Wiener Filter-Based Error Resilient Time-Domain Lapped Transform

Jie Liang, *Member, IEEE*, Chengjie Tu, *Member, IEEE*, Lu Gan, *Member, IEEE*, Trac D. Tran, *Member, IEEE*, and Kai-Kuang Ma, *Senior Member, IEEE*

***Abstract*—In this paper, the design of the error resilient time-domain lapped transform is formulated as a linear minimal mean-squared error problem. The optimal Wiener solution and several simplifications with different tradeoffs between complexity and performance are developed. We also prove the persymmetric structure of these Wiener filters. The existing mean reconstruction method is proven to be a special case of the proposed framework. Our method also includes as a special case the linear interpolation method used in DCT-based systems when there is no pre/post-filtering and when the quantization noise is ignored. The design criteria in our previous results are scrutinized and improved solutions are obtained. Various design examples and multiple description image coding experiments are reported to demonstrate the performance of the proposed method.**

***Index Terms*—Estimation, image coding, image communication, information theory.**

I. Introduction

With the rapid development of Internet, computer, and wireless communications technologies, there have been growing demands for delivering compressed images over Internet and wireless networks. This poses new challenges to conventional image compression algorithms, which are extremely vulnerable to transmission errors. On the other hand, perfect reception of all data is usually not necessary due to the intrinsic structures present in most natural images. Special algorithms known as error concealment can be employed to produce reasonable visual quality in the presence of transmission error.

Among the error concealment techniques that have been proposed [1], some methods, such as the reversible variable length coding [2], introduce error resilience at the encoder. Some of them focus on the error concealment at the decoder side by estimating the lost data with methods such as interpolation and projection onto convex sets [3]–[8]. Other approaches tackle the problem by a joint design of the encoder and decoder, for which the lapped transform provides a useful framework [9].

In the original lapped transform [9], a postfilter is applied at block boundaries after the DCT. The postfilter is usually designed to remove the remaining redundancy between neighboring blocks, thereby improving the coding efficiency of the DCT and reducing the blocking artifact associated with DCT-based schemes. On the other hand, the postfilter can also be designed to spread out the information of a block to its neighboring blocks. This is helpful when we need to recover a lost block during image transmission.

In [10], various techniques were proposed to estimate the lost data when the lapped orthogonal transform (LOT) was used. A mean reconstruction method and a nonlinear sharpening method were found to be quite effective, under the assumption that DC coefficients were intact. In particular, in the mean reconstruction method, each lost block was estimated by averaging its available neighboring blocks. In [11], it was found that the extended lapped transform (ELT) has better robustness against transmission error than the conventional LOT. However, the complexity of the ELT is higher than the LOT and it does not have linear phase [9]; thus, its application in image coding is limited.

The methods in [10] and [11] performed error concealment at the decoder. The transforms used there were still optimized for the best compression performance. In [12], the optimization of the error resilient LOT for a specific error concealment technique was addressed. It was shown that the lost blocks can be better recovered this way. In addition, the transform can be designed to achieve different tradeoffs between compression efficiency and error resilience.

Since the main purpose of [12] was to verify the feasibility of error resilient lapped transform, it still used the simple mean reconstruction method in [10]. To improve the smoothness of the reconstructed images, the maximally smooth recovery (MSR) method proposed in [13] was used in [14], and a multiple description codec was developed using the transforms designed in [12]. However, the transforms in [12] were not optimal for the MSR method because they were designed for the mean reconstruction method. This problem was resolved in [15] by incorporating the maximally smooth recovery constraint in the objective function of the transform optimization. Better transforms were obtained to achieve the same reconstruction quality with lower bit rate.

Manuscript received January 16, 2006; revised August 4, 2006. This work was supported in part by the SFU President's Research Grant, in part by the NSERC Discovery Grant, and in part by the National Science Foundation CAREER Grant CCR-0093262. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ljubisa Stankovic.

J. Liang is with the School of Engineering Science, Simon Fraser University, Burnaby, BC V5A 1S6 Canada (e-mail: jiel@sfu.ca).

C. Tu is with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: chentu@microsoft.com).

L. Gan is with the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool L69 3GJ, U.K. (e-mail: lu.gan@liv.ac.uk).

T. D. Tran is with the Electrical and Computer Engineering Department, The Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: trac@jhu.edu).

K.-K. Ma is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 (e-mail: ekkma@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2006.888332

Despite the improvements in [12], [14], and [15], some limitations still exist. First, only orthogonal lapped transform was considered. Therefore, the MSE of the reconstructed image is always the same [12], which seriously confines the error concealment capability of the system. Second, the entire M-band, $2M$-tap ($M \times 2M$ for short) lapped transform matrix is optimized directly (M is the block size), which increases the complexity of both optimization and implementation.

Recently, a new family of lapped transforms, the time-domain lapped transform (TDLT) [16], has been developed. In the TDLT, a prefilter is applied at each block boundary before the DCT. At the decoder side, a postfilter is applied at the same location after the inverse DCT. This framework is more compatible to DCT-based infrastructures, since the pre/postfilters can be easily incorporated into existing DCT software or hardware implementations. The TDLT also offers competitive compression performance compared to JPEG 2000 [17]. As a result, it has been used in Microsoft Windows Media Video 9 codec (WMV9) [18]. WMV9 has been accepted by the DVD Forum as one of the three mandatory formats for the next-generation HD DVD. It is also being standardized by the Society of Motion Picture and Television Engineers (SMPTE) as its VC-1 video coding standard. Therefore, the TDLT will play an important role in future image and video coding applications, and it is necessary to develop error resilient TDLT so that it can be used in error prone environments.

Error resilient TDLT is first considered in [19] and [20]. Thanks to the structure of the TDLT, the design of error resilient lapped transform reduces to that of the pre/postfilters. The problem is, therefore, more tractable. Biorthogonal solutions with lower MSE can be easily obtained by using biorthogonal pre/postfilters. In addition, the decoder can employ two postfilters—one for perfectly received blocks and another for lost blocks.

One remaining problem in [19] and [20] is that the mean reconstruction method is still used to estimate the lost blocks. In this paper, we present a general framework of error resilient TDLT. We formulate the filter design as a linear minimal mean-squared error (LMMSE) problem, and derive the corresponding Wiener filter solution, which unleashes the full potential of the error resilient lapped transform. Several simplifications of the general structure and their optimal solutions are then developed. The mean reconstruction method is revealed to be a trivial special case of the general solution. As a by-product, we also show that the linear interpolation method used in DCT systems is a special case of the proposed scheme when the pre/postfilters are disabled and when quantization error is ignored. In addition, we prove that these Wiener filters have persymmetric structure and can be implemented efficiently. We also revisit some of the design criteria used in our preliminary results in [21] and [22], such as the reconstruction gain and postfilter switching, and improved solutions are presented.

Compared to the mean reconstruction method in [19] and [20], the reconstruction error can be reduced by as much as 80% by the Wiener filter method, and up to $7 \sim 8$ dB improvement can be achieved in multiple description image coding experiments. Our method also shows considerable improvements over the results in [14] and [15].

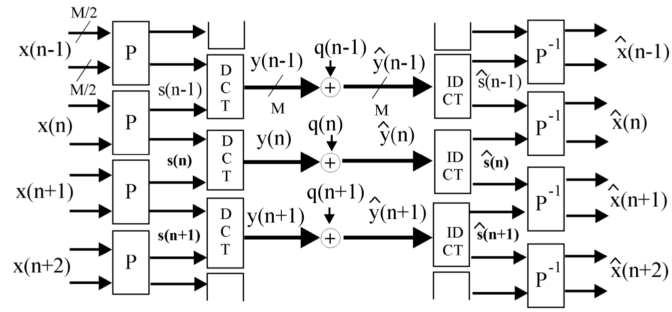


Fig. 1. Forward and inverse time-domain lapped transform.

II. General Pre/Postfiltering Structure for Error Concealment

Fig. 1 illustrates the time-domain lapped transform-based image compression system with block size of M (M is even). An $M \times M$ prefilter $\mathbf{P}$ is applied at the boundary of two neighboring blocks before the DCT. Therefore, the basis functions of the forward transform cover two blocks. Correspondingly, a postfilter $\mathbf{T}$ is applied by the decoder at each block boundary after inverse DCT. In this paper, we choose the following structure of $\mathbf{P}$ and $\mathbf{T}$ so that they yield linear-phase perfect reconstruction filter bank [16]

$$\begin{aligned}\mathbf{P} &= \mathbf{W}_M \mathrm{diag}\{\mathbf{I}_{M/2}, \mathbf{V}_{M/2}\}\mathbf{W}_M \\ \mathbf{T} = \mathbf{P}^{-1} &= \mathbf{W}_M \mathrm{diag}\{\mathbf{I}_{M/2}, \mathbf{V}_{M/2}^{-1}\}\mathbf{W}_M\end{aligned} \tag{1}$$

where

$$\mathbf{W}_M = \frac{1}{\sqrt{2}}\begin{bmatrix}\mathbf{I}_{M/2} & \mathbf{J}_{M/2} \\ \mathbf{J}_{M/2} & -\mathbf{I}_{M/2}\end{bmatrix}. \tag{2}$$

The $\mathbf{I}_{M/2}$ and $\mathbf{J}_{M/2}$ above are $M/2 \times M/2$ identity matrix and reversal identity matrix, respectively. The matrix $\mathbf{V}_{M/2}$ can be optimized to improve the compression performance of the system.

In this paper, we use $\mathbf{x}(i)$, $\mathbf{s}(i)$, $\mathbf{y}(i)$, and $\mathbf{q}(i)$ to denote the ith block of prefilter input, DCT input, DCT output, and quantization noise, respectively. Notice that $\mathbf{x}(i)$ is aligned with the prefilter, whereas the rest are aligned with the DCT. Compared with the notations in [12], [14], [15], [19], and [20], the definition here can simplify the problem formulation and the derivation of the optimal solution.

The structure of the time-domain lapped transform allows an effective strategy for error concealment. In this paper, we assume that all coefficients of a block are either received perfectly or lost entirely. This scenario can happen in, for example, multiple description coding [23]. In [19] and [20], a mean reconstruction method as in [12] is used, where the lost block is estimated by averaging the received neighboring blocks. It is shown in [19] and [20] that the pre/postfilters can be jointly designed such that the reconstructed quality is improved in the presence of transmission error. Moreover, two postfilters can be designed—one for perfectly received blocks and another one for lost blocks, as shown in Fig. 2(a).

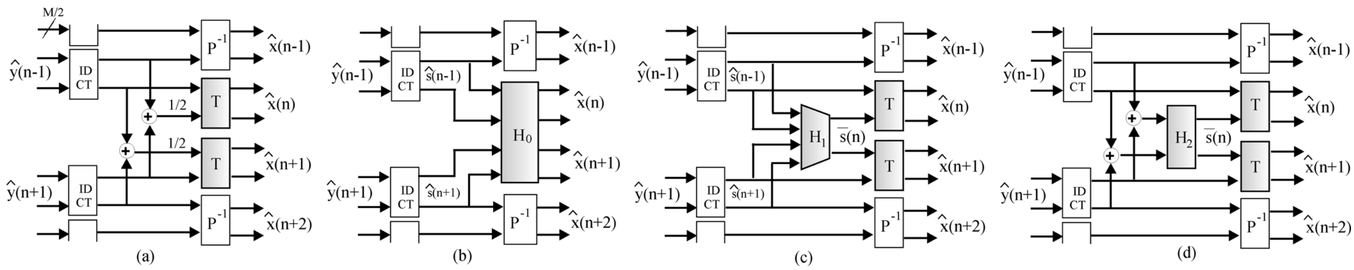


Fig. 2. (a) Existing decoder side error concealment design. (b) General structure for error concealment. (c) A close approximation of (b). (d) Further simplification of (c).

To further improve the reconstruction quality, notice that when $\hat{\mathbf{y}}(n)$ is lost, the error concealment problem can be viewed as the estimation of $\mathbf{x}(n)$ and $\mathbf{x}(n+1)$ from the observations $\hat{\mathbf{y}}(n-1)$ and $\hat{\mathbf{y}}(n+1)$, or equivalently $\hat{\mathbf{s}}(n-1)$ and $\hat{\mathbf{s}}(n+1)$. Therefore, the general filter should be a $2M \times 2M$ matrix $\mathbf{H}_0$, as shown in Fig. 2(b). If we define

$$\begin{aligned}
\mathbf{x}_2 &= [\,\mathbf{x}^T(n) \quad \mathbf{x}^T(n+1)\,]^T \\
\mathbf{x}_4 &= [\,\mathbf{x}^T(n-1) \quad \mathbf{x}^T(n) \quad \mathbf{x}^T(n+1) \quad \mathbf{x}^T(n+2)\,]^T \\
\hat{\mathbf{s}}_2 &= [\,\hat{\mathbf{s}}^T(n-1) \quad \hat{\mathbf{s}}^T(n+1)\,]^T \\
\mathbf{s}_3 &= [\,\mathbf{s}^T(n-1) \quad \mathbf{s}^T(n) \quad \mathbf{s}^T(n+1)\,]^T \\
\hat{\mathbf{s}}_3 &= [\,\hat{\mathbf{s}}^T(n-1) \quad \hat{\mathbf{s}}^T(n) \quad \hat{\mathbf{s}}^T(n+1)\,]^T \\
\mathbf{q}_3 &= [\,\mathbf{q}^T(n-1) \quad \mathbf{q}^T(n) \quad \mathbf{q}^T(n+1)\,]^T
\end{aligned} \tag{3}$$

the auto-correlation of the reconstruction error can be written as

$$\mathbf{R}_{\mathbf{ee}} = E\{(\mathbf{H}_0\hat{\mathbf{s}}_2 - \mathbf{x}_2)(\mathbf{H}_0\hat{\mathbf{s}}_2 - \mathbf{x}_2)^T\}. \tag{4}$$

The linear MMSE solution of $\mathbf{H}_0$ is one that minimizes the following MSE expression:

$$\mathcal{E} = \frac{1}{2M}\mathrm{trace}\{\mathbf{R}_{\mathbf{ee}}\}. \tag{5}$$

The optimal solution is given by the Wiener filter [24]

$$\mathbf{H}_0^* = \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}. \tag{6}$$

Matrices $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}$ and $\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}$ in (6) can be obtained as follows. We first partition $\mathbf{P}$ into

$$\mathbf{P} = [\,\mathbf{P}_0^T \quad \mathbf{P}_1^T\,]^T \tag{7}$$

where $\mathbf{P}_0$ and $\mathbf{P}_1$ contain the top and the bottom $M/2$ rows of the prefilter $\mathbf{P}$, respectively. Let $\mathbf{C}_3 = \mathrm{diag}\{\mathbf{C}, \mathbf{C}, \mathbf{C}\}$ represents the DCT operations of three neighboring blocks ($\mathbf{C}$ is the M-point DCT). We have

$$\hat{\mathbf{s}}_3 = \mathbf{s}_3 + \mathbf{C}_3^T\mathbf{q}_3 = \mathbf{P}_{34}\mathbf{x}_4 + \mathbf{C}_3^T\mathbf{q}_3 \tag{8}$$

where

$$\mathbf{P}_{34} = \mathrm{diag}\{\mathbf{P}_1, \mathbf{P}, \mathbf{P}, \mathbf{P}_0\}. \tag{9}$$

From (8), we get

$$\begin{aligned}
\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3} &= \mathbf{R}_{\mathbf{x}_4\mathbf{s}_3} = \mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T \\
\mathbf{R}_{\mathbf{s}_3\hat{\mathbf{s}}_3} &= \mathbf{R}_{\mathbf{s}_3\mathbf{s}_3} = \mathbf{P}_{34}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T \\
\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3} &= \mathbf{P}_{34}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T + \mathbf{C}_3^T\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}\mathbf{C}_3
\end{aligned} \tag{10}$$

where $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ and $\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}$ are correlation matrices of $\mathbf{x}_4$ and $\mathbf{q}_3$, respectively. In deriving (10), we assume that the input is uncorrelated with the quantization noise. Matrices $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}$ and $\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}$ in (6) can, thus, be obtained from the appropriate submatrices of $\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3}$ and $\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3}$. In this paper, $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ in (10) is obtained by assuming the input follows an AR(1) model. We also assume that the quantization noises of different subbands are uncorrelated, i.e., $\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}$ is a diagonal matrix. At high rates, the noise variance of the kth subband can be written as [25]

$$\sigma_{q_k}^2 = c\,\sigma_{y_k}^2\,2^{-2R_k} \tag{11}$$

where c is a constant that depends on the input statistics, $\sigma_{y_k}^2$ is the subband variance that can be obtained from the input statistics and the forward transform, and R_k is the bit rate allocated to the kth subband. The bit allocation can be optimized during the filter design. We will demonstrate in Section VI-C that the quantization noise can be safely ignored in the Wiener filter expression, since the final error is dominated by the transmission loss.

When applied to image error concealment, an important requirement is that the Wiener filter should maintain the DC component of the local region, i.e.,

$$\mathbf{H}_0\mathbf{d} = \mathbf{d} \tag{12}$$

where $\mathbf{d} = [1, 1, \ldots, 1]^T$. Therefore, we need to minimize the MSE subject to the constraint in (12). The solution can be found by the Lagrangian method, which constructs the following objective function:

$$\mathcal{E}' = \frac{1}{2M}\mathrm{trace}\{\mathbf{R}_{\mathbf{ee}}\} + \sum_{i=0}^{2M-1}\lambda_i(\mathbf{h}_{0,i}\mathbf{d} - 1) \tag{13}$$

where $\mathbf{h}_{0,i}$ is the ith row of $\mathbf{H}_0$. The corresponding LMMSE solution is

$$\bar{\mathbf{H}}_0^* = (\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2} - M\boldsymbol{\lambda}\mathbf{d}^T)\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}, \tag{14}$$

where

$$\begin{aligned}\boldsymbol{\lambda} &= [\lambda_0, \lambda_1, \ldots, \lambda_{2M-1}]^T \\ &= \frac{1}{M(\mathbf{d}^T \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} \mathbf{d})} \left(\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} - \mathbf{I}\right) \mathbf{d}.\end{aligned} \tag{15}$$

Another way to ensure the DC condition is to scale each row of the Wiener filter such that the sum of every row is unity, i.e.,

$$\hat{\mathbf{H}}_0^* = \boldsymbol{\Theta} \mathbf{H}_0^* \tag{16}$$

where $\boldsymbol{\Theta}$ is a $2M \times 2M$ diagonal matrix whose ith diagonal entry is $\boldsymbol{\Theta}(i, i) = 1/\sum_{j=0}^{2M-1} \mathbf{H}_0^*(i, j)$. Our experimental results show that the difference between the two solutions is negligible. Therefore, the scaling method in (16) is chosen in the rest of this paper due to its simplicity.

III. Simplifications and Factorizations

In this section, we show that the general solution in (6) can be approximately factorized into two stages, which can simplify the implementation. Further simplifications are also presented, leading to the conclusion that the mean reconstruction method is a special case of our method. We also prove that these Wiener filters are persymmetric, a property that can be used to further reduce the implementation cost.

A. A Close Approximation of the General Solution

In [19] and [20], the lost blocks are first estimated by the simple average of neighboring blocks before applying postfilter. This is clearly not optimal in the light of estimation theory. To find the optimal estimate of $\mathbf{s}(n)$ in this two-stage approach, we define an $M \times 2M$ matrix $\mathbf{H}_1$, i.e.,

$$\bar{\mathbf{s}}(n) = \mathbf{H}_1 \hat{\mathbf{s}}_2. \tag{17}$$

The optimal solution for $\mathbf{H}_1$ can be found by minimizing

$$\mathcal{E}_1 = \frac{1}{M} \mathrm{trace}\{E\{(\mathbf{H}_1\hat{\mathbf{s}}_2 - \mathbf{s}(n))(\mathbf{H}_1\hat{\mathbf{s}}_2 - \mathbf{s}(n))^T\}\} \tag{18}$$

and the solution is also a Wiener filter, which can be written as

$$\mathbf{H}_1^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} \tag{19}$$

where $\mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_2}$ is a submatrix of $\mathbf{R}_{\mathbf{s}_3\hat{\mathbf{s}}_3}$.

Once $\bar{\mathbf{s}}(n)$ is obtained, the postfilter is applied as usual. A different postfilter can be used around lost blocks to further improve the visual quality. This scheme, as shown in Fig. 2(c), can be viewed as imposing the following structure to the general matrix $\mathbf{H}_0$ in Fig. 2(b):

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{T}_M & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{T}_M \end{bmatrix} \begin{bmatrix} \mathbf{I}_{b1} \\ \mathbf{H}_1 \\ \mathbf{I}_{b2} \end{bmatrix} \tag{20}$$

where $\mathbf{I}_{b1}$ and $\mathbf{I}_{b2}$ are defined by

$$\begin{aligned}\mathbf{I}_{b1} &= [\mathbf{0}_{M/2} \quad \mathbf{I}_{M/2} \quad \mathbf{0}_{M/2} \quad \mathbf{0}_{M/2}] \\ \mathbf{I}_{b2} &= [\mathbf{0}_{M/2} \quad \mathbf{0}_{M/2} \quad \mathbf{I}_{M/2} \quad \mathbf{0}_{M/2}].\end{aligned} \tag{21}$$

When applied to $\hat{\mathbf{s}}_2, \mathbf{I}_{b1}$ simply extracts the second half of $\hat{\mathbf{s}}(n-1)$ and $\mathbf{I}_{b2}$ extracts the first half of $\hat{\mathbf{s}}(n+1)$.

Due to the structure of the lapped transform, the result given by the two-stage method in (20) approximates the general solution (6) very well. In fact, the two structures are equivalent if $\mathbf{T} = \mathbf{P}^{-1}$ and if we ignore the quantization noise. In this case, $\hat{\mathbf{x}}(n) = \mathbf{x}(n)$, $\hat{\mathbf{s}}(n) = \mathbf{s}(n)$. It can be seen from Fig. 1 that

$$\mathbf{x}_2 = \hat{\mathbf{x}}_2 = \begin{bmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{s}}_1(n-1) \\ \hat{\mathbf{s}}(n) \\ \hat{\mathbf{s}}_0(n+1) \end{bmatrix} \tag{22}$$

where $\hat{\mathbf{s}}_1(n-1)$ and $\hat{\mathbf{s}}_0(n+1)$ denote the second half of $\hat{\mathbf{s}}(n-1)$ and the first half of $\hat{\mathbf{s}}(n+1)$, respectively. Plugging (22) into the definition of $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}$, the general Wiener filter in (6) becomes

$$\mathbf{H}_0^* = \begin{bmatrix} \mathbf{T} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{\hat{\mathbf{s}}_1(n-1)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} \\ \mathbf{R}_{\hat{\mathbf{s}}(n)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} \\ \mathbf{R}_{\hat{\mathbf{s}}_0(n+1)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} \end{bmatrix}. \tag{23}$$

Since $\hat{\mathbf{s}}_1(n-1)$ and $\hat{\mathbf{s}}_0(n+1)$ are parts of $\hat{\mathbf{s}}_2$, we have $\mathbf{R}_{\hat{\mathbf{s}}_1(n-1)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} = \mathbf{I}_{b1}$ and $\mathbf{R}_{\hat{\mathbf{s}}_0(n+1)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} = \mathbf{I}_{b2}$; thus, the general Wiener filter reduces to the two-stage method given by (19) and (20).

In [20], it is found that applying two postfilters can improve the performance of the mean reconstruction method. One postfilter is for the correctly received blocks and another one is for the lost blocks. When the $M \times 2M$ Wiener filter (19) is used, the analysis above shows that there is no need to switch between two postfilters. The perfect reconstruction postfilter $\mathbf{P}^{-1}$ is sufficient. Therefore, the implementation can be simplified.

B. Further Simplifications

The complexity of (20) and Fig. 2(c) can be further reduced by imposing the following constraint on $\mathbf{H}_1$:

$$\mathbf{H}_1 = [\mathbf{H}_2 \quad \mathbf{H}_2] = \mathbf{H}_2 [\mathbf{I}_M \quad \mathbf{I}_M] \tag{24}$$

where the size of $\mathbf{H}_2$ is $M \times M$. This is equivalent to estimating $\mathbf{s}(n)$ by

$$\bar{\mathbf{s}}(n) = \mathbf{H}_2 (\hat{\mathbf{s}}(n-1) + \hat{\mathbf{s}}(n+1)) \triangleq \mathbf{H}_2 \hat{\mathbf{s}}_a. \tag{25}$$

The structure is shown in Fig. 2(d). Again, Wiener solution exists in this case and is given by

$$\mathbf{H}_2^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_a} \mathbf{R}_{\hat{\mathbf{s}}_a\hat{\mathbf{s}}_a}^{-1}. \tag{26}$$

The matrices involved can be obtained from (10) by simple manipulations.

In this case, our experimental results show that using two postfilters is still necessary, because the performance of the $M \times M$ filter (26) is far below that of (19), and a special postfilter is required to further reduce the error.

It is clear from (24) that the mean reconstruction method used in [19] and [20] is simply a special case of the already suboptimal approach in (24) with $\mathbf{H}_2 = (1/2)\mathbf{I}_M$. Therefore, it can be expected that the error concealment performance can be improved considerably if $\mathbf{H}_2^*$, $\mathbf{H}_1^*$, or $\mathbf{H}_0^*$ are used.

C. Persymmetry of the Wiener Filters

The Wiener filters $\mathbf{H}_{L\times N}$, derived in the last two sections, including the constrained solutions (14) and (16), satisfy the following persymmetric condition (also known as centrosymmetric) [26]:

$$\mathbf{H}_{L\times N} = \mathbf{J}_{L\times L}\mathbf{H}_{L\times N}\mathbf{J}_{N\times N} \tag{27}$$

which means $\mathbf{H}_{i,j} = \mathbf{H}_{L-1-i,N-1-j}$, for $i = 0,\ldots,L-1$, $j = 0,\ldots,N-1$, i.e., the matrix is symmetric with respect to its center. The proof is given in the Appendix and relies on the linear-phase property of the lapped transform. The persymmetric structure allows the Wiener filters to be factorized as [27]

$$\mathbf{H}_{L\times N} = \mathbf{W}_{L\times L}\mathrm{diag}\{\mathbf{A}_{L/2\times N/2}, \mathbf{B}_{L/2\times N/2}\}\mathbf{W}_{N\times N}. \tag{28}$$

This factorization can be exploited to reduce the complexity of implementation by roughly 50%.

IV. Estimation From One Neighboring Block

Sometimes it is necessary to estimate a lost block from only one neighboring block. This can happen at the image boundary or when contiguous blocks are lost. A Wiener solution can also be obtained for this scenario. When only the previous block $\hat{\mathbf{s}}(n-1)$ is available after the inverse DCT, the objective is to find an $M\times M$ matrix $\mathbf{H}_3^*$ which minimizes

$$\mathrm{trace}\{E\{(\mathbf{H}_3\hat{\mathbf{s}}(n-1)-\mathbf{s}(n))(\mathbf{H}_3\hat{\mathbf{s}}(n-1)-\mathbf{s}(n))^T\}\}. \tag{29}$$

The solution is, thus, given by

$$\mathbf{H}_3^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}(n-1)}\mathbf{R}^{-1}_{\hat{\mathbf{s}}(n-1)\hat{\mathbf{s}}(n-1)}. \tag{30}$$

The matrices involved can be readily obtained from $\mathbf{R}_{\mathbf{s}_3\hat{\mathbf{s}}_3}$ and $\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3}$ in (10).

Similarly, if only the next block $\hat{\mathbf{s}}(n+1)$ is available, the corresponding Wiener solution $\mathbf{H}_4^*$ becomes

$$\mathbf{H}_4^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}(n+1)}\mathbf{R}^{-1}_{\hat{\mathbf{s}}(n+1)\hat{\mathbf{s}}(n+1)}. \tag{31}$$

From that facts that $\mathbf{R}_{\hat{\mathbf{s}}(n-1)\hat{\mathbf{s}}(n-1)} = \mathbf{R}_{\hat{\mathbf{s}}(n+1)\hat{\mathbf{s}}(n+1)}$, $\mathbf{R}_{\hat{\mathbf{s}}(n-1)\hat{\mathbf{s}}(n-1)}$ is persymmetric, and $\mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}(n-1)} = \mathbf{J}\mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}(n+1)}\mathbf{J}$, as shown in the Appendix, it is straightforward to verify that $\mathbf{H}_3^* = \mathbf{J}\mathbf{H}_4^*\mathbf{J}$, i.e., the two filters are persymmetric to each other. However, each of them is not persymmetric; therefore, they cannot be factorized as in (28).

V. Relationship With Linear Interpolation in DCT Systems

Another special case of the proposed framework is when the pre/postfilters are turned off. In this case, the lapped transform reduces to the DCT. If we still estimate a lost block $\mathbf{x}(n)$ by its two neighboring blocks $\hat{\mathbf{x}}_B = \left[\hat{\mathbf{x}}^T(n-1), \hat{\mathbf{x}}^T(n+1)\right]^T$, the Wiener filter (19) becomes

$$\mathbf{H}_{\mathrm{DCT}} = \mathbf{R}_{\mathbf{x}(n)\hat{\mathbf{x}}_B}\mathbf{R}^{-1}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}. \tag{32}$$

An interesting fact is that when the quantization noise in $\mathbf{R}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}$ is ignored, only the Mth and the $(M+1)$th columns of the Wiener filter are nonzero, and the result is, therefore, a linear interpolation of the missing block using the last pixel of $\hat{\mathbf{x}}(n-1)$ and the first pixel of $\hat{\mathbf{x}}(n+1)$. This justifies the wide adoptions of linear interpolation method in, for example, [4], [5], and [7].

To illustrate this property and to gain more insights, we assume $\mathbf{R}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B} = \mathbf{R}_{\mathbf{x}_B\mathbf{x}_B}$ and consider the expression of the Wiener filter when the block size is 2. Assuming an AR(1) input with unit variance and correlation coefficient ρ, one can show that

$$\mathbf{R}_{\mathbf{x}_n\hat{\mathbf{x}}_B} = \begin{bmatrix}\rho^2 & \rho & \rho^2 & \rho^3\\ \rho^3 & \rho^2 & \rho & \rho^2\end{bmatrix} \tag{33}$$

$$\mathbf{R}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B} = \begin{bmatrix}1 & \rho & \rho^4 & \rho^5\\ \rho & 1 & \rho^3 & \rho^4\\ \rho^4 & \rho^3 & 1 & \rho\\ \rho^5 & \rho^4 & \rho & 1\end{bmatrix}. \tag{34}$$

Using the adjoint matrix method, the first column of the inverse matrix $\mathbf{R}^{-1}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}$ can be found to be

$$\frac{1}{\det(\mathbf{R}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B})}\begin{bmatrix}1-\rho^2-\rho^6+\rho^8\\ -\rho(1-\rho^2-\rho^6+\rho^8)\\ 0\\ 0\end{bmatrix}. \tag{35}$$

It can be easily verified that the first column of the product $\mathbf{R}_{\mathbf{x}_n\hat{\mathbf{x}}_B}\mathbf{R}^{-1}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}$ becomes all zero. Since the Wiener filter is still persymmetric in this case, the last column is also zero. When $\rho = 0.95$, the filter becomes

$$\mathbf{H}_{\mathrm{DCT}} = \begin{bmatrix}0 & 0.67 & 0.33 & 0\\ 0 & 0.33 & 0.67 & 0\end{bmatrix} \tag{36}$$

which represents the linear interpolation between the two immediate neighboring pixels of the lost block.

When $M = 8$, the Wiener filter (32) only has the following nonzero entries in the eighth and the ninth columns

$$\begin{bmatrix}0.88 & 0.77 & 0.65 & 0.54 & 0.43 & 0.32 & 0.21 & 0.11\\ 0.11 & 0.21 & 0.32 & 0.43 & 0.54 & 0.65 & 0.77 & 0.88\end{bmatrix}^T \tag{37}$$

which is also a linear interpolation operator.

However, if quantization noise is included in $\mathbf{R}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}$, its inverse $\mathbf{R}^{-1}_{\hat{\mathbf{x}}_B\hat{\mathbf{x}}_B}$ would lose the nice analytical expression in (35), and the Wiener filter would in general be a full matrix, meaning that the linear interpolation is no longer optimal in the DCT systems.

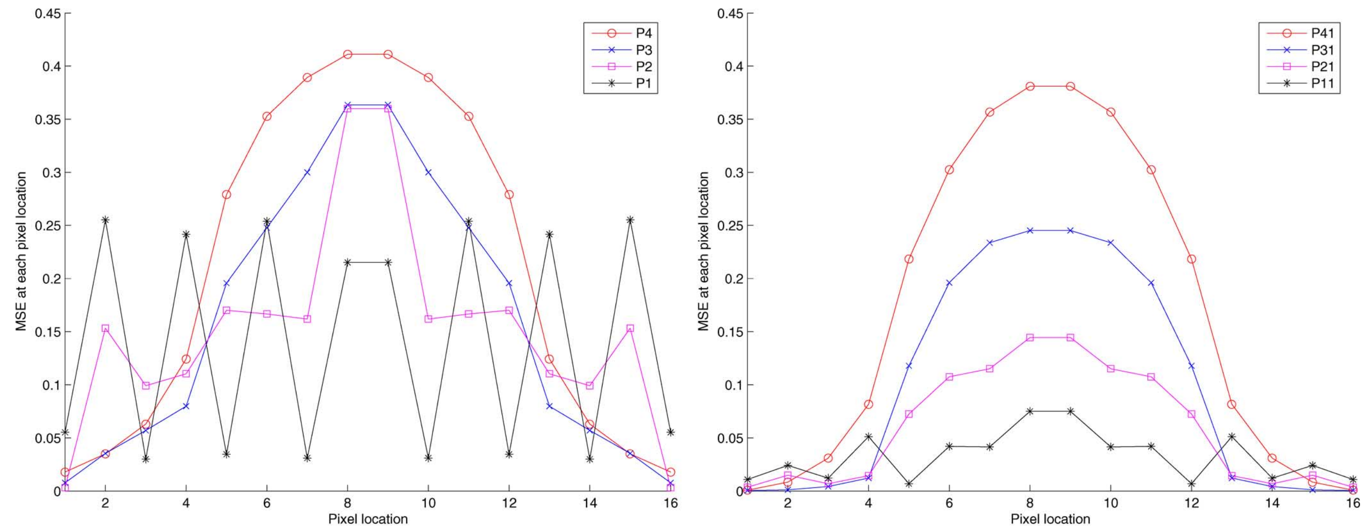
VI. Design Examples and Applications

A. Design Criteria

In this section, we show various design examples and their applications in multiple description image coding. In [20] and our preliminary results in [21] and [22], a Matlab optimization program is used to find the optimal TDLT that maximizes the following objective function

$$J = G_{TC} - \alpha\,\mathcal{E} + \beta\,G_R \tag{38}$$

which is a weighted average of the coding gain G_{TC} of the transform, the residual MSE $\mathcal{E}$ in (5) after transmission error and Wiener filter-based error concealment, and the reconstruction gain G_R. Notice that the classical lapped transform is only optimized for coding gain. The coding gain measures the maximum



(a) (b)

Fig. 3. (a) Error distributions of the TDLT designs in [20]. (b) Error distributions of new designs.

distortion reduction of a transform over PCM scheme. With optimal bit allocation the coding gain reduces to [25]

$$G_{TC} \triangleq 10\log_{10} \frac{\sigma_x^2}{\left(\prod_{i=0}^{M-1} \sigma_{y_i}^2 \|f_i\|^2\right)^{1/M}} \tag{39}$$

where σ_x^2 is the variance of the input, $\sigma_{y_i}^2$ is the variance of the ith subband, and $\|f_i\|^2$ is the norm of the ith synthesis basis function. The input is assumed to follow an *AR* (1) model with correlation coefficient $\rho = 0.95$.

The reconstruction gain is first defined in [12] to control the error distribution across transform coefficients when a block is lost. This is pivotal to the performance of orthogonal transforms because the MSE is always the same in different methods. The reconstruction gain in [12] attains its maximal value when the error is uniformly distributed across all transform coefficients.

Since biorthogonal transform is used in [20], the estimation error in the transform domain is no longer equal to the final reconstruction error. Therefore, the following spatial domain reconstruction gain is defined [20]:

$$G_R = \frac{\left(\prod_{i=0}^{2M-1} \sigma_{e_i}^2\right)^{1/2M}}{\frac{1}{2M}\sum_{i=0}^{2M-1} \sigma_{e_i}^2} \tag{40}$$

where $\sigma_{e_i}^2$ is the ith diagonal entry of $\mathbf{R_{ee}}$ in (4), i.e., the final expected reconstruction error of the ith pixel in the two blocks $[\mathbf{x}^T(n) \quad \mathbf{x}^T(n+1)]^T$.

When Wiener filter is applied, our experimental results show that the reconstruction error can be reduced so dramatically that there is no need to control the error distribution explicitly. Therefore, we always fix $\beta = 0$ in this paper. In addition, our experiments indicate that the error distribution control imposes serious constraints on the filter design. Removing this criterion in the optimization can improve the coding gain or reconstruction quality, as will be shown later.

In fact, higher spatial domain reconstruction gain as defined above does not always lead to pleasant visual quality, since it tends to generate similar error at each pixel, thereby creating a clear visual artifact around the lost block, especially if the neighboring blocks have very small quantization error. This suggests that visually pleasant error distribution should have a smooth transition between healthy blocks and the error concealed blocks. In [14] and [15], this is achieved by solving for each block a underdetermined equation with a smoothness constraint. In our framework, if the control of error distribution is necessary, we can include a curve fitting term in the objective function such that the optimized reconstruction error across the two blocks follows a desired shape, for example, a Gaussian curve. Moreover, the template can be defined to match the neighboring quantization noise at the block boundary. Further discussion can be found in [22]. However, when the error distribution constraint is introduced, the coding gain or the MSE has to be sacrificed; therefore, it becomes more complicated to find a good tradeoff.

B. Comparison With Mean Reconstruction Method

In this section, we compare the performance of the Wiener filter-based TDLT design with that of the mean reconstruction-based results in [20]. The entries in matrix $\mathbf{V}$ of (1) are the optimization parameters. Different solutions can be obtained by varying α in (38) (β is fixed as 0). The quantization noise is ignored in Wiener filter expressions, for example, (19). We will show in Section VI-C that this is a reasonable choice.

Four families with $M = 8$ are designed. Their coding gains and residual MSE after error concealment are summarized in Table I. Some optimized matrices $\mathbf{V}$ in the prefilter are shown in Table II, and the error distribution of some designs across the two affected blocks $\mathbf{x}(n)$ and $\mathbf{x}(n+1)$ are plotted in Fig. 3. The Wiener filters $\mathbf{H}_0^*$, $\mathbf{H}_1^*$, and $\mathbf{H}_2^*$ in (6), (19), and (26) are used

TABLE I
DESIGN EXAMPLES OF DIFFERENT CONFIGURATIONS WITH $M = 8$, $\beta = 0$ AND DIFFERENT α IN (38)

Cfg.	P1	P10	P11	P12	P2	P20	P21	P22	P3	P30	P31	P32	P4	P40	P41	P42
α	-	92	92	111	-	26	26	37.3	-	12.5	12.5	18.5	-	1	1	1
G_{TC}	6.96	6.96	6.96	6.96	8.41	8.42	8.42	8.42	9.17	9.17	9.17	9.17	9.61	9.61	9.61	9.61
MSE	0.14	0.03	0.03	0.08	0.15	0.06	0.06	0.10	0.16	0.10	0.10	0.13	0.21	0.17	0.17	0.18

TABLE II
SOME OPTIMIZED RESULTS FOR THE FREE MATRIX $\mathbf{V}_{M/2}$ IN (1) WITH $M = 8$

Cfg.	P11	P21
$\mathbf{V}_{\frac{M}{2}}$	$\begin{bmatrix} -0.9672 & 0.4692 & 0.3425 & 0.1418 \\ -0.1651 & 0.6042 & -0.7937 & 0.0394 \\ 0.4408 & 0.5179 & 0.0867 & 0.7004 \\ -0.2171 & -0.5711 & -0.3456 & 0.7543 \end{bmatrix}$	$\begin{bmatrix} 0.9448 & 0.1452 & -0.7385 & -0.0213 \\ -0.0809 & 1.0449 & -0.2754 & -0.2339 \\ 0.7003 & 0.2656 & 0.6545 & -0.1442 \\ 0.1620 & 0.2895 & -0.0939 & 0.9174 \end{bmatrix}$
Cfg.	P31	P41
$\mathbf{V}_{\frac{M}{2}}$	$\begin{bmatrix} 0.9588 & 0.6245 & -0.0743 & -0.1501 \\ -0.4781 & 0.9028 & 0.3171 & -0.1448 \\ 0.1968 & -0.2192 & 0.9954 & 0.0638 \\ 0.0278 & 0.1226 & -0.0712 & 1.0049 \end{bmatrix}$	$\begin{bmatrix} 0.9490 & 0.7662 & 0.3260 & 0.2031 \\ -0.5546 & 0.8881 & 0.5931 & 0.2007 \\ 0.1099 & -0.3708 & 1.0738 & 0.3596 \\ -0.0307 & 0.0048 & -0.1267 & 1.1664 \end{bmatrix}$

to obtain configurations Pi0, Pi1, and Pi2 ($i = 1, \ldots, 4$), respectively. Two postfilters are optimized for configurations Pi2, since $\mathbf{H}_2^*$ is suboptimal. The configurations in each family are designed to have similar coding gains so that their performance can be compared fairly. The first group yields the lowest coding gain but also the lowest MSE in the presence of transmission error. On the contrary, the last group has the highest coding gain (close to the best result in [16]) but is the most vulnerable to transmission loss.

Also included in Table I are the configurations P1 to P4 in [20] with the mean reconstruction method ($\mathbf{H}_2 = (1/2)\mathbf{I}$). Among them, P3 and P4 use two postfilters. The first one is the inverse of the prefilter and is applied to healthy blocks, and the second one is applied around lost blocks to further reduce the reconstruction error.

Compared with P1 to P4, we can see from these results that the final reconstruction error in the presence of transmission error can be reduced substantially by the $2M \times 2M$ and $M \times 2M$ Wiener filters. The improvement is more pronounced as coding gain decreases, since more correlations among neighboring blocks are introduced. For example, compared with P1, the MSE after error concealment is reduced by 80% in P10 and P11. It can still be reduced by 20% even when the coding gain is at its highest value, as given by P40 and P41. Notice also that even the MSE of P3i is less than that of P1, although the coding gain of P3i is much higher than P1.

Table I also shows that the MSE given by $\mathbf{H}_1^*$ in (19) is identical to that of the general $2M \times 2M$ solution $\mathbf{H}_0^*$ in (6). This verifies their relationship proved in Section III. The performance of the $M \times M$ Wiener filter $\mathbf{H}_2^*$ lies roughly halfway between the mean reconstruction and the $M \times 2M$ Wiener filter.

Some results in Table I are also better than our preliminary results in [21] and [22], due to the elimination of reconstruction gain in the objective function and due to the switch of two postfilters for Pi2 only. The improvement can also be observed in the image coding experiments below.

To verify the error concealment performance of the Wiener filters, we implement a multiple description codec following the approach in [14], where the transformed image is split into four descriptions at block level. For example, all (even, even)-indexed blocks are grouped into the first description, all (even, odd)-indexed blocks are grouped into the second description, and so on. The context-adaptive binary arithmetic coding in [17] is applied to each description independently. At the decoder side, the inverse DCT is applied to all received blocks first. After that, the Wiener filter is used to estimate each lost block before applying postfiltering. The Wiener filters derived in this paper are based on 1-D signal model. To apply them to 2-D images, we first estimate each row of a lost block from its horizontal neighboring blocks, and then estimate each column from its vertical neighbors. The final result is the weighted average of the two estimations, and the weighting factors are decided by the number of available neighbors in each direction.

The scenario of losing three out of the four descriptions requires special treatments. In this case, a three-step method is employed to estimate the lost data. We first estimate those blocks that have available horizontal neighbors, followed by those blocks with available vertical neighbors. The rest are then estimated using previously estimated neighbors. We will show later that this approach achieves significant improvement over the maximally smooth recovery method in [14] and [15].

Fig. 4 plots the R-D curves of different TDLT configurations with different number of available descriptions when the 512×512 Lena image is used. Fig. 4(a) shows that all new designs yield better compression performance than their counterparts in [20] with the same coding gain. In addition, they provide much better reconstruction results than the mean reconstruction method. For example, at 2 bits/pixel (bpp), the PSNRs by P11 and P21 is about $7 \sim 8$ dB and 4 dB higher than that of P1 and P2, respectively. We can also see that the improvement of P1 over P4 is less than 3 dB in most cases, whereas P11 can be 10 dB better than P41. This agrees with the theoretical MSE in Table I. Moreover, it can be seen that even P31 can achieve better R-D performance than P1. Overall, P21 offers a good tradeoff between the compression performance and error resilience.

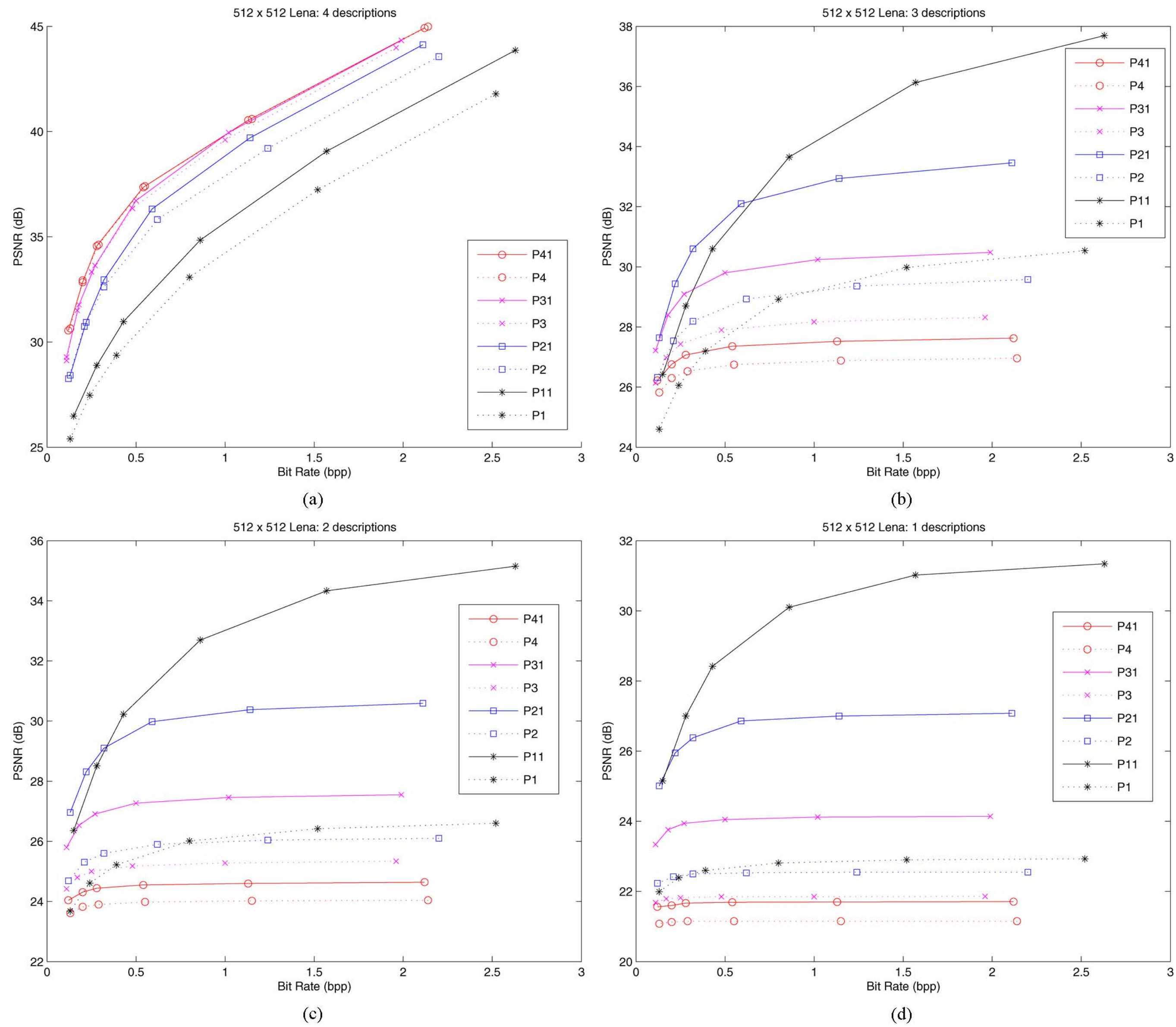


Fig. 4. R-D curves of the 512×512 Lena image under different loss patterns and different TDLT configurations: (a) with four descriptions, (b) with three descriptions, (c) with two descriptions, and (d) with one description.

Fig. 5 shows the reconstruction results of different methods when two descriptions of the 512×512 Lena image are lost. The average bit rate is 1 bpp. The PSNRs with and without error are reported for each case. Satisfactory results are produced by P11 and P21. Notice that the coding performance of P21 is only 0.9 dB below the compression-optimized TDLT, but the reconstruction after error is 6.5 dB higher. The results of P1 to P4 are quite blurred and exhibit strong ghost artifacts near the edges, due to the average operator in mean reconstruction method.

Portions of the multiple description decoding results for the 512×512 Barbara image are given in Fig. 6, when P11 is used. It can be seen that the quality of the reconstruction decreases gradually when more descriptions are lost. This shows that the AR(1) model-based Wiener filter is quite robust even for images of rich textures.

C. Robustness of the Wiener Filter to Quantization Noise

As mentioned in Section VI-B, the Wiener filter examples in this paper are obtained by ignoring the quantization noise. In this subsection, we investigate the robustness of the Wiener filter to quantization noise. A prefilter and the corresponding Wiener filter are designed by including the quantization noise (11) in (19). This requires us to choose a typical average bit rate, because in practice we prefer to have a prefilter and a Wiener filter that can be used at all bit rates. In our optimization, the average bit rate is fixed as 1 bit/pixel. Notice that the bit rate does not have to be specified in the coding gain optimization part, because the bit rate can be canceled out under optimal bit allocation [25].

The parameter α in (38) is chosen as 30 so that the coding gain of the optimized lapped transform is the same as P21 in

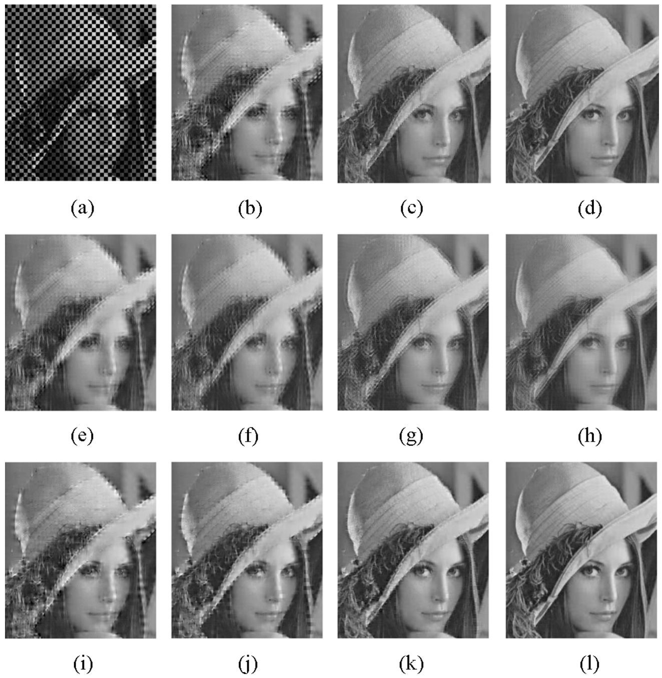


Fig. 5. Decoding results at 1 bpp and 50% loss. (a) Loss pattern. (b) Original TDLT (23.87/39.94 dB). (c) P22 (27.54/38.38 dB). (d) P12 (28.68/35.59 dB). (e) P4 (24.01/39.92 dB). (f) P3 (25.28/39.61 dB). (g) P2 (26.00/38.11 dB). (h) P1 (26.19/34.47 dB). (i) P41 (24.59/39.96 dB). (j) P31 (27.45/39.84 dB). (k) P21 (30.32/39.03 dB). (l) P11 (33.20/35.91 dB).

Table I. The multiple description coding performances of the two configurations for the images Lena and Barbara are compared in Fig. 7, which shows that the result depends on the nature of the image. In particular, P21 gives better concealment performance for the Lena image, whereas considering the quantization noise improves the performance of the Barbara image. Therefore, a better quantization noise model is necessary in order to get a consistent gain. However, the fact that the difference of the two methods is less than 0.4 dB in all cases suggests that the Wiener filter is not sensitive to the quantization noise, since the error caused by transmission loss is usually much higher than the quantization noise.

D. Comparison With MSR Method

In this section, we compare the performance of our Wiener filter-based TDLT with the results in [14] and [15]. The method in [14] uses the maximally smooth recovery in the estimation part and examples T6 to T9 in [12] in the transform part. Since our codec uses arithmetic coding, whereas adaptive Huffman coding is used in [14], we will only compare the performance of the two methods in the absence of quantization error. This is to ensure fair comparison of the transform and the error concealment parts. Reconstructed PSNRs for the 256×256 Lena image with different number of received descriptions are reported in [14] and is duplicated in Table III. For fair comparison, we design four $M \times 2M$ Wiener filters using $\mathbf{H}_1^*$ in (19) to match the coding gains of T6 to T9, respectively. As discussed before, β is fixed as 0 in the objective function (38) and the postfilter is chosen as the inverse of the prefilter.

As shown in Table III, the Wiener filter based TDLT achieves an average of 2.98 dB improvement over [14]. More improvements are achieved at higher transform coding gain. In addition, the most significant improvement, 3.94 dB on average, happens when only one description is received. Therefore, our method offers more graceful quality degradation when more descriptions are lost.

Fig. 6. Portions of reconstruction results with P11 design and the 512×512 Barbara image at 1 bpp. (a) four descriptions (32.72 dB), (b) three descriptions (29.09 dB), (c) two descriptions (27.12 dB), and (d) one description (24.11 dB).

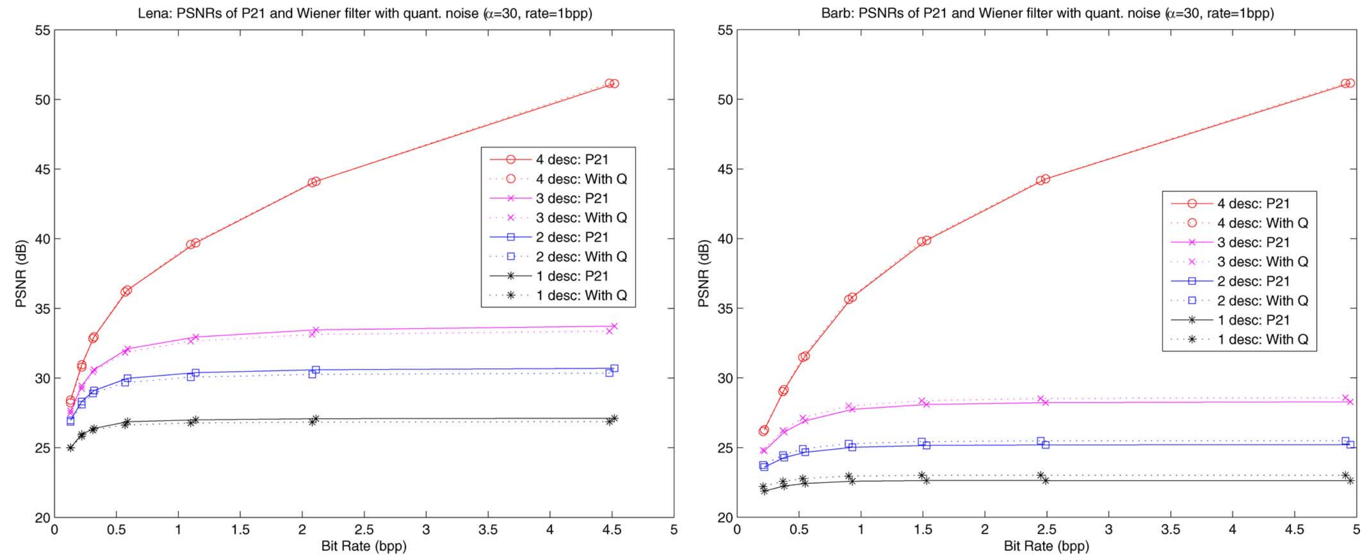
Next, we compare our method with the results in [15], where the maximally smooth recovery constraint is applied in the transform design. The re-designed transforms in [15] can achieve the same error resilience with lower bit rate than the transforms used in [14].

In Table IV, we use the 256×256 Lena image to compare the quality degradation behavior of our method with that of the MSR method in [15] when the error-free reconstruction quality is 33 dB. Although different entropy coding methods are used, we believe the performance degradations of the two methods with respect to the same error-free quality can be compared fairly. The first part in Table IV is taken from [15]. The second part is given by the $M \times 2M$ Wiener filter $\mathbf{H}_1^*$-based TDLT with the same coding gains as those in [15]. The average bit rate of our method is 19% lower than that in [15] (up to 33% in the case of M2). Note that this fact alone should not be used to judge against the MSR method since arithmetic coding is used in our method and adaptive Huffman coding is used in [15] (in fact, the actual bit rate saving is even higher since the table overhead for the adaptive Huffman code is not counted in [15]). In terms of error concealment quality, our average PSNR improvement over [15] is 0.74 dB, with 1.07 dB at the highest coding gain and 1.46 dB when only one description is available. These behaviors are also consistent with Table III, confirming the more graceful performance degradation of our method when more data are lost.

VII. Conclusion

This paper analyzes the error concealment design of the time-domain lapped transform from the perspective of estimation theory. The general LMMSE solution and various simplifications are proposed. Design examples and image coding results show that the reconstruction error can be reduced dramatically, compared to the mean reconstruction method and the maximally smooth recovery method. We also show that both the mean reconstruction method and the linear interpolation method are special cases of the proposed framework.

The performance of the proposed method can be further improved in several ways. First of all, the Wiener filters in this paper are designed based on 1-D signal model. Better performance can be achieved by designing 2-D Wiener filter directly.



(a) (b)

Fig. 7. Performance comparison between P21 (with quantization noise ignored) and a configuration without ignoring quantization noise. The two configurations have the same coding gain. (a) Results with the image Lena. (b) Results with the image Barbara.

TABLE III
PSNR (IN DECIBELS) UNDER DIFFERENT LOSSES FOR THE 256 $\times$ 256 LENA IMAGE WITHOUT QUANTIZATION ERROR

Available	Results in [14]				α in (38) ($\beta = 0$)				Average
Descriptions	T6	T7	T8	T9	69.0	74.0	112.0	143.0	Improvement (dB)
3	29.33	31.90	32.95	33.45	32.15	34.72	35.31	35.56	2.48
2	26.22	28.90	30.02	30.57	29.16	31.71	32.32	32.58	2.52
1	21.08	23.57	24.75	25.34	25.50	27.75	28.46	28.78	3.94
Avg. Improvement (dB)	-	-	-	-	3.39	3.27	2.79	2.52	2.98
Coding Gain (dB)	7.83	7.18	6.76	6.51	7.83	7.18	6.76	6.51	-

TABLE IV
PSNR (IN DECIBELS) 256 $\times$ 256 LENA IMAGE WHEN THE CODING DISTORTION IS 33 dB

Available	Results in [15]				α in (38) ($\beta = 0$)				Average
Descriptions	M2	M8	M49	M58	15.7	30.8	81.6	102.2	Improvement (dB)
4	33.05	33.07	33.04	33.04	33.05	33.07	33.04	33.04	-
3	26.99	28.99	30.96	31.11	27.62	29.37	31.00	31.20	0.29
2	24.37	26.77	29.47	29.67	25.22	27.35	29.65	29.93	0.47
1	20.20	22.34	25.57	26.14	21.94	24.26	26.74	27.13	1.46
Average Improvement (dB)	-	-	-	-	1.07	0.96	0.46	0.45	0.74
Bit Rate (bpp)	0.77	0.88	1.14	1.27	0.51	0.65	1.08	1.14	-19%
Coding Gain (dB)	8.96	8.26	7.08	6.85	8.96	8.26	7.08	6.85	-

Some preliminary result have been reported in [28]. Second, the closed-form Wiener solutions lend themselves naturally to adaptive error concealment. How to implement the adaptive algorithm with reasonable complexity is our ongoing work.

APPENDIX

In this section, we prove the persymmetric relationship in (27) for the $2M \times 2M$ Wiener filter $\mathbf{H}_0^*$ in (6) and the constrained solutions in (14) and (16). The proof for other simplified Wiener filers can be obtained similarly.

The expression of $\mathbf{H}_0^*$ is reproduced as follows:

$$\mathbf{H}_0^* = \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}. \tag{41}$$

Since $\mathbf{JJ} = \mathbf{I}$, $\mathbf{H}_0^*$ satisfies (27) if the following are true:

$$\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2} = \mathbf{J}\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}\mathbf{J} \tag{42}$$
$$\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1} = \mathbf{J}\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}\mathbf{J}. \tag{43}$$

Since $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}$ is a submatrix of $\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3}$, (42) can be obtained if $\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3} = \mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T = \mathbf{J}\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3}\mathbf{J}$, which, in turn, needs $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4} = \mathbf{J}_{4M}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{J}_{4M}$ and $\mathbf{P}_{34}^T = \mathbf{J}_{4M}\mathbf{P}_{34}^T\mathbf{J}_{3M}$. The former is obvious since $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ is a symmetric Toeplitz matrix. The latter is a direct result of $\mathbf{P} = \mathbf{JPJ}$, which has already been reinforced in the TDLT in order to obtain linear-phase filer bank [16].

Similarly, (43) can be established if $\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3} = \mathbf{J}\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3}\mathbf{J}$. Recall that $\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3}$ is given by

$$\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3} = \mathbf{P}_{34}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T + \mathbf{C}_3^T\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}\mathbf{C}_3. \quad (44)$$

The first part is persymmetric since both $\mathbf{P}_{34}$ and $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ are persymmetric. To show the persymmetry of the second part, we assume the subband quantization noises of different blocks are uncorrelated. Thus

$$\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3} = \mathrm{diag}\{\mathbf{R}_{\mathbf{qq}}, \mathbf{R}_{\mathbf{qq}}, \mathbf{R}_{\mathbf{qq}}\}. \quad (45)$$

The block noise correlation after the inverse DCT can be rewritten as

$$\mathbf{R}_{\mathbf{vv}} \triangleq \mathbf{C}^T\mathbf{R}_{\mathbf{qq}}\mathbf{C} = \sum_{i=0}^{M-1} \sigma_{\mathbf{q}_i}^2 \mathbf{c}_i^T\mathbf{c}_i \quad (46)$$

where $\mathbf{c}_i$ is the ith basis function of the DCT, which has linear phase. Therefore, $\mathbf{c}_i$ is either symmetric or anti-symmetric. Let

$$\mathbf{B}_i \triangleq \mathbf{c}_i^T\mathbf{c}_i. \quad (47)$$

By the symmetry of $\mathbf{c}_i$, we have

$$\begin{aligned}\mathbf{B}_i(j, k) &= \mathbf{c}_{i,j}\,\mathbf{c}_{i,k} = \mathbf{c}_{i,M-1-j}\,\mathbf{c}_{i,M-1-k} \\ &= \mathbf{B}_i(M-1-j, M-1-k). \end{aligned} \quad (48)$$

This shows that $\mathbf{B}_i$ is also persymmetric, which leads to the persymmetry of $\mathbf{R}_{\mathbf{vv}}$ and $\mathbf{C}_3^T\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}\mathbf{C}_3$, and, therefore, the persymmetry of the Wiener filter $\mathbf{H}_0^*$. The proof for other simplified Wiener filters can be conducted in a similar fashion.

Once the persymmetry of the Wiener filter is established, the persymmetry of the constrained Wiener solution in (14) becomes evident after substituting (15) into (14) and noticing that $\mathbf{dd}^T$ is persymmetric. The scaled Wiener filter in (16) is also persymmetric since $\mathbf{\Theta}(i, i) = \mathbf{\Theta}(L-1-i, L-1-i)$ for an $L \times N$ persymmetric Wiener filter.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their suggestions that have significantly enhanced the quality of this paper.

REFERENCES

[1] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: A review," *Proc. IEEE*, vol. 86, no. 5, pp. 974–997, May 1998.

[2] Y. Takishima, M. Wada, and H. Murakami, "Reversible variable length codes," *IEEE Trans. Commun.*, vol. 43, no. 2/3/4, pp. 158–162, Feb./Mar./Apr. 1995.

[3] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projections onto convex sets," *IEEE Trans. Image Process.*, vol. 4, no. 4, pp. 470–477, Apr. 1995.

[4] P. Salama, N. B. Shroff, E. J. Coyle, and E. J. Delp, "Error concealment techniques for encoded video streams," in *Proc. IEEE Conf. Image Processing*, Oct. 1995, vol. 1, pp. 9–12.

[5] J.-W. Suh and Y.-S. Ho, "Error concealment based on directional interpolation," *IEEE Trans. Consum. Electron.*, vol. 43, no. 3, pp. 295–302, Aug. 1997.

[6] J. W. Park and S. U. Lee, "Recovery of corrupted image data based on the NURBS interpolation," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1003–1008, Oct. 1999.

[7] Z. Alkachouh and M. G. Bellanger, "Fast DCT-based spatial domain interpolation of blocks in images," *IEEE Trans. Image Process.*, vol. 9, no. 4, pp. 729–732, Apr. 2000.

[8] X. Li and M. Orchard, "Novel sequential error concealment using orientation adaptive interpolation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 857–864, Oct. 2002.

[9] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.

[10] P. Haskell and D. Messerschmitt, "Reconstructing lost video data in a lapped orthogonal transform based coder," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Apr. 1990, vol. 4, pp. 1985–1988.

[11] R. L. de Queiroz and K. R. Rao, "Extended lapped transform in image coding," *IEEE Trans. Image Process.*, vol. 4, no. 6, pp. 828–832, Jun. 1995.

[12] 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.

[13] Y. Wang, Q.-F. Zhu, and L. Shaw, "Maximally smooth image recovery in transform coding," *IEEE Trans. Commun.*, vol. 41, no. 10, pp. 1544–1551, Oct. 1993.

[14] 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.

[15] ——, "Lapped orthogonal transform designed for error resilient image coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 9, pp. 752–764, Sep. 2002.

[16] 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.

[17] C. Tu and T. D. Tran, "Context based entropy coding of block transform coefficients for image compression," *IEEE Trans. Image Process.*, vol. 11, no. 11, pp. 1271–1283, Nov. 2002.

[18] S. Srinivasan, P. Hsu, T. Holcomb, K. Mukerjee, S. Regunathan, B. Lin, J. Liang, M. Lee, and J. Ribas-Corbera, "Windows media video 9: Overview and applications," *Signal Process.: Image Commun.*, vol. 19, no. 9, pp. 851–875, Oct. 2004.

[19] C. Tu, T. D. Tran, and J. Liang, "Error resilient pre-/post-filtering for DCT-based block coding systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Apr. 2003, vol. 3, pp. 273–276.

[20] ——, "Error resilient pre-/post-filtering for DCT-based block coding systems," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 30–39, Jan. 2006.

[21] J. Liang, C. Tu, T. D. Tran, and L. Gan, "Wiener filtering for generalized error resilient time domain lapped transform," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Mar. 2005, vol. 2, pp. 181–184.

[22] ——, "Error resilient DCT image coding with pre/post-filtering and wiener filtering," in *Proc. SPIE Conf. Visual Communications and Image Processing*, Jul. 2005, vol. 5960, pp. 1215–1225.

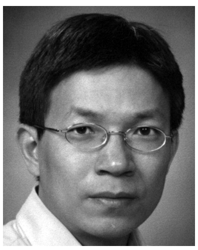
[23] V. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74–93, Sep. 2001.

[24] S. Haykin, *Adaptive Filtering*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[25] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[26] A. Cantoni and P. Butler, "Properties of the eigenvectors of persymmetric matrices with applications to communication theory," *IEEE Trans. Commun.*, vol. 24, no. 8, pp. 804–809, Aug. 1976.

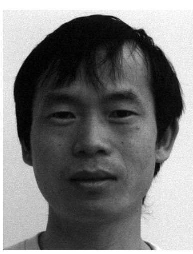
[27] A. K. Soman, P. P. Vaidyanathan, and T. Q. Nguyen, "Linear phase paraunitary filter banks: Theory, factorizations and designs," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3480–3496, Dec. 1993.
[28] J. Liang, X. Li, G. Sun, and T. D. Tran, "Two-dimensional wiener filters for error resilient time domain lapped transform," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Toulouse, France, May 2006, vol. III, pp. 241–244.



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.

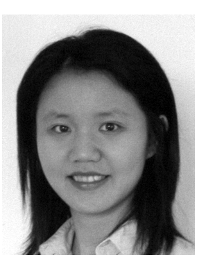
Since May 2004, he has been an Assistant Professor at the School of Engineering Science, Simon Fraser University, Burnaby, BC, Canada. From 2003 to 2004, he was with the Video Codec Group of Microsoft Digital Media Division. He was a summer intern at the same group in 2001. From 1997 to 1999, he was with Hewlett-Packard Singapore and the Center for Wireless Communications, NUS. His research interests include multimedia communications, multirate signal processing, and digital communications.

Dr. Liang is a member of the Multimedia Systems and Applications Technical Committee (MSATC) of the IEEE Circuits and Systems Society.



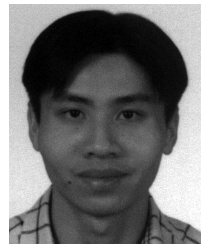
Chengjie Tu (S'02–M'04) received the B.E. and M.E. degrees from University of Science and Technology of China (USTC) in 1994 and 1997, respectively, the M.S.E and Ph.D. degrees from The Johns Hopkins University, Baltimore, MD, in 2002 and 2003, respectively.

He is currently with the Video Codec Group of Microsoft Digital Media Division, Microsoft Corporation, Redmond, WA. His current research interests include image/video compression, multirate signal processing, and error control and concealment for image/video communication.



Lu Gan (S'02–M'04) received the B.E. and M.E. degrees from Southeast University, Nanjing, China, in 1998 and 2000, respectively, and the Ph.D. degree from Nanyang Technological University, Singapore, in 2004, all in electrical and electronic engineering.

Since July 2006, She has been a Lecturer in the Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, U.K. From 2003 to 2004, she was with the Center for Signal Processing, Nanyang Technological University, Singapore. From 2004 to 2006, she was a Lecturer with the School of Electrical Engineering and Computer Science, University of Newcastle, NSW, Australia. Her research interests are in wavelets and filterbanks and their applications in image/video processing and communications.



Trac D. Tran (S'94–M'98) received the B.S. and M.S. degrees from the Massachusetts Institute of Technology, Cambridge, in 1993 and 1994, respectively, and the Ph.D. degree from the University of Wisconsin, Madison, in 1998, all in electrical engineering.

In July of 1998, he joined the Department of Electrical and Computer Engineering, The Johns Hopkins University, Baltimore, MD, where he currently holds the rank of Associate Professor. His research interests are in the field of digital signal processing, particularly in multirate systems, filter banks, transforms, wavelets, and their applications in signal analysis, compression, processing, and communications.

Dr. Tran was the co-director (with Prof. J. L. Prince) of the 33rd Annual Conference on Information Sciences and Systems (CISS'99), Baltimore, in March 1999. He received the National Science Foundation CAREER award in 2001. In the summer of 2002, he was an ASEE/ONR Summer Faculty Research Fellow at the Naval Air Warfare Center Weapons Division (NAWCWD), China Lake, CA. He currently serves as Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING as well as the IEEE TRANSACTIONS ON IMAGE PROCESSING.

Kai-Kuang Ma (S'80–M'84–SM'95) received the B.E. degree in electronic engineering from Chung Yuan Christian University, Chung-Li, Taiwan, R.O.C., the M.S. degree in electrical engineering from Duke University, Durham, NC, and the Ph.D. degree in electrical engineering from North Carolina State University, Raleigh.

He is currently a Professor of electrical and electronic engineering at Nanyang Technological University, Singapore. From 1992 to 1995, he was a Member of Technical Staff at the Institute of Microelectronics, National University of Singapore, working on digital video coding and the MPEG standards. From 1984 to 1992, he was with the IBM Corporation, Kingston, NY, and then Research Triangle Park, NC, engaging in various DSP and VLSI advanced product development. His research interests are in the areas of digital signal, image, and video processing. These include digital image/video coding and standards, wavelets and filter banks, joint source/channel image coding, denoising, super resolution, error resilience and concealment, and content-based image/video indexing and retrieval.

Dr. Ma is the Technical Program Co-Chair of two international conferences: the IEEE International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS) 2007 and the IEEE International Conference on Image Processing (ICIP) 2004. He currently serves as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, Associate Editor of the IEEE TRANSACTIONS ON MULTIMEDIA, and Editorial Board Member of the *Journal of Visual Communication and Image Representation* and the *International Journal of Image and Graphics*. He is an elected Technical Committee Member of the IEEE Signal Processing Society Image and Multidimensional Signal Processing (IMDSP) Technical Committee, the IEEE Communications Society Multimedia Communications Technical Committee, and the IEEE Circuits and Systems Society Digital Signal Processing (DSP) Technical Committee. He has been actively serving as Technical Program Committee member of numerous international conferences each year. He is Chairman of IEEE Signal Processing Singapore Chapter (2000 to 2002). He is Singapore MPEG Chairman and Head of Delegation (1997 to 2001) and the General Chair of organizing a series of international standard meetings and events held in Singapore (2001).