Optimal Multichannel L Filters Based on Reduced Ordering

N. Nikolaidis I. Pitas

Department of Informatics
University of Thessaloniki, Thessaloniki 540 06, GREECE
fax: +3031-996304 e-mail: pitas@zeus.csd.auth.gr

ABSTRACT — **In this paper we propose a new family of multichannel L filters that are based on the R-ordering principle and provide expressions for the optimal filter coefficients. Furthermore, we calculate the moments of the R-ordered vectors that are involved in these expressions. Experiments with noisy two-channel vector fields and color images are presented in order to demonstrate the superiority of the proposed filters over other multichannel filters.**

1 Introduction

Multichannel signals appear in a variety of signal processing applications including multispectral satellite imaging, multiple antenna transmission systems, color image processing, motion vector field processing e.t.c. The most straightforward way to process multichannel signals is to treat each channel independently using single-channel techniques. However this approach does not exploit the correlation between channels. Multichannel techniques that take into account this correlation have been reported to be more effective [1],[2].

Nonlinear signal processing techniques based on data ordering have been very successful in the single-channel case [3]. The single-channel L filters [3] whose output is a linear combination of the ordered input samples belong to this category. Unfortunately the extension of L filters to the multichannel case is not straightforward due to the fact that no globally agreeable ordering scheme exists for multichannel observations. Several sub-ordering principles (marginal, conditional, partial, reduced ordering) have been proposed in the literature [3]. In this paper we present a new family of multichannel L filters that are based on the reduced ordering scheme. The coefficients of the proposed filters can be optimized for a specific noise distribution with respect to the mean square error between the filter output and the desired, noise-free signal, provided that the latter is constant within the filter window.

2 Reduced ordering

Reduced ordering, also known as R-ordering, orders vector-valued observations $\mathbf{x}_1, \ldots, \mathbf{x}_N$, $\mathbf{x}_i = [x_{i1}, \ldots, x_{ip}]^T$ according to their distance d_i from some reference vector $\overline{\mathbf{a}}$. As a consequence, multivariate ordering is reduced to one-dimensional ordering. Reduced ordering is rather easy to implement and is the ordering principle that is the most "natural" for vector-valued observations. The i-th order statistic according to R-ordering will be denoted $\mathbf{x}_{[i]} = [x_{[i]1}, \ldots, x_{[i]p}]^T$ and is the vector whose distance d_i from the reference vector is ranked i-th among distances $d_1 \ldots, d_N$. The distance measure d_i can be the L_2 norm $(\mathbf{x}_i - \overline{\mathbf{a}})^T(\mathbf{x}_i - \overline{\mathbf{a}})$, the Mahalanobis distance $[(\mathbf{x}_i - \overline{\mathbf{a}})^T \Gamma^{-1} (\mathbf{x}_i - \overline{\mathbf{a}})]^{1/2}$ (Γ being the covariance matrix of the input data), or some other distance measure. The arithmetic mean, the marginal or vector median, or any other suitable multivariate estimator of location evaluated over a subset of the input data, can be used as reference vector $\overline{\mathbf{a}}$. R-ordering has been used in [1] to introduce the R_E, R_M filters.

3 Multichannel L filters based on R-ordering

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

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

where $\mathbf{A}_i$ are $(p \times p)$ coefficient matrices and $\mathbf{x}_{[i]}(k)$ are the R-ordered input vectors $\mathbf{x}(k - \nu), \ldots, \mathbf{x}(k), \ldots, \mathbf{x}(k + \nu)$, $N = 2\nu + 1$. Let us suppose that the input signal $\mathbf{x}(k)$ is a constant vector-valued signal $\mathbf{s} = [s_1, \ldots, s_p]^T$ corrupted by additive zero-mean p-channel noise $\mathbf{n}(k) = [n_1(k), \ldots, n_p(k)]^T$. The noise components are distributed according to some joint distribution

function $f_{\mathbf{n}}$. Furthermore the noise vectors at different time instants are assumed to be iid and uncorrelated to the constant signal. The mean square error (MSE) between the filter output and the constant signal $\mathbf{s}$ can be expressed in the following way :

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

Suppose that $\mathbf{R}_{ji}, \quad i,j = 1\ldots,N$ denotes the $(p\times p)$ correlation matrix of the j-th and i-th order statistics and $\underline{\mu}_i, \ i = 1,\ldots,N$ the mean vector of the i-th order statistic :

$$\mathbf{R}_{ji} = \mathrm{E}[\mathbf{x}_{[j]}\mathbf{x}_{[i]}^T], \quad \underline{\mu}_i = E[\mathbf{x}_{[i]}]. \quad (3)$$

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

$$\mathbf{a}_i = \left[\begin{array}{cccc}\mathbf{A}_{1i} & \mathbf{A}_{2i} & \cdots & \mathbf{A}_{Ni}\end{array}\right]^T \ i = 1,\ldots,p \quad (4)$$

where $\mathbf{A}_{ji}$ denotes the i-th row of the matrix $\mathbf{A}_j$. Finally, let us define the $(Np\times Np)$ matrix $\check{\mathbf{R}}_N$ and the $(Np\times 1)$ vector $\underline{\check{\mu}}_N$ in the following way:

$$\check{\mathbf{R}}_N = \left[\begin{array}{ccc}\mathbf{R}_{11} & \cdots & \mathbf{R}_{1N} \\ \vdots & \ddots & \vdots \\ \mathbf{R}_{N1} & \cdots & \mathbf{R}_{NN}\end{array}\right] \quad (5)$$

$$\underline{\check{\mu}}_N = \left[\begin{array}{cccc}\underline{\mu}_1^T & \underline{\mu}_2^T & \cdots & \underline{\mu}_N^T\end{array}\right]^T. \quad (6)$$

After some manipulation and using the previous notation, the MSE (2) takes the following form :

$$\varepsilon = \sum_{i=1}^{p}\mathbf{a}_i^T\check{\mathbf{R}}_N\mathbf{a}_i - 2\mathbf{s}^T\left[\begin{array}{c}\mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_p^T\end{array}\right]\underline{\check{\mu}}_N + \mathbf{s}^T\mathbf{s}. \quad (7)$$

3.1 Unconstrained filter coefficients

A multichannel L filter with no constraint imposed on its coefficients is called an unconstrained L filter. The optimum coefficients in this case can be found by differentiating (7) with respect to $\mathbf{a}_i$ and equating the result to zero. This optimization problem has a unique solution provided that $\check{\mathbf{R}}_N$ is positive definite. By solving the resulting sets of equations we obtain the following expressions for the optimum filter coefficients :

$$\mathbf{a}_i = s_i\check{\mathbf{R}}_N^{-1}\underline{\check{\mu}}_N \quad i = 1,\ldots,p. \quad (8)$$

3.2 Unbiased filter coefficients

A multichannel L filter is said to be unbiased estimator of location if $\mathrm{E}[\mathbf{y}(k)] = \mathbf{s}$. This is equivalent to :

$$\mathbf{a}_i^T\underline{\check{\mu}}_N = s_i \quad i = 1,\ldots,p. \quad (9)$$

The minimization of MSE (7) when coefficients satisfy (9) is a constrained optimization problem and can be solved using the method of Lagrange multipliers. The resulting optimum coefficients are given by :

$$\mathbf{a}_i = \frac{s_i}{\underline{\check{\mu}}_N^T\check{\mathbf{R}}_N^{-1}\underline{\check{\mu}}_N}\check{\mathbf{R}}_N^{-1}\underline{\check{\mu}}_N \quad i = 1,\ldots,p. \quad (10)$$

3.3 Location invariant filter coefficients

A multichannel L filter is said to be location invariant if an input of the form $\mathbf{x}'(k) = \mathbf{x}(k)+\mathbf{b}$ produces an output $\mathbf{y}'(k) = \mathbf{T}[\mathbf{x}'(k)] = \mathbf{y}(k)+\mathbf{b}$, where $\mathbf{y}(k) = \mathbf{T}[\mathbf{x}(k)]$. This property implies that the following constraint holds for the filter coefficients :

$$\sum_{i=1}^{N}\mathbf{A}_i = \mathbf{I} \quad (11)$$

where $\mathbf{I}$ is the $(p\times p)$ identity matrix. The evaluation of the optimum filter coefficients can be done using Lagrange multipliers. Let us define the $(Np\times Np)$ matrix $\tilde{\mathbf{R}}_N$ that has the same structure as the matrix $\check{\mathbf{R}}_N$ but whose submatrices $\tilde{\mathbf{R}}_{ji}$ are the $(p\times p)$ correlation matrices of the j-th and i-th order statistics of the noise vectors :

$$\tilde{\mathbf{R}}_{ji} = E[\mathbf{n}_{[j]}\mathbf{n}_{[i]}^T], \quad i,j = 1,\ldots,N. \quad (12)$$

Let us also denote by $\tilde{\mathbf{R}}_p$ the matrix that results from $\tilde{\mathbf{R}}_N^{-1}$ through the following rearrangement of its elements :

$$\begin{aligned}[\tilde{R}_p]_{N\cdot(i-1)+j,N\cdot(k-1)+l} = [\tilde{R}_N^{-1}]_{i+p\cdot(j-1),k+p\cdot(l-1)} \\ i,k = 1,\ldots,p \ \ j,l = 1,\ldots,N\end{aligned} \quad (13)$$

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

$$\tilde{\mathbf{R}}_p = \left[\begin{array}{cccc}\mathbf{P}_{11} & \mathbf{P}_{12} & \cdots & \mathbf{P}_{1p} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{P}_{p1} & \mathbf{P}_{p2} & \cdots & \mathbf{P}_{pp}\end{array}\right] \quad (14)$$

where the submatrices $\mathbf{P}_{ij}$ are of dimensions $(N\times N)$. Also we define :

$$\mathbf{G} = \left[\begin{array}{ccc}\mathbf{e}^T\mathbf{P}_{11}\mathbf{e} & \mathbf{e}^T\mathbf{P}_{12}\mathbf{e} & \cdots \ \mathbf{e}^T\mathbf{P}_{1p}\mathbf{e} \\ \vdots & \ddots & \vdots \\ \mathbf{e}^T\mathbf{P}_{p1}\mathbf{e} & \mathbf{e}^T\mathbf{P}_{p2}\mathbf{e} & \cdots \ \mathbf{e}^T\mathbf{P}_{pp}\mathbf{e}\end{array}\right] \quad (15)$$

where $\mathbf{e} = [1, \ldots, 1]^T$. The optimum location invariant L filter coefficients were found to be :

$$\mathbf{a}_i = \tilde{\mathbf{R}}_N^{-1} \begin{bmatrix} \mathbf{c}_i \\ \vdots \\ \mathbf{c}_i \end{bmatrix} \quad i = 1, \ldots, p \tag{16}$$

where $\mathbf{c}_i$ is the i-th column of $\mathbf{G}^{-1}$.

4 Moments of R-ordered data

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

$$d_i = \sum_{k=1}^{p} (x_{ik} - s_k)^2 = \sum_{k=1}^{p} n_{ik}^2. \tag{17}$$

From the previous expression it is obvious that d_i are iid random variables. By ordering d_i we also imply an ordering of $\mathbf{x}_i$. The notation $d_i \leftrightarrow d_{(j)}$ will be used in the sequel to denote that d_i occupies the j-th position in the ordered sequence $d_{(1)}, \ldots, d_{(N)}$. The same meaning is given to the notation $\mathbf{x}_i \leftrightarrow \mathbf{x}_{[j]}$, for the R-ordered vectors. The following formulae have been derived :

$$f_{x_{[i]k},x_{[j]l}}(u,v) = \int_0^{+\infty} \int_0^{w_2} \frac{f_{x_{mk},d_m}(u,w_1)}{f_{d_m}(w_1)} \cdot \frac{f_{x_{nl},d_n}(v,w_2)}{f_{d_n}(w_2)} f_{d_{(i)},d_{(j)}}(w_1,w_2) dw_1 dw_2, \; i < j \tag{18}$$

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

$$f_{x_{[i]k},x_{[i]l}}(u,v) = \int_0^{+\infty} \frac{f_{x_{nk},x_{nl},d_n}(u,v,w)}{f_{d_n}(w)} f_{d_{(i)}}(w) dw, \; k \neq l \tag{19}$$

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

$$f_{x_{[i]k}}(u) = \int_0^{+\infty} \frac{f_{x_{nk},d_n}(u,w)}{f_{d_n}(w)} f_{d_{(i)}}(w) dw \tag{20}$$

where $\mathbf{x}_n \leftrightarrow \mathbf{x}_{[i]}$ and $d_n \leftrightarrow d_{(i)}$. The next step that we have to undertake is to evaluate the pdfs $f_{d_{(i)}}(w)$, $f_{d_{(i)},d_{(j)}}(w_1,w_2)$, $f_{d_i}(w)$, $f_{x_{ik},d_i}(u,w)$, $f_{x_{ik},x_{il},d_i}(u,v,w)$.

To do so, we start by evaluating the joint pdf g of $d_i, n_{i1}, ., n_{im-1}, n_{im+1}, ., n_{ip}$ using the well-known method for the evaluation of the joint pdf of functions of random variables [4] :

$$g = \{f_{\mathbf{n}}(n_{i1}, .., n_{im-1}, h, n_{im+1}, .., n_{ip}) + f_{\mathbf{n}}(n_{i1}, .., n_{im-1}, -h, n_{im+1}, .., n_{ip})\}/2h, \quad h = \sqrt{(d_i - \sum_{l=1, l\neq m}^{p} n_{il}^2)}. \tag{21}$$

The previous formula holds for $d_i \geq \sum_{l=1,l\neq m}^{p} n_{il}^2$. In the opposite case, the joint pdf equals to zero. The pdf $f_{d_i}(w)$ can then be found by integrating (21) over $n_{i1}, .., n_{im-1}, n_{im+1}, .., n_{ip}$. The joint pdf of d_i, n_{il} for all $l \neq m$ can be found in a similar manner by integrating (21) over all n_{ir}, $r \neq m$, $r \neq l$. The evaluation of the joint pdf of d_i, n_{im} can be done in the same way. Having calculated $f_{n_{il},d_i}(u,w)$ we can easily derive the joint pdf of d_i, x_{il} :

$$f_{x_{il},d_i}(u,w) = f_{n_{il},d_i}(u - s_l, w). \tag{22}$$

The joint pdf of d_i, n_{il}, n_{ik}, $l \neq m$, $k \neq m$ can be evaluated by integrating (21) over all n_{ir}, $r \neq m$, $r \neq l$, $r \neq k$. $f_{x_{ik}x_{il}d_i}(u,v,w)$ is given by :

$$f_{x_{ik}x_{il}d_i}(u,v,w) = f_{n_{ik}n_{il}d_i}(u - s_k, v - s_l, w). \tag{23}$$

Finally, $f_{d_{(i)}}$, $f_{d_{(i)},d_{(j)}}$ are given by [3] :

$$f_{d_{(i)}}(w) = \frac{N!}{(i-1)!(N-i)!} F_{d_i}^{i-1}(w)[1 - F_{d_i}(w)]^{N-i} f_{d_i}(w) \tag{24}$$

$$f_{d_{(i)},d_{(j)}}(u,w) = \frac{N!}{(i-1)!(j-i-1)!(N-j)!} F_{d_i}^{i-1}(u) f_{d_i}(u)[F_{d_i}(w) - F_{d_i}(u)]^{j-i-1} f_{d_i}(w) [1 - F_{d_i}(w)]^{N-j}, \; i < j, \; u < w \tag{25}$$

where F_{d_i} denotes the cdf of d_i. By substituting the derived pdfs in (18)-(20) we can calculate the elements of the matrix $\check{\mathbf{R}}_N$ and the vector $\underline{\check{\mu}}_N$ and therefore proceed to the evaluation of the filter coefficients. The calculation of the moments that are required in the case of the location invariant filter can be done in a similar way.

It must be noted that in real situations the constant signal $\mathbf{s}$ as well as the pdf $f_{\mathbf{n}}$ of the noise vector is not known. However $\mathbf{s}$ can be estimated using the arithmetic mean or the marginal median on a relatively large window in the constant signal area. $f_{\mathbf{n}}$ can also be estimated from the input data. However it is much more easier to estimate directly the moments $\mathrm{E}[x_{[i]k}]$, $\mathrm{E}[x_{[i]k}, x_{[j]l}]$. In the case of non-constant vector-valued signals,

a segmentation of the signal must be performed prior to the filtering procedure in order to obtain areas with almost constant vectors. Estimates for $\mathbf{s}$, $\mathrm{E}[x_{[i]k}]$, $\mathrm{E}[x_{[i]k}, x_{[j]l}]$ must then be found for each region in order to evaluate the filter coefficients that correspond to this region. Unlike the unbiased and unconstrained multichannel L filters, the evaluation of optimal location invariant filter coefficients does not require a priori knowledge of the noise-free input signal but depends only on the noise statistics. Therefore the signal or image has not to be segmented since a single set of coefficients suffices for the entire image.

5 Experimental results

In order to test the performance of the proposed filters we conducted an experiment on a noisy two-channel vector field of dimensions (64×64). The original, noise-free vector field was composed of two constant signal regions. The first region was a rectangle of dimensions (32×32) placed in the center of the vector field which was surrounded by the second region. The constant vectors corresponding to the two regions were $\mathbf{s}_1 = [2, -5]^T$ and $\mathbf{s}_2 = [1, 2]^T$ respectively. The original vector field was corrupted by additive bivariate zero-mean noise whose components were distributed according to the Laplacian-Morgenstern distribution [2] with equal standard deviations $\sigma_i = \sqrt{2}$ and correlation coefficient $r = 9/32$. The reference vector $\overline{\mathbf{a}}$ was chosen to be the output of a 5×5 marginal α-trimmed mean filter with trimming coefficient $\alpha = 0.4$, centered on the current filter position. The window size for all filters was 3×3. The Noise Reduction Index (NRI), in dB, obtained by the proposed filters can be found in Table 1 along with the NRI obtained by other filters. The superiority of the proposed multichannel L-filters is obvious.

The second experiment concerned the filtering of noisy color images using location invariant multichannel L filters, which do not require the constant signal to be known. The color image "Pepper" has been used in the experiments. The above-mentioned image was corrupted by zero-mean additive multivariate contaminated Gaussian noise in all three RGB channels :

$$0.8 \cdot \mathbf{N}(m_R, m_G, m_B, \sigma_R, \sigma_G, \sigma_B, r_{RG}, r_{GB}, r_{RB}) + 0.2 \cdot \mathbf{N}(m'_R, m'_G, m'_B, \sigma'_R, \sigma'_G, \sigma'_B, r'_{RG}, r'_{GB}, r'_{RB})$$

where $m_R = m_G = m_B = m'_R = m'_G = m'_B = 0$, $\sigma_R = \sigma_G = \sigma_B = 20$, $\sigma'_R = \sigma'_G = \sigma'_B = 40$, $r_{RG} = r_{GB} = r_{RB} = r'_{RG} = r'_{GB} = r'_{RB} = 0.5$. The reference vector for the location invariant filter was chosen to be the output of a 5×5 α-trimmed filter with $\alpha = 0.2$. The performance of the various 3×3 filters was compared with respect to the Noise Reduction Index (Table 2). The proposed multichannel filter (marked #1 in Table 2) achieved the highest noise reduction among all the filters. A further performance enhancement was achieved by combining the multichannel L filter with a marginal α-trimmed mean filter. An edge map was used to switch between the multichannel L filter (homogeneous regions) and the 3×3 α-trimmed mean filter ($\alpha = 0.222$) (edge areas). By this simple technique the performance of the location invariant filter was improved (entry marked #2 in Table 2).

References

[1] R.C Hardie, G.R Arce, "Ranking in R^p and its use in multivariate image estimation", IEEE Trans. CSVT, vol 1, No. 2, pp 197-209, June 1991.

[2] C. Kotropoulos, I. Pitas "Multichannel L-filters based on marginal data ordering", IEEE Transactions on Signal Processing, vol 42, No 10, Oct. 1994.

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

[4] A. Papoulis, *Probability, random variables and stochastic processes*, McGraw Hill, 1984.

Unconstrained L	-12.40
Unbiased L	-12.39
Loc. Invariant L	-10.98
Arith. Mean	-7.04
Marginal Median	-9.70
a-trimmed, a=0.33	-10.05
L_1 Vect. Median	-8.25

Table 1: NRI (dB) obtained by various 3×3 filters in the filtering of a noisy vector field.

Loc. Invariant L #1	-7.85
Loc. Invariant L #2	-8.14
Arith. Mean	-7.28
Marginal Median	-6.82
L_1 Vect. Median	-5.29
α-trimmed, α=0.11	-7.56
single-channel L	-7.58

Table 2: NRI (dB) obtained by various 3×3 filters in the filtering of noisy color image "Pepper".