ADAPTIVE LMS ORDER STATISTIC FILTERS WITH VARIABLE STEP-SIZES

C. Kotropoulos *S. Tsekeridou* *M. Gabrani*† *I. Pitas*

Department of Informatics, Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE.
E-mail: {costas, sofia, pitas}@zeus.csd.auth.gr
† Department of Electrical and Computer Engineering, Drexel University,
32nd and Chestnut Streets, Philadelphia, PA 19104, U.S.A.
E-mail: maria@cbis.ece.drexel.edu

ABSTRACT

In this paper, adaptive LMS filters based on order statistics employing variable step-sizes are proposed. A novel step-size selection is developed. The performance of the designed filters in noise suppression is compared to the one of adaptive filters that use other step-size selection procedures for still images. It is demonstrated by experiments that the proposed step-size selection yields the best performance for a wide range of noise types including the Gaussian noise, the impulsive noise and the very impulsive one.

1. INTRODUCTION

Adaptive signal processing has been an active research topic for more than two decades. It has found numerous applications in system identification, channel equalisation, echo cancellation etc. [1]. The most widely known adaptive filters are the linear ones that have the form of either finite impulse response (FIR) filters or lattice filters. However, linear filters may not be suitable for applications where the transmission channel is nonlinear or the noise is impulsive or the signal is strongly non-stationary (e.g. in image processing).

On the contrary, a multitude of nonlinear techniques has been proved 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 [2]. One of the best known families is based on the order statistics [3]. It uses the concept of sample ordering. The power of the ordering concept is well illustrated by the median filters which preserve the edges and are the 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 [4].

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. It has been proved both in theory and in practice that adaptive techniques can cope with non-stationary and/or time-varying signals. In this paper we deal with adaptive filters whose coefficients are chosen by employing the Least Mean Squares (LMS) iterative algorithm for the minimisation of the mean squared error (MSE) between the filter output and the desired response. Several authors have used the LMS algorithm to design nonlinear filters. For example, the LMS algorithm has been extensively used in the design of L-filters [5, 6, 7, 8]. A survey on adaptive order statistic filters can also be found in [9]. Adaptive LMS L-filters that employ variable step-sizes are designed and their performance in noise suppression is compared in the case of still images. A variable step-size selection mechanism that is reminiscent to the one proposed in [10] is derived. However, it is worth noting that although our approach is inspired by the method discussed in [10], the details of the selection algorithm are new. It is found that such a variable step-size selection can accelerate the convergence of the filter coefficients towards the optimal ones, especially at the beginning of the filtering session, without deteriorating the noise reduction achieved by the filter at convergence (that is, the mean squared error at convergence). Furthermore, the proposed variable step-size selection algorithm yields always a stable adaptive filtering algorithm which is not the case with normalised LMS (NLMS) algorithm [8], or the method proposed in [10].

2. ADAPTIVE LMS L-FILTERS WITH NON-HOMOGENEOUS STEP-SIZES

Let the observed image $x(\mathbf{k})$ be expressed as the sum of a noise-free image $d(\mathbf{k})$ plus zero-mean two-dimensional additive white noise, i.e., $x(\mathbf{k}) = d(\mathbf{k}) + \eta(\mathbf{k})$ where $\mathbf{k} = (k, l)$ denotes the pixel coordinates. In image processing, a neighbourhood is defined around each pixel $\mathbf{k}$. Among the several neighbourhoods (i.e., filter masks, e.g. cross, x-shape, square, circle) that are used in digital image processing [2], we shall rely on the square window of dimensions $(2\nu + 1) \times (2\nu + 1)$. Let $N = (2\nu + 1)^2$. Since we intend to apply a filter based on sample ordering let us rearrange the $(2\nu + 1) \times (2\nu + 1)$ filter window in a lexicographic order (i.e., row by row) to a $N \times 1$ vector $\mathbf{x}(\mathbf{k})$. If $\mathcal{K}$ and $\mathcal{L}$ denote the image rows and columns respectively, depending on the image scanning method, each pixel (k, l), $k = 1, \ldots, \mathcal{K}$, $l = 1, \ldots, \mathcal{L}$ can be represented by a single running index n. Henceforth, a one-dimensional (1-D) notation is adopted for simplicity.

Let $\mathbf{x}_r(n)$ be the ordered input vector at pixel k given by $\mathbf{x}_r(n) = \left(x_{(1)}(n), x_{(2)}(n), \ldots, x_{(N)}(n)\right)^T$ where $x_{(1)}(n) \leq x_{(2)}(n) \leq \ldots \leq x_{(N)}(n)$ denote the order statistics in the $N \times 1$ input vector. The output of the L-filter is defined by $y(n) = \mathbf{a}^T \mathbf{x}_r(n)$ where $\mathbf{a}$ is the L-filter coefficient vector. The coefficient vector that minimises the MSE between the filtered output and the desired response is simply given by the Wiener solution, i.e.: $\mathbf{a}_o = \mathbf{R}_{x_r}^{-1}\mathbf{p}_r$. $\mathbf{R}_{x_r} = \mathrm{E}\left[\mathbf{x}_r(n)\,\mathbf{x}_r^T(n)\right]$ is the correlation matrix of the observed ordered image pixel values and $\mathbf{p}_r = \mathrm{E}[d(n)\ \mathbf{x}_r(n)]$ denotes the cross-correlation vector between the ordered input vector $\mathbf{x}_r(n)$ and the desired image pixel value $d(n)$. If instantaneous estimates for $\mathbf{R}_{x_r}$ and $\mathbf{p}_r$ are used, then the LMS updating equation for the filter coefficients results in:

$$\mathbf{a}(n+1) = \mathbf{a}(n) + \mu\, e(n)\, \mathbf{x}_r(n) \tag{1}$$

with $e(n)$ denoting the estimation error at pixel n, i.e., $e(n) = d(n) - y(n)$. μ is the adaptation step-size that should satisfy the inequality $0 < \mu < \frac{2}{\mathrm{tr}[\mathbf{R}]}$ so that the average MSE converges to a steady-state value [1] where tr[] stands for the trace of the bracketed matrix. We have also used the identity $\mathrm{tr}[\mathbf{R}_{x_r}] = \mathrm{tr}[\mathbf{R}]$, where $\mathbf{R} = \mathrm{E}[\mathbf{x}(n)\mathbf{x}^T(n)]$ is the correlation matrix of the input observations.

It is well known that the slow convergence rate of the LMS algorithm compared to the convergence rate of the recursive least squares (RLS) algorithm is attributed to the fact that only one parameter, the step-size μ, controls the convergence of all the filter coefficients. On the contrary, in the case of the RLS algorithm, the convergence of each filter coefficient is controlled by a separate element of the Kalman gain vector [1]. In addition, at each iteration, the Kalman gain vector is updated utilising all the information contained in the input data, extending back to the algorithm initialisation. This observation led us to employ different step-size parameters for the various LMS L-filter coefficients in their updating equations, i.e.:

$$\mathbf{a}(n+1) = \mathbf{a}(n) + e(n)\, \mathbf{M}\, \mathbf{x}_r(n) \tag{2}$$

where $\mathbf{M}$ is the following diagonal matrix $\mathbf{M} = \mathrm{diag}[\mu_1, \ldots, \mu_N]$. In the following, a design procedure that enables the selection of μ_i, $i = 1, \ldots, N$ is developed. The discussion has been motivated by the step-size selection proposed in [10]. However, the adaptation of the method to the problem under study is novel.

Let $\mathbf{R}_{x_r} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ where $\mathbf{U} = \{U_{ij}\}$ is the modal matrix of $\mathbf{R}_{x_r}$ whose j-th column is the eigenvector associated with the j-th eigenvalue of $\mathbf{R}_{x_r}$ and $\mathbf{\Lambda}$ is a diagonal matrix composed of the eigenvalues of $\mathbf{R}_{x_r}$. Let $\mathbf{e}_a(n) = \mathbf{a}(n) - \mathbf{a}_o$ denote the coefficient-error vector at n. It is more convenient to work with the transformed coefficient-error vector at iteration n, $\underline{\xi}(n) = \mathbf{U}^T\mathbf{e}_a(n)$. Following similar lines with [10] it can be shown that:

$$\underline{\xi}(n+1) = (\mathbf{I} - \hat{\mathbf{M}}\mathbf{\Lambda})\underline{\xi}(n) \quad ; \quad \hat{\mathbf{M}} = \mathbf{U}^T\mathbf{M}\mathbf{U} \tag{3}$$

where $\mathbf{I}$ is the $N \times N$ identity matrix. It can be easily seen that $\hat{\mathbf{M}}$ is no more diagonal. Its ij-element is:

$$\hat{\mu}_{ij} = \sum_{k=1}^{N} \mu_k U_{ki} U_{kj}. \tag{4}$$

The evolution of the coefficient-error covariance matrix results in [1, 10]:

$$\begin{aligned}\mathbf{K}(n+1) \quad = \quad & \mathbf{K}(n) - \hat{\mathbf{M}}\mathbf{\Lambda}\mathbf{K}(n) - \mathbf{K}(n)\mathbf{\Lambda}\hat{\mathbf{M}} + \hat{\mathbf{M}}\mathbf{\Lambda}\mathrm{tr}\,[\mathbf{\Lambda} \\ & \mathbf{K}(n)]\,\hat{\mathbf{M}} + J_{\min}\hat{\mathbf{M}}\mathbf{\Lambda}\hat{\mathbf{M}}\end{aligned} \tag{5}$$

where $J_{\min}$ denotes the minimum MSE. For a moment, we shall assume that $\mathrm{E}[e_{a_i}e_{a_j}] = \sigma_{e_a}^2\delta_{ij}$ with δ_{ij} denoting Kronecker delta. Such an assumption implies that $\mathbf{K}(n) = \sigma_{e_a}^2\mathbf{I}$. If we also assume that $\hat{\mathbf{M}} = \mu\mathbf{I}$, then (5) can be simplified to a more tractable form than (5), i.e.:

$$\begin{aligned}\mathbf{K}(n+1) \quad = \quad & \mathbf{K}(n)\left(\mathbf{I} - 2\mathbf{\Lambda}\hat{\mathbf{M}}\right) + \mathrm{tr}\,[\mathbf{\Lambda}\mathbf{K}(n)]\,\hat{\mathbf{M}}\mathbf{\Lambda}\hat{\mathbf{M}} + \\ & J_{\min}\hat{\mathbf{M}}\mathbf{\Lambda}\hat{\mathbf{M}}.\end{aligned} \tag{6}$$

In addition, we shall assume that the eigenvalues of $\mathbf{R}_{x_r}$ are equal. Clearly, such an assumption does not hold for the correlation matrix of the order statistics. It can be considered only as a *design assumption*. Then, the step-sizes can be chosen so that the excess mean squared error $\mathrm{E}[J_{\mathrm{ex}}(\mathbf{a}(n))] = \mathrm{tr}[\mathbf{\Lambda}\mathbf{K}(n)]$ is minimised. Such a minimisation problem has been solved by Bershad [11] when (6) holds and all the eigenvalues are equal. To minimise $\mathrm{E}[J_{\mathrm{ex}}(\mathbf{a}(n))]$ the diagonal elements of $\mathbf{K}(n)$ should be minimised. Following similar reasoning with [10] it can be shown that $\hat{\mu}_{ii}$ should be chosen as $\hat{\mu}_{ii} = \frac{1}{\mathrm{tr}[\mathbf{R}_{x_r}]} = \frac{1}{\mathrm{tr}[\mathbf{R}]}$. In the remaining analysis, all the assumptions that yield (6) will be dropped out. By substituting $\hat{\mu}_{ii}$ into (4) we obtain:

$$\sum_{k=1}^{N} \mu_k U_{ki}^2 = \frac{1}{\mathrm{tr}[\mathbf{R}]}. \tag{7}$$

One may write N equations like (7) for $i = 1, \ldots, N$. Then, the set of N equations can be solved for μ_k. However, the solution of the set of equations does not guarantee that each μ_k is less than $1/\mathrm{tr}[\mathbf{R}]$. On the contrary, it is trivial to show that the set of equations (7) implies:

$$\sum_{k=1}^{N} \mu_k = \frac{N}{\mathrm{tr}[\mathbf{R}]}. \tag{8}$$

Accordingly, we shall follow a different approach. Let us assume that the step-size μ_k which controls the adaptation of coefficient $a_k(n)$ (i.e., the weight of the k-th order statistic $x_{(k)}(n)$) is given by $\mu_k = f(\mu_{k1}, \ldots, \mu_{ki}, \ldots, \mu_{kN})$ where $f(\cdot)$ stands for an appropriate function and each μ_{ki} is chosen by taking into consideration only the i-th eigenvector of $\mathbf{R}_{x_r}$. Let us denote by $\mathbf{R}_k$ the matrix:

$$\mathbf{R}_k = \{R_{x_r;ij}\} \qquad i,j = 1, \ldots, k. \tag{9}$$

Each μ_{ki} can be defined as follows:

$$\mu_{ki} = G_k\ \mu_{\min_i} \qquad ; \qquad G_k = \frac{\mathrm{tr}[\mathbf{R}]}{\mathrm{tr}[\mathbf{R}_k]} \tag{10}$$

By substituting (10) into (7) we obtain:

$$\mu_{\min_i} = \left(\mathrm{tr}^2[\mathbf{R}] \sum_{k=1}^{N} \frac{U_{ki}^2}{\mathrm{tr}[\mathbf{R}_k]}\right)^{-1}. \tag{11}$$

Accordingly,

$$\mu_{ki} = \begin{cases} \left(\mathrm{tr}[\mathbf{R}]\mathrm{tr}[\mathbf{R}_k]\sum_{l=1}^{N}\frac{U_{li}^2}{\mathrm{tr}[\mathbf{R}_l]}\right)^{-1} = \psi & \text{if } \psi \le \frac{1}{\mathrm{tr}[\mathbf{R}]}, \\ \frac{1}{\mathrm{tr}[\mathbf{R}]} & \text{otherwise.} \end{cases} \tag{12}$$

Having computed μ_{ki}, the step-size μ_k can be obtained, for example, by computing the average value of μ_{ki}:

$$\mu_k = \frac{1}{N}\sum_{i=1}^{N} \mu_{ki}. \tag{13}$$

The major difficulties of the variable step-size selection analysed above are the following: (i) It requires the computation of the eigenvalues and eigenvectors of the correlation matrix at every image pixel. Therefore, it increases dramatically the computational complexity of the algorithm. (ii) There is no guarantee that the $\mu_{\min_i}$ determined in (11) yields always a stable filter operation. We have to check if $\mu_{\min_i} < 1/\mathrm{tr}[\mathbf{R}]$ before accepting the value computed in (11). It has been found that the most crucial term in the variable step-size selection algorithm proposed is the ratio G_k. Accordingly, we propose to compute a sequence of variable step-sizes by employing a run-time estimate of G_k and a constant step-size $\mu_0 < 1/\mathrm{tr}[\mathbf{R}]$ (e.g. $\mu_0 = 10^{-8}$) as follows:

$$\mu_k(n) = \begin{cases} \hat{G}_k(n)\,\mu_0 & \text{if } \mu_k(n) < \mu_{\max} \\ \mu_0^{'}/\left(\mathbf{x}^T(n)\mathbf{x}(n)\right) & \text{otherwise} \end{cases} \quad (14)$$

where $\mu_{\max}$ is an upper bound on the variation of μ, e.g. 10^{-6}, $\mu_0^{'}$ is a positive real number less than 1 and $\hat{G}_k(n)$ is an estimate for the ratio G_k for every image pixel that is computed by:

$$\hat{G}_k(n) = \frac{\sum_{i=1}^{N} Q_i(n)}{\sum_{i=1}^{k} Q_i(n)} \quad (15)$$

$$Q_i(n) = \frac{1}{n-N}\sum_{j=N}^{n} x_{(i)}^2(j), \quad n = N+1, \ldots \quad (16)$$

Obviously, $Q_i(n)$ can be computed recursively.

3. EXPERIMENTAL RESULTS

The following quantitative criteria are considered in order to quantify the quality of filtering: (i) the Noise Reduction index (NR):

$$NR = 10\log\frac{\sum_{k=1}^{\mathcal{K}}\sum_{l=1}^{\mathcal{L}}(x(k,l)-d(k,l))^2}{\sum_{k=1}^{\mathcal{K}}\sum_{l=1}^{\mathcal{L}}(y(k,l)-d(k,l))^2} \quad (17)$$

and the Mean-Squared Error (MSE):

$$MSE = \frac{1}{\mathcal{K}\mathcal{L}}\sum_{k=1}^{\mathcal{K}}\sum_{l=1}^{\mathcal{L}}(d(k,l)-y(k,l))^2 \quad (18)$$

In (17)-(18), $d(k,l)$ denotes the noise-free image, $x(k,l)$ is corrupted image and $y(k,l)$ is the filtered image. $\mathcal{K}$ and $\mathcal{L}$ are the number of image rows and columns, respectively. The experiments are performed using a filter window size of 3×3.

Subsequently, the performance of the variable step-size selection algorithm proposed in Section 2 is studied. First, we compare the performance of adaptive L-filters that employ the following step-size selections: (i) constant step-size ($\mu_0 = 10^{-8}$), (ii) the normalised LMS algorithm [8], (iii) the algorithm proposed in [10], and (iv) the proposed algorithm. A noisy "Airfield" [12] image produced by adding mixed Gaussian noise having zero mean and standard deviation equal to 50 and impulsive noise having probability of impulse occurrence 10% with an equal percentage of positive and negative impulses has been used. The initial filter coefficients are set equal to zero. We shall also assume that the noise-free image is available. Performance results have been obtained either during the adaptation by using the coefficients determined at each image pixel or by using a set of coefficients that is produced by averaging the filter coefficients computed in the last image row throughout the entire image. For comparison purposes the same figures of merit for median filter are also included in Table 1. By inspecting the entries of Table 1 it is found that all step-size selection algorithms provide almost the same results with respect to the quantitative criteria computed. However, the averaged filter coefficients in the last image row when the proposed algorithm is applied yield the best results.

As has already been explained, the motivation for introducing a variable step-size selection algorithm is in accelerating the convergence rate. This is clearly seen by examining the learning curves in Figure 1. Each plot represents an approximation of the ensemble-averaged learning curve of each adaptive L-filter under study for the first five rows of the image. Each row corresponds to 510 samples. It has been obtained following the procedure described in [1]. That is, the squared norm of the estimation error $e(n)$ has been computed at each image pixel. This experiment has been repeated 100 times, each time using an independent realization of the process $\{\eta(n)\}$. The averaged squared norm of the estimation error is then determined by computing the ensemble average of $e(n)$ over the 100 independent trials of the experiment. It is evident that the rate of convergence for the adaptive L-filter that employs the proposed step-size selection algorithm (Figure 1(b)) is faster than the corresponding rate of the adaptive L-filter with a constant step-size $\mu_0 = 10^{-8}$ (Figure 1(a)). The rate of convergence of the adaptive L-filter with the proposed step-size selection algorithm is almost identical to the one that employs the method proposed in [10] (Figure 1(c)). However, the proposed method does not rely on eigenvalue-eigenvector computations. It is seen that the normalised LMS L-filter achieves the fastest initial rate of convergence (Figure 1(d)). Furthermore, large errors are observed at the beginning of each row which is an undesirable effect. It is worth noting that this type of algorithm may yield bias in estimating the mean of the output which is not the case with the proposed algorithm. Figures 2(a) and 2(b) show the variation of the step-size at the last pixel of each row along all image rows for the proposed algorithm and the algorithm in [10], respectively.

The last experiment in this set aims at studying the dependence of adaptive L-filter performance on the noise-free image that is used as desired signal when several step-size selection algorithms are employed. Towards this goal we split each filtering procedure in two sessions, namely, the training session and the test session. A different pair of noisy and noise-free images is used in the training and in the test session. More specifically, in the training session, a pair of images originated from image "Bridge" has been used while a pair of images originated from image "Airfield" has been used in the test session. The noisy images in both sessions have been corrupted by the same kind of noise. All images have been extracted from the TUT database [12]. The objective in the training phase is to determine a set of L-filter coefficients by averaging those coefficients found at the last image row. This set of L-filter coefficients is used subsequently in the test session. Moreover, we have tested several kinds of noise that yield the same SNR in the noisy input image. Table 2 summarises the MSE found in each case examined. By inspecting Table 2, it is seen that the performance of the adaptive L-filters under study is almost identical for very impulsive and impulsive noise. For in-

Table 1: Performance indices achieved by adaptive L-filters that employ different step-size selections in still image filtering.

Step-size selection	Running coefficients		Averaged coefficients	
	NR (in dB)	MSE	NR (in dB)	MSE
Constant	8.140	586.885	8.098	592.563
Normalised LMS	7.974	609.755	8.233	574.446
[10]	8.136	587.360	8.231	574.646
Proposed	8.131	588.147	8.281	568.107
Median filter	7.446	688.545		

Table 2: Dependence of adaptive L-filter performance on the noise-free image that is used as desired signal when several step-size selection algorithms are employed. Training image: "Bridge". Test image: "Airfield". SNR in the noisy input image 9 dB in all cases.

Step-size selection/Filter	Noise Type	MSE
Median filter	very impulsive	140.879
Normalised LMS		152.867
Constant step-size		149.425
[10]		149.757
Proposed		146.496
Median filter	impulsive	170.913
Normalised LMS		230.033
Constant step-size		201.634
[10]		188.513
Proposed		182.166
Median filter	Gaussian	207.794
Normalised LMS		622.438
Constant step-size		195.056
[10]		194.056
Proposed		192.908

put images with poor quality (e.g. SNR 3 dB) the maximum gain in the SNR at the output is 0.5 dB. For Gaussian noise, the maximum gain in the SNR at the output is approximately 1 dB. By considering the entry for the normalised LMS L-filter in the case of Gaussian noise it becomes evident that the filter results in a significantly biased estimate of the mean of the output image. This bias is measured and found equal to 30.38 which manifests the strong dependence of this filter on the noise-free signal that is used to yield the filter coefficients. Such a conclusion is in par with the experimental results reported in [8]. For comparison purposes the MSE in the output of the median filter is also included in Table 2.

4. CONCLUSIONS

The use of adaptive nonlinear filters based on order statistics with variable step-sizes has been examined in this paper. An LMS adaptation algorithm was used for the adaptation of the L-filter coefficients. A variable step-size selection algorithm has been proposed and its performance has been studied. It has been shown that such a selection algorithm can accelerate the rate of convergence in the first image rows. The proposed algorithm has an almost identical performance to the method proposed in [10]. However, no eigenvector/eigenvalue computations are needed in our framework. Moreover, it has been found that the proposed algorithm yields an adaptive L-filter that does not depend on the desired image used when the filter coefficients are determined. This is not always the case with the normalised LMS L-filter.

5. REFERENCES

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

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

[3] 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.

[4] 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.

[5] I. Pitas, and A.N. Venetsanopoulos, "Adaptive filters based on order statistics," *IEEE Trans. on Signal Processing*, vol. 39, no. 2, pp. 518–522, February 1991.

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

[7] 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.

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

[9] 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.

[10] S. Chen, and G.R. Arce, "Microstatistic LMS filters," *IEEE Trans. on Signal Processing*, vol. 41, no. 3, pp. 1021–1034, March 1993.

[11] N.J. Bershad, "On the optimum gain parameter in LMS adaptation," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. ASSP-35, no. 7, pp. 1065–1068, July 1987.

[12] M. Gabbouj, and I. Tabus, "TUT noisy image database," *Technical Report ISBN 951-722-281-5*, Signal Processing Lab., Tampere Univ. of Technology, Tampere, Finland, December 1994.

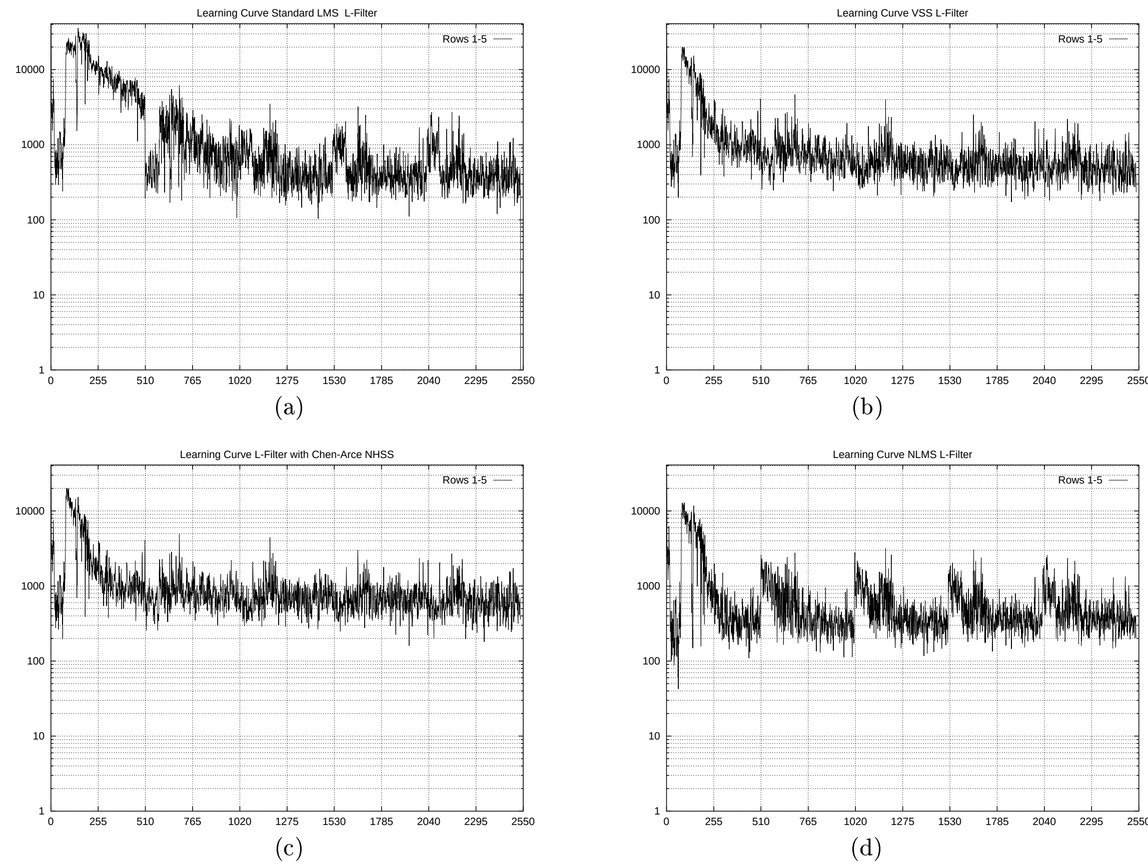

Figure 1: Learning curves for several adaptive L-filters with different step-size selection algorithms: (a) Constant step-size $\mu_0 = 10^{-8}$. (b) Proposed step-size selection. (c) Non-homogeneous step-size selection proposed by Chen and Arce. (d) Normalised LMS L-filter.

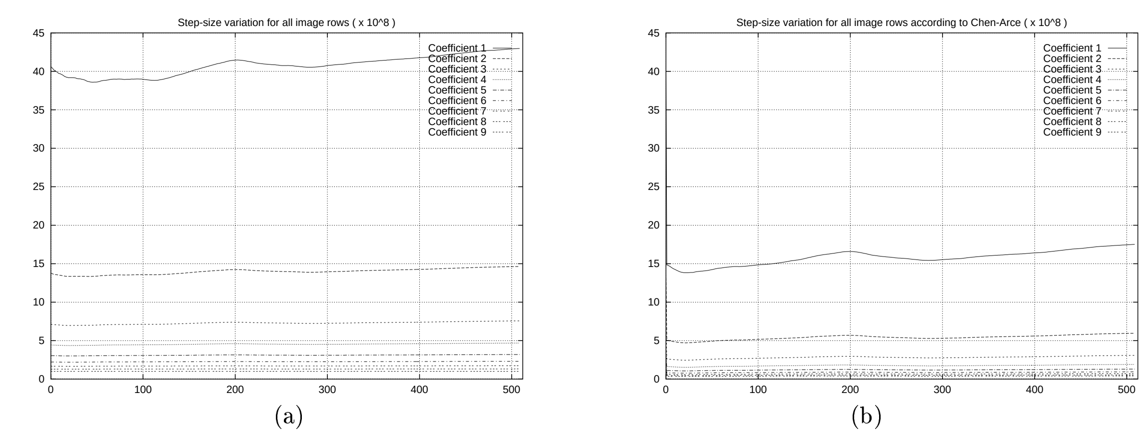

Figure 2: Variation of the step-size at the last pixel of each row along all image rows. (a) Proposed step-size selection. (b) Non-homogeneous step-size selection proposed by Chen and Arce.