k -NN) filter [10] are also based on the reduced ordering concept. The vector median filter has been proposed in [1] as an extension of the median in the multichannel case. In [11] the class of $L^j l$ filters is extended to handle multivariate data and an average gradient adaptive technique is used to find the optimal filter coefficients. Robust multichannel filters that are based on the minimum covariance determinant (MCD) estimator [12] are presented in [13]. Multichannel filters whose output is a weighted sum of the input vectors using weights that depend either on the distance between the sample vector and a reference vector or on the sum of distances of the sample vector from all the other vectors are presented in [14], [15]. Finally, the generalized vector directional filter (GVDF) that ranks vectors according to their direction is proposed in [16].

In this paper we present a new family of multichannel L filters that are based on the reduced ordering scheme. The coefficients of the proposed filters can be optimized for a specific noise distribution with respect to the mean squared error between the filter output and the desired, noise-free signal, provided that the latter is constant within the filter window. The structural constraints of unbiasedness and

Manuscript received January 23, 1995; revised September 1, 1995. This paper was recommended by Associate Editor J. Juhola.

N. Nikolaidis is with the Department of Electrical & Computer Engineering, University of Thessaloniki, Thessaloniki 540 06, Greece.

I. Pitas is with the Department of Informatics, University of Thessaloniki, Thessaloniki 540 06, Greece.

Publisher Item Identifier S 1051-8215(96)05197-X.

location invariance are also incorporated in the filter design, giving rise to two filter variants, namely the unbiased and location invariant multichannel L filters.

The structure of the paper is the following. The reduced ordering principle is presented in Section II. The definition of the L filters based on reduced- (R-) ordering and the derivation of the optimal filter coefficients are given in Section III. Section IV deals with the calculation of the moments of the R-ordered vectors that are involved in the evaluation of the filter coefficients. Implementation issues are discussed in Section V. Experiments involving noisy two-channel vector fields and color images are presented in Section VI and conclusions are drawn in Section VII.

II. REDUCED ORDERING

Reduced ordering, also known as R-ordering, orders p -channel vector-valued observations $\mathbf{x}_1, \dots, \mathbf{x}_N$, $\mathbf{x}_i = [x_{i1}, \dots, x_{ip}]^T$ according to their distance d_i from some reference vector $\bar{\mathbf{a}}$ (Fig. 1). As a consequence, multivariate ordering is reduced to one-dimensional ordering. Reduced ordering is rather easy to implement, it can provide clues about outlying observations and is the ordering principle that is the most "natural" for vector-valued observations. The i th order statistic according to R-ordering will be denoted $\mathbf{x}_{[i]} = [x_{[i]1}, \dots, x_{[i]p}]^T$ and is the vector whose distance from the reference vector is ranked i th among distances d_1, \dots, d_N . The distance measure d_i can be the L_2 norm $((\mathbf{x}_i - \bar{\mathbf{a}})^T(\mathbf{x}_i - \bar{\mathbf{a}}))^{1/2}$, the Mahalanobis distance $((\mathbf{x}_i - \bar{\mathbf{a}})^T \Gamma^{-1}(\mathbf{x}_i - \bar{\mathbf{a}}))^{1/2}$ (Γ being the covariance matrix of the input data), or any other distance measure. It is obvious that the choice of an appropriate reference vector is crucial for the reduced ordering scheme. Ideally, the reference vector should be the true value of the underlying vector that is to be estimated. In practice, any suitable multivariate estimator of location (e.g., the arithmetic mean, the marginal or vector median) evaluated over a subset of the input data, can be used as reference vector $\bar{\mathbf{a}}$. The choice of the appropriate estimator depends on the data characteristics.

III. MULTICHANNEL L FILTERS BASED ON R-ORDERING

A p -channel L filter of length N that is based on R-ordering is defined by the following input-output relation:

$$\mathbf{y}(k) = \sum_{i=1}^N \mathbf{A}_i \mathbf{x}_{[i]}(k) \quad (1)$$

where \mathbf{A}_i are $(p \times p)$ coefficient matrices and $\mathbf{x}_{[i]}(k)$ are the R-ordered input vectors $\mathbf{x}(k-\nu), \dots, \mathbf{x}(k), \dots, \mathbf{x}(k+\nu)$, $N = 2\nu + 1$. According to (1), each component $y_j(k)$ of the output vector is a linear combination of all $x_{[i]l}(k)$, $i = 1, \dots, N$, $l = 1, \dots, p$. Let us suppose that the input signal $\mathbf{x}(k)$ is a constant vector-valued signal $\mathbf{s} = [s_1, \dots, s_p]^T$ corrupted by additive zero-mean p -channel noise $\mathbf{n}(k) = [n_1(k), \dots, n_p(k)]^T$

$$\mathbf{x}(k) = \mathbf{s} + \mathbf{n}(k). \quad (2)$$

The noise components are distributed according to some joint distribution function $f_{\mathbf{n}}$. Furthermore, the noise vectors at

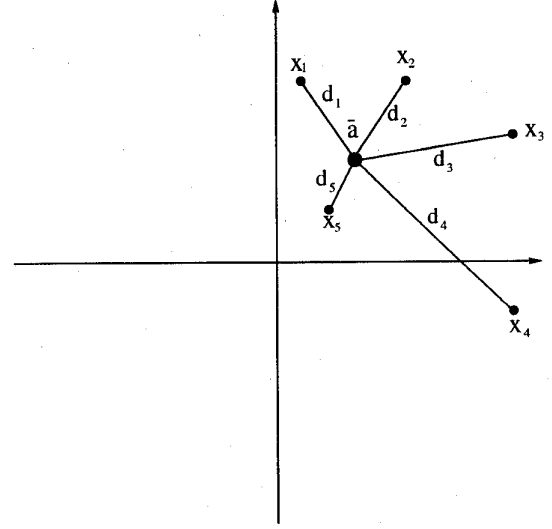


Fig. 1. The reduced ordering principle.

different instants are assumed to be iid and uncorrelated to the constant signal. The optimization procedure that will be presented in the sequel is similar to the one used in [7] for multichannel L filters that are based on marginal ordering. The mean squared error (MSE) between the filter output and the constant signal \mathbf{s} can be expressed in the following way:

$$\begin{aligned} \varepsilon &= E[(\mathbf{y}(k) - \mathbf{s})^T(\mathbf{y}(k) - \mathbf{s})] \\ &= E \left[\sum_{i=1}^N \sum_{j=1}^N \mathbf{x}_{[i]}^T(k) \mathbf{A}_i^T \mathbf{A}_j \mathbf{x}_{[j]}(k) \right] \\ &\quad - 2\mathbf{s}^T E \left[\sum_{i=1}^N \mathbf{A}_i \mathbf{x}_{[i]}(k) \right] + \mathbf{s}^T \mathbf{s}. \end{aligned} \quad (3)$$

The time index k can be dropped without loss of generality. After some manipulation, the previous equation becomes

$$\varepsilon = \sum_{i=1}^N \sum_{j=1}^N \text{tr}[\mathbf{A}_j \mathbf{R}_{ji} \mathbf{A}_i^T] - 2\mathbf{s}^T \sum_{i=1}^N \mathbf{A}_i \underline{\mu}_i + \mathbf{s}^T \mathbf{s} \quad (4)$$

where \mathbf{R}_{ji} is the $(p \times p)$ correlation matrix of the j th and i th order statistics

$$\mathbf{R}_{ji} = E[\mathbf{x}_{[j]} \mathbf{x}_{[i]}^T], \quad i, j = 1, \dots, N \quad (5)$$

and $\underline{\mu}_i$, $i = 1, \dots, N$ denotes the $(p \times 1)$ mean vector of the i th order statistic

$$\underline{\mu}_i = E[\mathbf{x}_{[i]}]. \quad (6)$$

Let us now denote by \mathbf{a}_i the $(Np \times 1)$ vector that is made up of the i th rows of matrices $\mathbf{A}_1, \dots, \mathbf{A}_N$, i.e.,

$$\mathbf{a}_i = [\mathbf{A}_{1i} \ \mathbf{A}_{2i} \ \dots \ \mathbf{A}_{Ni}]^T \quad i = 1, \dots, p \quad (7)$$

where \mathbf{A}_{ji} denotes the i th row of the matrix \mathbf{A}_j . It is obvious that the following relation holds between the elements A_{kij} of matrices \mathbf{A}_k and the elements a_{in} of vectors \mathbf{a}_i :

$$A_{kij} = a_{in}, \quad \text{for } n = (k-1) \cdot p + j. \quad (8)$$

Let us also define the $(Np \times Np)$ matrix $\tilde{\mathbf{R}}_N$ and the $(Np \times 1)$ vector $\tilde{\underline{\mu}}_N$ in the following way:

$$\tilde{\mathbf{R}}_N = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1N} \\ \mathbf{R}_{21} & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{R}_{N1} & \mathbf{R}_{N2} & \cdots & \mathbf{R}_{NN} \end{bmatrix} = E[\tilde{\mathbf{x}}\tilde{\mathbf{x}}^T] \quad (9)$$

$$\tilde{\underline{\mu}}_N = \begin{bmatrix} \underline{\mu}_1 \\ \vdots \\ \underline{\mu}_N \end{bmatrix} = E[\tilde{\mathbf{x}}] \quad (10)$$

where

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x}[1] \\ \vdots \\ \mathbf{x}[N] \end{bmatrix} \quad (11)$$

is the $(Np \times 1)$ vector of all the R-ordered input vectors. After some manipulation and by using the previous notation, the MSE (4) takes the following form:

$$\varepsilon = \sum_{i=1}^p \mathbf{a}_i^T \tilde{\mathbf{R}}_N \mathbf{a}_i - 2\mathbf{s}^T \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p^T \end{bmatrix} \tilde{\underline{\mu}}_N + \mathbf{s}^T \mathbf{s}. \quad (12)$$

A. Unconstrained Filter Coefficients

A multichannel L filter with no constraint imposed on its coefficients is called an *unconstrained L filter*. The optimal coefficients in this case can be found by differentiating (12) with respect to \mathbf{a}_i and equating the result to zero. This optimization problem has a unique solution provided that $\tilde{\mathbf{R}}_N$ is nonsingular. From (9) it is obvious that $\tilde{\mathbf{R}}_N$ is the correlation matrix of the random vector $\tilde{\mathbf{x}}$. Therefore $\tilde{\mathbf{R}}_N$ is positive definite (and thus nonsingular), unless the random variables that form $\tilde{\mathbf{x}}$ are linearly dependent. However, no such case has been observed in the experiments presented in Section VI. The resulting set of equations has the form

$$\tilde{\mathbf{R}}_N \mathbf{a}_i = s_i \tilde{\underline{\mu}}_N, \quad i = 1, \dots, p. \quad (13)$$

By solving these equations for \mathbf{a}_i we obtain the following expression for the optimal unconstrained filter coefficients:

$$\mathbf{a}_i = s_i \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N, \quad i = 1, \dots, p. \quad (14)$$

The minimum MSE obtained using the above coefficients can be easily calculated by incorporating (14) in (12)

$$\varepsilon_{\min} = \mathbf{s}^T \mathbf{s} (1 - \tilde{\underline{\mu}}_N^T \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N). \quad (15)$$

B. Unbiased Filter Coefficients

A multichannel L filter is said to be *unbiased* estimator of location if the following constraint holds:

$$E[\mathbf{y}(k)] = \mathbf{s}. \quad (16)$$

This is equivalent to

$$\mathbf{a}_i^T \tilde{\underline{\mu}}_N = s_i, \quad i = 1, \dots, p. \quad (17)$$

By incorporating the previous constraint into (12), we obtain the following expression for the mean squared error:

$$\varepsilon = \sum_{i=1}^p \mathbf{a}_i^T \tilde{\mathbf{R}}_N \mathbf{a}_i - \mathbf{s}^T \mathbf{s}. \quad (18)$$

The minimization of MSE (18) when coefficients satisfy (17) is a constrained optimization problem and can be solved using the method of Lagrange multipliers. The Lagrangian function to be minimized is the following:

$$\Phi = \sum_{i=1}^p \mathbf{a}_i^T \tilde{\mathbf{R}}_N \mathbf{a}_i - \mathbf{s}^T \mathbf{s} + \sum_{i=1}^p \lambda_i (s_i - \mathbf{a}_i^T \tilde{\underline{\mu}}_N). \quad (19)$$

Differentiating the previous equation with respect to the filter coefficients and equating the result to zero we get

$$\tilde{\mathbf{R}}_N \mathbf{a}_i = \frac{\lambda_i}{2} \tilde{\underline{\mu}}_N, \quad i = 1, \dots, p. \quad (20)$$

The Lagrange multipliers in these equations can be evaluated by solving them for \mathbf{a}_i and substituting to the constraint equations (17). By doing so, the following expression results:

$$\lambda_i = \frac{2s_i}{\tilde{\underline{\mu}}_N^T \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N}. \quad (21)$$

Substituting this result back to (20) we get the optimal unbiased filter coefficients

$$\mathbf{a}_i = \frac{s_i}{\tilde{\underline{\mu}}_N^T \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N} \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N, \quad i = 1, \dots, p. \quad (22)$$

By incorporating (22) in (18) we can calculate the minimum MSE obtained by the optimal unbiased multichannel L filters

$$\varepsilon_{\min} = \mathbf{s}^T \mathbf{s} \left(\frac{1}{\tilde{\underline{\mu}}_N^T \tilde{\mathbf{R}}_N^{-1} \tilde{\underline{\mu}}_N} - 1 \right). \quad (23)$$

C. Location Invariant Filter Coefficients

A multichannel L filter is said to be *location* or *translation invariant* if an input of the form $\mathbf{x}'(k) = \mathbf{x}(k) + \mathbf{b}$ produces an output $\mathbf{y}'(k)$ such that

$$\mathbf{y}'(k) = \mathbf{T}[\mathbf{x}'(k)] = \mathbf{y}(k) + \mathbf{b} \quad (24)$$

where $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$. This property implies that the following constraint holds for the filter coefficients:

$$\sum_{k=1}^N \mathbf{A}_k = \mathbf{I} \quad (25)$$

where \mathbf{I} is the $(p \times p)$ identity matrix. The set of equations (25) can be rewritten in the equivalent form

$$\begin{aligned} \sum_{k=1}^N A_{kii} &= 1, & i &= 1, \dots, p \\ \sum_{k=1}^N A_{kij} &= 0, & i, j &= 1, \dots, p, i \neq j. \end{aligned} \quad (26)$$

Using (25) the output error can be written in the following form:

$$\mathbf{y}(k) - \mathbf{s} = \sum_{i=1}^N \mathbf{A}_i(\mathbf{n}_{[i]}(k) + \mathbf{s}) - \mathbf{s} = \sum_{i=1}^N \mathbf{A}_i \mathbf{n}_{[i]}(k). \quad (27)$$

Therefore, the MSE in the case of the location invariant multichannel L filters is given by

$$\varepsilon = \sum_{i=1}^p \mathbf{a}_i^T \tilde{\mathbf{R}}_N \mathbf{a}_i. \quad (28)$$

The $(Np \times Np)$ matrix $\tilde{\mathbf{R}}_N$ has the same structure as the matrix $\tilde{\mathbf{R}}_N$. However, its submatrices $\tilde{\mathbf{R}}_{ji}$ are the $(p \times p)$ correlation matrices of the j th and i th order statistics of the noise vectors

$$\tilde{\mathbf{R}}_{ji} = E[\mathbf{n}_{[j]} \mathbf{n}_{[i]}^T], \quad i, j = 1, \dots, N. \quad (29)$$

The constrained optimization problem can be solved using Lagrange multipliers. The Lagrangian function that we have to minimize is the following:

$$\begin{aligned} \Phi = & \sum_{i=1}^p \mathbf{a}_i^T \tilde{\mathbf{R}}_N \mathbf{a}_i + \sum_{i=1}^p \lambda_{ii} \left(1 - \sum_{k=1}^N A_{kii} \right) \\ & - \sum_{i=1}^p \sum_{j=1}^p \sum_{i \neq j} \lambda_{ij} \sum_{k=1}^N A_{kij}. \end{aligned} \quad (30)$$

The derivative of Φ with respect to A_{kil} is given by

$$\frac{\partial \Phi}{\partial A_{kil}} = \frac{\partial \Phi}{\partial a_{in}} = \sum_{j=1}^{Np} 2a_{ij} \tilde{R}_{nj} - \lambda_{il}, \quad i, l = 1, \dots, p, \quad k = 1, \dots, N, \quad n = 1, \dots, Np \quad (31)$$

where \tilde{R}_{ij} are the elements of $\tilde{\mathbf{R}}_N$ and $A_{kil} = a_{in}$ (8). By equating the derivatives to zero and writing the set of equations in matrix notation we get

$$\tilde{\mathbf{R}}_N \mathbf{a}_i = \frac{1}{2} \begin{bmatrix} \lambda_i \\ \vdots \\ \lambda_i \end{bmatrix}, \quad i = 1, \dots, p \quad (32)$$

where

$$\lambda_i = [\lambda_{i1}, \dots, \lambda_{ip}]^T. \quad (33)$$

In order to proceed, we have to eliminate the Lagrange multipliers from (32). Let us denote by $\tilde{\mathbf{R}}_p$ the matrix that results from $\tilde{\mathbf{R}}_N^{-1}$ through the following rearrangement of its elements:

$$[\tilde{\mathbf{R}}_p]_{N \cdot (i-1) + j, N \cdot (k-1) + l} = [\tilde{\mathbf{R}}_N^{-1}]_{i+p \cdot (j-1), k+p \cdot (l-1)} \quad i, k = 1, \dots, p, \quad j, l = 1, \dots, N \quad (34)$$

where $[\tilde{\mathbf{R}}_p]_{i,j}$, $[\tilde{\mathbf{R}}_N^{-1}]_{i,j}$ denote the elements of $\tilde{\mathbf{R}}_p$, $\tilde{\mathbf{R}}_N^{-1}$, respectively. $\tilde{\mathbf{R}}_p$ can be partitioned in the following way:

$$\tilde{\mathbf{R}}_p = \begin{bmatrix} \mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1p} \\ \mathbf{P}_{21} & \mathbf{P}_{22} & \cdots & \mathbf{P}_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{p1} & \mathbf{P}_{p2} & \cdots & \mathbf{P}_{pp} \end{bmatrix} \quad (35)$$

where the submatrices \mathbf{P}_{ij} are of dimensions $(N \times N)$. By solving (32) for \mathbf{a}_i and substituting the result in the constraint equations (26), we get the following set of equations:

$$\mathbf{G} \underline{\lambda}_i = [0, \dots, 2, \dots, 0]^T, \quad i = 1, \dots, p \quad (36)$$

where

$$\mathbf{G} = \begin{bmatrix} \mathbf{e}^T \mathbf{P}_{11} \mathbf{e} & \mathbf{e}^T \mathbf{P}_{12} \mathbf{e} & \cdots & \mathbf{e}^T \mathbf{P}_{1p} \mathbf{e} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{e}^T \mathbf{P}_{p1} \mathbf{e} & \mathbf{e}^T \mathbf{P}_{p2} \mathbf{e} & \cdots & \mathbf{e}^T \mathbf{P}_{pp} \mathbf{e} \end{bmatrix} \quad (37)$$

$$\mathbf{e} = [1, \dots, 1]^T \quad (38)$$

and the nonzero element of the $(p \times 1)$ vector in the right side of (36) is in the i th position. By solving (36) for $\underline{\lambda}_i$ and substituting back to (32), we get the optimal location invariant L filter coefficients

$$\mathbf{a}_i = \tilde{\mathbf{R}}_N^{-1} \begin{bmatrix} \mathbf{c}_i \\ \vdots \\ \mathbf{c}_i \end{bmatrix} \quad i = 1, \dots, p \quad (39)$$

where $\mathbf{c}_i = [c_{1i}, \dots, c_{pi}]^T$ is the i th column of \mathbf{G}^{-1} . The minimum MSE for the location invariant filter can be easily evaluated to be

$$\varepsilon_{\min} = \sum_{i=1}^p \begin{bmatrix} \mathbf{c}_i \\ \vdots \\ \mathbf{c}_i \end{bmatrix}^T \tilde{\mathbf{R}}_N^{-1} \begin{bmatrix} \mathbf{c}_i \\ \vdots \\ \mathbf{c}_i \end{bmatrix}. \quad (40)$$

IV. EVALUATION OF MOMENTS OF R-ORDERED DATA

From the previous analysis it is obvious that, in order to proceed with the design of the unconstrained and unbiased multichannel L filters, we have to evaluate the moments $E[x_{[i]k} x_{[j]l}]$, $E[x_{[i]k}]$ where $x_{[i]k}$ denotes the k th element of $\mathbf{x}_{[i]}$. The joint pdf $f_{x_{[i]k} x_{[j]l}}(u, v)$ as well as the pdf $f_{x_{[i]k}}(u)$ must be evaluated for this purpose. In the following, we shall derive formulas for these pdf's when the squared L_2 norm (i.e., the sum of the squared componentwise differences) is used as a distance measure. In order to proceed in our analysis we shall make the assumption that the reference vector is the constant signal vector \mathbf{s} , which we shall consider to be known. Under this assumption the distances d_i are given by

$$d_i = \sum_{k=1}^p (x_{ik} - s_k)^2 = \sum_{k=1}^p n_{ik}^2 \quad (41)$$

where x_{ik} , n_{ik} are the elements of \mathbf{x}_i , \mathbf{n}_i , respectively. From the previous expression and the assumptions that we have made about the noise, it is obvious that d_i are iid random variables. Note also that x_{ik} are independent from all d_j , $j \neq i$. By ordering the distances d_i

$$d_{(1)} \leq d_{(2)} \leq \dots \leq d_{(N)} \quad (42)$$

we impose the same ordering on their corresponding vectors \mathbf{x}_i . The notation $d_i \leftrightarrow d_{(j)}$ will be used in the sequel to denote that d_i occupies the j th position in the ordered sequence $d_{(1)}, \dots, d_{(N)}$. The same meaning is given to the notation $\mathbf{x}_i \leftrightarrow \mathbf{x}_{[j]}$, for the R-ordered vectors.

The following cases must be considered for the evaluation of $f_{x_{[i]k}, x_{[j]l}}(u, v)$.

- 1) $i \neq j$. In this case, the pdf under consideration is the joint pdf of elements from the same ($k = l$) or different ($k \neq l$) channels that belong to different ordered vectors $\mathbf{x}_{[i]}, \mathbf{x}_{[j]}$. The joint pdf can be shown to have the following form:

$$\begin{aligned} f_{x_{[i]k}, x_{[j]l}}(u, v) &= \int_0^{+\infty} \int_0^{w_2} \frac{f_{x_{mk}, d_m}(u, w_1)}{f_{d_m}(w_1)} \cdot \frac{f_{x_{nl}, d_n}(v, w_2)}{f_{d_n}(w_2)} \\ &\cdot f_{d_{(i)}, d_{(j)}}(w_1, w_2) dw_1 dw_2, \quad i < j \end{aligned} \quad (43)$$

where $\mathbf{x}_m \leftrightarrow \mathbf{x}_{[i]}, \mathbf{x}_n \leftrightarrow \mathbf{x}_{[j]}, d_m \leftrightarrow d_{(i)}, d_n \leftrightarrow d_{(j)}$.

- 2) $i = j$, i.e., when we deal with the joint pdf of elements from the same ordered vector $\mathbf{x}_{[i]}$. In this case, the joint pdf has the form

$$\begin{aligned} f_{x_{[i]k}, x_{[i]l}}(u, v) &= \int_0^{+\infty} \frac{f_{x_{nk}, x_{nl}, d_n}(u, v, w)}{f_{d_n}(w)} \\ &\cdot f_{d_{(i)}}(w) dw, \quad k \neq l \end{aligned} \quad (44)$$

where $\mathbf{x}_n \leftrightarrow \mathbf{x}_{[i]}$ and $d_n \leftrightarrow d_{(i)}$.

Furthermore, the pdf $f_{x_{[i]k}}(u)$ can be proven to be

$$f_{x_{[i]k}}(u) = \int_0^{+\infty} \frac{f_{x_{nk}, d_n}(u, w)}{f_{d_n}(w)} f_{d_{(i)}}(w) dw \quad (45)$$

where $\mathbf{x}_n \leftrightarrow \mathbf{x}_{[i]}$ and $d_n \leftrightarrow d_{(i)}$. The proof of these formulas is given in Appendix A. The next step that we have to undertake is to evaluate the pdf's $f_{d_{(i)}}(w)$, $f_{d_{(i)}, d_{(j)}}(w_1, w_2)$, $f_{d_i}(w)$, $f_{x_{ik}, d_i}(u, w)$, $f_{x_{ik}, x_{il}, d_i}(u, v, w)$ that are involved in (43)–(45). We start by evaluating the joint pdf g of $d_i, n_{i1}, \dots, n_{im-1}, n_{im+1}, \dots, n_{ip}$ which is given by (46), shown at the bottom of the page. The previous formula holds for $d_i \geq \sum_{l=1, l \neq m}^p n_{il}^2$. In the opposite case, the joint pdf is equal to zero. The proof for this formula is given in Appendix B. The pdf $f_{d_i}(w)$ can then be found by integrating (46) over $n_{i1}, \dots, n_{im-1}, n_{im+1}, \dots, n_{ip}$

$$\begin{aligned} f_{d_i}(w) &= \int_{C_{i1}}^{D_{i1}} \dots \int_{C_{ip}}^{D_{ip}} \\ &g(n_{i1}, \dots, n_{im-1}, w, n_{im+1}, \dots, n_{ip}) \\ &\cdot dn_{i1} \dots dn_{im-1} dn_{im+1} \dots dn_{ip} \end{aligned} \quad (47)$$

where $[C_{ik}, D_{ik}]$ is the domain of the random variable n_{ik} . The joint pdf of d_i, n_{il} for all $l \neq m$ can be found in a similar

manner by integrating (46) over all $n_{ir}, r \neq m, r \neq l$. The evaluation of the joint pdf of d_i, n_{im} can be done in the same way. The only difference is that, instead of using the joint pdf of $d_i, n_{i1}, \dots, n_{im-1}, n_{im+1}, \dots, n_{ip}$ we must evaluate the joint pdf of $d_i, n_{i1}, \dots, n_{ih-1}, n_{ih+1}, \dots, n_{ip}, h \neq m$ which is given by a formula exactly analogous to (46). Having calculated $f_{n_{il}, d_i}(u, w)$ and taking under consideration that

$$x_{il} = s_l + n_{il}, \quad l = 1, \dots, p \quad (48)$$

we can easily derive the joint pdf of d_i, x_{il}

$$f_{x_{il}, d_i}(u, w) = f_{n_{il}, d_i}(u - s_l, w). \quad (49)$$

The joint pdf of $d_i, n_{il}, n_{ik}, l \neq m, k \neq m$ can be evaluated by integrating (46) over all $n_{ir}, r \neq m, r \neq l, r \neq k$. By combining this result with (48) we get the joint pdf of d_i, x_{il}, x_{ik}

$$f_{x_{ik}, x_{il}, d_i}(u, v, w) = f_{n_{ik}, n_{il}, d_i}(u - s_k, v - s_l, w). \quad (50)$$

Finally, $f_{d_{(i)}}(w)$, $f_{d_{(i)}, d_{(j)}}(w_1, w_2)$ are given by [17]

$$\begin{aligned} f_{d_{(i)}}(w) &= \frac{N!}{(i-1)!(N-i)!} F_{d_i}^{i-1}(w) \\ &\cdot [1 - F_{d_i}(w)]^{N-i} f_{d_i}(w) \end{aligned} \quad (51)$$

$$\begin{aligned} f_{d_{(i)}, d_{(j)}}(w_1, w_2) &= \frac{N!}{(i-1)!(j-i-1)!(N-j)!} \cdot F_{d_i}^{i-1}(w_1) f_{d_i}(w_1) \\ &\cdot [F_{d_i}(w_2) - F_{d_i}(w_1)]^{j-i-1} f_{d_i}(w_2) [1 - F_{d_i}(w_2)]^{N-j}, \\ &i < j, w_1 < w_2 \end{aligned} \quad (52)$$

where $F_{d_i}(w)$ denotes the cdf of d_i . By substituting the derived pdf's in (43)–(45), we can calculate the elements of the matrix $\hat{\mathbf{R}}_N$ and the vector $\hat{\boldsymbol{\mu}}_N$ and, therefore, proceed to the evaluation of the unconstrained and unbiased filter coefficients.

In the case of location invariant filters we must evaluate the moments $E[n_{[i]k} n_{[j]l}], E[n_{[i]k}]$. The corresponding pdf's $f_{n_{[i]k}, n_{[j]l}}(u, v)$, $f_{n_{[i]k}}(u)$ have the same form with (43)–(45) and can be deduced in a straightforward manner from these equations by substituting x_{ik}, x_{il} with n_{ik}, n_{il} . The pdf's $f_{n_{ik}, d_i}(u, w)$, $f_{n_{ik}, n_{il}, d_i}(u, v, w)$ that appear in the resulting expressions have already been calculated as intermediate stages in the evaluation of the moments $E[x_{[i]k} x_{[j]l}], E[x_{[i]k}]$.

$$\begin{aligned} &g(n_{i1}, \dots, n_{im-1}, d_i, n_{im+1}, \dots, n_{ip}) \\ &= \frac{f_{\mathbf{n}}\left(n_{i1}, \dots, n_{im-1}, \sqrt{\left(d_i - \sum_{l=1, l \neq m}^p n_{il}^2\right)}, n_{im+1}, \dots, n_{ip}\right) + f_{\mathbf{n}}\left(n_{i1}, \dots, n_{im-1}, -\sqrt{\left(d_i - \sum_{l=1, l \neq m}^p n_{il}^2\right)}, n_{im+1}, \dots, n_{ip}\right)}{2\sqrt{d_i - \sum_{l=1, l \neq m}^p n_{il}^2}} \end{aligned} \quad (46)$$

V. IMPLEMENTATION ISSUES

In real situations, the constant signal \mathbf{s} as well as the pdf $f_{\mathbf{n}}$ of the noise vector are unknown. The signal \mathbf{s} can be estimated using a suitable multivariate estimator of location acting upon all the vectors in the constant signal area. $f_{\mathbf{n}}$ can also be estimated from the multivariate histogram of the input data in homogeneous (constant) signal regions. In this case, the equations that have been derived in the previous section must be appropriately modified to handle discrete data, i.e., the integrals must be replaced by sums. Furthermore, in most cases the integrals in (43)–(45) and (47) cannot be evaluated analytically and thus, we have to resort to rather tedious numerical integrations. The computational burden required for these calculations depends on the vector dimension p (which controls the size of $\hat{\mathbf{R}}_N, \hat{\boldsymbol{\mu}}_N$ and the number of the nested integrals e.g., in (47)) and on the filter size N (which controls the size of $\hat{\mathbf{R}}_N, \hat{\boldsymbol{\mu}}_N$). Even for moderate values of these parameters (e.g., $p = 3, N = 9$) the computational complexity is significantly large. Therefore, it is much more preferable to estimate directly the moments $E[x_{[i]k}], E[x_{[i]k}x_{[j]l}]$ from the input data using the following unbiased estimators:

$$\hat{E}[x_{[i]k}] = \frac{1}{M} \sum_{m=1}^M x_{[i]k}(m) \quad (53)$$

$$\hat{E}[x_{[i]k}, x_{[j]l}] = \frac{1}{M-1} \sum_{m=1}^M x_{[i]k}(m)x_{[j]l}(m) \quad (54)$$

where $x_{[i]k}(m)$ are the elements of the R-ordered vectors within a window that moves through the data (m being the current window position). The evaluation of the coefficients of all the unconstrained and unbiased multichannel L filters in Section VI was done in this way. In the case of nonconstant vector-valued signals, a segmentation of the signal must be performed prior to the filtering procedure in order to obtain areas with almost constant vectors. For each of these regions, a different set of coefficients must be evaluated, using region-specific estimates for $\mathbf{s}, E[x_{[i]k}], E[x_{[i]k}, x_{[j]l}]$.

Unlike the unbiased and unconstrained multichannel L filters, the evaluation of optimal location invariant filter coefficients does not require *a priori* knowledge of the noise-free input signal but depends only on the noise statistics. Therefore, the signal does not need to be segmented since a single set of coefficients suffices for the entire signal, provided, of course, that all parts of the signal are corrupted by the same noise model. This noise model must be known in order to evaluate the filter coefficients. A convenient way to evaluate the moments $E[n_{[i]k}], E[n_{[i]k}, n_{[j]l}]$ in this case is to generate M sets of N noise vectors, sort them according to R-ordering and then use formulas analogous to (53), (54). If the noise model is not known, one can estimate \mathbf{s} in some homogeneous region, calculate the noise vectors in the region by solving (2) for \mathbf{n} , sort them, and use estimates analogous to (53), (54).

It must be clarified here that, in the case where estimates are used for the signal \mathbf{s} and/or some or all of the statistics required for the computation of the filter coefficients, the corresponding filters are suboptimal. However, as will be seen in Section VI,

even in this case the proposed filters perform better than other multichannel filters.

Since reduced ordering requires a reference vector $\bar{\mathbf{a}}$ to be evaluated on each filter position, the L filters can be considered as a two-stage procedure. In the first stage, the signal is filtered using an appropriate multichannel filter (multichannel α -trimmed, marginal median, arithmetic mean filter, etc.) in order to obtain an estimate for $\bar{\mathbf{a}}$. This estimate is used in the second stage for the evaluation of distances d_i . Alternatively, the proposed filters can be considered as double-window filters like the multichannel double-window modified trimmed mean (DW-MTM) filter or the multichannel k -NN filter [10]. Obviously, this two-stage/double-window procedure leads to increased computational complexity. However, the enhanced performance of the proposed filters makes them an attractive alternative to other multichannel filters, especially in applications where execution time is not critical.

VI. EXPERIMENTAL RESULTS

In order to test the performance of the proposed filters we conducted four sets of experiments involving noisy motion fields and color images. The first two sets of experiments dealt with the filtering of artificially generated two-channel vector fields of dimensions 64×64 . In both these experiments the original, noise-free vector field was composed of two constant signal regions (Fig. 2). The vectors corresponding to the two regions were $\mathbf{s}_1 = [2, -5]^T$ (central region) and $\mathbf{s}_2 = [1, 2]^T$ (surrounding area). In the first experiment the original vector field was corrupted by additive bivariate zero-mean noise whose components were distributed according to the Laplacian–Morgenstern distribution [7] with equal standard deviations $\sigma_i = \sqrt{2}$ on both channels and correlation coefficient $r = 9/32$. In order to apply the proposed filters, the observations in each region were used to calculate the constant signal \mathbf{s} and the moments $E[x_{[i]k}], E[x_{[i]k}, x_{[j]l}], E[n_{[i]k}], E[n_{[i]k}, n_{[j]l}]$ that are required for the evaluation of the 3×3 filter coefficients for this region. The optimal unconstrained and location invariant coefficient matrices \mathbf{A}_i are listed in Fig. 3. Note however that the listed coefficients are optimal only for the particular noise parameters (variance, correlation coefficient) described above. Furthermore, as it was mentioned in Section III-A, the coefficients of the unconstrained filter are signal-dependent and, therefore, they are optimal only for filtering a specific constant vector, i.e., $\mathbf{s} = [1, 2]^T$.

The quantitative criterion that was used to evaluate the performance of the proposed filters was the noise reduction index (NRI)

$$\text{NRI} = 10 \log \frac{\sum_k (\mathbf{y}(k) - \mathbf{s})^T (\mathbf{y}(k) - \mathbf{s})}{\sum_k (\mathbf{x}(k) - \mathbf{s})^T (\mathbf{x}(k) - \mathbf{s})} \quad (55)$$

where $\mathbf{y}(k), \mathbf{x}(k)$, and \mathbf{s} are the filtered, noisy, and reference (noiseless) vectors, respectively. The NRI obtained by the three variants of the proposed filters can be found in Table I along with the NRI obtained by other filters, i.e., the arith-

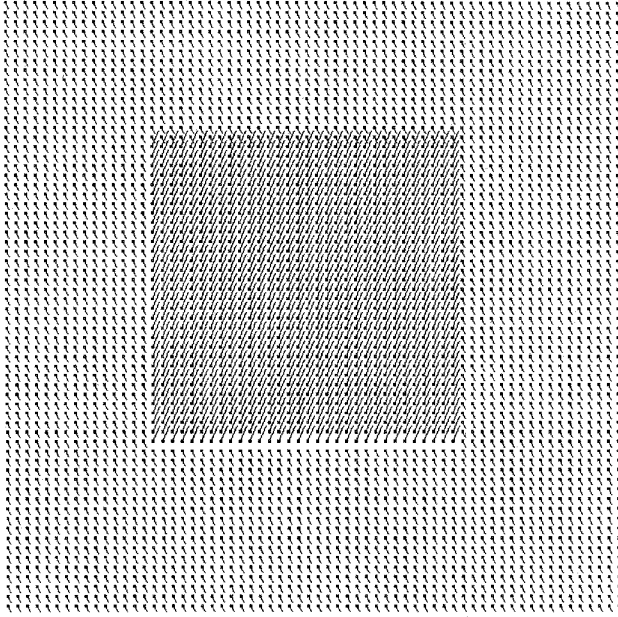


Fig. 2. Original, noise-free vector field.

| | | |
|--|--------------------------------------|---------------------------------------|
| 0.08834 0.20656 0.17696 0.41380 | 0.03666 0.08398 0.07345 0.16825 | 0.01591 0.04687 0.03187 0.09391 |
| A ₁ | A ₂ | A ₃ |
| 0.01013 0.03030 0.02030 0.06070 | 0.00757 0.01884 0.01517 0.03775 | 0.00527 0.01179 0.01057 0.02363 |
| A ₄ | A ₅ | A ₆ |
| 0.00400 0.00605 0.00801 0.01212 | 0.00204 0.00305 0.00409 0.00612 | 0.00065 0.00168 0.00131 0.00338 |
| A ₇ | A ₈ | A ₉ |
| Unconstrained filter coefficients | | |
| 0.48028 0.02182 0.01581 0.49515 | 0.20626 0.00031 0.00330 0.20510 | 0.11500 -0.00385 -0.00874 0.11620 |
| A ₁ | A ₂ | A ₃ |
| 0.07187 -0.00126 -0.00517 0.07472 | 0.05104 -0.00691 -0.00262 0.04871 | 0.03420 -0.00591 -0.00141 0.03121 |
| A ₄ | A ₅ | A ₆ |
| 0.02192 -0.00042 0.00051 0.01522 | 0.01355 -0.00247 -0.00076 0.00878 | -0.00584 -0.00129 -0.00091 0.00487 |
| A ₇ | A ₈ | A ₉ |
| Location invariant filter coefficients | | |

Fig. 3. Optimal coefficient matrices A_i for two-channel unconstrained and location invariant L filters of window size 3×3 and for the noise model (Laplacian noise) described in the first set of experiments in Section VI.

metric mean, marginal median, L_1 vector median, marginal α -trimmed mean filter with trimming coefficient $\alpha = 0.33$ (the coefficient that gave the best results among the others that have been tested) and the multichannel k -NN filter [10] using $k = (N + 1)/2$, generalized distances, and the output of a 5×5 marginal median filter as reference vector. The window size for all filters was set to 3×3 . The CPU time required for the execution of the above filtering algorithms implemented on the same machine (Silicon Graphics Indigo Workstation having a 50 MHz MIPS R4000 CPU) is also included in this table as well as in all the other tables in this section. Note that two different entries are given for each of the three variants of the multichannel L filters. Each entry

TABLE I
NRI OBTAINED BY VARIOUS 3×3 FILTERS
IN THE FILTERING OF A NOISY VECTOR FIELD

| Filter | NRI (db) | | CPU time (sec) |
|----------------------------------|-----------|----------------|----------------|
| | Laplacian | Cont. Gaussian | |
| Unconstrained L #1 | -18.73 | -17.86 | 0.20 |
| Unconstrained L #2 | -12.40 | -11.65 | 0.57 |
| Unbiased L #1 | -18.71 | -17.77 | 0.20 |
| Unbiased L #2 | -12.39 | -11.62 | 0.57 |
| Loc. Invariant L #1 | -16.07 | -15.52 | 0.20 |
| Loc. Invariant L #2 | -10.98 | -9.66 | 0.57 |
| Arith. Mean | -7.04 | -7.66 | 0.03 |
| Marginal Median | -9.70 | -7.85 | 0.07 |
| L_1 Vect. Median | -8.25 | -6.99 | 0.29 |
| α -trimmed, $\alpha=0.33$ | -10.05 | -8.35 | 0.08 |
| 5×5 α -trimmed | -10.91 | -9.44 | 0.39 |
| k -NN | -10.32 | -9.06 | 1.00 |

corresponds to a different method in the evaluation of the reference vector. According to the first approach (entry marked #1), a single estimate of \bar{a} was evaluated for each of the two constant signal regions by taking the arithmetic mean of all the observations belonging to this region. In the second, more realistic approach (entry #2), the reference vector was chosen to be the output of a 5×5 marginal α -trimmed mean filter [6] with trimming coefficient $\alpha = 0.4$, centered on the current filter position. The degradation of the filter performance in the latter case demonstrates the importance of an appropriately chosen reference vector for the proposed filters. However, even in this case, the multichannel L filters are superior to the other listed filters. The performance of a 5×5 marginal α -trimmed mean filter like the one used for the evaluation of the L filter reference vector is also listed in this table. This was done in order to demonstrate the fact that the proposed multichannel L filter, which can be considered to act as a post-processing stage on the α -trimmed mean filter, improves the output of this filter. Regarding computation time, it can be easily seen that multichannel L filters are more computationally intensive than the other filters. Only the multichannel k -NN filter takes more CPU time to execute. By comparing the CPU time for the two entries provided for each multichannel L filter variant (precalculated \bar{a} and \bar{a} evaluated on each filter position separately) one can see that most of the CPU time is consumed for the evaluation of the reference vector.

In the second set of experiments, the same vector field was corrupted by bivariate contaminated Gaussian noise of the form

$$(1 - \epsilon) \cdot N(m_{11}, m_{12}, \sigma_{11}, \sigma_{12}, r_1) + \epsilon \cdot N(m_{21}, m_{22}, \sigma_{21}, \sigma_{22}, r_2) \quad (56)$$

where $m_{11} = m_{12} = m_{21} = m_{22} = 0$, $\sigma_{11} = 1$, $\sigma_{12} = 2$, $\sigma_{21} = 2$, $\sigma_{22} = 4$, $r_1 = 0.5$, $r_2 = 0.7$, $\epsilon = 0.1$. The noisy vector field was filtered using 3×3 multichannel L filters as well as other filters of the same window size. The optimal unconstrained and location invariant coefficient matrices A_i for this noise distribution are listed in Fig. 4. The unconstrained filter coefficients in this figure are optimal for $s = [1, 2]^T$. The evaluation of moments that are necessary for the calculation of the two sets of multichannel

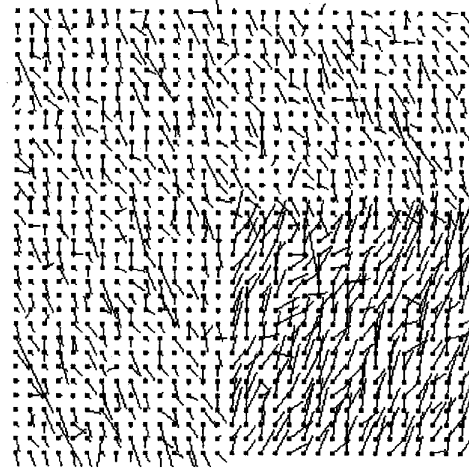
| | | |
|--|--|--|
| A_1 | A_2 | A_3 |
| $\begin{bmatrix} 0.10589 & 0.18413 \\ 0.21051 & 0.36606 \end{bmatrix}$ | $\begin{bmatrix} 0.04943 & 0.07657 \\ 0.09826 & 0.15223 \end{bmatrix}$ | $\begin{bmatrix} 0.03070 & 0.04268 \\ 0.06104 & 0.08486 \end{bmatrix}$ |
| A_4 | A_5 | A_6 |
| $\begin{bmatrix} 0.02287 & 0.02354 \\ 0.04546 & 0.04680 \end{bmatrix}$ | $\begin{bmatrix} 0.01634 & 0.01513 \\ 0.03248 & 0.03009 \end{bmatrix}$ | $\begin{bmatrix} 0.01208 & 0.00938 \\ 0.02402 & 0.01865 \end{bmatrix}$ |
| A_7 | A_8 | A_9 |
| $\begin{bmatrix} 0.01016 & 0.00501 \\ 0.02020 & 0.00997 \end{bmatrix}$ | $\begin{bmatrix} 0.00837 & 0.00233 \\ 0.01665 & 0.00465 \end{bmatrix}$ | $\begin{bmatrix} 0.00779 & -0.00038 \\ 0.01549 & -0.00076 \end{bmatrix}$ |
| Unconstrained filter coefficients | | |
| A_1 | A_2 | A_3 |
| $\begin{bmatrix} 0.36195 & 0.03810 \\ 0.03004 & 0.48364 \end{bmatrix}$ | $\begin{bmatrix} 0.18070 & 0.00544 \\ 0.00627 & 0.20914 \end{bmatrix}$ | $\begin{bmatrix} 0.11968 & 0.00039 \\ -0.00132 & 0.11859 \end{bmatrix}$ |
| A_4 | A_5 | A_6 |
| $\begin{bmatrix} 0.08890 & -0.00798 \\ -0.00125 & 0.07184 \end{bmatrix}$ | $\begin{bmatrix} 0.07116 & -0.00834 \\ -0.00720 & 0.04860 \end{bmatrix}$ | $\begin{bmatrix} 0.05948 & -0.00885 \\ -0.01049 & 0.03314 \end{bmatrix}$ |
| A_7 | A_8 | A_9 |
| $\begin{bmatrix} 0.04901 & -0.00619 \\ -0.00827 & 0.01940 \end{bmatrix}$ | $\begin{bmatrix} 0.03822 & -0.00574 \\ -0.00621 & 0.01140 \end{bmatrix}$ | $\begin{bmatrix} 0.03085 & -0.00683 \\ -0.00153 & 0.00421 \end{bmatrix}$ |
| Location invariant filter coefficients | | |

Fig. 4. Optimal coefficient matrices A_i for two-channel unconstrained and location invariant L filters of window size 3×3 and for the noise model (contaminated Gaussian noise) described in the second set of experiments in Section VI.

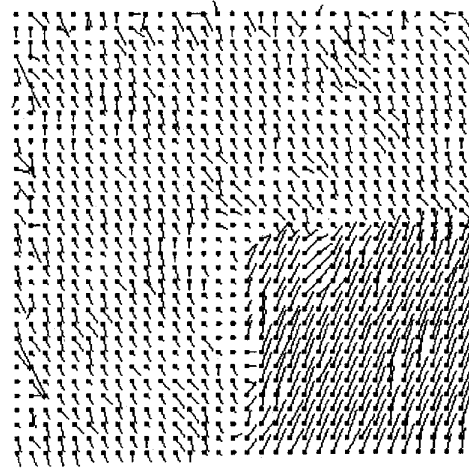
L filter coefficients (one for every region) was done using the observations on the corresponding region. The reference vector was evaluated using the two approaches described in the previous experiment. The trimming coefficient of the 5×5 α -trimmed filter that was used to estimate the reference vector was experimentally chosen to be $\alpha = 0.32$. The results, which are tabulated in Table I, verify once more that the proposed multichannel L filters outperform the other filters.

The performance of the proposed filters in the previous set of experiments can be also seen in Figs. 5 and 6. Only the upper-left quarter of the vector field in Fig. 2 is depicted in these figures in order to have a more detailed visualization. The corrupted vector field is depicted in Fig. 5(a). The output of the 3×3 multichannel k -NN filter (which gave the best results among the filters that were used for comparison) can be seen in Fig. 5(b). It must be noted that the vectors on the borders of the vector field are noisy because they have not been filtered. The result obtained by the 3×3 unconstrained multichannel L filter which uses the arithmetic mean of all the vectors in each region to estimate the reference vector (entry #1 in Table I) can be seen in Fig. 6(a). The filtering results are almost perfect. The effect of filtering with an unconstrained multichannel L filter that uses the output of a 5×5 α -trimmed mean filter as reference vector is presented in Fig. 6(b). It is obvious that, although the filter performance is worse in this case, the results are far better than those obtained by the multichannel k -NN filter.

The third set of experiments dealt with the filtering of noisy motion vector fields. In order to obtain the motion field, a full



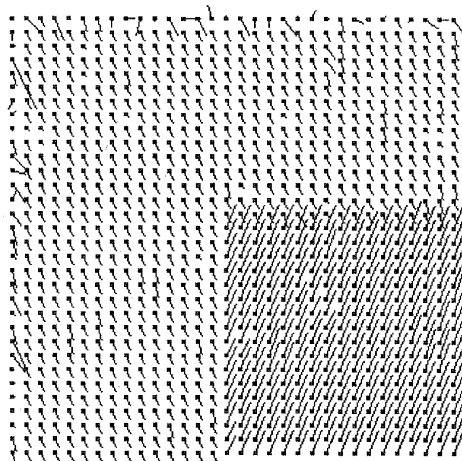
(a)



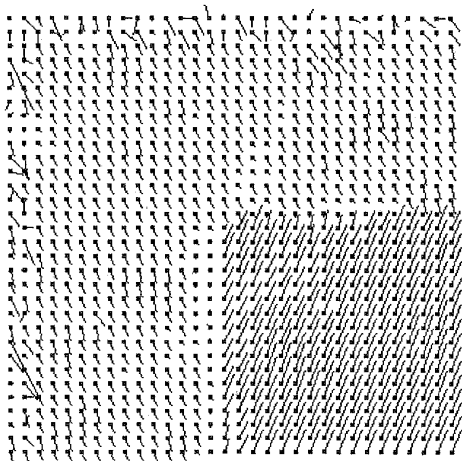
(b)

Fig. 5. (a) Vector field corrupted by contaminated Gaussian noise. (b) Noisy vector field filtered by a 3×3 multichannel k -NN filter.

search block-matching motion estimation algorithm [19] was applied on the frames #1 [Fig. 7(a)] and #10 of the "Trevor White" image sequence. This initial vector field was then smoothed using a 3×3 marginal median filter. The smoothed motion field [Fig. 7(b)] was used as the original, noise-free vector field in this experiment. We have used this procedure to create an uncorrupted motion field since its existence is essential in assessing the filter performance by NRI. The original motion field was corrupted by additive bivariate zero-mean noise whose components were distributed according to the Laplacian-Morgenstern distribution with equal standard deviations $\sigma_1 = \sigma_2 = 6\sqrt{2}$ and correlation coefficient $r = 9/32$ [Fig. 8(a)]. The noisy field was filtered with a location invariant multichannel L filter because, as it was pointed out in the previous section, optimal location invariant filter coefficients depend only on the noise characteristics, which were assumed to be known. Hence, a single set of coefficients suffices for the entire vector field. The method of generating sets of N noise vectors that has been described in Section V



(a)



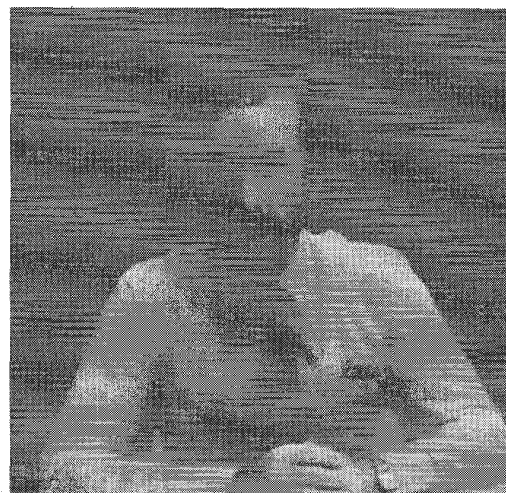
(b)

Fig. 6. Noisy vector field filtered by a 3×3 unconstrained multichannel L filter. (a) The reference vector is evaluated once for each of the two homogeneous areas. (b) The reference vector is evaluated using an α -trimmed mean filter ($\alpha = 0.32$) on a 5×5 window.

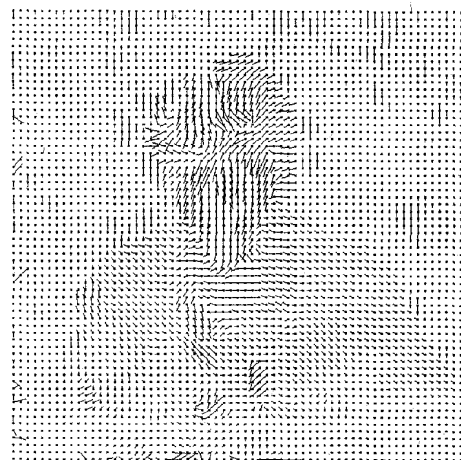
TABLE II
NRI OBTAINED BY VARIOUS 3×3 FILTERS IN THE FILTERING OF
A MOTION VECTOR FIELD CORRUPTED BY LAPLACIAN NOISE

| Filter | NRI (db) | | CPU time (sec) |
|----------------------------------|----------------------|----------------------|-------------------|
| | $\sigma = 4\sqrt{2}$ | $\sigma = 6\sqrt{2}$ | |
| Loc. Invariant L | -9.62 | -11.38 | 0.58 |
| Arith. Mean | -8.31 | -9.04 | 0.03 |
| Marginal Median | -8.93 | -9.83 | 0.07 |
| L_1 Vect. Median | -7.61 | -8.38 | 0.29 |
| α -trimmed, $\alpha=0.25$ | -9.23 | -10.19 | 0.08 |
| k-NN | -8.84 | -10.15 | 1.00 |

was used to evaluate the moments of the ordered noise vectors. The output of the optimal 3×3 location invariant L filter that uses a 5×5 marginal α -trimmed mean filter ($\alpha = 0.25$) to estimate the reference vector \bar{a} is presented in Fig. 8(b). The performance of this filter along with the performance of the marginal median, arithmetic mean, L_1 norm vector median,



(a)



(b)

Fig. 7. (a) First frame of "Trevor White" image sequence. (b) Smoothed motion field obtained from frames #1 and #10.

marginal α -trimmed mean ($\alpha = 0.33$) and multichannel k -NN filter ($k = (N+1)/2$, reference vector: 5×5 marginal median output, generalized distances) of the same window size, are summarized in Table II. The NRI obtained by these filters for a smaller noise standard deviation i.e., $\sigma_1 = \sigma_2 = 4\sqrt{2}$ is also presented in the same table. In both cases, the location invariant multichannel L filter outperforms the other listed filters.

The final set of experiments concerned the filtering of noisy color images using location invariant multichannel L filters. The color image "Pepper" [Fig. 9(a)] has been used in the experiments. The above-mentioned image was corrupted by zero-mean additive multivariate contaminated Gaussian noise in all three RGB channels [Fig. 9(b)]

$$(1 - \epsilon) \cdot \mathcal{N}(m_{1R}, m_{1G}, m_{1B}, \sigma_{1R}, \sigma_{1G}, \sigma_{1B}, r_{RG}, r_{GB}, r_{RB}) \\ + \epsilon \cdot \mathcal{N}(m_{2R}, m_{2G}, m_{2B}, \sigma_{2R}, \sigma_{2G}, \sigma_{2B}, r'_{RG}, r'_{GB}, r'_{RB}) \quad (57)$$

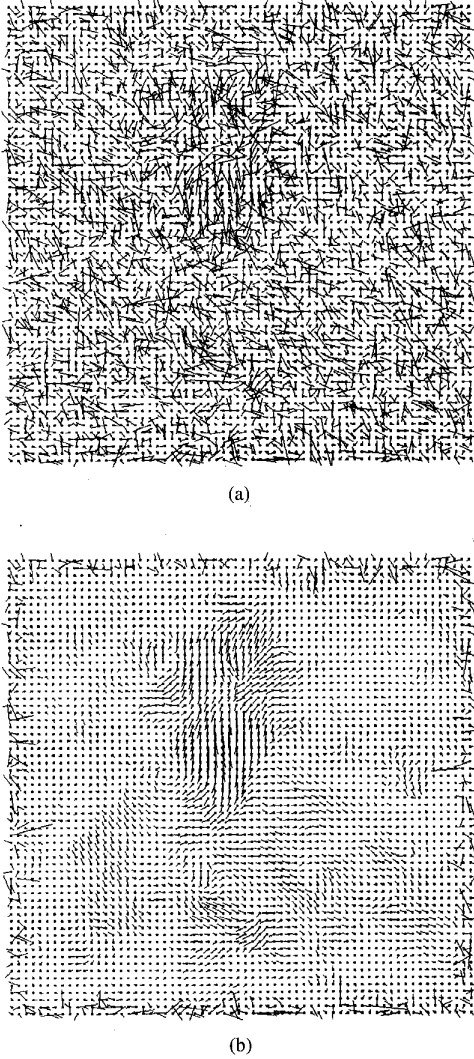


Fig. 8. (a) Motion field corrupted by Laplacian noise. (b) Noisy motion field filtered by a 3×3 location invariant multichannel L filter.

where $m_{1R} = m_{1G} = m_{1B} = m_{2R} = m_{2G} = m_{2B} = 0$, $\sigma_{1R} = \sigma_{1G} = \sigma_{1B} = 20$, $\sigma_{2R} = \sigma_{2G} = \sigma_{2B} = 40$, $r_{RG} = r_{GB} = r_{RB} = r'_{RG} = r'_{GB} = r'_{RB} = 0.5$ and $\epsilon = 0.2$. In order to improve the accuracy of the comparisons, float numbers were used for the representation of the pixel triplets throughout this experiment. The reference vector for the location invariant filter was chosen to be the output of a 5×5 α -trimmed filter with $\alpha = 0.2$. The multichannel location invariant L filter was compared to the arithmetic mean, marginal median (using the L_1 norm), marginal α -trimmed mean ($\alpha = 0.11$), multichannel k -NN filter ($k = (N + 1)/2$, reference vector: 5×5 marginal median output, generalized distances) and optimal single-channel location invariant L filter acting separately on each channel. The performance of the various 3×3 filters was compared with respect to the NRI (Table III). The proposed multichannel filter (marked #1 in Table III) achieved the highest noise reduction among all the filters. A further performance enhancement was achieved by combining the multichannel L filter with a marginal α -trimmed mean

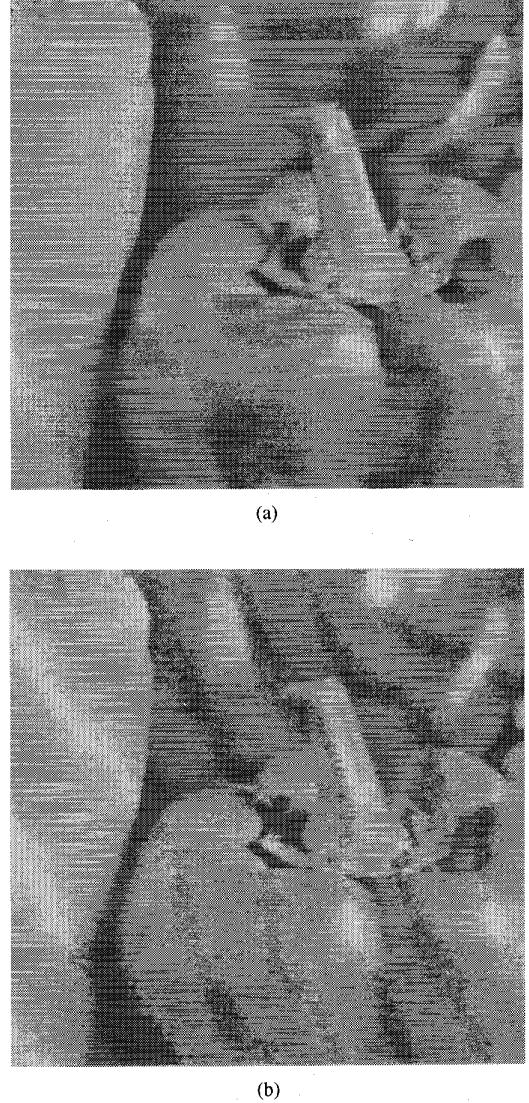


Fig. 9. (a) Original image "Pepper." (b) Image "Pepper" corrupted by contaminated Gaussian noise.

filter. An edge map produced by applying an edge detector on the intensity component of the color image was used to switch between the multichannel L filter (homogeneous regions) and the 3×3 α -trimmed mean filter ($\alpha = 0.222$) (edge areas). By this simple technique the performance of the location invariant filter was improved (entry marked #2 in Table III). The output of the single-channel L filter and the two variants of the multichannel L filter can be seen in Figs. 10 and 11. Note that the multichannel L filter that uses no edge information blurs the edges. However, the multichannel L and α -trimmed filter combination can successfully overcome this drawback.

VII. CONCLUSION

A new class of multichannel L filters that are based on the reduced ordering principle have been presented in this paper. Expressions for the optimal coefficients (with respect to the output MSE) in the case of a constant input signal

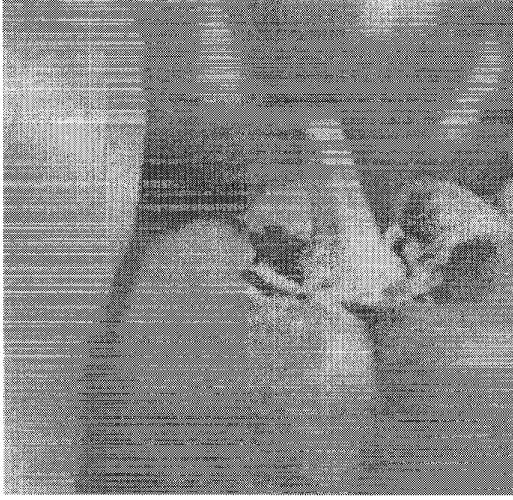


Fig. 10. Noisy image "Pepper" smoothed by componentwise filtering using 3×3 single-channel optimal L filter.

TABLE III
NRI OBTAINED BY VARIOUS 3×3 FILTERS IN THE FILTERING OF COLOR
IMAGE "PEPPER" CORRUPTED BY CONTAMINATED GAUSSIAN NOISE

| Filter | NRI (db) | CPU time (sec) |
|----------------------------------|----------|----------------|
| Loc. Invariant L #1 | -7.85 | 12.80 |
| Loc. Invariant L #2 | -8.14 | 12.96 |
| Arith. Mean | -7.28 | 0.80 |
| Marginal Median | -6.82 | 1.73 |
| L_1 Vect. Median | -5.29 | 5.44 |
| α -trimmed, $\alpha=0.11$ | -7.56 | 2.42 |
| single-channel L | -7.58 | 2.20 |
| k-NN | -7.03 | 26.21 |

corrupted by additive noise have been provided. Experiments with artificially generated vector fields as well as motion vector fields and color images have verified the superior performance of the proposed filters in multichannel signal filtering.

APPENDIX A

□ Evaluation of $f_{x_{[i]k}}(u)$:

We start by evaluating the joint pdf $f_{x_{[i]k}d_{(i)}}(u, w)$ which can be expressed in the following way:

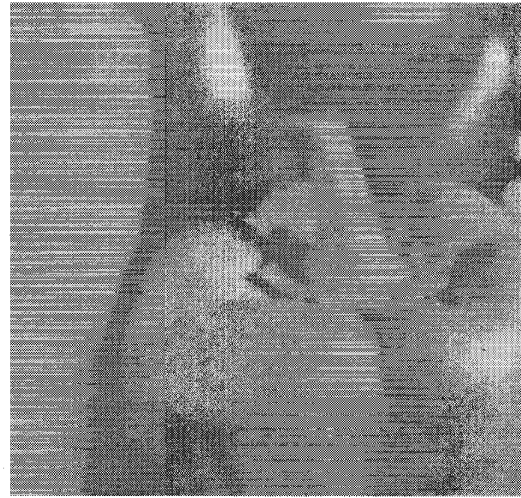
$$f_{x_{[i]k}d_{(i)}}(u, w) = f_{x_{[i]k}}(u|d_{(i)} = w)f_{d_{(i)}}(w), \quad w > 0. \quad (\text{A-1})$$

Since the input vectors x_1, \dots, x_N are independent random vectors, the only dependence of $x_{[i]k}$ on the elements of the other ordered vectors $x_{[j]}, j \neq i$, is through the ordered distance $d_{(i)}$. Therefore, for a fixed value of $d_{(i)}$, the distribution of $x_{[i]k}$ is the same as the distribution of the corresponding, not ordered component. In other words, the conditional pdf of $x_{[i]k}$ subject to the condition that $d_{(i)} = w$ is identical to the conditional pdf of x_{nk} subject to the condition $d_n = w$, where $x_n \leftrightarrow x_{[i]}$ and $d_n \leftrightarrow d_{(i)}$

$$f_{x_{[i]k}}(u|d_{(i)} = w) = f_{x_{nk}}(u|d_n = w) = \frac{f_{x_{nk}d_n}(u, w)}{f_{d_n}(w)}. \quad (\text{A-2})$$



(a)



(b)

Fig. 11. (a) Noisy image "Pepper" filtered by a 3×3 location invariant multichannel L filter without edge information. (b) Noisy image "Pepper" filtered by a 3×3 location invariant multichannel L filter on homogeneous areas and a 3×3 α -trimmed filter on edges.

By incorporating the previous relation in (A-1) we get

$$f_{x_{[i]k}d_{(i)}}(u, w) = \frac{f_{x_{nk}d_n}(u, w)}{f_{d_n}(w)}f_{d_{(i)}}(w), \quad w > 0. \quad (\text{A-3})$$

The previous formula can be alternatively obtained starting from the definition of the joint pdf of two random variables

$$f_{x_{[i]k}d_{(i)}}(u, w) du dw = \text{Prob}(\mathcal{A}) \quad (\text{A-4})$$

where \mathcal{A} is the following event:

$$\mathcal{A} = \{u < x_{[i]k} < u + du, w < d_{(i)} < w + dw\}. \quad (\text{A-5})$$

This event occurs if $u < x_{nk} < u + du, w < d_n < w + dw$ for one of the random variables $d_n, d_n < w$ for $i-1$ of the d_n and $d_n > w + dw$ for the rest $N-i$ of d_n . The number of ways in

which this combined event can happen is $N!/(i-1)!(N-i)!$ and each way has a probability of occurrence

$$f_{x_{nk}d_n}(u, w) du dw F_{d_n}^{i-1}(w)[1 - F_{d_n}(w)]^{N-i}.$$

Therefore

$$f_{x_{[i]k}d_{(i)}}(u, w) du dw = \frac{N!}{(i-1)!(N-i)!} \cdot f_{x_{nk}d_n}(u, w) du dw F_{d_n}^{i-1}(w)[1 - F_{d_n}(w)]^{N-i}. \quad (\text{A-6})$$

By eliminating $du dw$ from both sides of (A-6) and taking into account the expression (51) for the pdf of the ordered distances $d_{(i)}$ we obtain (A-3).

Having evaluated $f_{x_{[i]k}d_{(i)}}(u, w)$ the pdf $f_{x_{[i]k}}(u)$ can be easily obtained by integrating (A-3) over w . By doing so, (45) results.

□ *Evaluation of $f_{x_{[i]k}x_{[j]l}}(u, v)$, $i < j$:*

The evaluation of this joint pdf will be done through the evaluation of $f_{x_{[i]k}x_{[j]l}d_{(i)}d_{(j)}}(u, v, w_1, w_2)$ which can be expressed in the following way:

$$\begin{aligned} f_{x_{[i]k}x_{[j]l}d_{(i)}d_{(j)}}(u, v, w_1, w_2) \\ = f_{x_{[i]k}x_{[j]l}}(u, v | d_{(i)} = w_1, d_{(j)} = w_2) \\ \cdot f_{d_{(i)}d_{(j)}}(w_1, w_2), 0 < w_1 < w_2. \end{aligned} \quad (\text{A-7})$$

However, the conditional joint pdf of $x_{[i]k}, x_{[j]l}$ subject to the condition that $d_{(i)} = w_1, d_{(j)} = w_2$ is identical to the conditional joint pdf of x_{mk}, x_{nl} subject to the condition $d_m = w_1, d_n = w_2$, where $x_m \leftrightarrow x_{[i]k}, x_n \leftrightarrow x_{[j]l}, d_m \leftrightarrow d_{(i)}, d_n \leftrightarrow d_{(j)}$

$$\begin{aligned} f_{x_{[i]k}x_{[j]l}}(u, v | d_{(i)} = w_1, d_{(j)} = w_2) \\ = f_{x_{mk}x_{nl}}(u, v | d_m = w_1, d_n = w_2). \end{aligned} \quad (\text{A-8})$$

Furthermore, since x_i are independent from all $d_j, j \neq i$, the right part of (A-8) can be rewritten as follows:

$$\begin{aligned} f_{x_{mk}}(u | d_m = w_1) f_{x_{nl}}(v | d_n = w_2) \\ = \frac{f_{x_{mk}d_m}(u, w_1)}{f_{d_m}(w_1)} \frac{f_{x_{nl}d_n}(v, w_2)}{f_{d_n}(w_2)}. \end{aligned} \quad (\text{A-9})$$

Therefore

$$\begin{aligned} f_{x_{[i]k}x_{[j]l}d_{(i)}d_{(j)}}(u, v, w_1, w_2) \\ = \frac{f_{x_{mk}d_m}(u, w_1)}{f_{d_m}(w_1)} \frac{f_{x_{nl}d_n}(v, w_2)}{f_{d_n}(w_2)} f_{d_{(i)}d_{(j)}}(w_1, w_2), \\ 0 < w_1 < w_2. \end{aligned} \quad (\text{A-10})$$

By integrating (A-10) over w_1, w_2 equation (43) results.

□ *Evaluation of $f_{x_{[i]k}x_{[i]l}}(u, v)$, $k \neq l$:*

The joint pdf $f_{x_{[i]k}x_{[i]l}d_{(i)}}(u, v, w)$ can be expressed in the following way:

$$\begin{aligned} f_{x_{[i]k}x_{[i]l}d_{(i)}}(u, v, w) \\ = f_{x_{[i]k}x_{[i]l}}(u, v | d_{(i)} = w) f_{d_{(i)}}(w), w > 0. \end{aligned} \quad (\text{A-11})$$

The conditional joint pdf of $x_{[i]k}, x_{[i]l}$ subject to the condition that $d_{(i)} = w$ is identical to the conditional joint pdf of

x_{nk}, x_{nl} subject to the condition $d_n = w$, where $x_n \leftrightarrow x_{[i]}$ and $d_n \leftrightarrow d_{(i)}$

$$\begin{aligned} f_{x_{[i]k}x_{[i]l}}(u, v | d_{(i)} = w) \\ = f_{x_{nk}x_{nl}}(u, v | d_n = w) = \frac{f_{x_{nk}x_{nl}d_n}(u, v, w)}{f_{d_n}(w)}. \end{aligned} \quad (\text{A-12})$$

Therefore

$$\begin{aligned} f_{x_{[i]k}x_{[i]l}d_{(i)}}(u, v, w) \\ = \frac{f_{x_{nk}x_{nl}d_n}(u, v, w)}{f_{d_n}(w)} f_{d_{(i)}}(w), w > 0. \end{aligned} \quad (\text{A-13})$$

Equation (44) can be easily deduced by integrating the previous expression over w .

APPENDIX B

In order to evaluate the joint pdf g of $d_i, n_{i1}, \dots, n_{im-1}, n_{im+1}, \dots, n_{ip}$ we will make use of $p-1$ auxiliary variables u_1, \dots, u_{p-1} such that

$$\begin{aligned} u_1 &= n_{i1}, \dots, u_{m-1} = n_{im-1}, \\ u_m &= n_{im+1}, \dots, u_{p-1} = n_{ip}. \end{aligned} \quad (\text{B-1})$$

The joint pdf g' of d_i, u_1, \dots, u_{p-1} can be evaluated using the well-known method for the evaluation of the pdf of functions of random variables [18]. First of all, we solve the set of equations (B-1), (41) for n_{i1}, \dots, n_{ip}

$$\begin{aligned} n_{i1} &= u_1, \dots, n_{im-1} = u_{m-1}, \\ n_{im}, n'_{im} &= \pm \sqrt{d_i - \sum_{l=1}^{p-1} u_l^2}, \\ n_{im+1} &= u_m, \dots, n_{ip} = u_{p-1}. \end{aligned} \quad (\text{B-2})$$

The joint pdf of d_i, u_1, \dots, u_{p-1} is given by the following relation:

$$\begin{aligned} g'(d_i, u_1, \dots, u_{p-1}) &= \frac{f_{\mathbf{n}}(n_{i1}, \dots, n_{im}, \dots, n_p)}{|J(n_{i1}, \dots, n_{im}, \dots, n_p)|} \\ &+ \frac{f_{\mathbf{n}}(n_{i1}, \dots, n'_{im}, \dots, n_p)}{|J(n_{i1}, \dots, n'_{im}, \dots, n_p)|} \end{aligned} \quad (\text{B-3})$$

where n_{i1}, \dots, n_{ip} are given by (B-2) and $J(\cdot)$ is the following Jacobian:

$$\begin{vmatrix} \frac{\partial u_1}{\partial n_{i1}} & \dots & \frac{\partial u_1}{\partial n_{ip}} \\ \vdots & \vdots & \vdots \\ \frac{\partial d_i}{\partial n_{i1}} & \dots & \frac{\partial d_i}{\partial n_{ip}} \\ \vdots & \vdots & \vdots \\ \frac{\partial u_{p-1}}{\partial n_{i1}} & \dots & \frac{\partial u_{p-1}}{\partial n_{ip}} \end{vmatrix} = \begin{vmatrix} 1 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 2n_{i1} & \dots & 2n_{ip} \\ \vdots & \vdots & \vdots \\ 0 & \dots & 1 \end{vmatrix} = 2n_{im}. \quad (\text{B-4})$$

By substituting (B-2), (B-4) in (B-3) and using $n_{i1}, \dots, n_{im-1}, n_{im+1}, \dots, n_{ip}$ instead of their equivalent variables u_1, \dots, u_{p-1} , (46) results.

REFERENCES

- [1] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters," *IEEE Proc.*, vol. 78, no. 4, pp. 678-689, Apr. 1990.
- [2] R. Machuchá and K. Phillips, "Applications of vector fields to image processing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-5, no. 3, pp. 316-329, May 1983.
- [3] A. C. Bovik, T. S. Huang, D. C. Munson, Jr., "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP-31, no. 6, pp. 1342-1350, Dec. 1983.
- [4] V. Barnett, "The ordering of multivariate data," *J. Royal Statist. Soc. A*, vol. 139, part 3, pp. 318-354, 1976.
- [5] V. Barnett and T. Lewis, *Outliers in Statistical Data*. New York: Wiley, 1984.
- [6] I. Pitas, "Marginal order statistics in color image filtering," *Opt. Eng.*, vol. 29, no. 5, pp. 495-503, May 1990.
- [7] C. Kotropoulos and I. Pitas, "Multichannel L -filters based on marginal data ordering," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2581-2595, Oct. 1994.
- [8] R. C. Hardie and G. R. Arce, "Ranking in R^p and its use in multivariate image estimation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 2, pp. 197-209, June 1991.
- [9] I. Pitas and P. Tsakalides, "Multivariate ordering in color image restoration," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 3, pp. 247-259, Sept. 1991.
- [10] V. Koivunen and S. Kassam, "Nearest neighbor filters for multivariate data," in *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, vol. II, June 1995, pp. 734-737.
- [11] S. Basu, G. Arce, and K. Barner, "Multichannel permutation filters: Definition and optimization," in *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*, vol. II, June 1995, pp. 746-749.
- [12] P. Rousseeuw, "Least median of squares regression," *J. American Statistical Assoc.*, vol. 79, pp. 871-880, 1980.
- [13] V. Koivunen, "A robust approach to enhancement of multivariate images," in *Proc. Int. Conf. Image Processing*, vol. II, pp. 462-466, Nov. 1994.
- [14] A. Buchowicz and I. Pitas, "Multichannel distance filters," in *Proc. IEEE Int. Conf. Image Processing*, vol. II, pp. 575-579, Nov. 1994.
- [15] G. Economou and S. Fotopoulos, "A family of adaptive nonlinear low complexity filters," in *Proc. European Conf. Circuit Theory and Design*, Sept. 1993, pp. 521-524.
- [16] P. E. Trahanias and A. N. Venetsanopoulos, "Vector directional filters—A new class of multichannel image processing filters," *IEEE Trans. Image Processing*, vol. 2, no. 4, pp. 528-534, Oct. 1993.
- [17] H. A. David, *Order statistics*. New York: Wiley, 1980, pp. 8-11.
- [18] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 1984, pp. 172-173.
- [19] A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation and Compression*. New York: Plenum, 1988, pp. 335-340.



Nikos Nikolaidis (S'93) was born in Kozani, Greece in 1968. He received the Diploma in Electrical Engineering (highest honors) from the Aristotle University of Thessaloniki, Greece in 1991.

He is currently a research and teaching assistant and studies towards the Ph.D. degree at the Department of Electrical and Computer Engineering, Aristotle University of Thessaloniki. His research interests include multichannel signal processing, nonlinear image processing and analysis, and copyright protection of digital images.

Mr. Nikolaidis is a scholar of the State Scholarships Foundation of Greece and member of the Technical Chamber of Greece.



Ioannis Pitas received the Diploma of Electrical Engineering in 1980 and the Ph.D. degree in electrical engineering in 1985, both from the University of Thessaloniki, Greece.

Since 1994, he has been a Professor at the Department of Informatics, University of Thessaloniki. From 1980 to 1993 he served as Scientific Assistant, Lecturer, Assistant Professor, and Associate Professor in the Department of Electrical and Computer Engineering at the same University. He served as a Visiting Research Associate at the University

of Toronto, Canada, University of Erlangen-Nuernberg, Germany, Tampere University of Technology, Finland, and as Visiting Assistant Professor at the University of Toronto. He was a lecturer in short courses for continuing education. His current interests are in the areas of digital image processing, multidimensional signal processing, and computer vision. He has published over 190 papers and contributed in eight books in his area of interest. He is the co-author of the book, *Nonlinear Digital Filters: Principles and Applications* (Kluwer, 1990) and author of *Digital Image Processing Algorithms* (Prentice-Hall, 1993). He is the editor of the book *Parallel Algorithms and Architectures for Digital Image Processing, Computer Vision and Neural Networks* (Wiley, 1993).

Dr. Pitas has been a member of the European Community ESPRIT Parallel Action Committee. He has also been an invited speaker and/or member of the program committee of several scientific conferences and workshops. He is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS and co-editor of Multidimensional Systems and Signal Processing. He was chair of the 1995 IEEE Workshop on Nonlinear Signal and Image Processing (NSIP95).