Two-Dimensional Block Diagonal LMS Adaptive Filtering

Mahmood R. Azimi-Sadjadi, *Senior Member, IEEE,* and Hongye Pan

Abstract—This paper is concerned with the development of a two-dimensional (2-D) adaptive filters using the block diagonal least mean squared (BDLMS) method. In this adaptive filtering scheme, the image is scanned and processed block by block in a diagonal fashion, and the filter weights are adjusted once per block rather than once per pixel. The diagonal scanning is adopted to avoid the problems inherent in the 1-D standard scanning schemes and to account for the correlations in two directions. The weight updating equation for 2-D BDLMS is derived, and the convergence properties of the algorithms are investigated. Simulation results that indicate the effectiveness of the 2-D BDLMS when used for image enhancement, estimation, and detection applications are presented.

I. INTRODUCTION

ONE-DIMENSIONAL gradient-based adaptive schemes have found applications in numerous areas such as system identification, adaptive control and filtering, channel equalization, and neural network learning [1]. The least mean squares (LMS) algorithm originally developed by Widrow and Hopf [1], [2] uses gradient descent method to update the weights in a transversal adaptive filter structure. In contrast to the nonadaptive schemes, this approach takes into account the nonstationarity inherent in practical signals. In addition, it does not make any *a priori* assumptions regarding the statistics of the data and the degradations.

A variant of the LMS algorithm known as block LMS or BLMS was derived in [3] based on the application of the block processing scheme [4]. Using this algorithm, the weights of a transversal filter are updated upon the arrival of a new block of data. This is in contrast with the conventional LMS adaptive algorithm which adjusts the weights for each new sample of the input data. In addition to the inherent benefits of the block processing including parallelism (ideally suited for systolic array implementation), reduced roundoff error effects, improved stability performance, etc., BLMS can be implemented very efficiently using fast transform techniques.

Mikhael and Wu [5], [6] introduced a 1-D fast algorithm for block adaptive digital filtering with application to time-varying system identification. To track the time-varying parameters of the unknown system, the step size is updated using the Taylor series expansion of the error signal. This method accelerates the convergence especially for the nonstationary signals.
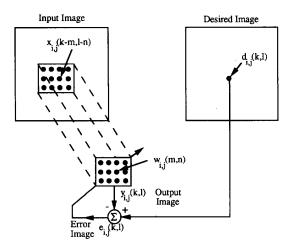
Fig. 1. 2-D adaptive transversal filtering process.

The conventional LMS scheme was first extended to the 2-D case by Hadhoud and Thomas [7]. Their 2-D LMS algorithm is particularly useful in various image processing areas such as image enhancement and image data compression where the variations in local statistics of the image must be taken into account. This method processes an image pixel by pixel using a 1-D scanning scheme, and consequently, it considers the correlation of pixels in only one direction. Soni *et al.* [8] used a 2-D adaptive LMS filter to detect and isolate small objects with broad-band spectra from background clutter with a narrow-band spectrum. More recently, several other authors [9]–[11] have developed different 2-D LMS-based adaptive algorithms. In [10], a new 2-D sequential adaptive filtering scheme was proposed which uses variable step-size to improve the convergence behavior. This algorithm was then extended in [11] for block processing using a scalar adaptation rule.

This paper is concerned with the development of a 2-D block diagonal LMS (BDLMS) algorithm based on the application of the 2-D block processing method [12]. A 2-D diagonal scanning is employed to preserve the local correlational information of pixels in both directions. The convergence behavior of the 2-D BDLMS filter is studied. To consider the nonstationarity inherent in real-world images, a variable step size rule is given. Applications of the 2-D BDLMS in image estimation, filtering, and detection areas are studied. Simulation results for filtering additive noise from a corrupted image are presented, and a comparison is made between the proposed 2-D BDLMS and the standard Wiener
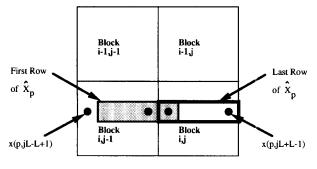
Fig. 2. Support region for $\mathbf{X}_{i,j}$.

filter. In addition, the 2-D BDLMS was used as a predictor to detect and isolate small objects from the background clutter [8]. This can be particularly useful in applications such as target detection and medical imaging where the detection of subtle abnormalities is of special importance.

II. 2-D BLOCK DIAGONAL LMS ADAPTIVE FILTER

A. 2-D Block Implementation

Consider a 2-D adaptive transversal filter, as shown in Fig. 1. In this figure, $d(m, n)$ is the desired image of size M by N, $x(m, n)$ is the reference input image of the same size, and $w(m, n)$'s represent the weights of the transversal (nonrecursive) filter of the order K by L.

Let us assume that the images are partitioned into nonoverlapping blocks each of size K by L. The (k, ℓ)th element of the output block (i, j) can be obtained from the linear convolution of the filter coefficients with the input image. Since the filter is assumed to be linear space-invariant (LSI) within each block and space-varying interblock wise, we can write

$$y_{i,j}(k, \ell) = \sum_{m=0}^{K-1} \sum_{n=0}^{L-1} w_{i,j}(m, n) \ x_{i,j}(k - m, \ell - n)$$

(1)

$$k\epsilon[0, K - 1]$$
$$\ell\epsilon[0, L - 1]$$

where $w_{i,j}(m, n)$'s represent the weights of the filter in block (i, j), and $y_{i,j}(k, \ell) = y(iK + k, jL + \ell)$ is the (k, ℓ)th element of (i, j)th output block, and similarly, we have $x_{i,j}(k - m, \ell - n) = x(iK + k - m, jL + \ell - n)$. Equation (1) can easily be written in matrix form by arranging all $y_{i,j}(k, \ell)$'s in a block and the associated $w_{i,j}(m, n)$'s in row-ordered vector form and $x_{i,j}(k - m, n - \ell)$'s in a matrix form. This process yields

$$Y_{i,j} = \mathbf{X}_{i,j} W_{i,j}$$

(2)

where $Y_{i,j}$ represents the (i, j)th output block arranged in a row-ordered vector of size $KL \times 1$ defined as

$$Y_{i,j} = [\hat{y}_{iK} \quad \hat{y}_{iK+1} \cdots \hat{y}_{iK+K-1}]^t$$

(3a)

where

$$\hat{y}_p = [y(p, jL) \quad y(p, jL + 1) \cdots y(p, jL + L - 1)]$$

(3b)

and $p\epsilon[iK, iK + K - 1]$. Matrix $\mathbf{X}_{i,j}$ of size $KL \times KL$ is defined by

$$\mathbf{X}_{i,j} = \begin{bmatrix} \hat{X}_{iK} & \hat{X}_{iK-1} & \cdots & \hat{X}_{iK-K+1} \\ \hat{X}_{iK+1} & \hat{X}_{iK} & \cdots & \hat{X}_{iK-K+2} \\ \vdots & \vdots & \vdots & \vdots \\ \hat{X}_{iK+K-1} & \hat{X}_{iK+K-2} & \cdots & \hat{X}_{iK} \end{bmatrix}$$

(4a)

where we have (4b), which appears at the bottom of this page, and $p\epsilon[iK - K + 1, iK + K - 1]$. The region in which $\mathbf{X}_{i,j}$ is defined is shown in Fig. 2. The last row of matrix $\hat{X}_p$ represents the pth, $p\epsilon[iK - K + 1, iK + K - 1]$ row of block $(i-1, j)$ or (i, j). The other rows of $\hat{X}_p$ are obtained by shifting this row into blocks $(i - 1, j - 1)$ or $(i, j - 1)$ one pixel at a time. Consequently, the last row of $\mathbf{X}_{i,j}$ corresponds to block (i, j), and the other rows are obtained by shifting this block into other neighboring three blocks as shown in the figure. The weight vector $W_{i,j}$ of size $KL \times 1$ is defined by

$$W_{i,j} = [\hat{W}_{i,j}(0) \quad \hat{W}_{i,j}(1) \quad \cdots \quad \hat{W}_{i,j}(K - 1)]^t$$

(5a)

where

$$\hat{W}_{i,j}(q) = [w_{i,j}(q, 0) \quad w_{i,j}(q, 1) \cdots w_{i,j}(q, L - 1)]$$

(5b)

and $q\epsilon[0, K - 1]$. The expression for the error $e_{i,j}(k, \ell) := d_{i,j}(k, \ell) - y_{i,j}(k, \ell)$ in block form becomes

$$E_{i,j} = D_{i,j} - Y_{i,j} = D_{i,j} - \mathbf{X}_{i,j} W_{i,j}$$

(6)

where $D_{i,j}$ and $E_{i,j}$ are, respectively, the (i, j)th blocks of the desired image and error image defined in a fashion similar to $Y_{i,j}$ in (3).

B. Block Diagonal Adaptive Filtering

Having formulated the 2-D adaptive filter equations in matrix form, the estimation of the parameters or weights of the filter can be accomplished using the standard Wiener filter theory [1], [2] by choosing the weights such that the error $E_{i,j}$ is minimized in the mean squared sense. The mean squared value of the estimation error, which is referred to as the "index of performance" for the 2-D block LMS adaptive filter or the "block mean squared error" (BMSE)

$$\text{BMSE} = \Xi := \frac{1}{KL} E[E_{i,j}^t E_{i,j}].$$

(7a)

$$\hat{X}_p = \begin{bmatrix} x(p, jL) & x(p, jL - 1) & \cdots & x(p, jL - L + 1) \\ x(p, jL + 1) & x(p, jL) & \cdots & x(p, jL - L + 2) \\ \vdots & \vdots & \vdots & \vdots \\ x(p, jL + L - 1) & x(p, jL + L - 2) & \cdots & x(p, jL) \end{bmatrix}$$
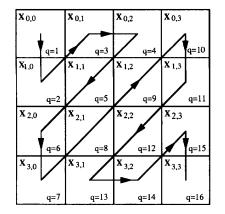
(4b)

Fig. 3.   Block diagonal processing.

Fig. 5.   2-D BDLMS adaptive smoother structure.

Fig. 4.   Original Lena image.

Fig. 6.   Corrupted Lena image (SNR = 1.8 dB).

where $E[\cdot]$ represents the expectation operator [13], can be minimized for the weight vector by taking its derivative with respect to the weight vector, i.e.

$$\Delta_{i,j} := \frac{d\Xi}{dW_{i,j}} = -2P_{i,j} + 2R_{i,j}W_{i,j}. \qquad (7b)$$

and setting it equal to zero. This yields the analytical optimal Wiener-Hopf solution [1], [2] for the weight vector of the filter as

$$W^*_{i,j} = R^{-1}_{i,j}P_{i,j}. \qquad (8)$$

where $R_{i,j}$ and $P_{i,j}$ are the autocorrelation and cross-correlation matrices defined by

$$R_{i,j} := E[X^t_{i,j}X_{i,j}] \qquad (9a)$$

$$P_{i,j} := E[X^t_{i,j}d_{i,j}] \qquad (9b)$$

and $X_{i,j}$ is the (i,j)th block of the input image given by

$$X_{i,j} = [\hat{\mathbf{x}}_{iK} \quad \hat{\mathbf{x}}_{iK+1} \quad \cdots \quad \hat{\mathbf{x}}_{iK+K-1}]^t \qquad (10a)$$

where

$$\hat{\mathbf{x}}_p = [x(p,jL) \quad x(p,jL+1) \quad \cdots \quad x(p,jL+L-1)].$$
$$(10b)$$

For this optimal choice of the weight vector, the BMSE for the (i,j)th block takes its minimum, which is

$$\Xi_{min} = \frac{1}{KL}E[D^t_{i,j}D_{i,j}] - P^t_{i,j}W^*_{i,j}. \qquad (11)$$

Note that if the input is globally stationary, we then have $W^* = R^{-1}P$.

Although the above procedure is quite straightforward in nature, for a large amount of data, the process becomes computationally laborious. An alternative way to minimize the BMSE in (7a) is to use a gradient search technique such as the steepest descent method [2], which leads to the LMS algorithm. However, the extension of the 1-D LMS rule to

(a)

(b)

Fig. 7. Processed Lena image: (a) Fixed μ_B (SNR = 9.1 dB). (b) variable μ_B, (SNR = 9.8 dB).

images creates several problems in the adaptation process at the end of rows (or columns) as the 2-D spatial correlations of a group of pixels in a neighborhood are generally ignored. To overcome these problems, the idea of 2-D diagonal processing is proposed. The procedure is depicted in Fig. 3. Using this diagonal scanning scheme, the updating equation for the weight vector becomes

$$W_{i',j'} = W_{i,j} - \mu_B \Delta_{i,j} \qquad (12)$$

where $W_{i',j'}$ is the updated weight vector at block location (i',j'), μ_B is the step size, and $\Delta_{i,j}$ is the gradient at block

(i,j) defined in (7b). The block indices (i,j) and (i',j') satisfy the condition that the sum of i and j, and the sum of i' and j' along each diagonal line is a constant. Six different cases for changing the block indices exist and are given below.

For $i + j$ an even number

$i' = i + 1, j' = j - 1$, when $i < i_{max}$ and $j \neq 0$
$i' = i + 1, j' = j$, when $i < i_{max}$ and $j = 0$
$i' = i, j' = j + 1$, when $i = i_{max}$.

For $i + j$ an odd number

$i' = i - 1, j' = j + 1$, when $j < j_{max}$ and $i \neq 0$
$i' = i, j' = j + 1$, when $j < j_{max}$ and $i = 0$
$i' = i + 1, j' = j$, when $j = j_{max}$

where i_{max} and j_{max} are the last row and column block indices, respectively.

As discussed previously, the calculation of an accurate gradient $\Delta_{i,j}$ involves a large amount of ensemble average computation. Based on the method of steepest descent, an estimate of the gradient can be obtained using [2]

$$\hat{\Delta}_{i,j} := \frac{1}{KL} \frac{\partial[E_{i,j}^t E_{i,j}]}{\partial W_{i,j}} = \frac{-2}{KL} X_{i,j}^t E_{i,j}. \qquad (13)$$

Substituting this estimate into the updating equation (12) gives the 2-D BDLMS adaptation rule as

$$W_{i',j'} = W_{i,j} - \mu_B \hat{\Delta}_{i,j} = W_{i,j} + \frac{2\mu_B}{KL} X_{i,j}^t E_{i,j}. \qquad (14)$$

As in the 1-D case [3], the best choice for the block size is the filter order. In addition, matrix $X_{i,j}$ in this case is doubly Toeplitz. Thus, the operations involving this matrix such as those in (2) and the adaptation equation (14) can be performed efficiently using FFT algorithms [3], [12]. In the next section, the convergence properties of the 2-D BDLMS are given.

III. CONVERGENCE PROPERTIES OF 2-D BDLMS ADAPTIVE FILTER

The convergence properties of the 2-D BDLMS adaptive filter to be discussed in this section include the bounds on the step size μ_B, adaptation speed, learning curve, and adaptation accuracy [1]-[3]. For the purpose of this section, we assume that the input image is a stationary process.

The bounds on the step size for the weight vector to converge to the Wiener solution is given [1]-[3] by

$$0 < \mu_B < \frac{1}{\lambda_{max}}, \qquad (15)$$

where λ_{max} is the maximum eigenvalue of R. In practice, however, it is not easy to find the eigenvalues of the matrix R. Instead, the trace of the matrix R, which represents the power of the input image and is always greater than λ_{max}, i.e., $\mathrm{tr}(R) > \lambda_{max}$ can be used [2]. Thus, we get

$$0 < \mu_B < \frac{1}{\mathrm{tr}(R)}, \qquad (16)$$

which is more useful than that in (15).

Adaptation speed is used to measure the speed of the adaptive process (or learning curve). In 2-D BDLMS, the learning ability is measured by the speed at which BMSE approaches Ξ_{min}. Usually, this adaptation speed is given in
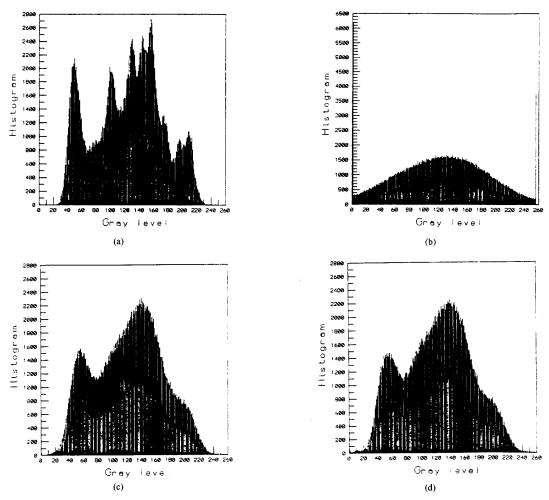
Fig. 8. Histograms: (a) Original; (b) corrupted; (c) processed, fixed μ_B; (d) processed, variable μ_B.

terms of 2-D BDMSE time constant, which indicates the speed at which the weight vector converges to the Wiener solution when the method of steepest descent is used. For the special case in which all eigenvalues of the input auto-correlation matrix R are equal, the block time constant T_{BMSE} of the 2-D BDLMS adaptive filter is found [3], [14] to be

$$T_{\text{BMSE}} = \frac{KL}{4\mu_B\lambda}. \tag{17}$$

Alternatively, we can use the trace of R to determine T_{BMSE} as

$$T_{\text{BMSE}} = \frac{(KL)^2}{4\mu_B\text{tr}(R)} \tag{18}$$

where $\text{tr}(R) = KL\lambda$ when all the eigenvalues are equal. As can be seen, the adaptation speed is inversely proportional to the step size μ_B. Increasing the step size μ_B above the convergence range may cause the adaptation process to exhibit instability problems.

Adaptation accuracy describes how close BMSE (or Ξ) gets to the minimum BMSE Ξ_{min} after the adaptation. The

deviation from this minimum is caused by gradient noise or inaccuracy inherent in the steepest descent method [2]. For the 2-D BDLMS filter, the misadjustment was derived [3], [14] to be

$$M = \frac{\mu_B}{KL}\text{tr}(R). \tag{19}$$

When all the eigenvalues of R are equal, the misadjustment can be expressed in terms of the block time constant T_{BMSE} as

$$M = \frac{1}{4T_{\text{BMSE}}}. \tag{20}$$

Again, this shows that there is tradeoff between the adaptation speed and the misadjustment [1]–[3].

IV. 2-D BDLMS FOR NONSTATIONARY PROCESSES

For nonstationary processes, the objective is to achieve fast convergence and tracking and small misadjustment when the statistical properties change in different blocks [5], [6]. In other words, the filtering process should adapt and track fast enough to capture the statistical changes in the image. Since the key factor for the convergence of an adaptive process is the step
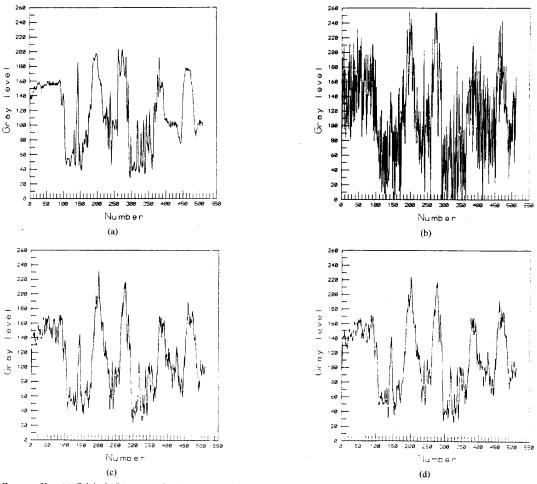
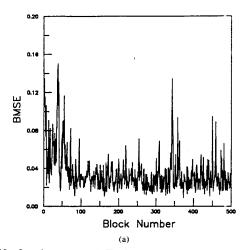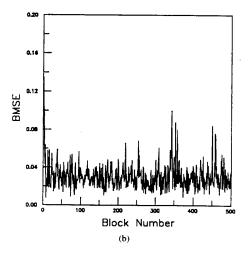Fig. 9. Cross profiles: (a) Original; (b) corrupted; (c) processed, fixed μ_B; (d) processed, variable μ_B.

Fig. 10. Learning curves: (a) Fixed μ_B; (b) variable μ_B.

size μ_B, in order to obtain better results for nonstationary images, this parameter has to be adjusted based on the spatial information in each block.

Let us consider a nonstationary process in the block sense, i.e., the block statistics—block mean, and block covariances change from one block to another. In other words, the mean

Fig. 11.   Processed Lena image using standard Wiener filter (SNR = 7.4 dB).

vector and the block covariance matrix of the process are functions of the spatial coordinates or block indices (i, j). To account for the non-stationarity during the 2-D BDLMS filtering, the filter must track the spatial changes in each block of the input image data. This is accomplished by varying the step size $\mu_B(i, j)$. The optimal step size $\mu_B^*(i, j)$ at each block is obtained based on minimizing the mean squared of the error block with respect to this parameter. In this method, the idea is to express the future error block $E_{i',j'}$ in terms of the current error block $E_{i,j}$ by applying the Taylor series expansion and trying to minimize $E_{i',j'}$ in mean squared sense [10]. Thus, considering only the first two terms in the expansion, the error block $E_{i',j'}$ can be expressed as

$$E_{i',j'} = E_{i,j} + \frac{dE_{i,j}}{dW_{i,j}} \Delta W_{i,j}, \qquad (21a)$$

where $\Delta W_{i,j}$ is given by

$$\Delta W_{i,j} = W_{i',j'} = W_{i,j} = \frac{2\mu_B(i,j)}{KL} \mathbf{X}_{i,j}^t E_{i,j}. \qquad (21b)$$

The differential term in (21a) can be obtained by taking the derivative of (6) i.e.

$$\frac{dE_{i,j}}{dW_{i,j}} = -\mathbf{X}_{i,j}. \qquad (22)$$

Thus, using the expression for the weight difference $\Delta W_{i,j}$ in (21b), we can write

$$E_{i',j'} = E_{i,j} - \frac{2\mu_B(i,j)}{KL} \mathbf{X}_{i,j} \mathbf{X}_{i,j}^t E_{i,j}. \qquad (23)$$

Taking the derivative of the mean squared of the error block with respect to the step size, and then setting it to zero, gives the optimal step size $\mu_B^*(i, j)$ as

$$\mu_{B^*}(i, j) = \frac{KL}{2} \times \frac{E_{i,j}^t \mathbf{X}_{i,j} \mathbf{X}_{i,j}^t E_{i,j}}{E_{i,j}^t \mathbf{X}_{i,j} \mathbf{X}_{i,j}^t \mathbf{X}_{i,j} \mathbf{X}_{i,j}^t E_{i,j}}. \qquad (24)$$

Now, this optimal step size at block (i, j) can be used in conjunction with the adaptation equation (14). Obviously, this method requires considerable computations compared with the fixed step size method since $\mu_B(i, j)$ should be adjusted at each block. As a result, there is a tradeoff between better adaptation characteristic and the computational efforts. The experimental results indicated that $\mu_B^*(i, j)$ often changed abruptly and exceeded the boundaries for convergence. In order to overcome this problem, an acceptable upper limit for the convergence condition for $\mu_B^*(i, j)$ must be determined based on a trial-and-error procedure. In the next section, the simulation results for both the fixed and the variable step size cases are presented.

V. IMPLEMENTATION AND RESULTS

A. Image Estimation and Filtering

In this section, the 2-D BDLMS is used as a smoother to remove the effects of additive noise from an image. The test image "Lena" in Fig. 4 is of size 512 × 512 and has 256 gray levels. The size of the blocks is chosen to be 4 × 4, which is equal to the order of the filter. The block diagram of the 2-D BDLMS when configured as smoother is shown in Fig. 5. The desired image is formed by delaying the reference input image by one pixel in each direction. This is because in practical situations, the desired image is not normally available. Additionally, it can be shown that this provides a fixed-lag smoother with better noise removal capabilities than the filter.

To see this, let us consider the (k, ℓ)th element of the (i, j)th output block. If the reference image $x(m, n)$ consists of a clean image $s(m, n)$ and additive noise $\eta(m, n)$, then using (1), this output pixel becomes

$$\begin{aligned}
y(iK &+ k, jL + \ell) \\
&= \sum_{m=0}^{K-1} \sum_{n=0}^{L-1} s(iK + k - m, jL + \ell - n) w_{i,j}(m, n) \\
&\quad + \sum_{m=0}^{K-1} \sum_{n=0}^{L-1} \eta(iK + k - m, jL + \ell - n) w_{i,j}(m, n) \\
&= \hat{s}(iK + k, jL + \ell) + \hat{\eta}(iK + k, jL + \ell).
\end{aligned} \qquad (25)$$

Note that $y(iK + k, jL + \ell)$ is the fixed-lag smoothed estimate of $d(iK + k, jL + \ell)$, which is actually $x(iK + k - 1, jL + \ell - 1)$. This is because in this process, all observations up to $x(iK + k, jL + \ell)$ are used to generate the estimate of $x(iK + k - 1, jL + \ell - 1)$. The error signal that is the input to the adaptive system is

$$\begin{aligned}
e(iK &+ k, jL + \ell) \\
&= d(iK + k, jL + \ell) - y(iK + k, jL + \ell) \\
&= s(iK + k - 1, jL + \ell - 1) - \hat{s}(iK + k, jL + \ell) \\
&\quad + \eta(iK + k - 1, jL + \ell - 1) - \hat{\eta}(iK + k, jL + \ell) \\
&= \tilde{s}(iK + k - 1, jL + \ell - 1) + \tilde{\eta}(iK + k, jL + \ell).
\end{aligned} \qquad (26)$$

The first term corresponds to the smoothing error in the estimation of the original image, and the second term is the
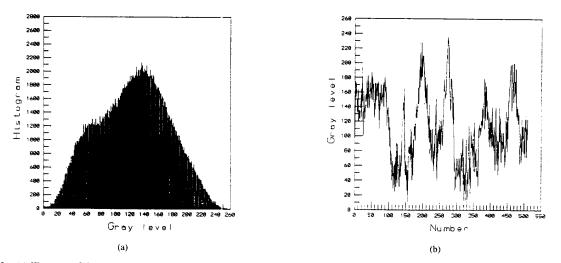
Fig. 12. (a) Histogram of the processed Lena image using Wiener filter; (b) cross profile of the processed Lena image using Wiener filter.
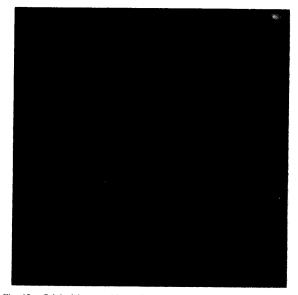

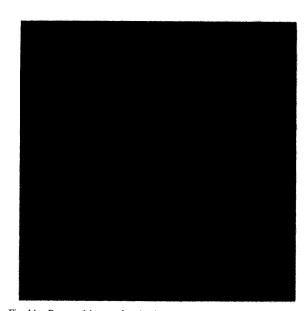
Fig. 13. Original image with small objects.

Fig. 14. Processed image after the detection.

difference between the actual noise and the smoothed (MA processed) noise. As a result, these feedback terms are used to drive the adaptive system and to achieve better image restoration.

The original image was globally corrupted by adding a white Gaussian noise with zero mean to produce signal-to-noise ratio (SNR) of 1.8 dB. The corrupted image is shown in Fig. 6. The processed images using the 2-D BDLMS smoother for both the fixed and the variable step size, are given in Fig. 7(a) and (b), respectively. For the fixed step size case, the SNR of the filtered image was measured to be 9.1 dB, which indicates substantial improvement in the quality of the image. For the variable step size case, the SNR of the image was found to be 9.8 dB, which is slightly better than that of the fixed case.

The histograms of the original, corrupted, and processed images are shown in Fig. 8(a)–(d), respectively, for both the fixed and the variable step size. The cross-section profiles of the original and corrupted images at row 256 (middle of the image) are shown in Fig. 9(a) and (b), respectively. The cross profiles for the processed images using both the fixed and the variable step size are given in Fig. 9(c) and (d), respectively. Comparison of the histograms and the cross profiles of the corrupted and the processed images with those of the original image shows the restoration capability of the adaptive system. The learning curves that describe the variation of BMSE during the operation are also shown in Fig. 10(a) and (b) for the fixed and the variable step size cases, respectively.

The results of the 2-D BDLMS adaptive filter in this paper are compared with those of the standard frequency domain Wiener filter [16], and the processed image is shown in Fig. 11. The SNR of this image is found to be 7.4 dB, which is lower than that of the 2-D BDLMS smoother. The comparison of the histogram and cross-profile plots in Fig. 12 for the standard Wiener filter with those of 2-D BDLMS indicates better restoration and noise removal capability of our approach.

B. Detection of Small Objects in Clutter

The 2-D BDLMS can be configured as a predictor or a prewhitening filter [8] to detect and separate small objects in the background clutter. Such systems can be very useful in medical-imaging applications such as digital mammography where the detection of subtle abnormalities is of prime importance for early detection and treatment of breast cancer. These abnormalities usually have very small spatial spread and are often surrounded and/or obscured by the background clutter with considerably larger spatial extent. In the frequency domain, these abnormalities manifest themselves in the form of wide-band components, whereas the background will have a narrow-band spectrum. A 2-D LMS-based adaptive filter can be used as a prewhitening filter to separate the wide-band components from those of the narrow-band thus isolating the small objects. The output of the system is the estimate of the background image, which is then subtracted from the input image to provide the error image that captures the small objects. The receiver operating characteristic (ROC) curves of this system generated for detection and localization of small objects showed significant improvements over the matched filtering particularly for very low signal-to-background (clutter) ratios [8].

Fig. 13 shows part of a digitized mammogram of size 244 × 244 for a moderately dense normal case. Several (that is, 15) small objects or "abnormalities" of dimensions 1–2 pixels with various intensities were synthetically generated and inserted in the image at different locations. This image was then applied as a reference input to the adaptive system. The desired image in this case is formed by advancing the reference input in both directions in order to construct a predictor system. The block size of 4 × 4 and a two-step predictor were empirically determined to provide acceptable results. The thresholded resultant image (error image) is shown in Fig. 14. As can be observed, all the objects were successfully detected in this case. Of course, the detectibility of the objects depends on several factors including signal-to-clutter ratio, competing clutter with similar shapes as those of the actual objects, obscuration of the objects in clutter, etc. A total of two false positives were detected for this case.

Future work should focus on studying real normal/abnormal cases and on determining the robustness and the ROC [8] of the proposed approach.

VI. CONCLUSION

A new 2-D BDLMS adaptive algorithm is developed in this paper by extending 1-D block LMS adaptive filter to the 2-D case. The advantage of this 2-D BDLMS algorithm over the scalar 2-D LMS algorithm is that it takes into account image correlations along both directions. Convergence properties of the 2-D BDLMS are discussed. Although the convergence behavior of the 2-D BDLMS is equivalent to that of its scalar counterpart, the BDLMS algorithm can be implemented faster than the scalar ones as in the 1-D case [3]. To account for nonstationarity in the image, an adaptation rule for step size is derived using a similar approach as in [10], which adjusts this parameter based on the spatial activities within each processing window. This provides better tracking characteristics of the adaptive filter. Simulation results that show the usefulness of the proposed scheme for image filtering and detection applications are presented.

REFERENCES

[1] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[2] B. Widrow, J. M. McCool, M. G. Larimore, and C. R. Johnson, Jr., "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
[3] G. A. Clark, S. K. Mitra, and S. R. Parker, "Block implementation of adaptive digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, no. 3, pp. 774–752, June 1981.
[4] C. Burrus, "Block implementation of digital filters," *IEEE Trans. Circuit Theory*, vol. CT-18, pp. 697–701, Nov. 1971.
[5] W. B. Mikhael and F. H. Wu, "Fast algorithms for block FIR adaptive digital filtering," *IEEE Trans. Circuits Syst.*, vol. CAS-34, no. 10, pp. 1152–1160, Oct. 1987.
[6] W. B. Mikhael and F. H. Wu, "A fast block FIR adaptive digital filtering algorithm with individual adaptation of parameters," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1–10, Jan. 1989.
[7] M. M. Hadhoud and D. W. Thomas, "The two-dimensional adaptive LMS (TDLMS) algorithm," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 485–494, May 1988.
[8] T. Soni, J. R. Zeidler, and W. H. Ku, "Performance evaluation of 2-D adaptive prediction filters for detection of small objects in image data," *IEEE Trans. Image Processing*, vol. 2, pp. 327–339, July 1993.
[9] T. Soni, B. D. Rao, J. R. Zeidler, and W. H. Ku, "Enhancement of images using the 2-D LMS adaptive algorithm," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* (Toronto), May 1991, pp. 3029–3032.
[10] W. B. Mikhael and S. M. Ghosh, "Two-dimensional variable step-size sequential adaptive gradient algorithms with applications," *IEEE Trans. Circuits Systems*, vol. 38, pp. 1577–1580, Dec. 1991.
[11] ———, "Two-dimensional block adaptive filtering algorithms," *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS'92)* (San Diego), May 1992, pp. 1219–1222.
[12] M. R. Azimi-Sadjadi and R. A. King, "Two-dimensional block processors—Structures and implementations," *IEEE Trans. Circuits Systems*, vol. CAS-33, pp. 42–50, Jan. 1986.
[13] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*. New York: McGraw-Hill, 1977.
[14] H. Pan, "Two-dimensional block diagonal least mean square adaptive filtering," M.S. Thesis, Colorado State Univ., Fort Collins, CO, Dec. 1991.
[15] T. C. Hsia, "Convergence analysis of LMS and NLMS adaptive algorithm," *Proc. IEEE Int. Conf. Acoust. Speech Signal Processing*, Apr. 1983, pp. 667–670.
[16] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

Mahmood R. Azimi-Sadjadi (SM'89) was born in Tehran, Iran, in 1952. He received the B.S. degree from the University of Tehran, Iran, in 1977 and the M.Sc. and Ph.D. degrees from the Imperial College, University of London, England, in 1978 and 1982, respectively, all in electrical engineering.

He served as an Assistant Professor in the Department of Electrical and Computer Engineering at the University of Michigan–Dearborn. Since July 1986, he has been with the Department of Electrical Engineering at Colorado State University, Fort Collins, where he is now an Associate Professor. His areas of interst are in digital signal/image processing, multidimensional system theory and analysis, adaptive filtering, system identification, and neural networks. He coauthored the book *Digital Filtering in One and Two Dimensions* (New York: Plenum, 1989).

Dr. Azimi-Sadjadi was the recipient of the 1990 BATTELLE Faculty Felowship Award and the 1984 Dow Chemical Outstanding Young Faculty Award of the American Society for Engineering Education.

Hongye Pan received the B.Sc. degree in comunication engineering from Chengdu Institute of Electrical Engineering, Chengdu, China, in 1982 and the M.Sc. degree in electrical engineering from Colorado State University, Fort Collins, in 1991.

She is currently with Analytical Technologies, Inc., as a data analyst. Her research interests include multidimensional digital filters, adaptive filters for digital signal processing, and software design for circuit and system simulations.