Adaptive multichannel marginal L-filters

C. Kotropoulos I. Pitas*

Department of Informatics

Aristotle University of Thessaloniki

Thessaloniki 540 06, GREECE

E-mail: {costas, pitas}@zeus.csd.auth.gr

Abstract

Three adaptive multichannel L-filters based on marginal data ordering are proposed. They rely on well-known algorithms for the iterative minimization of the mean square error (MSE), namely, the least mean squares (LMS), the normalized LMS (NLMS), and the LMS-Newton (LMSN) algorithms. We treat both the unconstrained minimization of the MSE and the minimization of the MSE when structural constraints are imposed on the filter coefficients. The performance of the proposed adaptive multichannel L-filters is compared to that of other multivariate nonlinear filters in color image filtering. Adaptive multichannel linear filters and adaptive single-channel L-filters are considered as well. Performance comparisons are made in both RGB and $U^*V^*W^*$ color spaces. The proposed adaptive multichannel L-filters outperform the other candidates in noise suppression for color images corrupted by mixed impulsive and additive white contaminated Gaussian noise.

*Corresponding author: I. Pitas, Dept. of Informatics, Aristotle University of Thessaloniki, Box 451, GR-540 06 Thessaloniki, Greece, tel.: +30-31-996304, fax: +30-31-996304

Adaptive signal processing has grown tremendously in recent decades. Adaptive filters are applied in a wide variety of problems including system identification, channel equalization, and echo cancellation in telephone channels [1, 2]. These problems all involve one-dimensional (1-D) signals. Progress has been much slower, however, in the development of adaptive algorithms for two-dimensional (2-D) problems (e.g., image filtering). Extension of the popular 1-D least mean squares (LMS) algorithm to two dimensions was reported in [3] at the 1988. An optimality criterion governing the choice of the convergence factor for the 2-D LMS (TDLMS) algorithm is developed in [4]. A local-mean estimation procedure is incorporated into the TDLMS filter so that the filter preserves image edges [5]. However, linear filters are not suitable for applications where the noise is impulsive, e.g., in image processing, as demonstrated later in the paper.

In the latter case, a multitude of nonlinear techniques have proven to be successful alternatives to the linear techniques [6]. One of the best known nonlinear filter classes is based on *order statistics* [7] and uses the concept of data ordering. A plethora of nonlinear filters are now based on data ordering. Among these are L-filters whose output is defined as a linear combination of the order statistics [8]. The L-filters have found extensive applications in digital signal and image processing because they have a well-defined design methodology. They are the estimators that minimize the mean-square error (MSE) between the filter output and the noise-free signal. An L-filter design that relies on a noniterative minimization of the MSE yields very tedious expressions for computing the marginal and joint cumulative functions of the order statistics (cf. [9]). On the contrary, adaptive L-filters are appealing because they avoid the computational burden of the noniterative methods [10].

Recently, increasing attention has been given to nonlinear processing of vector-valued signals. It is well known that there is no unambiguous or universally accepted method to order multivariate data. Therefore, several *subordering principles* have emerged, namely, marginal ordering (M-ordering), reduced ordering (R-ordering), partial ordering and conditional ordering [11]. The

α-trimmed mean [13], and the multichannel L-filters [14] are a few examples of filters based on the M-ordering principle. The R-ordering principle is another competitive subordering principle. The following multichannel filters based on R-ordering have been proposed: the vector median [15], the ranked-order type estimator $\mathcal{R}_E$ [16], L-filters based on radial medians [17], the multichannel modified trimmed mean (MTM) and its double window extension (DWMTM) [13], the weighted vector median, and the α-trimmed vector median [18]. Other multivariate extensions of the MTM filter and M-filters stemming from M-estimators are investigated in [19], where the influence function of the mentioned filters is derived as well. A multichannel L-filter based on R-ordering was proposed in [20]. New filters that utilize Bayesian techniques and nonparametric methodologies to adapt to local data in image processing are proposed in [21]. A new approach to vector median filtering, where, first, absolute sorting of the vectorial space based on Peano space filling curves is used and, second, a scalar median filtering operation is applied, can be found in [22].

The main contribution of this paper is the design of adaptive multichannel L-filters based on marginal data ordering using the MSE as the fidelity criterion as well as the assessment of their performance in color image filtering. Three novel adaptive multichannel L-filters are proposed that are based on well-known algorithms for the iterative minimization of the MSE, namely, the least mean squares (LMS), the normalized LMS (NLMS) and the LMS-Newton (LMSN) algorithms. Both the unconstrained minimization of the MSE and its minimization under structural constraints on the filter coefficients (e.g., location-invariance) are treated. For a constant multichannel signal, the convergence of the adaptive multichannel L-filters to the optimal solution reported in [14] is demonstrated. Moreover, the performance of the adaptive multichannel L-filters under study in color image filtering is compared to that of other multichannel nonlinear filters including the multichannel DWMTM [13], the multichannel MTM [13], the marginal median [6], the vector median that is based on both the L_1 norm (VM L_1) and on the L_2 norm [15], the ranked-order estimator $\mathcal{R}_E$ [16] and the marginal α-trimmed mean [13]. The case of a color image corrupted

The marginal arithmetic mean is also included in the comparative study, because it is usually an ad hoc choice. Both adaptive multichannel/single-channel linear filters and adaptive single-channel L-filters are considered. The comparative study is conducted first in the RGB color space. However, it is well known that color distances are not Euclidean in this color space [23]. Motivated by this fact, we investigate the performance of the filters under study in a color space where color distances are approximated by Euclidean ones, e.g., in $U^*V^*W^*$ space [23].

It is evident that in this paper no distinction is made between additive white noise and impulsive noise. Another interesting method would be to detect first the pixels that have been corrupted by impulsive noise and subsequently to apply an appropriate technique to deal with the additive white noise. A self-organizing neural network is used to detect the pixels corrupted by impulsive noise and a noise-exclusive median filter is then applied to eliminate the additive white noise in [24]. However, such a method is appropriate for an "off-line" filtering scheme mainly due to the time needed for the self-organizing neural network to converge. In this paper, our interest is focused on "on-line" filtering algorithms.

Adaptive filtering relies generally on the availability of a reference signal, i.e., a reference image. In certain cases (e.g., in image sequences), it is reasonable to assume that a previous noise-free frame can act as a reference image for a number of subsequent image frames. Moreover, full advantage of the temporal correlations that exist between the reference image and the actual desired noise-free image can be taken into account by using motion estimation and compensation in a separate step. Motivated by the success of motion-compensated filtering [25], the use of motion compensation in determining the reference image pixel at each iteration is proposed. When a reference image is not available, we can employ an adaptive location-invariant multichannel L-filter that is modified so that it minimizes the total output power subject to the structural constraints. Indeed such an adaptive filter does not rely on a reference signal, as we can seen later.

The work presented in this paper extends previously reported work [10, 13, 14]. Novel adaptive

The outline of the paper is as follows. The problem statement and our motivation for the design of adaptive multichannel L-filters is given in Section 2. The unconstrained adaptive multichannel L-filters are derived in Section 3. Section 4 is devoted to the design of constrained adaptive multichannel L-filters. Experimental results are included in Section 5 and conclusions are drawn in Section 6.

2 Problem statement

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a random sample of N observations of a p-dimensional random variable $\mathbf{X}$. Each vector-valued observation $\mathbf{x}_j = (x_{j1}, x_{j2}, \ldots, x_{jp})^T$ belongs to a p-dimensional space denoted by $\mathcal{R}^p$. The M-ordering scheme orders the vector components independently, thus yielding

$$x_{i(1)} \leq x_{i(2)} \leq \cdots \leq x_{i(N)} \qquad i = 1, \ldots, p. \tag{1}$$

It has been stated that the output of a p-channel L-filter of length N operating on a sequence of p-dimensional vectors $\{\mathbf{x}(k)\}$ for N odd is given by [14]

$$\mathbf{y}(k) \stackrel{\triangle}{=} \mathbf{T}[\mathbf{x}(k)] = \sum_{i=1}^{p} \mathbf{A}_i \tilde{\mathbf{x}}_i(k), \tag{2}$$

where $\mathbf{A}_i$ are $(p \times N)$ coefficient matrices. Let $\mathbf{a}_{il}^T$, $l = 1, \ldots, p$ denote the l-th row of matrix $\mathbf{A}_i$ and

$$\tilde{\mathbf{x}}_i(k) = \left(x_{i(1)}(k), \ldots, x_{i(N)}(k)\right)^T \qquad i = 1, \ldots, p \tag{3}$$

be the $(N \times 1)$ vector of the order statistics along the i-th channel.

Let us suppose that the observed p-dimensional signal $\{\mathbf{x}(k)\}$ can be expressed as a sum of a p-dimensional noise-free signal $\{\mathbf{s}(k)\}$ and a vector-valued noise sequence $\{\mathbf{n}(k)\}$ having a zero mean vector, i.e., $\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{n}(k)$. The noise vector $\mathbf{n}(k) = (n_1(k), \ldots, n_p(k))^T$ is a p-dimensional vector of random variables characterized by the joint probability density function (pdf) of its components. The noise vector components are assumed to be uncorrelated in the general case.

identically distributed (i.i.d.) and that at each value of index k the signal vector $\mathbf{s}(k)$ and the noise vector $\mathbf{n}(k)$ are uncorrelated. We want to find the multichannel L-filter coefficient matrices $\mathbf{A}_i$, $i = 1, \ldots, p$ that minimize the MSE between the filter output $\mathbf{y}(k)$ and the noise-free signal $\mathbf{s}(k)$. Strictly speaking, we define MSE as the trace of the MSE matrix. Following reasoning similar to that in [14], but without invoking the assumption of a constant signal $\mathbf{s}$, it can be shown that the MSE is expressed as

$$\varepsilon(k) = \mathrm{E}\left[(\mathbf{y}(k) - \mathbf{s}(k))^T (\mathbf{y}(k) - \mathbf{s}(k))\right] = \sum_{i=1}^{p} \{\mathbf{a}_{(i)}^T \tilde{\mathbf{R}}_p \mathbf{a}_{(i)} - 2\mathbf{a}_{(i)}^T \tilde{\mathbf{q}}_{(i)}\} + \mathrm{E}\left[\mathbf{s}^T(k)\mathbf{s}(k)\right] \tag{4}$$

where[1]

$$\begin{aligned} \mathbf{a}_{(i)} &= \left(\mathbf{a}_{1i}^T \mid \mathbf{a}_{2i}^T \mid \cdots \mid \mathbf{a}_{pi}^T\right)^T, \qquad \tilde{\mathbf{R}}_p = \mathrm{E}\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\right], \qquad \tilde{\mathbf{q}}_{(i)} = \mathrm{E}\left[s_i(k)\tilde{\mathbf{X}}(k)\right], \\ \tilde{\mathbf{X}}(k) &= \left(\tilde{\mathbf{x}}_1^T(k) \mid \tilde{\mathbf{x}}_2^T(k) \mid \ldots \mid \tilde{\mathbf{x}}_p^T(k)\right)^T. \end{aligned} \tag{5}$$

It can easily be seen that $\tilde{\mathbf{R}}_p$ is a composite matrix that consists of the correlation matrices of the ordered input samples from the same channel - e.g., $\mathbf{R}_{ii} = \mathrm{E}[\tilde{\mathbf{x}}_i(k)\tilde{\mathbf{x}}_i^T(k)]$ - as well as from different channels - e.g., $\mathbf{R}_{ij} = \mathrm{E}[\tilde{\mathbf{x}}_i(k)\tilde{\mathbf{x}}_j^T(k)]$, $i \neq j$:

$$\tilde{\mathbf{R}}_p = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} & \cdots & \mathbf{R}_{1p} \\ \mathbf{R}_{12}^T & \mathbf{R}_{22} & \cdots & \mathbf{R}_{2p} \\ \vdots & & \ddots & \vdots \\ \mathbf{R}_{1p}^T & \mathbf{R}_{2p}^T & \cdots & \mathbf{R}_{pp} \end{bmatrix}. \tag{6}$$

Minimizing (4) over $\mathbf{a}_{(i)}$ is a quadratic minimization problem that has a unique solution under the condition that $\tilde{\mathbf{R}}_p$ is positive definite. It is easily deduced that the minimum MSE coefficient vector is:

$$\mathbf{a}_{(i)}^o = \tilde{\mathbf{R}}_p^{-1} \tilde{\mathbf{q}}_{(i)}. \tag{7}$$

Eq. (7) explicitly yields the filter coefficients provided that we are able to calculate the moments of the order statistics from univariate populations that appear in $\mathbf{R}_{ii}$, as well as the product moments

[1] We take for granted that $\tilde{\mathbf{X}}(k)$ and $s_i(k)$ are stationary stochastic processes.

is fairly easy for i.i.d. input variates, i.e., in the case of a constant signal $\mathbf{s}(k) = \mathbf{s}$ as demonstrated in [14]. Even for independent, nonidentically distributed input variates, the framework tends to become very complicated (cf. [14]). The difficulties are increased in color image processing, where the observations $\check{\mathbf{X}}(k)$ and the desired signal $\mathbf{s}(k)$ are strongly nonstationary. To overcome these obstacles, we resort to the use of iterative algorithms for the minimization of $\varepsilon(k)$ in (4).

3 Unconstrained minimization of the Mean Squared Error

In this section, three adaptive multichannel L-filters are derived that iteratively minimize the MSE (4) without imposing any constraints on the filter coefficients. These algorithms are (i) the LMS, (ii) the NLMS, and (iii) the LMSN filters. The rationale underlying the choice of each algorithm is stated explicitly.

3-A LMS adaptive multichannel L-filter

The filter coefficient vectors $\mathbf{a}_{(i)}$, $i = 1, \ldots, p$ that minimize the MSE (4) can be computed recursively using the steepest descent algorithm as follows

$$\mathbf{a}_{(i)}(k+1) = \mathbf{a}_{(i)}(k) + \frac{1}{2}\mu[-\nabla\varepsilon_{(i)}(k)], \tag{8}$$

where $\nabla\varepsilon_{(i)}(k) = \frac{\partial\varepsilon(k)}{\partial\alpha_{(i)}(k)}$ and μ denotes the adaptation step-size parameter. Using $\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)$ and $s_i(k)\tilde{\mathbf{X}}(k)$ as instantaneous estimates of $\tilde{\mathbf{R}}_p$ and $\mathbf{q}_{(i)}$, respectively, the LMS adaptive multichannel L-filter is obtained, i.e.,

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \mu\left[s_i(k) - \tilde{\mathbf{X}}^T(k)\hat{\mathbf{a}}_{(i)}(k)\right]\tilde{\mathbf{X}}(k) = \hat{\mathbf{a}}_{(i)}(k) + \mu\, e_i(k)\,\tilde{\mathbf{X}}(k). \tag{9}$$

The bracketed term in (9) is the a priori estimation error $e_i(k)$ between the i-th component of the desired signal and the filter output, because $y_i(k) = \sum_{j=1}^{p}\hat{\mathbf{a}}_{ji}^T\tilde{\mathbf{x}}_j(k) = \tilde{\mathbf{X}}^T(k)\hat{\mathbf{a}}_{(i)}$. We see that the LMS algorithm yields a very simple recursive relation for updating the L-filter coefficients. This is the rationale underlying its choice for minimizing the MSE. Eq. (9) employs the composite

LMS (linear) filter uses the vector of input observations. Accordingly, the convergence properties of the LMS adaptive multichannel L-filter depend on the eigenvalue distribution of the composite correlation matrix $\tilde{\mathbf{R}}_p(k)$. However, an important question arises: Which is the appropriate range of the adaptation step μ that guarantees the convergence of the adaptive filter in both the mean and the mean-square sense? Up-to-date results are available only for stationary environments and for simple cases, such as time-variant system identification in nonstationary environments [1, 2]. To obtain an optimal sequence of adaptation step parameters, we follow the approach proposed in [4]. Let $\mathbf{M}_i$ be a diagonal matrix of dimensions $(pN \times pN)$ associated with the updating equation for the coefficient vector $\mathbf{a}_{(i)}(k)$, i.e.,

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \frac{1}{2}\mathbf{M}_i[-\nabla \varepsilon_{(i)}(k)]. \tag{10}$$

The MSE $\varepsilon(k)$ can be approximated by its instantaneous value, i.e., $\varepsilon(k) = \sum_{i=1}^{p} e_i^2(k)$. Moreover, the a priori estimation error at iteration $(k+1)$ can be expressed in the form of a Taylor series in terms of the a priori estimation error at k, i.e.,

$$e_i(k+1) = e_i(k) + \sum_{j=1}^{pN} \frac{\partial e_i(k)}{\partial a_{(i)j}} \Delta a_{(i)j} + \sum_{j=1}^{pN}\sum_{l=1}^{pN} \frac{\partial^2 e_i(k)}{\partial a_{(i)j}\, \partial a_{(i)l}} \Delta a_{(i)j}\, \Delta a_{(i)l} + \mathcal{O}(3), \tag{11}$$

where $a_{(i)j}$ denotes the j-th element of $\mathbf{a}_{(i)}$, $\Delta a_{(i)j} = a_{(i)j}(k+1) - a_{(i)j}(k)$ and $\mathcal{O}(3)$ are the higher order terms. From (10), we obtain $\Delta a_{(i)j} = \mu_{i,\,jj} e_i(k) \tilde{X}_j(k)$, where $\mu_{i,\,jj}$ is the jj-diagonal element of matrix $\mathbf{M}_i$ and $\tilde{X}_j(k)$ is the j-th element of the composite vector of the ordered observations $\tilde{\mathbf{X}}(k)$. We also have $\frac{\partial e_i(k)}{\partial a_{(i)j}} = -\tilde{X}_j(k)$. Furthermore, due to the definition of filter output (2) the second and higher order derivatives in (11) are zero. Accordingly,

$$e_i(k+1) = e_i(k) \left\{ 1 - \sum_{j=1}^{pN} \mu_{i,\,jj}(k) \tilde{X}_j^2(k) \right\}. \tag{12}$$

By squaring both sides of (12) and setting $\frac{\partial e_i^2(k+1)}{\partial \mu_{i,\,jj}}$ to zero, we find that the optimal step-size sequence $\mu_{i,\,jj}$ satisfies the equation

$$e_i^2(k) \tilde{X}_j^2(k) \left\{ 1 - \sum_{j=1}^{pN} \mu_{i,\,jj} \tilde{X}_j^2(k) \right\} = 0. \tag{13}$$

$$\sum_{j=1}^{pN} \mu_{i,\,jj}(k)\tilde{X}_j^2(k) = 1. \tag{14}$$

It is reasonable to assume that the adaptation step more heavily weighs the filter coefficients that have larger gradients than those having smaller gradients. That is, the adaptation step is given by

$$\mu_{i,\,jj}(k) = \beta|\frac{\partial e_i(k)}{\partial a_{(i)j}(k)}| = \beta|\tilde{X}_j(k)| \tag{15}$$

where β is a proportionality constant. By combining (14) and (15) and solving for β, we obtain the following optimal step-size sequence

$$\mu^*_{i,\,jj}(k) = \frac{|\tilde{X}_j(k)|}{\sum_{l=1}^{pN} |\tilde{X}_l(k)|^3}. \tag{16}$$

We see that (16) does not depend on index i that refers to channels. Therefore, the same step-size sequence can be applied to all equations that update the filter coefficients.

3-B NLMS adaptive multichannel L-filter

Our motivation in employing the NLMS algorithm to iteratively minimize the MSE (4) stems from the following reasons: (i) It provides a way to automate the choice of the adaptation step-size parameter in order to speed up the convergence of the algorithm. (ii) Its design is based on a limited knowledge of the input-signal statistics. It is able to track the varying signal statistics. The updating equations result as follows. Let $\mu_{i,\,jj}(k) = \mu(k)$ be a single adaptation step-size parameter for all the elements of coefficient vectors $\mathbf{a}_{(i)}$. From (14) we obtain:

$$\mu(k) = \frac{1}{\tilde{\mathbf{X}}^T(k)\tilde{\mathbf{X}}(k)}. \tag{17}$$

The substitution of (17) into (9) yields the updating equations for the coefficients of the normalized LMS adaptive multichannel L filter, i.e.:

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \frac{\mu_0}{\tilde{\mathbf{X}}^T(k)\tilde{\mathbf{X}}(k)} e_i(k)\tilde{\mathbf{X}}(k), \qquad i = 1, \ldots, p \tag{18}$$

where $\mu_0 \in (0, 1]$ is a parameter that is introduced for additional control.

The eigenvalue spread of the composite correlation matrix $\tilde{\mathbf{R}}_p(k)$ is large in principle. In such a case, the LMSN algorithm is a powerful alternative to LMS [26]. The LMSN algorithm is an approximate implementation of Newton's method for minimizing a cost function of several variables. It employs computationally efficient estimates for the autocorrelation matrix of the input signal (in our case, of the composite vector of the ordered observations) and for the gradient of the objective function (i.e., the MSE). The updating formula for the LMSN multichannel L-filter is given by

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}_{(i)}}(k) + \frac{1}{2}\mu\, \hat{\mathbf{R}}_p^{-1}(k)\left[-\hat{\nabla}\varepsilon_{(i)}(k)\right] = \hat{\mathbf{a}}_{(i)}(k) + \mu \hat{\mathbf{R}}_p^{-1}(k) e_i(k)\tilde{\mathbf{X}}(k), \qquad i = 1, \ldots, p. \quad (19)$$

Eq. (19) is similar to the coefficient vector recursion described in [1] with the difference that it encompasses an estimate for the composite correlation matrix and the composite vector of the ordered observations instead of the correlation matrix of the input signal and the input multichannel signal itself. An estimate of the composite matrix $\tilde{\mathbf{R}}_p(k)$ can be calculated by using the Robbins-Monro procedure that solves the equation $\mathrm{E}\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) - \tilde{\mathbf{R}}_p(k)\right] = \mathbf{0}$. An iterative solution of this equation is given by $\hat{\mathbf{R}}_p(k) = \hat{\mathbf{R}}_p(k-1) + \zeta\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) - \hat{\mathbf{R}}_p(k-1)\right]$. Using the matrix inversion lemma, we obtain

$$\hat{\mathbf{R}}_p^{-1}(k) = \frac{1}{1-\zeta}\left\{\hat{\mathbf{R}}_p^{-1}(k-1) - \frac{\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)}{\left(\frac{1-\zeta}{\zeta}\right) + \tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)}\right\}. \quad (20)$$

The LMSN algorithm for updating the coefficients of the multichannel marginal L-filter is summarized in Table 1.

3-D Computational complexity of the unconstrained adaptive multichannel L-filters

Having derived the updating equations for the three adaptive multichannel L-filters, we proceed to the study of their computational complexity. A common characteristic of the proposed adaptive algorithms is that they require data ordering. If the vector of the order statistics along the i-th channel is known at $k-1$, $\tilde{\mathbf{x}}_i(k-1)$, the same vector at k can be found by deleting the last

of comparisons for such a running scheme is approximately $2\lfloor \log_2 N \rfloor + 2$ [27]. In the case of multichannel image processing, we delete/insert elements of a vector that contains the image pixel values that are discarded from/appended to the running window. The number of comparisons in the case of of $W \times W$ square window is of the order of $\mathcal{O}(W \log_2 W)$ per channel. Table 2 depicts the storage requirements, the number of comparisons and arithmetic operations for the adaptive multichannel L-filters under study. We see that the storage requirements and the number of arithmetic operations for the LMS adaptive multichannel L-filter with fixed step-size (9), or variable step-size (16) and the NLMS (18) is of the order of $\mathcal{O}(p^2\ N)$ whereas that for the LMSN (19) is of $\mathcal{O}(p^2\ N^2)$.

In the following section, we study the design of constrained adaptive multichannel L-filters that minimize MSE (4) subject to structural constraints imposed on the filter coefficients.

4 Constrained minimization of the MSE

Frequently, structural constraints are imposed on the filter coefficients. For example, in the single-channel case, the sum of the filter coefficients must be equal to one. This is true for both linear adaptive filters such as the TDLMS [3, 5] and for nonlinear adaptive filters such as the location-invariant LMS L-filter [10]. Moreover, as we see in (9), (18), and (19), all the unconstrained adaptive filters derived so far depend on the knowledge of a reference signal $\mathbf{s}(k)$. In certain cases, a reference signal can easily be found. For example, in an image sequence one can always choose a previous noise-free frame as a reference image. In cases where this is not possible, the need to develop an adaptive filter structure that does not rely on a reference signal emerges. It will be shown that the adaptive location-invariant multichannel L-filter can be modified so that it does not depend on a reference signal. Moreover, experiments demonstrated that the performance of the adaptive location-invariant multichannel L-filters is practically independent of the reference signal that is used. Therefore, they possess robustness properties in this sense.

[14]. Let us recall the definition of the location-invariant multichannel L-filter first. A multichannel marginal L-filter is said to be *location-invariant* if its output is able to track small perturbations of its input, i.e., if $\mathbf{x}'(k) = \mathbf{x}(k) + \mathbf{b}$ then $\mathbf{y}'(k) = \mathbf{T}[\mathbf{x}'(k)] = \mathbf{y}(k) + \mathbf{b}$, where $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$. The definition of a location-invariant multichannel L-filter yields the following set of constraints imposed on the filter coefficients

$$\mathbf{G}^T \mathbf{a}_{(i)} = \mathbf{b}_i \quad i = 1, \ldots, p \tag{21}$$

where $\mathbf{1}$ denotes the $(N \times 1)$ unitary vector, i.e., $\mathbf{1} = (1, 1, \ldots, 1)^T$, $\mathbf{G}^T$ is a $(p \times pN)$ matrix having the structure

$$\mathbf{G}^T = \begin{bmatrix} \mathbf{1}^T & \mathbf{0}^T & \cdots & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{1}^T & \cdots & \mathbf{0}^T \\ \vdots & & \ddots & \vdots \\ \mathbf{0}^T & \mathbf{0}^T & \cdots & \mathbf{1}^T \end{bmatrix} \tag{22}$$

where $\mathbf{0}$ is a $(N \times 1)$ vector of zeroes and $\mathbf{b}_i$ is the i-th basis vector in $\mathcal{R}^p$, i.e., a vector whose elements are zero except the i-th element which equals 1. In the following, we derive two constrained adaptive multichannel L-filters based on LMS and LMSN algorithms.

4-A LMS location-invariant multichannel L-filter

We are seeking the L-filter whose output minimizes the MSE (4) subject to (21). A well-established methodology for minimizing a cost function subject to constraints was proposed by Frost [28]. This approach is adopted in our analysis. The problem under study is formulated as the minimization of the following Lagrangian function

$$H(\mathbf{a}) = \frac{1}{2}\varepsilon(k) + \mathbf{\Lambda}^T \begin{bmatrix} \mathbf{G}^T \mathbf{a}_{(1)} - \mathbf{b}_1 \\ \vdots \\ \mathbf{G}^T \mathbf{a}_{(p)} - \mathbf{b}_p \end{bmatrix}, \tag{23}$$

where $e(k)$ is given by (7) and $\Lambda = (\underline{\lambda}_1 | \cdots | \underline{\lambda}_p)$ is a $(p \times 1)$ vector. By differentiating $H(\mathbf{a})$ with respect to $\mathbf{a}_{(i)}$ we obtain

$$\frac{\partial H(\mathbf{a})}{\partial \mathbf{a}_{(i)}} = \tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)} - \tilde{\mathbf{q}}_{(i)}(k) + \mathbf{G}\underline{\lambda}_i. \tag{24}$$

Accordingly the steepest descent solution is given by

$$\mathbf{a}_{(i)}(k+1) = \mathbf{a}_{(i)}(k) - \mu \frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)} = \mathbf{a}_{(i)}(k) - \mu \left[\tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)}(k) - \tilde{\mathbf{q}}_{(i)}(k) + \mathbf{G}\underline{\lambda}_i\right]. \tag{25}$$

We demand $\mathbf{a}_{(i)}(k+1)$ to satisfy the set of constraints (21). By substituting (25) into (21) and solving for $\underline{\lambda}_i$, we get

$$\underline{\lambda}_i = \frac{1}{\mu}(\mathbf{G}^T\mathbf{G})^{-1}\left[\mathbf{G}^T\mathbf{a}_{(i)}(k) - \mu\mathbf{G}^T\left(\tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)}(k) - \tilde{\mathbf{q}}_{(i)}(k)\right) - \mathbf{b}_i\right]. \tag{26}$$

By combining (25) and (26), the steepest descent solution is rewritten as follows

$$\begin{aligned} \mathbf{a}_{(i)}(k+1) &= \left[\mathbf{I} - \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\right]\left\{\mathbf{a}_{(i)}(k) - \mu\left[\tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)}(k) - \tilde{\mathbf{q}}_{(i)}(k)\right]\right\} + \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{b}_i \\ &= \mathbf{P}\left\{\mathbf{a}_{(i)}(k) + \mu\left[\tilde{\mathbf{q}}_{(i)}(k) - \tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)}(k)\right]\right\} + \mathbf{f}_i \quad i = 1, \ldots, p, \end{aligned} \tag{27}$$

where $\mathbf{P}$ is the *projection matrix* of dimensions $(pN \times pN)$ defined by:

$$\mathbf{P} = \left[\mathbf{I} - \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T\right] = \left[\mathbf{I} - \frac{1}{N}\mathbf{G}\mathbf{G}^T\right] \tag{28}$$

and $\mathbf{f}_i$ is a $(pN \times 1)$ vector given by

$$\mathbf{f}_i = \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{b}_i = \frac{1}{N}\mathbf{G}\mathbf{b}_i \quad i = 1, \ldots, p. \tag{29}$$

In (28) and (29) we exploit the fact that $(\mathbf{G}^T\mathbf{G})^{-1} = \frac{1}{N}\mathbf{I}$ where $\mathbf{I}$ denotes the identity matrix of appropriate dimensions. The algorithm is initialized by

$$\mathbf{a}_{(i)}(0) = \mathbf{f}_i \quad i = 1, \ldots, p. \tag{30}$$

By using instantaneous estimates for $\tilde{\mathbf{R}}_p(k)$ and $\tilde{\mathbf{q}}_{(i)}(k)$, the LMS location-invariant multichannel L-filter is obtained

$$\hat{\mathbf{a}}_{(i)}(k+1) = \mathbf{P}\left(\hat{\mathbf{a}}_{(i)}(k) + \mu e_i(k)\tilde{\mathbf{X}}(k)\right) + \mathbf{f}_i \quad i = 1, \ldots, p. \tag{31}$$

be minimized is the total output power subject to constraints (21) that prevent the filter coefficients from becoming identically zero. It can easily be shown that the LMS location-invariant L-filter coefficients are now updated as follows

$$\hat{\mathbf{a}}_{(i)}(k+1) = \mathbf{P}\left(\hat{\mathbf{a}}_{(i)}(k) - \mu y_i(k)\tilde{\mathbf{X}}(k)\right) + \mathbf{f}_i \quad i = 1, \ldots, p. \tag{32}$$

It is evident that by replacing $e_i(k)$ by $-y_i(k)$ the same filter structure can be used for the minimization of the total output power subject to the constraints (21).

4-B LMSN location-invariant multichannel L-filter

The vast majority of constrained adaptive algorithms rely on the LMS algorithm. To the authors' knowledge no attempt has been made to design constrained adaptive filters based on other adaptive algorithms, such as the recursive least squares (RLS) or the LMSN algorithm. The case is much simpler for the LMSN than the RLS algorithm, because LMSN shares the same framework with LMS in the sense that LMSN employs the gradient of the error function (i.e., the Lagrangian (23)) given by (24). By premultiplying both sides of (24) by $\tilde{\mathbf{R}}_p^{-1}(k)$, we obtain

$$\tilde{\mathbf{R}}_p^{-1}(k)\frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)} = \mathbf{a}_{(i)}(k) - \tilde{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{q}}_{(i)}(k) + \tilde{\mathbf{R}}_p^{-1}(k)\mathbf{G}\underline{\lambda}_i \quad i = 1, \ldots, p. \tag{33}$$

By identifying that

$$\mathbf{a}^*_{(i)}(k) = \tilde{\mathbf{R}}_p^{-1}(k)\left[\tilde{\mathbf{q}}_{(i)}(k) - \mathbf{G}\underline{\lambda}_i\right] \quad i = 1, \ldots, p \tag{34}$$

is the solution to the minimization of the cost function (23) (i.e., $\frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)} \mid_{\mathbf{a}^*_{(i)}(k)} = \mathbf{0}$), (34) is rewritten as

$$\mathbf{a}^*_{(i)}(k) = \mathbf{a}_{(i)}(k) - \tilde{\mathbf{R}}_p^{-1}(k)\frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)}. \tag{35}$$

The steepest descent solution is obtained from (35) by adding an additional step-size parameter μ

$$\mathbf{a}_{(i)}(k+1) = \mathbf{a}_{(i)}(k) + \mu\left(\mathbf{a}^*_{(i)}(k) - \mathbf{a}_{(i)}(k)\right) = \mathbf{a}_{(i)}(k) - \mu\tilde{\mathbf{R}}_p^{-1}(k)\frac{\partial H(\mathbf{a}(k))}{\partial \mathbf{a}_{(i)}(k)}. \tag{36}$$

expected values involved in the gradient of $H(\mathbf{a}(k))$ with respect to $\mathbf{a}_{(i)}(k)$, the following recursions result

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \mu \hat{\mathbf{R}}_p^{-1}(k) \left[e_i(k) \tilde{\mathbf{X}}(k) - \mathbf{G} \underline{\lambda}_i \right] \qquad i = 1, \ldots, p. \tag{37}$$

By demanding that (37) satisfies the set of constraints (21), we get

$$\underline{\lambda}_i = \frac{1}{\mu} \left(\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right)^{-1} \left[\mathbf{G}^T \hat{\mathbf{a}}_{(i)}(k) - \mathbf{b}_i \right] + \left(\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right)^{-1} \mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) e_i(k) \tilde{\mathbf{X}}(k). \tag{38}$$

The substitution of (38) into (37) yields

$$\begin{aligned} \hat{\mathbf{a}}_{(i)}(k+1) &= \left\{ \mathbf{I} - \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \left[\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right]^{-1} \mathbf{G}^T \right\} \left[\hat{\mathbf{a}}_{(i)}(k) + \mu \hat{\mathbf{R}}_p^{-1}(k) e_i(k) \tilde{\mathbf{X}}(k) \right] \\ &+ \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \left[\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right]^{-1} \mathbf{b}_i. \end{aligned} \tag{39}$$

The comparison between (39) and (27) reveals that the structure of the LMSN location-invariant multichannel L-filter is the same with that of the LMS location-invariant one but with a different matrix $\mathbf{P}$ and a different vector $\mathbf{f}_i$. The new matrix $\mathbf{P}'$ and vector $\mathbf{f}_i'$ are now given by

$$\mathbf{P}' = \mathbf{I} - \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \left[\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right]^{-1} \mathbf{G}^T \tag{40}$$

$$\mathbf{f}_i' = \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \left[\mathbf{G}^T \hat{\mathbf{R}}_p^{-1}(k) \mathbf{G} \right]^{-1} \mathbf{b}_i \qquad i = 1, \ldots, p. \tag{41}$$

The updating equation (39) is initialized by:

$$\mathbf{a}_{(i)}(0) = \mathbf{f}_i' \qquad i = 1, \ldots, p. \tag{42}$$

4-C Computational complexity of the adaptive location-invariant multichannel L-filters

The computational complexity of the constrained adaptive multichannel L-filters under study was derived as well. Table 3 summarizes the additional storage requirements and arithmetic operations needed in the LMS and LMSN location-invariant multichannel L-filters, respectively. We assume that the step-size parameter in LMS algorithm is chosen as in (17). It is seen for both algorithms the additional storage requirements are of the order of $\mathcal{O}(p^2\ N^2)$ and the additional arithmetic

filter requires the inversion of a $(p \times p)$ matrix whose computational complexity is of order $\mathcal{O}(p^3)$. The inspection of Table 3 reveals that the LMSN location-invariant multichannel L-filter is more computationally demanding than the LMS location-invariant multichannel L-filter.

5 Experimental results

In this section, we present two sets of experiments in order to assess the performance of the adaptive multichannel L-filters that we have discussed so far. In the first set of experiments, two-channel 1-D input signal sequences generated by corrupting a constant signal by additive white bivariate contaminated Gaussian noise are used. Both unconstrained and constrained adaptive multichannel L-filters are considered. The purpose is to demonstrate that the adaptive filters yield the same results with the optimal solutions reported in [14]. In the second set of experiments, we deal with color images, i.e., three-channel 2-D signals. Our aim is to test the performance of the proposed adaptive multichannel L-filters in noisy color image filtering.

5-A Two-channel 1-D signals

First, the case of a two-channel 1-D signal $\mathbf{s}(k) = \mathbf{s}$ corrupted by additive white bivariate contaminated Gaussian noise is treated, because for such a signal, the optimal multichannel L-filter coefficients were derived in [14]. Let $\mathcal{N}(\xi_1, \xi_2 \ ; \sigma_1, \sigma_2 \ ; r)$ denote a joint bivariate Gaussian distribution, where the parameters ξ_i and σ_i, $i = 1, 2$, are the expected values and the standard deviations of the components, respectively. Let r be the correlation coefficient. A vector valued signal $\mathbf{s} = (1.0, 2.0)^T$ corrupted by additive white bivariate noise $\mathbf{n}(k)$ with pdf given by $(1 - \varrho)\mathcal{N}(0, 0 \ ; 1, 3 \ ; 0.5) + \varrho\mathcal{N}(0, 0 \ ; 3, 9 \ ; 0.7)$, for $\varrho = 0.1$ was used as a test signal, as in [14]. We demonstrate the convergence of the ensemble averaged squared error for each adaptive algorithm. Subsequently, the noise reduction index (NR) defined as the ratio of the output noise power to the

$$\mathrm{NR} = 10 \log \frac{\sum_k (\mathbf{y}(k) - \mathbf{s}(k))^T (\mathbf{y}(k) - \mathbf{s}(k))}{\sum_k (\mathbf{x}(k) - \mathbf{s}(k))^T (\mathbf{x}(k) - \mathbf{s}(k))}. \tag{43}$$

is measured and is compared to that achieved by the nonadaptive multichannel L-filter.

To begin with, let us consider the performance of the LMS, NLMS and LMSN adaptive multichannel L-filters. A sequence $\{\mathbf{x}(k)\}$ of 10000 samples was created and the squared norm of the estimation error $\|\mathbf{e}(k)\|^2 = \|\mathbf{s}(k) - \mathbf{y}(k)\|^2$ was computed. This experiment was repeated 200 times, each time using an independent realization of the process $\{\mathbf{n}(k)\}$. The averaged squared norm of the estimation error is then determined by computing the ensemble average of $\|\mathbf{e}(k)\|^2$ over the 200 independent trials of the experiment. Thus, an approximation of the ensemble-averaged learning curve of each adaptive algorithm was obtained and is plotted in Figure 1. The filter length N is 9 in all cases. All the filter coefficient recursions were initialized by $\mathbf{a}_{(i)}(0) = \mathbf{0}$. The tunable parameters (e.g., the adaptation step-size etc.) in each algorithm were chosen so that the adaptive filters yield an NR index approximately equal to that of the optimal non-adaptive multichannel L-filter reported in [14]. For example, μ has been equal to 5×10^{-5} in the LMS multichannel L-filter. In the NLMS algorithm, the best choice for parameter μ_0 that weighs the adaptation step size sequence was found to be $\mu_0 = 0.1$. For the LMSN algorithm, μ, ζ and δ were set to 0.0008, 0.001, and 0.01, respectively. From Figure 1d, it is obvious that NLMS adaptive multichannel L-filter attains the fastest convergence rate. The estimates of the multichannel L-filter coefficients were obtained by averaging the steady state values of $\hat{\mathbf{a}}_{(i)}(k)$, $i = 1, \ldots, p$ over the 200 independent trials of the experiment. Using these filter coefficients the NR index for each adaptive multichannel L-filter was calculated. The results are tabulated in Table 4. To facilitate the comparisons, the NR index achieved by the nonadaptive multichannel L-filter designed in [14] is also given. By comparing the NR indices tabulated in Table 4 and the learning curves plotted in Figure 1, we conclude that: (i) All algorithms converge toward the optimal solution. (ii) LMSN adaptive multichannel L-filter more closely matches the NR achieved by the nonadaptive design. (iii) LMS algorithm is the second best. (iv) Although NLMS attains the fastest convergence rate, we see that its NR is approximately

Subsequently, we examine the performance of the adaptive location-invariant multichannel L-filters under study. The already described experimental setup was used in this case as well. The learning curves of the LMS and LMSN location-invariant multichannel L-filters are given in Figure 2a and 2b, respectively. We see that LMSN exhibits a faster convergence rate. Moreover, its steady state MSE is lower than the LMS, as can be deduced from Table 5, where the NR achieved by both adaptive algorithms is tabulated, and Figure 2. The NR achieved by the nonadaptive design [14] is included in Table 5 for comparison purposes. The NR achieved by the marginal median is included for the same purposes as well. The LMSN location-invariant multichannel L-filter outperforms the nonadaptive one by 1 dB. This discrepancy is attributed to the errors occurring in the estimation of the moments of the marginal order statistics employed in [14].

5-B Color image filtering

The second set of experiments deals with color images, i.e., three-channel 2-D signals. To begin with, let us relate the present task to the problem formulation adopted thus far. An observed color image is denoted by $\mathbf{x}(\mathbf{k})$ where $\mathbf{k} = (i, j)$ is the vector of the pixel coordinates. It can refer to any color space e.g. RGB, XYZ, $U^*V^*W^*$ etc. A neighborhood is defined around each pixel $\mathbf{k}$. Our purpose is to design a filter defined on this neighborhood (to be called the filter window hereafter) that aims at estimating the noise-free central pixel value $\mathbf{s}(\mathbf{k})$. For each color component, we rely on the square window of dimensions $W \times W$, where w is assumed to be an odd number ($W = 2w + 1$). Let $N = W^2$. By rearranging the preceding $W \times W$ filter window in a lexicographic order (i.e., row by row) to a $N \times 1$ vector, we obtain

$$\mathbf{x}_l(k) = (x_l(i-w, j-w), x_l(i-w, j-w+1), \ldots, x_l(i-w, j+w), \ldots, x_l(i+w, j+w))^T . \quad (44)$$

We assume that the window is sliding over a color image plane in a raster scan fashion. If K and L denote the image rows and columns, respectively, a scalar running index $k = (i-1)K + j$ where $1 \leq i \leq K$ and $1 \leq j \leq L$ can replace the pixel coordinates $\mathbf{k}$. Therefore, the 1-D notation adopted

Our goal is to test the ability of the proposed adaptive multichannel L-filters in suppressing the noise in color images. Both unconstrained and constrained adaptive multichannel L-filters are considered. We compare the noise reduction capability of the adaptive multichannel L-filters under study to the other multichannel nonlinear filters as well as to their single-channel counterparts. The following nonlinear filters were considered: the vector median [15], the marginal median [12, 13], the marginal α-trimmed mean [12, 13], the multichannel MTM [13], the multichannel DWMTM [13] and the ranked-order estimator $\mathcal{R}_E$ [16]. Whenever the filter window is not specified explicitly, a window of dimensions 3×3 is assumed.

The multichannel DWMTM filter [13] uses two window sizes 3×3 and 5×5, as in the single-channel case [29]. The multichannel MTM filter is a generalization of the single-channel one proposed in [29]. Its output is evaluated as for the DWMTM filter with the exception that only one window of size 3×3 is used [13]. The trimming parameter for α-trimmed mean filters has been 0.2 in all experiments. For the $\mathcal{R}_E$ filter [16], only the best result is tabulated. We have also included the arithmetic mean in the comparative study, because it is a straightforward choice for noise filtering in many practical applications. The performance of three adaptive single-channel L-filters that are used to filter the noise in each primary color component (i.e., channel) was independently taken into consideration as well. A multichannel extension of the 2-D LMS (TDLMS) algorithm (i.e., adaptive multichannel linear filter) and the TDLMS (i.e., adaptive single-channel linear filter) [3] that is used in each primary color component separately were included in the comparative study. We employed the NR index defined by (43) as an objective figure of merit in the performance comparisons. Moreover, the visual quality of the filtered images was used as a subjective figure of merit.

In all experiments, the adaptive linear/nonlinear filter coefficients were initialized by zero. The LMS and LMSN location-invariant multichannel L-filters are initialized by (30) and (42), respectively. Moreover, the filter coefficients determined recursively by the adaptive algorithm at each

sidered in this paper depend explicitly on the knowledge of a reference image (e.g., the original image $\mathbf{s}(k)$) as we can see in the updating equations for the filter coefficients. However, a reference image is seldom available in practice. In the context of an image sequence (e.g. video), we can assume that a previous noise-free frame can act as a reference image for a number of subsequent image frames. Furthermore, in such a context, it is also possible to exploit motion compensation (MC) in determining the desired response at each pixel assuming that the displacement vectors between the frame acting as a reference image and the actual noise-free image are known beforehand. Let $s_L(\mathbf{k}\,;t)$ and $s_L(\mathbf{k}\,;t-\tau)$ be the luminance components (i.e., the Y component in XYZ color space) of two color image frames that are τ time instants apart. Motivated by the success of motion-compensated filtering [25], we propose the following choice of a reference image $\hat{\mathbf{s}}(\mathbf{k}\,;t) = \mathbf{s}(\mathbf{k}-\mathbf{d}^*\,;t-\tau)$ where $\mathbf{d}^*$ is the displacement vector that minimizes a prediction error of the form $PE(\mathbf{d}) = \sum_{\mathcal{S}} |s_L(\mathbf{k}\,;t) - s_L(\mathbf{k}-\mathbf{d}\,;t-\tau)|$ in a neighborhood $\mathcal{S}$ around each pixel $\mathbf{k}$. A block of 8×8 pixels was used in the estimation of the displacement vectors between the 50th frame and frames 45th – 49th of "Trevor White". The displacement vector field between the 50th and 48th frames produced by the block-matching algorithm, is shown in Figure 3a. The displacement vector field between the 50th and the 45th frames is also shown in Figure 3b.

Another point that requires some further clarification is the choice of the color space where the performance comparisons are to be made. It is well known that color distances are not Euclidean in the RGB primary system [23]. Color distances are approximated by Euclidean distances in the so called *uniform color spaces* e.g. the modified universal camera site (USC), the $L^*a^*b^*$, the $L^*u^*v^*$, and the $U^*V^*W^*$ [23]. To guarantee that the measured NR indices correspond to perceived color differences, we felt the need to test the performance of the several filters in a uniform color space. We chose the $U^*V^*W^*$ space for this purpose.

Let us consider the 50th frame of color image sequence "Trevor White". This frame is corrupted by additive white trivariate contaminated Gaussian noise having the probability distribution (1 –

primary color component are replaced by impulses of value 0 or 255 (i.e., positive and negative impulses). Here $\mathbf{C}_i$, $i = 1,2$, denotes the covariance matrix of each trivariate joint Gaussian distribution. The following covariance matrices were used

$$\mathbf{C}_1 = \begin{bmatrix} 100 & 100 & 210 \\ 100 & 400 & 180 \\ 210 & 180 & 900 \end{bmatrix} \quad \mathbf{C}_2 = \begin{bmatrix} 900 & -300 & -210 \\ -300 & 400 & 60 \\ -210 & 60 & 100 \end{bmatrix}. \tag{45}$$

The description of the experiments in this set is organized as follows. First, we justify why multichannel nonlinear filtering is worth pursuing for color images corrupted by the noise model already described. Subsequently, filtering results are presented in RGB and $U^*V^*W^*$ color spaces. Next we demonstrate that by employing motion compensation, the dependence of the results obtained on the reference image used to determine the filter coefficients can be alleviated. Finally, we focus on the performance of the LMS/LMSN location-invariant multichannel L-filters in both color spaces and we compare the visual quality of the filtered images they produce to that of the other filtered images.

To begin with we answer the following questions: (i) Is multichannel filtering preferable to single-channel filtering in the problem examined? (ii) Shall we rely on nonlinear adaptive filtering techniques or on linear ones? Let us assume for the moment that the original noise-free image is available. Table 6 summarizes the performance of the NLMS and LMSN algorithms for L-filtering and linear filtering in a multichannel as well as in a single-channel approach in both color spaces, i.e., in RGB and in $U^*V^*W^*$. The experiments in $U^*V^*W^*$ space were conducted as follows. (a) The noisy input color image and the reference color image was transformed to $U^*V^*W^*$ color space. (b) Filtering was performed in $U^*V^*W^*$ color space. (c) The filter performance was measured in this domain. In NLMS adaptive algorithm, the adaptation step-size parameter μ_0 is chosen as 0.05. In LMSN algorithm, we chose $\mu = 5 \times 10^{-4}$, $\delta = 0.01$, and $\zeta = 0.001$. From a careful inspection of Table 6 the following conclusions are drawn:

color spaces. There is a gain of 2-2.5 dB in NR by employing multichannel adaptive filtering techniques in RGB. The NR gain in $U^*V^*W^*$ ranges between 0.4-2.5 dB.

2. The adaptive multichannel L-filters outperform the adaptive multichannel linear filters by 2.5-3 dB in RGB space. They offer a 0.85-1.54 dB higher NR than the adaptive multichannel linear techniques in $U^*V^*W^*$.

A clear difference between the approach and the type of filtering is seen. In the following, the 48th color image frame of "Trevor White" is used as the reference image for the adaptive filtering techniques.

The NR achieved in RGB color space by the filters under study is given in Table 7. The adaptation step-size parameter that yields the best result in terms of the visual quality of the filtered image has been used in NLMS algorithms. It is included in the corresponding entry of Table 7. For LMSN algorithms, the parameters used in the previous experiment yield filtered images with the best visual quality. Note that the best $\mathcal{R}_E$ estimator corresponds to index $J = 2$ [16]. By examining Table 7 we conclude:

1. Nonlinear filters are ranked as the four best filters, namely, the multichannel DWMTM filter, the MC LMSN adaptive multichannel L-filter, the MC NLMS multichannel L-filter and the MC LMSN location-invariant multichannel L-filter. The MC LMSN/NLMS adaptive multichannel L-filter attains an almost identical performance to that of the multichannel DWMTM filter. Recall, however, that the multichannel DWMTM filter employs a window of dimensions 5×5.

2. The adaptive multichannel L-filters outperform the adaptive multichannel linear filters by approximately 1.6 dB in the case of MC NLMS algorithm and by 2.1 dB in the case of MC LMSN algorithm.

3. The multichannel techniques are found superior to single-channel ones by 1.668 dB in MC NLMS adaptive L-filters and by 1.086 dB in MC LMSN adaptive L-filters.

4. The multichannel techniques yield a much better NR index than single-channel ones in linear

5. MC LMSN adaptive algorithms give identical results to MC NLMS ones with respect to NR index. However, the LMSN adaptive multichannel L-filter without MC yields an almost 2 dB higher NR than the NLMS adaptive multichannel L-filter without MC.

Figure 4a shows the noise-free 50th frame of color image sequence "Trevor White" in RGB color space that is used as a test image. The 48th frame of Trevor White that is used as a reference image is shown in Figure 4b. The test image corrupted by mixed impulsive and additive trivariate contaminated Gaussian noise is depicted in Figure 4c. The output of the MC NLMS and the MC LMSN adaptive multichannel L-filter is shown in Figure 4d and 4e, respectively. We argue that MC NLMS gives a slightly superior filtered image than the MC LMSN in terms of the visual quality. For comparison purposes, the output of the marginal median filter is depicted in Figure 4f. The marginal median filter preserves the edges but fails to remove the noise in the homogeneous regions. On the contrary, in homogeneous regions, the MC NLMS/MC LMSN multichannel L-filter performs better than the marginal median filter. The output of the multichannel DWMTM filter and the vector median filter based on the L_1 norm are given in Figure 4g and 4h, respectively. We see that the output of the vector median filter is clearly poorer than the output of the proposed adaptive multichannel L-filters. The image filtered by the multichannel DWMTM filter exhibits a visual quality that is ranked between that of the MC NLMS and the MC LMSN adaptive multichannel L-filters. The performance of the adaptive location-invariant L-filter is studied separately later.

The performance of the several filters included in the comparisons has been measured in $U^*V^*W^*$ color space as well. As before, the adaptation step-size parameter that yields the best result in terms of the visual quality of the filtered image has been used in NLMS algorithms. It can be found in the corresponding entry of Table 8. For LMSN algorithms, one may use the parameters $\delta = 0.01$ and $\zeta = 0.001$ and either the fixed step size $\mu = 5 \times 10^{-4}$ or a variable step-size $\mu(k)$ given by $\mu(k) = \frac{\mu_0}{\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{X}}(k)}$ as in algorithm II [26]. In the latter case, the parameter μ_0 used is given in the corresponding entry of Table 8. The inspection of Table 8 reveals that:

2. They outperform the multichannel linear techniques by 0.5 dB in the case of the MC NLMS algorithm and by 1.27 dB in the case of the MC LMSN algorithm.

3. The MC NLMS adaptive multichannel L-filter is superior to the MC NLMS single-channel L-filters by 0.9 dB. The MC LMSN adaptive multichannel L-filter is slightly better than using three separate LMSN adaptive single-channel L-filters by 0.682 dB.

4. Multichannel techniques yield a much higher NR than the single-channel techniques in linear filtering.

5. Three out of the four best filtering techniques are found to be multichannel and nonlinear. It is surprising that the MC multichannel normalized TDLMS linear filtering has proven efficient in $U^*V^*W^*$ color space.

The output of several filters included in our comparative study was transformed from $U^*V^*W^*$ to RGB for display purposes. The output of the MC NLMS and MC LMSN adaptive multichannel L-filter is shown in Figure 5a and 5b, respectively. By comparing Figures 5a and 5b to Figures 4d, and 4e, we conclude that by processing the color image in a uniform color space, color information is preserved better than by processing the image in RGB color space. This is self-evident in the uniform greenish background of the image. Moreover, the inability of the marginal median to smooth the noise in homogeneous regions is clearly depicted in Figure 5c. It is remarkable that although the multichannel DWMTM attains a high NR close enough to the NR achieved by the proposed adaptive multichannel L-filters, the visual quality of its output (Figure 5d) is worse than that of the MC NLMS (Figure 5a) and the MC LMSN adaptive multichannel L-filters (Figure 5b). The poor performance of the vector median filter is easily identified in Figure 5e. Figure 5f depicts the output of the MC multichannel normalized TDLMS filter. By comparing the visual quality of Figure 5f to that of Figure 5a, we find the former to be slightly poorer than that of the latter one. Note that the NR for the MC normalized TDLMS filter drops faster than that of the NLMS multichannel L-filter when the 45th frame is chosen as the reference image (i.e., -13.624 dB against

Therefore, for a color image corrupted by mixed additive white trivariate contaminated Gaussian plus impulsive noise, very good performance in terms of NR was obtained by employing MC adaptive multichannel L-filters. Furthermore, the aforementioned measured improvement in terms of the NR index corresponds to a perceived subjective improvement as manifested by the experiments performed in the perceptually uniform color space $U^*V^*W^*$. We found that the MSE in such a space better describes the subjective improvement than the MSE in the RGB space.

Table 9 summarizes the NR achieved by the NLMS and the LMSN adaptive multichannel L-filters in RGB space, when motion compensation is exploited. Our objective is to demonstrate that the results reported previously do not depend strongly on the reference image used. Indeed by employing MC it is seen that the NR decreases only 0.74 dB when the 45th frame is chosen as the reference image instead of the 49th frame for the NLMS in RGB. The corresponding decrease is approximately 0.5 dB when the LMSN algorithm is used in RGB color space. An inspection of Table 9 shows that MC can improve the NR achieved by the NLMS adaptive multichannel L-filter by approximately 3 dB in the case of the 48th frame. However, the gain is smaller for LMSN adaptive multichannel L-filter. In the latter case, an improvement of 0.8 dB was found. The gain for both algorithms increases when the reference image becomes more distant from the filtered image. However, the range of the gain observed is narrower for the LMSN adaptive multichannel L-filter. By comparing the NR indices achieved by the MC LMSN adaptive multichannel L-filter and the MC normalized TDLMS multichannel (linear) filter, we find that the best adaptive multichannel L-filter that employs MC outperforms the best adaptive multichannel linear filter that also uses MC by 1.72 dB. Similarly, we see that the best adaptive multichannel L-filter (i.e., the MC LMSN multichannel L-filter) is superior to the best adaptive single-channel L-filter (i.e., the MC LMSN single-channel L-filter) by 1.086 dB when MC is used. The role of MC in $U^*V^*W^*$ space is marginal in terms of the NR achieved. However, we find it to be crucial in tracking moving edges and preserving edge detail information. The outputs of the proposed MC NLMS and MC LMSN adaptive multichannel

respectively. The filtered images in Figures 6a and 6b were obtained by performing filtering in RGB color space. By comparing these figures with Figure 4d and 4e we cannot identify an observable deterioration in visual quality. Moreover, the visual quality of Figures 6a and 6b is still superior to that of the images filtered by the marginal median (Figure 4f) or the vector median (Figure 4h). By performing adaptive noise filtering in $U^*V^*W^*$ space, we obtained the filtered images shown in Figure 6c and 6d with MC when the 45th frame is used as the reference image for the NLMS and the LMSN algorithm, respectively. A comparison of the visual quality of the Figures 6c and 6d to the one of Figures 5a-b reveals similar information. Therefore, thanks to MC the proposed algorithms still offer a filtered image of higher quality than a broad set of alternatives including the vector median, the marginal median, the α-trimmed mean and the multichannel MTM filters. Moreover, the dependence of the reference image in the derivation of the filter coefficients has been diminished, yielding realistic filter designs.

Finally, we examine the performance of the LMS/LMSN adaptive location-invariant multichannel L-filter in more detail. Table 10 summarizes the NR achieved by these two filter structures in both color spaces with/without MC when either the 48th frame or the 45th one is chosen as the reference image. From an inspection of Table 10, we see that the MC adaptive location-invariant multichannel L-filters are proven the most robust adaptive filter structures because there is almost no deterioration in NR when the reference image changes. This point is valid in both color spaces. Moreover, there is no deterioration in the visual quality of the filtered images produced in either case as we can see by comparing Figures 7a and 7c with Figures 7b and 7d, respectively. The comparison of Figures 7a and 7c with the output of the marginal median filter (Figure 4f) or the output of the vector median filter (Figure 4h) reveals the superiority of the visual quality of the filtered images by the MC adaptive location-invariant multichannel L-filters. Note that such observations extend the good robustness properties of the location-invariant adaptive filter structure described in [10] for the single-channel case.

In this paper three adaptive multichannel L-filters were proposed, namely the LMS, the NLMS and the LMSN adaptive multichannel L-filters. The design of both unconstrained and constrained filters was studied. The performance of both the unconstrained and the constrained proposed adaptive multichannel L-filters has been tested in noise removal in RGB and in $U^*V^*W^*$ color spaces. It was also compared to that of other well-known multichannel nonlinear filters. Adaptive multichannel linear filters and single-channel either nonlinear or linear filters were considered as well. We found by experiments that NLMS and LMSN adaptive multichannel L-filters have the best performance in noise suppression for color images corrupted by mixed impulsive and additive white trivariate contaminated Gaussian noise. Moreover, we found that thanks to MC the dependence of the filter performance on the reference image can be diminished, yielding practical filtering schemes in image sequence processing. We demonstrate that the MC LMS/LMSN adaptive location-invariant multichannel L-filters are the most robust adaptive filter structures in the sense that they achieve practically the same NR index independent of the reference image chosen. These conclusions were also verified by processing color images in a uniform color space (e.g., $U^*V^*W^*$). Therefore, the proposed adaptive multichannel L-filters achieve noise suppression and preserve the color information faithfully.

Acknowledgment

This work was supported by the European Community ESPRIT project (20229) "NOBLESSE".

References

[1] B. Widrow and S.D. Stearns, *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, N.J., 1985.

[2] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englewood Cliffs, N.J., 1986.

rithm", *IEEE Transactions on Circuits and Systems*, vol. 35, no. 5, pp. 485–494, May 1988.

[4] W.B. Mikhael and S.M. Ghosh, "Two-dimensional variable step-size sequential adaptive gradient algorithms with applications", *IEEE Transactions on Circuits and Systems*, vol. 38, no. 12, pp. 1577–1580, December 1991.

[5] J.N. Lin, X. Nie, and R. Unbehauen, "Two-dimensional LMS adaptive filter incorporating a local mean estimator for image processing", *IEEE Transactions on Circuits and Systems, Part II*, vol. 40, no. 7, pp. 417–428, July 1993.

[6] I. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publ., Norwell, MA, 1990.

[7] I. Pitas and A.N. Venetsanopoulos, "Order statistics in digital image processing", *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1893–1921, December 1992.

[8] A.C. Bovik, T.S. Huang, and D.C. Munson, Jr., "A generalization of median filtering using linear combinations of order statistics", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-31, no. 6, pp. 1342–1350, December 1983.

[9] L. Naaman and A.C. Bovik, "Least-squares order statistic filters for signal restoration", *IEEE Transactions on Circuits and Systems*, vol. 38, no. 3, pp. 244–257, March 1991.

[10] C. Kotropoulos and I. Pitas, "Adaptive LMS L-filters for noise suppression in images", *IEEE Transactions on Image Processing*, vol. 5, no. 12, pp. 1596–1609, December 1996.

[11] V. Barnett, "The ordering of multivariate data", *J. R. Statist. Soc. A*, vol. 139, Part 3, pp. 318–354, 1976.

[12] I. Pitas, "Marginal order statistics in color image filtering", *Optical Engineering*, vol. 29, no. 5, pp. 495–503, May 1990.

actions on Circuits and Systems for Video Technology, vol. 1, no. 3, pp. 247–259, September 1991.

[14] C. Kotropoulos and I. Pitas, "Multichannel L-filters based on marginal data ordering", *IEEE Transactions on Signal Processing*, vol. 42, no. 10, pp. 2581–2595, October 1994.

[15] J. Astola, P. Haavisto, and Y. Neuvo, "Vector median filters", *Proceedings of the IEEE*, vol. 78, no. 4, pp. 678–689, April 1990.

[16] R.C. Hardie and G.R. Arce, "Ranking in R^p and its use in multivariate image estimation", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 1, no. 2, pp. 197–209, June 1991.

[17] S.A. Kassam and M. Aburdene, "Multivariate median filters and their extensions", in *Proc. of the IEEE Int. Symp. on Circuits and Systems (ISCAS 91)*, 1991, pp. 85–88.

[18] T. Viero, K. Öistämö, and Y. Neuvo, "Three dimensional median-related filters for color image sequence filtering", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 4, no. 2, pp. 129–142, April 1994.

[19] V. Koivunen, N. Himayat, and S.A. Kassam, "Nonlinear filtering techniques for multivariate images-Design and robustness characterization", *Signal Processing*, vol. 57, no. 1, pp. 81–91, February 1997.

[20] N. Nikolaidis and I. Pitas, "Multichannel L-filters based on reduced ordering", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 5, pp. 470–482, October 1996.

[21] K.N. Plataniotis, D. Androutsos, S. Vinayagamoorthy, and A.N. Venetsanopoulos, "Color image processing using adaptive multichannel filters", *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 933–949, July 1997.

filling curves", *IEEE Transactions on Image Processing*, vol. 6, no. 7, pp. 1025–1037, July 1997.

[23] A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, N.J., 1989.

[24] H. Kong and L. Guan, "A neural network adaptive filter for the removal of impulsive noise in digital images", *Neural Networks*, vol. 9, no. 3, pp. 373–378, 1996.

[25] J.C. Brailean, R.P. Kleihorst, S. Efstratiadis, A.K. Katsaggelos, and R.L. Lagendijk, "Noise reduction for dynamic images sequences: A review", *Proceedings of the IEEE*, vol. 83, no. 9, pp. 1272–1292, September 1995.

[26] P.S.R. Diniz, M.L.R. deCampos, and A. Antoniou, "Analysis of LMS-Newton adaptive filtering algorithms with variable convergence factor", *IEEE Transactions on Signal Processing*, vol. 43, no. 3, pp. 617–627, March 1995.

[27] I. Pitas, "Fast algorithms for running order and max/min calculations", *IEEE Transactions on Circuits and Systems*, vol. 36, no. 6, pp. 795–904, June 1989.

[28] O.L. Frost, III, "An algorithm for linearly constrained adaptive array processing", *Proceedings of the IEEE*, vol. 60, pp. 926–935, 1972.

[29] Y.H. Lee and S.A. Kassam, "Generalized median filtering and related nonlinear filtering techniques", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 3, pp. 672–683, June 1985.

List of Figures

Table 1: LMS-Newton adaptive multichannel L-filter

Initialization:

$\hat{\mathbf{R}}_p^{-1}(0) = \delta^{-1}\mathbf{I}$, where δ is a small real number.

Initial L-filter coefficients: $\hat{\mathbf{a}}_{(i)}(0) = \mathbf{0}$, $i = 1, \ldots, p$.

For $k \geq 1$:

Multichannel L-filter output: $y_i(k) = \hat{\mathbf{a}}_{(i)}^T(k)\tilde{\mathbf{X}}(k)$, $i = 1, \ldots, p$.

A priori estimation error: $e_i(k) = s_i(k) - y_i(k)$, $i = 1, \ldots, p$.

Coefficient updating equations:

$$
\begin{aligned}
\mathbf{t}(k) &= \hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k) \\
\gamma(k) &= \left(\frac{1-\zeta}{\zeta}\right) + \tilde{\mathbf{X}}^T(k)\mathbf{t}(k) \\
\hat{\mathbf{R}}_p^{-1}(k) &= \frac{1}{1-\zeta}\left[\hat{\mathbf{R}}_p^{-1}(k-1) - \frac{\mathbf{t}(k)\mathbf{t}^T(k)}{\gamma(k)}\right] \\
\hat{\mathbf{a}}_{(i)}(k+1) &= \hat{\mathbf{a}}_{(i)}(k) + \mu\hat{\mathbf{R}}_p^{-1}(k)e_i(k)\tilde{\mathbf{X}}(k)
\end{aligned}
$$

Table 2: Computational complexity of the unconstrained adaptive multichannel L-filters.

Algorithm / Parameter	LMS (9)	LMS (16)	NLMS (18)	LMSN (19)
Number of variables	$p^2N + pN + 2p + 1$	$p^2N + 2pN + 2p + 1$	$p^2N + pN + 2p + 2$	$p^2N^2 + p^2N + pN + 2p + 3$
Number of comparisons per iteration	$2p\lfloor \log_2 N \rfloor + 2p$			
Number of multiplications per iteration	$2p^2N + p$	$3p^2N + 3pN$	$2p^2N + pN + +2p$	$4p^2N^2 + 2p^2N + +pN + 2p$
Number of additions per iteration	$2p^2N$	$2p^2N + pN - 1$	$2p^2N + pN - 1$	$3p^2N^2 + 2p^2N - pN + 1$

Table 3: Computational requirements of the location-invariant multichannel L-filters in addition to those tabulated in Table 2.

Algorithm / Parameter	(31)	(39)
Number of additional variables	$p^2\,N^2 + p^2\,N + p\,N$	$p^2\,N^2 + p^2\,N + 2p\,N + p^2$
Number of multiplications upon initialization	$p^2\,N^2 + p\,N$	
Number of additions upon initialization	$p^2\,N^2$	
Number of additional multiplications per iteration	$p^3\,N^2$	$3p^3\,N^2 + 2p^3\,N + \mathcal{O}(p^3)$
Number of additional additions per iteration	$p^3\,N^2$	$2p^3\,N^2 + p^2\,N^2 + 3p^3\,N - 3p^2\,N - p^2 + \mathcal{O}(p^3)$

Table 4: Noise reduction (in dB) achieved by the adaptive multichannel L-filters for the bivariate contaminated Gaussian noise model (Filter length $N = 9$).

Filter	NR
LMS adaptive multichannel L-filter	-18.057
NLMS adaptive multichannel L-filter	-17.721
LMSN adaptive multichannel L-filter	-18.564
non-adaptive multichannel L-filter	-18.564

Table 5: Noise reduction (in dB) achieved by the location-invariant multichannel L-filters for the bivariate contaminated Gaussian noise model (Filter length $N = 9$).

Filter	NR
LMS location-invariant multichannel L-filter	-11.853
LMSN location-invariant multichannel L-filter	-11.998
non-adaptive location-invariant multichannel L-filter	-10.997
marginal median	-9.8209

Table 6: Performance comparisons between multichannel versus single-channel and L-filtering versus linear filtering adaptive algorithms both in RGB and $U^*V^*W^*$ for the mixed additive trivariate contaminated Gaussian plus impulsive noise model (Filter window 3×3).

			NR (dB) in color space	
Approach	Type of filtering	Adaptive algorithm	RGB	$U^*V^*W^*$
multichannel	L-filtering	NLMS	-13.838	-15.180
		LMSN	-14.176	-14.822
	linear filtering	NLMS	-11.408	-14.047
		LMSN	-11.196	-13.278
single-channel	L-filtering	NLMS	-11.738	-14.337
		LMSN	-12.495	-14.412
	linear filtering	NLMS	-8.830	-11.647
		LMSN	-9.166	-12.073

Table 7: Noise reduction (in dB) achieved in (NTSC) RGB color space by several filters in the restoration of the 50th color frame of "Trevor White" corrupted by mixed additive white trivariate contaminated Gaussian plus impulsive noise (Filter window 3×3).

Filter	NR	Ranking
marginal median	-11.748	[7]
vector median L_1	-10.043	[13]
vector median L_2	-8.508	[18]
$\mathcal{R}_E$-filter	-8.696	[16]
α-trimmed mean ($\alpha = 0.2$)	-11.250	[11]
arithmetic mean	-8.514	[17]
multichannel MTM filter	-11.425	[10]
multichannel DWMTM filter	-13.362	[1]
NLMS adaptive multichannel L-filter ($\mu_0 = 0.1$)	-10.246	
MC NLMS adaptive multichannel L-filter ($\mu_0 = 0.5$)	-13.228	[3]
LMSN adaptive multichannel L-filter	-12.529	
MC LMSN adaptive multichannel L-filter	-13.320	[2]
LMS location-invariant multichannel L-filter ($\mu_0 = 0.1$)	-11.376	
MC LMS location-invariant multichannel L-filter ($\mu_0 = 0.2$)	-12.076	[6]
LMSN location-invariant multichannel L-filter	-11.935	
MC LMSN location-invariant multichannel L-filter	-12.403	[4]
normalized TDLMS multichannel filter ($\mu_0 = 0.1$)	-9.271	
MC normalized TDLMS multichannel filter ($\mu_0 = 0.1$)	-11.600	[8]
TDLMS-Newton multichannel filter	-10.762	
MC TDLMS-Newton multichannel filter	-11.169	[12]
NLMS adaptive single-channel L-filters ($\mu_0 = 0.05$)	-9.574	
MC NLMS adaptive single-channel L-filters ($\mu_0 = 0.05$)	-11.560	[9]
LMSN adaptive single-channel L-filters	-12.010	
MC LMSN adaptive single-channel L-filters	-12.234	[5]
normalized TDLMS filters ($\mu_0 = 0.05$)	-7.699	
MC normalized TDLMS filters ($\mu_0 = 0.05$)	-9.264	[15]
TDLMS-Newton filters	-9.080	
MC TDLMS-Newton filters	-9.823	[14]

Table 8: Noise reduction (in dB) achieved in $U^*V^*W^*$ color space by several filters in the restoration of the 50th color frame of "Trevor White" corrupted by mixed additive white trivariate contaminated Gaussian plus impulsive noise (Filter window 3×3).

Filter	NR	Ranking
marginal median	-11.270	[12]
vector median L_1	-9.775	[15]
vector median L_2 ($\mathcal{R}_E$-filter)	-8.553	[17]
α-trimmed mean ($\alpha = 0.2$)	-10.904	[13]
arithmetic mean	-9.212	[16]
multichannel MTM filter	-11.291	[11]
multichannel DWMTM filter	-13.920	[4]
NLMS adaptive multichannel L-filter ($\mu_0 = 0.1$)	-14.121	
MC NLMS adaptive multichannel L-filter ($\mu_0 = 0.5$)	-14.412	[1]
LMSN adaptive multichannel L-filter ($\mu_0 = 0.1$)	-13.992	
MC LMSN adaptive multichannel L-filter (fixed $\mu = 0.001$)	-14.159	[2]
LMS location-invariant multichannel L-filter ($\mu_0 = 0.1$)	-12.033	
MC LMS location-invariant multichannel L-filter ($\mu_0 = 0.1$)	-12.020	[9]
LMSN location-invariant multichannel L-filter	-12.265	
MC LMSN location-invariant multichannel L-filter	-12.183	[8]
normalized TDLMS multichannel filter ($\mu_0 = 0.5$)	-13.725	
MC normalized TDLMS multichannel filter ($\mu_0 = 0.5$)	-13.922	[3]
TDLMS-Newton multichannel filter ($\mu_0 = 0.1$)	-12.701	
MC TDLMS-Newton multichannel filter ($\mu_0 = 0.1$)	-12.889	[7]
NLMS adaptive single-channel L-filters ($\mu_0 = 0.5$)	-12.902	
MC NLMS adaptive single-channel L-filters ($\mu_0 = 0.5$)	-13.513	[5]
LMSN adaptive single-channel L-filters ($\mu_0 = 0.1$)	-13.407	
MC LMSN adaptive single-channel L-filters (fixed $\mu = 0.001$)	-13.477	[6]
normalized TDLMS filters ($\mu_0 = 0.5$)	-10.002	
MC normalized TDLMS filters ($\mu_0 = 0.5$)	-10.668	[14]
TDLMS-Newton filters ($\mu_0 = 0.1$)	-11.418	
MC TDLMS-Newton filters (fixed $\mu = 0.001$)	-11.501	[8]

Table 9: Noise reduction (in dB) achieved by the adaptive multichannel L-filters in the restoration of the 50th color frame of "Trevor White" corrupted by mixed additive white trivariate contaminated Gaussian plus impulsive noise for several past frames acting as reference images with/without motion compensation (MC).

	NR (dB) for NLMS multichannel L-filter		NR (dB) for LMSN multichannel L-filter	
Reference Image	with MC	without MC	with MC	without MC
49th frame	-13.552	-12.223	-13.613	-13.326
48th frame	-13.228	-10.264	-13.320	-12.529
47th frame	-13.066	-9.475	-13.248	-11.577
46th frame	-13.201	-8.478	-13.178	-10.870
45th frame	-12.814	-7.995	-13.115	-9.684

Table 10: Noise reduction (in dB) achieved by the adaptive location-invariant multichannel L-filters in the restoration of the 50th color frame of "Trevor White" corrupted by mixed additive white trivariate contaminated Gaussian plus impulsive noise with/without motion compensation (MC) in color spaces RGB and $U^*V^*W^*$. Either the 48th frame or the 45th one is used as the reference image.

	NR (dB) in RGB		NR (dB) in $U^*V^*W^*$	
Reference Image	48	45	48	45
Filter				
LMS location-invariant	-11.376	-10.220	-12.033	-11.028
MC LMS location-invariant	-12.076	-12.052	-12.020	-11.926
LMSN location-invariant	-11.935	-10.783	-12.265	-12.078
MC LMSN location-invariant	-12.403	-12.335	-12.183	-12.134

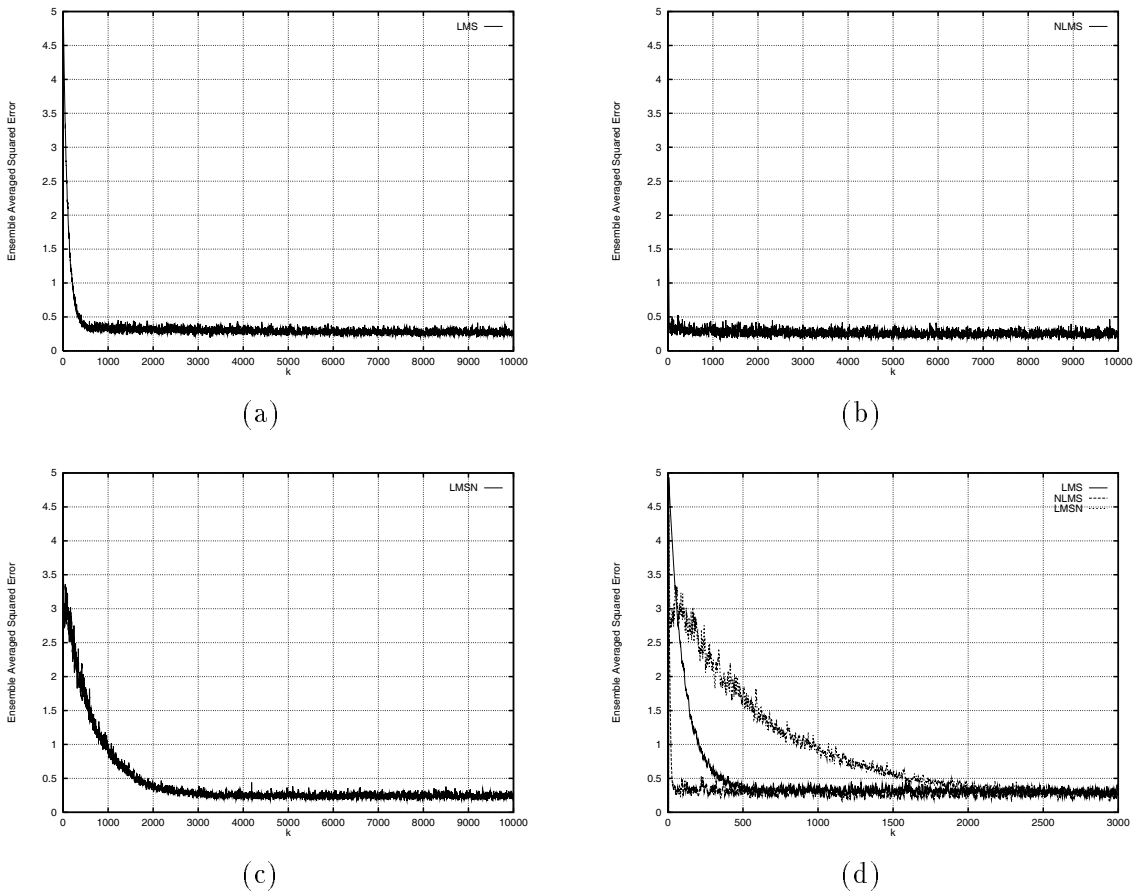
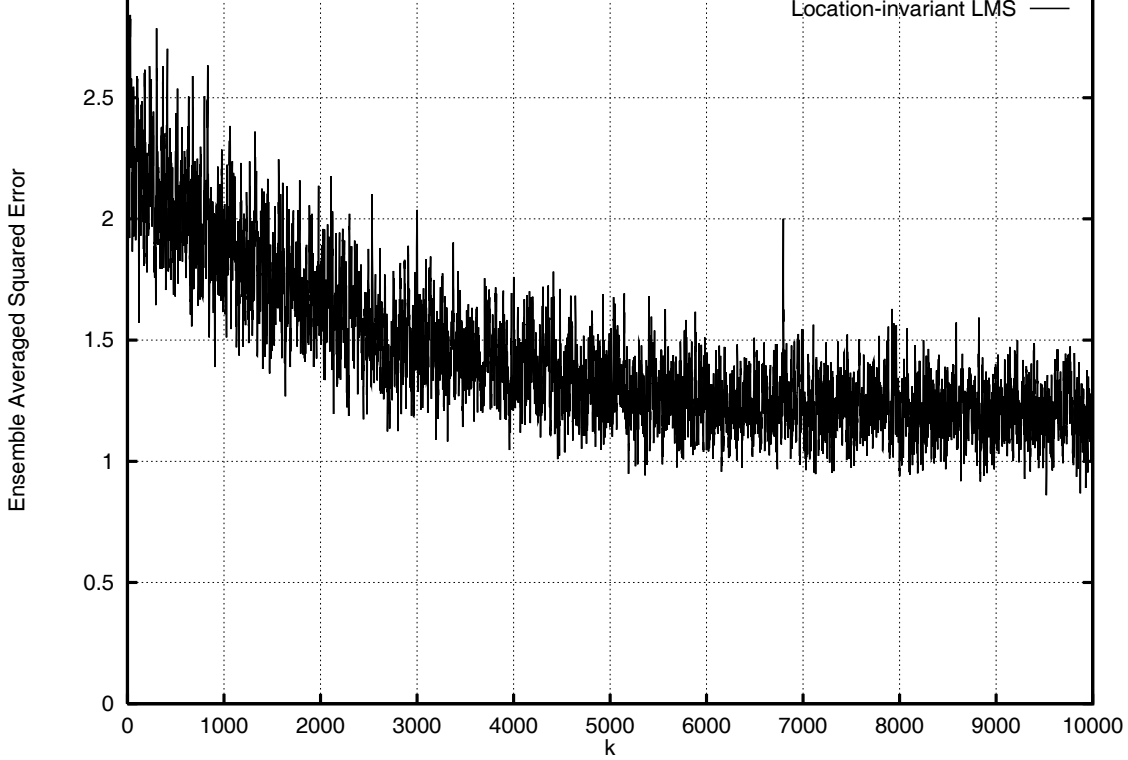
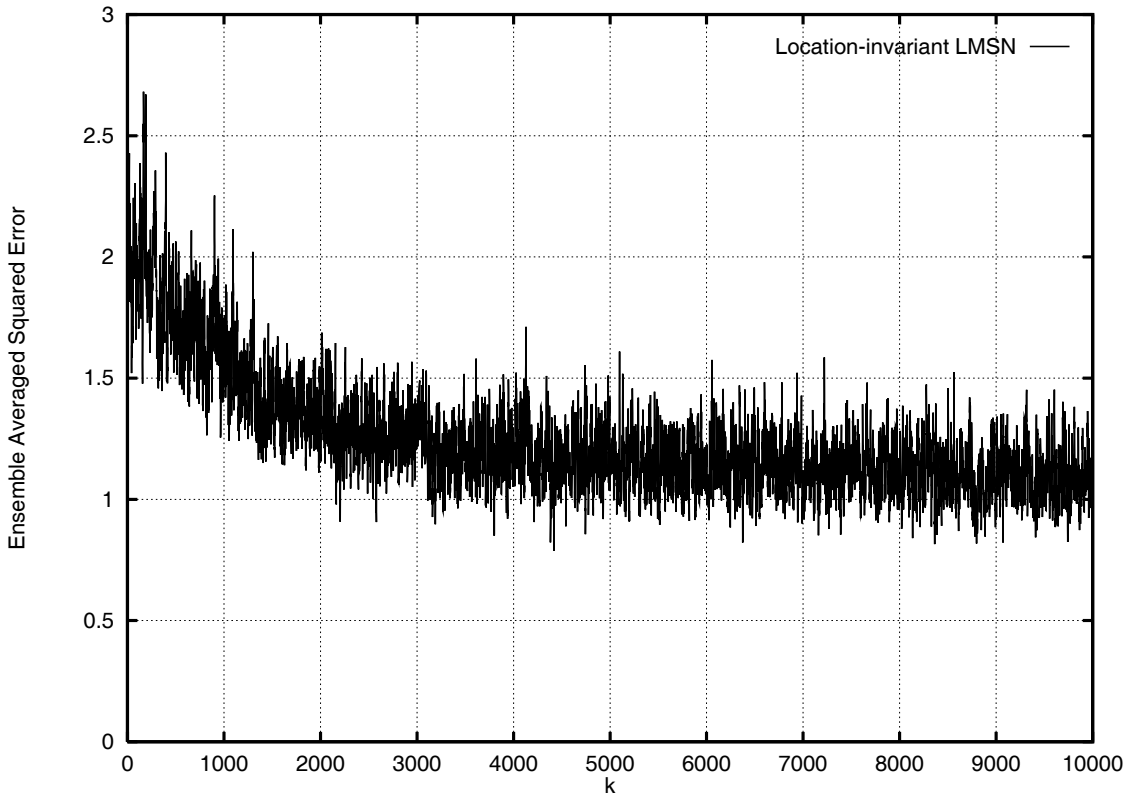


Figure 1: Learning curves for (a) LMS, (b) NLMS, and (c) LMSN adaptive multichannel L-filters. (d) Zooming in the first 3000 iterations for the three adaptive multichannel L-filters.



(a)



(b)

Figure 2: Learning curves of the (a) LMS adaptive location-invariant multichannel L-filter; (b) LMSN adaptive location-invariant multichannel L-filter.

(a) (b)

Figure 3: Displacement vector fields (a) between the 50th and the 48th noise-free frames of image sequence "Trevor White"; (b) between the 50th and the 45th noise-free frames of the same image sequence.

(a) (b)

(c)

(d)

(e)

(f)

(g)

(h)

Figure 4: Output of various filters in smoothing the mixed impulsive and additive white trivariate contaminated Gaussian noise that corrupts the 50-th frame of color image sequence "Trevor White". Image filtering has been performed in RGB color space.

(a)

(b)

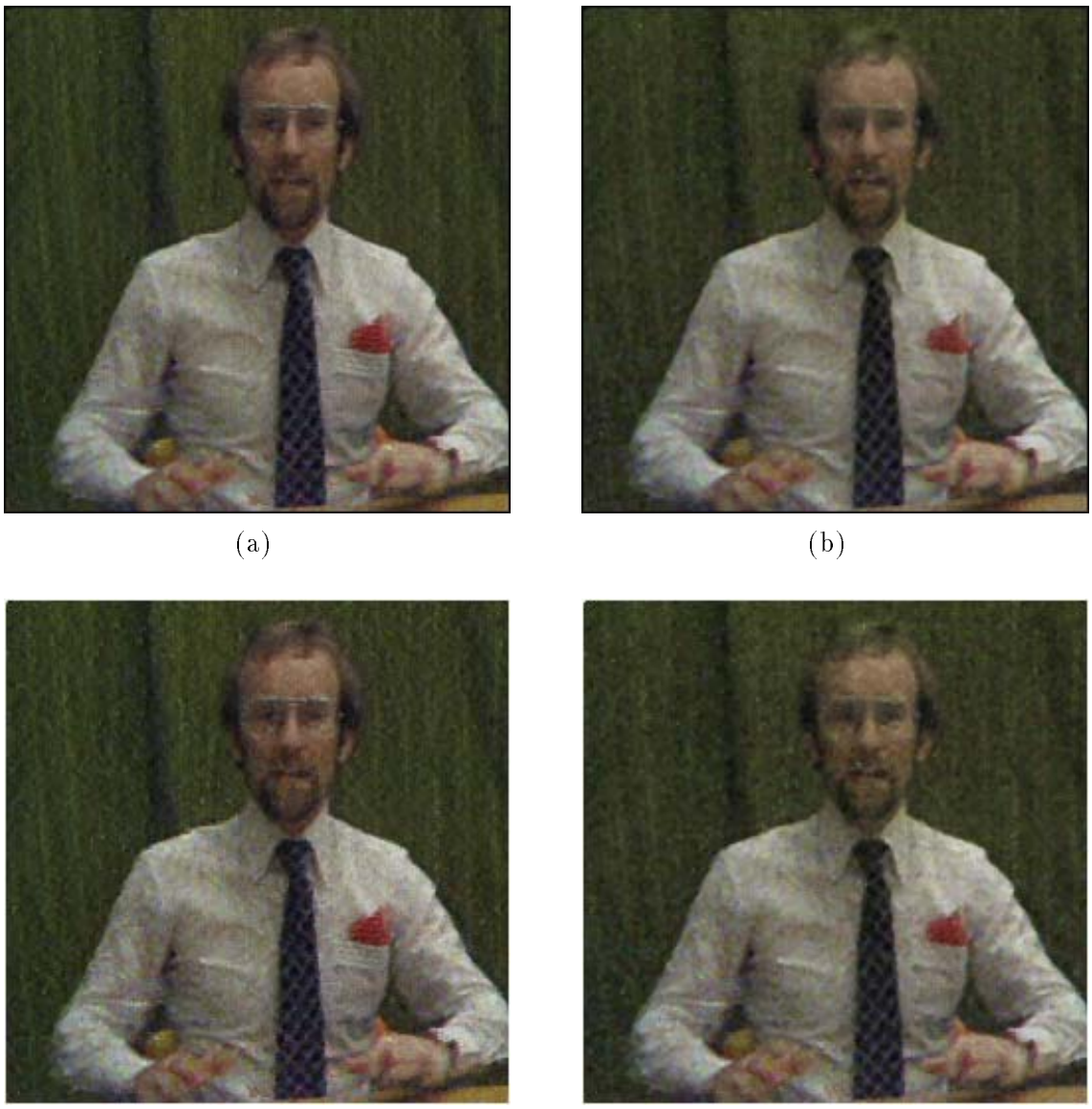
(c)

(d)

(e)

(f)

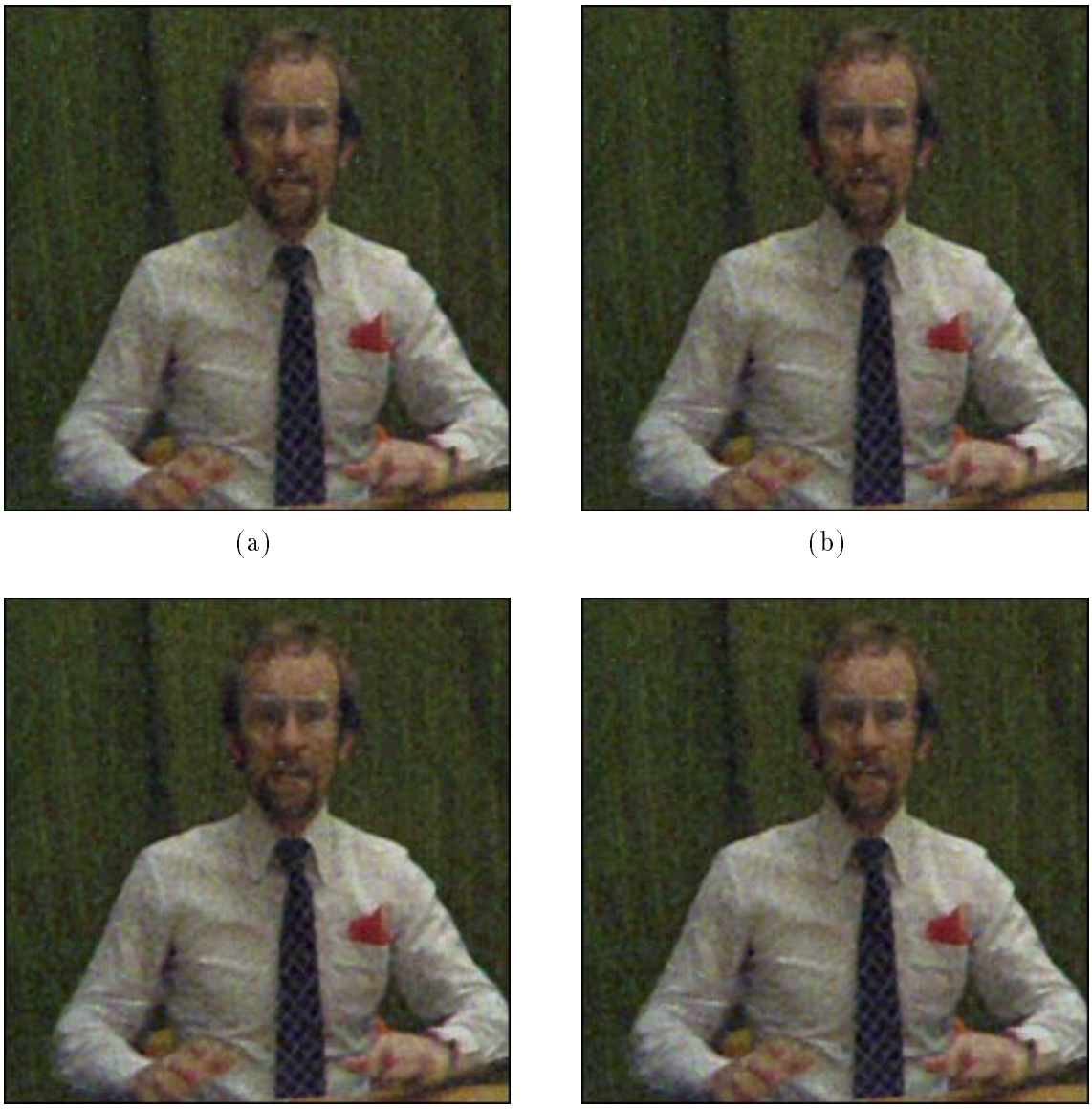
Figure 5: Output of various filters in smoothing the mixed impulsive and additive white trivariate contaminated Gaussian noise that corrupts the 50-th frame of color image sequence "Trevor White". Image filtering has been performed in $U^*V^*W^*$ color space.



(a) (b)

(c) (d)

Figure 6: Output of the MC NLMS/LMSN adaptive multichannel L-filters when the 45th frame of color image sequence "Trevor White" is chosen as the reference image.



(a) (b)

(c) (d)

Figure 7: Output of the MC LMS/LMSN adaptive location-invariant multichannel L-filters when either the 48th or the 45th frame of color image sequence "Trevor White" is used as the reference image. Filtering is performed in RGB.