Filter Banks for Prediction-Compensated Multiple Description Coding

Jing Wang, *Student Member, IEEE*, and Jie Liang, *Member, IEEE*

***Abstract*—In this paper, a prediction-compensated multiple description (MD) coding framework for two-band filter banks is proposed, in which the coefficients in each subband are split into two descriptions. Each description also includes the prediction residuals of the data in the other description. The designs of the optimal orthogonal and biorthogonal filter banks are formulated in a unified framework, and both one-level and multiple-level decompositions are analyzed. Contrary to the existing MD filter banks in the literature, the optimal filter banks in the proposed scheme are quite similar to those in single description coding. Therefore, the method can be applied to systems with single-description-optimized filter banks and still attain near-optimal performance. Image coding results show that this method achieves better performance and lower complexity than the latest JPEG 2000 based MDC.**

***Index Terms*—Filter bank, image coding, linear prediction, multiple description coding, wavelet transform.**

I. Introduction

MULTIPLE description coding (MDC) [1] is an attractive technique of combating transmission errors. In MDC, the source signal is encoded into several coded streams called descriptions, which are sent to the receiver via different network paths. Judiciously designed redundancies are introduced in all descriptions such that the reconstruction quality degrades gracefully when some of them are lost. In this paper, we focus on MDC with two descriptions and three decoders, where the decoder receiving one and two descriptions is called the *side decoder* and the *central decoder*, respectively. In particular, we are interested in wavelet transform based MDC and its application in image coding.

Many transform coding based MDC schemes have been developed, which can be classified into two categories. In the first category, each description contains the same number of coefficients as the transform outputs, whereas in the second category, a description only has half of transform outputs.

Some schemes of the first category are based on the multiple description scalar quantizer (MDSQ) [2], [3], which introduces redundancy implicitly via a central quantizer and an index assignment. The index assignment creates two side quantizers such that each can produce an acceptable side distortion, and their combination yields the finer central quantizer. The MDSQ is asymptotically near optimal at high rates [4], [5], and has been used in wavelet image coding in [6]. However, the design and implementation of the index assignment are quite challenging, and the redundancy is not easy to adjust. A two-stage modified MDSQ (MMDSQ) with the same asymptotical performance as the MDSQ is developed in [5], which only involves scalar quantizers and avoids the index assignment. Together with the wavelet transform and Tarp filter, the MMDSQ achieves one of the best MD image coding results.

In [7], the theoretical performance of the MDSQ-based multiple description transform coding (MDTC) is studied by high resolution analysis, which shows that the Karhunen–Loève transform (KLT) is still the optimal MD block transform for stationary sources. An information-theoretic scheme is recently developed in [8], where a vector MD quantizer is used after the block transform instead of MDSQs. The analysis in it confirms that the KLT is indeed rate-distortion optimal.

Although the MDTC approach is asymptotically optimal for stationary signals, in practice it may not always have the best performance. Hence, it is worthwhile to investigate alternative approaches. One reason is that the KLT is only optimal when the block size goes to infinite. For finite block sizes, practical fast approximations of the KLT such as the DCT cannot decorrelate the signals perfectly, and filter banks (FBs) with longer filters are needed to get better coding efficiency, making it necessary to study the optimal filter bank for a particular MDC scheme. In addition, when a finite block size is used, the correlation between neighboring blocks is not utilized by the MDTC. Better results can be achieved by taking full advantage of the source correlation. Moreover, different technologies from the MDTC might be more suitable when the source is nonstationary, such as natural images.

Other methods in the first MDC category that work well in practice include the direct coding method in [9] and [10], where the redundancy is introduced explicitly by grouping the wavelet transform outputs into two parts, and each description includes one part with a fine quantization and another part with a coarse quantization. The low rate part is the redundant information, which is discarded when two descriptions are received. If one description is lost, all data in the description are used in the reconstruction. Recently this method is applied to JPEG 2000 under the name of RD-MDC [11], where the Lagrangian method is used to find the optimal data partition and bit allocation. However, to get balanced descriptions, the RD-MDC needs to classify all JPEG 2000 codeblocks into two subsets, such that any codeblock in one subset has similar characteristics to another codeblock in the other subset. This procedure is quite time consuming. The complexity can be reduced by searching a small subset of all codeblocks, but at the price of lower performance. In addition, the side distortion of this approach at low redundancies is not satisfactory.

Manuscript received September 10, 2007; accepted December 08, 2008. First published January 27, 2009; current version published April 15, 2009. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Gerald Schuller. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada under Grants RGPIN312262-05, EQPEQ330976-2006, and STPGP350416-07.

The authors are with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6, Canada (e-mail: jingw@sfu.ca; jiel@sfu.ca).

Digital Object Identifier 10.1109/TSP.2009.2013896

In the second MDC category, each description only includes half of the transform outputs. Interdescription prediction is usually used in the side decoder. This splitting approach was first developed in [12]. However, the performance of its side decoder suffers from the inherent prediction residual, especially at high rates [1], [12]. To mitigate this problem, a pairwise correlating transform (PCT) is developed in [13] to introduce a controlled amount of redundancy between two DCT coefficients before splitting. This can reduce the prediction error when one of them is lost. Nevertheless, in [14], it is shown that the PCT at high redundancies is still not satisfactory, and a generalized PCT (GPCT) is proposed. At low redundancies it is the same as the PCT. At high redundancies, in addition to the PCT, each description also encodes the prediction residual of the other half of PCT outputs to reduce the side distortion. Note that by encoding the prediction residual, each description of the GPCT essentially has the same number of coefficients as the transform outputs. Therefore, it becomes a member of the first MDC category.

Some drawbacks still exist in the PCT and GPCT systems. First, the PCT can only be applied to coefficients with large variances relative to the quantization error [13]. Other coefficients are directly split and are estimated as zero in the side decoder, which limits the low redundancy performance of the system. Second, the PCT only uses the correlation it inserts between the two parts of a block, but does not exploit the correlation among neighboring blocks. Finally, the implementation of the system is not easy. In [13], all image blocks are classified into four classes. Coefficient variances of each class need to be calculated and sorted to find the PCT parameters. Also, existing entropy coding cannot be used for the PCT outputs due to different statistics and block sizes. The GPCT further complicates the system, because new transform and entropy coding are needed to code the residual.

In [15] and [16], the lapped orthogonal transform is used to introduce redundancy between neighboring blocks, and the transformed coefficients are split at the block level. In [17], the time-domain lapped transform [18] is used, which simplifies the design and also enables more efficient prediction. The Wiener filter is applied in [19] to improve the estimation of missing blocks. In [20], a prediction-compensated MDC (PC-MDC) framework is developed by encoding the prediction residual in [19]. Superior image coding performance over the MMDSQ, RD-MDC, and PCT is achieved in [20], especially at low redundancies. Note that the PC-MDC in [20] also becomes a special case of the first MDC category due to the prediction compensation.

Although the prediction compensation method achieves good performance in [20] under the block transform framework, this approach has not been well studied in wavelet based MDC. Therefore, one of the motivations of this paper is to develop a prediction-compensated MDC framework for two-band filter banks, denoted as PC-MDFB. Such a generalization is not straightforward, because the tree structure of the wavelet transform poses many challenges to the design of the optimal MD filter bank (MDFB), the design of interdescription prediction, and the partition of the wavelet outputs.

The problem of optimal MDFB design has been studied in [21] and [22] for other MD frameworks. In [21], the output of each subband is used to form one description. The lost description is estimated from the received one using linear prediction. In [22], the pairwise correlating transform in [13] is applied to each frequency in the Fourier domain before splitting. Clearly, both schemes belong to the second MDC category. Since they rely solely on the interdescription prediction in the side decoder, their side distortions are not satisfactory, as in the PCT. In addition, by creating one description from each subband, the two subband filters increasingly resemble each other as the increase of the redundancy, making them quite different from the existing FBs for single description coding (SDC). Therefore, the methods in [21] and [22] cannot be applied satisfactorily to existing wavelet-based systems.

Given that the prediction compensated MDC essentially belongs to the first MDC category, we expect that the optimal MDFBs in our case would be closer to the optimal filter banks in SDC, compared to those in [21] and [22]. The results in this paper verify that our optimal FBs are indeed very similar to their SDC counterparts, although not identical. This is because the filter bank and prediction compensation in our scheme are different from the block transform and MD quantizer in the MDTC. Nevertheless, these results suggest that our method can be applied to systems with single-description-optimized filter banks and attain near-optimal performance.

In Section II, we formulate the design of one-level PC-MDFBs. Section III generalizes the results to multiple levels, which are not considered in [21], [22]. Various design examples are given in Section IV. The application of our method in JPEG 2000 based MD image coding is reported in Section V, which shows that our method achieves better performance than the RD-MDC, MMDSQ, and PCT, especially at low redundancies. Our method also has lower complexity than the RD-MDC.

II. Design of One-Level MD Filter Banks

A. System Overview

Fig. 1 shows the block diagram of generating two descriptions in the proposed prediction-compensated multiple description filter bank framework (PC-MDFB). Only one-level decomposition is considered in this section. Multiple levels are studied in Section III. The input signal $x(n)$ is processed by a two-band perfect reconstruction filter bank. The subband outputs are denoted as $y_0(n)$ and $y_1(n)$, which are further split into even-indexed and odd-indexed parts, denoted as $y_{ij}(n), i, j = 0, 1$. The even (odd) parts of the two subbands are coded together to form the base layer of Description 1 (2), with a bit rate of R_{00} and R_{10} for coefficients from Subband 0 and Subband 1, respectively.

In addition, each description also uses the reconstructed base layer to predict the base layer coefficients of the other description, and the prediction residuals are encoded in each description as an enhancement layer, with bit rate R_{01} and R_{11} for the residuals of Subband 0 and Subband 1, respectively. R_{i1} is usually much lower than R_{i0} for correlated sources.

This scheme creates two balanced descriptions, i.e., the bit rate of each description is

$$R = (R_{00} + R_{01} + R_{10} + R_{11})/4. \tag{1}$$

Unbalanced descriptions can be generated by allowing the two descriptions to have different bit rates.

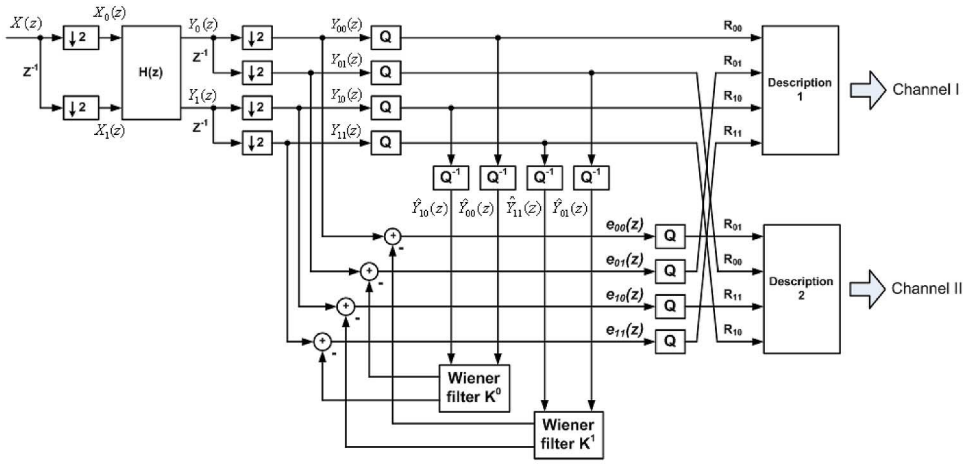


Fig. 1. Encoders of filter bank-based prediction-compensated MDC.

At the decoder, if both descriptions are received, the decoded base layer coefficients from the two descriptions are combined to obtain the reconstructed signal. The enhancement layer in each description is discarded. If only one description is received, the missing coefficients are first estimated from the received base layer coefficients by linear filtering. The decoded prediction residuals are then added to the estimation before applying the synthesis filter bank. Since a strong correlation usually exists between neighboring subband coefficients (especially in low-frequency subbands), good MDC performance can be expected by exploiting this correlation.

Our method represents a different MDFB paradigm from [21] and [22]. First, encoding the prediction residual improves the side decoder performance. Second, the data in each description in [21] and [22] come from only one subband. Therefore, the filter banks in them are drastically different from the existing ones in, e.g., JPEG 2000. In contrast, each description in our method includes transform coefficients from both subbands. Moreover, our data partition is designed such that the prediction between the two descriptions reduces largely to the prediction of neighboring coefficients within each subband. Thus, the two subbands can still be highly uncorrelated. As shown in Section IV, the optimal filter banks in our method are quite similar to their SDC counterparts. Hence, our method can be applied to existing systems such as JPEG 2000 without having to change the filter bank.

A feature-oriented method is developed recently in [23], which is a hybrid of the prediction-only method and the direct coding method. In [23], the wavelet transform coefficients in the other description that can be easily predicted are not encoded in each description, and other coefficients are encoded directly at a lower rate. However, prediction compensation is not used in this method. Also, the method requires side information to be sent to the decoder to identify the coding mode for each coefficient. The bit allocation also involves complicated optimization.

Our method also has various advantages over the PCT and GPCT. First, it can have better performance at very low redundancy, where the prediction compensation can still be applied, especially in the low-frequency subband, as shown by the examples in Section V. Second, the source correlation is fully utilized in our method by more general linear prediction, as detailed in Section II-B and Section II-E, whereas in PCT the prediction is limited between two coefficients in a block. In addition, at high redundancy, the theoretical performance of our scheme is as good as the GPCT, as shown in [20]. Moreover, our scheme can be easily implemented, and existing entropy coding can be used directly.

The optimization of the MDFB in our scheme needs two steps. We first fix the MDFB, and find the closed-form expressions of the corresponding optimal predictor, bit allocation and objective function. An unconstrained numerical optimization program, such as the simplex-based Matlab function *fminsearch*, is then used to search the optimal MDFB that minimizes the objective function among all possible MDFBs, by treating the filter bank parameters as unknown variables. Numerical optimization is needed because the objective function is a complicated nonlinear function of the filter bank parameters. Such an approach has been widely used in filter bank optimizations [24].

B. Optimal Interdescription Prediction

We now derive the optimal interdescription prediction. First, assuming the input signal is a wide-sense stationary (WSS) correlated Gaussian random sequence with a known power spectral density (p.s.d.) $S_X(\omega)$, its 2×2 polyphase p.s.d. matrix can be written as

$$\mathbf{S_{XX}}(\omega)=\begin{bmatrix} S_{X_0X_0}(\omega) & S_{X_0X_1}(\omega) \\ S_{X_1X_0}(\omega) & S_{X_1X_1}(\omega)\end{bmatrix} \tag{2}$$

where $S_{X_iX_j}(\omega)$ is the p.s.d. function between its polyphases $x_i(n)$ and $x_j(n)$ in Fig. 1.

Let the analysis and synthesis filters be [24]

$$\begin{aligned} H_i(z)&=H_{i0}(z^2)+H_{i1}(z^2)z^{-1},\ i=0,1\\ G_i(z)&=G_{i1}(z^2)+G_{i0}(z^2)z^{-1},\ i=0,1. \end{aligned} \tag{3}$$

It can be seen from Fig. 1 that

$$\underbrace{\begin{bmatrix} Y_0(\omega)\\ Y_1(\omega)\end{bmatrix}}_{\mathbf{Y}(\omega)}=\underbrace{\begin{bmatrix} H_{00}(\omega) & H_{01}(\omega)\\ H_{10}(\omega) & H_{11}(\omega)\end{bmatrix}}_{\mathbf{H}(\omega)}\underbrace{\begin{bmatrix} X_0(\omega)\\ X_1(\omega)\end{bmatrix}}_{\mathbf{X}(\omega)}. \tag{4}$$

Similarly, the synthesis filter bank yields

$$\underbrace{\begin{bmatrix}\hat{X}_0(\omega)\\ \hat{X}_1(\omega)\end{bmatrix}}_{\hat{\mathbf{X}}(\omega)} = \underbrace{\begin{bmatrix}G_{00}(\omega) & G_{10}(\omega)\\ G_{01}(\omega) & G_{11}(\omega)\end{bmatrix}}_{\mathbf{G}(\omega)} \underbrace{\begin{bmatrix}\hat{Y}_0(\omega)\\ \hat{Y}_1(\omega)\end{bmatrix}}_{\hat{\mathbf{Y}}(\omega)} \tag{5}$$

where the hat symbol denotes the reconstruction of a signal at the synthesis side, and $\mathbf{G}(\omega)\mathbf{H}(\omega) = \mathbf{I}$. Denote $\mathbf{S_{YY}}(\omega)$ as the p.s.d. matrix of the analysis filter bank output, and $\mathbf{S}_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega)$ and $\mathbf{S}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}(\omega)$ as the p.s.d. matrices of the synthesis filter bank input and output, respectively. We have from (4) and (5)

$$\begin{aligned}\mathbf{S_{YY}}(\omega) &= \mathbf{H}(\omega)\mathbf{S_{XX}}(\omega)\mathbf{H}(\omega)^H,\\ \mathbf{S}_{\hat{\mathbf{X}}\hat{\mathbf{X}}}(\omega) &= \mathbf{G}(\omega)\mathbf{S}_{\hat{\mathbf{Y}}\hat{\mathbf{Y}}}(\omega)\mathbf{G}(\omega)^H.\end{aligned} \tag{6}$$

Given these relationships, the optimal linear prediction for the missing coefficients in Description 1 is the following Wiener filter [21] (due to space limitation, we not always show the parameter ω in the p.s.d. function)

$$\begin{aligned}\mathbf{K}^0(\omega) &= \begin{bmatrix}S_{Y_{01}Y_{00}} & S_{Y_{01}Y_{10}}\\ S_{Y_{11}Y_{00}} & S_{Y_{11}Y_{10}}\end{bmatrix}\begin{bmatrix}S_{Y_{00}Y_{00}} & S_{Y_{00}Y_{10}}\\ S_{Y_{10}Y_{00}} & S_{Y_{10}Y_{10}}\end{bmatrix}^{-1}\\ &\triangleq \begin{bmatrix}K^0_{00}(\omega) & K^0_{01}(\omega)\\ K^0_{10}(\omega) & K^0_{11}(\omega)\end{bmatrix}\end{aligned} \tag{7}$$

where $S_{Y_{ij}Y_{kl}}(\omega)$ is the p.s.d. function between $Y_{ij}(\omega)$ and $Y_{kl}(\omega)$. In this paper, as discussed in [19], [20], and [25], the quantization noise is ignored in the Wiener filter.

To find $S_{Y_{ij}Y_{kl}}(\omega)$, define the autocorrelation function of a WSS sequence $X(n)$ as $R_{XX}(m) = E[X(n+m)X^*(n)]$. It is easy to show that after downsampling $y_0(n)$ and $y_1(n)$ in Fig. 1, we have

$$\begin{aligned}&R_{Y_{ij}Y_{kl}}(m)\\ &\quad= \begin{cases}R_{Y_iY_k}(2m), & (j,l) = (0,0),\ \text{or}\ (1,1),\\ R_{Y_iY_k}(2m+1), & (j,l) = (0,1),\\ R_{Y_iY_k}(2m-1), & (j,l) = (1,0)\end{cases}\end{aligned} \tag{8}$$

where $i,k = 0,1$. Therefore, the $S_{Y_{ij}Y_{kl}}(\omega)$ needed by the Wiener filter can be written as

$$\begin{aligned}S_{Y_{i0}Y_{k0}}(\omega) &= S_{Y_{i1}Y_{k1}}(\omega)\\ &= \frac{1}{2}\left(S_{Y_iY_k}\left(\frac{\omega}{2}\right) + S_{Y_iY_k}\left(\frac{\omega}{2}+\pi\right)\right)\\ S_{Y_{i0}Y_{k1}}(\omega) &= S^*_{Y_{k1}Y_{i0}}(\omega)\\ &= \frac{e^{j\frac{\omega}{2}}}{2}\left(S_{Y_iY_k}\left(\frac{\omega}{2}\right) - S_{Y_iY_k}\left(\frac{\omega}{2}+\pi\right)\right).\end{aligned} \tag{9}$$

C. Various Distortions of the System

Given the expression of interdescription prediction, the various distortions in the scheme can be obtained. In Description 1, $y_{i0}(n)$ is coded at the base layer and $y_{i1}(n)$ is coded in the enhancement layer. Since $y_{i0}(n)$ is WSS Gaussian, its distortion-rate (D-R) function at high rates is [21], [25]

$$D'_{i0} = 2^{-2R_{i0}}\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_{Y_{i0}Y_{i0}}(\omega)d\omega\right) \triangleq 2^{-2R_{i0}}\psi^2_{i0}. \tag{10}$$

Let

$$\mathbf{F}(\omega) = \mathbf{G}^H(\omega)\mathbf{G}(\omega) = \begin{bmatrix}F_{00}(\omega) & F_{01}(\omega)\\ F_{10}(\omega) & F_{11}(\omega)\end{bmatrix} \tag{11}$$

and define $g_i^2 \triangleq (1)/(2\pi)\int_{-\pi}^{\pi}F_{ii}(\omega)d\omega$. For paraunitary filter banks, $\mathbf{F}(\omega) = \mathbf{I}$ in (11). Thus, $F_{ii}(\omega) = g_i^2 = 1$. After synthesis filtering, the reconstruction error per sample caused by the quantization of $y_{i0}(n)$ is [24]

$$D_{i0} = \frac{1}{2}g_i^2 D'_{i0} = \frac{1}{2}g_i^2 2^{-2R_{i0}}\psi^2_{i0} \triangleq 2^{-2R_{i0}}w^2_{i0}. \tag{12}$$

To find the reconstruction error caused by the enhancement layer, we start from the prediction error

$$\begin{bmatrix}e_{01}(\omega)\\ e_{11}(\omega)\end{bmatrix} = \begin{bmatrix}Y_{01}(\omega)\\ Y_{11}(\omega)\end{bmatrix} - \mathbf{K}^0(\omega)\begin{bmatrix}\hat{Y}_{00}(\omega)\\ \hat{Y}_{10}(\omega)\end{bmatrix} \tag{13}$$

from which we can get the p.s.d. of the prediction error

$$S_{e_{i1}}(\omega) = S_{Y_{i1}Y_{i1}} - K^0_{i0}S_{Y_{00}Y_{i1}} - K^0_{i1}S_{Y_{10}Y_{i1}}. \tag{14}$$

Since $y_{i1}(n)$ is predictively coded, its quantization error equals that of the residual $e_{i1}(n)$ [25]. Therefore, the reconstruction error per sample after synthesis filtering is

$$\begin{aligned}D_{i1} &= \frac{1}{2}g_i^2 2^{-2R_{i1}}\exp\left(\frac{1}{2\pi}\int_{-\pi}^{\pi}\log S_{e_{i1}}(\omega)d\omega\right)\\ &\triangleq \frac{1}{2}g_i^2 2^{-2R_{i1}}\psi^2_{i1} \triangleq 2^{-2R_{i1}}w^2_{i1}.\end{aligned} \tag{15}$$

In Description 2, $y_{i1}(n)$ is coded in the base layer, while $y_{i0}(n)$ is coded in the enhancement layer. The average distortions for Description 2 and 1 are equal since our method generates balanced descriptions.

When both descriptions are received, only the base layer coefficients are used in the reconstruction, and the corresponding average distortion per sample (*the central distortion*) is

$$D_0 = \left(g_0^2 D'_{00} + g_1^2 D'_{10}\right)/2 = D_{00} + D_{10}. \tag{16}$$

If only one description is available, half of $y_0(n)$ and $y_1(n)$ are in the base layer and the rest are in the enhancement layer. The average distortion (*the side distortion*) is

$$D_1 = (D_{00} + D_{01} + D_{10} + D_{11})/2. \tag{17}$$

Let p be the probability of losing one description, the expected distortion is thus

$$D = p^2\sigma_x^2 + 2p(1-p)D_1 + (1-p)^2 D_0. \tag{18}$$

In [21], a *single-channel distortion* is also defined by ignoring the base layer error in D_1, i.e.,

$$D_s = (D_{01} + D_{11})/2. \tag{19}$$

D. Optimal Bit Allocation and Distortions

We now derive the optimal bit allocation and distortions for a given filter bank. Two approaches can be used: minimizing the expected distortion under a bit rate constraint, or minimizing the single-channel distortion under a redundancy constraint. At high rates, the two approaches lead to the same optimal solution. In this paper, both methods are adopted to facilitate comparison with different schemes.

1) Approach I: In this method, we minimize the expected distortion D, subject to the bit rate constraint (1). Using Lagrangian method with high rate assumption, we get

$$R_{i0} = R + \frac{1}{4}\log_2\left(\frac{w_{i0}^3}{p\prod_{k,l=0,1,(k,l)\neq(i,0)} w_{kl}}\right), \quad i = 0,1$$
$$R_{i1} = R + \frac{1}{4}\log_2\left(\frac{p\, w_{i1}^3}{\prod_{k,l=0,1,(k,l)\neq(i,1)} w_{kl}}\right), \quad i = 0,1. \tag{20}$$

$$D_{00} = D_{10} = (p\, w_{00}w_{01}w_{10}w_{11})^{\frac{1}{2}} 2^{-2R}$$
$$D_{01} = D_{11} = \left(\frac{1}{p} w_{00}w_{01}w_{10}w_{11}\right)^{\frac{1}{2}} 2^{-2R} \tag{21}$$
$$D = p^2\sigma_x^2 + 4(1-p)\sqrt{p\, w_{00}w_{01}w_{10}w_{11}}\, 2^{-2R}. \tag{22}$$

2) Approach II: In the SDC of correlated Gaussian sources, the minimum required bit rate to achieve the distortion D^* is given by [25]

$$R^* = \frac{1}{2}\log_2\frac{w^2}{D^*} \tag{23}$$

where $w^2 = \exp((1)/(2\pi)\int_{-\pi}^{\pi}\log S_X(\omega)d\omega)$. To get the same central distortion, i.e., $D_0 = D_{00} + D_{10} = D^*$, a MDC scheme needs more bits. The redundancy is $\rho = (R_{00} + R_{10} + R_{01} + R_{11})/2 - R^*$. This leads to the optimization approach that minimizes D_s in (19) for given ρ and D_0. The corresponding Lagrangian objective function is $J = D_s + \lambda_1\rho + \lambda_2 D_0$.

At high rate, for a fixed D_0, the bit allocation for R_{00} and R_{10} is independent of that of R_{01} and R_{11}. Therefore, the optimization can be decoupled into two steps. The first yields

$$D_{00} = D_{10} = D^*/2 \tag{24}$$
$$R_{i0} = \frac{1}{2}\log_2\frac{2w_{i0}^2}{D^*},\ i = 0,1. \tag{25}$$

The redundancy becomes

$$\rho = \left(\frac{R_{00}+R_{10}}{2} - R^*\right) + \frac{R_{01}+R_{11}}{2} = \frac{1}{2}\log_2\frac{2w_{00}w_{10}}{w^2} + \frac{R_{01}+R_{11}}{2} \tag{26}$$

and the objective function reduces to $J = (1)/(2)(D_{01}+D_{11}) + \lambda\rho$. As in [21], we assume high redundancy, and the solution to this part can be found to be

$$R_{i1} = \rho + \frac{1}{2}\log_2\frac{w_{i1}w^2}{2\prod_{k,l=0,1,(k,l)\neq(i,1)} w_{kl}} \tag{27}$$
$$D_{01} = D_{11} = \frac{2w_{00}w_{01}w_{10}w_{11}}{w^2}2^{-2\rho}. \tag{28}$$

The optimal single-channel distortion is

$$D_s = \frac{1}{2}(D_{01} + D_{11}) = \frac{2w_{00}w_{01}w_{10}w_{11}}{w^2}2^{-2\rho}. \tag{29}$$

Another measure of redundancy is the redundancy ratio

$$\rho_r = (R_{01} + R_{11})/(R_{00} + R_{10}). \tag{30}$$

From (1) and (26), we can get the following relationship between ρ and ρ_r:

$$\rho = \frac{1}{2}\log_2\frac{2w_{00}w_{10}}{w^2} + \frac{2\rho_r}{1+\rho_r}R. \tag{31}$$

As shown by (22) and (29), both approaches need to minimize $w_{00}w_{01}w_{10}w_{11}$, which depends only on the filter bank and the input. Therefore, they lead to the same optimal filter bank. In addition, the optimal solution remains unchanged for different channel loss probabilities and redundancies, under the assumptions of high rate and high redundancy.

E. Simplified Interdescription Predictions

The ideal interdescription predictor in (7) is given in frequency domain, and it uses all available data to predict each coefficient in the other description. Two simplifications can be used. The first one predicts a missing coefficient using only the data available in the same subband. The Wiener filter in (7) is then simplified to be

$$\mathbf{K}^0(\omega) = \begin{bmatrix} S_{Y_{01}Y_{00}} & 0 \\ 0 & S_{Y_{11}Y_{10}} \end{bmatrix}\begin{bmatrix} S_{Y_{00}Y_{00}} & 0 \\ 0 & S_{Y_{10}Y_{10}} \end{bmatrix}^{-1} \triangleq \begin{bmatrix} K_{00}^0(\omega) & 0 \\ 0 & K_{11}^0(\omega) \end{bmatrix}. \tag{32}$$

In the second simplification, the FIR Wiener filter can be applied to the FB output coefficients directly, which is more suitable for practical applications. In this case, let $\mathbf{y}_{B_i}(n) = [y_{0i}(n)\, y_{1i}(n)]^T$. Define the $2L$ blocks of neighboring $\mathbf{y}_{B_i}(n)$'s as

$$\mathbf{y}_{B_{i,L}}(n) = \left[\mathbf{y}_{B_i}^T(n-L+1)\ldots\mathbf{y}_{B_i}^T(n+L)\right]^T. \tag{33}$$

In Description 1, we use the reconstructed $\mathbf{y}_{B_{0,L}}(n)$ to predict $\mathbf{y}_{B_1}(n)$. The optimal Wiener filter is

$$\mathbf{K}_L^0 = \mathbf{R}_{\mathbf{y}_{B_1}\mathbf{y}_{B_{0,L}}}\mathbf{R}^{-1}_{\mathbf{y}_{B_{0,L}}\mathbf{y}_{B_{0,L}}} \tag{34}$$

where $\mathbf{R}_{\mathbf{y}_{B_{0,L}}\mathbf{y}_{B_{0,L}}}$ is the autocorrelation of $\mathbf{y}_{B_{0,L}}(n)$, and $\mathbf{R}_{\mathbf{y}_{B_1}\mathbf{y}_{B_{0,L}}}$ is the cross-correlation between $\mathbf{y}_{B_1}(n)$ and

$\mathbf{y}_{B_{0,L}}(n)$. The corresponding autocorrelation of the residual for $\mathbf{y}_{B_1}(n)$ is

$$\mathbf{R}_{e_{B_1}} = \mathbf{R}_{\mathbf{y}_{B_1}\mathbf{y}_{B_1}} - \mathbf{K}_L^0 \mathbf{R}_{\mathbf{y}_{B_{0,L}}\mathbf{y}_{B_1}} \triangleq \begin{bmatrix} \sigma_{01}^2 & * \\ * & \sigma_{11}^2 \end{bmatrix}. \quad (35)$$

All matrices involved above can be expressed in terms of the input correlation matrix and the FIR analysis filter bank. Finally, let

$$\mathbf{R}_{\mathbf{y}_{B0}\mathbf{y}_{B0}} = \mathbf{R}_{\mathbf{y}_{B1}\mathbf{y}_{B1}} \triangleq \begin{bmatrix} \sigma_{00}^2 & * \\ * & \sigma_{10}^2 \end{bmatrix}. \quad (36)$$

With these expressions of σ_{ij}^2, the various distortions needed by (16) and (17) can be found to be [24]

$$D_{ij} = c \cdot 2^{-2R_{ij}} g_i^2 \sigma_{ij}^2 \triangleq 2^{-2R_{ij}} w_{ij}^2 \quad (37)$$

where g_i^2 is given in (Section II-C). For Gaussian inputs, c is a constant that depends on the input statistics and quantization scheme. The value of c does not affect the optimized FB and it is chosen as 1 here. Equation (37) has the same format as (12) and (15). Therefore, the optimization in Section II-D is still applicable.

III. Design of Multilevel MD Filter Banks

In this section, we study the optimal filter bank for the proposed MDC when M-level FB decompositions are used. In this case, there are $M+1$ subband signals y_i, $i = 0, \ldots, M$. By the noble identity [24], the $M+1$ equivalent subband filters are obtained as follows:

$$\bar{H}_k(z) = \begin{cases} H_1\left(z^{2^k}\right) \prod_{n=0}^{k-1} H_0\left(z^{2^n}\right), & k = 0, \ldots, M-1 \\ \prod_{n=0}^{k-1} H_0\left(z^{2^n}\right), & k = M. \end{cases} \quad (38)$$

$$\bar{G}_k(z) = \begin{cases} G_1\left(z^{2^k}\right) \prod_{n=0}^{k-1} G_0\left(z^{2^n}\right), & k = 0, \ldots, M-1 \\ \prod_{n=0}^{k-1} G_0\left(z^{2^n}\right), & k = M. \end{cases} \quad (39)$$

To get two descriptions, we split each subband into the even-indexed part $y_{i0}(n)$ and the odd-indexed part $y_{i1}(n)$, $i = 0, \ldots, M$. $y_{i0}(n)\,(y_{i1}(n))$ is coded in the base layer in Description 1 (2) with bit rate R_{i0}. The prediction residual of $y_{i1}(n)\,(y_{i0}(n))$ is also encoded as the enhancement layer in Description 1 (2) with bit rate R_{i1}, where each $y_{i1}(n)$ is predicted using all reconstructed base layer coefficients $\hat{y}_{i0}(n)$'s. Similarly, $y_{i0}(n)$ is predicted using all $\hat{y}_{i1}(n)$'s.

The Wiener filter in Description 1 is given by

$$\mathbf{K}^0(\omega) = \begin{bmatrix} S_{Y_{01}Y_{00}}(\omega) & \cdots & S_{Y_{01}Y_{M,0}}(\omega) \\ \vdots & \ddots & \vdots \\ S_{Y_{M,1}Y_{00}}(\omega) & \cdots & S_{Y_{M,1}Y_{M,0}}(\omega) \end{bmatrix} \times \begin{bmatrix} S_{Y_{00}Y_{00}}(\omega) & \cdots & S_{Y_{00}Y_{M,0}}(\omega) \\ \vdots & \ddots & \vdots \\ S_{Y_{M,0}Y_{00}}(\omega) & \cdots & S_{Y_{M,0}Y_{M,0}}(\omega) \end{bmatrix}^{-1} \quad (40)$$

where $S_{Y_{ij}Y_{kl}}(\omega)$ can be derived from the input statistics by recursively computing the p.s.d. of a subband using (6) and (9). The Wiener filter for Description 2 can be obtained similarly.

The p.s.d. of the prediction errors of $y_{i1}(n)$ is

$$S_{e_{i1}}(\omega) = S_{Y_{i1}Y_{i1}}(\omega) - \mathbf{K}_i^0 [S_{Y_{00}Y_{i1}}(\omega), \ldots, S_{Y_{M,0}Y_{i1}}(\omega)]^T \quad (41)$$

where $\mathbf{K}_i^0$ is the ith row of $\mathbf{K}^0(\omega)$. Define $\bar{g}_i^2 \triangleq 1/(2\pi) \times \int_{-\pi}^{\pi} |\bar{G}_i(\omega)|^2 d\omega$. As in Section II-B, we can derive the D-R function of each part of Description 1, which is the same as that in Description 2 since we have balanced descriptions:

$$D_{ij} = w_{ij}^2 2^{-2R_{ij}},\ i = 0, \ldots, M,\ j = 0, 1 \quad (42)$$

where D_{ij} is the reconstruction error per sample contributed by the quantization of $y_{ij}(n)$ or $e_{ij}(n)$, and

$$w_{ij}^2 = \begin{cases} \frac{1}{2^{(i+1)}} \bar{g}_i^2 \psi_{ij}^2, & i = 0, \ldots, M-1 \\ \frac{1}{2^i} \bar{g}_i^2 \psi_{ij}^2, & i = M. \end{cases} \quad (43)$$

where ψ_{ij}^2 are defined in (10) and (12). The central distortion and side distortion are

$$D_0 = \sum_{i=0}^{M} D_{i0}, D_1 = \frac{1}{2} \sum_{i=0}^{M} (D_{i0} + D_{i1}). \quad (44)$$

Here, we use the criterion of minimizing the expected distortion D in (18) under the bit rate constraint (1). Using Lagrangian method, the optimal bit allocation is found to be

$$R_{i0} = R + \frac{1}{4(M+1)} \log_2 \frac{w_{i0}^{4(M+1)}}{p^{(M+1)} \prod_{j=0}^{M} w_{j0}^2 w_{j1}^2}$$
$$R_{i1} = R + \frac{1}{4(M+1)} \log_2 \frac{p^{(M+1)} w_{i1}^{4(M+1)}}{\prod_{j=0}^{M} w_{j0}^2 w_{j1}^2}. \quad (45)$$

The corresponding distortions are

$$D_{i0} = \left(p^{(M+1)} \prod_{j=0}^{M} w_{j0}^2 w_{j1}^2 \right)^{\frac{1}{2(M+1)}} 2^{-2R}$$
$$D_{i1} = \left(p^{-(M+1)} \prod_{j=0}^{M} w_{j0}^2 w_{j1}^2 \right)^{\frac{1}{2(M+1)}} 2^{-2R}. \quad (46)$$

Given the probability p and rate R, the minimal expected distortion D becomes

$$D = p^2 \sigma_x^2 + 2(1-p)(M+1) \sqrt{p w_0 w_1} 2^{-2R} \quad (47)$$

where w_0 and w_1 are geometric means of all w_{j0}^2 and w_{j1}^2, respectively. For $M = 1$, this reduces to (22).

IV. Design Examples

In this section, we design various optimal PC-MDFBs for a first-order autoregressive [AR(1)] sequence with correlation coefficient $r = 0.95$. The bit rate is chosen to be $R = 2$

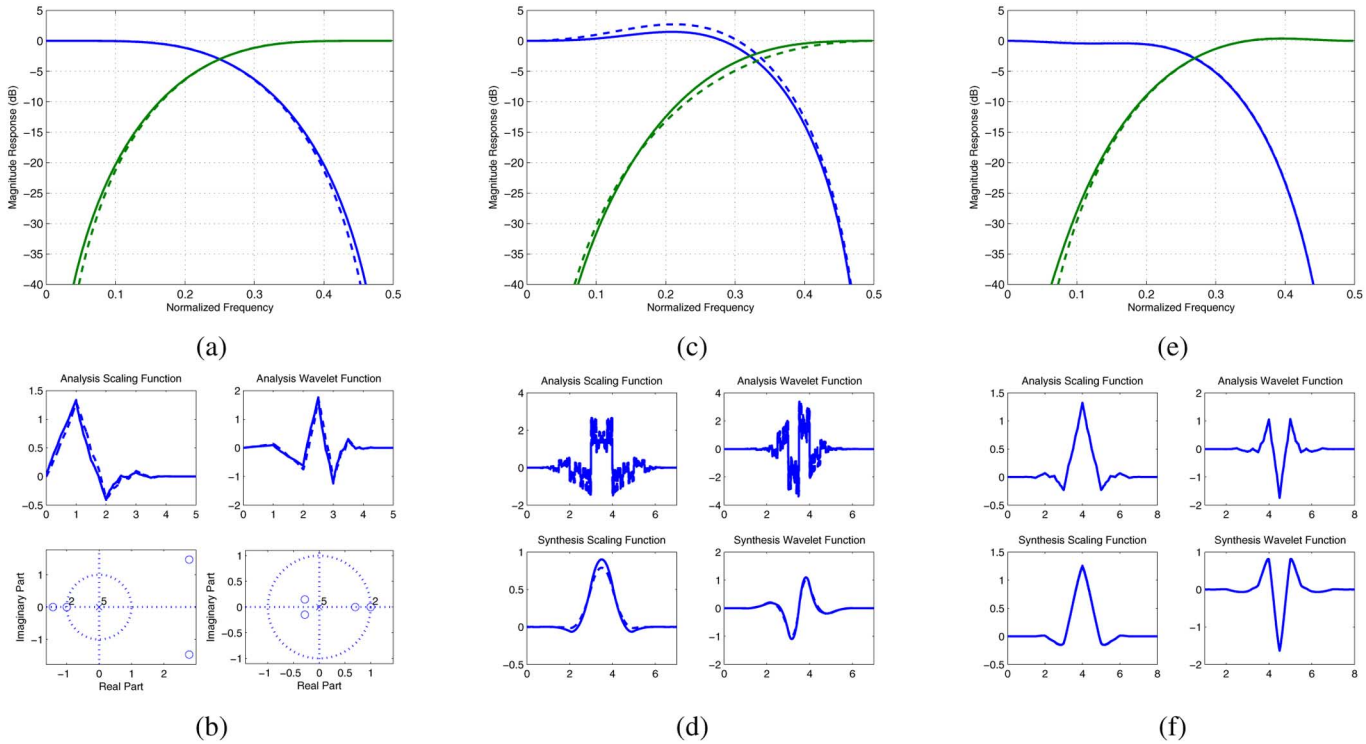


Fig. 2. (a)–(b) Example 1: 6-tap PUFB for PC-MDFB (solid lines) and D6 wavelet (dashed lines). (a) Frequency responses. (b) Top: analysis scaling and wavelet functions. Bottom: zero and pole distributions of H_0 and H_1. (c)–(d) Example 2: 8-tap LPPRFB for PC-MDFB (solid) and SDC (dashed). (c) Frequency responses. (d) Scaling and wavelet functions. (e)–(f) Example 3: 9/7-tap PC-MDFB (solid) and JPEG 2000 9/7 wavelet (dashed). (e) Frequency responses. (f) Scaling and wavelet functions.

bits/sample/description. As described in Section II-A, a numerical optimization routine is used to search the optimal filter bank that minimizes the objective function.

The p.s.d. function of AR(1) signals with correlation coefficient r is [25]

$$S_X(\omega) = \frac{\sigma_x^2(1-r^2)}{1-2r\cos(\omega)+r^2}. \tag{48}$$

The downsampled signal is still an AR(1) signal with correlation r^2, so its p.s.d. function is given by

$$S_{X_0X_0}(\omega) = S_{X_1X_1}(\omega) = \frac{\sigma_x^2(1-r^4)}{1-2r^2\cos(\omega)+r^4}. \tag{49}$$

Let $\alpha = r/(1+r^2)$, the 2×2 polyphase p.s.d. matrix of the input sequence can be shown to be

$$S_{\mathbf{XX}}(\omega) = \begin{bmatrix} 1 & \alpha(1+e^{-j\omega}) \\ \alpha(1+e^{j\omega}) & 1 \end{bmatrix} \times \frac{\sigma_x^2(1-r^4)}{1-2r^2cos(\omega)+r^4}. \tag{50}$$

A. One-Level Filter Banks

1) Two-Band 6-Tap Paraunitary Filter Banks: In this example, we use the lattice structure of two-band paraunitary filter banks (PUFB) in [24]. In this case, $\mathbf{G}(\omega) = \mathbf{H}^H(\omega)$ and

$$\mathbf{H}(\omega) = \mathbf{R}(\theta_n)\mathbf{\Lambda}(\omega)\mathbf{R}(\theta_{n-1})\mathbf{\Lambda}(\omega)\cdots\mathbf{\Lambda}(\omega)\mathbf{R}(\theta_0) \tag{51}$$

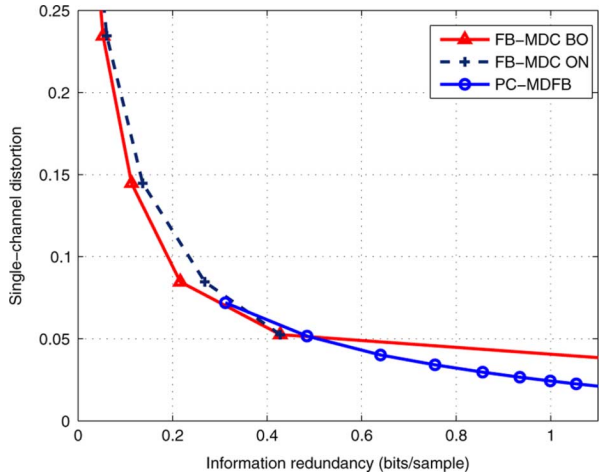


Fig. 3. Redundancy (ρ) versus single-channel distortion (D_s) curves of PC-MDFB and the two solutions in [21].

where

$$\mathbf{R}(\theta_i) = \begin{bmatrix} \cos\theta_i & \sin\theta_i \\ -\sin\theta_i & \cos\theta_i \end{bmatrix}$$

and

$$\mathbf{\Lambda}(\omega) = \begin{bmatrix} 1 & 0 \\ 0 & e^{-j\omega} \end{bmatrix}.$$

The frequency domain Wiener filter (7) is used. To enhance the smoothness of reconstructed signals, two vanishing moments [25] are imposed on the filter bank. The frequency responses and scaling/wavelet functions of the optimized 6-tap PUFB for

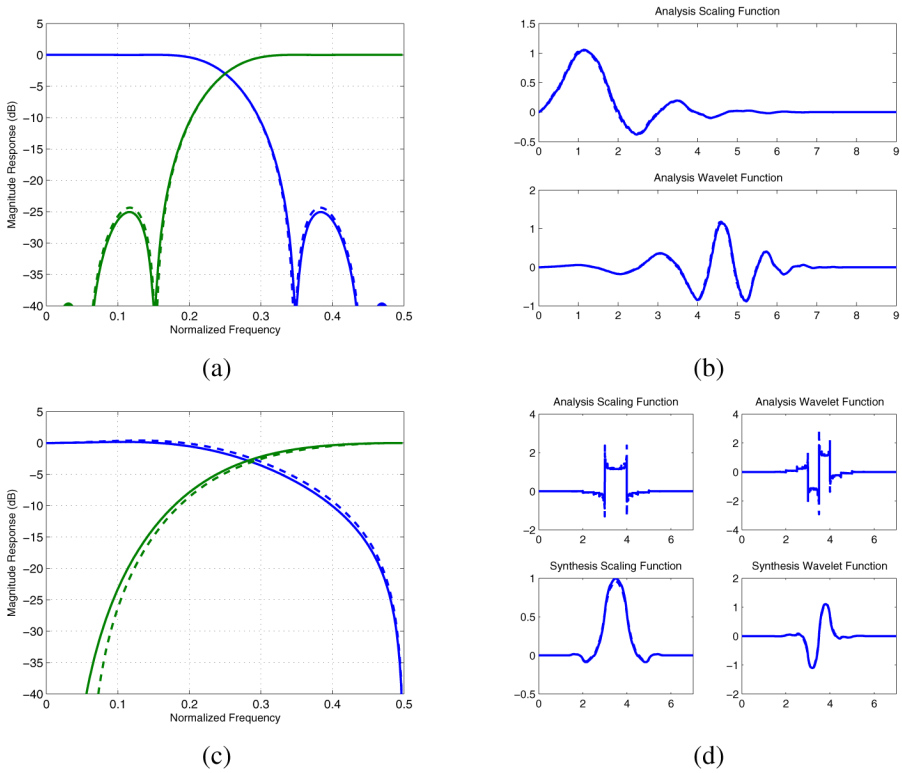


Fig. 4. (a)–(b) Optimal 10-tap, five-level PUFB for PC-MDFB (solid lines) and SDC (dashed lines). (a) Frequency responses. (b) Scaling and wavelet functions. (c)–(d) The optimal 8-tap, four-level LPPRFB for PC-MDFB (solid) and SDC (dashed). (c) Frequency responses. (d) Scaling and wavelet functions.

the PC-MDFB and Daubechies D6 wavelet are compared in Fig. 2(a)–(b). The coding gain (a measure of the compression efficiency of the filter bank [24]) of the 6-tap PUFB for PC-MDFB is 5.7367 dB, which is very close to the 5.7505 dB of the D6 wavelet.

2) Two-Band 8-Tap Linear Phase Perfect Reconstruction Filter Banks: In this biorthogonal FB example, we design the optimal linear phase perfect reconstruction filter bank (LPPRFB) using the type-II lattice in [26] for the PC-MDFB framework. The filter bank has one vanishing moment in the analysis filter bank and two vanishing moments in the synthesis filter bank. The FIR Wiener filter (34) is used with $L = 1$. The frequency responses and scaling/wavelet functions of the optimal 8-tap LPPRFBs for PC-MDFB and SDC are compared in Fig. 2(c)–(d). The coding gain of the optimal LPPRFB for PC-MDFB is 6.24 dB, and the coding gain of the optimal LPPRFB for SDC is 6.39 dB.

3) Two-Band 9/7-Tap Biorthogonal Filter Banks: This example uses the lifting structure of two-band biorthogonal filter banks with four lifting steps such that the analysis low-pass and high-pass filters have nine taps and seven taps, respectively, as in the 9/7 wavelet in JPEG 2000 [25]. The FIR Wiener filter (34) with $L = 1$ is used. The frequency responses and scaling/wavelet functions are given in Fig. 2(e)–(f). The coding gain of our 9/7-tap PC-MDFB is 5.913 dB, which is also close to the 5.916 dB of the 9/7 wavelet in JPEG 2000.

4) Comparison With [21]: To further study the performance of the proposed method, we redesign the two-band, 8-tap LPPRFB in example II using the general Wiener filter in (7) for AR(1) sources with correlation coefficient $r = 0.9$, and compare it with the optimal IIR orthogonal and biorthogonal solutions in Fig. 9 and Table II of [21], denoted as FB-MDC ON and FB-MDC BO, respectively. The comparison is fair since the same R-D function and input signal model are used. The results are shown in Fig. 3. It can be seen that our result has similar single-channel distortion D_s at low redundancy while achieves lower D_s at high redundancy, despite the FIR and linear phase constraints in our method. As shown in [21], the D_s of their method cannot be arbitrarily reduced due to the presence of the prediction error. For example, the smallest D_s of the orthogonal solution in [21] is 0.0525, with a redundancy of 0.2967 nats/sample. The smallest D_s of the biorthogonal solution in [21] is 0.0374, but it can only be achieved at infinite redundancy [21]. This problem is resolved in our method by encoding the prediction residual. Finally, although our method has similar performance to that in [21] at low redundancy for this strongly correlated signal, the coding of the prediction residual in our method is more helpful for nonstationary signals, as shown in the image coding results in Section V.

B. Multiple-Level Filter Banks

Two design examples of the optimal M-level orthogonal and biorthogonal filter banks for AR(1) signals with the FIR Wiener filter are given in Fig. 4. The PC-MDFB in Fig. 4(a)–(b) is a 10-tap PUFB optimized for five-level decomposition and the one in Fig. 4(c)–(d) is a 8-tap LPPRFB optimized for four-level decomposition. Each has one vanishing moment. As in the one-level case, the optimal filter banks for the proposed MDC are still very close to the optimal SDC filter banks. This suggests that the proposed method can be applied to practical wavelet based image coding systems such as the JPEG 2000 without having to change the wavelet transform.

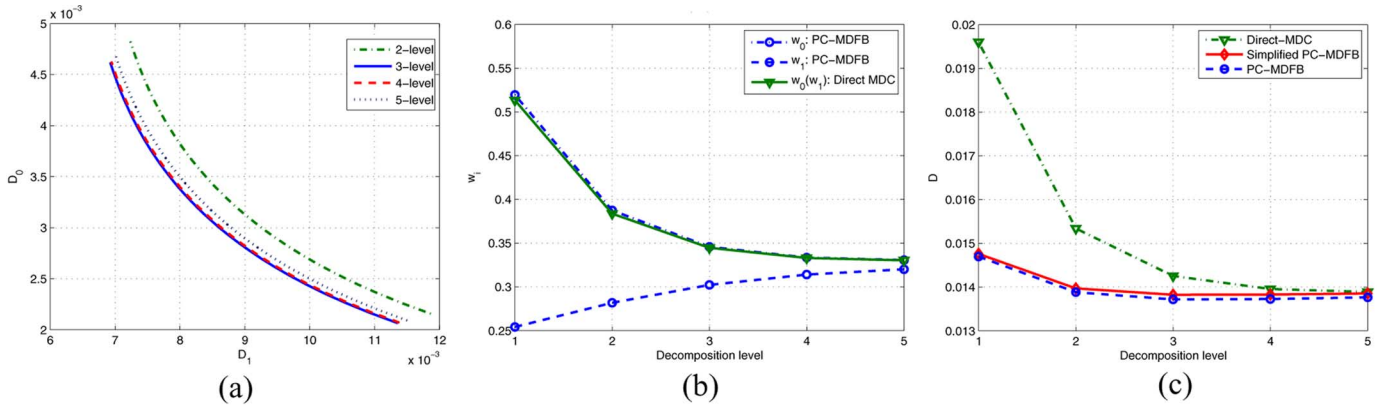


Fig. 5. (a) D_0 versus D_1 curves of the 6-tap PUFB for PC-MDFB with multiple levels. (b) w_0 and w_1 curves. (c) Expected distortions when $R = 2$ bits/sample/description and $p = 0.1$.

To gain further insights, we compare the performance of a 6-tap PUFB for the PC-MDFB with different levels in Fig. 5(a). An important observation is that the performance does not always increase with the levels of decomposition. In this example, the optimal performance is achieved by three levels. Similar behaviors can also be observed using optimized biorthogonal FBs and FBs from SDC.

To explain this, we start from (47), which shows that the optimized expected distortion increases with $w_0 w_1$. The w_0 and w_1 curves of our method are given in Fig. 5(b), which shows that as the number of decomposition levels M increases, the w_0 of our method decreases due to improved decorrelation, whereas w_1 increases, since the reduced correlation also decreases the prediction efficiency. Therefore, the product $w_0 w_1$ does not necessarily decrease as the increase of M.

In Fig. 5(b), our method is also compared with the direct MDC, where each subband is split into even/odd-indexed coefficients. The even (odd) part is coded at high rate in Description 1 (2). A low rate version of the odd (even) part is also included in Description 1 (2). This method can be viewed as an application of the methods in [9]–[11] to our framework. The w_0 curve of the direct MDC is slightly better than our method, but its w_1 curve is much worse, leading to worse overall performance. This illustrates the tradeoff between the filter bank coding performance and the prediction efficiency.

Fig. 5(c) shows the expected distortions for an AR(1) source with different levels. The direct MDC, the PC-MDFB, and a simplified PC-MDFB are compared. In the simplified PC-MDFB, a suboptimal Wiener filter is used, where only reconstructed base layer coefficients of the same subband are used to predict a lost coefficient, similar to (32). The figure shows that our method has significant advantage over the direct MDC when the decomposition level is from 1 to 3. Above that the difference becomes very small. Note that this result is only for AR(1) sequences. For nonstationary signals such as natural images, the advantages of our method will be more pronounced, as verified in the next section.

V. Applications in Image Coding

In this section, we apply the proposed prediction-compensated multiple description coding scheme to JPEG 2000 without changing its wavelet transform. The JPEG 2000 codec OpenJPEG [27] is used. Our source codes are available at [28]. We compare our method with the RD-MDC in [11], whose codes are at [29]. The comparison is fair, since both methods use the same OpenJPEG codec. We also include as reference the results of the MMDSQ in [5], which is another state-of-the-art method, but it uses different quantization and entropy coding from the JPEG 2000.

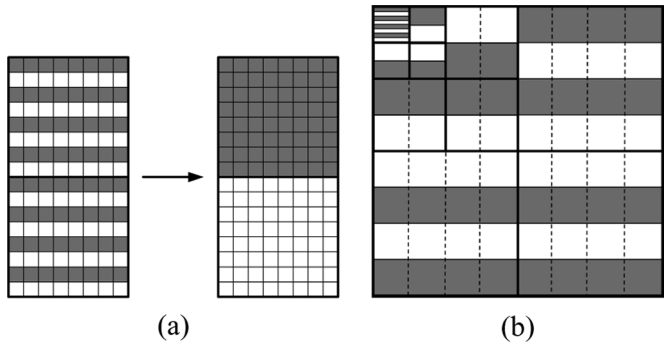


Fig. 6. Subband partitions: (a) LL subband and (b) Other subbands.

In source splitting based MDCs with block transforms, there are mainly two ways of splitting and prediction, i.e., coefficient-level or block-level approach, as in [13] and [20], respectively, and the latter has been shown to give better results [20]. However, the splitting and interdescription prediction in the wavelet framework is more challenging, because the tree structure enables a lot more choices, and it is not clear which one is the optimal in practical applications. In our preliminary method in [30], the HL, LH, and HH subbands are encoded via a scaling based prediction in a parent-children subtree, but such a simple linear prediction is not effective enough, and adaptive prediction is needed to take full advantage of the parent–children dependency, as in SPIHT.

In this paper, we develop another low complexity scheme, where the prediction compensation is only applied to the LL subband. Image coding results show that this is sufficient to yield favorable performance over other methods, because the prediction residual in the LL subband can benefit the entire image. Since the basic unit of the JPEG 2000 entropy coding is a four-row stripe [25], the LL subband is split into even and odd-indexed rows. The two parts are then grouped together, partitioned into JPEG 2000 codeblocks, and encoded into the base layer of each description, as shown in Fig. 6(a). To generate the enhancement layer, a simple linear prediction is used between the two parts, where each coefficient is predicted by the average

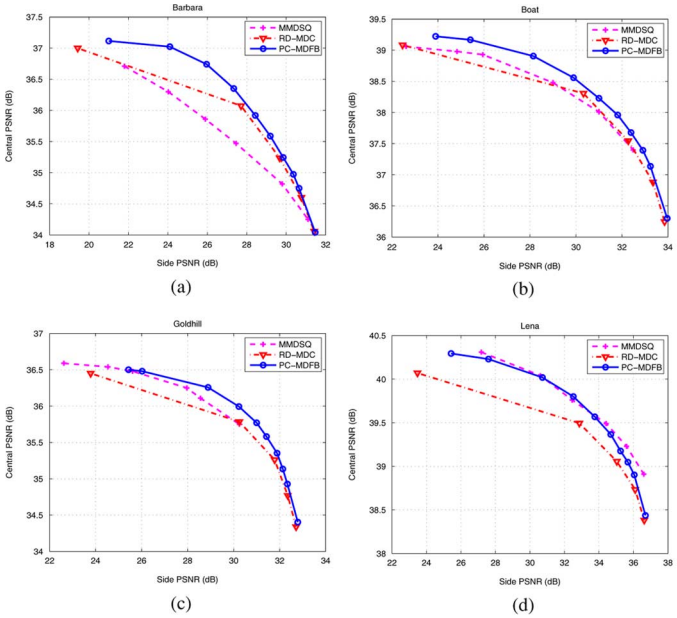


Fig. 7. Performances of PC-MDFB, RD-MDC, and MMDSQ at $\mathrm{R} = 0.5$ bit/pixel/description. (a) Barbara. (b) Boat. (c) Goldhill. (d) Lena.

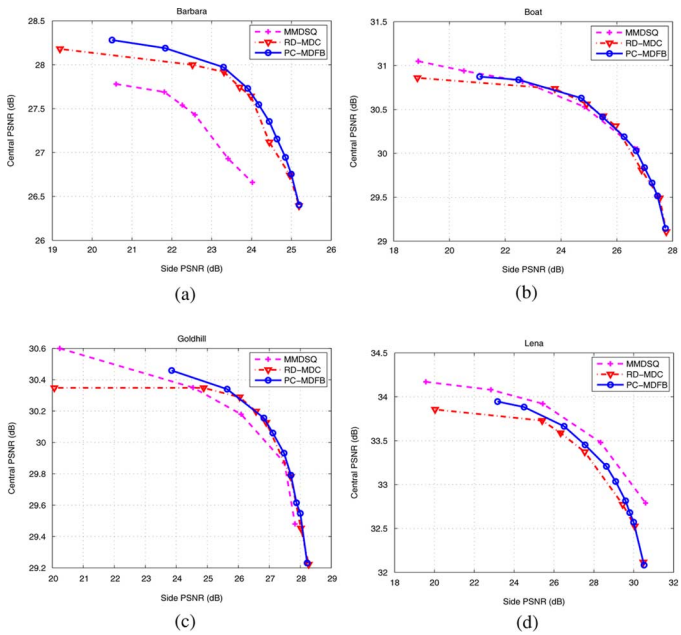


Fig. 8. Performances of PC-MDFB, RD-MDC and MMDSQ at $\mathrm{R} = 0.125$ bit/pixel/description. (a) Barbara. (b) Boat. (c) Goldhill. (d) Lena.

of the two nearest reconstructed base layer coefficients, and the prediction residual is encoded in the enhancement layer. Such a partition provides a good tradeoff between the prediction performance and the adverse impact to the JPEG 2000 entropy coding.

Since the correlations between a pair of parent–child subbands and within a high-frequency subband are not as strong as in the LL subband, no prediction is used in the HL, LH and HH subbands, i.e., each codeblock in them is coded at two rates directly, one for each description. Different from the RD-MDC approach that dynamically determines the codeblocks and their rates for a description, we use a deterministic alternating partition as illustrated in Fig. 6(b), where codeblocks in white slices are coded at high rate in one description, and low rate in another description, and *vice versa* for gray slices. Moreover, if a slice is coded in base (enhancement) layer, its immediate children slice in the next subband will be coded in enhancement (base) layer. This partition generates two descriptions with roughly the same amount of information about each region of the image. Balanced descriptions are then ensured by the JPEG 2000 rate control. The slice height in each subband is chosen to be the same as the codeblock size of each subband. If the subband size is smaller than the predetermined codeblock size, the subband is split into two horizontal slices and each of them forms a small codeblock. Although this scheme does not have the inter-subband prediction in [30], the partition in Fig. 6(b) is more friendly to the JPEG 2000 entropy coding.

The rate control in our method is as follows. Given the bit rate and redundancy ratio, the corresponding base layer and enhancement layer bit rates are first calculated. The JPEG 2000 rate control scheme is then used to encode all base layer codeblocks at the target rate, and then the reconstructed base layer coefficients in LL subband are used to encode the residual codeblocks at its target bit rate. The codeblocks in other subbands are encoded directly into the corresponding enhancement layer. Since we use symmetric splitting with predefined patterns, balanced descriptions are achieved with very low complexity.

Compare with the RD-MDC, our method has lower complexity. The only extra work in our method is to partition and produce the necessary base layer codeblocks and residual codeblocks. All modules of JPEG 2000 can be reused, and good performance and balanced descriptions are easily achieved. As a result, our encoder only needs 0.70 s to generate two descriptions from an 512×512 image on a PC with 2.13-GHz Intel Core 2 Duo CPU and 2-GB memory. The MMDSQ also takes less than one second. In contrast, to find the optimal partition, the RD-MDC needs to exhaustively search all possible codeblock combinations. A suboptimal algorithm is suggested in [11] to reduce the complexity by searching a smaller group of codeblocks, but the performance is reduced as the decrease of the group size. A similar method is proposed in [31], where the search is constrained over sets with periodically repeating codeblock partition pattern. With a codeblock group size of 25, the suboptimal RD-MDC in [11] needs 34 s to create two descriptions. When the group size is 20, it still takes 2.85 s, which is four times of our method. In this paper, the results of RD-MDC are obtained using the group size 25.

The tradeoffs between the central PSNR (D_0) and the average side PSNR (D_1) of the three methods at rate $\mathrm{R} = 0.5$ and 0.125 bits per pixel (bpp) per description are given in Figs. 7 and 8, respectively. Four 512×512 test images with different characteristics are used. The average MSE of the two side decoder output is used to calculate the average side PSNR in the figures. Four levels of wavelet transform are used [i.e., $M = 4$ in (38)], and the codeblock size is 64×64. At low redundancies, our method achieves better results than the RD-MDC and the MMDSQ in most cases, where the enhancement layer of our method still contains the most important edge information of the image, which improves the side decoder PSNR and visual quality. As the increase of redundancy, the three methods become very close to each others.

Fig. 9 shows the side decoder result of our PC-MDFB method and the RD-MDC. The two methods are compared at the same

Fig. 9. Results of one side PSNR D_1 with $R = 0.5$ bpp per description. The other side PSNR D_2 and the central PSNR D_0 are listed in the parentheses. (a) Goldhill by RD-MDC: 30.25 dB ($D_2 = 31.00$ dB, $D_0 = 35.78$ dB); (b) Goldhill by PC-MDFB: 31.26 dB ($D_2 = 30.81$ dB, $D_0 = 35.77$ dB); (c) Lena by RD-MDC: 33.15 dB ($D_2 = 33.17$ dB, $D_0 = 39.49$ dB); (d) Lena by PC-MDFB: 34.12 dB ($D_2 = 34.36$ dB, $D_0 = 39.49$ dB).

bit rate and same central PSNRs. It can be seen that our method yields better PSNR and visual quality.

Finally, we give a brief comparison with the PCT. In Fig. 9 of [13], the D_0 of image Lena is kept at 35.78 dB. At this D_0, the JPEG-based single description coder in [13] needs a rate of 0.60 bpp, whereas the JPEG 2000 codec in [27] only needs 0.36 bpp. For a fair comparison, we focus on the redundancy ratio instead of the actual value of the rate. The lowest PCT redundancy ratio in Fig. 9 of [13] is about 10% over 0.60 bpp, with $D_1 \approx 25$ dB. In our method, when $R_1 = 0$, the lowest redundancy is 3.8%, with $D_1 = 22.90$ dB. Therefore, our method reaches a lower redundancy range than the PCT. In addition, when the rate is 10% over 0.36 bpp, the side PSNR of our method is $D_1 = 27.95$ dB, which is higher than that of the PCT at the same redundancy ratio.

VI. Conclusion

This paper studies the optimal filter bank for a prediction-compensated MDC scheme. We formulate the problem for both one-level and multiple-level decompositions for correlated Gaussian sources. Orthogonal and biorthogonal design examples show that the optimal solutions are very close to the optimal single description filter banks. Image coding results in the JPEG 2000 framework show that our method achieves similar or better performances than existing methods in the literature with very low complexity.

The image coding performance of our method can be further improved by refining the data partition and using adaptive intrasubband and intersubband predictions. Finally, although only balanced MDC is studied in this paper, unbalanced MDC can be easily generated by our scheme. In the theory, we can allow the two descriptions to have different base and enhancement layer rates, as mentioned in Section II-A, then the closed-form optimal bit allocations similar to (20) and (27) can still be obtained by the Lagrangian method. In JPEG 2000, this can be achieved by assign different target bit rates when encoding codeblocks of different descriptions. Alternatively, unbalanced descriptions can be created by partitioning the subbands unevenly.

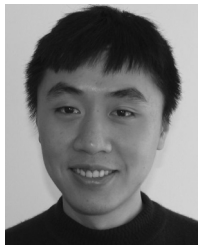
Acknowledgment

The authors would like to thank Dr. C. Tian for providing the MMDSQ codec in [5], and the anonymous reviewers for their suggestions that have significantly enhanced the presentation of this paper.

References

[1] V. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sep. 2001.
[2] V. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.
[3] V. Vaishampayan and J. Domaszewicz, "Design of entropy-constrained multiple description scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 40, no. 1, pp. 245–250, Jan. 1994.
[4] V. Vaishampayan and J. Batllo, "Asymptotic analysis of multiple description quantizers," *IEEE Trans. Inf. Theory*, vol. 44, no. 1, pp. 278–284, Jan. 1998.
[5] C. Tian and S. Hemami, "A new class of multiple description scalar quantizer and its application to image coding," *IEEE Signal Process. Lett.*, vol. 12, no. 4, pp. 329–332, Apr. 2005.
[6] S. Servetto, K. Ramchandran, V. Vaishampayan, and K. Nahrstedt, "Multiple description wavelet based image coding," *IEEE Trans. Image Process.*, vol. 9, no. 5, pp. 813–826, May 2000.
[7] J.-C. Batllo and V. A. Vaishampayan, "Asymptotic performance of multiple description transform codes," *IEEE Trans. Inf. Theory*, vol. 43, no. 2, pp. 703–707, Mar. 1997.
[8] J. Chen, C. Tian, and S. Diggavi, "Multiple description coding for stationary and ergodic sources," in *Proc. Data Compression Conf.*, Mar. 2007, pp. 73–82.
[9] W. Jiang and A. Ortega, "Multiple description coding via polyphase transform and selective quantization," in *Proc. SPIE Conf. Visual Comm. Image Processing*, Feb. 1999, vol. 3653, pp. 998–1008.
[10] A. Miguel, A. Mohr, and E. Riskin, "SPIHT for generalized multiple description coding," in *Proc. IEEE Conf. Image Process.*, Oct. 1999, vol. 3, pp. 842–846.
[11] T. Tillo, M. Grangetto, and G. Olmo, "Multiple description image coding based on Lagrangian rate allocation," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 673–683, Mar. 2007.
[12] N. Jayant, "Subsampling of a DPCM speech channel to provide two self-contained half-rate channels," *Bell Syst. Tech. J.*, vol. 60, pp. 501–509, Apr. 1981.
[13] Y. Wang, M. Orchard, V. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 351–366, Mar. 2001.
[14] Y. Wang, A. Reibman, M. Orchard, and H. Jafarkhani, "An improvement to multiple description transform coding," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2843–2854, Nov. 2002.
[15] S. S. Hemami, "Reconstruction-optimized lapped transforms for robust image transmission," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 2, pp. 168–181, Apr. 1996.
[16] D. Chung and Y. Wang, "Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 6, pp. 895–908, Sep. 1999.
[17] C. Tu, T. D. Tran, and J. Liang, "Error resilient pre-post-filtering for DCT-based block coding systems," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 30–39, Jan. 2006.

[18] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-processing," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1557–1571, Jun. 2003.
[19] J. Liang, C. Tu, L. Gan, T. D. Tran, and K.-K. Ma, "Wiener filter-based error resilient time domain lapped transform," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 428–441, Feb. 2007.
[20] G. Sun, U. Samarawickrama, J. Liang, C. Tian, C. Tu, and T. D. Tran, "Multiple description coding with prediction compensation," *IEEE Trans. Image Process.* Feb. 2007 [Online]. Available: http://www.ensc.sfu.ca/jiel/papers/MDLTPC.pdf, submitted for publication
[21] X. Yang and K. Ramchandran, "Optimal subband filter banks for multiple description coding," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2477–2490, Nov. 2000.
[22] P. Dragotti, S. Servetto, and M. Vetterli, "Optimal filter banks for multiple description coding: Analysis and synthesis," *IEEE Trans. Inf. Theory*, vol. 48, no. 7, pp. 2036–2052, Nov. 2002.
[23] Y. Liu and S. Oraintara, "Feature-oriented multiple description wavelet-based image coding," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 121–131, Jan. 2007.
[24] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
[25] D. Taubman and M. Marcellin, *JPEG 2000: Image Compression Fundamentals, Standards, and Practice*. Boston, MA: Kluwer, 2002.
[26] T. D. Tran, "M-channel linear phase perfect reconstruction filter bank with rational coefficients," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 49, no. 7, pp. 914–927, Jul. 2002.
[27] OpenJPEG. [Online]. Available: http://www.openjpeg.org
[28] PC-MDFB Source Code. [Online]. Available: http://www.ensc.sfu.ca/~jiel/MDJ2K.html
[29] RD-MDC Source Code. [Online]. Available: http://www.telematica.polito.it/sas-ipl
[30] J. Wang and J. Liang, "Filter banks for prediction-compensated multiple description coding," in *Proc. Data Compression Conf.*, Mar. 2008, pp. 392–401.
[31] E. Akyol, A. M. Tekalp, and R. Civanlar, "A flexible multiple description coding framework for adaptive peer-to-peer video streaming," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 2, pp. 231–245, Aug. 2007.



Jing Wang (S'07) received the B.E. and M.E. degrees from Zhejiang University, China, in 2004 and 2006, respectively, both in electrical engineering. He is currently working towards the Ph.D. degree at Simon Fraser University, Burnaby, BC, Canada.

His research interests include image and video coding, multimedia communications, and information theory.

Jie Liang (S'99–M'04) received the B.E. and M.E. degrees from Xi'an Jiaotong University, China, in 1992 and 1995, the M.E. degree from National University of Singapore (NUS), in 1998, and the Ph.D. degree from The Johns Hopkins University, Baltimore, MD, in 2003, respectively.

From 2003 to 2004, he was with the Video Codec Group of Microsoft Digital Media Division, Redmond, WA. Since May 2004, he has been an Assistant Professor at the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. His research interests include image and video coding, multirate signal processing, and joint source channel-coding.