

MINLAB: Minimum Noise Structure for Ladder-Based Biorthogonal Filter Banks

See-May Phoong, *Member, IEEE*, and Yuan-Pei Lin, *Member, IEEE*

Abstract—In this paper, we introduce a minimum noise structure for ladder-based biorthogonal (MINLAB) filter bank. The minimum noise structure ensures that the quantization error has a unity noise gain, even though the filter bank is biorthogonal. The coder has a very low design and implementation cost. Perfect reconstruction property is structurally preserved. Optimal bit allocation and coding gain formulas are derived. We show that the coding gain of the optimal MINLAB coder is always greater than or equal to unity. For both AR(1) process and MA(1) process, the MINLAB coder with two taps has a higher coding gain than the optimal orthonormal coder with an infinite number of taps. In addition to its superior decorrelation ability, it has many other desired features that make it a potentially valuable and attractive alternative to the orthonormal coder, especially for the high-fidelity compression.

Index Terms—Biorthogonal, compression, filter bank, minimum noise, subband coding, wavelet coding.

I. INTRODUCTION

RECENTLY, there has been considerably interest in applying the ladder structure to data compression [1]–[17]. Fig. 1 shows a simple two-channel filter bank (FB) that uses only one ladder. Such a structure is also known as the lifting scheme in wavelet coding [6], [7], [16]. In the absence of the quantizers, such a biorthogonal system always has the perfect reconstruction (PR) [18], that is, $y(n) = x(n)$ for all possible $x(n)$, regardless of the choice of $P(z)$. In other words, the FB is structurally PR. The analysis filters $H_i(z)$ and synthesis filters $F_i(z)$ are, respectively

$$\begin{aligned} H_0(z) &= 1 - z^{-1}P(z^2), & H_1(z) &= z^{-1} \\ F_0(z) &= 1, & F_1(z) &= z + P(z^2). \end{aligned} \quad (1)$$

The implementation and design of the biorthogonal system involves only $P(z)$, and hence, the design and computational cost is very low. Even though the filters in (1) are simple, their coding performance is comparable with that of orthonormal coders. Note that such a FB can never be orthonormal unless $P(z)$ is zero.

The ladder structure is first applied to progressive coding in [1]. It is shown in [2] how roundoff noise of the ladder structure at the encoder is cancelled by that at the decoder. A special

useful class of ladder structure is used as a framework for the construction of causal stable PR IIR FB [3], [4]. The same framework can be generalized to the quincunx (two-dimensional) 2-D FB by using a simple one-dimensional (1-D)-to-2-D mapping [4], [5]. In [6] and [7], the ladder structure or so-called lifting scheme is employed to construct biorthogonal wavelets and wavelets defined on irregular sampling grid. It was shown in [8] that both orthogonal and biorthogonal FB can be factorized into lifting steps. The nonlinear operator is used in the ladder structure to generate nonlinear FB with PR property [9], [10]. A low-cost and useful nonlinear operation is introduced in [10], and it is shown that the nonlinear FB coders have the ability to preserve edges and remove the blocking and ringing effects of compressed images. The ladder structure has also been applied to lossless and lossy coding of images, and satisfactory coding results can be obtained. In [11] and [12], the authors apply the ladder structure for the high-quality lossy compression and lossless coding of medical images. The proposed hierarchical interpolation (HINT) compression enables progressive resolution transmission. In [13] and [15], the authors introduce a new transform called the S+P transform. The S+P transform is a combination of the Haar transform and the predictive transform (which is in the form of a ladder). The compression algorithm proposed in [15] can support both progressive fidelity and progressive resolution transmission. It was demonstrated [15] that in the application of both lossy and lossless image coding, the S+P transform produces excellent compression results. In [14], the optimal predictor with certain zero constraint is used as $P(z)$, and the filter is obtained through the optimization of Bernstein polynomial. In [16], the authors proposed a ladder structure with integer-to-integer transform for the lossless coding of images. The 2-D four-channel ladder structure was studied in [17]. However, none of the ladder-based coders considered above have the unity noise gain property. Therefore, in the case of lossy compression, like most biorthogonal coder, the coding gain \mathcal{CG} of the ladder structure FB is not guaranteed to be greater than unity.

On the other hand, the class of orthonormal FB is known to have coding gain $\mathcal{CG} \geq 1$ [18]. There has been a lot of interest in finding the optimal orthonormal FB that yields a maximum coding gain for a given input statistics [21]–[24]. The theory of optimal orthonormal coder is closely related to the principle component FB [21], [22]. The principle component FB problem is solved in [21], whereas the solution of the optimal orthonormal coder is given in [22]. It is shown in [22] that the analysis and synthesis filters of an optimal orthonormal FB are the ideal filters that satisfy the majorization and decorrelation properties. The optimal FIR case is solved in [23] and [24].

Manuscript received September 16, 1998; revised May 30, 1999. This work was supported by the NSC under Grants 88-2213-E-002-080 and NSC 88-2218-E-009-016, Taiwan, R.O.C. The associate editor coordinating the review of this paper and approving it for publication was Dr. Xiang-Gen Xia.

S.-M. Phoong is with the Department of Electrical Engineering, Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan, R.O.C. (e-mail: smp@cc.ee.ntu.edu.tw).

Y.-P. Lin is with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

Publisher Item Identifier S 1053-587X(00)01009-6.

In this paper, we introduce a minimum noise structure for the ladder-based biorthogonal FB shown in Fig. 1. We will call such a coder the minimum noise ladder-based biorthogonal (MINLAB) coder. The MINLAB coder has the unity noise gain property. Optimal bit allocation and coding formulas are derived. The coding gain of the optimal MINLAB coder is equal to the square root of the prediction gain, and hence, it is guaranteed to be greater than or equal to unity. The optimal biorthogonal coder can be solved using Levinson recursion. For both autoregressive (AR) process and moving average (MA) process of order one, the proposed biorthogonal coder with two taps has a higher coding gain than *any* optimal orthonormal FB (with any number of taps). Preliminary results of this work have been presented in [19] and [20].

Outline of the Paper: Our presentation will go as follows: In Section II, we briefly discuss the coding performance of the traditional ladder-based coder. The MINLAB coder is introduced in Section III. The optimal bit allocation and coding gain formulas will be derived. We will show that the optimal MINLAB coder can be obtained from linear prediction theory. In Section IV, we will derive the MMSE predictor for the minimum noise structure. The merits of the MINLAB coders will be discussed in Section V. In Section VI, we provide examples such as AR(1) and MA(1) inputs to demonstrate the performance of the MINLAB coder. In Section VII, we will derive the results for tree structure MINLAB coder. The case of biorthogonal coder using more than one ladder will be studied in Section VIII.

Notations and Signal Model: Boldfaced characters represent vectors and matrices. The symbols \mathbf{I}_N and \mathbf{J}_N denote, respectively, the identity matrix and the reversal matrix of dimension N . For example

$$\mathbf{J}_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

In this paper, we assume that the input signal $x(n)$ is a real-valued zero-mean wide sense stationary (WSS) process with autocorrelation coefficients $r(k)$. Therefore, its variance is the same as $r(0)$. The statistical expectation of a random process $x(n)$ is expressed as $E\{x(n)\}$.

II. THE TRADITIONAL LADDER-BASED SUBBAND CODER

In a traditional subband coder, quantizers Q_i are placed directly after the subband signals $x_i(n)$, as shown in Fig. 1. In this paper, we make some commonly used assumptions on the quantizers.

Noise Model: Assume that the quantizers are scalar uniform quantizers and can be modeled as an additive noise source. Therefore, $\hat{x}(n) = x(n) + q(n)$ (as indicated by the dashed line in Fig. 1), where $x(n)$ and $\hat{x}(n)$ are, respectively, the input and output of the quantizer. We assume that for a b -bit quantizer, the variance of quantization noise $q(n)$ satisfies

$$\sigma_q^2 = c2^{-2b}\sigma_x^2 \quad (2)$$

where σ_x^2 is the variance of the input $x(n)$, and c is some constant depending only on the statistics of $x(n)$.

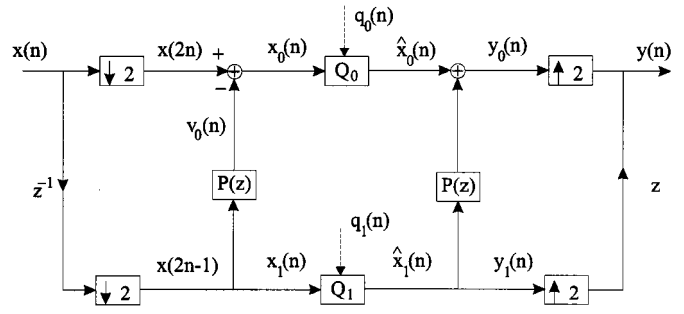


Fig. 1. Conventional subband coder using ladder.

Bit Rate: Let b_0 and b_1 be the number of bits assigned to Q_0 and Q_1 , respectively. The average bit rate in this case is

$$b = \frac{1}{2}(b_0 + b_1). \quad (3)$$

Average Output Noise Variance: Consider the synthesis bank of Fig. 1. The output noise is defined as $q_{\text{out}}(n) = y(n) - x(n)$. It contains contribution from both $q_0(n)$ and $q_1(n)$. Due to the upsampler, the output noise is not a WSS process. To quantify the error, we use the average noise variance

$$\sigma_{q_{\text{out}}}^2 = \frac{1}{2}(E\{(y_0(n) - x(2n))^2\} + E\{(y_1(n) - x(2n-1))^2\}).$$

It is clear that $E\{(y_1(n) - x(2n-1))^2\} = \sigma_{q_1}^2$. Assume that $q_1(n)$ is white and uncorrelated with $q_0(n)$. Then, we have $E\{(y_0(n) - x(2n))^2\} = \sigma_{q_0}^2 + \sigma_{q_1}^2 E_p$, where $E_p = \int_0^{2\pi} |P(e^{j\omega})|^2 d\omega/2\pi$ is the energy of the filter $P(z)$. Substituting these results into the above equation, we get

$$\sigma_{q_{\text{out}}}^2 = \frac{1}{2}\sigma_{q_0}^2 + \frac{1}{2}\sigma_{q_1}^2(1 + E_p).$$

The noise gain for $q_0(n)$ is unity, whereas $q_1(n)$ is amplified by $1 + E_p$. Due to this noise amplification, it is not guaranteed that the coding gain $\mathcal{CG} \geq 1$. In the next section, we will show how to eliminate the noise amplification by judiciously placing the quantizer Q_1 .

III. MINIMUM NOISE STRUCTURE FOR LADDER-BASED BIORTHOGONAL (MINLAB) CODERS

Consider Fig. 1. Note that the input to $P(z)$ at the analysis end is $x(2n-1)$, whereas the input to $P(z)$ at the synthesis end is its quantized version $\hat{x}(2n-1)$. This means that in the reconstruction process, $q_1(n)$ is added to the top branch through the filter $P(z)$. To avoid this, we can move the quantizer Q_1 to the left, as shown in Fig. 2. This has the dramatic effect of making the noise gain unity. We will refer to Fig. 2 as the MINLAB coder. To explain the unity noise gain property of this structure, note that from Fig. 2, we have

$$\begin{aligned} x_0(n) &= x(2n) - v_0(n) \\ \hat{x}_0(n) &= x_0(n) + q_0(n) \\ y_0(n) &= \hat{x}_0(n) + v_0(n). \end{aligned}$$

From the above equations and Fig. 2, we can conclude that the errors on the top and bottom branches are, respectively

$$\begin{aligned} y_0(n) - x(2n) &= q_0(n) \\ y_1(n) - x(2n-1) &= q_1(n). \end{aligned}$$

Therefore, the average variance of output error in the MINLAB coder is given by

$$\sigma_{q_{\text{out}}}^2 = 0.5 (\sigma_{q_0}^2 + \sigma_{q_1}^2).$$

The above equation is valid for any additive noise source. We *do not make any assumptions* on $q_0(n)$ and $q_1(n)$. That means, the noise gain is *always* one, even though the FB is never orthonormal. If the quantizers used are scalar quantizers that satisfy (2), then the above expression for $\sigma_{q_{\text{out}}}^2$ can be rewritten as

$$\sigma_{q_{\text{out}}}^2 = 0.5c (2^{-2b_0} \sigma_x^2 + 2^{-2b_1} \sigma_x^2) \quad (4)$$

where we have used the fact that $\sigma_{x_1}^2 = \sigma_x^2$. Applying the arithmetic mean (AM) geometric mean (GM) inequality to the above equation and using (3), we get

$$\sigma_{q_{\text{out}}}^2 \geq c2^{-2b} [\sigma_{x_0}^2 \sigma_x^2]^{1/2}$$

with equality if and only if the bits are allocated as

$$b_i = b + \frac{1}{2} \log \sigma_{x_i}^2 - \frac{1}{2} \log [\sigma_{x_0}^2 \sigma_x^2]^{1/2}. \quad (5)$$

Equal Step Size Rule: From the above derivation, we see that the average output noise variance $\sigma_{q_{\text{out}}}^2$ is minimized when the two quantizers have the same noise variance. The noise variances $\sigma_{q_i}^2$ and the quantization stepsize Δ_i are related as $\sigma_{q_i}^2 = \text{const} * \Delta_i^2$. Therefore, we conclude that the MINLAB coder continues to be optimal if the stepsizes of the quantizers are equal.

Coding Gain: If we define the coding gain of the coder as the ratio of the error variance in direct quantization [as in (2)] over that of the coder $\sigma_{q_{\text{out}}}^2$, then under the optimal bit allocation (5), the coding gain can be written as

$$\mathcal{CG} = \frac{\sigma_x^2}{[\sigma_{x_0}^2 \sigma_x^2]^{1/2}} = \sqrt{\frac{\sigma_x^2}{\sigma_{x_0}^2}}. \quad (6)$$

A. Optimal Biorthogonal Coders

In this subsection, we will find $P(z)$ such that the coding gain in (6) is maximized. We will first consider the FIR case and then the IIR case.

1) **Optimal FIR MINLAB Coder:** From (6), the coding gain \mathcal{CG} is maximized if $\sigma_{x_0}^2$ is minimized. The optimal solution of $P(z)$, such that $\sigma_{x_0}^2$ is minimized, can be obtained from linear prediction theory [25]. To see this, let $P(z)$ be an FIR filter of the form

$$P(z) = \sum_{n=-N}^{N-1} p(n)z^{-n}. \quad (7)$$

Then, the optimal solution is precisely the optimal predictor of $x(2n)$ based on the observations of $x(2n-2N+1), x(2n-2N+3), \dots, x(2n+2N-1)$. The noncausal predictor can

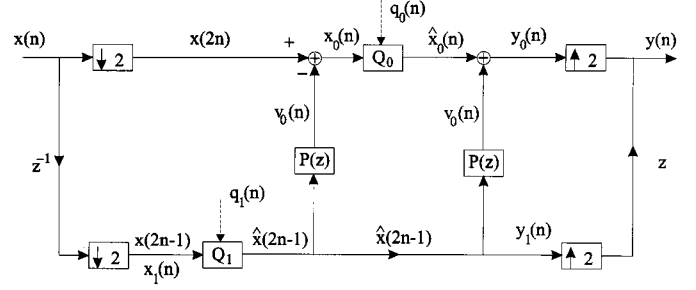


Fig. 2. MINLAB encoder and decoder.

be used here since we are predicting the even samples from the odd samples. A causal implementation of such a system is always possible by inserting enough delays at appropriate places in Fig. 2. Let $x(n)$ be a real-valued WSS process with autocorrelation coefficients $r(k)$. Then, using the orthogonality principle [18], [25], the optimal $p(n)$ that minimizes $\sigma_{x_0}^2$ is the solution of

$$\mathbf{R}_x \mathbf{p} = \mathbf{r} \quad (8)$$

where the matrix \mathbf{R}_x and the vectors \mathbf{p}, \mathbf{r} are as shown by \mathbf{R}_x at the top of the next page. Note that the matrix \mathbf{R}_x above is the autocorrelation matrix of the signal $x(2n-1)$, and hence, it is positive definite (except for the special case when $x(2n-1)$ is a line spectral process). Therefore, the above normal equation can be solved in $\mathcal{O}(N^2)$ by using the Levinson fast algorithm. The optimal predictor \mathbf{p} is given by

$$\mathbf{p} = \mathbf{R}_x^{-1} \mathbf{r}. \quad (9)$$

In addition, the minimum achievable variance $\sigma_{x_0, \text{min}}^2$ is given by

$$\mathcal{E} = r(0) - \mathbf{r}^T \mathbf{R}_x^{-1} \mathbf{r} = r(0) - \sum_{k=-N}^{N-1} p(k)r(2k+1) \quad (10)$$

and the prediction gain is

$$G_p = \frac{\sigma_x^2}{\sigma_{x_0, \text{min}}^2} = \frac{\sigma_x^2}{\mathcal{E}} \geq 1. \quad (11)$$

The above inequality follows from the linear prediction theory [25]. The prediction gain is unity if and only if all the observations are uncorrelated to the target of prediction $x(2n)$. In this case, the optimal bit allocation formula in (5) reduces to

$$b_0 = b - \frac{1}{2} \log \sqrt{G_p}, \quad b_1 = b + \frac{1}{2} \log \sqrt{G_p}.$$

Therefore, the better the prediction is, the more bits are assigned to Q_1 , and the fewer bits are assigned to Q_0 . The coding gain in (6) then becomes $\mathcal{CG} = \sqrt{G_p}$. Note that in the derivation of (8), we have assumed that the autocorrelation matrix of the quantized observations $\hat{x}(2n-1)$ is very close to that of the original observation. This assumption is valid only when the bit rate is high so that the quantization noise variance is small. In the case of low bit rate coding, the autocorrelation matrix of $\hat{x}(2n-1)$ can differ significantly from that of $x(2n-1)$. This can result in a substantial loss in coding performance. In next

$$\mathbf{R}_x = \begin{pmatrix} r(0) & r(2) & r(4) & \cdots & r(4N-2) \\ r(2) & r(0) & r(2) & \cdots & r(4N-4) \\ r(4) & r(2) & r(0) & \cdots & r(4N-6) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r(4N-2) & r(4N-4) & r(4N-6) & \cdots & r(0) \end{pmatrix}, \quad \mathbf{p} = \begin{pmatrix} p(-N) \\ p(-N+1) \\ \vdots \\ p(N-1) \end{pmatrix}, \quad \mathbf{r} = \begin{pmatrix} r(2N-1) \\ r(2N-3) \\ \vdots \\ r(1) \\ r(1) \\ \vdots \\ r(2N-1) \end{pmatrix}.$$

section, we will show how to obtain the minimum mean-square-error (MMSE) predictor.

Linear-Phase Property: The optimal predictor $P(z)$ obtained from solving (8) has linear phase, i.e., $p(n) = p(-n-1)$. To see this, note that the matrix \mathbf{R}_x on the left-hand side of (8) satisfies

$$\mathbf{J}_{2N}\mathbf{R}_x\mathbf{J}_{2N} = \mathbf{R}_x$$

where \mathbf{J}_{2N} is the reversal matrix of size $2N$, as defined in Section I. Since the vector \mathbf{r} is symmetric, we have $\mathbf{J}_{2N}\mathbf{r} = \mathbf{r}$. Using these properties and the fact that $\mathbf{J}_{2N}\mathbf{J}_{2N} = \mathbf{I}_{2N}$, we can rewrite (8) as

$$\mathbf{R}_x(\mathbf{J}_{2N}\mathbf{p}) = \mathbf{r}.$$

Comparing the above equation and (8), we conclude that $\mathbf{J}_{2N}\mathbf{p} = \mathbf{p}$. The vector \mathbf{p} is symmetric and hence $P(z)$ has linear phase. Summarizing all the results, we have

Theorem 1: Consider the MINLAB coder in Fig. 2, where $p(n)$ is as in (7). The coding gain of the coder is maximized when $p(n)$ is chosen as the optimal prediction filter in (9). The optimal prediction filter has linear phase, and the maximum coding gain \mathcal{CG}_{\max} is given by

$$\mathcal{CG}_{\max} = \sqrt{G_p} \geq 1$$

where G_p is the prediction gain in (11). The coding gain is always greater than or equal to unity with equality if and only if the autocorrelation coefficients of $x(n)$ satisfy $r(2k+1) = 0$ for $0 \leq k \leq N-1$.

2) *Optimal IIR Biorthogonal Coders:* The predictor $P(z)$ can be taken as the more general IIR filter. The predictor $P(z)$ can be optimized such that $\sigma_{x_0}^2 = \int_0^{2\pi} S_x(e^{j\omega})|H_0(e^{j\omega})|^2 d\omega/2\pi$ is minimized. The special case when $P(z)$ is an allpass filter is studied in [24]. It was shown that for a wide class of random process, a first-order IIR prediction filter provides satisfactory coding results. To be more specific, the prediction filter $P(z)$ is taken as a first-order allpass function

$$P(z) = \frac{z^{-1} - \alpha}{1 - \alpha z^{-1}}.$$

The analysis filter in (1) becomes $H_0(z) = z^{-2} - z^{-1}(z^{-2} - \alpha)/(1 - \alpha z^{-2})$. IIR filters of this form are studied in detail in [4], and it was shown that they have many good features. For

example, they always have at least one zero at $z = 1$. In this case, the variance of $x_0(n)$ has the form [26]

$$\sigma_{x_0}^2 = 2 \left(r(0) + (-1 + \alpha + \alpha^2)r(1) - (\alpha - \alpha^3) \sum_{n=0}^{\infty} \alpha^n r(2n+3) \right)$$

To get the optimal IIR predictor, we can find α_{opt} such that the above quantity is minimized. If the input is a MA(1) process, then $\alpha_{\text{opt}} = -0.5$. If the input is an AR(1) process, then $\alpha_{\text{opt}} = (1 - \sqrt{1 + \rho^2})/\rho^2$ [26].

B. Connections and Comparisons with Other Coding Systems

Comparison with DPCM: The differential pulse coding modulation (DPCM) is also a prediction-based coding technique. The coding gain of a DPCM system is the prediction gain. The proposed biorthogonal coder differs from DPCM in a number of ways.

- 1) In DPCM (either open-loop or closed-loop), the decoding process always involves a feedback path. In the MINLAB coder, the reconstruction process uses only FIR filter if $P(z)$ is FIR.
- 2) In DPCM, the predictor can only make use of the past data for prediction. In the MINLAB coder, the predictor can use past and future data for prediction. Hence, $\sqrt{G_p}$ in (11) can be larger than the prediction gain in DPCM. That means, for certain inputs, the coding gain of the biorthogonal coder can be larger than that of a DPCM of the same complexity, as we will demonstrate in Section VI-B.
- 3) The optimal predictor in a DPCM system has minimum phase and, hence, cannot be linear phase. On the other hand, the optimal filter in the MINLAB coder always has linear phase for real-valued WSS processes. Moreover, the predictor is working at half of the input data rate. Therefore, the complexity of the MINLAB coder is about one fourth that of DPCM with the same prediction filter length.
- 4) Unlike DPCM, the MINLAB coder has a hierarchical structure. Therefore, progressive resolution transmission can be done, and coding schemes like the zerotree algorithms [15] and [27] can be used to further exploit the correlation among different scales.

A Very Low Delay Coder: In many applications such as speech coding, we need coders with low delay. To obtain a low delay coder, we can take the filter $P(z)$ as a causal FIR filter of the form $\sum_{n=0}^{N-1} p(n)z^{-n}$. Then, a causal implementation of Fig. 2 can be obtained by replacing the advanced chain by a

delay chain. In such a causal implementation of Fig. 2, regardless of the filter length N , the coder has a delay of only one sample (i.e., $y(n) = x(n-1)$). The optimal solution of $P(z)$ is given by the optimal causal predictor of $x(2n)$ based on the observations $x(2n-2N+1), x(2n-2N+3), \dots, x(2n-1)$.

Connection to the Optimal Orthonormal Coder: It was shown in [23] and [24] that the optimal FIR orthonormal coder (with order $N-1$) has a unity coding gain if and only if the autocorrelation coefficients of the input satisfy $r(2k+1) = 0$ for $0 \leq k \leq N-1$. Comparing this result with Theorem 1, we conclude that the optimal biorthogonal coder has a unity coding gain if and only if the optimal orthonormal coder of the same order has a unity coding gain. In fact, for most inputs, the biorthogonal coder outperforms the orthonormal coder, as we will see in Section VI.

Connection to the KLT: It is known that the Karhunen-Loeve transform (KLT) for the two-channel case has the form

$$\mathbf{T} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

and the coding gain of the KLT is given by

$$\mathcal{CG}_{\text{KLT}} = \frac{1}{\sqrt{1 - r^2(1)/r^2(0)}}.$$

For the proposed biorthogonal coder, if we take the filter $P(z) = p(0)$, then we call the coder a biorthogonal transform. The optimal predictor will be $P(z) = r(1)/r(0)$, and the coding gain for the biorthogonal transform is identical to $\mathcal{CG}_{\text{KLT}}$.

IV. MMSE PREDICTION FILTER FOR QUANTIZED OBSERVATION SAMPLES

With high bit rate assumption, $\hat{x}(2n-1) \approx x(2n-1)$. The optimal predictor in Section III is designed based on unquantized observations $x(2n-1)$. In this section, we will derive the MMSE predictor by taking into account that the observation samples are quantized data $\hat{x}(2n-2k-1)$. We will assume that the quantization noise is uncorrelated with the quantizer input. Let \mathbf{R}_q be the autocorrelation matrix of the quantization noise $q_1(n)$. Following steps similar to those in Section III, we can derive the optimal MMSE predictor

$$\mathbf{p}_{\text{mmse}} = [\mathbf{R}_x + \mathbf{R}_q]^{-1} \mathbf{r} \quad (12)$$

where \mathbf{R}_x and \mathbf{r} are defined in (8). The matrix $[\mathbf{R}_x + \mathbf{R}_q]$ continues to be positive definite and Toeplitz, and the optimal solution in (12) can be obtained by the Levinson fast algorithm. Moreover, $\mathbf{J}[\mathbf{R}_x + \mathbf{R}_q]\mathbf{J} = [\mathbf{R}_x + \mathbf{R}_q]$. Therefore, the MMSE prediction filter \mathbf{p}_{mmse} also has linear phase. To carry on the derivation, we will assume that the quantization noises $q_0(n)$ and $q_1(n)$ are uncorrelated, and their variances are equal (this is a reasonable assumption as the stepsizes of the quantizers are equal)

$$\mathbf{R}_q = \sigma_q^2 \mathbf{I}.$$

To further simplify (12), we need the matrix inversion formula

$$[\mathbf{A} + \mathbf{B}]^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}[\mathbf{A}^{-1} + \mathbf{B}^{-1}]^{-1} \mathbf{A}^{-1}$$

where \mathbf{A} and \mathbf{B} are nonsingular matrices of the same dimension. Applying the above identity to (12), we have

$$\begin{aligned} \mathbf{p}_{\text{mmse}} &= \mathbf{R}_x^{-1} \mathbf{r} - \sigma_q^2 \mathbf{R}_x^{-1} [\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1} \mathbf{R}_x^{-1} \mathbf{r} \\ &= \mathbf{p} - \sigma_q^2 \mathbf{R}_x^{-1} [\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1} \mathbf{p} \end{aligned} \quad (13)$$

where $\mathbf{p} = \mathbf{R}_x^{-1} \mathbf{r}$ is the optimal predictor in (9). Recall that \mathbf{p} is the optimal predictor when the observation samples used are unquantized original data $x(2n-2k-1)$. In the case of quantized observation samples, the optimal predictor \mathbf{p}_{mmse} has a correction term proportional to σ_q^2 . In the case of high bit rate coding, this term will be insignificant, and $\mathbf{p}_m \approx \mathbf{p}$. From (13), we can verify that the MMSE is given by

$$\mathcal{E}_{\text{mmse}} = \mathcal{E} + \sigma_q^2 \mathbf{p}^T [\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1} \mathbf{p} \quad (14)$$

where \mathcal{E} is the prediction error variance for the case of unquantized observation in (10). Since the matrix $[\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1}$ is a positive definite matrix, we have $\mathcal{E}_{\text{mmse}} \geq \mathcal{E}$ for all $\mathbf{p} \neq \mathbf{0}$. Therefore, the prediction error variance increases if the observation samples are the quantized data. This increase is proportional to the quantization error variance σ_q^2 . This explains why the coding gain of MINLAB decreases as the bit rate decreases.

1) Comparison of Performances of the Predictors in (8) and in (12) for Coarse Quantization: Consider the case when the predictor used is $\mathbf{p} = \mathbf{R}_x^{-1} \mathbf{r}$ but the observation samples are in fact the quantized data. Therefore, there is a mismatch between the predictor and its observation samples. In this case, we can show that the error variance is

$$\mathcal{E}_{\text{mismatch}} = \mathcal{E} + \sigma_q^2 \mathbf{p}^T \mathbf{p}.$$

Note that the amount of increase due to the mismatch is proportional to the quantization error variance and the prediction filter's energy. When the data are quantized coarsely, $\sigma_q^2 \mathbf{p}^T \mathbf{p}$ can become the dominant term, and the prediction error $\mathcal{E}_{\text{mismatch}}$ can even be larger than the original signal $x(2n)$ (a prediction loss). The performance of the predictor $\mathbf{p} = \mathbf{R}_x^{-1} \mathbf{r}$ can degrade significantly at low bit rate coding.

Comparing the above equation with (14), we have

$$\Delta \mathcal{E} = \mathcal{E}_{\text{mismatch}} - \mathcal{E}_{\text{mmse}} = \sigma_q^2 \mathbf{p}^T (\mathbf{I} - [\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1}) \mathbf{p}.$$

Since the matrix $\mathbf{I} - [\sigma_q^2 \mathbf{R}_x^{-1} + \mathbf{I}]^{-1}$ is positive definite, the difference $\Delta \mathcal{E} > 0$ for any nonzero predictor. In fact, we can verify that this difference is bounded by

$$\sigma_q^2 \mathbf{p}^T \mathbf{p} \frac{1}{1 + \lambda_{\max}/\sigma_q^2} \leq \Delta \mathcal{E} \leq \sigma_q^2 \mathbf{p}^T \mathbf{p} \frac{1}{1 + \lambda_{\min}/\sigma_q^2}$$

where λ_{\max} and λ_{\min} are, respectively, the maximum and minimum eigenvalues of \mathbf{R}_x^{-1} . We see that when the quantization error is large, the increase of prediction error caused by the mismatch of the predictor and the observation data can become very significant. Therefore, it is important to design a predictor that matches the observation samples, especially at low bit rate coding.

V. MERITS OF THE MINLAB CODER

The MINLAB coder in Fig. 2 enjoys many advantages. It has many other good features that make it attractive in various applications. In the following, we list some of its advantages.

- 1) **Structurally PR:** Similar to the orthonormal FB, the proposed biorthogonal FB has a structurally PR implementation, as in Fig. 2.
- 2) **Equal stepsize rule:** Like the orthonormal coder, the equal stepsize rule is an optimal quantization procedure, as discussed in the last section. Entropy coding can be applied to further compress the quantizer output.
- 3) **Unity noise gain:** The synthesis bank does not amplify the quantization noise. Hence, the optimal MINLAB coder has a coding gain $\mathcal{CG} \geq 1$. Moreover, we will show later that for two important classes of inputs [that is, the AR(1) process and MA(1) process], the optimal MINLAB coder with two taps outperforms the optimal orthonormal coder of any order.
- 4) **Low design cost:** The design of the optimal MINLAB coder is simple. Unlike the optimal orthonormal coder, neither constrained optimization nor spectral factorization is needed. Optimal MINLAB coder can be obtained by using Levinson algorithm.
- 5) **Low complexity:** To implement the analysis or synthesis bank, we need only one filter $P(z)$. Moreover, the optimal $P(z)$ has *linear phase*. Therefore, the complexity of the biorthogonal coder is roughly one fourth that of an orthonormal coder of the same order.
- 6) **Coding gain increases with N :** It is well-known that the prediction gain G_p is a nondecreasing function of N . Hence, the coding gain increases when the filter order increases.
- 7) **Low delay:** It is known that the delay of an orthonormal coder is proportional to the filter order [18]. The longer the filters are, the larger the system delay is. In the MINLAB coder, if $P(z)$ is a causal filter (either IIR or FIR), then the system delay is only one sample, regardless of the filter order. As the prediction gain increases with filter order, so does the coding gain. Therefore, we can improve the performance of such a biorthogonal coder without introducing extra system delay.
- 8) **Lossy/lossless compression:** Let the input $x(n)$ be a discrete amplitude signal with stepsize Δ_x . For many applications, the inputs are integers, and $\Delta_x = 1$. Suppose the output of $P(z)$ is quantized using a quantizer Q_p . Then, the MINLAB coder can be modified for lossless compression as follows.
 - a) Set the stepsize of Q_p , $\Delta_p = n\Delta_x$. Any type of quantizer (round off or truncation or ceiling) can be used as Q_p .
 - b) Set the stepsizes of the subband quantizers as $\Delta_0 = \Delta_1 = \Delta_x$. Use entropy coding to encode the outputs of Q_0 and Q_1 .

When $\Delta_0 = \Delta_1 > \Delta_x$, we have a lossy MINLAB coder. It becomes lossless when $\Delta_0 = \Delta_1 = \Delta_x$.

- 9) **Coding of finite length signal:** In a conventional subband coder, when the input has finite length L , the total

number of samples in the subband increases due to linear convolution with the analysis filters, unless these filters have length 2. Therefore, periodic extension is used to solve this problem. In the MINLAB coder in Fig. 2, to reconstruct the output signal, we need only to retain L samples in the subbands $\lfloor (L+1)/2 \rfloor$ samples of $x_0(n)$ and $\lfloor L/2 \rfloor$ samples of $x_1(n)$ where $\lfloor a \rfloor$ denotes the largest integer $\leq a$. No periodic extension is needed. To see this, let us ignore the quantizers in Fig. 2. It is clear that $x(2n)$ has $\lfloor (L+1)/2 \rfloor$ nonzero samples, and $x(2n-1)$ has $\lfloor L/2 \rfloor$ nonzero samples. Note that $x_1(n) = x(2n-1)$. Therefore, to reconstruct $x(2n-1)$, we need to retain only $\lfloor L/2 \rfloor$ samples of $x_1(n)$. Moreover, since $y_0(n) = x_0(n) + v(n) = x(2n)$ for $0 \leq n \leq \lfloor (L+1)/2 \rfloor$, we need to retain only the first $\lfloor (L+1)/2 \rfloor$ samples of $x_0(n)$ for the reconstruction of $x(2n)$.

- 10) **Incorporation of EZW and SPIHT algorithm:** As we will see in Section VII, the MINLAB coder in Fig. 2 can be generalized to obtain a tree structure MINLAB coder. Such a system continues to enjoy all of the properties listed above. Using this wavelet-type MINLAB coder, zerotree algorithms such as EZW and SPIHT [15], [27] can be applied.

VI. PERFORMANCE ANALYSIS

In this section, we will provide several examples to demonstrate the coding performance of the proposed biorthogonal coder. The results will be compared with the optimal orthonormal coder.

A. AR(1) Inputs

Let the input be an AR(1) process with $r(k) = \rho^{|k|}$ for $0 < \rho < 1$. For this AR(1) process, we compare the performance of the following various coders:

- 1) Let $P(z) = p(-1)z + p(0)$. From the normal equation (8), we get the optimal predictor as $p(0) = p(-1) = \rho/(1+\rho^2)$. In this case, the optimal coding gain has the closed-form expression

$$\mathcal{CG}_{\text{MINLAB}(2)} = \sqrt{\frac{1+\rho^2}{1-\rho^2}}$$

where the index 2 indicates that the predictor has two taps. In this case, however, only one multiplier is needed. We can verify that there is no need to use a longer predictor because the coding gain cannot be increased by using a longer filter. Therefore, the gain given in the above equation is the maximum coding gain that can be attained by MINLAB coder with any prediction filter order.

- 2) Take $P(z) = p(0)$. The optimal predictor is simply $P(z) = \rho$, and the coding gain is

$$\mathcal{CG}_{\text{MINLAB}(1)} = \frac{1}{\sqrt{1-\rho^2}}.$$

- 3) Consider the coding gain for optimal orthonormal coders with infinite taps and four taps. It was shown in [22]–[24] that the coding gains are, respectively

$$\mathcal{CG}_{\text{ortho}(\infty)} = \frac{1}{\sqrt{1 - (16/\pi^2)(\tan^{-1} \rho)^2}}$$

$$\mathcal{CG}_{\text{ortho}(4)} = \sqrt{\frac{1 + 1/3\rho^2}{1 - \rho^2}}.$$

- 4) The DPCM of order one is optimal in this case as the input is an AR(1) process. Its coding gain is given by

$$\mathcal{CG}_{\text{DPCM}(1)} = \frac{1}{1 - \rho^2}.$$

- 5) Suppose that we use the traditional biorthogonal coder in Fig. 1. Then, it can be shown that the maximum achievable coding gain for a two-tap filter $P(z)$ is given by

$$\mathcal{CG}_{\text{tradit}(2)} = \sqrt{\frac{1 + \rho^2}{1 - \rho^2}} \sqrt{\frac{1}{1 + E_p}}, \quad \text{where } E_p = \frac{2\rho^2}{(1 + \rho^2)^2}.$$

These gains are shown in Fig. 3. It is clear from the figure that $\mathcal{CG}_{\text{DPCM}(1)} > \mathcal{CG}_{\text{MINLAB}(2)} > \mathcal{CG}_{\text{ortho}(\infty)} > \mathcal{CG}_{\text{ortho}(4)} > \mathcal{CG}_{\text{MINLAB}(1)}$ for all possible ρ . Therefore, we see that for AR(1) process, the optimal MINLAB coder with two taps (one multiplier) outperforms the optimal orthonormal coder with infinite number of taps. From Fig. 3, we also note that for the traditional biorthogonal coder in Fig. 1, its coding gain $\mathcal{CG}_{\text{tradit}(2)} < 1$ for $\rho < 0.5$.

We also compare the performance of the MINLAB coder with the widely used 9/7 filters. Fig. 4 shows that the coding gain of MINLAB coder has a higher coding gain for AR(1) process with all ρ . In other words, the MINLAB coder has a better decorrelation ability than the 9/7 filters. The MINLAB coder has only two taps (one multiplication), whereas the 9/7 filters have 16 taps (eight multiplications). However, the 9/7 filters are signal independent, whereas the MINLAB coder is signal dependent.

B. MA(1) Inputs

Let the input be an MA(1) process with $r(0) = 1$, $r(\pm 1) = \rho$ for $0 < \rho < 0.5$, and $r(k) = 0$ for all the other k . Then, the optimal prediction filters of two taps and one tap are, respectively, $P(z) = \rho + \rho z$ and $P(z) = \rho$. The coding gain for these two biorthogonal coder are, respectively

$$\mathcal{CG}_{\text{MINLAB}(2)} = \frac{1}{\sqrt{1 - 2\rho^2}} \quad \mathcal{CG}_{\text{MINLAB}(1)} = \frac{1}{\sqrt{1 - \rho^2}}.$$

Note that in this case, there is no need to use a predictor longer than two taps. The coding gain cannot be increased by using a longer filter [this can be seen from the normal equation (8)]. Therefore, $\mathcal{CG}_{\text{MINLAB}(2)}$ above is the maximum coding gain that can be attained by MINLAB coder with any prediction filter order. For optimal orthonormal coder, it was shown [23], [24] that the coding gain is

$$\mathcal{CG}_{\text{ortho}(\infty)} = \frac{1}{\sqrt{1 - (16/\pi^2)\rho^2}}$$

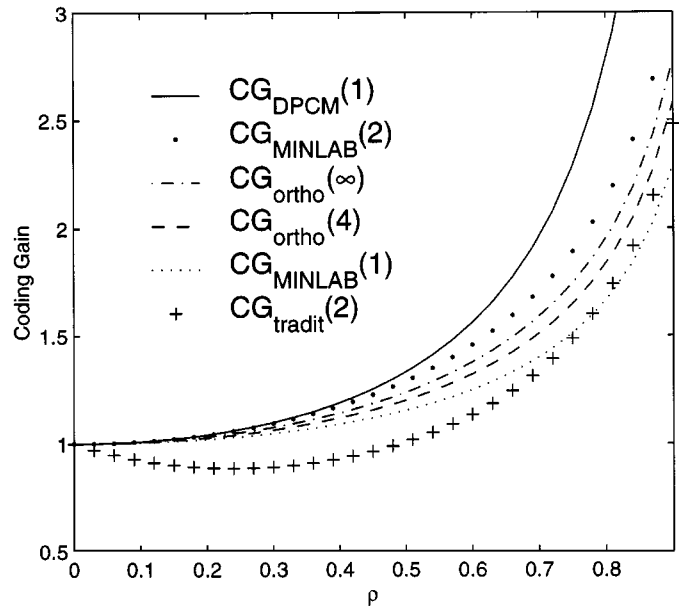


Fig. 3. Coding gain comparison for AR(1) process.

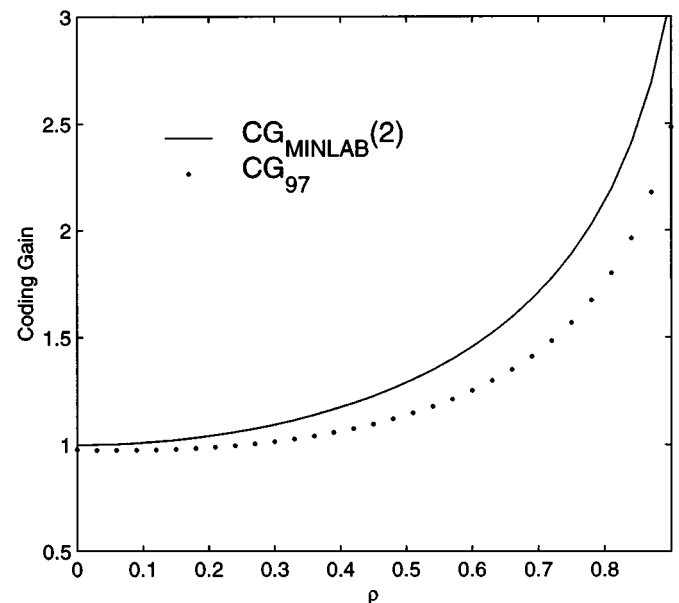


Fig. 4. Comparison of MINLAB coder and the 9/7 filters for AR(1) process.

$$\mathcal{CG}_{\text{ortho}(4)} = \frac{1}{\sqrt{1 - (4/3)\rho^2}}.$$

For the DPCM with one multiplier, its coding gain is

$$\mathcal{CG}_{\text{DPCM}(1)} = 1/(1 - \rho^2).$$

The coding gain for the traditional coder in Fig. 1 is

$$\mathcal{CG}_{\text{tradit}(2)} = 1/\sqrt{(1 - 2\rho^2)(1 + 2\rho^2)}.$$

From these coding gain expressions, it is not difficult to prove that

$$\mathcal{CG}_{\text{MINLAB}(2)} > \mathcal{CG}_{\text{DPCM}(1)} > \mathcal{CG}_{\text{ortho}(\infty)} > \mathcal{CG}_{\text{ortho}(4)} > \mathcal{CG}_{\text{MINLAB}(1)} > \mathcal{CG}_{\text{tradit}(2)}$$

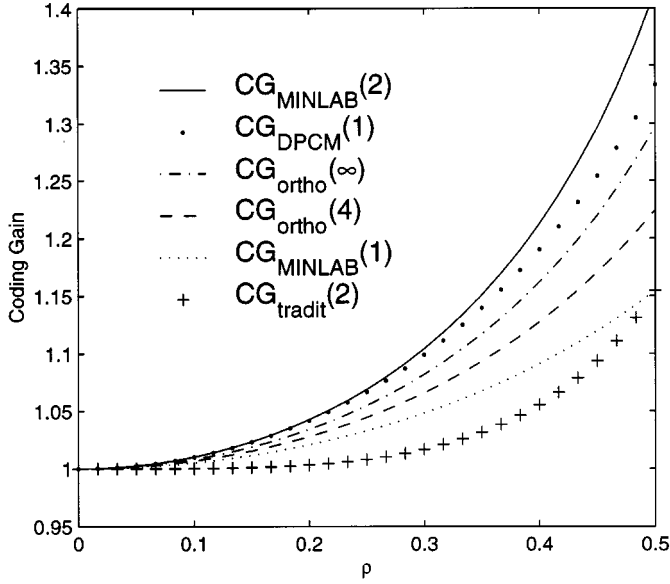


Fig. 5. Coding gain comparison for MA(1) process.

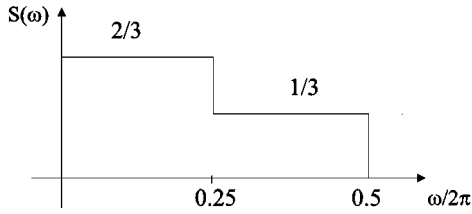


Fig. 6. Example of piecewise constant spectrum.

for all $0 < \rho < 0.5$. The optimal MINLAB coder with two taps again outperforms than the optimal orthonormal coder. Moreover the optimal MINLAB coder with two taps (one multiplication) is superior to the DPCM with one multiplication. All these gains are shown in Fig. 5.

C. Inputs With $r(2k) = \delta(k)$

For orthonormal coders, the coding gain is independent of the even autocorrelation coefficients $r(2K)$ [22]. In this subsection, we consider the case when the input $x(n)$ is a WSS process with its autocorrelation coefficients satisfying $r(2k) = \delta(k)$. One example that satisfies this condition is shown in Fig. 6. Since $r(2k) = \delta(k)$, we have $\mathbf{R}_x = \mathbf{I}$, and we have $p(n) = r(2n+1)$. Therefore, the optimal predictor in (9) has the closed form

$$P(z) = \sum_{n=-\infty}^{\infty} r(2n+1)z^{-n} = \sum_{n=0}^{\infty} r(2n+1)(z^{-n} + z^{n+1}). \quad (15)$$

It is shown in Appendix A that the variance of $x_0(n)$ can be expressed as

$$\sigma_{x_0, \min}^2 = \int_0^{2\pi} S(\omega)S(\omega + \pi) \frac{d\omega}{2\pi}. \quad (16)$$

Using Theorem 1, the coding gain of the MINLAB coder is

$$\mathcal{CG}_{\text{MINLAB}}(\infty) = \frac{1}{\sqrt{\int_0^{2\pi} S(\omega)S(\omega + \pi) d\omega/2\pi}} \quad (17)$$

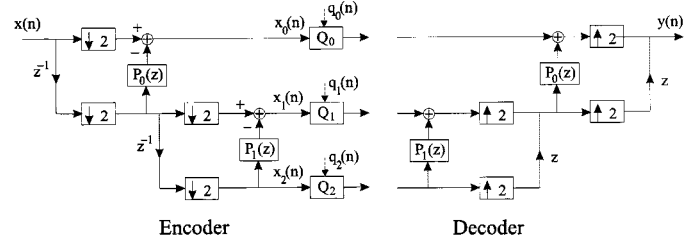


Fig. 7. Traditional tree structure subband coder using ladder.

where the argument ∞ indicates that there are infinite number of taps. If $x(n)$ is a Gaussian process, then the rate distortion theoretic bound [28] on the coding gain of any compression system is given by (see Appendix B)

$$\mathcal{CG}_{\text{RD}} = \frac{1}{\sqrt{\exp\left[\int_0^{2\pi} \ln(S(\omega)S(\omega + \pi)) d\omega/2\pi\right]}}. \quad (18)$$

This rate distortion theoretic bound can be achieved by DPCM of infinite order [25]. Using the fact that $\exp\left[\int_0^{2\pi} \ln A(\omega) d\omega/2\pi\right] \leq \int_0^{2\pi} A(\omega) d\omega/2\pi$ for any non-negative function $A(\omega)$ (with equality if and only if $A(\omega)$ is a constant), we can conclude from (17) and (18) that

$$\mathcal{CG}_{\text{MINLAB}}(\infty) \leq \mathcal{CG}_{\text{RD}}$$

and \mathcal{CG}_{RD} can be achieved by the MINLAB coder if and only if the product $S(\omega)S(\omega + \pi)$ is a constant. For the optimal orthonormal coder with an infinite order, it is shown [22] that $\mathcal{CG}_{\text{ortho}}(\infty)$ achieves \mathcal{CG}_{RD} in (18) if and only if $r(2k) = \delta(k)$ and $S(\omega)S(\omega + \pi)$ is a constant. Therefore, both the $\mathcal{CG}_{\text{MINLAB}}(\infty)$ and $\mathcal{CG}_{\text{ortho}}(\infty)$ achieve the rate distortion theoretic bound for the same class of spectrum. For the example shown in Fig. 6, we can verify that $\mathcal{CG}_{\text{MINLAB}}(\infty) = \mathcal{CG}_{\text{ortho}}(\infty) = \mathcal{CG}_{\text{RD}} = \sqrt{9/8}$.

VII. TREE STRUCTURE MINLAB CODERS

In the MINLAB FB shown in Fig. 2, the energy is mostly in the lower branch. Therefore, we decompose the lower subband signal $x_1(n)$ to obtain a biorthogonal wavelet decomposition, as shown in Fig. 7. In a traditional wavelet coding, the quantizers are placed directly after the subband signals as in Fig. 7. There will be a mismatch between the encoder and decoder. At the encoder, the inputs to $P_0(z)$ and $P_1(z)$ are the original data, whereas at the decoder, the inputs to $P_0(z)$ and $P_1(z)$ are the quantized data. It can be verified that this mismatch will cause the amplification of quantization noise in the reconstruction process. In the following, we will introduce a minimum noise structure with unity noise gain for the ladder-based tree structure biorthogonal coder. We will first derive the results for the two-level decomposition in Fig. 7, and then the results of the more general L -level decomposition will be stated without proof.

A. Two-Level Tree Structure MINLAB Coder

To avoid mismatch, we can use the quantized data as the input to $P_0(z)$ and $P_1(z)$ at the encoder. This can be done by modifying the encoder as in Fig. 8. Using an analysis similar to the

two-channel case, it can be shown that the following relations continue to hold:

$$\begin{aligned} y_0(n) - x(2n) &= q_0(n) \\ y_1(n) - x(4n-1) &= q_1(n) \\ y_2(n) - x(4n-3) &= q_2(n). \end{aligned}$$

Therefore, the average output noise variance is given by

$$\begin{aligned} \sigma_{q_{\text{out}}}^2 &= \frac{1}{2}\sigma_{q_0}^2 + \frac{1}{4}\sigma_{q_1}^2 + \frac{1}{4}\sigma_{q_2}^2 \\ &= \frac{c}{4} (2 * 2^{-2b_0}\sigma_{x_0}^2 + 2^{-2b_1}\sigma_{x_1}^2 + 2^{-2b_2}\sigma_x^2) \end{aligned} \quad (19)$$

where we have used (2) and the fact that $\sigma_{x_2}^2 = \sigma_x^2$. The average bit rate in this case is

$$b = \frac{1}{2}b_0 + \frac{1}{4}b_1 + \frac{1}{4}b_2.$$

Using the above equation and applying the AM-GM inequality to (19), we have

$$\sigma_{q_{\text{out}}}^2 \geq c2^{-2b}\sigma_{\text{GM}}^2, \quad \text{where } \sigma_{\text{GM}}^2 = \left[(\sigma_{x_0}^2)^2 \sigma_{x_1}^2 \sigma_x^2 \right]^{1/4} \quad (20)$$

with equality if and only if the bits are allocated as

$$b_i = b + \frac{1}{2} \log \sigma_{x_i}^2 - \frac{1}{2} \log \sigma_{\text{GM}}^2.$$

Note that there is no feedback loop in the minimum noise structure shown in Fig. 8. Therefore, stability is always guaranteed.

Optimizing the Filters $P_0(z)$ and $P_1(z)$: Let $P_0(z)$ and $P_1(z)$ be FIR filters of the form

$$P_0(z) = \sum_{n=-N_0}^{N_0-1} p_0(n)z^{-n}, \quad P_1(z) = \sum_{n=-N_1}^{N_1-1} p_1(n)z^{-n}.$$

From (20), the lower bound is minimized if $P_0(z)$ and $P_1(z)$ are designed to minimize $\sigma_{x_0}^2$ and $\sigma_{x_1}^2$, respectively. To minimize the variance of $x_0(n)$, the filter $P_0(z)$ is chosen as the optimal predictor obtained in Section III. To minimize the variance of $x_1(n)$, the filter $P_1(z)$ is designed as the optimal predictor of $x(4n-1)$ based on the observations of $x(4n-4N_1+1), x(4n-4N_1+5), \dots, x(4n+4N_1-3)$. The optimal solution of $P_1(z)$ can be obtained from a normal equation similar to (8). We can show that both $P_0(z)$ and $P_1(z)$ have linear phase. The prediction gains of these filters are

$$\begin{aligned} G_{p_0} &= \frac{\sigma_x^2}{\sigma_{x_0}^2} \geq 1 \\ G_{p_1} &= \frac{\sigma_x^2}{\sigma_{x_1}^2} \geq 1. \end{aligned}$$

The inequalities follow from linear prediction theory. The gain $G_{p_0} = 1$ if and only if the autocorrelation coefficients $r(2k+1) = 0$ for $0 \leq k \leq N_0 - 1$. The gain $G_{p_1} = 1$ if and only if $r(4k+2) = 0$ for $0 \leq k \leq N_1 - 1$. Under the optimal bit allocation, the coding gain is given by

$$\mathcal{CG}_{\text{max}} = (G_{p_0})^{1/2}(G_{p_1})^{1/4} \geq \sqrt{G_{p_0}}.$$

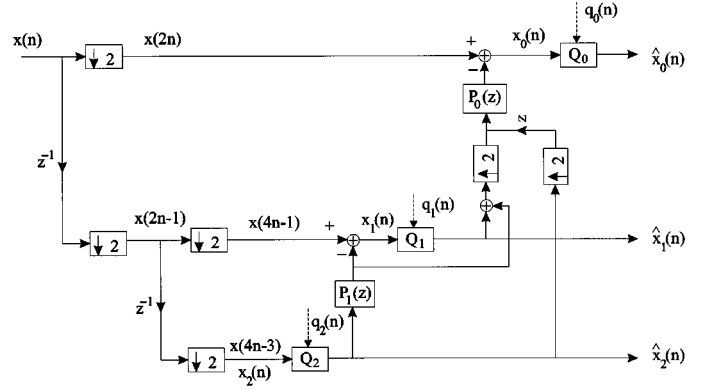


Fig. 8. Tree structure MINLAB encoder. Its decoder is the same as that in Fig. 7.

Therefore, the maximum coding gain of the two-level MINLAB coder in Fig. 8 is never less than that of the one-level MINLAB coder in Fig. 2.

B. L -Level Tree Structure MINLAB Coder

The same idea can be extended to the more general L -level tree structure FB. If the inputs to all the prediction filters $P_k(z)$ at the encoder are quantized data, the noise gain will be unity, and such a L -level tree structure will have the minimum noise property. Since all the derivations are very similar to the two-level decomposition, we will simply state the results below.

Let G_{p_k} be the prediction gain of the predictor at the $(k+1)$ th level. Then, the optimal bit allocation formula becomes

$$b_k = b - \frac{1}{2} \log G_{p_k} + \frac{1}{2} \log \left[G_{p_0}^{2^{L-1}} G_{p_1}^{2^{L-2}} \dots G_{p_{L-1}} \right]^{1/2^L}.$$

In addition, under the optimal bit allocation, the coding gain is

$$\mathcal{CG}_{\text{max}} = (G_{p_0})^{1/2}(G_{p_1})^{1/4} \dots (G_{p_{L-1}})^{1/2^L}. \quad (21)$$

Since all $G_{p_k} \geq 1$, we conclude that the coding gain is unity if and only if all of the prediction gains are unity. Moreover, the coding gain satisfies the recursive formula

$$\mathcal{CG}_{\text{max},l} = \mathcal{CG}_{\text{max},l-1}(G_{p_{l-1}})^{1/2^l}$$

where $\mathcal{CG}_{\text{max},l}$ is the coding gain of a l -level tree structure FB. Therefore, the coding gain always improves when we increase the number of levels. However, the coding gain saturates as L increases, as we will see in the example at the end of this section.

Comments on the Complexity: If the predictor $P_i(z)$ is chosen as $\sum_{n=-N_i}^{N_i} p_i(n)z^{-n}$, then the optimal predictors will have linear phase. To implement one level of the MINLAB encoder, we need to implement one linear phase predictor and two adders (except for the L th level, which needs only one adder). Therefore, the complexity of an L th-level MINLAB encoder is that of $P_i(z)$ and $(2L-1)$ additions. Comparing the non-MINLAB encoder with traditional structure in Fig. 7, the MINLAB encoder needs $(L-1)$ more additions. The complexity of decoder for both cases is the same.

Asymptotic Results as L Approaches Infinity: Let $P_k(z) = \sum_{n=-\infty}^{\infty} p_k(n)z^{-n}$. We will study the asymptotic results of this ideal case as the number of level L approaches infinity.

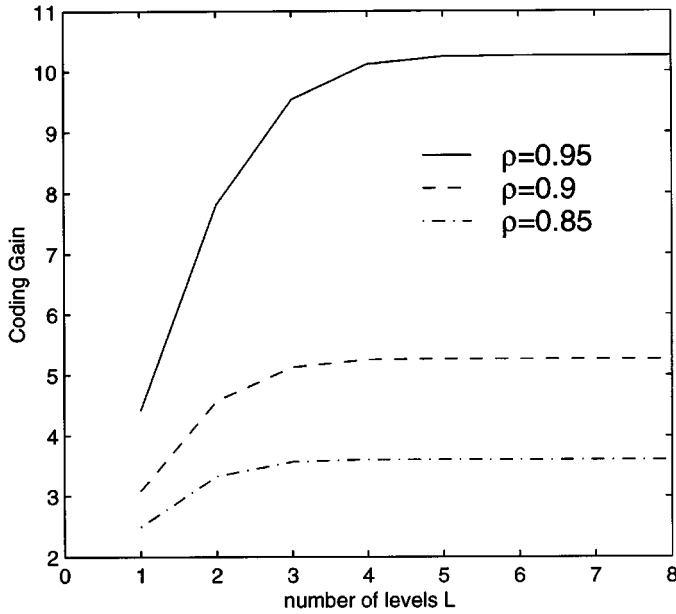


Fig. 9. Coding gain of tree structure MINLAB coder.

From Theorem 1, we know that $G_{P_k} = 1$ if and only if the autocorrelation coefficients satisfy $r(2^{k+1}n - 2^k)$ for all $n = 0, \pm 1, \pm 2, \dots$. By using

$$\bigcup_{k=0}^{\infty} \{2^{k+1}n - 2^k\} = \{\pm 1, \pm 2, \pm 3, \dots\}$$

it is not difficult to see that the coding gain of a L -level ideal MINLAB coder \mathcal{CG}_L satisfies $\lim_{L \rightarrow \infty} \mathcal{CG}_L = 1$ if and only if the input signal is white.

Example: Let the input be an AR(1) process with $r(k) = \rho^{|k|}$. Using the normal equation (8), we can get the prediction gain of the k predictor as

$$G_{P_k} = \frac{1 + \rho^{2^{k+1}}}{1 - \rho^{2^{k+1}}}.$$

Substituting this result into (21), the coding gain of an L -level is

$$\mathcal{CG}_L = \left[\frac{1 + \rho^2}{1 - \rho^2} \right]^{1/2} \left[\frac{1 + \rho^4}{1 - \rho^4} \right]^{1/4} \cdots \left[\frac{1 + \rho^{2^L}}{1 - \rho^{2^L}} \right]^{1/L}.$$

This coding gain is plotted in Fig. 9 for three different values of ρ . Compared with the case of the one-level MINLAB in Fig. 2, the coding gain of the two-level or three-level MINLAB coder is much larger. In addition, note that the coding gain saturates as the number of levels L increases. The saturation point L_{sat} depends on the correlation of the data. For AR(1) process, $L_{\text{sat}} = 3, 4, 5$ for $\rho = 0.85, 0.9, 0.95$, respectively.

VIII. A MORE GENERAL MINLAB CODER

The biorthogonal coder in Fig. 2 uses only one ladder. Such a system is a special case of the more general biorthogonal FB shown in Fig. 10. The FB in Fig. 10 is studied in detailed in [4]. In the simple coder of Fig. 2, only the variance of the even samples is reduced by prediction from odd samples, but the variance

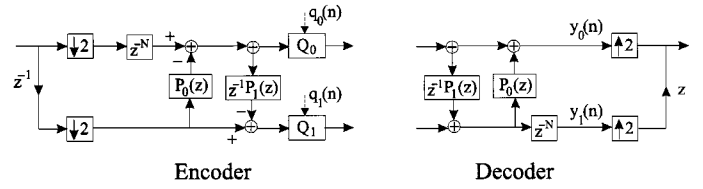


Fig. 10. More general ladder-based biorthogonal coder.

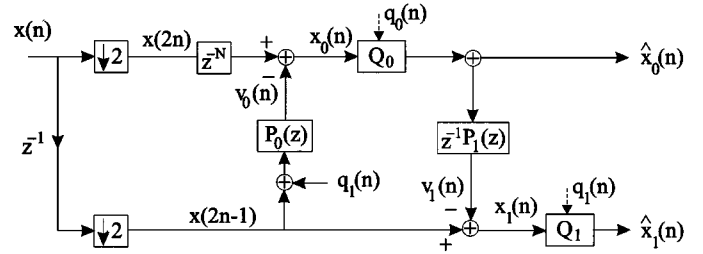


Fig. 11. MINLAB encoder for the system in Fig. 10. Its decoder is the same as that in Fig. 10.

of the odd samples is unchanged. Hence, its performance is limited by the lower branch of Fig. 2. In the following, we will show how to apply the more general structure of Fig. 10 to reduce the variance of $x_1(n)$.

In this section, we assume that the filters $P_0(z)$ and $P_1(z)$ are causal filters of the form

$$P_0(z) = \sum_{n=0}^{N_0-1} p_0(n)z^{-n}, \quad P_1(z) = \sum_{n=0}^{N_1-1} p_1(n)z^{-n}.$$

The choice of N , N_0 , and N_1 will be discussed later in the design. Note that the filter on the second ladder is $z^{-1}P_1(z)$ instead of $P_1(z)$. This delay is inserted so that there will be no delay-free loop in the MINLAB structure. For the coder in Fig. 10, its minimum noise structure is shown in Fig. 11. In the minimum noise structure, the quantization noise $q_1(n)$ is added to the first ladder so that the input to $P_0(z)$ is the quantized data $\hat{x}(2n-1)$. Since $v_1(n)$ depends only on the past values of $q_1(n)$, we do not have a delay-free loop. To see why the noise gain is unity for the MINLAB coder, we can carry out the following analysis. From Fig. 11, we have

$$\begin{aligned} \hat{x}_1(n) &= x(2n-1) - v_1(n) + q_1(n) \\ y_1(n) &= \hat{x}_1(n-N) + v_1(n-N) \\ &= x(2n-2N-1) + q_1(n-N) \\ \hat{x}_0(n) &= x(2n-2N) - v_0(n) + q_0(n) \\ y_0(n) &= \hat{x}_0(n) + v_0(n) = x(2n-2N) + q_0(n). \end{aligned}$$

From the above equations, we get $y_0(n) - x(2n-2N) = q_0(n)$ and $y_1(n) - x(2n-2N-1) = q_1(n-N)$. The noise gains for the upper and lower branch are both unity. By carrying out the similar derivation as in Section III, we can verify that the coding gain is maximized if the product $\sigma_{x_0}^2 \sigma_{x_1}^2$ is minimized, and the maximized coding gain can be expressed as

$$\mathcal{CG}_{\text{max}} = \frac{\sigma_x^2}{\min_{P_0, P_1} \sqrt{\sigma_{x_0}^2 \sigma_{x_1}^2}}. \quad (22)$$

Designing the Predictors to Maximize the Coding Gain in (22): From Fig. 11, it is clear that $\sigma_{x_0}^2$ depends only on $P_0(z)$, whereas $\sigma_{x_1}^2$ depends on both $P_0(z)$ and $P_1(z)$. To find the global optimality, we need nonlinear optimization. Therefore, we consider a suboptimal solution. Since $\sigma_{x_0}^2$ depends only on $P_0(z)$, we can design $P_0(z)$ to be the optimal predictor of $x(2n - 2N)$. After designing $P_0(z)$, the subband signal $x_0(n)$ is known. Its autocorrelation coefficients $r_{x_0}(k)$ and its cross correlation coefficients with $x(2n - 1)$ can be calculated. Using this information, the optimal predictor $P_1(z)$ can be derived from the normal equation. Note that we are using the residue of the even samples for prediction of $x(2n - 1)$. In general, the choice of N , N_0 , and N_1 will affect the coding gain. One way to decide these parameters is the following.

- 1) *Choice of N_0 :* Note that we are using the odd samples on the two sides of $x(2n - 2N)$ for the prediction of $x(2n - 2N)$. To get a linear-phase filter $P_0(z)$, we can choose $N_0 = 2N - 1$.
- 2) *Choice of N :* When N increases, the prediction gain of $P_0(z)$ increases. However, the correlation between the observation $x_0(n)$ and the target $x(2n - 1)$ typically becomes weaker as N grows. The prediction gain of $P_1(z)$ decreases as N increases. Therefore, there is a tradeoff between the gains of the two predictors $P_0(z)$ and $P_1(z)$.
- 3) *Choice of N_1 :* As N_1 increases, the prediction gain of $P_1(z)$ increases, but the complexity also grows. Note that in general, $P_1(z)$ is not linear phase.

A Note on the Stability Issue: Note that there is no feedback loop in the minimum noise structure in Fig. 11. The system is stable if both $P_0(z)$ and $P_1(z)$ are stable filters. We can verify that the noise transfer matrix from the quantizers to the subbands has the form

$$\begin{pmatrix} 1 & -P_0(z) \\ -z^{-1}P_1(z) & 1 + z^{-1}P_0(z)P_1(z) \end{pmatrix}.$$

The noise transfer matrix is FIR. Moreover, its determinant is unity. Hence, it is a unimodular matrix, and its inverse is also FIR.

IX. CONCLUDING REMARKS

In this paper, we have proposed a minimum noise structure (Fig. 2) for the class of ladder-based FB. The coding performance of the proposed biorthogonal coder is analyzed in detail. For both AR(1) process and MA(1) process, the optimal biorthogonal coder with two taps (with a complexity of only one multiplication) outperforms *any* optimal orthonormal coders. In addition to its excellent coding performance, the coder has many other desired features (see Section V). These features make the biorthogonal coder a potentially valuable and attractive alternative to the orthonormal coder. We have also generalized the minimum noise structure to the following two cases: 1) Tree structure biorthogonal FB and 2) more general biorthogonal FB with more than one ladder. In [29], we introduce a novel prediction-based lower triangular transform (PLT). The new transform is biorthogonal and has a minimum noise structure with unity noise gain. The PLT has an identical coding gain as the KLT, but it has a much lower design and implementation cost.

Given the input statistics, the PLT can be obtained by using the Levinson fast algorithm. In the special case of AR(1) process, the optimal biorthogonal transform coder has a closed-form expression, and no optimization is required.

APPENDIX A

PROOF OF (16) IN EXAMPLE 1

Since $P(z)$ is given in (15), the analysis filter $H_0(e^{j\omega})$ in (1) becomes

$$\begin{aligned} H_0(\omega) &= 1 - \sum_{n=-\infty}^{\infty} r(2n+1)e^{-j\omega(2n+1)} \\ &= \sum_{n=-\infty}^{\infty} (-1)^n r(n)e^{-j\omega n} \end{aligned}$$

where we have used the fact that $r(2k) = \delta(k)$. Comparing the above equation with the power spectrum $S(\omega) = \sum r(n)e^{-j\omega n}$, we have $H_0(\omega) = S(\omega + \pi)$. Since the downsampler does not change the variance, the minimum variance of $x_0(n)$ can be expressed as

$$\sigma_{x_0, \min}^2 = \int_0^{2\pi} S(\omega)S^2(\omega + \pi) \frac{d\omega}{2\pi}. \quad (23)$$

Since $r(2k) = \delta(k)$, we have $S(\omega) + S(\omega + \pi) = 2$ for all ω . Using this relation and the fact that $S(\omega)$ is periodic, we can further simplify (23) as

$$\begin{aligned} \sigma_{x_0, \min}^2 &= \int_0^{\pi} S(\omega)S^2(\omega + \pi) \frac{d\omega}{2\pi} + \int_{\pi}^{2\pi} S(\omega)S^2(\omega + \pi) \frac{d\omega}{2\pi} \\ &= \int_0^{\pi} S(\omega)S^2(\omega + \pi) \frac{d\omega}{2\pi} + \int_0^{\pi} S(\omega + \pi)S^2(\omega) \frac{d\omega}{2\pi} \\ &= \int_0^{\pi} S(\omega)S(\omega + \pi)[S(\omega + \pi) + S(\omega)] \frac{d\omega}{2\pi} \\ &= 2 \int_0^{\pi} S(\omega)S(\omega + \pi) \frac{d\omega}{2\pi} = \int_0^{2\pi} S(\omega)S(\omega + \pi) \frac{d\omega}{2\pi}. \end{aligned}$$

APPENDIX B

THE RATE DISTORTION THEORETIC BOUND

For a Gaussian WSS process with power spectrum $S(\omega)$, the minimum distortion that can be achieved by any b -bit coding system is given by [28]

$$D_{\min}(b) = 2^{-2b} \exp \left[\int_0^{2\pi} \ln S(\omega) \frac{d\omega}{2\pi} \right].$$

To obtain the coding gain \mathcal{CG}_{RD} of the form in (18), we carry out the derivation

$$\begin{aligned} \int_0^{2\pi} \ln S(\omega) \frac{d\omega}{2\pi} &= \int_0^{\pi} \ln S(\omega) \frac{d\omega}{2\pi} + \int_{\pi}^{2\pi} \ln S(\omega) \frac{d\omega}{2\pi} \\ &= \int_0^{\pi} \ln S(\omega) \frac{d\omega}{2\pi} + \int_0^{\pi} \ln S(\omega + \pi) \frac{d\omega}{2\pi} \\ &= \frac{1}{2} \int_0^{2\pi} \ln(S(\omega)S(\omega + \pi)) \frac{d\omega}{2\pi}. \end{aligned}$$

ACKNOWLEDGMENT

The authors are thankful to Prof. T. Q. Nguyen, Boston University, Boston, MA, for bringing our attention to the recent developments on lifting schemes, which inspired this work.

REFERENCES

- [1] T. Endoh and Y. Tamazaki, "Progressive coding scheme for multi level images," in *Proc. Picture Coding Symp.*, Tokyo, Japan, Apr. 1986, pp. 21–22.
- [2] A. M. Bruekers and A. W. M. van den Enden, "New networks for perfect inversion and perfect reconstructions," *IEEE J. Select. Areas Commun.*, vol. 12, Jan. 1992.
- [3] C. W. Kim and R. Ansari, "FIR/IIR exact reconstruction filter banks with applications to subband coding of images," in *Proc. Midwest CAS Symp.*, Monterey, CA, May 1991.
- [4] 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.
- [5] C. W. Kim and R. Ansari, "Subband decomposition procedure for quincunx sampling grids," in *Proc. SPIE-Visual Commun. Image Process.*, Boston, MA, Nov. 1991.
- [6] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Appl. Comput. Harmon. Anal.*, vol. 3, no. 2, pp. 186–200, 1996.
- [7] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM J. Math. Anal.*, vol. 29, no. 2, pp. 511–546, 1997.
- [8] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 247–269, 1998.
- [9] R. Claypoole, G. Davis, W. Sweldens, and R. Baraniuk, "Nonlinear wavelet transform for image coding," in *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, vol. 1, 1997, pp. 662–667.
- [10] R. L. de Queiroz, D. A. F. Florencio, and R. W. Schafer, "Nonexpansive pyramid for image coding using nonlinear filterbank," *IEEE Trans. Image Processing*, vol. 7, pp. 246–252, Feb. 1998.
- [11] P. Roos, A. Viergever, and M. C. A. van Dijke, "Reversible intraframe compression of medical images," *IEEE Trans. Med. Imag.*, vol. 7, pp. 328–336, Sept. 1988.
- [12] G. R. Kuduvali and R. M. Rangayyan, "Performance analysis of reversible image compression techniques for high-resolution digital telerradiology," *IEEE Trans. Med. Imag.*, vol. 11, pp. 430–445, Sept. 1992.
- [13] A. Said and W. A. Pearlman, "Reversible image compression via multiresolution and predictive coding," *Proc. SPIE*, pp. 664–674, Nov. 1993.
- [14] W. J. Ho and W. T. Chang, "Multiresolution interpolative DPCM for data compression," *Proc. SPIE Vis. Commun. Image Process.*, 1995.
- [15] A. Said and W. A. Pearlman, "An image multiresolution representation for lossless and lossy compression," *IEEE Trans. Image Processing*, vol. 5, pp. 1303–11, Sept. 1996.
- [16] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. Yeo, "Lossless image compression using integer to integer wavelet transforms," in *Proc. IEEE Int. Conf. Image Process.*, 1997, p. 596.
- [17] M. Iwahashi, S. Fukuma, and N. Kambayashi, "Lossless coding of still images with four channel prediction," in *Proc. IEEE Int. Conf. Image Process.*, 1997, p. 266.
- [18] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [19] S.-M. Phoong and Y. P. Lin, "A new class of optimal biorthogonal coder," *IEEE Signal Processing Lett.*, vol. 6, pp. 4–6, Jan. 1999.

- [20] S.-M. Phoong and Y. P. Lin, "Optimal ladder-based biorthogonal coder," in *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process.*, Mar. 1999.
- [21] M. K. Tsatsanis and G. B. Giannakis, "Principle component filter banks for optimal multiresolution analysis," *IEEE Trans. Signal Processing*, vol. 43, pp. 1766–1777, Aug. 1995.
- [22] P. P. Vaidyanathan, "Theory of optimal orthonormal coder," *IEEE Trans. Signal Processing*, vol. 47, pp. 1528–1543, June 1998.
- [23] A. Kirac and P. P. Vaidyanathan, "Theory and design of optimal FIR compaction filters," *IEEE Trans. Signal Processing*, vol. 47, pp. 903–919, Apr. 1998.
- [24] J. Tuqan and P. P. Vaidyanathan, "Globally optimal two channel FIR orthonormal filter banks adapted to the input signal statistics," presented at the the IEEE Int. Conf. Acoust., Speech Signal Process., 1998.
- [25] J. Makhoul, "Linear prediction: A tutorial review," *Proc. IEEE*, vol. 63, pp. 561–580, 1975.
- [26] J. Tuqan and P. P. Vaidyanathan, "Optimum low cost two channel IIR orthonormal filter bank," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 1998, pp. 2425–2428.
- [27] J. M. Shapiro, "Embedded image coding using zerotrees of wavelets," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.
- [28] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
- [29] S.-M. Phoong and Y. P. Lin, "PLT versus KLT," in *Proc. Int. Symp. Circuits Syst.*, May 1999.



See-May Phoong (M'96) was born in Johor, Malaysia, in 1968. He received the B.S. degree in electrical engineering from the National Taiwan University (NTU), Taipei, Taiwan, R.O.C., in 1991 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 1992 and 1996, respectively.

He was with the Faculty of the Department of Electronic and Electrical Engineering, Nanyang Technology University, Singapore, from September 1996 to September 1997. Since September 1997, he has been an Assistant Professor with the Institute of Communication Engineering and Electrical Engineering, NTU. His interests include signal compression, transform coding, and filter banks and their applications to communications.

Dr. Phoong was the recipient of the 1997 Wilts Prize at Caltech for Outstanding Independent Research in Electrical Engineering.



Yuan-Pei Lin (S'93–M'97) was born in Taipei, Taiwan, R.O.C., in 1970. She received the B.S. degree in control engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 1992 and the M.S. and Ph.D. degrees, both in electrical engineering, from the California Institute of Technology, Pasadena, in 1993 and 1997, respectively.

She joined the Department of Electrical and Control Engineering, National Chiao-Tung University, in 1997. Her research interests include multirate filter banks and applications to communication systems.

Dr. Lin is currently an Associate Editor of the journal *Multidimensional Systems and Signal Processing*.