# Adaptive LMS $L$-filters for noise suppression in images

C. Kotropoulos I. Pitas*

Department of Informatics

Aristotle University of Thessaloniki

Thessaloniki 540 06, GREECE

**Abstract**

Several adaptive LMS $L$-filters, both constrained and unconstrained ones, are developed for noise suppression in images and being compared in this paper. First, the location-invariant LMS $L$-filter for a nonconstant signal corrupted by zero-mean additive white noise is derived. It is demonstrated that the location-invariant LMS $L$-filter can be described in terms of the generalized linearly constrained adaptive processing structure proposed by Griffiths and Jim. Subsequently, the normalized and the signed error LMS $L$-filters are studied. A modified LMS $L$-filter with nonhomogeneous step-sizes is also proposed in order to accelerate the rate of convergence of the adaptive $L$-filter. Finally, a signal-dependent adaptive filter structure is developed to allow a separate treatment of the pixels that are close to the edges from the pixels that belong to homogeneous image regions.

*Corresponding author: I. Pitas, Dept. of Informatics, Aristotle University of Thessaloniki, P.O. Box 451, GR-540 06 Thessaloniki, Greece, tel.: +30-31-996304, fax: +30-31-996304, email: pitas@zeus.csd.auth.gr

Adaptive signal processing has exhibited a significant development in the past two decades. Adaptive filters have been applied in a wide variety of problems including system identification, channel equalization, echo cancellation in telephone channels, suppression of narrowband interference in wideband signals and adaptive arrays [1]–[4]. The most widely known adaptive filters are linear ones having the form of either FIR or lattice filters. However, such filters may not be suitable for applications where the transmission channel is nonlinear or where the noise is impulsive or where the signal is strongly nonstationary (e.g. in image processing).

On the contrary, a multitude of nonlinear techniques has been proven a successful alternative to the linear techniques in all the above-mentioned cases. For a review of the nonlinear filter classes the reader may consult [5]. One of the best known families is based on the order statistics [6]. It uses the concept of sample ordering. The power of the ordering concept is well illustrated by median filters which preserve edges and are optimal estimators for impulsive noise. There is now a multitude 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 of the input sequence [7]. The $L$-filters have found extensive application in digital signal and image processing, because they have a well-defined design methodology as the estimators which minimize the Mean-Squared Error (MSE) between the filter output and the noise-free signal [8]–[11]. It is well-known that digital image filtering techniques must take into account the local image content (i.e., the local statistics), because image statistics vary throughout an image. They should exploit the correlation between image pixels as is the case in homogeneous regions. Furthermore, they should meet different objectives. For example, in homogeneous regions the objective is noise smoothing, whereas close to the

the edges than in homogeneous regions. Such requirements are fulfilled by adaptive filtering techniques, because it has been proved both theoretically as well as in practice that adaptive techniques can cope with nonstationary and/or time varying signals. Moreover, adaptive techniques do not make any a priori assumptions regarding the statistics of the data and the degradations (i.e., the noise) as opposed to nonadaptive ones. In our case, a design of $L$-filters which relies on a non-iterative minimization of the MSE between the filter output and the desired response yields very tedious expressions for computing the marginal and joint cumulative functions of the order statistics (cf. [9]). Consequently, the design of adaptive $L$-filters is proved appealing in avoiding the computational burden of the non-iterative methods. Therefore, nonlinear adaptive techniques should be pursued. Signal-adaptive filters, i.e., filters that change their smoothing properties at each image point according to the local image content, have been used in image processing applications where impulsive or signal-dependent noise is present [12]–[16]. Most of these filters change their coefficients close to the image edges or close to impulses, so that their performance becomes similar to the median filter. However, the main disadvantage of such filters is that the update of their coefficients is rather heuristic and usually does not come from the minimization of an error norm.

Recently, the adaptation of the coefficients employed in order statistic filters by using linear adaptive signal processing techniques has received much attention in the literature [17]–[28]. In this paper, we shall confine ourselves to the design of adaptive $L$-filters. Many authors have tried to design adaptive $L$-filters either by using the Least Mean Squares (LMS) or the Recursive Least Squares (RLS) algorithm [18]–[21] or by using constrained LMS adaptive algorithms [22]–[28]. The majority of the above-mentioned algorithms have been derived assuming a constant signal corrupted by zero-mean additive white noise. All of

[27] that have been tested on images. But the later algorithms, are mainly steepest-descent algorithms that use the correlation matrix of the order statistics and the crosscorrelation vector between the order statistics and the desired response themselves and not estimates of them.

Digital images can be degraded either by motion blur, defocus blur and/or by introducing noise during acquisition or transmission. In this paper, we shall examine the case when images are corrupted by additive white noise. The main contribution of the paper is in the design and comparison of several adaptive $L$-filters for noise suppression in images. The properties of the developed adaptive $L$-filters are studied as well. Another primary goal is to establish links between the adaptive $L$-filters under study and other algorithms developed elsewhere. Two novel classes of adaptive LMS $L$-filters are discussed and a generalization to a signal-dependent adaptive filter structure is proposed. All the adaptive $L$-filters stem from the same algorithm, the *Least Mean Squares algorithm* that is used to minimize a cost function (e.g. Mean Squared Error, Mean Absolute Error, overall output power).

First, the location-invariant LMS $L$-filter for a nonconstant signal corrupted by zero-mean additive white noise is derived. Our interest in deriving the location-invariant LMS $L$-filter relies on the observation that this filter structure can be modified so as it performs even when a reference signal is not available. It is shown that the updating formula for the location-invariant LMS $L$-filter coefficients is the same with the one derived in the case of a constant signal corrupted by additive noise. Furthermore, we demonstrate that the derived constrained LMS $L$-filter can be described in terms of the general structure proposed by Griffiths and Jim [30]. Next, the normalized LMS $L$-filter is studied. Our motivation in studying the normalized LMS $L$-filter is justified for three reasons, namely: (i) This LMS variant provides a way to automate the choice for the step-size parameter in order to speed

of the input-signal statistics, as is in practice [38]. (iii) It is able to track the varying signal statistics [3]. The derivation of the majority of adaptive filter algorithms relies on the minimization of MSE criterion. Another optimization criterion that is encountered in image processing is the Mean Absolute Error (MAE) criterion. The so called signed error LMS $L$-filter that minimizes the MAE between the filter output and the desired response is derived. The signed-error LMS $L$-filter shares the benefits of the signed error LMS algorithm, i.e., the reduced numerical requirements due to the elimination of any multiplication in the coefficient updating equation. It is shown that for a certain choice of the step-size parameter the signed error LMS $L$-filter and the normalized LMS $L$-filter turn to be identical. Subsequently, a modified LMS $L$-filter with nonhomogeneous step-sizes is introduced in order to accelerate the rate of convergence of the adaptive $L$-filter by allowing the convergence of each $L$-filter coefficient to be controlled by a separate step-size parameter. Finally, a signal-dependent adaptive filter structure is developed aiming at a different treatment of the image pixels close to the edges from the pixels that belong to homogeneous regions. It is shown by experiments that the normalized LMS $L$-filter attains the best results with respect to both criteria. Furthermore, it is found that even better results are obtained by using the signal-dependent adaptive filter structure that employs two normalized LMS $L$-filters with different window sizes.

Throughout the paper two crucial aspects in the performance of an adaptive $L$-filter are highlighted and addressed properly:

1) *the selection of a step-size sequence* in a nonstationary environment (as is in image filtering) from a deterministic point of view (i.e., seeking to minimize always the a posteriori estimation error) without relying on the minimization of

[37, 38].

2) *what can be done when a reference image is not available.* Two solutions are proposed: (i) We show in theory that the location-invariant LMS $L$-filter (Section 2) can be modified so that it minimizes the overall output power. The variants of the unconstrained LMS $L$-filter cannot handle this case. (ii) We demonstrate by experiments that the location-invariant LMS $L$-filter coefficients are robust in the sense that they actually do not depend on the reference image which has been used in the training session. In the experimental results (Section 5) we show that the location-invariant $L$- filter coefficients that are determined at the end of a training session on a noisy version of "Trevor White" can be applied to a noisy version of "Lenna" without any significant performance deterioration. (iii) The requirement for the reference image is to be strongly correlated with the noise-free image that yields the noisy one which is filtered. In certain cases (e.g. in image sequences), it is reasonable to assume that one noise-free frame can act as a reference image for a number of image frames. In such a case, although the normalized LMS $L$-filter and the signal-dependent adaptive filter structure depend on the reference image, it is demonstrated (Section 5) that they can filter several noisy image frames using a single reference image without a great loss in the noise reduction capability of the filter (1.3 dB at most).

The outline of the paper is as follows. Section 2 is devoted to the design of the location-invariant $L$-filter. The normalized LMS $L$-filter, the signed error LMS $L$-filter and the modified LMS $L$-filter with nonhomogeneous step-sizes are studied in Section 3. The signal-dependent adaptive filter structure is described in Section 4. Experimental results are in-

## 2 Location-invariant LMS $L$-filter

In this section, we shall derive the updating formula for the location-invariant LMS $L$-filter in the case of a nonconstant signal corrupted by zero-mean additive white noise.

Let us consider that the observed image $x(\mathbf{k})$ can be expressed as a sum of an arbitrary image $s(\mathbf{k})$ plus zero-mean 2-d additive white noise, i.e.,

$$x(\mathbf{k}) = s(\mathbf{k}) + n(\mathbf{k}) \tag{1}$$

where $\mathbf{k} = (i, j)$ denotes the pixel coordinates. In image processing, 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 image pixel value $s(\mathbf{k})$ by minimizing a certain criterion. Among the several filter masks (e.g. cross, x-shape, square, circle) that are used in digital image processing [5], we shall rely on the square window of dimensions $\Xi \times \Xi$ where $\Xi$ is generally assumed to be an odd number, i.e., $\Xi = 2\xi + 1$:

$$\mathbf{X}(\mathbf{k}) = \begin{bmatrix} x(i-\xi, j-\xi) & x(i-\xi, j-\xi+1) & \dots & x(i-\xi, j+\xi) \\ \vdots & & & \vdots \\ x(i+\xi, j-\xi) & x(i+\xi, j-\xi+1) & \dots & x(i+\xi, j+\xi) \end{bmatrix}. \tag{2}$$

Let $N = \Xi^2$. Since we intend to apply a filter based on sample ordering let us rearrange the above $\Xi \times \Xi$ filter window in a lexicographic order (i.e., row by row) to a $N \times 1$ vector

$$\mathbf{x}(\mathbf{k}) = (x(i-\xi, j-\xi), x(i-\xi, j-\xi+1), \dots, x(i-\xi, j+\xi), \dots, x(i+\xi, j+\xi))^T. \tag{3}$$

We shall assume that the filter window is sliding over the image in a raster scan fashion. If

$$k = (i-1)\,K + j \qquad 1 \le i \le K \quad 1 \le j \le L \tag{4}$$

can be used instead of the pixel coordinates $\mathbf{k}$. Henceforth, a 1-D notation will be adopted for simplicity. Let $\mathbf{x}_r(k)$ be the ordered input vector at time pixel $k$ given by

$$\mathbf{x}_r(k) = \left(x_{(1)}(k), x_{(2)}(k), \ldots, x_{(N)}(k)\right)^T \tag{5}$$

where $x_{(i)}(k)$ denotes the $i$-th largest observation in the $N \times 1$ input vector (3). We are seeking the $L$-filter [7] whose output at $k$

$$y(k) = \mathbf{a}^T(k)\,\mathbf{x}_r(k) \tag{6}$$

minimizes the MSE

$$J(k) = \mathrm{E}\left[(y(k) - s(k))^2\right] = \mathrm{E}\left[s^2(k)\right] - 2\mathbf{a}^T(k)\,\mathbf{p}(k) + \mathbf{a}^T(k)\,\mathbf{R}(k)\,\mathbf{a}(k). \tag{7}$$

$\mathbf{R}(k) = \mathrm{E}\left[\mathbf{x}_r(k)\,\mathbf{x}_r^T(k)\right]$ is the correlation matrix of the observed ordered image pixel values and $\mathbf{p}(k) = \mathrm{E}\left[s(k)\,\mathbf{x}_r(k)\right]$ denotes the crosscorrelation vector between the ordered input vector $\mathbf{x}_r(k)$ and the desired image pixel value $s(k)$ subject to the constraint

$$\mathbf{1}_N^T\,\mathbf{a}(k) = 1. \tag{8}$$

In (8), $\mathbf{1}_N$ is the $N \times 1$ unitary vector, i.e., $\mathbf{1}_N = (1, 1, \ldots, 1)^T$. The dependence of the correlation matrix and the crosscorrelation vector on $k$ signifies that none assumption about stationarity has been made. The constraint (8) ensures that the filter preserves the zero-frequency or dc signals [7, 9]. Let $\nu = (N+1)/2$. By employing (8), we can partition the $L$-filter coefficient vector as follows

$$\mathbf{a}(k) = \left(\mathbf{a}_1^T(k) | a_\nu(k) | \mathbf{a}_2^T(k)\right)^T \tag{9}$$

$$\mathbf{a}_1(k) = (a_1(k), \ldots, a_{\nu-1}(k))^T \quad \mathbf{a}_2(k) = (a_{\nu+1}(k), \ldots, a_N(k))^T . \tag{10}$$

The coefficient for the median input sample is evaluated then by:

$$a_\nu(k) = 1 - \mathbf{1}_{\nu-1}^T \mathbf{a}_1(k) - \mathbf{1}_{\nu-1}^T \mathbf{a}_2(k). \tag{11}$$

In the sequel, we shall drop the subscript from the unitary vector of dimensions $(N-1)/2 \times 1$ for notation simplicity. Similarly, the ordered input vector can be rewritten as

$$\mathbf{x}_r(k) = \left(\mathbf{x}_{r1}^T(k) | x_{(\nu)}(k) | \mathbf{x}_{r2}^T(k)\right)^T . \tag{12}$$

Let $\mathbf{a}'(k)$ be the reduced $L$-filter coefficient vector

$$\mathbf{a}'(k) = \left(\mathbf{a}_1^T(k) | \mathbf{a}_2^T(k)\right)^T \tag{13}$$

and $\hat{\mathbf{x}}_r(k)$ be the $(N-1) \times 1$ vector

$$\hat{\mathbf{x}}_r(k) = \begin{bmatrix} \mathbf{x}_{r1}(k) - x_{(\nu)}(k)\mathbf{1} \\ \\ \mathbf{x}_{r2}(k) - x_{(\nu)}(k)\mathbf{1} \end{bmatrix} . \tag{14}$$

Following the analysis in [25] it can be proven that the LMS recursive relation for updating the reduced $L$-filter coefficient vector is given by

$$\hat{\mathbf{a}}'(k+1) = \hat{\mathbf{a}}'(k) + \mu\, \varepsilon(k)\, \hat{\mathbf{x}}_r(k) \tag{15}$$

where $\varepsilon(k)$ is the estimation error at pixel $k$, i.e., $\varepsilon(k) = s(k) - y(k)$. The recursive formula (15) is identical to the one reported in [25], but without invoking the assumption of a constant signal. Eq. (15) along with (11) constitute the *location-invariant LMS L-filter*. The convergence properties of the location-invariant LMS $L$-filter depend on the eigenvalue distribution of the correlation matrix of the vector $\hat{\mathbf{x}}_r(k)$.

Recently, several researchers have proposed adaptive algorithms for the design of $L$-filters that fulfill structural constraints such as the location-invariance that is discussed in this paper

has been a hot research topic in adaptive array beamforming for two decades [29]–[34]. Two are the main classes of constrained LMS algorithms, namely the gradient projection algorithm and Frost's algorithm. A generalized structure that includes Frost's algorithm as well as other linearly constrained adaptive algorithms has been proposed by Griffiths and Jim [30]. Therefore, the need emerges to connect the constrained adaptive $L$-filter structures with these main classes of linear constrained LMS algorithms. The algorithms reported in [24, 26, 27, 28] utilize Frost's algorithm. The possibility of using the gradient projection algorithm has also been proposed in [27]. Between those two, Frost's algorithm is preferable, because it does not accumulate errors as opposed to gradient projection algorithm [29, 30]. We shall demonstrate that the proposed algorithm (15) falls into Griffiths' and Jim's general structure shown in Figure 1. This general structure consists of two paths. In the upper path, the input observations are filtered by a fixed filter whose coefficients have a sum of unity, i.e.,

$$\mathbf{1}_N^T \mathbf{a}_c = 1. \tag{16}$$

In the lower path, the input observations are multiplied by matrix preprocessor $\mathbf{B}$ of dimensions $(N-1) \times N$ at most, and the resulted data are fed to an adaptive filter whose coefficients are updated according to the unconstrained LMS algorithm. In addition, the rows of matrix $\mathbf{B}$, $\mathbf{b}_m$, should sum up to zero, i.e.,

$$\mathbf{1}_N^T \mathbf{b}_m = 0 \qquad \forall m. \tag{17}$$

In our case, both the fixed filter as well as the adaptive one are $L$-filters that are driven by the observed image ordered pixel values. If we choose $\mathbf{a}_c$ to be a median filter, i.e.,

$$a_c(i) = \delta_{i\nu} \quad i = 1, \ldots, N \tag{18}$$

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & \dots & 0 & -1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & -1 & 0 & \dots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & \dots & 1 & -1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & -1 & 1 & \dots & 0 \\ \vdots & & & & & & & \vdots \\ 0 & 0 & \dots & 0 & -1 & 0 & \dots & 1 \end{bmatrix} \tag{19}$$

the location-invariant LMS $L$-filter results.

If a desired image is not available, (15) is modified as follows:

$$\hat{\mathbf{a}}'(k+1) = \hat{\mathbf{a}}'(k) - \mu\, y(k)\, \hat{\mathbf{x}}_r(k) \tag{20}$$

and the resulting filter minimizes the overall output power

$$J(k) = \mathbf{a}^T(k)\mathbf{R}(k)\mathbf{a}(k) \tag{21}$$

subject to (8).

# 3 Variants of the unconstrained LMS $L$-filter

In this section, we deal with the unconstrained LMS adaptive $L$-filter [19]–[21] whose coefficients are updated by using the following recursive formula:

$$\hat{\mathbf{a}}(k+1) = \hat{\mathbf{a}}(k) + \mu\, \varepsilon(k)\, \mathbf{x}_r(k). \tag{22}$$

Three modifications of the unconstrained LMS adaptive $L$-filter are discussed, namely, the normalized LMS $L$-filter, the signed error LMS $L$-filter and the modified LMS $L$-filter with nonhomogeneous step-sizes.

A difficult problem frequently met in the design of adaptive filters that are based on the LMS algorithm, such as the algorithm (22), is the selection of the step-size parameter $\mu$. Although in theory sufficient conditions on $\mu$ exist that guarantee the convergence of the LMS algorithm, these conditions depend on the knowledge of the eigenvalues of the correlation matrix $\mathbf{R}$ [2, 35, 36]. In the case of the LMS adaptive $L$-filter (22), $\mathbf{R}$ is the correlation matrix of the ordered input vector. For example, the necessary and sufficient condition for the average $L$-filter coefficient vector $\mathrm{E}[\hat{\mathbf{a}}(k)]$ to be convergent is

$$0 < \mu < \frac{2}{\lambda_{\max}} \tag{23}$$

where $\lambda_{\max}$ denotes the maximal eigenvalue of matrix $\mathbf{R}$ [2]. Furthermore, $\mu$ should satisfy the following more strict condition

$$0 < \mu < \frac{2}{3\,\mathrm{tr}[\mathbf{R}]} = \frac{2}{3 \times \text{ total input power}} \tag{24}$$

in order to achieve convergence of the average Mean-Squared Error $\mathrm{E}[J(k)]$ to a steady state value [36]. In (24), $\mathrm{tr}[\cdot]$ stands for the trace of the matrix inside brackets. In addition, when the adaptive filter is going to operate in a nonstationary environment (as is in image processing), inequalities (23) and (24) turn to be useless, since the correlation matrix $\mathbf{R}$ is time/space varying. In such cases, it is reasonable to employ a time/space varying step-size parameter $\mu(k)$. Let us evaluate the a posteriori estimation error at pixel $k$, defined as follows

$$\varepsilon'(k) = s(k) - \hat{\mathbf{a}}^T(k+1)\mathbf{x}_r(k). \tag{25}$$

It can easily be shown that

$$\varepsilon'(k) = \varepsilon(k)\left(1 - \mu(k)\,\mathbf{x}_r^T(k)\mathbf{x}_r(k)\right). \tag{26}$$

$$\mu(k) = \frac{1}{\mathbf{x}_r^T(k)\mathbf{x}_r(k)} = \frac{1}{||\mathbf{x}_r(k)||^2} \tag{27}$$

then, $\varepsilon'(k)$ becomes zero. A step-size sequence of the form (27) motivated us to modify (22) as follows:

$$\hat{\mathbf{a}}(k+1) = \hat{\mathbf{a}}(k) + \frac{\mu_0}{||\mathbf{x}_r(k)||^2}\,\varepsilon(k)\,\mathbf{x}_r(k). \tag{28}$$

The recursive equation (28) describes the adaptation of the coefficients of the normalized LMS $L$-filter. It is equivalent to the linear normalized LMS algorithm [3]. The only difference is that (28) employs the vector of the ordered observations $\mathbf{x}_r(k)$ to update the adaptive $L$-filter coefficients, whereas the linear normalized LMS algorithm employs the input vector $\mathbf{x}(k)$ given by (3). Eq. (28) may be interpreted as an ordinary LMS updating formula that operates on the normalized ordered noisy-input observations $\mathbf{x}_r(k)/||\mathbf{x}_r(k)||$. Therefore according to (24) $\mu_0$ should be chosen to satisfy the inequality

$$0 < \mu_0 \leq \frac{2}{3}, \tag{29}$$

since we deal now with the correlation matrix of the normalized ordered input observations given by

$$\hat{\mathbf{R}} = \mathrm{E}\left[\frac{\mathbf{x}_r(k)\mathbf{x}_r^T(k)}{||\mathbf{x}_r(k)||^2}\right] \simeq \frac{1}{\mathrm{tr}[\mathbf{R}]}\,\mathbf{R} \tag{30}$$

whose trace is equal to 1. In practice, it has been found that (29) is still conservative, because the best results have been obtained for $\mu_0 = 0.8$, as can be seen in Section 5. However, if $\mu_0$ is chosen to be equal to 2/3, we are still close to the optimum.

## 3-B Signed error LMS $L$-filter

Another criterion that is frequently encountered in nonlinear image processing is the Mean Absolute Error (MAE) criterion (also called Least Mean Absolute Value criterion [3]). The

timal MAE estimator of location for the double-exponential distribution (i.e., the Laplacian distribution) [5]. In the sequel, we shall derive the LMS adaptive $L$-filter that is optimum under the MAE criterion.

The MAE criterion to be minimized is defined as

$$J'(k) = \mathrm{E}\left[|s(k) - y(k)|\right]. \tag{31}$$

In this case, the coefficient vector $\mathbf{a}(k)$ must be updated at $k$ so that (31) is minimized. The method of steepest-descent yields the following recursive relation for updating the filter coefficients:

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \mu\ \left[-\nabla J'(k)\right] \tag{32}$$

where

$$\nabla J'(k) = \frac{\partial J'(k)}{\partial \mathbf{a}(k)} = \frac{\partial}{\partial \mathbf{a}(k)}\,\mathrm{E}\left[|s(k) - \mathbf{a}^T(k)\mathbf{x}_r(k)|\right]. \tag{33}$$

An unbiased estimate for the gradient $\nabla J'(k)$ can be obtained, if we drop out the expectation operator, i.e.,

$$\hat{\nabla} J'(k) = \frac{\partial}{\partial \mathbf{a}(k)}\,|s(k) - \mathbf{a}^T(k)\mathbf{x}_r(k)| = -\mathrm{sgn}\left[s(k) - \mathbf{a}^T(k)\mathbf{x}_r(k)\right]\mathbf{x}_r(k) \tag{34}$$

where $\mathrm{sgn}\,[\cdot]$ denotes the sign of the bracketed expression:

$$\mathrm{sgn}\,[x] = \begin{cases} 1 & \text{if } x > 0 \\ \\ -1 & \text{if } x < 0. \end{cases} \tag{35}$$

Therefore, the updating formula for the coefficients of the signed error LMS $L$-filter is

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \mu\ \mathrm{sgn}\,[\varepsilon(k)]\ \mathbf{x}_r(k). \tag{36}$$

algorithm [3]. The only difference is that (36) employs the vector of the ordered observed image pixel values instead of the input vector itself.

The step-size parameter can be chosen as follows. The a posteriori estimation error at $k$ is simply given by:

$$\varepsilon'(k) = s(k) - \hat{\mathbf{a}}^T(k+1)\mathbf{x}_r(k) = \varepsilon(k) - \mu \, \mathrm{sgn}\,[\varepsilon(k)] \; \mathbf{x}_r^T(k)\mathbf{x}_r(k). \tag{37}$$

By employing the following identity

$$\mathrm{sgn}\,[\varepsilon(k)] = \frac{\varepsilon(k)}{|\varepsilon(k)|} \tag{38}$$

the a posteriori estimation error is rewritten as

$$\varepsilon'(k) = \varepsilon(k) \left\{ 1 - \frac{\mu(k)}{|\varepsilon(k)|} \, \mathbf{x}_r^T(k)\mathbf{x}_r(k) \right\}. \tag{39}$$

It can easily be seen that in order $|\varepsilon'(k)|$ to be less than $|\varepsilon(k)|$, the step-size sequence should be chosen so that

$$0 < \mu(k) < \frac{2 \; |\varepsilon(k)|}{\mathbf{x}_r^T(k)\mathbf{x}_r^T(k)} \tag{40}$$

If the step-size sequence $\mu(k)$ is chosen as

$$\mu(k) = \frac{\mu_0 |\varepsilon(k)|}{\mathbf{x}_r^T(k)\mathbf{x}_r(k)} \qquad 0 < \mu_0 < 1 \tag{41}$$

then $|\varepsilon'(k)|$ will be $(1 - \mu_0)|\varepsilon(k)|$, i.e., the a posteriori absolute estimation error is smaller that the (a priori) absolute estimation error. It is worth noting that by substituting (41) into (36), the updating formula for the normalized LMS $L$-filter coefficients is obtained. It is true that one cannot argue that the normalized LMS $L$-filter minimizes the MAE criterion. Recall that normalized LMS $L$-filter design is based on the minimization of the MSE. The above-described discussion can serve only as an indication to explain the very good performance of this filter with respect to MAE criterion in the experiments reported in Section 5.

It is well known that the rate of convergence of the LMS algorithm is one order of magnitude slower that that of other adaptive algorithms (e.g. the Recursive Least-Squares or the Recursive Least-Squares Lattice algorithm) [2]. This slow rate of convergence may be attributed to the fact that only one parameter, the step-size $\mu$, controls the convergence of all the filter coefficients. On the contrary, in the case of the Recursive Least-Squares algorithm, the convergence of each filter coefficient is controlled by a separate element of the Kalman gain vector. In addition, at each time instant the Kalman gain vector is updated utilizing all the information contained in the input data, extending back to the instant of time when the algorithm is initiated. This observation motivated us to employ different step-size parameters for the various LMS $L$-filter coefficients in the recursive relation that describes their adaptation. We have found by experiments that the following step-size sequence

$$\mu_i(k) = \mu_0 \frac{\sum_{j=0}^{k} x_{(i)}(k-j)}{\sum_{j=0}^{k} x_{(1)}(k-j)} \tag{42}$$

gives results comparable to those obtained by using the normalized LMS $L$-filter algorithm. By using (42), the modified LMS $L$-filter updating formula is written as follows:

$$\mathbf{a}(k+1) = \mathbf{a}(k) + \varepsilon(k)\, \mathbf{M}(k)\mathbf{x}_r(k) \tag{43}$$

where $\mathbf{M}(k)$ denotes the following diagonal matrix:

$$\mathbf{M}(k) = \mathrm{diag}\left[\mu_1(k), \mu_2(k), \ldots, \mu_N(k)\right]. \tag{44}$$

It can be shown that the selection of the step-size parameters according to (42) guarantees a smaller a posteriori estimation error for every $k$. Indeed, since for non-negative input observations

$$x_{(i)}(l) \geq x_{(1)}(l) \qquad \forall l \tag{45}$$

$$\varepsilon'(k) = \varepsilon(k)\left\{1 - \sum_{i=1}^{N} \mu_i(k) x_{(i)}^2(k)\right\} < \varepsilon(k)\left\{1 - \mu_0 \sum_{i=1}^{N} x_{(i)}^2(k)\right\}. \tag{46}$$

Therefore, if $\mu_0$ is chosen so that

$$0 < \mu_0 < \frac{1}{\max_k\{\sum_{i=1}^{N} x_{(i)}^2(k)\}} = \frac{1}{\text{peak input power}} \tag{47}$$

$\varepsilon'(k) < \varepsilon(k)$. Another property that the modified LMS $L$-filter (43) possesses is drawn from (46). The right-hand side of the inequality in (46) is identified to be the a posteriori error at $k$ of an LMS $L$-filter that uses a constant step-size parameter $\mu = \mu_0$. Thus, the modified LMS $L$-filter (43) produces always a smaller a posteriori estimation error than the ordinary LMS $L$-filter that employs a constant step-size $\mu = \mu_0$.

# 4 Signal-dependent adaptive $L$-filters

In the sequel, we shall describe a signal-dependent adaptive $L$-filter structure that adjusts its smoothing properties at each point according to the local image content in order to achieve edge preservation as well as maximum noise suppression in homogeneous regions. The signal-dependent adaptive $L$-filter structure consists of two LMS adaptive $L$-filters whose outputs $y_L(k)$ and $y_H(k)$ are combined to give the final response as follows:

$$y(k) = y_L(k) + \beta(k)\{y_H(k) - y_L(k)\} = \beta(k) y_H(k) + [1 - \beta(k)]\, y_L(k) \tag{48}$$

where $\beta(k)$ is a signal-dependent weighting factor. $\beta(k)$ can be chosen so as it minimizes the Mean-Squared Error between the filter output $y(k)$ given by (48) and the desired response $s(k)$ [14]. Let $\mathbf{a}_H(k)$ and $\mathbf{a}_L(k)$ denote the $L$-filter coefficient vectors driven by the high frequency and the low-frequency data respectively. It can easily be shown that $\mathrm{E}\left[(s(k) - y(k))^2\right]$ is minimized for

$$\beta(k) = \frac{(\mathbf{R}(k)\mathbf{a}_L(k) - \mathbf{p}(k))^T (\mathbf{a}_H(k) - \mathbf{a}_L(k))}{(\mathbf{a}_H(k) - \mathbf{a}_L(k))^T \mathbf{R}(k) (\mathbf{a}_H(k) - \mathbf{a}_L(k))}. \tag{49}$$

as follows. In homogeneous regions, we should have

$$\mathbf{R}(k)\mathbf{a}_L(k) = \mathbf{p}(k) \tag{50}$$

i.e., $\mathbf{a}_L(k)$ is the optimal $L$-filter that minimizes the MSE between the filter output and the desired response. Therefore, $\beta(k)$ equals 0. Similarly, close to the edges we should have

$$\mathbf{R}(k)\mathbf{a}_H(k) = \mathbf{p}(k). \tag{51}$$

By replacing (51) into (49) it is seen that $\beta(k)$ equals 1. In order to avoid the estimation of the correlation matrix $\mathbf{R}(k)$ and the crosscorrelation vector $\mathbf{p}(k)$, we replace the optimal signal-dependent weighting factor (49) by another one that shares the same favorable properties, namely, by the local Signal-to-Noise Ration measure

$$\beta(k) = 1 - \frac{\sigma_n^2}{\hat{\sigma}_x^2(k)} \tag{52}$$

where $\sigma_n^2$ is the noise variance and $\hat{\sigma}_x^2(k)$ is the variance of the noisy input observations. The adaptive $L$-filters $\mathbf{a}_L(k)$ and $\mathbf{a}_H(k)$ may use different window sizes. In such a case, the coefficient $\beta(k)$ given by (52) can be used as a signal-dependent switch between the two LMS adaptive $L$-filters, i.e.,

$$y(k) = \begin{cases} y_H(k) & \text{if } \beta(k) > \beta_t \\ \\ y_L(k) & \text{otherwise} \end{cases} \tag{53}$$

where $0 < \beta_t < 1$ is a threshold that determines a trade-off between noise suppression and edge preservation. In the reported experiments $\beta_t$ is chosen to be either 0.6 or 0.75. As far as the threshold $\beta_t$ is concerned, a same parameter is also found in Signal-Adaptive Median filters [14] and a selection of a proper value does not add any difficulty.

We shall present five sets of experiments in order to assess the performance of the adaptive $L$-filters we have discussed so far. All the sets of experiments have been conducted on images. Four of these sets of experiments presuppose that a reference image (e.g. the original image) is available. In practice, reference images are usually transmitted through TV telecommunication channels to measure the performance of the channel. In such cases, the proposed adaptive $L$-filters can be proven very useful, if the design of an optimal filter for the specific channel characteristics is required. However, in certain cases (e.g. in image sequences) it is reasonable to assume that one noise-free frame can act as a reference image for a number of image frames. The first set of experiments deals with the second frame of the image sequence called "Trevor White". Its purpose is to demonstrate the superiority of adaptive $L$-filters over the adaptive linear filters in noise removal applications. The second set of experiments has been conducted on "Lenna". Our goal is to compare the performance of the adaptive $L$-filters under study. In the third set of experiments, we treat the case where a reference image is not available. Next, we investigate the ability of the proposed adaptive $L$-filters to cope with the nonstationarity in image data. Finally, we study the performance of the proposed filters in environments with time-varying statistics (e.g. image sequence filtering).

To begin with, let us describe the experimental set-up and the quantitative criteria we shall use in the experiments. Both the linear and the $L$-filter coefficients have been randomly initialized in the interval (0,1) and they have been normalized by their sum so as their sum equals unity. The entire noisy image has been used in the training of the adaptive filter. A single run on the training noisy image has been performed. In the first three sets of experiments, the coefficients derived during the operation of the algorithm on the last image

dependent filter structure, the coefficients $\mathbf{a}_L(k)$ and $\mathbf{a}_H(k)$ that have been derived during the operation of the algorithm throughout the entire image have been averaged and have been applied subsequently to filter the pixel blocks that fall into the homogeneous regions or close to the image edges. In the remaining two sets of experiments, the $L$-filter coefficients determined recursively by the adaptive algorithm at each image pixel have been used to filter the noisy input image. Two criteria have been employed, namely, the noise reduction index (NR) defined as the ratio of the mean output noise power to the mean input noise power, i.e.,

$$\text{NR} = 10 \log \frac{\frac{1}{KL} \sum_{i=1}^{K} \sum_{j=1}^{L} (y(i,j) - s(i,j))^2}{\frac{1}{KL} \sum_{i=1}^{K} \sum_{j=1}^{L} (x(i,j) - s(i,j))^2} \quad \text{(in dB)} \tag{54}$$

and the Mean Absolute Error Reduction (MAER) defined as the ratio of the mean absolute error in the output to the mean absolute error in the input, i.e.,

$$\text{MAER} = 20 \log \frac{\frac{1}{KL} \sum_{i=1}^{K} \sum_{j=1}^{L} |y(i,j) - s(i,j)|}{\frac{1}{KL} \sum_{i=1}^{K} \sum_{j=1}^{L} |x(i,j) - s(i,j)|} \quad \text{(in dB)}. \tag{55}$$

In (54) and (55), $s(i,j)$ is the original image pixel, $x(i,j)$ denotes the same image pixel corrupted by noise and $y(i,j)$ is the filter output at the same image pixel. $K, L$ are the number of image rows and columns respectively.

Figure 2a shows the second frame of "Trevor White". We shall compare the performance of the normalized LMS $L$-filter to the one of the normalized LMS linear filter when the original image is corrupted by

(a) zero-mean additive white Gaussian noise having standard deviation $\sigma_n$=20,

(b) impulsive noise with probability $p = 10\%$ (both positive and negative impulses with equal probability), and,

(c) mixed impulsive ($p = 10\%$) and additive Gaussian white noise ($\sigma_n = 20$).

small variance (e.g. $\sigma_n = 1$) in case (b) in order to avoid instabilities. Table 1 summarizes the noise reduction index achieved in all the above-mentioned cases. It is seen that the normalized LMS $L$-filter outperforms its linear counterpart for all noise models. Even for Gaussian noise, it performs slightly better than the adaptive linear filter. The same superior performance has been obtained in a variety of experiments employing different original images (such as "Lenna", "Baboon", "Car" etc.). That is, the NR index achieved is not image-dependent. Due to lack of space, only one experiment is reported. Therefore, its use for noise suppression in image processing is fully justified. The original image corrupted by mixed impulsive and Gaussian noise is shown in Figure 2b. The output of the normalized LMS $L$-filter of dimensions $3 \times 3$ may be found in Figure 2c. For comparison purposes, the output of the normalized LMS linear filter of the same dimensions is shown in Figure 2d. The superiority of the adaptive nonlinear filter is self-evident. Similar conclusions can also be drawn if a modified LMS $L$-filter with nonhomogeneous step-sizes is employed. What remains is to test the performance of the described LMS $L$-filters, both constrained and unconstrained ones, against that of the median filter that is a straightforward choice in the context of nonlinear image processing. To that end, we proceed to the description of the second set of experiments.

Figure 4a shows another original image frequently encountered in image processing literature, the so-called "Lenna". We shall consider the following two cases:

(a) severely corrupting the original image by adding zero-mean additive white Gaussian noise of standard deviation $\sigma_n = 50$ plus impulsive noise with probability $p = 10\%$,

(b) corrupting the original image by adding only Gaussian noise with standard de-

The NR as well as the MAER achieved by the location-invariant LMS $L$-filter, the normalized LMS $L$-filter and the modified LMS $L$-filter with nonhomogeneous step-sizes, all of dimensions $3 \times 3$ in the case of mixed impulsive and additive Gaussian noise are listed in Table 2. In the same table, we have also included the corresponding figures of merit for the $3 \times 3$ median filter. As can be seen, the condition for location-invariant estimation is strict enough and the resulting adaptive $L$-filter is only 1 dB better than the median filter with respect to both quantitative measures. The modified LMS $L$-filter with nonhomogeneous step-sizes is the second best adaptive $L$-filter yielding an almost 2.5 dB better NR and MAER compared to the median filter. The normalized LMS $L$-filter achieves the best performance both in the MSE sense as well as in MAE sense. Especially for the MAE criterion, this performance was expected, since we have already demonstrated the connection between the signed error LMS $L$-filter and the normalized one. The optimal value of parameter $\mu_0$ has been found experimentally. Figure 3 shows the NR and the MAER achieved for several values of the parameter $\mu_0$. It can be seen that the optimal value of $\mu_0$ is different for the two figures of merit. In our experiments we have used the value of $\mu_0$ for which NR attains a minimum, i.e., $\mu_0 = 0.8$.

In addition, we have tested the performance of two signal-dependent adaptive $L$-filter structures. The first one uses two $3 \times 3$ adaptive $L$-filters that are trained by different regions of the corrupted input image. More specifically, the pixels that belong to the homogeneous image regions are used to adapt the coefficients of the one adaptive $L$-filter, while those that are close to the image edges are used to adapt the coefficients of the second adaptive $L$-filter. Any of the adaptive $L$-filters that have been described in this paper (e.g. the location-invariant LMS $L$-filter, the normalized LMS $L$-filter or the modified LMS $L$-filter

experiments described, we have used the normalized LMS $L$-filter. By inspecting Table 2, it is seen that the signal-dependent adaptive $L$-filter structure provides the best results, when the window of the normalized LMS $L$-filter that is used in homogeneous image regions is of larger dimensions (e.g. $5 \times 5$) than that of the adaptive LMS $L$-filter that is trained by pixels close to the image edges. The observed superior performance is due to the larger window size that is used to filter the noise in homogeneous regions. In such a case, the local SNR measure (52) is computed twice by calculating the local variance of the noisy input observations for both window sizes. An image pixel is declared as an edge pixel if either $\beta_{3\times3}(k)$ or $\beta_{5\times5}(k)$ exceeds the threshold. For the sake of completeness, it should also be noted that when the image corruption is not too severe, the impulse detection and removal mechanism described in [14] has also to be included in the filter structure in order not to misinterpret pixels that pass the test (53) as edges although they are actually impulses.

The performance of the proposed adaptive $L$-filter has also been tested when the original image is corrupted by zero-mean additive white Gaussian noise with $\sigma_n = 50$. The noise reduction index and the reduction in MAE achieved in the output of the filters under study is summarized in Table 3. The performance of the arithmetic mean filter is also included in Table 3 for comparison purposes. By inspecting Table 3, it is seen that the location-invariant LMS $L$-filter has an almost identical performance with the arithmetic mean filter and it is much better than the median filter. This was expected, because the location-invariant LMS $L$-filter is able to adapt to the noise distribution. The performance of the modified LMS $L$-filter with nonhomogeneous step-sizes approximates that of the normalized LMS $L$-filter. The later is proven to be an almost 1 dB superior to the performance of the arithmetic mean filter and 2.5 dB superior to the performance of the median filter. The signal-dependent filter structure with a different window size in homogeneous regions and close to the edges

adaptive $L$-filters may be attributed to their slightly better performance close to the image edges than that of the arithmetic mean filter.

The original image corrupted by mixed impulsive and additive Gaussian noise is shown in Figure 4b. The output of the location-invariant LMS $L$-filter of dimensions $3 \times 3$ may be found in Figure 4c. The filtered image by using the modified LMS $L$-filter with nonhomogeneous step-sizes and the output of the normalized LMS $L$-filter are shown in Figures 4d and 4e respectively. The best result obtained, i.e., the output of the signal-dependent normalized LMS $L$-filter structure with a different window size in homogeneous regions and close to the edges, is shown in Figure 4f.

When a reference image is not available, the straightforward choice is to apply the location-invariant LMS $L$-filter, because only that design can be modified to work without a reference signal, as has already been discussed in Section 2. In Table 4, the noise reduction index achieved at the output of the modified location-invariant adaptive $L$-filter (20) is tabulated, when the original image "Lenna" is corrupted by mixed impulsive and additive Gaussian noise as well as by additive Gaussian noise. In parentheses, we have included the same figures of merit for the median filter and the arithmetic mean filter respectively. The step-size parameter has been chosen according to

$$\mu(k) = \frac{\mu_0}{||\mathbf{x}_r(k)||^2} \quad ; \quad \mu_0 = 1 \times 10^{-10}. \tag{56}$$

It is seen that the modified location-invariant LMS $L$-filter (20) is able to adapt to the noise distribution and to attain an almost identical performance to that of the maximum likelihood estimator of location for each noise model. Furthermore, we have tested the robustness of the $L$-filter coefficients that are determined at the end of a training session and are applied to filter a noisy image which has been produced by corrupting a different

location-invariant LMS $L$-filtering algorithm (15) on "Trevor White" corrupted by impulsive ($p = 10\%$) and additive Gaussian noise ($\sigma_n = 50$) using as a reference image the original "Trevor White". Subsequently, we have averaged the $L$-filter coefficients that were derived during the operation of the algorithm on the last image row. The resulting $L$-filter coefficients have been applied to filter the original image "Lenna" corrupted by the same mixed impulsive and additive Gaussian noise, as before. It has been found that only the location-invariant $L$-filter (15) coefficients are robust, in the sense that they do not depend on the reference image which has been used in the training session. In Table 4, we have tabulated the noise reduction achieved by filtering the original image "Lenna" by using the $L$-filter coefficients determined at the end of a training session on "Trevor White" and vice versa. It is seen that the attained noise reduction is close to the one achieved when the reference image is available.

Next, we study the ability of the proposed adaptive $L$-filters to cope with the nonstationarity in the image data. The experiment will be focused on the performance of the normalized LMS $L$-filter and the signal-dependent normalized LMS $L$-filter structure. Both these structures have been applied to "Lenna" that has been severely corrupted by adding zero-mean additive white Gaussian noise of standard deviation $\sigma_n = 50$ plus impulsive noise with probability $p = 10\%$. We suppose that the noise-free image is available and is used as reference image. The $L$-filter coefficients determined recursively by the adaptive algorithm at each image pixel have been used to filter the noisy image. Figure 5a shows image row #128 of the noisy input image. To facilitate the exposition we have plotted $s(128, j)$ in all related figures. For comparison purposes, row #128 of the median filtered image is shown in Figure 5b. It is seen that the noise is smoothed enough. However, the edge jitter introduced by a median filter of dimensions $3 \times 3$ is clear. A further noise reduction is achieved by

that both the normalized LMS $L$-filter and the signal-dependent adaptive $L$-filter structure introduce less jitter than the median filter. By comparing Figures 5b – 5d, it is seen that the best results are obtained by the signal-dependent adaptive filter structure.

The final set of experiments aims at demonstrating the performance of the proposed algorithms when the statistics of the input image are time-varying. Towards this end, we have tried to filter a noisy version of the 2nd frame in image sequence "Trevor White" by using as reference image the 2nd (noise-free) frame itself or the 5th frame or the 9th frame. Obviously, both the 5th and the 9th frames differ from the 2nd frame that is being filtered. Table 5 summarizes the performance of the different filters. As expected, when the image which is being filtered results by corrupting an original image that is different from the reference image which is used in the adaptive algorithm, the performance of the normalized LMS $L$-filter and the signal-dependent normalized LMS $L$-filter structure is degraded. However, the loss in performance is small (i.e., 1.3 dB at most) and their NR and MAER indices are still the highest. It is also seen that the performance of the location-invariant LMS $L$-filter is almost the same independently of the reference image which has been used in the adaptive algorithm. This result is in par with the discussion made in the third experiment. Moreover, the ability of the signal-dependent adaptive $L$-filter in tracking a time-varying edge is shown in Figure 6. In Figure 6, we focus on a part of row #128 in the 2nd and 9th frames of "Trevor White". Due to motion, a different edge appears at the two time instants. Although the signal-dependent normalized LMS $L$-filter has been trained using as reference an edge different from the one that appears in the noisy input image, it tends to track the later one demonstrating its ability to track slow time-variations in the characteristics of the input signal.

In this paper, we have described several adaptive LMS $L$-filters both constrained and unconstrained ones, and we have compared their efficiency in noise suppression in images. We have shown that the updating formula for the location invariant LMS $L$-filter that had previously been derived in the case of a constant signal corrupted by zero-mean additive white noise is still valid for nonconstant signals. We have demonstrated that the derived location-invariant LMS $L$-filter can be described in terms of the generalized linearly constrained adaptive processing structure proposed by Griffiths and Jim. We have shown that under a certain choice for the step-size parameter, the sign and the normalized LMS $L$-filters turn to be identical. The later property explains the excellent performance of the normalized LMS $L$-filter with respect to the Mean Absolute Error criterion. A modified LMS $L$-filter with nonhomogeneous step-sizes has also been proposed. It has been found that it gives comparable results to the normalized LMS $L$-filter. Finally, a signal-dependent adaptive $L$-filter structure has been developed to allow a separate treatment of the edges and the homogeneous image regions. By experiments, it has been shown that the signal-dependent filter structure which employs two normalized LMS $L$-filters with different window sizes yields the best results.

# References


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

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

[3] M. Bellanger, *Adaptive Digital Filters and Signal Analysis*. New York: Marcel Dekker, 1987.

*Processing Algorithms*. London, U.K.: Prentice Hall International, 1993.

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

[6] 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.

[7] 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. 1342–1349, December 1983.

[8] F. Palmieri, and C.G. Boncelet, Jr., "Optimal MSE Linear Combination of Order Statistics for Restoration of Markov Processes," in *Proc. 20th Ann. Conf. Info. Sci. Syst.*, Princeton, N.J., 1986.

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

[10] C. Kotropoulos, and I. Pitas, "Optimum nonlinear signal detection and estimation in the presence of ultrasonic speckle," *Ultrasonic Imaging*, vol. 14, no. 3, pp. 249–275, July 1992.

[11] 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.

[12] J.S. Lee, "Digital image enhancement and filtering by use of local statistics," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. PAMI-2, pp. 165–168, March 1980.

for images with signal-dependent noise," *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol. PAMI-7, no. 2, pp. 165–177, March 1985.

[14] R. Bernstein, "Adaptive nonlinear filters for simultaneous removal of different kinds of noise in images," *IEEE Trans. on Circuits and Systems*, vol. CAS-34, no. 11, pp. 1275–1291, November 1987.

[15] X.Z. Sun, and A.N. Venetsanopoulos, "Adaptive schemes for noise filtering and edge detection by use of local statistics," *IEEE Trans. on Circuits and Systems*, vol. CAS-35, no. 1, pp. 57–69, January 1988.

[16] T. Song, M. Gabbouj, and Y. Neuvo, "Adaptive $L$-filters based on local statistics," in *Proc. of the IEEE Winter Workshop on Nonlinear Digital Signal Processing*, pp. 3.2-2.1–3.2-2.4, Tampere, Finland, 1993.

[17] C. Kotropoulos, and I. Pitas, "Adaptive nonlinear filters for digital signal/image processing," in *Control and Dynamic Systems* (C.T. Leondes, ed.), vol. 67, pp. 263–318, San Diego, CA: Academic Press, 1994.

[18] F. Palmieri, and C.G. Boncelet, Jr., "A class of adaptive nonlinear filters," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 1483–1486, New York, 1988.

[19] I. Pitas, and A.N. Venetsanopoulos, "Adaptive $L$-filters," in *Proc. of the European Conf. on Circuit Theory and Design*, Brighton, England, 1989.

[20] I. Pitas, and A.N. Venetsanopoulos, "LMS and RLS adaptive $L$-filters," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 1389–1392, Albuquerque, 1990.

*Trans. on Signal Processing*, vol. 39, no. 2, pp. 518–522, February 1991.

[22] C. Kotropoulos, and I. Pitas, "Constrained adaptive LMS $L$-filters," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 1665–1668, Toronto, Canada, 1991.

[23] C. Kotropoulos, and I. Pitas, "Design of the constrained adaptive LMS $L$-filters," in *Proc. of the IEEE Mediterranean Electrotechnical Conf. (MELECON '91)*, Ljubliana, Slovenia, 1991.

[24] S. Roy, "On $L$-filter design using the gradient search algorithm," in *Proc. SPIE Nonlinear Image Processing II*, vol. 1451, pp. 254–256, 1991.

[25] C. Kotropoulos, and I. Pitas, "Constrained adaptive LMS $L$-filters," *Signal Processing*, vol. 26, no. 3, pp. 335–358, 1992.

[26] P.M. Clarkson, and G.A. Williamson, "Constrained adaptive order statistic filters for minimum variance signal estimation," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. IV, pp. 253–256, San Francisco, 1992.

[27] Q. Cai, W. Yang, Z. Liu, and D. Gu, "Recursion adaptive order statistics filters for signal restoration," in *Proc. of the IEEE Int. Symp. on Circuits and Systems*, pp. 129–132, San Diego, California, 1992.

[28] G.A. Williamson, and P.M. Clarkson, "On signal recovery with adaptive order statistic filters," *IEEE Trans. on Signal Processing*, vol. SP-40, no. 10, pp. 2622–2626, October 1992.

*ceedings of the IEEE*, vol. 60, pp. 926–935, 1972.

[30] L.J. Griffiths, and C.W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. on Antennas and Propagation*, vol. AP-30, no. 1, pp. 27–34, January 1982.

[31] M.J. Rude, and L.J. Griffiths, "A linearly constrained adaptive algorithm for constant modulus signal processing," in *Signal Processing V: Theories and Applications* (L. Torres, E. Masgrau and M.A. Lagunas, eds.), pp. 237–240, 1990.

[32] C.Y. Tseng, and L.J. Griffiths, "On the implementation of adaptive filters with adjustable linear constraints," in *Proc. of the IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 1449–1452, Albuquerque, 1990.

[33] L. Castedo, C.Y. Tseng, and L.J.Griffiths, "An adjustable constraint approach for robust adaptive beamforming," in *Signal Processing VI: Theories and Applications* (J. Vandewalle, R. Boite, M. Moonen and J. Oosterlinck, eds.), pp. 1121–1124, 1992.

[34] C.Y. Tseng, and L.J. Griffiths, "A unified approach to the design of linear constraints in minimum variance adaptive beamformers," *IEEE Trans. on Antennas and Propagation*, vol. 40, no. 12, pp. 1533–1542, December 1992.

[35] L.L. Horowitz, and K.D. Senne, "Performance advantage of complex LMS for controlling narrow-band adaptive arrays," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-29, no. 3, pp. 722–736, June 1981.

[36] A. Feuer, and E. Weinstein, "Convergence analysis of LMS filters with uncorrelated Gaussian data," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 1, pp. 222–230, February 1985.

*Acoustics Speech and Signal Processing*, vol. ASSP-35, no. 7, pp. 1065-1068, July 1987.

[38] D. T. M. Slock, "On the convergence behavior of the LMS and the normalized LMS algorithms," *IEEE Trans. on Signal Processing*, vol. 41, no. 3, pp. 2811-2825, September 1993.

Table 1 Noise reduction (in dB) achieved by the normalized LMS $L$-filter and its linear counterpart in the restoration of "Trevor White".

Table 2 Noise reduction and Mean Absolute Error reduction (in dB) achieved by various $L$-filters in the restoration of "Lenna" corrupted by mixed impulsive and additive Gaussian noise.

Table 3 Noise reduction and Mean Absolute Error reduction (in dB) achieved by various $L$-filters in the restoration of "Lenna" corrupted by additive Gaussian noise.

Table 4 Noise reduction (in dB) achieved by the location-invariant adaptive $L$-filter when a reference image is not available.

Table 5 Noise reduction and Mean Absolute Error reduction (in dB) achieved by various $L$-filters in smoothing the second frame of sequence "Trevor White" that has been corrupted by mixed impulsive and additive Gaussian noise for several reference images.

Figure 1 Generalized linearly constrained adaptive processing structure.

Figure 2 Comparison between the normalized LMS $L$-filter and its linear counterpart in suppressing mixed impulsive and additive Gaussian noise.

(a) Original image "Trevor White".

(b) Original image corrupted by mixed impulsive and additive Gaussian noise.

(c) Output of the $3 \times 3$ normalized LMS $L$-filter.

(d) Output of the $3 \times 3$ normalized LMS linear filter.

Figure 3 Plot of the Noise Reduction and Mean Absolute Error Reduction indices achieved by the normalized LMS $L$-filter versus the step-size parameter $\mu_0$.

Figure 4 Comparison between the adaptive LMS $L$-filters under study in suppressing mixed impulsive and additive Gaussian noise.

(a) Original image "Lenna".

(b) Original image corrupted by mixed impulsive and additive Gaussian noise.

(c) Output of the $3 \times 3$ location-invariant LMS $L$-filter.

(d) Output of the $3 \times 3$ modified LMS $L$-filter that employs nonhomogeneous step-sizes.

(e) Output of the $3 \times 3$ normalized LMS $L$-filter.

(f) Output of the signal-dependent adaptive $L$-filter structure that employs two normalized LMS $L$-filters of dimensions $5 \times 5$ and $3 \times 3$ to

respectively.

Figure 5 Output of various filters in smoothing image row #128 of "Lenna" corrupted by mixed impulsive and additive Gaussian noise. The same row of the noise-free "Lenna" is shown overlaid in all figures.

(a) Image row #128 in "Lenna" corrupted by mixed impulsive and additive Gaussian noise.

(b) Image row #128 in the filtered image by using the 3 $\times$3 median filter.

(c) Image row #128 in the filtered image by using the 3 $\times$3 normalized LMS (NLMS) $L$-filter.

(d) Image row #128 in the filtered image by using the signal-dependent normalized LMS $L$-filter structure.

Figure 6 Performance of the signal-dependent LMS $L$-filter structure in tracking a time-varying edge.

Table 1: Noise reduction (in dB) achieved by the normalized LMS $L$-filter and its linear counterpart in the restoration of "Trevor White".

| Noise Type | Filter | NR |
|---|---|---|
| Gaussian | $L$-filter | -7.586 |
| | linear filter | -7.225 |
| Impulsive | $L$-filter | -16.82 |
| | linear filter | -9.34 |
| Mixed | $L$-filter | -12.853 |
| | linear filter | -8.96 |

Table 2: Noise reduction and Mean Absolute Error reduction (in dB) achieved by the various $L$-filters in the restoration of "Lenna" corrupted by mixed impulsive and additive Gaussian noise.

| Method | NR | MAER |
|---|---|---|
| median $3 \times 3$ | -8.756 | -8.147 |
| location-invariant LMS $L$-filter $3 \times 3$ ($\mu = 5 \times 10^{-7}$) | -9.747 | -9.192 |
| modified LMS $L$-filter $3 \times 3$ with nonhomogeneous step-sizes ($\mu_0 = 5 \times 10^{-7}$) | -11.216 | -10.867 |
| normalized LMS $L$-filter $3 \times 3$ ($\mu_0 = 0.8$) | -11.281 | -11.071 |
| signal-dependent normalized LMS $L$-filter structure (equal dimensions $3 \times 3$; $\beta_t = 0.75$) | -9.024 | -9.552 |
| signal-dependent normalized LMS $L$-filter structure (L: $5 \times 5$, H: $3 \times 3$; $\beta_t = 0.75$) | -13.224 | -13.928 |

Table 3: Noise reduction and Mean Absolute Error reduction (in dB) achieved by the various $L$-filters in the restoration of "Lenna" corrupted by additive Gaussian noise.

| Method | NR | MAER |
|---|---|---|
| median $3 \times 3$ | -7.365 | -7.34 |
| arithmetic mean filter $3 \times 3$ | -8.949 | -9.01 |
| location-invariant LMS $L$-filter $3 \times 3$ ($\mu = 5 \times 10^{-7}$) | -8.971 | -9.04 |
| modified LMS $L$-filter $3 \times 3$ with nonhomogeneous step-sizes ($\mu_0 = 5 \times 10^{-7}$) | -9.60 | -9.77 |
| normalized LMS $L$-filter $3 \times 3$ ($\mu_0 = 0.8$) | -9.78 | -10.15 |
| signal-dependent normalized LMS $L$-filter structure (equal dimensions $3 \times 3$; $\beta_t = 0.6$) | -9.66 | -10.08 |
| signal-dependent normalized LMS $L$-filter structure (L: $5 \times 5$, H: $3 \times 3$; $\beta_t = 0.75$) | -11.32 | -12.76 |

Table 4: Noise reduction (in dB) achieved by the location-invariant adaptive $L$-filter when a reference image is not available.

| Filtered Image | NR | |
|---|---|---|
| "Lenna" corrupted by mixed impulsive and additive Gaussian noise | -8.40 | (-8.756) |
| "Lenna" corrupted by additive Gaussian noise | -8.65 | (-8.949) |
| "Lenna" using the $L$-filter coefficients of "Trevor White" | -9.477 | |
| "Trevor White" using the $L$-filter coefficients of "Lenna" | -9.293 | |

Table 5: Noise reduction and Mean Absolute Error reduction (in dB) achieved by various $L$-filters in smoothing the second frame of sequence "Trevor White" that has been corrupted by mixed impulsive and additive Gaussian noise for several reference images.

| Method | Ref: 2nd frame | | Ref: 5th frame | | Ref: 9th frame | |
|---|---|---|---|---|---|---|
| | NR | MAER | NR | MAER | NR | MAER |
| median $3 \times 3$ | -8.716 | -8.171 | - | - | - | - |
| location-invariant LMS $L$-filter $3 \times 3$ ($\mu = 5 \times 10^{-7}$) | -9.825 | -9.301 | -9.783 | -9.26 | -9.77 | -9.245 |
| normalized LMS $L$-filter $3 \times 3$ ($\mu_0 = 0.5$) | -11.454 | -11.27 | -10.91 | -10.87 | -10.45 | -10.66 |
| signal-dependent normalized LMS $L$-filter structure (L: $5 \times 5$, $3 \times 3$; $\beta_t = 0.65$) | -14.268 | -14.58 | -13.36 | -13.947 | -12.98 | -13.93 |

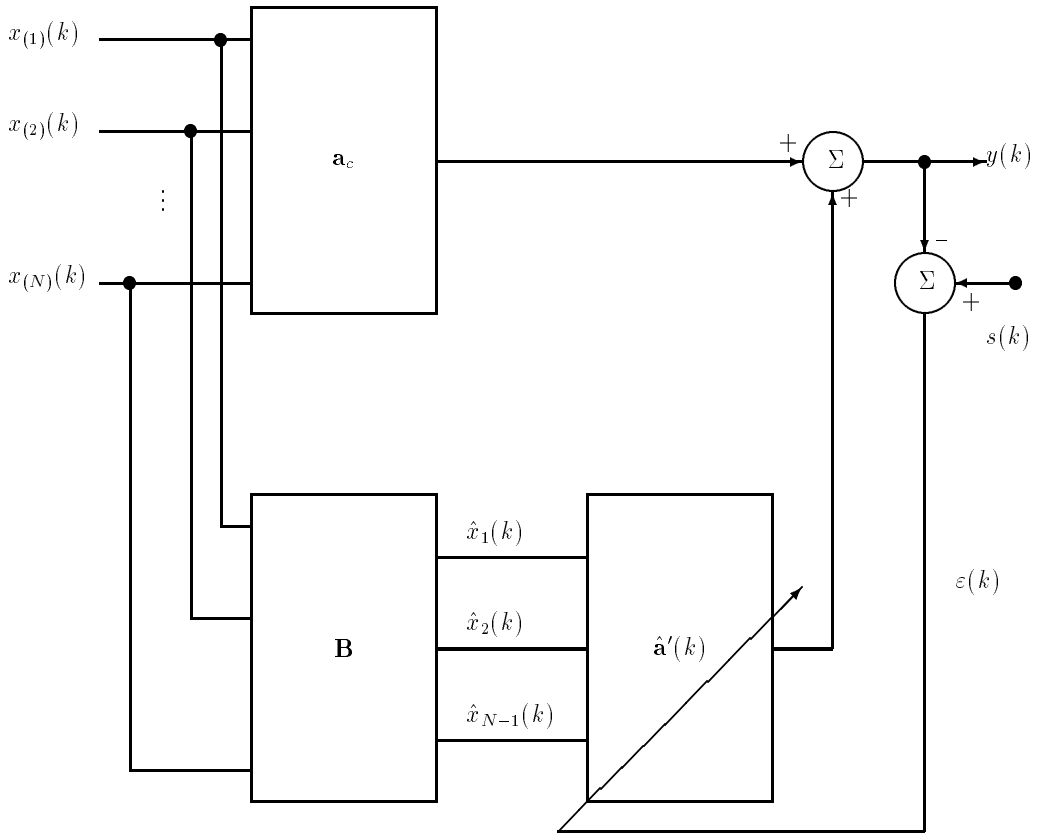



Figure 1: Generalized linearly constrained adaptive processing structure.

Figure 2 (a)

Figure 2 (b)

Figure 2 (c)

Figure 2 (d)

Figure 2: Comparison between the normalized LMS $L$-filter and its linear counterpart in suppressing mixed impulsive and additive Gaussian noise.

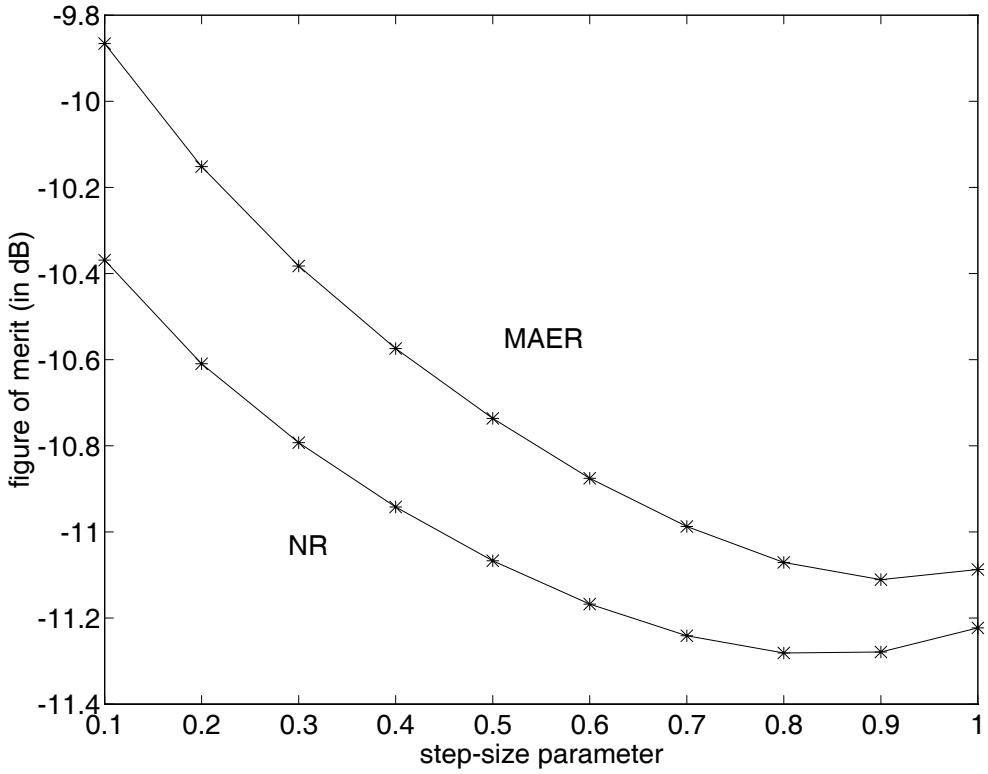



Figure 3: Plot of the Noise Reduction and Mean Absolute Error Reduction indices achieved by the normalized LMS $L$-filter versus the step-size parameter $\mu_0$.

Figure 4 (a)

Figure 4 (b)

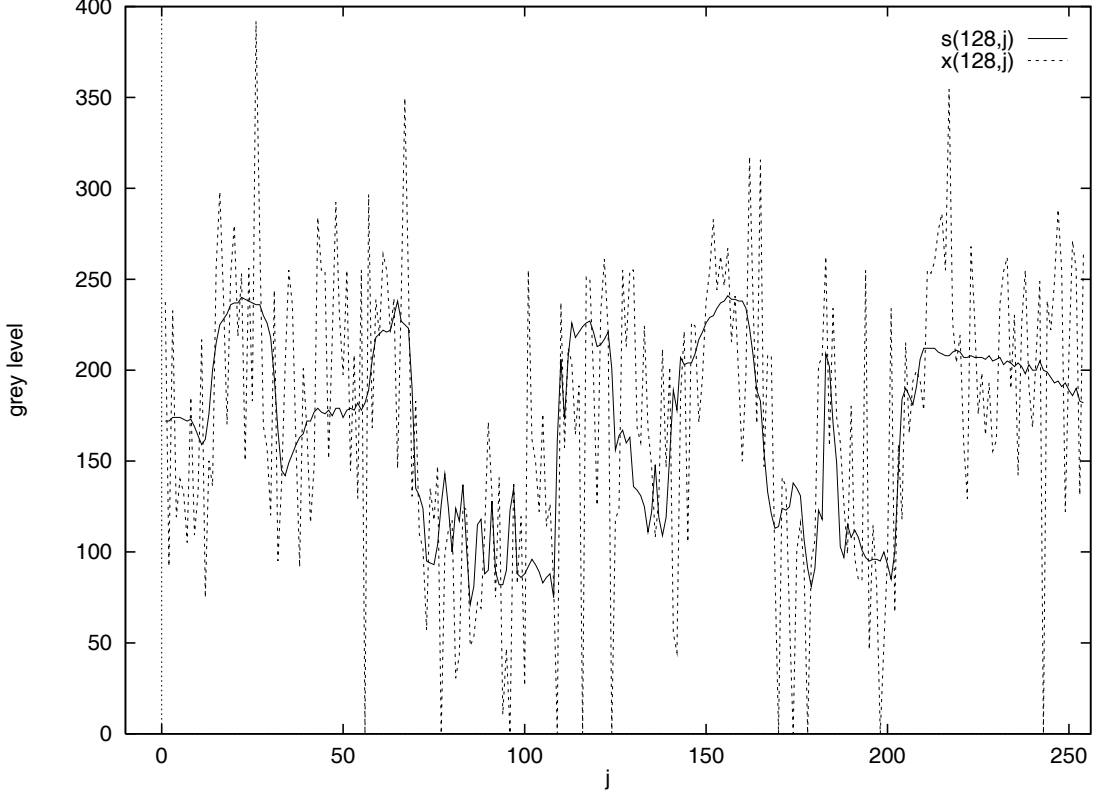
Figure 4 (c)

Figure 4 (d)

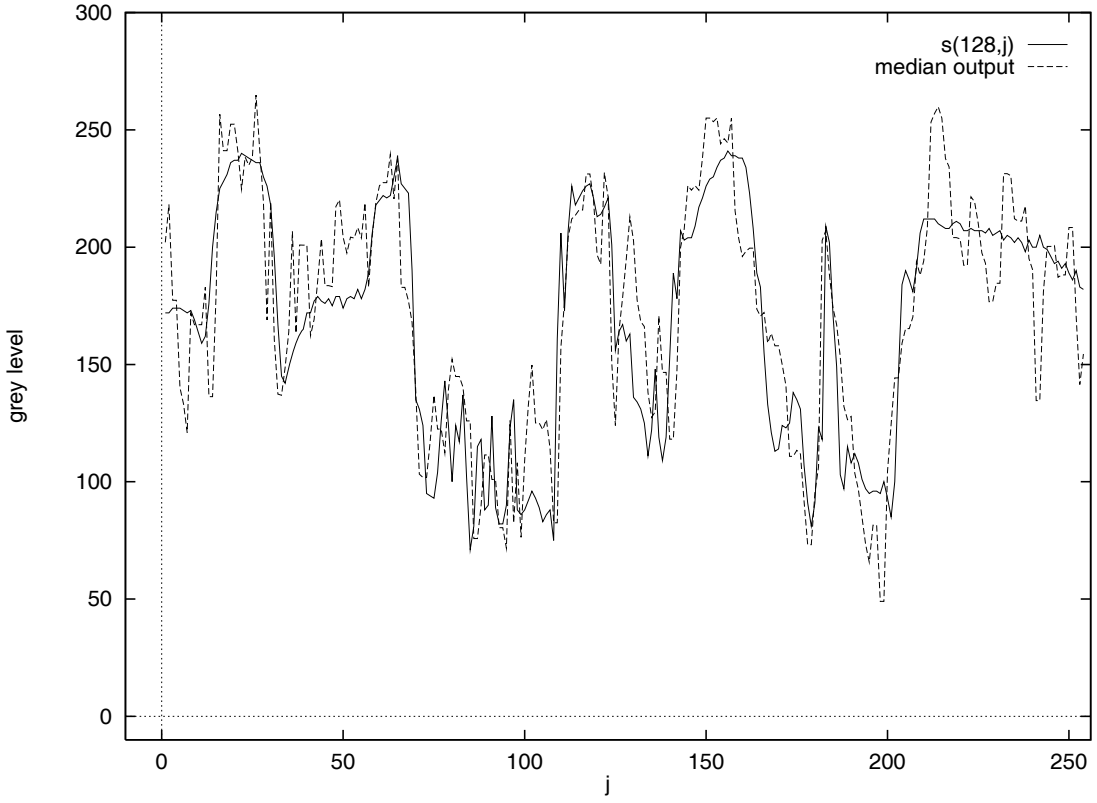
Figure 4 (e)

Figure 4 (f)

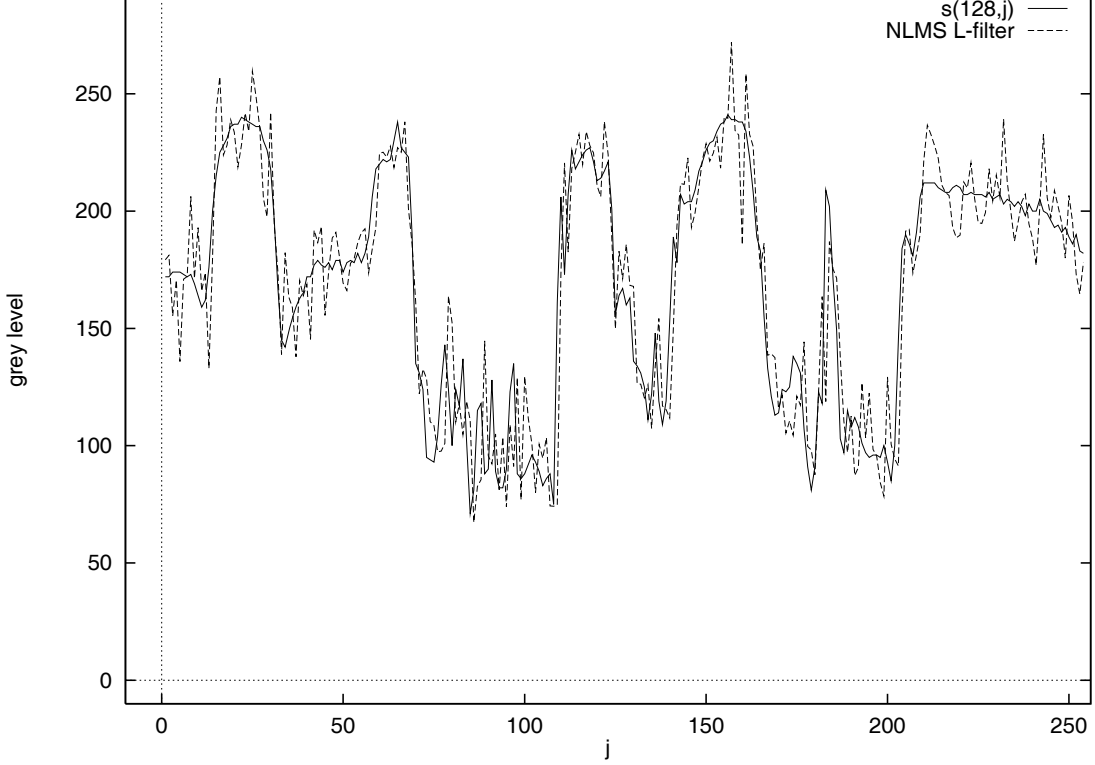
Figure 4: Comparison between the adaptive LMS $L$-filters under study in suppressing mixed impulsive and additive Gaussian noise.


s(128,j)
x(128,j)
grey level
j
0
50
100
150
200
250
300
350
400

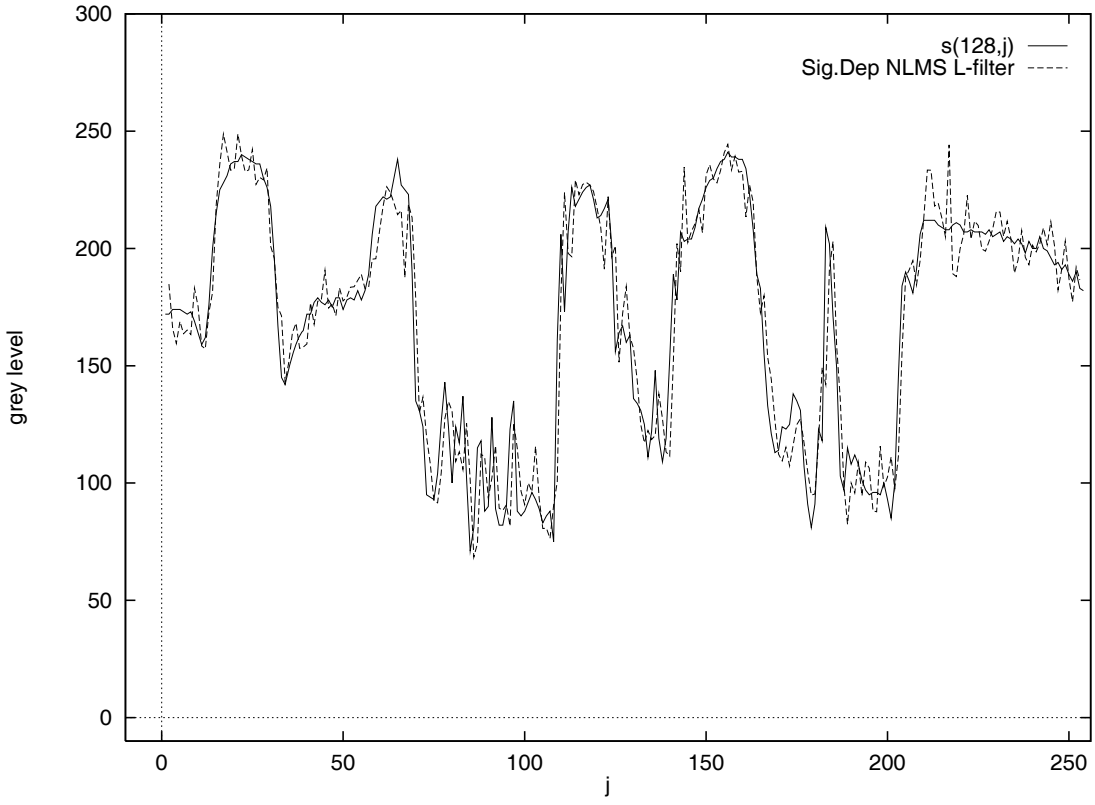
5(a)


s(128,j)
median output
grey level
j
0
50
100
150
200
250
300

5(b)




5(c)




5(d)

Figure 5: Output of various filters in smoothing image row #128 of "Lenna" corrupted by mixed impulsive and additive Gaussian noise. The same row of the noise-free "Lenna" is shown overlaid in all figures.

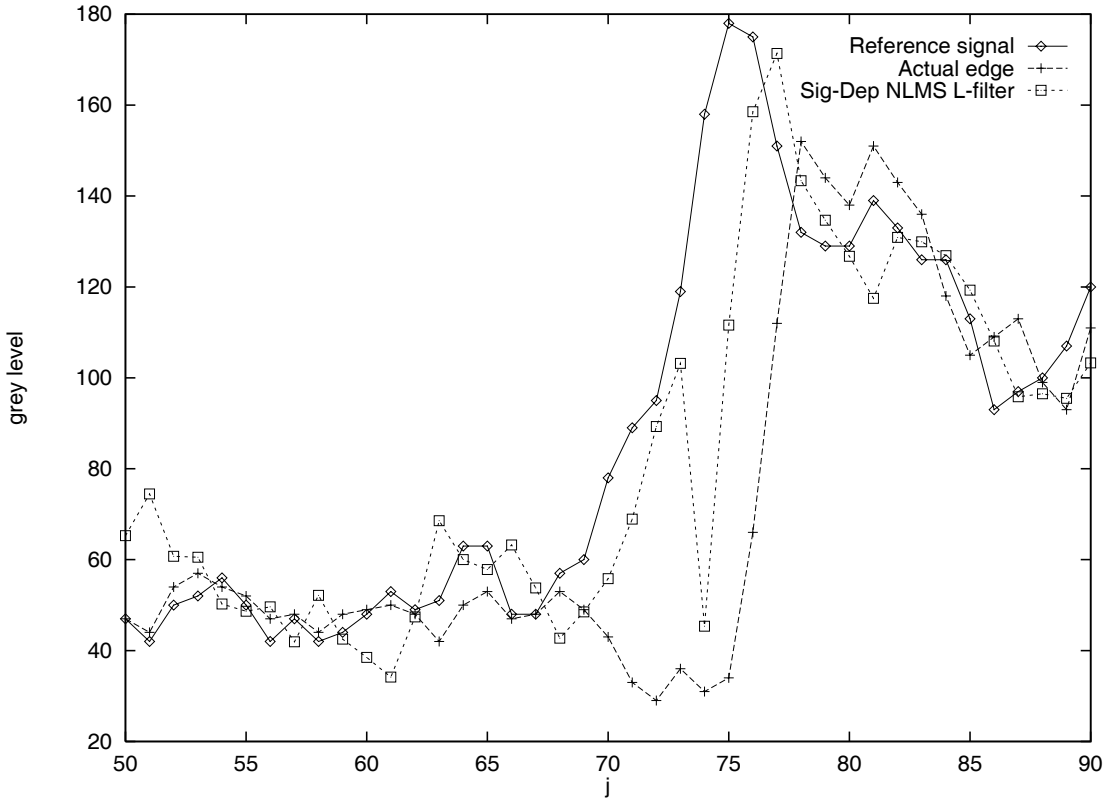



Figure 6: Performance of the signal-dependent LMS $L$-filter structure in tracking a time-varying edge.