# AFFINE EQUIVARIANCE IN MULTICHANNEL OS-FILTERING

V. Koivunen<sup>1</sup> S. Luukkonen<sup>2</sup>, H.  $Oja^2$ 

<sup>1</sup>Signal Processing Laboratory Tampere Univ. of Technology P.O. Box 553, FIN-33101 Tampere Finland

ABSTRACT

Nonlinear multichannel filters have successfully been applied to biomedical signals, multichannel images as well as processing of vector fields. In multichannel signals, component variances and correlations among components may be unequal and time-varying. Such changes can be expressed as an affine transformation of the input signal. In this paper, we investigate how the performance and statistical properties of multichannel filters stemming from order statistics (OS) change under affine transformations. An affine equivariant multichannel filter is introduced and the use of affine equivariant performance metric replacing the Mean Square Error is proposed. Advantages of affine equivariance are demonstrated in simulation, and filtering examples using real data are given.

## 1. INTRODUCTION

A growing number of signal processing applications use multichannel data. Components in multichannel signals may have different variances and may be correlated and statistical properties of the signal may be time-varying. As an example, a correlated elliptically symmetric noise density can be viewed as an affine transformation of a spherically symmetric density. Consequently, it's desirable to develop filtering techniques that yield estimates that commute with changes in variances and correlation structure.

In this paper, we study a multichannel OS-filter for noise attenuation that behaves properly under affine transformations, i.e., is affine equivariant. Formally, an estimator T is affine equivariant if

$$T(AX + \mathbf{v}) = AT(X) + \mathbf{v}$$

where A is a nonsingular  $k \times k$  matrix and v is a k-component translation vector and X are the k-variate sample points. An affine equivariant multivariate generalization of the median operator, the Oja median (OM) [6], is defined and an approximate algorithm is introduced. This new algorithm is then used in investigating the implications of equivariance on optimality. Commonly used multichannel OS-filters such as Vector Median (VM) and Marginal Median (MM) filters do not possess this equivariance property. The OM filter does not require covariance estimation. Other affine equivariant robust filters such as [3, 4], estimate the dispersion matrix. Both the signal value and dispersion matrix need to be estimated in a robust manner.

The paper is organized as follows. In Section 2 the affine equivariant Oja median [6] is defined and an approximate filtering algorithm is introduced. The affine equivariance of different error criteria is discussed as well. In section 3, the importance of affine equivariance is illustrated in terms of efficiency. Finally in Section Finland 4, filtering examples are given using simulated data and 3-channel image data. Component variances and correlation structure vary

<sup>2</sup>Dept. of Math. Sciences University of Oulu

FIN-90570 Oulu

## 2. EQUIVARIANT MULTICHANNEL OS-FILTER

in test signals. The properties of the Oja median are compared to

#### 2.1. Definitions

VM, MM and the mean vector.

The affine equivariant multichannel filter studied in this paper is based on the Oja median [6] defined as follows. Let

$${f x_1,...,x_n} = ({f x_{11},...,x_{1k}})^T,...,({f x_{n1},...x_{nk}})^T$$

be a random sample from a k-variate distribution. Let

$$P = \{ p = (i_1, ..., i_k) : 1 \le i_1 < ... < i_k \le n \}$$

be the set of  $N = {n \choose k}$  different k-tuples of the index set  $\{1, ..., n\}$ . Index  $p \in P$  refers to a k-subset of the *n* original observations. The volume of the simplex formed by  $p = (i_1, ..., i_k)$  and a candidate estimate  $\theta = (\theta_1, ..., \theta_k)^T$  is then

$$V_p(\theta) = \frac{1}{k!} abs \left\{ \det \left[ \begin{array}{cccc} 1 & 1 & \cdots & 1 \\ \theta & \mathbf{x}_{i_1} & \cdots & \mathbf{x}_{i_k} \end{array} \right] \right\}$$
(1)

where the index of  $\mathbf{x}_{i_i}$  refers to the index set *P*. The multivariate Oja median  $\hat{\theta}$  minimizes the sum of the volumes of the simplices

$$D_n(\theta) = \sum_{p \in P} V_p(\theta), \qquad (2)$$

where the sum is taken over all subsets  $p \in P$ . In case of 3-variate signals the simplex  $V_p(\theta)$  is a tetrahedron and in bivariate case its a triangle. The idea can be traced to the univariate median which minimizes the sum of absolute distances between the points  $x_i$  and  $\theta$ , i.e., the sum of volumes of the univariate simplices:

$$D_n( heta) = \sum_{i=1}^n V_i( heta) = \sum_{i=1}^n |x_i - heta|$$

Note that the sample mean vector minimizes the sum of the squared volumes. The influence function of Oja median is uniformly bounded (see [5]). The breakdown point of the Oja median depends on the dispersion of the corrupted points. The breakdown point is at least  $1 - 2^{-1/k}$  and 1/k if the corrupted points equal.

#### 2.2. Approximate algorithm

Any algorithm for solving  $L_1$  regression problems can be used for computing Oja median as can be seen from (1) and (2). However, an exact algorithm for computing the OM would require extensive computation. Therefore, an approximate algorithm which selects the solution among the original observations is presented:

In processing window of n vectors x<sub>i</sub>, i = 1, ...n, for each vector x<sub>i</sub> calculate the sum (2)

$$D_i = \sum_{p \in P} V_p\left(\mathbf{x_i}\right)$$

- 2. The output is the vector  $\mathbf{x}_i$  yielding the minimum of sums  $D_i$ .
- If the minimum is not unique, i.e., a tie occurs, the output have to be chosen from the set of input vectors x<sub>i</sub> yielding the minimum sum D<sub>i</sub>.

The observations are assumed to be in general position. In case of k-channel signals this means that no collection of more than k samples is allowed to fall into (k - 1)-dimensional subspace. However, if signal values in each channel are quantized to discrete values as in the case of color images, ties may occur. The ties are resolved as follows: If we have multiple candidates yielding the minimum for the error criterion, we choose the most frequently occurring one. If the tie is not yet resolved we choose the candidate which is closest to the value at the center of the processing window.

In order to illustrate how to apply the filter defined above, we use 3-channel data such as RGB color image. Let R, G, B denote the 3 channels. Then each sample is given by  $\mathbf{x_i} = (\mathbf{R_i}, \mathbf{G_i}, \mathbf{B_i})^{\mathrm{T}}$ . If a 3 × 3 processing window of 9 samples is used, the volume of the first simplex formed with the output candidate  $\hat{\theta} = \mathbf{x_j}$  may be obtained by

$$V_1(\mathbf{x_j}) = rac{1}{3!} \mathbf{abs} \left\{ \det \left[ egin{array}{ccccc} 1 & 1 & 1 & 1 \ R_j & R_{i_1} & R_{i_2} & R_{i_3} \ G_j & G_{i_1} & G_{i_2} & G_{i_3} \ B_j & B_{i_1} & B_{i_2} & B_{i_3} \end{array} 
ight] 
ight\},$$

where j = 1, ..., 9. The constant  $\frac{1}{3!}$  can be left out.

#### 2.3. Affine equivariant performance measure

In multichannel filtering applications the Mean Square Error (MSE) is most commonly used performance measure. This measure, however, is equivariant only under orthogonal, rigid body transformations. Therefore, does not take into account different scales and correlations among channels. Bensmail and Celeux [2] proposed eigenvalue decomposition for the error covariance matrix in the form

$$C = \lambda U \Sigma U^T$$
,

where U is the matrix of eigenvectors,  $\Sigma$  is a diagonal matrix with the normalized eigenvalues ( $det(\Sigma) = 1$ ) on the diagonal and  $\lambda$  is the Wilk's generalized variance. Terms *scale, shape and orientation* are used for items  $\lambda$ ,  $\Sigma$  and U. Generalized variance is related to the determinant of the error covariance matrix C whereas the Mean Square Error is related to the trace of C. The determinant behaves properly under affine transformations described by nonsingular matrix A because

$$det(ACA^{T}) = det(A)det(C)det(A^{T}) = (det(A))^{2}det(C).$$

### 3. IMPLICATIONS OF EQUIVARIANCE

In this section we study the benefits of affine equivariant filters in case the component variances or correlationg among components change. The performance is measured in terms of relative *efficiency* which is the ratio of error variances of two estimators. The other estimator may also be the optimal estimator yielding the smallest possible error variance. The efficiency of affine equivariant Oja median (OM) is compared to marginal median which is scaling equivariant, and to VM which is equivariant under orthogonal transformations.

The limiting distribution of  $n^{1/2}(\hat{\theta} - \theta)$  for OM has been shown to be k-variate normal. This holds for mean, VM and MM as well. Hence, one can compare the asymptotic efficiencies using asymptotic covariance matrices. For an underlying multivariate normal distribution the OM has good efficiency properties with respect to an optimal estimator (sample mean). For example, some values of the asymptotic marginal variance ratios (efficiencies) e as function of the dimension k in this comparison are.

$$(k, e(k)) = (2, 0.785), (3, 0.849), (6, 0.920).$$

If the underlying distribution is elliptic, the asymptotic efficiency of Oja median is superior to that of VM. In the spherical case the efficiencies equal. As an example, write  $e_1$  and  $e_2$  for the marginal efficiencies of the bivariate Oja median with respect to the bivariate VM. In the spherical case,

$$e_1 = e_2 = 1.0$$

If we then rescale the first component by multiplying it by  $\sigma$ , the efficiencies as a function of  $\sigma$ ,  $e_1(\sigma)$  for the first component and  $e_2(\sigma)$  for the second component, are

$$e_1(2) = 1.02 \qquad e_2(2) = 1.04 \\ e_1(5) = 1.10 \qquad e_2(5) = 1.23 \\ e_1(100) = 1.22 \qquad e_2(100) = 4.90$$

 $e_1$  goes to 1, 23 and  $e_2$  to the infinity as  $\sigma$  goes to the infinity. This means that if the variance of the first component is large as compared to the second component variance, the efficiency of the second component estimate may be really poor.

### 4. EXAMPLES

Simulated multichannel data and multichannel RGB color images are used in filtering examples. The performance Oja median (OM), VM and MM filters is evaluated quantitatively using Generalized Variance (GV) and MSE measures.

#### 4.1. Illustration of benefits of equivariance

The benefits of affine equivariance are demonstrated in a simulation. The simulation is performed using 100 realizations of 49 observations from bivariate Gaussian distribution. In Figure 1, spherically distributed data are transformed into elliptically distributed data by applying an affine transformation. As a result, the eigenvalues and eigenvectors of the data covariance matrix change. Scatter plot of 100 filter outputs is given in case no transformation is performed ( $\sigma_1^2 = \sigma_2^2 = 1.0$ , left column), one component is scaled ( $\sigma_1^2 = 100$ , middle column) and eigenvectors of the covariance matrix are rotated by  $\pi/4$  (right column). The output is here restricted to be one of the input values both in the OM and VM



Figure 1: Affine equivariant OM filter behaves properly under affine transformations whereas the variance of the estimates increases for VM and MM because they do not possess the equivariance property. Scatter plots of 100 estimates are given in case no transformation is performed (left column), one component is scaled (middle column) and eigenvectors of the covariance matrix are rotated by  $\pi/4$  (right column). (I) Change of variance (middle column) in one signal component causes increase in the variance of estimates obtained by VM (middle row) filter whereas the MM (I.a1-a2) and OM (I.c1-c3) filters behave properly. In 3rd column, also rotation is applied and the scatter of MM estimates increases significantly. (II) Transformations by scaling an eigenvalue (middle column) and rotating eigenvectors of the covariance matrix (right column) increase the variance of the estimates obtained by MM filter (II.a3) and VM (II.b2, the other component) whereas the OM (bottom) filter commutes.

Table 1: Generalized variances and MSE for the simulated data.

	No transf.		Scaling		Affine	
	GV	MSE	GV	MSE	GV	MSE
MM	0.0012	0.036	0.0012	0.036	0.0053	0.118
VM	0.0020	0.046	0.0034	0.074	0.0034	0.074
OM	0.0018	0.044	0.0018	0.044	0.0018	0.044

Table 2: Generalized variance and average component MSE: Multivariate Gaussian noise with unequal component variances ( $\sigma_1^2 = 400, \sigma_2^2 = 225, \sigma_3^2 = 100, \rho = 0.7$ ) is added to the image and 10% of the samples are replaced by outliers, which have minimum or maximum signal value with equal probability.

Algorithm	GV	MSE
OM	$78.4 \times 10^{4}$	126.8
VM	$51.5 \times 10^{4}$	99.4

filtering algorithms. In Fig. 1 (I), obtained estimates are retransformed back to original coordinate system. The dispersion of the outputs obtained by Oja median remains the same because of its affine equivariance property as can be seen in (I.c2-c3). MM commutes with scaling as can be seen from Fig.1 (I.a2) but dispersion increases when eigenvectors are rotated (I.a3). In VM filtering, the scaling of the first signal component makes it a dominant one and the variance in the estimates of the other signal component is significantly increased which can easily be observed from Fig.1 (I.b2, II.b2). VM commutes with the rotation of eigenvectors (I.b3). Fig. 1. (II) shows the scatter plots in case where no retransformation back to the original coordinate system is made. For the VM (II.b1b3) and OM, the scatter in the right column remains otherwise the same as in the middle except for the rotation (II.c1-c3) whereas the variance in the estimates produced by MM (II.a1-a3) is increased.

# Quantitative results for the simulation above are given using GV and MSE performance measures. In the first case no transformation is applied to the data, in the second case data are rescaled and in the third case full affine transformation is applied by rotating and rescaling data. The filters are then applied and retransformation back to original coordinate system performed.

#### 4.2. Example using 3-channel RGB-data

In case of RGB image data, correlated Gaussian noise with unequal component variances is added to original image. Furthermore, contaminated data are created by replacing 10% samples with outliers having minimum or maximum signal value with equal probability. The filter is compared to VM filter quantitatively using MSE and GV. The results are given in Table 2. The results indicate that VM has lower MSE because it has lower bias. The larger bias in OM is due to its low breakdown point. In qualitative comparison we study detail preservation, in particular, edge preservation. Fig. 2 depicts the contaminated image data and a subimage where several edges appear. The Oja median appears to preserve edges better. In practice ties occur quite frequently, hence the procedure for resolving ties has an impact on both the quantitative and qualitative results.

## 5. CONCLUSION

In this paper, we investigated the affine equivariance property of nonlinear multichannel filters. The OM filter possess such property. We illustrated that the efficiency of OM filter is not lost even if the signal component variances or correlations among the com-



Figure 2: (a) The original Room  $256 \times 256$  RGB color image contaminated by multivariate Gaussian noise with unequal component variances ( $\sigma_1^2=400$ ,  $\sigma_2^2=225$ ,  $\sigma_3^2=100$ ,  $\rho = 0.7$ ). Moreover, 10 % of the samples are replaced by outliers which have minimum or maximum signal value with equal probability. (b) A detail from the original image without added noise and (c) from the contaminated image. The filter outputs obtained by (d) Oja median and (e) Vector Median (VM) filters using  $3 \times 3$  processing window are shown as well. The output of Oja median appears less blurred.

ponents change whereas the efficiency of scaling (MM) and orthogonal equivariant (VM) filters decreases. We also introduced an approximate algorithm for computing the affine equivariant OM output. The behavior of MSE and GV performance measures under affine transformations was investigated as well. Benefits of affine equivariance were illustrated in a simulation and an example using 3-channel RGB image data was presented. In the latter task, the OM filter preserves edges better than the VM filter but has higher MSE because of bias caused by outliers. The computational complexity of OM is high which suggests the use of filters that employ covariance estimation [3] in case signal components are correlated and have unequal variances.

## 6. REFERENCES

- Astola, J., Haavisto, P., and Neuvo, Y., "Vector median filters", Proceedings of the IEEE, Vol. 78, No. 4., pp. 678-689, 1990.
- [2] Bensmail, H. and Celeux, G. "Regularized Gaussian Discriminant Analysis Through Eigenvalue Decomposition", JASA 91, 1741-1748.

- [3] Koivunen, V., "Nonlinear Filtering of Multivariate Images Under Robust Error Criterion", IEEE Trans.Image Processing, vol. 5, no. 6, pp. 1054-1060, 1996.
- [4] Koivunen, V., Himayat, N., and Kassam, S.A., "Nonlinear Filtering Techniques for Multivariate Images - Design and Robustness Characterization", Signal Processing, Vol. 57, pp. 81-91, 1997.
- [5] Niinimaa, A., and Oja, H., "On the Influence Functions of Certain Bivariate Medians", J. R. Statist. Soc., B, 57, pp. 565-574, 1995
- [6] Oja, H., "Descriptive Statistics for Multivariate distributions", Statist. and Prob. Letters, 1, pp.327-332, 1983.
- [7] Pitas I., Venetsanopoulos, A. "Order statistics in digital image processing", Proceedings of the IEEE. Vol. 80, No. 12, pp. 1893-1921, 1992.
- [8] Small, C., "A survey of multidimensional medians", International Statistical Review, Vol. 58, No. 3, pp. 263-277, 1990.