POLYPHASE ADAPTIVE FILTER BANKS FOR SUBBAND DECOMPOSITION

Ömer N. Gerek, A. Enis Çetin

Bilkent University, Dept. of Electrical and Electronics Engineering,
Bilkent, Ankara TR-06533, Turkey
E-mail: gerek@ee.bilkent.edu.tr
Phone: (90) 312-266 4307 Fax: (90) 312-266 4126

ABSTRACT

Subband decomposition is widely used in signal processing applications including image and speech compression. In most practical cases, the goal is to obtain subband signals that are suitable for data compression. In this paper, we present Perfect Reconstruction (PR) polyphase filter bank structures in which the filters adapt to the changing input conditions. This leads to higher compression results for images containing sharp edges, text, and subtitles.

1. ADAPTIVE PREDICTION FILTERS IN POLYPHASE FORM

Consider the simple two band PR polyphase decomposition structure shown in Figure 1 in which the filter P_1 can be
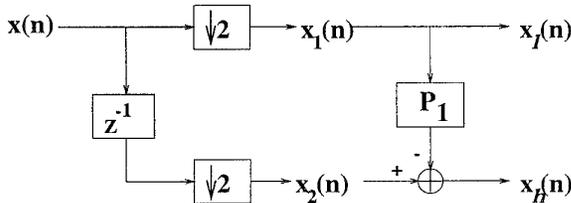
Figure 1: *Polyphase analysis and synthesis structure*

either linear or nonlinear [1],[2]. This structure has a simple transform matrix:

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

A good P_1 filter would be the one that can predict the samples of $x_2(n)$ as close as possible. In binary image coding, successful coding results are obtained by selecting P_1 as a median filter [3]. In [4], the P_1 is chosen as a FIR low-pass and a median filter pair which are interchanged according to characteristics of the input signal. In this paper, the adaptive FIR filters and adaptive order statistics filters are used in the filter bank and image coding algorithms are developed based on the resulting structure. In all cases, the inverse of the $\mathbf{P}$ matrix in Eq. (1) is given as:

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

A predictor should be adaptive for image and video signals as they are unstationary in nature. This reasoning
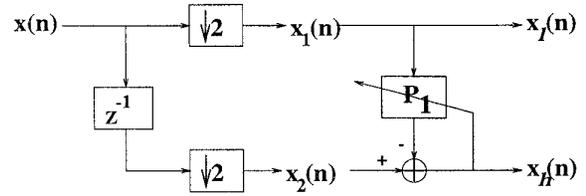
Figure 2: *Adaptive structure analysis stage*
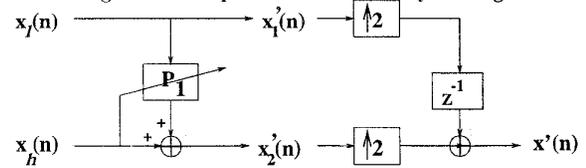
Figure 3: *Adaptive structure synthesis stage*

leads to the polyphase structure shown in Fig. 2 in which the prediction filter adapts itself to minimize the highband signal $x_h(n)$. In this way, unnecessary information in $x_2(n)$ is removed and an uncorrelated signal is transmitted to the receiver. This is especially useful when there are sharp transition regions in an image such as subtitles, text and graphics.

The polyphase adaptive filter bank structure is shown in Figures 2, 3. In this figure, the signal $x_2(n)$ is predicted from $x_1(n)$ (or from the even samples of $x(n)$) as follows:

$$\hat{x}_2(n) = \sum_{k=-N}^{N} a_k x_1(n-k) = \sum_{k=-N}^{M} a_k x(2n-2k) \qquad (3)$$

where the filter coefficients a_k's are updated using an LMS-type algorithm [5],[6], and the subsignal x_h is given by

$$x_h(n) = x_2(n) - \mathbf{P_1}(x_1) = x_2(n) - \hat{x}_2(n). \qquad (4)$$

Since $x_h(n)$ corresponds to the error signal in the adaptation algorithm, the decoder can also update the filter coefficients without any side information. It can easily be shown that the PR property is preserved as long as the same adaptation strategy is used in both encoding and decoding.

Another choice for P_1 is the adaptive Order Statistics (OS) filter [7]-[10]. Similar to the adaptive FIR case, the OS filters can be adapted by minimizing the subsignal x_h using an adaptation strategy.

2. MULTICHANNEL EXTENSIONS OF THE BASIC ADAPTIVE FILTER BANK

The structure in Fig. 1 and 2 can be extended to contain cascaded PR blocks as follows:

$$\mathbf{P} = \begin{bmatrix} 1 & -P_1(.) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ G_1(.) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & -P_2(.) \\ 0 & 1 \end{bmatrix} \times \cdots \quad (5)$$

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

$$\mathbf{P}^{-1} = \cdots \times \begin{bmatrix} 1 & P_2(.) \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} 1 & 0 \\ -G_1(.) & 1 \end{bmatrix} \times \begin{bmatrix} 1 & P_1(.) \\ 0 & 1 \end{bmatrix} \quad (6)$$

Multichannel and multidimensional extensions are also possible in a straightforward manner. An example of the PR multichannel extension is illustrated in Fig. 4. In this
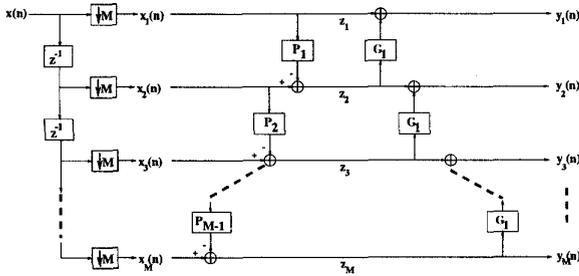
Figure 4: *Multi-band analysis structure*

figure, 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,

$$\begin{aligned}
z_1 &= x_1 \\
z_i &= x_i - P_{i-1}(z_{i-1}), \quad i = 2, 3, ..., M \\
y_i &= z_i + G_i(z_{i+1}), \quad i = 1, 2, ..., M - 1 \\
y_M &= z_M
\end{aligned} \quad (7)$$

The corresponding $\mathbf{P}$ matrix for this example can be given as :

$$\begin{bmatrix} 1 & -P_1 & 0 & 0 & \cdots \\ 0 & 1 & -P_2 & 0 & \cdots \\ 0 & 0 & 1 & -P_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ G_1 & 1 & 0 & 0 & \cdots \\ 0 & G_2 & 1 & 0 & \cdots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{bmatrix} \quad (8)$$

The PR can be achieved with any choice of the nonlinear operators as the $\mathbf{P}$ matrix can be inverted regardless of the filters P_i's and G_i's. This leads to the following synthesis equations:

$$\begin{aligned}
z'_M &= z_M \\
z'_i &= y_i - G_i(z'_{i+1}) &= z_i, \quad i = M - 1, ..., 1 \\
x'_1 &= z_1 &= x_1 \\
x'_i &= z'_i + P_{i-1}(z'_{i-1}) &= z_i + P_{i-1}(z_{i-1}) = x_i
\end{aligned} \quad (9)$$

The outputs, x'_i, of the synthesis filters are the same as the polyphase components, x_i, of the analysis filter bank. Notice that there are no restrictions on the filters P_i and G_i for perfect reconstruction.
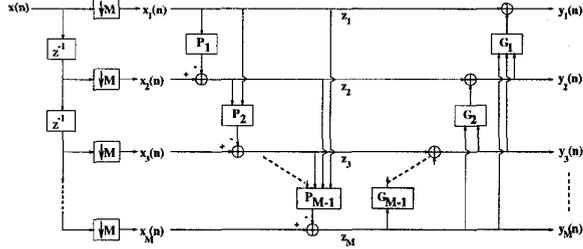
Figure 5: *Multi-band analysis structure - 2*

Another multichannel extension structure is illustrated in Fig. 5. The previous structure was considering the nearest polyphase component value of the input data stream. However, the structure in Fig. 5 is using increasing amount of polyphase component values for prediction as you go from up to down in the figure. The analysis and synthesis equations for this structure can be given as follows.

Analysis:

$$\begin{aligned}
z_1 &= x_1 \\
z_i &= x_i &- P_{i-1,i-1}\big(-P_{i-2,i-2}(-P_{i-3,i-3}(\cdots(-P_{1,1})))\big) \\
&&- P_{i-1,i-2}\big(-P_{i-3,i-3}(\cdots(-P_{1,1}))\big) \\
&& \vdots \quad \vdots \\
&&- P_{i-1,1}, \quad i = 2, 3, ..., M \\
&= x_i &- P_{i-1}(z_1, z_2, ..., z_{i-1}) \\
y_M &= z_M \\
y_i &= z_i &+ G_i(y_M, ..., y_{i+1}), \quad i = 1, 2, ..., M - 1
\end{aligned} \quad (10)$$

Synthesis:

$$\begin{aligned}
z'_i &= y_i - G_i(y_M, ..., y_{i+1}) &= z_i, \quad i = 1, 2, ..., M - 1 \\
z'_M &= y_M &= z_M \\
x'_1 &= z_1 &= x_1 \\
x'_i &= z_i + P_{i-1}(z_1, z_2, ..., z_{i-1}) &= x_i, \quad i = 2, 3, ..., M
\end{aligned} \quad (11)$$

This later structure also yields analysis matrices which can be composed to upper and lower triangular matrices with elements containing P_i's and G_i's only. In this structure, for predicting z_i's, the number of data used is increased with increasing index i. Conversely, more z_i samples are used for predicting y_i's when the index i is small. The computational and analysis complexity of this structure is high as compared to the structure in Fig. 4. Furthermore, no coding gain over the structure in Fig. 5 has been observed for the structure in Fig. 4.

3. CODING GAIN ANALYSIS

The structure in Fig. 2 can be considered as a transformation on the input signals x_1 and x_2. For an appropriate transform, the goal is to minimize the total variance of the transform coefficients. It has been shown in [12] that if the

coefficients before transformation are quantized at an average bit rate R, the PCM quantization variance can be stated as

$$\sigma_{r,PCM}^2 = \epsilon^2 2^{-2R} \sigma_x^2 \tag{12}$$

where ϵ^2 is a scaling factor proportional to the quantizer performance factor [12], and σ_x^2 is the variance of the input signal. Similarly, if the transform coefficients are also quantized at the same rate, R, then the minimum transform domain quantization variance becomes

$$\sigma_{r,TC}^2 = \epsilon^2 2^{-2R}(\sigma_1 \sigma_2) \tag{13}$$

The ratio of $\sigma_{r,PCM}^2$ to $\sigma_{r,TC}^2$ gives the coding gain:

$$G_{TC} = \frac{\sigma_x^2}{(\sigma_1^2 \sigma_2^2)^{(1/2)}} \tag{14}$$

The maximization of the gain term is equivalent to minimization of the variances of the transform domain coefficients. The transform matrix in Eq. (1) shows that transform domain output x_1 corresponding to the input x_1 is identically equal to the input signal. In this case, we can maximize the coding gain only by minimizing the variance of x_h.

The adaptive LMS algorithm minimizes the mean squared error at the output, hence it minimizes the variance σ_2^2. This minimization, therefore, maximizes the coding gain at each sample point. Since our OS adaptive algorithm also uses LMS type adaptation, the same variance minimization arguments hold.

In our simulations, the normalized LMS algorithm is used with normalization factor $\|x_1\|^2$. With different choices of the norms (L^1, L^2, and L^∞), different convergence performances are observed [13]. The L^1 and L^2 norms proved to perform good for most images.

4. SIMULATION STUDIES

The two dimensional extension of the proposed scheme is obtained by separately applying the one dimensional adaptive filter bank to the image data in a separable manner. The filter support, however, is chosen to exploit two dimensional characteristics better. The filter support is illustrated in Fig. 6.
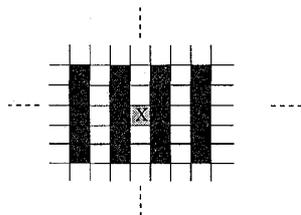
Figure 6: *Two dimensional separable prediction.*

In all of the simulations, we used the Embedded ZeroTree (EZT) coder to encode the transform coefficients [14]. Due to the characteristics of this coder, the best coding results were obtained by tree-structured two-band decompositions.

At compression ratios from 1:1 to 30:1, the adaptive algorithm produces acceptable outputs at the synthesis stage. Especially for the images with sharp variations, this algorithm gives better PSNRs at CR from 1: to 15:1. Due to corrupting the synthesis filter adaptation at lower bit rates, the reconstructed image starts getting deteriorated more than the conventional wavelet quantization after compression ratios of 15:1. The synthesis stage starts to diverge after the CRs around 30:1. The CR/PSNR plot for the adaptive algorithm applied to the first image in Table 1 is shown in Fig 7. A quick remedy to avoid the divergence of
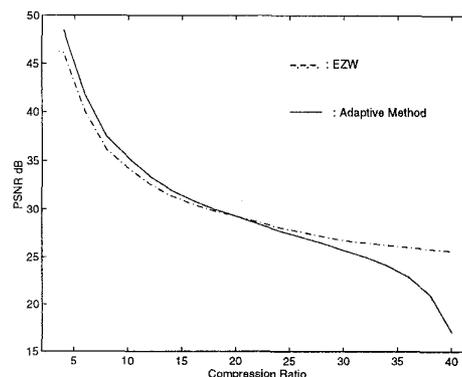
Figure 7: *EZW versus adaptive method at different CR's*

the the synthesis stage at very low bit rates is to substitute a fixed filter in the place of the adaptive filter whenever the bit rate is too low.

The 672×560 "barbara" image was compressed to 1 bits/pixel at a PSNR of 35.88dB with the adaptive OS type prediction filter. This PSNR is also comparable to the conventional EZW compression scheme which produces 35.90dB PSNR. Furthermore, our algorithm produces perceptually better images with eliminated ringing effects at the edges.

A set of various images have been compressed using the adaptive algorithm. Some of these images are presented in Table 1.

The coding results of these images in Table 1 at 1bpp are presented in table 2.

On the average, our algorithm outperforms the EZW for images with sharp variations. More robust adaptation algorithms for lower bit rates will be studied as a future work.
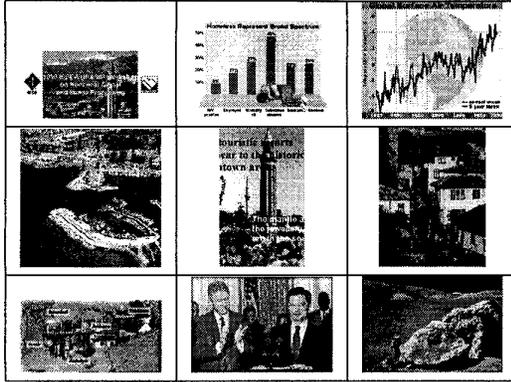
Table 1: *Test images.*

Image Description	EZW	Adaptive FIR	Adaptive OS
Call_for_papers(*)	36.96	36.87	36.99
Sci_Tech1(*)	36.41	36.33	36.48
Sci_Tech2(*)	31.61	31.60	31.65
Tourism1	30.18	30.12	30.19
Tourism2(*)	28.02	27.92	28.07
Tourism3	32.23	32.18	32.31
TR_map(*)	31.86	31.70	31.92
News0	34.14	34.15	34.16
News1	32.19	32.05	32.23

Table 2: *Experiment results (PSNR) of test images at 1bpp.*

5. REFERENCES

[1] F. J. Hampson and J. C. Pesquet, "A nonlinear subband decomposition with perfect reconstruction," *IEEE Int. Symp. on Image Proc.* 1996.

[2] Ömer N. Gerek, Metin Nafi Gürcan, A. Enis Çetin, "Binary Morphological Subband Decomposition For Image Coding," *IEEE Int. Symp. on Time-Frequency and Time Scale Analysis,* 1996.

[3] Ömer N. Gerek, A. Enis Çetin, "Binary nonlinear subband decomposition structures for textual image compression," *Int. Symp. on Nonlinear Signal and Image Processing, NSIP'97,* Sep. 1997.

[4] O. Egger, W. Li, and M. Kunt, "High Compression Image Coding Using an Adaptive Morphological Subband Decomposition," *Proc IEEE,* vol. 83, no. 2, pp.272-287, Feb. 1995.

[5] O. Arıkan, A. E. Çetin, Engin Erzin, 'Adaptive Filtering for non-Gaussian stable processes,' *IEEE Signal Processing Letters,* vol. 1, No. 11, pp. 163-165, November 1994.

[6] Gül Aydın, O. Tanrıkulu, A. Enis Çetin, "Robust least mean mixed norm adaptive filtering algorithms for α-stable random processes," *IEEE- ISCAS'97,* Hong Kong, June 1997.

[7] Gonzalo. R. Arce and M. Tian, "Order-statistic filter banks," *IEEE Transactions on Image Processing,* 5, June 1996.

[8] I. Pitas and A. Venetsanopoulos, "Adaptive filters based on order statistics," *IEEE Trans. Signal Processing,* vol. 39, Feb. 1991.

[9] P. Salembier, "Adaptive rank order based filters," *EURASIP Signal Processing,* 27(1):1-25, April 1992.

[10] P. Salembier and L. Jaquenoud, "Adaptive morphological multiresolution decomposition," *In Dougherty Gader, editor, Image Algebra and Mathematical Morphology,* volume 1568, pages 26-37, San Diego, USA, July 1991.

[11] S-M. Phoong, C. W. Kim, P.P Vaidyanathan, R. Ansari, "A new class of two channel biorthogonal filter banks and wavelet bases," *IEEE Trans. Signal Proc.,* Vol.43, No.3, pp. 649-665, March 1995.

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

[13] S. C. Douglas, "A family of normalized LMS algorithms," *Signal Processing LEtters,* vol. 1, no.3, pp. 49-51, Mar. 1994.

[14] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Trans. on Signal Processing,* vol. 41, no. 12, pp. 3445 - 3462, Dec. 1993.