Adaptive Polyphase Subband Decomposition Structures for Image Compression

Ömer Nezih Gerek, *Member, IEEE,* and A. Enis Çetin, *Member, IEEE*

Abstract—**Subband decomposition techniques have been extensively used for data coding and analysis. In most filter banks, the goal is to obtain subsampled signals corresponding to different spectral regions of the original data. However, this approach leads to various artifacts in images having spatially varying characteristics, such as images containing text, subtitles, or sharp edges. In this paper, adaptive filter banks with perfect reconstruction property are presented for such images. The filters of the decomposition structure which can be either linear or nonlinear vary according to the nature of the signal. This leads to improved image compression ratios. Simulation examples are presented.**

Index Terms—**Adaptive polyphase structures, adaptive subband decomposition, image coding, lifting structures.**

I. INTRODUCTION

SUBBAND decomposition is widely used in signal processing applications including speech, image and video compression. In most practical cases, the goal is to obtain subband signals corresponding to different spectral regions of the original signal. The frequency content of most audio and visual data are suitable for this kind of frequency selective coding. However, this approach leads to ringing artifacts in image and video signals containing text, subtitles or sharp edges. The ringing is mainly due to constant analysis filter banks which cannot cope with the sudden changes in the input signal.

In this paper, we present perfect reconstruction (PR) polyphase filter bank structures in which the filters adapt to the changing input conditions. The adaptation of the analysis filter bank leads to higher compression results for images containing sharp edges, text, and subtitles. Since most of the disturbing ringing artifacts occur on the boundaries of subtitles, texts and sharp edges, an adaptive filter bank can update its coefficients accordingly and can reduce the disturbing overshoots at the edges. Furthermore, most images and video signals are nonstationary in nature, therefore an adaptive filter bank can achieve higher efficiency than a fixed filter bank. The polyphase filter bank structures that we introduce allow the use of both the least mean squares (LMS) type FIR [1], [2] and nonlinear order

statistics [3] based adaptive filters to cope with the changes in the input signal.

The concepts of adaptive filtering and subband decomposition [4], [5] have been previously used together by a number of researchers [6]–[11]. Most of the proposed adaptation algorithms for subband decomposition filter banks consider the problem of system identification and noise removal [6]–[11]. In these works, the adaptive filtering problem is considered in the subband domain. The issues of efficient complex or real valued filter design methods to increase subband domain adaptive filtering performance are also investigated [6], [7]. In this approach, the design of the filter bank which satisfies pre-specified requirements for adaptive filtering in subbands is studied. In [12]–[17] the unknown system outputs are used for adapting the analysis filter bank coefficients so that the filter bank approximates an unknown system.

The choice of subband filter banks according to the input signal is also considered by some researchers [12]–[15]. The main goal of these works is to find the best wavelet basis for decomposing the data. For example, the autocorrelation matrix of the image data is used for determining a good basis for decomposition [13], [18], [19]. In [16] and [17], optimal coding after decomposition is considered, and optimum quantizers and optimum entropy coders are studied.

In [12]–[15], fixed filters chosen according to an optimality criterion are used throughout the entire duration or extent of the signal whereas, in this paper, the filters vary as the nature of the input signal changes. We propose an adaptation scheme for updating the filter coefficients of the subband decomposition filter bank which has a polyphase format with the prediction portion of a lifting stage [20]–[22]. In this aspect, it is different from the subband adaptive filter structures which perform adaptive filtering in the subbands [6]–[11]. The adaptation scheme in our method neither tries to estimate an unknown system nor uses a fixed filter bank throughout the entire duration of the signal. Since the problem we address here is the coding of the input data, the filter coefficients are updated to remove the unnecessary information among the neighboring subsignals. The aim is to obtain decorrelated subsignals. Due to the nonstationary characteristics of most image data, this improves the coding efficiency. In this aspect, the work in [24] is related with our work. In [24], previously determined linear and nonlinear filters were used in a switchable manner in different regions of the image. In our work, there is no need to select a pre-determined filter in different regions of the image because the proposed adaptive subband decomposition scheme inherently updates the filter banks and finds ideal filters for each signal sample while preserving the perfect reconstruction property.

Fig. 1.   Polyphase decomposition.

Fig. 2.   Analysis stage of the adaptive filter bank structure.
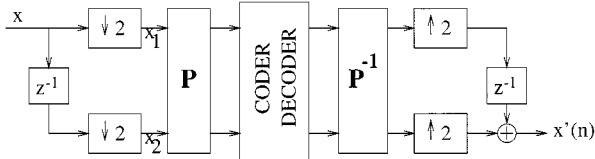
In Section II, we review the PR polyphase structure concept [4], [5] and present a procedure to make it adaptive. In this section, multichannel extensions of adaptive filter banks are also presented. As discussed above, either linear or nonlinear filters can be used in the decomposition structure without disturbing the PR property. In Section III, we describe another adaptive polyphase structure which contains a fixed antialiasing filter for the upper branch and an adaptive prediction filter for the lower branch. This structure is especially useful when a multiresolution viewing feature is needed. The coding gain analysis of the adaptive structure by means of the reduction of the variance and various adaptation schemes is described in Section IV. Simulation examples and image compression results are given in Section V. The incorporation of the quantization effects in the adaptive decomposition stage to obtain improved low bit rate performance is introduced in Section VI, and conclusions are presented in Section VII.

II. ADAPTIVE PREDICTION FILTERS IN POLYPHASE FORM

The block diagram of the basic two-band polyphase subband structure is shown in Fig. 1. In this structure, the input polyphase components x_1 and x_2 are multiplied by a 2×2 matrix, $\boldsymbol{P}_1$. For perfect reconstruction, the only constraint on this matrix is invertibility. One can try to optimize the $\boldsymbol{P}$ matrix according to the application. In fact, this way of utilizing the polyphase components is investigated in the framework of lifting schemes [20]–[22] and perfect-inversion polyphase networks [23].

In the next subsection, we introduce a class of polyphase structures in which the $\boldsymbol{P}$ matrix is not fixed, and describe how the filters that form $\boldsymbol{P}$ can be chosen.

A. Basic Polyphase Filter Bank Structure

Consider the following choice for the matrix $\boldsymbol{P}$:

$$\boldsymbol{P} = \begin{bmatrix} 1 & -P_1(.) \\ 0 & 1 \end{bmatrix}. \tag{1}$$

In (1), the filter P_1 need neither be a fixed nor a linear operator for perfect reconstruction as $\boldsymbol{P}$ is invertible regardless of the nature of P_1. Furthermore, the PR property is preserved as $\boldsymbol{P}$ is invertible at all time instants. Therefore, both nonlinear filters and time varying filters can be used in this structure. This stage can be considered as the prediction portion of the lifting stage [20]–[22].

The inverse of the $\boldsymbol{P}$ matrix in (1) is given as

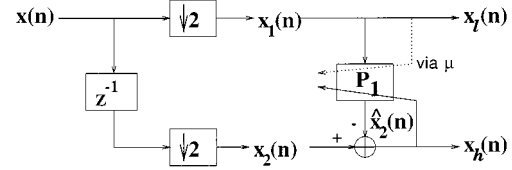$$\boldsymbol{P^{-1}} = \begin{bmatrix} 1 & P_1(.) \\ 0 & 1 \end{bmatrix}. \tag{2}$$

In this case, the low-band signal x_l is obtained by down-sampling the original signal, x and it is directly passed to the encoder. Therefore, a good way of obtaining the subsignal, $x_h(n)$, is to predict the samples of the second polyphase component x_2 from the first polyphase component x_1 which is equal to x_l. This approach is suitable for coding applications, in which the goal is to remove the correlated portion of the original signal as much as possible. In this way, the correlation between the channels is eliminated. The prediction branch of the polyphase system can be interpreted as the half stage of a lifting scheme, without the update branch.

Usually, the prediction filters are of low-pass nature, because the samples of x_2 are strongly correlated to x_1 for image data.

A predictor should be adaptive for image and video signals as they are nonstationary in nature. This reasoning leads to the polyphase structure shown in Fig. 2 in which the prediction filter adapts itself to minimize the high-band signal $x_h(n)$. This is especially useful when there are sharp transition regions in an image such as subtitles, text and graphics.

B. Adaptive Filter Bank Structure

The adaptive estimator for $x_h(n)$ is shown in Fig. 2. This FIR estimator is obtained by predicting $x_2(n)$ from $x_1(n)$ in a linear minimum mean squared error (LMMS) sense as follows:

$$\hat{x}_2(n) = \sum_{k=-N}^{N} w_{n,k} x_1(n-k) = \sum_{k=-N}^{N} w_{n,k} x(2n-2k) \tag{3}$$

where the filter coefficients $w_{n,k}$'s are updated using an LMS-type algorithm [27].

The FIR LMS adaptation is performed in a conventional manner as follows:

$$\hat{\boldsymbol{w}}(n+1) = \hat{\boldsymbol{w}}(n) + \mu \frac{\tilde{\boldsymbol{x}}_n e(n)}{\|\tilde{\boldsymbol{x}}_n\|^2} \tag{4}$$

where $\hat{\boldsymbol{w}}(n) = [w_{n,-N}, \cdots, w_{n,N}]$ is the weight vector at time instant n

$$\tilde{\boldsymbol{x}}_n = [x_1(n-N), x_1(n-N+1), \cdots, x_1(n+N-1),$$
$$x_1(n+N)]^T. \tag{5}$$

The subsignal x_h is given by

$$x_h(n) = x_2(n) - \hat{x}_2(n). \tag{6}$$

and

$$e(n) = x_h(n) = x_2(n) - \tilde{\boldsymbol{x}}_n^T(n)\hat{\boldsymbol{w}}(n). \tag{7}$$

Both $\mathcal{L}^1$ and $\mathcal{L}^2$ norms can be used in normalizing the update equations depending on the characteristics of the signal [27]. In our simulations, these norms were successfully used [27].

The scalar μ determines the step size of the adaptive algorithm. It is well known that the convergence speed of adaptation is low when μ is small, but the steady state error is smaller.

For large values of μ, the opposite happens and the convergence speed increases with a higher steady-state error. There are various methods to change the value of μ during adaptation in the LMS algorithm [29], [30]. Usually, the value of μ can be set to a large number between 1 and 2 in the beginning and, it can be gradually decreased to a smaller value between 0 and 1. In our case, the value of μ is determined according to the range of the input. Since the input data $\tilde{x}_1$ is available at the decoder side, the decoder can alter the parameter of its μ value for reconstruction, accordingly. The actual update equation is given by

$$\hat{w}(n+1) = \hat{w}(n) + \mu(\tilde{x}_n)\frac{\tilde{x}_n e(n)}{\|\tilde{x}_n\|^2} \qquad (8)$$

where

$$\mu(\tilde{x}_1) = \begin{cases} 0.4, & \mathbf{\Delta_{\tilde{x}}} < 10 \\ 0.6, & 10 \le \mathbf{\Delta_{\tilde{x}}} < 30 \\ 0.8, & 30 \le \mathbf{\Delta_{\tilde{x}}} < 80 \\ 1.0, & 80 \le \mathbf{\Delta_{\tilde{x}}} < 200 \\ 1.2, & 200 \le \mathbf{\Delta_{\tilde{x}}} < 256 \end{cases} \qquad (9)$$

and

$$\mathbf{\Delta_{\tilde{x}}} = \max(\tilde{x}_n) - \min(\tilde{x}_n). \qquad (10)$$

The piecewise constant values of the μ parameter are determined by minimizing the experimental error over a range of artificially generated signals producing the above Δ values. The dependency of the overall minimization on the μ value can be more precisely determined, but it is out of the scope of this paper.

In order to avoid extreme overshoots in filter tap updates, thresholds are used [31] both in the encoder and in the decoder. The reason to put such thresholds is to avoid divergence at very low bit rates which require coarse quantization of the transform data. In our simulation studies, we used a limiting threshold of -256 and 256 for each filter tap for image coding applications.

The PR property is preserved in this structure as long as the same adaptation algorithm is used at the encoding and the decoding stage. Since, for the no-quantization case, the subsignal $x_h(n)$ as well as $\tilde{x}_n$ are available both at the encoder and at the decoder, the synthesis stage can adapt the filter P_1 with the same filter tap coefficients $\hat{w}(n)$. Therefore, no side information needs to be transmitted. Practically, the quantized versions of the subsignals are transmitted, therefore the reconstruction stage uses the quantized coefficients to achieve a lossy compression.

It was observed in [3], [24], and [25] that, in various applications, including coding applications, the order statistics (OS) filters and especially the median filter perform better than the linear FIR filters for the images containing sharp variations like text. Furthermore, in [26], it was shown that generalized OS decomposition is a useful tool in image restoration. This observation motivates the use of adaptive OS filters in the structures shown in Figs. 2 and 3. The rank ordering of the input elements produces better coding results especially for the images that contain sharp edges.

The implementation of the OS type adaptation is similar to the linear FIR filter coefficient update. For the OS case, the input vector $\tilde{x}_n$ is first rank ordered. The largest and the smallest values of the vector are removed. As a result, another vector with a shorter size is obtained. This vector is then used as an input
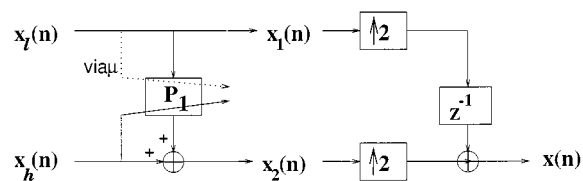

Fig. 3. Synthesis stage of the adaptive filter bank structure.

to the update equations (7) and (8) for adapting the filter coefficients. In our simulation studies, a region of support with nine elements is used.

C. Cascaded Adaptive PR Blocks

The structure described in Section II-A can be generalized by cascading matrices similar to the matrix in (1). The analysis and synthesis stages of the cascaded two band decomposition structure can be generated using Equations (1) and (2). The overall cascaded transformation matrix is obtained by multiplying triangular matrices which correspond to basic building blocks as follows:

$$P = \begin{bmatrix} 1 & -P_1(.) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ G_1(.) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -P_2(.) \\ 0 & 1 \end{bmatrix} \times \cdots \quad (11)$$

where the filters P_1, G_1, P_2, $\cdots$ can be linear, nonlinear or adaptive. In this way, the upper and lower branch subsignals can be filtered a number of times. The inverse matrix is given as

$$P^{-1} = \cdots \times \begin{bmatrix} 1 & P_2(.) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -G_1(.) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & P_1(.) \\ 0 & 1 \end{bmatrix}. \quad (12)$$

The synthesis filter bank can be easily reconstructed from P^{-1} as shown in Fig. 4.

D. Multichannel Extension of the PR Structure

The filter bank structures described in Section II can be extended to handle decompositions to bands other than the powers of two. The extension can be performed in various ways. Consider the multiband decomposition structure shown in Fig. 5.

In Fig. 5, an M band decomposition with two cascaded PR building blocks is illustrated. The PR property of this structure can be proved easily. In the analysis stage

$$\begin{aligned} v_1 &= x_1 \\ v_i &= x_i - P_{i-1}(v_{i-1}), & i = 2, 3, \cdots, M \\ y_i &= v_i + G_i(v_{i+1}), & i = 1, 2, \cdots, M-1 \\ y_M &= v_M. \end{aligned}$$

The corresponding P matrix for this case is given by

$$P = \begin{bmatrix} 1 & -P_1 & 0 & 0 & \cdots \\ 0 & 1 & -P_2 & 0 & \cdots \\ 0 & 0 & 1 & -P_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}$$

$$\times \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ G_1 & 1 & 0 & 0 & \cdots \\ 0 & G_2 & 1 & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix}. \qquad (13)$$
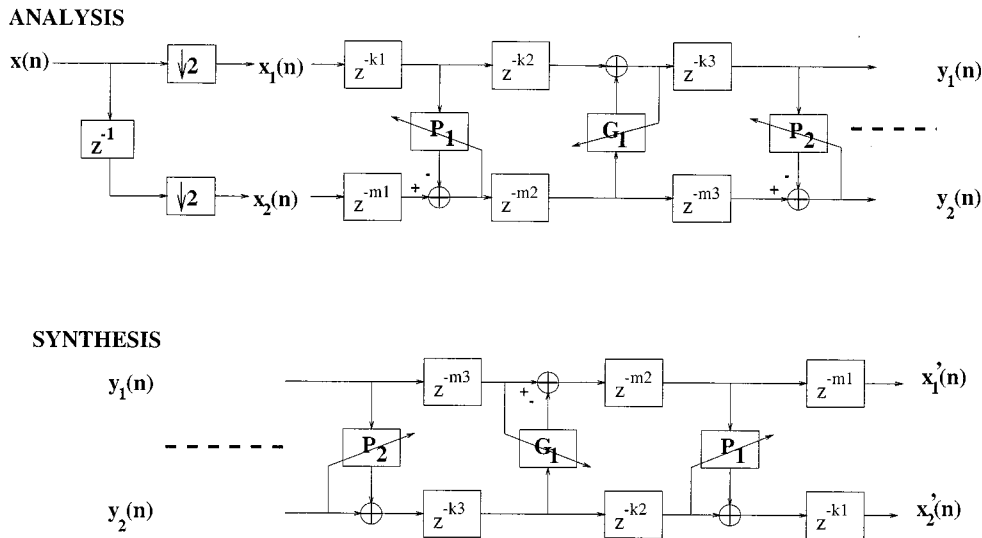
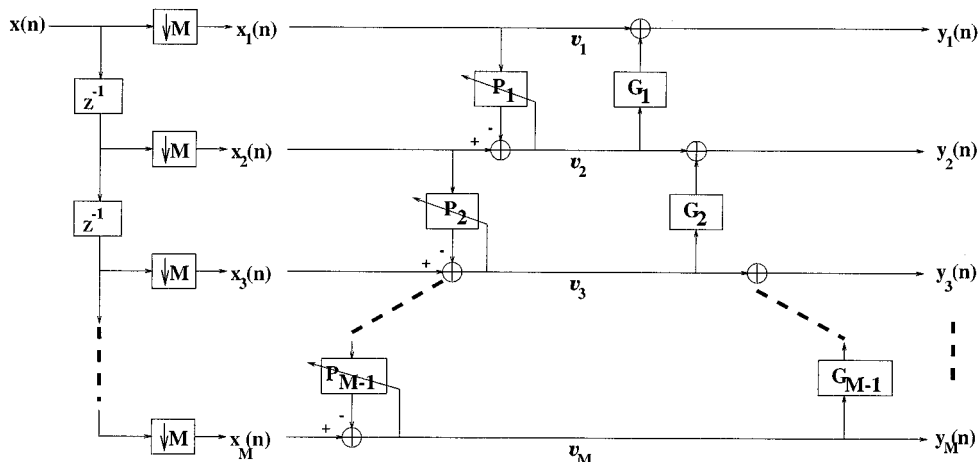ANALYSIS

SYNTHESIS

Fig. 4.   Cascaded polyphase filters.

Fig. 5.   Multichannel analysis structure-1 of the adaptive filter bank structure.

Since the matrix P is formed by multiplying upper and a lower triangular matrices, it can be inverted regardless of the filters P_i's and G_i's. Therefore, PR can be achieved with any choice of the nonlinear operators. This leads to the following synthesis equations:

$$v'_M = v_M$$
$$v'_i = y_i - G_i(v'_{i+1}) = v_i, \qquad i = M - 1, \cdots, 1$$
$$v'_1 = v_1 = x_1$$
$$x'_i = v'_i + P_{i-1}(v'_{i-1}) = v_i + P_{i-1}(v_{i-1}) = x_i,$$
$$\qquad i = 2, \cdots, M.$$

The outputs, x'_i, of the synthesis filters are the same as the polyphase components, x_i, of the analysis filter bank.

Another multichannel extension structure is shown in Fig. 6. In the previous structure, only the samples of x_k was considered to estimate x_{k+1}, $k = 2, 3, \cdots, M$. On the other hand, the structure in Fig. 6 uses all of the previous polyphase components

for prediction as the index of the subsignals increase in Fig. 6. The analysis equations for this structure are given as follows:

$$v_1 = x_1$$
$$v_i = x_i - P_{i-1}(v_1, v_2, \cdots, v_{i-1})$$
$$y_M = v_M$$
$$y_i = v_i + G_i(y_M, \cdots, y_{i+1}), \qquad i = 1, 2, \cdots, M - 1.$$

The synthesis equations are given by

$$v'_i = y_i - G_i(y_M, \cdots, y_{i+1}) = v_i,$$
$$i = 1, 2, \cdots, M - 1$$
$$v'_M = y_M = v_M$$
$$x'_1 = v_1 = x_1$$
$$x'_i = v_i + P_{i-1}(v_1, v_2, \cdots, v_{i-1}) = x_i,$$
$$i = 2, 3, \cdots, M.$$

This later structure also yields analysis matrices which can be decomposed to upper and lower triangular matrices with elements containing P_i's and G_i's only. In this structure, for predicting v_i's, the number of data used increases with increasing
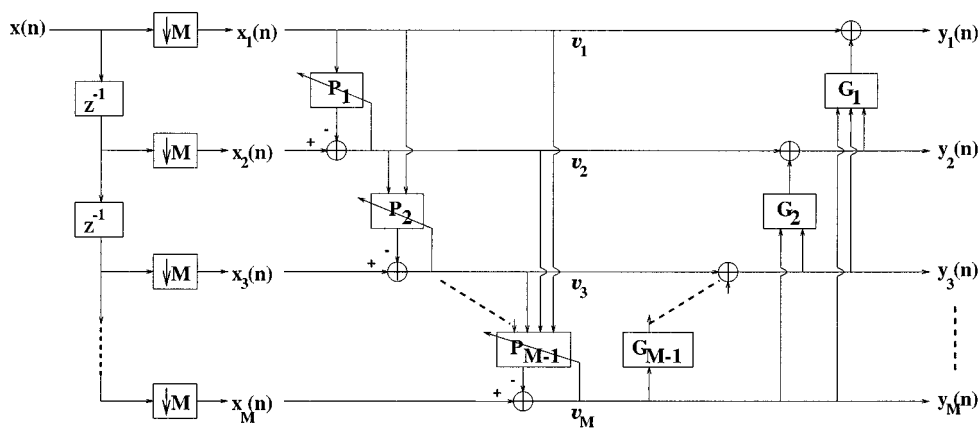
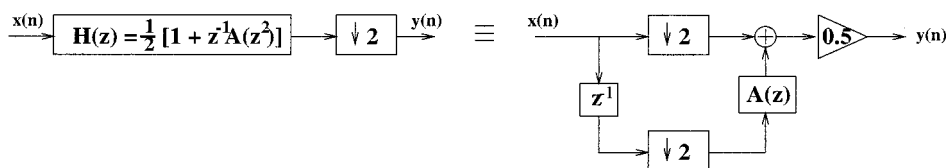Fig. 6. Multichannel analysis structure-2 of the adaptive filter bank structure.

Fig. 7. Two equivalent structures for low-pass filtering and downsampling.

index i. Conversely, more v_i samples are used for predicting y_i's when the index i is small. The computational complexity of this structure is high as compared to the structure in Fig. 5.

III. ADAPTIVE PR STRUCTURE WITH AN ANTI-ALIASING FILTER

In many applications, multiresolution display of an image is a desirable property. Since $x(n)$ is simply down-sampled in the upper branch of Fig. 2, the visual quality of the subsignal $x_l(n)$ is poor due to aliasing.

In order to remove the aliasing, a two-stage cascaded P matrix can be used. The matrix P should be designed in such a way that the first stage should reduce the aliasing and the second stage should produce a good "high-band" signal.

In a typical QMF structure, the input signal is low-pass filtered before down-sampling to eliminate aliasing. If the low-pass filter is a half-band filter [4], [33], [34], i.e., $H(z) = (1/2)[1 + z^{-1}A(z^2)]$, then the so called "noble identity" [4] can be used and the filtering operations can be carried out after down-sampling as shown in Fig. 7.

The second stage of the analysis structure consists of adaptive prediction of subsignal $x_h(n)$. In this case, the samples of the low-pass filtered subsignal $x_l(n)$ are used to predict $x_h(n)$. The overall analysis structure is shown in Fig. 8. Perfect reconstruction can be achieved using the synthesis block shown in Fig. 9.

The analysis polyphase structure has the following matrix:

$$P = \begin{bmatrix} 1 & 0 \\ A(z) & 1 \end{bmatrix} \times \begin{bmatrix} 0.5 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -P_1(.) \\ 0 & 1 \end{bmatrix} \quad (14)$$

and the synthesis matrix is simply

$$P^{-1} = \begin{bmatrix} 1 & P_1(.) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -A(z) & 1 \end{bmatrix}. \quad (15)$$
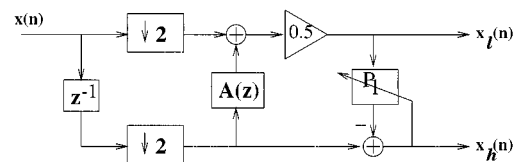

Fig. 8. Adaptive filter bank structure with an antialiasing filter.
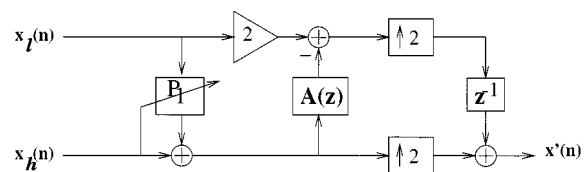
Fig. 9. Synthesis stage corresponding to Fig. 8.

In our simulation studies, we use the half-band Lagrange family for low-pass filtering [33], [34]. As an example, the $A(z)$ polynomials of the first two Lagrange filters with sizes three and seven are

$$A(z) = \tfrac{1}{2} + \tfrac{1}{2} z^1$$

and

$$A(z) = -\tfrac{1}{16} z^{-1} + \tfrac{9}{16} + \tfrac{9}{16} z^1 - \tfrac{1}{16} z^2$$

respectively.

IV. CODING GAIN ANALYSIS

The structure in Fig. 2 can be considered as a transformation of the input signals x_1 and x_2. For stationary signals and with an orthogonal transformation, it was shown in [35] that the ratio of the error variance of the original quantized signal σ_{PCM}^2 to the error variance of the transform domain quantized signal σ_{SBC}^2

gives the orthogonal transform coding gain of a stationary signal as

$$G_{SBC} = \frac{\sigma_{tot}^2}{(\sigma_l^2 \sigma_h^2)^{(1/2)}}. \tag{16}$$

In general, the coding performance is related with the reduced variance of the transform domain signals. In our case, the polyphase components $x_1(n)$ and $x_2(n)$ have the same variance as the input signal, each having half of the signal size. Using the adaptive decomposition scheme, one of the polyphase components is kept to have the same variance, and the variance of the other component is approximately minimized by a gradient method. The variance of the original sequence is (with zero mean assumption)

$$\sigma_{tot}^2 = \frac{1}{N} \left(\sum_{n=0}^{N/2-1} x_1^2(n) + \sum_{n=0}^{N/2-1} x_2^2(n) \right) = \frac{1}{N} \sum_{n=0}^{N-1} x^2(n)$$

and the variance of the transform domain signal is

$$\sigma_{trans}^2 = \frac{1}{N} \left(\sum_{n=0}^{N/2-1} x_l^2(n) + \sum_{n=0}^{N/2-1} x_h^2(n) \right).$$

Since $\sum_{n=0}^{N/2-1} x_h^2(n)$ is reduced by the gradient algorithm, $\sum_{n=0}^{N/2-1} x_h^2(n) \leq \sum_{n=0}^{N/2-1} x_2^2(n)$, so $\sigma_{trans}^2 \leq \sigma_{tot}^2$.

The minimization of the variance of x_h can be explained by the very nature of the LMS type adaptation. The adaptive subband decomposition scheme in Fig. 2 tries to predict the values of x_2 using an LMS type adaptive filter. The LMS filter is a gradient estimator which tries to minimize the mean squared error J

$$J = E\left[(x_2(n) - \boldsymbol{w}^T \tilde{\boldsymbol{x}}_n)(x_2(n) - \tilde{\boldsymbol{x}}_n^T \boldsymbol{w})\right] \tag{17}$$

$$= \sigma_{x2}^2 - \boldsymbol{w}^T \boldsymbol{p} - \boldsymbol{p}^T \boldsymbol{w} + \boldsymbol{w}^T \boldsymbol{R} \boldsymbol{w} \tag{18}$$

where

σ_{x2}^2 variance of the subsignal x_2;
$\boldsymbol{p}$ cross correlation vector between input vector $\tilde{\boldsymbol{x}}_n$ and signal x_2;
$\boldsymbol{R}$ correlation matrix of the input vector $\tilde{\boldsymbol{x}}_n$;
   vector instantaneous filter coefficients at the time instant n.
$\boldsymbol{w}$

The high-band signal is given by

$$x_h(n) = x_2(n) - \boldsymbol{w}^T(n)\tilde{\boldsymbol{x}}_n \tag{19}$$

where $\tilde{\boldsymbol{x}}_n = [x(n-N), x(n-N+1), \cdots, x(n+N-1), x(n+N)]$.

When the input signal is stationary, the LMS filter converges to the optimum vector $\boldsymbol{w}_0$ which gives the minimum mean-squared error

$$J_{\min} = \sigma_2^2 = \sigma_{x2}^2 - \boldsymbol{w}_0^T \boldsymbol{R}^T \boldsymbol{w}_0. \tag{20}$$

Therefore, the variance σ_2^2 is minimized, and the overall variance of the transform domain signal is reduced.

In a typical image, there is high correlation between the neighboring samples. The redundancy between the subbands is eliminated by the adaptive prediction. In general, signals may have parts that are predictable in different frequency bands. Therefore, the adaptive scheme is different from the regular QMF filter banks where the decomposition is only in terms of spectral separation. Our adaptive decomposition method is also
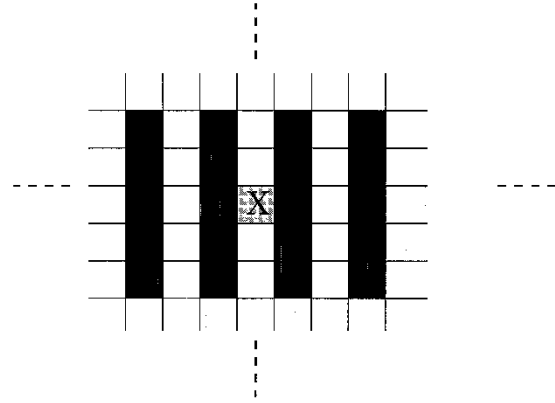

Fig. 10. Filter support for 2-D separable prediction. Gray-labeled pixel is to be predicted from the black-labeled pixels.
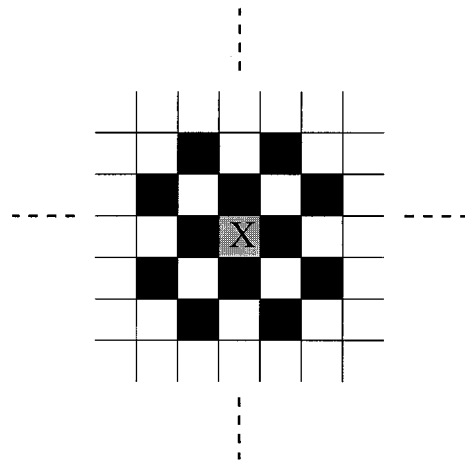

Fig. 11. Filter support for 2-D quincunx prediction. Gray-labeled pixel is to be predicted from the black-labeled pixels.

different from the eigen-decomposition based methods which require transformation matrices or correlation matrices both at the encoder and decoder [36] which makes compression very difficult and impractical.

For the filter bank with the antialiasing stage described in Section III, the coding gain is obtained by multiplying coding gains of the antialiasing filter block and the adaptive filter block. The antialiasing filter bank is a subband filter bank that applies a low-pass filter to the upper branch of the subband signals. In this case, some of the high-pass information content of the upper branch x_1 is eliminated from the signal by this filtering operation. This information loss appears as a slightly lower prediction performance at the adaptation stage. Yet, the decrease in compression ratio due to the lower prediction performance is very small because most of the information to predict x_2 from x_1 is in the low-pass portion of the signals for most images. The slight decrease in compression can be acceptable at the expense of a multiresolution previewing feature.

V. TWO-DIMENSIONAL FILTER BANK STRUCTURES

The extension of the adaptive structure to the two-dimensional (2-D) case is needed for image coding purposes. A straightforward 2-D generalization can be achieved by applying
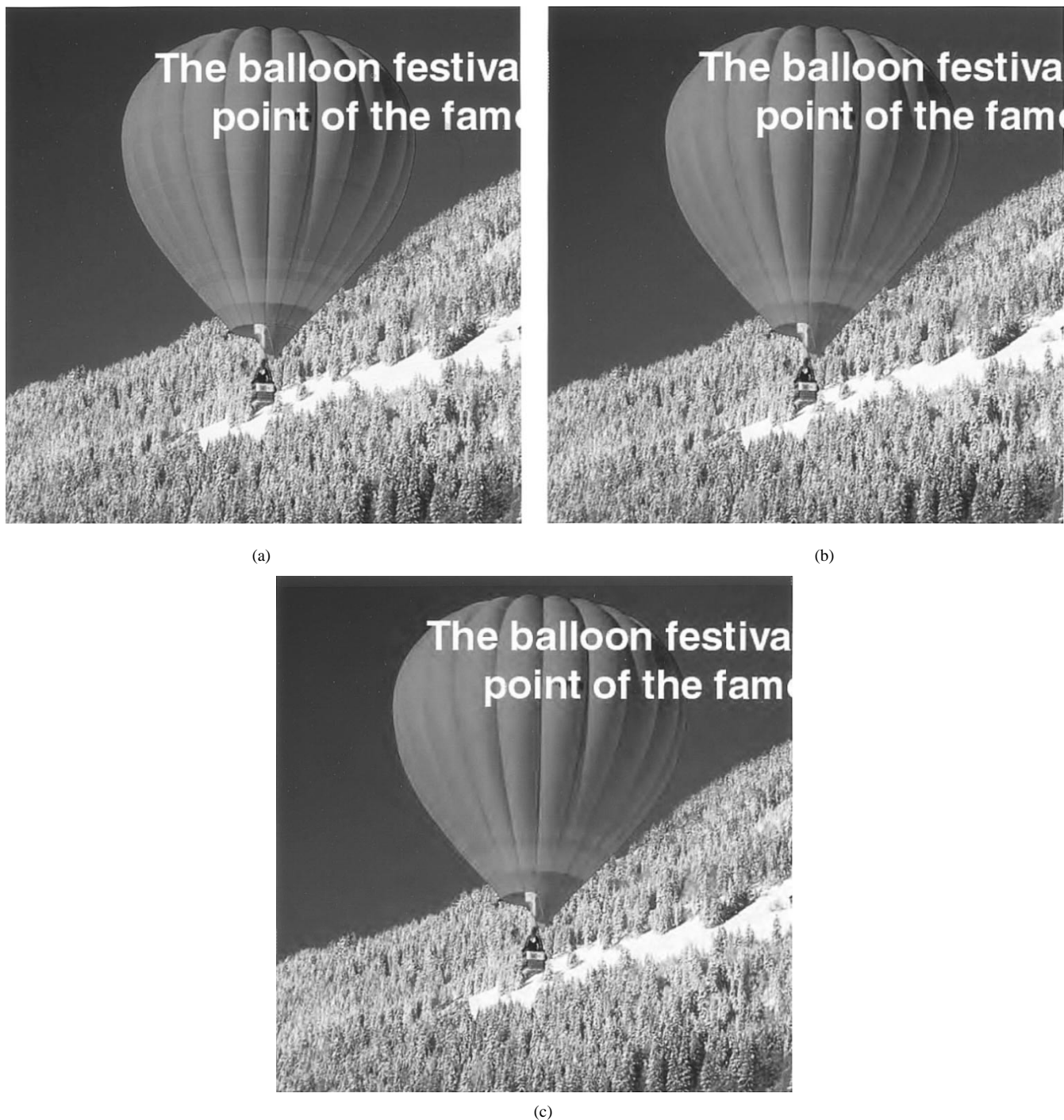
(a)

(b)

(c)

Fig. 12.   Test image containing text. (a) Original, (b) 2-D LMS adaptive at 1 bpp, and (c) EZW at 1 bpp.

one-dimensional (1-D) filters to the image data in a separable manner. In this way, first the columns of the image are filtered, then this data is row-wise processed. This is a conventional method to implement multidimensional filters with 1-D modules.

However, better prediction performance can be achieved than consecutive 1-D row-wise and column-wise processing. Consider the region of support shown in Fig. 10 for horizontal processing. The gray pixel can be predicted from the black pixels using an LMS adaptive algorithm. Since more samples are used in the support region, better prediction performance is achieved. Once the row-wise processing is finished, the

column-wise adaptive filtering is carried out. In our simulation studies, the region of support in Fig. 10 is used. It is also experimentally observed that this produces better coding results.

A frequently used nonseparable downsampling method for images is the "quincunx" downsampling. The region of support of the prediction filter can readily be extended to the quincunx downsampling method as shown in Fig. 11.

VI. SIMULATION STUDIES

In this section, image compression examples using the adaptive subband filter banks are presented. In high quality

TABLE I
EXPERIMENT RESULTS (PSNR) FOR 5 LEVEL DECOMPOSITION OF
THE TEST IMAGE AT 1 BPP

P_1 filter	Plain Downsampling	Antialiased Downsampling
Median	30.44	30.35
Adaptive FIR	32.08	31.96
Adaptive OS	32.28	32.21

image coding applications, the adaptive filter bank produces images with higher PSNR compared to fixed filter banks. This improvement is also visible due to the elimination of the ringing effects. For images containing text and sharp variations, the PSNR improvement is higher. Specifically, it was demonstrated in [37] that the nonlinear adaptive scheme is suitable for compressing fingerprint images which usually have sharp variations.

In the following simulation studies, we used the embedded zero-tree (EZT) coder to encode the transform coefficients [38]. Due to the characteristics of EZT, the best coding results were obtained by the tree-structured two-band decompositions.

The coding results for the 512×512 image shown in Fig. 12(a) at 1 bit/pixel (bpp) bit-rate are given in Table I. In all of the experiments presented in the table, a separable 2-D prediction geometry similar to Fig. 10 is used. Specifically, the resultant prediction filter was using 12 coefficients. The LMS type adaptation with 12-tap filter is a slower operation than using a fixed prediction filter of the same size, and on SUN ULTRA computers, the adaptive operation is approximately three times slower than the fixed linear prediction. However, obtaining pixels from the 2-D prediction geometry also adds some overhead. The fixed prediction scheme with 12-tap 2-D prediction is about 1.5 times slower than the conventional row-by-row filtering. The coder for compressing the transform coefficients is the same in all our experiments, therefore it contributes the same amount of computation time to the overall process. Typically, the overall computation time for compressing an image using the adaptive scheme is approximately three times the computation time for compressing the same image using the conventional EZW.

The first column of the table shows the results without using the antialiasing filter stage, and the second column shows the results with the antialiasing filter stage. The EZT coder [38] with fixed wavelet filter banks (called embedded zero-tree wavelet coder—EZW) of biorthogonal Barlaud filter [39], and nine-tap orthogonal Coiflet filter [15] produces PSNR's of 31.42 dB and 31.45 dB, respectively. The public domain EZW programs use mostly this Coiflet filter bank, and experimentally, it gives the best overall coding results among these types of coders. These PSNR's are 0.83 dB less than the PSNR obtained using the adaptive decomposition method. In addition to the improved PSNR, the adaptive filter bank eliminates the ringing effects in Fig. 12(b) which are apparent in the EZW coder as shown in Fig. 12(c). To give a better idea about the reduction of the ringing artifacts, Fig. 13(a) shows the enlarged detail of our decoder output, and Fig. 13(b) shows the EZW output of the same portion at 1 bpp.

(a)

(b)

Fig. 13. Detail parts from Fig. 12: (a) 2-D LMS adaptive at 1 bpp and (b) EZW at 1 bpp.

A set of 16 8-bpp gray-tone images is compressed using the adaptive subband coding scheme and the EZW with a fixed filter bank. In all cases, the adaptive method achieves a higher PSNR at 1 bpp. Similar to the EZT coder, the set partitioning in hierarchical trees (SPIHT) coder [40] gives better results with the adaptive filter bank. The coding results for these images are presented in Table II. The thumbnailed versions of the 8-bpp gray-tone test images are shown in Fig. 15.

In some of the test images, the adaptive filter bank still produces slightly lower PSNR's. This is due to the texture characteristics these particular images. When the image contains a texture which is difficult to follow with a gradient type adaptive algorithm, the error sequence stays large. In these situations, a filter bank optimized for the signal can be used. On the other hand, if the image contains a number of portions with different textures, then it is difficult to obtain a single filter bank suitable for all the portions of the image. In those situations, another adaptation which selects the appropriate filter bank in a specific region, should be applied.

The adaptive morphological subband decomposition described in [24] uses an alternating coding strategy for different regions in the image. Specifically, it chooses linear filters for textured regions and morphological filters otherwise. In the "baboon" image this algorithm achieves a PSNR of 25.858 dB at 0.49 bpp, whereas our adaptive algorithm gives a PSNR of

Image Description	EZW	1-D prediction		2-D prediction	
		Adaptive FIR	Adaptive OS	Adaptive FIR	Adaptive OS
Call_for_papers (390x190)	36.50	36.80	36.96	36.87	36.99
Sci_Tech1 (350x244)	36.19	36.30	36.41	36.33	36.48
Sci_Tech2 (350x286)	31.14	31.56	31.61	31.60	31.65
House (256x256)	38.80	38.97	39.10	39.08	39.22
Baboon (256x256)	30.50	30.46	30.56	30.50	30.61
Tourism1 (480x480)	30.20	30.14	30.20	30.15	30.22
Tourism2 (300x500)	27.50	27.88	28.02	27.92	28.07
Tourism3 (360x480)	32.25	32.20	32.27	32.22	32.31
TR_map (470x250)	31.50	31.66	31.86	31.70	31.92
News0 (350x240)	34.21	34.15	34.19	34.20	34.23
News1 (350x240)	32.25	32.19	32.22	32.22	32.25
News2 (160x240)	23.90	24.12	24.20	24.18	24.27
Map_Africa (260x320)	33.02	33.29	33.43	33.36	33.50
s_text1 (512x256)	34.10	36.10	36.28	36.30	36.33
Pepper (256x256)	38.15	38.18	38.17	38.50	38.44
Zelda (512x512)	39.87	39.79	39.71	39.91	39.85

26.91 dB at the same bit rate. Furthermore, the filter selection method used in [24] can also be used in our algorithm to switch between adaptive FIR and adaptive OS filter banks, as well. The adaptive nonlinear filter banks are successful in coding images with regions separated by sharp edges. On the other hand, the adaptive FIR filter bank is useful for regions with uniform textures like images of grass or forest. The texture detection algorithm described in [24] can be used for separating such regions and using alternating adaptive filter banks. The PR property is preserved since the texture detection is based on the low-low component of the decomposed signal which is available to both the encoder and the decoder side.

In a filter bank structure, the perfect reconstruction depends on the lossless transmission of the subsignals to the synthesis side. In adaptive filter banks, the high-band signal is also used for adapting the synthesis filter. Therefore at very low bit rates the performance of the adaptive filter bank coder deteriorates. Consider the PSNR versus CR plots shown in Fig. 14 for the image in top left corner of Fig. 15. Above CR = 20 level, fixed filter banks start producing better results than the adaptive filter bank. However, at these CR levels, the coding is not visually transparent, and disturbing coding artifacts become visible in the decoded images. For the low bit-rate coding, a modified adaptive scheme is proposed in Section VII. In the modified scheme, the quantization effects are incorporated in the adaptation algorithm, and the divergence cases for the synthesis stage are eliminated.

VII. LOW BIT-RATE IMAGE CODING

The adaptation block in the analysis stage produces an error signal which is used in the synthesis stage to reconstruct the
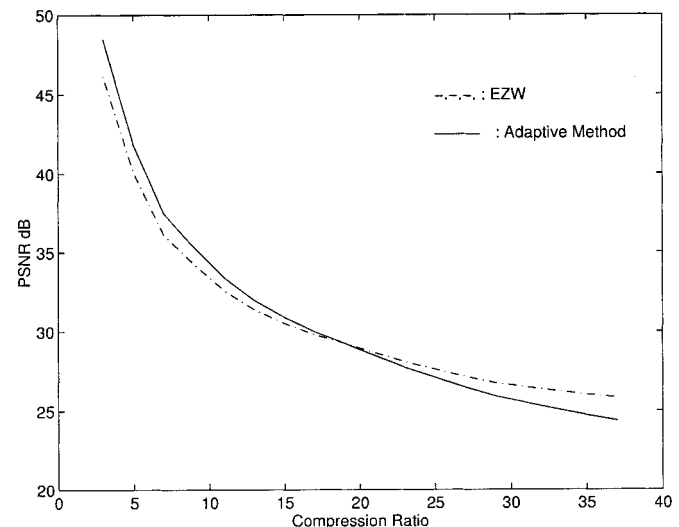

Fig. 14. EZW versus adaptive method at different CRs.

signal. During coding, both the input signal x_l and the error signal x_h are subject to quantization. The reconstruction algorithm in the synthesis stage tolerates the quantization effects up to the compression level of approximately 0.5 bpp. Below this bit rate, the change in the signal due to quantization causes the reconstruction algorithm to diverge. In other words, the analysis and synthesis adaptive blocks converge to completely different signals because the input signal to the synthesis block is different from the input signals to the analysis block.

In order to avoid the divergence situation, a method is proposed which incorporates the quantization effects before the adaptation block in the analysis stage. In Fig. 16, the upper
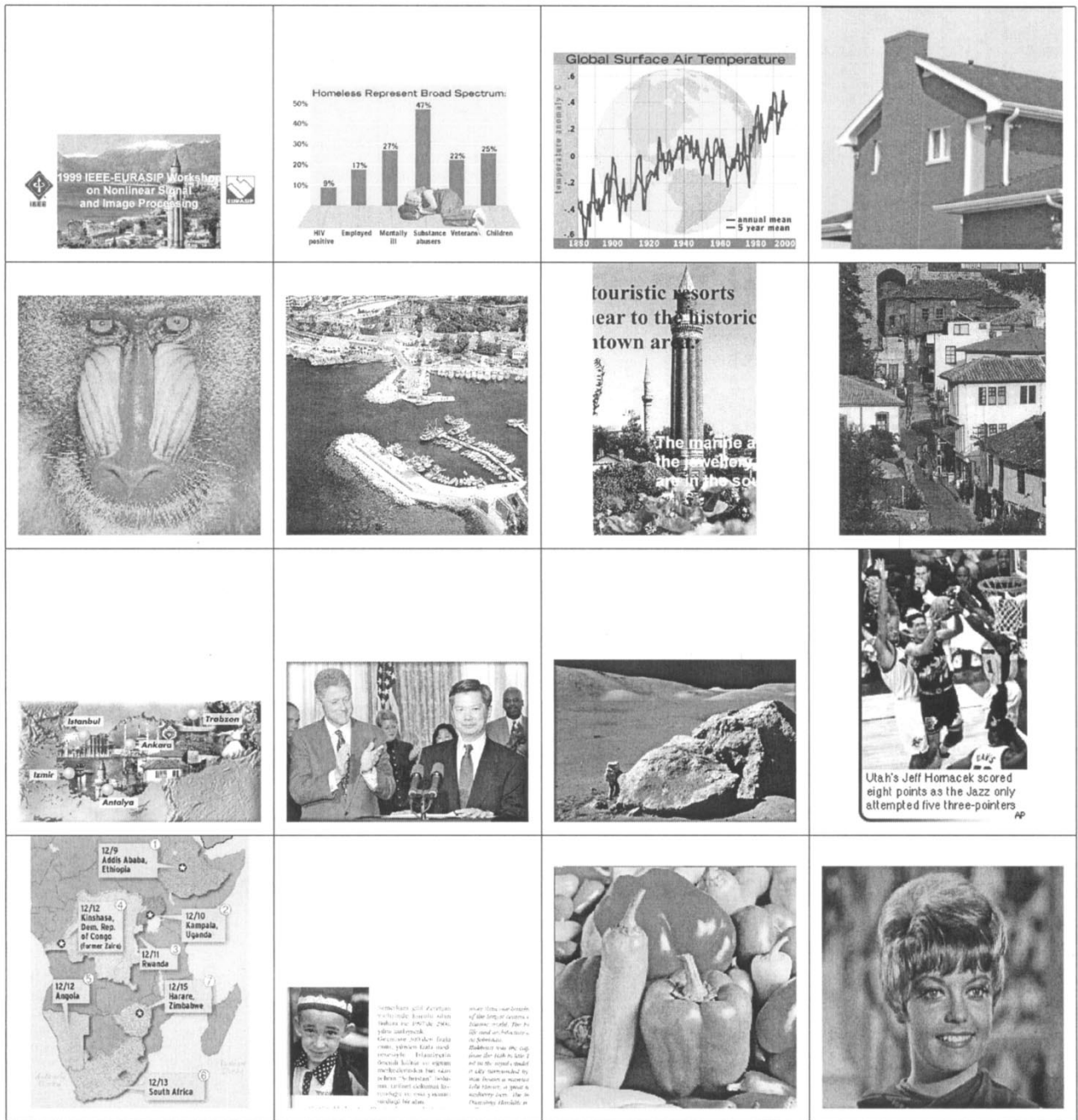
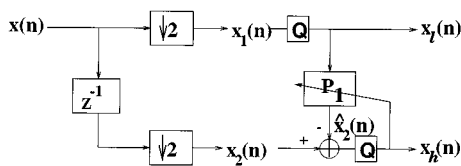Fig. 15.   Thumbnailed versions of 16 test images.

Fig. 16.   Analysis stage of the adaptive filter bank structure with quantization.

polyphase component is quantized before being fed to the adaptive block. Furthermore, the error signal used for determining the gradient vector in the LMS algorithm is also quantized. In this way, we obtain an adaptation scheme which uses quantized

coefficients which are exactly transmitted to the synthesis side without any further loss.

Apparently, the zerotree coder cannot be used in this structure because it quantizes the coefficients in an unpredictable manner depending on the chosen bit rate. In our simulations, we used a scalar quantizer in this structure to obtain bit rates lower than 1 bpp. In order to make a comparison, a fixed prediction filter and an LMS adaptive filter is used, and the outputs are encoded using an arithmetic coder. The fixed filters used in our experiments are the five-tap Lagrange low-pass filter, which gives reasonable prediction values, and a five-tap median filter. The coding results are presented in Table III. Since it is difficult to

TABLE III
COMPRESSION RATIOS (CR) OF THE TEST IMAGES AROUND 0.2 BPP. PSNR VALUES ARE GIVEN IN THE FIRST COLUMN

Image Description	Fixed filter	1-D prediction		2-D prediction	
		Adaptive FIR	Adaptive OS	Adaptive FIR	Adaptive OS
Call_for_papers (18.55 dB)	39.85	40.84	40.97	41.00	41.22
Sci_Tech1 (18.30 dB)	39.90	40.40	40.63	40.51	40.95
Sci_Tech2 (15.90 dB)	39.78	39.98	40.42	40.12	40.64
House (19.65 dB)	40.08	40.13	40.28	40.24	40.49
Baboon (15.50 dB)	39.09	39.48	39.60	39.56	39.86
Tourism1 (15.39 dB)	39.60	39.50	39.75	39.62	39.90
Tourism2 (14.05 dB)	39.02	39.45	39.82	39.63	39.98
Tourism3 (16.40 dB)	39.60	39.42	39.66	39.57	40.05
TR_map (16.00 dB)	39.50	39.72	39.98	39.72	40.24
News0 (17.40 dB)	40.10	39.81	40.05	39.94	40.26
News1 (16.40 dB)	40.20	39.83	40.18	39.99	40.33
News2 (12.33 dB)	39.10	39.90	40.05	40.05	40.18
Map_Africa (16.87)	39.20	39.41	40.02	39.60	40.23
s_text1 (17.25)	38.02	39.20	39.60	39.25	39.99
Pepper (19.41)	39.90	39.90	40.24	40.05	40.47
Zelda (20.22)	40.12	39.94	40.01	40.27	40.16

control the output rate in this system, the PSNRs are fixed, and the compression ratios are calculated.

VIII. CONCLUSIONS

Adaptive subband decomposition schemes for image coding are developed in this paper. The adaptive filters are embedded into a polyphase structure. In the two channel structure, the high-band subsignal is estimated from the low-band subsignal using an LMS type adaptation algorithm. The perfect reconstruction property is retained as long as the same adaptation algorithm is used at the analysis and synthesis stages.

We also introduced an adaptive filter bank with antialiasing filtering for the low-band signal. In this structure, the low-band signal is obtained from the input using half-band low-pass filters followed by downsampling. The high-band subsignal is then estimated from the low-band subsignal using adaptive prediction. In this structure, a high quality multiresolution viewing capability is possible due to the antialiasing filter.

The structure is also extended to 2-D image decomposition in various ways. It is shown that separable and quincunx downsampling extensions are possible.

In order to obtain low bit rates, a method that incorporates the quantization effects inside and before the adaptation is developed. Experimental results with this scenario indicate that the adaptive method gives better results than a fixed filter.

The overall structure is observed to be efficient for compressing images. The adaptive subband decomposition method produces visually better results by eliminating ringing artifacts. Specifically, for the images that contain sharp variations such as text, subtitles and graphics, our algorithm outperforms other methods in experimental studies.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
[2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
[3] G. R. Arce and M. Tian, "Order-statistic filter banks," *IEEE Trans. Image Processing*, vol. 5, June 1996.
[4] G. Strang, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1996.
[5] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
[6] S. Weiss, M. Harteneck, and R. W. Stewart, "On implementation and design of filter banks for subband adaptive systems," in *Proc. IEEE Workshop Signal Processing Systems*, Cambridge, MA, Oct. 1998, pp. 172–181.
[7] M. Harteneck, S. Weiss, and R. W. Stewart, "Design of near perfect reconstruction oversampled filter banks for subband adaptive filters," *IEEE Trans. Circuits Syst. II*, vol. 46, pp. 1081–1085, Aug. 1999.
[8] S. Hosur and A. H. Tewfik, "Wavelet transform domain adaptive FIR filtering," *IEEE Trans. Signal Processing*, vol. 45, pp. 617–630, Mar. 1997.
[9] N. Erdöl and F. Başbuğ, "Performance of wavelet transform based adaptive filters," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, vol. 3, pp. 500–503, Apr. 1993.
[10] S. Attallah and M. Najim, "On the convergence enhancement of the wavelet transform based LMS," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Detroit, MI, May 1995, pp. 973–976.
[11] N. Himayat and S. A. Kassam, "A structure for adaptive order statistic filtering," *IEEE Trans. Image Processing*, vol. 3, pp. 265–280, May 1994.
[12] Y. Higa, H. Ochi, and S. Kinjo, "A subband adaptive filter with a variable analysis filter bank," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Munich, Germany, Apr. 1997, pp. 2309–2312.
[13] P. Desarte, B. Acq, and D. T. M. Slock, "Signal-adapted multiresolution transform for image coding," *IEEE Trans. Inform. Theory*, vol. 38, pp. 897–903, Mar. 1992.
[14] A. H. Tewfik, D. Sinha, and P. Jorgensen, "On the optimal choice of a wavelet for signal representation," *IEEE Trans. Inform. Theory*, vol. 38, pp. 747–765, Mar. 1992.
[15] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 713–718, Mar. 1992.

[16] V. K. Goyal, J. Zhuang, M. Vetterli, and C. Chan, "Transform coding using adaptive bases and quantization," in *IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996.

[17] P. Moulin, K. Ramchandran, and V. Pavlovic, "Transform image coding based on joint adaptation of filter banks and tree structures," in *IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, Sept. 1996.

[18] P. Moulin and M. K. Mihcak, "Theory and design of signal-adapted FIR paraunitary filter banks," *IEEE Trans. Signal Processing*, vol. 46, pp. 920–929, Apr. 1998.

[19] A. Kirac and P. P. Vaidyanathan, "Theory and design of optimum FIR compaction filters," *IEEE Trans. Signal Processing*, vol. 46, pp. 903–919, Apr. 1998.

[20] W. Sweldens, "The lifting scheme: A new philosophy in biorthogonal wavelet constructions," in *Proc. SPIE Wavelet Applicat. Signal Image Processing III*, vol. 2569, A. F. Laine and M. Unser, Eds., 1995, pp. 68–79.

[21] R. Claypoole, G. Davis, W. Sweldens, and R. Baraniuk, "Nonlinear wavelet transform for image coding," in *Proc. 1997 Asilomar Conf. Signals, Systems, Computers*, 1997.

[22] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Applicat.*, vol. 4, no. 3, pp. 247–269, 1998.

[23] F. A. M. L. Bruekers and A. W. M. Van Den Enden, "New networks for perfect inversion and perfect reconstruction," *IEEE J. Select. Areas Commun.*, vol. 10, pp. 130–137, Jan. 1992.

[24] O. Egger, W. Li, and M. Kunt, "High compression image coding using an adaptive morphological subband decomposition," *Proc. IEEE*, vol. 83, pp. 272–287, Feb. 1995.

[25] F. J. Hampson and J. C. Pesquet, "A nonlinear subband decomposition with perfect reconstruction," in *IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, 1996.

[26] K. Anandakumar and S. A. Kassam, "Nonlinear multiresolution decomposition with applications in image restoration," *J. Electron. Imag.*, 1996.

[27] O. Arıkan, A. E. Çetin, and E. Erzin, "Adaptive filtering for non-Gaussian stable processes," *IEEE Signal Processing Lett.*, vol. 1, pp. 163–165, Nov. 1994.

[28] G. Aydın, O. Arıkan, and A. E. Çetin, "Robust adaptive filtering algorithms for alpha-stable random processes," *IEEE Trans Circuits Syst. II*, vol. 46, pp. 199–203, Feb. 1999.

[29] R. H. Kwong and E. W. Johnston, "A variable step size LMS algorithm," *IEEE Trans. Signal Processing*, vol. 40, pp. 1633–1642, July 1992.

[30] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithm: Analysis and simulations," *IEEE Trans. Signal Processing*, vol. 45, pp. 631–639, Mar. 1997.

[31] N. J. Bershad, "On weight update saturation nonlinearities in LMS adaptation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 623–630, Apr. 1990.

[32] P. Salembier, "Adaptive rank order based filters," *EURASIP Signal Process.*, vol. 27, no. 1, pp. 1–25, 1992.

[33] C. W. Kim, R. Ansari, and A. E. Çetin, "A class of linear-phase regular biorthogonal wavelets," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, San Francisco, CA, Mar. 1992.

[34] S.-M. Phoong, C. W. Kim, P. P. Vaidyanathan, and R. Ansari, "A new class of two channel biorthogonal filter banks and wavelet bases," *IEEE Trans. Signal Processing*, vol. 43, pp. 649–665, Mar. 1995.

[35] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[36] A. Karaş and B. S. Yarman, "A new approach for constructing base functions via Gram Schmidt orthogonalization method," in *Workshop Signal Processing*, 1994.

[37] Ö N. Gerek and A. E. Çetin, "Polyphase adaptive filter banks for fingerprint image compression," *Electron. Lett.*, vol. 34, pp. 1931–1932, Oct. 1, 1998.

[38] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.

[39] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transforms," *IEEE Trans. Image Processing*, vol. 1, pp. 205–220, Apr. 1992.

[40] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy image compression," *IEEE Trans. Image Processing*, vol. 5, pp. 1303–1310, Sept. 1996.

Ömer Nezih Gerek (S'89–M'98) was born in Eskisehir, Turkey, in 1969. He received the B.S., M.S., and Ph.D. degrees from Bilkent University, Ankara, Turkey, all in electrical engineering.

He held the Bilkent scholarship during his B.S. program. During his M.S. and Ph.D. studies, he was a Research Assistant with Bilkent University. From 1998 to 1999, he was a Technical Researcher with the Swiss Federal Institute of Technology, Lausanne, Switzerland. He is currently an Assitant Professor with the Department of Electrical Engineering, Eskisehir Anadolu University, Eskisehir, Turkey. His research interests include image processing, audio–visual signal coding and multiresolution analysis.

A. Enis Çetin (S'84–M'87) received the B.Sc, M.S.E., and Ph.D. degrees in systems engineering from the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia.

From 1987 to 1989, he was Assistant Professor of electrical engineering with the University of Toronto, Toronto, ON, Canada. Since then, he has been with Bilkent University, Ankara, Turkey, where he is currently a Full Professor. During the Summers of 1988, 1991, and 1992, he was with Bell Communications Research (Bellcore), Piscataway, NJ. During the 1994–1995 academic year, he was a Visiting Professor with the University of Minnesota, Minneapolis.