MULTICHANNEL ADAPTIVE L-FILTERS IN COLOR IMAGE FILTERING

Constantine Kotropoulos, Ioannis Pitas and Maria Gabrani

Department of Informatics, Aristotle University of Thessaloniki
Thessaloniki 540 06, Greece
E-mail: {costas, pitas}@zeus.csd.auth.gr

ABSTRACT

Three novel adaptive multichannel L-filters based on marginal data ordering are proposed. They rely on well-known algorithms for the unconstrained minimization of the Mean Squared Error (MSE), namely, the Least Mean Squares (LMS), the normalized LMS (NLMS) and the LMS-Newton (LMSN) algorithm. Performance comparisons in color image filtering have been made both in 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.

1. INTRODUCTION

Adaptive signal processing has exhibited a tremendous growth in the two past decades. Adaptive filters have been applied in a wide variety of problems including system identification, channel equalization, echo cancellation in telephone channels [1]. All the above-mentioned problems involve one-dimensional (1-D) signals and 1-D linear filters. However, linear filters are not suitable for applications where the noise is impulsive, e.g. in image processing. In the later case, a multitude of nonlinear techniques has been proved a successful alternative to the linear techniques [2]. One of the best known nonlinear filter classes is based on the *order statistics*. It uses the concept of data ordering. There is now a plethora of nonlinear filters based on data ordering. Among them are the L-filters whose output is defined as a linear combination of the order statistics [3]. A design of L-filters which relies on a non-iterative minimization of the MSE yields very tedious expressions for computing the marginal and joint cumulative functions of the order statistics (cf. [4]). On the contrary, adaptive L-filters are proved to be appealing because they avoid the computational burden of the non-iterative methods [5].

Recently, increasing attention has been given to nonlinear processing of vector-valued signals [6, 7, 8, 9]. The major difficulty in the definition of multichannel filters is the lack of an unambiguous and universally accepted definition of ordering for multivariate data [10]. Filters such as those proposed in [8, 9] are based on marginal ordering whereas other filters are based on reduced ordering [6, 7].

The main contribution of this paper is in the design of adaptive multichannel L-filters based on marginal data ordering using the MSE as fidelity criterion as well as in the assessment of their performance in color image filtering. Three novel adaptive multichannel L-filters are proposed in this paper that are based on well-known algorithms for the iterative unconstrained minimization of the MSE, namely, the Least Mean Squares, the normalized LMS (NLMS) and the LMS-Newton (LMSN) algorithm. The performance of the adaptive multichannel L-filters under study in color image filtering is compared to the one of other well-known multichannel nonlinear filters and of adaptive single-channel L-filters as well. The comparative study is conducted both in the RGB and in $U^*V^*W^*$ color spaces.

This work was partially supported by ESPRIT Project 20229 NOBLESSE.

2. PROBLEM STATEMENT

Let $\mathbf{x}_1, \ldots, \mathbf{x}_N$ be a random sample of N observations of a p-dimensional random variable $\mathbf{X}$. The marginal ordering scheme orders the vector components independently, thus yielding:

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

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 [9]:

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

where $\mathbf{A}_i$ is a $(p \times N)$ coefficient matrix. Let $\mathbf{a}_{il}^T$, $l = 1, \ldots, p$ denote the l-th row of matrix $\mathbf{A}_i$ and $\tilde{\mathbf{x}}_i(k) = \left(x_{i(1)}(k), \ldots, x_{i(N)}(k)\right)^T$, $i = 1, \ldots, p$ be the $(N \times 1)$ vector of the order statistics along the i-th channel. Let us also 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 noise vector sequence $\{\mathbf{n}(k)\}$ of zero mean vector having the same dimensionality, i.e., $\mathbf{x}(k) = \mathbf{s}(k) + \mathbf{n}(k)$. The noise vector components are assumed to be uncorrelated in the general case. In addition, we assume that the noise vectors at different values of index k are independent 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 $\mathbf{A}_i$, $i = 1, \ldots, p$ which minimize the MSE between the filter output $\mathbf{y}(k)$ and the noise-free signal $\mathbf{s}(k)$. Following similar reasoning as in [9], but without invoking the assumption of a constant signal $\mathbf{s}$, it can be shown that the MSE is expressed as:

$$\varepsilon(k) = \sum_{i=1}^{p} \{\mathbf{a}_{(i)}^T \tilde{\mathbf{R}}_p \mathbf{a}_{(i)} - 2\mathbf{a}_{(i)}^T \tilde{\mathbf{q}}_{(i)}(k)\} + \mathrm{E}\left[\mathbf{s}^T(k)\mathbf{s}(k)\right] \quad (3)$$

where $\mathbf{a}_{(i)} = \left(\mathbf{a}_{1i}^T \mid \mathbf{a}_{2i}^T \mid \cdots \mid \mathbf{a}_{pi}^T\right)^T$. Moreover, $\tilde{\mathbf{R}}_p(k) = \mathrm{E}\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\right]$ and $\tilde{\mathbf{q}}_{(i)}(k) = \mathrm{E}\left[s_i(k)\tilde{\mathbf{X}}^T(k)\right]$ with $\tilde{\mathbf{X}}(k) = (\tilde{\mathbf{x}}_1^T(k) \mid \tilde{\mathbf{x}}_2^T(k) \mid \ldots \mid \tilde{\mathbf{x}}_p^T(k))^T$.

Minimizing (3) over $\mathbf{a}_{(i)}$ is a quadratic minimization problem that has a unique solution under the condition that $\tilde{\mathbf{R}}_p(k)$ is positive definite. It is easily deduced that the minimum MSE coefficient vector is:

$$\mathbf{a}_{(i)}^o(k) = \tilde{\mathbf{R}}_p^{-1}(k)\tilde{\mathbf{q}}_{(i)}(k). \quad (4)$$

It is seen that (4) yields explicitly determined filter coefficients provided that we are able to calculate the moments of the order statistics from univariate populations that appear in $\mathbf{R}_{ii}(k)$ as well as the product moments of the order statistics from bivariate populations that appear in $\mathbf{R}_{ij}(k)$, $i \neq j$ and $i = 1, \ldots, p$. This is fairly easy for i.i.d. input variates, i.e., in the case of a constant signal $\mathbf{s}(k) = \mathbf{s}$ as has been demonstrated in [9]. Even for independent, non-identically distributed input variates the framework tends to become very complicated (cf. [9]). The difficulties are increased in color image processing, where the observations $\tilde{\mathbf{X}}(k)$ and the desired signal $\mathbf{s}(k)$ are strongly nonstationary. In order to overcome this obstacle, we shall resort on iterative algorithms for the minimization of $\varepsilon(k)$ in (3).

3. UNCONSTRAINED MINIMIZATION OF THE MEAN SQUARED ERROR

In this section, three adaptive multichannel L-filters are derived that iteratively minimize the MSE (3) without imposing any constraints on the filter coefficients. These algorithms are: (i) the LMS, (ii) the NLMS, and (iii) the LMSN. The rationale underlying the choice of each algorithm is stated explicitly.

The filter coefficient vectors $\mathbf{a}_{(i)}$, $i = 1, \ldots, p$ that minimize the MSE (3) can be computed recursively using the steepest descent algorithm as follows:

$$\mathbf{a}_{(i)}(k+1) = \mathbf{a}_{(i)}(k) + \mu\left[\mathbf{q}_{(i)}(k) - \tilde{\mathbf{R}}_p(k)\mathbf{a}_{(i)}(k)\right]. \quad (5)$$

Using $\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)$ and $s_i(k)\tilde{\mathbf{X}}(k)$ as instantaneous estimates of $\tilde{\mathbf{R}}_p(k)$ and $\mathbf{q}_{(i)}(k)$ respectively the LMS adaptive multichannel L-filter is obtained, i.e.:

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \mu\left[s_i(k) - \tilde{\mathbf{X}}^T(k)\hat{\mathbf{a}}_{(i)}(k)\right]\tilde{\mathbf{X}}(k) \quad (6)$$

where the bracketed term in (6) is the a priori estimation error $e_i(k)$ between the i-th component of the desired signal $s_i(k)$ and the filter output $y_i(k) = \tilde{\mathbf{X}}^T(k)\hat{\mathbf{a}}_{(i)}$. It is seen 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. (6) employs the composite vector of the ordered observations $\tilde{\mathbf{X}}(k)$. On the contrary, the ordinary adaptive multichannel 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 $\tilde{\mathbf{R}}_p(k)$.

Let $\mathbf{M}_i$ be a diagonal matrix of dimensions $(pN \times pN)$ associated with the updating equation for the coefficient vector $\mathbf{a}_{(i)}(k)$. 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.:

$$e_i(k+1) = e_i(k) + \sum_{j=1}^{pN} \frac{\partial e_i(k)}{\partial a_{(i)j}} \Delta a_{(i)j} + \sum_{j=1}^{pN}\sum_{l=1}^{pN} \frac{\partial^2 e_i(k)}{\partial a_{(i)j}\,\partial a_{(i)l}} \Delta a_{(i)j}\,\Delta a_{(i)l} + \mathcal{O}(3) \quad (7)$$

where $a_{(i)j}$ denotes the j-th element of $\mathbf{a}_{(i)}$. In (7), $\Delta a_{(i)j} = a_{(i)j}(k+1) - a_{(i)j}(k)$ and $\mathcal{O}(3)$ are the higher order terms. Due to the definition of filter output (2) the second and higher order derivatives in (7) are zero. Accordingly, the step-size sequence $\mu_{i,\,jj}$ that minimizes $e_i^2(k+1)$ satisfies the equation:

$$\sum_{j=1}^{pN} \mu_{i,\,jj}(k)\tilde{X}_j^2(k) = 1 \quad (8)$$

provided that $e_i(k) \neq 0$ and $\tilde{X}_j(k) \neq 0$ during the adaptation. If the adaptation step weighs more the filter coefficients that have larger gradients than those having smaller gradients, i.e., $\mu_{i,\,jj} = \beta|\frac{\partial e_i(k)}{\partial a_{(i)j}(k)}|$, the following optimal step-size sequence is obtained:

$$\mu_{i,\,jj}^* = \frac{|X_j(k)|}{\sum_{j=1}^{pN}|X_j(k)|^3}. \quad (9)$$

The normalized LMS (NLMS) algorithm provides a way to automate the choice of the adaptation step-size parameter in order to speed up the convergence of the algorithm. Its design is based on a quite limited knowledge of the input-signal statistics and it is able to track the varying signal statistics. 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 (8) we obtain $\mu(k) = 1/(\tilde{\mathbf{X}}^T(k)\tilde{\mathbf{X}}(k))$. Then, the substitution of $\mu(k)$ into (6) yields the updating equations for the coefficients of the normalized LMS adaptive multichannel L filter, i.e.:

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \frac{\mu_0}{\tilde{\mathbf{X}}^T(k)\tilde{\mathbf{X}}(k)} e_i(k)\tilde{\mathbf{X}}(k), \quad i = 1, \ldots, p \quad (10)$$

where $\mu_0 \in (0, 1]$ is a parameter that is introduced for additional control. The composite vector of the ordered observations $\tilde{\mathbf{X}}(k)$ is employed instead of the vector of input observations in (10).

It is well-known that the eigenvalue spread of the composite correlation matrix $\tilde{\mathbf{R}}_p(k)$ is large in principle. In such a case, LMS-Newton (LMSN) algorithm is a powerful alternative to LMS [12]. The LMSN algorithm employs computationally efficient estimates for the autocorrelation

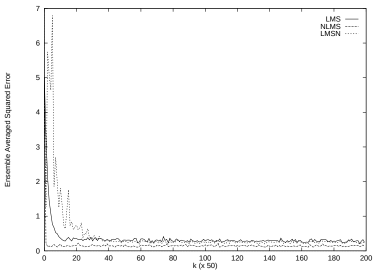

Figure 1: Learning curves for the multichannel adaptive L-filters under study.

Table 1: 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	-17.616
LMSN adaptive multichannel L-filter	-18.661
nonadaptive multichannel L-filter	-18.564

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:

$$\hat{\mathbf{a}}_{(i)}(k+1) = \hat{\mathbf{a}}_{(i)}(k) + \mu \tilde{\mathbf{R}}_p^{-1}(k) e_i(k) \tilde{\mathbf{X}}(k), \quad i = 1, \ldots, p. \tag{11}$$

An estimate of the composite matrix $\tilde{\mathbf{R}}_p(k)$ can be calculated by using the Robbins-Monro procedure which solves the equation $\mathrm{E}\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) - \tilde{\mathbf{R}}_p(k)\right] = \mathbf{0}$. The solution of this equation is given by:

$$\hat{\mathbf{R}}_p(k) = \hat{\mathbf{R}}_p(k-1) + \zeta\left[\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k) - \hat{\mathbf{R}}_p(k-1)\right]. \tag{12}$$

By using the matrix inversion lemma, we obtain:

$$\hat{\mathbf{R}}_p^{-1}(k) = \frac{1}{1-\zeta}\left\{\hat{\mathbf{R}}_p^{-1}(k-1) - \frac{\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)}{\left(\frac{1-\zeta}{\zeta}\right)\tilde{\mathbf{X}}^T(k)\hat{\mathbf{R}}_p^{-1}(k-1)\tilde{\mathbf{X}}(k)}\right\} \tag{13}$$

4. 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.

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 have been derived in [9]. A vector valued signal $\mathbf{s} = (1.0, 2.0)^T$ corrupted by additive white bivariate noise $\mathbf{n}(k)$ with probability density function (pdf) identical to the one employed in [9] has been used as a test signal. An approximation of the ensemble-averaged learning curve for each adaptive algorithm has been obtained following the procedure described in [1]. The learning curves are plotted in Figure 1. The filter length N has been 9 in all cases. In the plots of Figure 1 points every 50 time instants have been used. It is seen that NLMS adaptive multichannel L-filter attains the fastest convergence rate. Subsequently, the noise reduction index (NR) defined as the ratio of the output noise power to the input noise power, i.e.:

$$\mathrm{NR} = 10\log\frac{\sum_k(\mathbf{y}(k)-\mathbf{s}(k))^T(\mathbf{y}(k)-\mathbf{s}(k))}{\sum_k(\mathbf{x}(k)-\mathbf{s}(k))^T(\mathbf{x}(k)-\mathbf{s}(k))}. \tag{14}$$

is measured and is compared to the one achieved by the nonadaptive multichannel L-filter. The estimates of the multichannel L-filter coefficients have been 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. The results are tabulated in Table 1. To facilitate the comparisons, the NR index achieved by the nonadaptive multichannel L-filter designed in [9] is also given. By comparing the NR indices tabulated in Table 1, we conclude that: (i) All algorithms converge towards the optimal solution. (ii) The LMSN adaptive multichannel L-filter approaches better the NR achieved by the nonadaptive design. (iii) The LMS algorithm is the second best. (iv) Although, NLMS attains the fastest convergence rate, it is seen that its NR is approximately 1 dB less than the NR achieved by the nonadaptive design.

The second set of experiments deals with color images, i.e., three-channel two-dimensional signals. Let us consider the 50th frame of color image sequence "Trevor White". The original noise-free image is corrupted by additive white trivariate contaminated Gaussian noise having the probability distribution:

$$(1-\varrho)\mathcal{N}(\mathbf{0}\,;\mathbf{C}_1) + \varrho\mathcal{N}(\mathbf{0}\,;\mathbf{C}_2) \tag{15}$$

for $\varrho = 0.1$ plus impulsive noise such that 6 % of the image pixels in each primary color component are replaced by impulses of value 0 or 255 (i.e., positive and negative impulses). In (15), $\mathbf{C}_i$, $i = 1, 2$ denotes the covariance matrix of each trivariate joint Gaussian distribution. The following covariance matrices have been used:

$$\mathbf{C}_1 = \begin{bmatrix} 100 & 100 & 210 \\ 100 & 400 & 180 \\ 210 & 180 & 900 \end{bmatrix} \quad \mathbf{C}_2 = \begin{bmatrix} 900 & -300 & -210 \\ -300 & 400 & 60 \\ -210 & 60 & 100 \end{bmatrix}. \tag{16}$$

A point that needs 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 [11]. 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 $U^*V^*W^*$ [11]. In order 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 have chosen the $U^*V^*W^*$ space for this purpose. In all experiments, the 48th color image frame of "Trevor White" is used as a reference image for the adaptive filtering techniques. The NR achieved by the filters under

(a) (b) (c)

Figure 2: (a) Noisy 50th frame of "Trevor White". (b) Output of the 3 ×3 marginal median filter. (c) Output of the 3 × 3 LMSN multichannel adaptive L-filter. Filtering is performed in $U^*V^*W^*$.

study in both color spaces is given in Table 2. It is seen that the LMSN is ranked as the best filtering technique in $U^*V^*W^*$ and as the second best filtering technique in RGB. Figure 2a shows the noisy corrupted 50th frame of color image sequence "Trevor White" in RGB color space. The output of the 3 × 3 marginal median filter is shown in Figure 2b for comparison purposes. The output of the LMSN filter of the same dimensions is shown in Figure 2c.

Table 2: Noise reduction (in dB) achieved in RGB and $U^*V^*W^*$ color spaces by several filters in the restoration of the noisy 50th color frame of "Trevor White". (Filter window 3×3).

Filter	NR (dB) in color space	
	RGB	$U^*V^*W^*$
marginal median	-11.750	-11.200
vector median L_1	-10.006	-9.700
vector median L_2	-8.528	-8.550
$\mathcal{R}_E$-filter	-8.690	-8.850
α-trimmed mean ($\alpha = 0.2$)	-11.260	-10.820
arithmetic mean	-8.510	-9.156
multichannel MTM filter	-11.440	-11.248
multichannel DWMTM filter	-13.358	-13.849
NLMS multichannel L-filter	-11.245	-14.310
LMSN multichannel L-filter	-12.428	-14.490
NLMS single-channel L-filters	-9.687	-13.527
LMSN single-channel L-filters	-11.980	-14.225

5. REFERENCES

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

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

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

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

[5] C. Kotropoulos and I. Pitas, "Adaptive LMS L-filters for noise suppression in images," *IEEE Trans. on Image Processing*, accepted for publication, 1995.

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

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

[8] I. Pitas and P. Tsakalides, "Multivariate ordering in color image restoration," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 1, no. 3, pp. 247–259, September 1991.

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

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

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

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