WIENER FILTERING FOR GENERALIZED ERROR RESILIENT TIME DOMAIN LAPPED TRANSFORM

Jie Liang[1], *Chengjie Tu*[2], *Trac D. Tran*[3] *and Lu Gan*[4]

[1]Engineering Science, Simon Fraser University, Burnaby, BC, Canada, V5A 1S6
[2]Microsoft Corporation, Redmond, WA, USA, 98052
[3]ECE Department, The Johns Hopkins University, Baltimore, MD, USA 21218
[4]School of EECS, The University of Newcastle, Callaghan, NSW, Australia, 2308

ABSTRACT

In this paper, we revisit the design of time-domain lapped transform for error resilient image transmission. A general structure is first proposed whose solution is given by a Wiener filter. Two simplified schemes with different tradeoffs between complexity and performance are then developed, for which Wiener filter solutions also exist. We will show that the existing method is a special case of the general scheme. Design examples and image coding experiments verify that the performance of our new approach is significantly better than existing techniques.

1. 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 decoder side by estimating the lost data with methods such as interpolation and projection onto convex sets [3, 4]. Other approaches tackle the problem by a joint design of the encoder and decoder, for which the lapped transform is a powerful tool [5, 6, 7].

In the original lapped transform [5], a postfilter is applied at block boundaries after the DCT. The postfilter can be designed to remove the remaining redundancy between neighboring blocks, and thereby improving the coding efficiency of the DCT and reducing the blocking artifact associated with DCT-based schemes. Notice, however, that the postfilter also spreads out the information of a block to its neighboring blocks. This property is used in [6, 7] to achieve error concealment, where the lost blocks can be better recovered by a judicious design of the forward and inverse lapped transform.

This work is supported by Simon Fraser University (SFU) Start-up Grant, SFU President's Research Grant, NSF CAREER CCF-0093262, and NSF CCF-0430869. Author emails: jiel@sfu.ca, chentu@microsoft.com, trac@jhu.edu, lugan@ecemail.newcastle.edu.au.

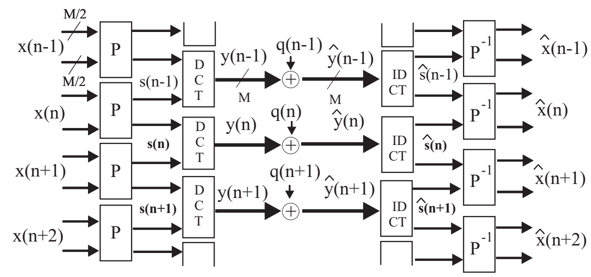

Fig. 1. Forward and inverse time-domain lapped transform.

Recently, a new family of lapped transform, called time-domain lapped transform (TDLT) [8], has been developed. In TDLT, a prefilter at block boundaries is applied before the DCT. This makes it more compatible to existing DCT-based infrastructures, because dramatic improvement can be achieved with minimum changes to existing software or hardware implementations.

The new lapped transform has also been applied to error concealment [9, 10]. More flexibilities and better performance have been demonstrated. In particular, the decoder can invoke two postfilters - one for perfectly received blocks and another for lost blocks to improve the visual quality.

In [6, 9, 10], the lost coefficient blocks are simply estimated by averaging its neighboring blocks. This greatly limits the error concealment capability of the lapped transform. In [7], a maximally smoothness recovery method is used, but its complexity is increased significantly. In this paper, we first present a general framework of TDLT-based error concealment, and give the corresponding Wiener filter solution. Several simplifications of the general structure and their optimal solutions are then developed. The existing averaging method is revealed to be a very special case of the general solution. Compared to the method in [9, 10], the reconstruction error can be reduced by as much as 80% given the same compression capability, and more than 4 dB improvement can be achieved in image coding experiments.

2. GENERAL PRE/POSTFILTERING STRUCTURE FOR ERROR CONCEALMENT

Fig. 1 illustrates the implementation of the time-domain lapped transform for a system with block size of M. An $M \times M$ prefilter $\mathbf{P}$ is applied at the boundary of two neighboring blocks be-

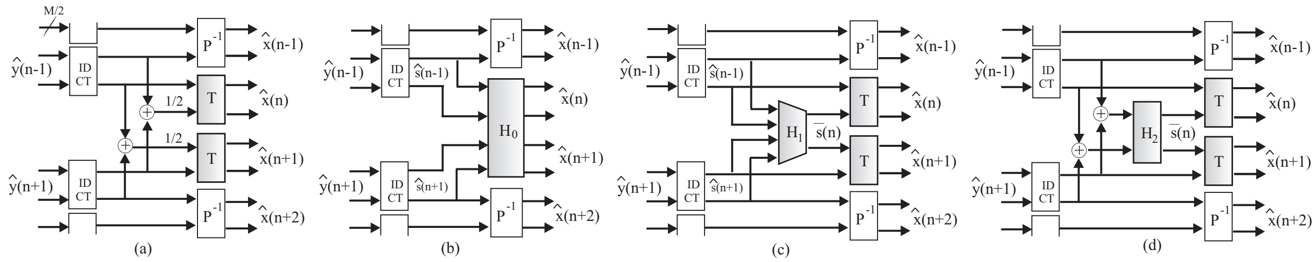

Fig. 2. (a) Existing decoder side error concealment design; (b) General structure for error concealment; (c) A special case of (b); (d) Further simplification of (c).

fore the DCT. As a result, the basis functions of the forward transform cover two blocks, and better compression performance can be achieved by exploiting the correlation between neighboring blocks. Correspondingly, a postfilter is applied by the decoder at each block boundary after inverse DCT. The postfilter is simply the inverse of the prefilter if there is no transmission error. In what follows, we use $\mathbf{x}(i)$, $\mathbf{s}(i)$, $\mathbf{y}(i)$ and $\mathbf{q}(i)$ to denote the i-th block of prefilter input, DCT input, DCT output, and quantization noise, respectively.

The structure of the time-domain lapped transform allows an effective strategy for error concealment. In this paper, we assume that the DCT coefficients of a block are either received perfectly or lost entirely. In [9, 10], the lost block is estimated by averaging the received neighboring blocks. This mean reconstruction method has also been used in [6, 7]. It is shown in [9, 10] that the pre- and postfilter can be jointly designed such that the reconstructed quality is still satisfactory in the case of transmission error. Moreover, two postfilters can be designed - one for perfectly received blocks and another one for lost blocks. The overall scheme is shown in Fig. 2 (a).

However, the approach in [9, 10] is only a special case of a more general framework. To see this, notice that when $\hat{\mathbf{y}}(n)$ is lost, the error concealment problem can be viewed as the estimation of $\hat{\mathbf{x}}(n)$ and $\hat{\mathbf{x}}(n+1)$ 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 solution should be a $2M \times 2M$ matrix $\mathbf{H}_0$, as shown in Fig. 2 (b).

Define

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

The auto-correlation of the reconstruction error is given by

$$\mathbf{R}_{ee} = E\{(\mathbf{H}_0\hat{\mathbf{s}}_2 - \mathbf{x}_2)(\mathbf{H}_0\hat{\mathbf{s}}_2 - \mathbf{x}_2)^T\}. \tag{2}$$

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

$$\mathcal{E} = \frac{1}{2M} trace\{\mathbf{R}_{ee}\}. \tag{3}$$

The optimal solution is given by the Wiener filter

$$\mathbf{H}_0^* = \mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}. \tag{4}$$

Matrices $\mathbf{R}_{\mathbf{x}_2\hat{\mathbf{s}}_2}$ and $\mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}$ in (4) can be obtained as follows. First of all, let us partition $\mathbf{P}$ into

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_0 \\ \mathbf{P}_1 \end{bmatrix}, \tag{5}$$

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

$$\hat{\mathbf{s}}_3 = \mathbf{s}_3 + \mathbf{C}_3^T\mathbf{q}_3 = \mathbf{P}_{34}\mathbf{x}_4 + \mathbf{C}_3^T\mathbf{q}_3, \tag{6}$$

where $\mathbf{P}_{34} = diag\{\mathbf{P}_1,\, \mathbf{P},\, \mathbf{P},\, \mathbf{P}_0\}$. From (6) we get

$$
\begin{aligned}
\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3} &= \mathbf{R}_{\mathbf{x}_4\mathbf{s}_3} = \mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T, \\
\mathbf{R}_{\mathbf{s}_3\hat{\mathbf{s}}_3} &= \mathbf{R}_{\mathbf{s}_3\mathbf{s}_3} = \mathbf{P}_{34}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T, \\
\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3} &= \mathbf{P}_{34}\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}\mathbf{P}_{34}^T + \mathbf{C}_3^T\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}\mathbf{C}_3.
\end{aligned} \tag{7}
$$

$\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ and $\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}$ in (7) are correlation matrices of $\mathbf{x}_4$ and $\mathbf{q}_3$, respectively, and it is assumed 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 (4) can be obtained from sub-matrices of $\mathbf{R}_{\mathbf{x}_4\hat{\mathbf{s}}_3}$ and $\mathbf{R}_{\hat{\mathbf{s}}_3\hat{\mathbf{s}}_3}$ once $\mathbf{R}_{\mathbf{x}_4\mathbf{x}_4}$ and $\mathbf{R}_{\mathbf{q}_3\mathbf{q}_3}$ are known.

When applied to image error concealment, an important requirement is that each estimated block has the same brightness as its neighbors. One way to achieve this is to make the sum of every row of the filter to be unity. In this paper, this is ensured by a simple normalization of the Wiener filter. Our results show that its effect on MSE is negligible.

We point out here that all Wiener filters derived in this paper have special structures, and can be factorized for fast implementations. The details are omitted due to space limitations.

Although the solution in (4) can be applied directly, several simplified structures can be developed to achieve different tradeoffs between complexity and performance. In particular, we are interested in finding the optimal solution for the two-stage approach studied in [9, 10], where the lost block is estimated first before applying postfilter. Two such simplifications are presented next, and their relationship with existing method will be discussed.

3. SIMPLIFIED PRE/POSTFILTERING STRUCTURES

In the two-stage error concealment method, we first estimate $\mathbf{s}(n)$ by an $M \times 2M$ matrix $\mathbf{H}_1$, *i.e.*,

$$\bar{\mathbf{s}}(n) = \mathbf{H}_1\hat{\mathbf{s}}_2. \tag{8}$$

Table 1. Design Examples of Different Configurations

	Without switching postfilter								Switching postfilter							
Cfg.	P1	P10	P11	P12	P2	P20	P21	P22	P3	P30	P31	P32	P4	P40	P41	P42
α	-	92	92	130	-	28	28	20	-	10.5	10.5	15.5	-	1	1	1
β	-	0	0	0	-	0.6	0.6	2	-	0.5	0.5	0.5	-	0	0.15	0.15
G_{TC}	6.96	6.96	6.96	6.95	8.41	8.42	8.42	8.43	9.17	9.17	9.18	9.19	9.61	9.61	9.61	9.61
G_R	0.67	0.75	0.75	0.94	0.64	0.67	0.67	0.84	0.59	0.57	0.60	0.37	0.62	0.34	0.63	0.69
MSE	0.14	0.03	0.03	0.09	0.15	0.06	0.06	0.12	0.16	0.11	0.11	0.13	0.21	0.17	0.19	0.19

The postfilter is then applied as usual. A different postfilter can be designed to improve the visual quality. The structure is shown in Fig. 2 (c). Clearly, it only covers a subset of the general structure in Fig. 2 (b), since $\mathbf{H}_0$ is restricted to be

$$\mathbf{H}_0 = \begin{bmatrix} \mathbf{T}_M & \mathbf{0}_M \\ \mathbf{0}_M & \mathbf{T}_M \end{bmatrix} \begin{bmatrix} \mathbf{I}_1 \\ \mathbf{H}_1 \\ \mathbf{I}_2 \end{bmatrix}, \tag{9}$$

where $\mathbf{I}_1$ and $\mathbf{I}_2$ are defined by

$$\begin{aligned} \mathbf{I}_1 &= \begin{bmatrix} \mathbf{0}_{\frac{M}{2}} & \mathbf{I}_{\frac{M}{2}} & \mathbf{0}_{\frac{M}{2}} & \mathbf{0}_{\frac{M}{2}}, \end{bmatrix}, \\ \mathbf{I}_2 &= \begin{bmatrix} \mathbf{0}_{\frac{M}{2}} & \mathbf{0}_{\frac{M}{2}} & \mathbf{I}_{\frac{M}{2}} & \mathbf{0}_{\frac{M}{2}}, \end{bmatrix}. \end{aligned} \tag{10}$$

When applied to $\hat{\mathbf{s}}_2$, $\mathbf{I}_1$ simply extracts the second half of $\hat{\mathbf{s}}(n-1)$ and $\mathbf{I}_2$ extracts the first half of $\hat{\mathbf{s}}(n+1)$.

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

$$\mathcal{E}_1 = \frac{1}{M} trace\{E\{(\mathbf{H}_1\hat{\mathbf{s}}_2 - \mathbf{s}(n))(\mathbf{H}_1\hat{\mathbf{s}}_2 - \mathbf{s}(n))^T\}\}, \tag{11}$$

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

$$\mathbf{H}_1^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_2} \mathbf{R}_{\hat{\mathbf{s}}_2\hat{\mathbf{s}}_2}^{-1}, \tag{12}$$

where $\mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_2}$ is a sub-matrix of $\mathbf{R}_{\mathbf{s}_3\hat{\mathbf{s}}_3}$. Since $\mathbf{H}_1^*$ is only a function of the input and prefilter, the postfilter can be optimized to improve the visual quality, as in [9, 10].

The complexity of the special case in (9) and Fig. 2 (c) can be further reduced by imposing the following constraint on $\mathbf{H}_1$:

$$\mathbf{H}_1 = \begin{bmatrix} \mathbf{H}_2 & \mathbf{H}_2 \end{bmatrix} = \mathbf{H}_2 \begin{bmatrix} \mathbf{I}_M & \mathbf{I}_M \end{bmatrix}, \tag{13}$$

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

$$\bar{\mathbf{s}}(n) = \mathbf{H}_2(\hat{\mathbf{s}}(n-1) + \hat{\mathbf{s}}(n+1)) \triangleq \mathbf{H}_2\hat{\mathbf{s}}_a. \tag{14}$$

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

$$\mathbf{H}_2^* = \mathbf{R}_{\mathbf{s}(n)\hat{\mathbf{s}}_a} \mathbf{R}_{\hat{\mathbf{s}}_a\hat{\mathbf{s}}_a}^{-1}. \tag{15}$$

The matrices involved can be obtained from (7) by simple manipulations.

It is clear from (13) that the mean reconstruction method used in [9, 10] is simply a special case of the already sub-optimal approach in (13) with $\mathbf{H}_2 = \frac{1}{2}\mathbf{I}_M$. Therefore it can be expected that the error concealment performance can be improved considerably if $\mathbf{H}_2^*$, $\mathbf{H}_1^*$ or $\mathbf{H}^*$ are used.

4. DESIGN EXAMPLES AND APPLICATIONS

In this section, we show design examples of different error concealment schemes. The following criteria are considered: coding gain of the perfect reconstruction system, the MSE (3) in the case of data loss, and the reconstruction gain. Among them, the coding gain is a measure of the compression capability, which is given by

$$G_{TC} \triangleq 10\log_{10} \frac{\sigma_x^2}{\left(\prod_{i=0}^{M-1} \sigma_{y_i}^2 ||f_i||^2\right)^{\frac{1}{M}}}, \tag{16}$$

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

The reconstruction gain measures the distribution of reconstruction error when a block is lost. It is defined as [6, 10]

$$G_R = \frac{\left(\prod_{i=0}^{2M-1} \sigma_{e_i}^2\right)^{\frac{1}{2M}}}{\frac{1}{2M}\sum_{i=0}^{2M-1} \sigma_{e_i}^2}, \tag{17}$$

where $\sigma_{e_i}^2$ is the i-th diagonal entry of $\mathbf{R}_{ee}$ in (2).

The final objective function is [10]

$$J = G_{TC} - \alpha\,\mathcal{E} + \beta\,G_R. \tag{18}$$

Different solutions can be obtained by varying α and β and maximizing the function above. In this paper, the quantization noise is ignored in Wiener filter expressions. Four families are obtained with Matlab 6.5 and the results are summarized in Table 1, where the configurations P1 to P4 are obtained from [10] with the mean reconstruction method ($\mathbf{H}_2 = \frac{1}{2}\mathbf{I}$). In the first two families, the postfilter is simply the inverse of the prefilter, whereas dynamic switching of two postfilters are used in the last two families. Wiener filters in (4), (12) and (15) are used to obtain configurations Pi0, Pi1, and Pi2 ($i = 1, \ldots, 4$), respectively. All solutions 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 enjoys the best resilience to error. On the contrary, the last group has the highest coding gain and the worst robustness to transmission loss.

It is clear from Table 1 that the MSE of error concealment can be reduced substantially by the proposed methods. The improvement becomes more prominent as coding gain decreases, since there is more correlation among neighboring blocks. For example, compared with P1, the MSE is reduced by an astonishing 80% by P10 and P11. It can still be reduced by 10-20% even when the coding gain is at its highest value, as in the P4 family. The table

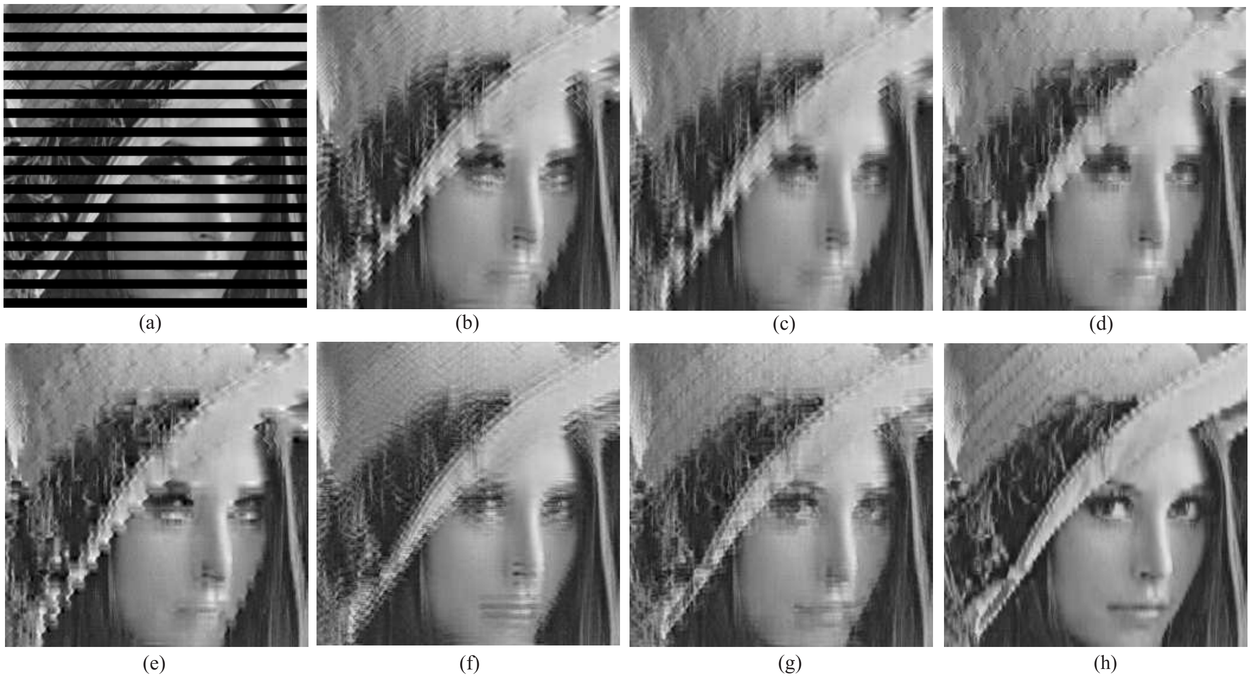

Fig. 3. Decoding results with 50% slice loss. (a) Loss pattern; (b) Original TDLT postfilter (24.34 / 40.06 dB); (c) P4 (24.40 / 40.06 dB); (d) P42 (24.64 / 40.09 dB); (e) P41 (24.73 / 40.09 dB); (f) P2 (25.99 / 38.24 dB); (g) P22 (26.74 / 38.27 dB); (h) P21 (30.11 / 38.88 dB).

also shows that the MSE given by (12) is very close to that of the general solution (4) in most cases.

Fig. 3 summarizes portions of the decoded images with different error concealment approaches. The 512×512 Lena image is coded at 1 bit/pixel by the L-CEB coder [11]. Data loss is assumed to be in slice mode, and the loss rate is chosen to be 50%, as illustrated in Fig. 3 (a). Mean reconstruction method is used in Fig. 3 (b), (c) and (f) for lost blocks, and various Wiener filters are used in other cases. The PSNRs of each method with and without data loss are included in the description of Fig. 3. Vertical neighbors are used to estimate the lost blocks. Family P4 only improves the PSNR slightly because of the limited correlation among neighboring blocks. However, family P2 makes a huge difference. The PSNR of P21 is 5.8 dB higher than the standard TDLT and 4.1 dB higher than P2. Notice that the coding performance of P21 is only sacrificed by 1.2 dB compared to the best TDLT, making it very promising for practical application.

5. CONCLUSION

This paper analyzes the error concealment problem of the time-domain lapped transform from the perspective of estimation theory. The general MMSE solution and various simplifications are proposed. Design examples and image coding results show that the reconstruction error can be reduced dramatically. The closed-form Wiener solutions also lend themselves naturally to adaptive error concealment, which is a topic of our ongoing research.

6. REFERENCES

[1] Y. Wang and Q. F. Zhu, "Error control and concealment for video communication: a review," *Proceedings of IEEE*, vol. 86(5), pp. 974–997, May 1998.

[2] Y. Takishima, M. Wada, and H. Murakami, "Reversible variable length codes," *IEEE Trans. Commun.*, vol. 43(234), pp. 158–162, 1995.

[3] H. Sun and W. Kwok, "Concealment of damaged block transform coded images using projections onto convex sets," *IEEE Trans. Image Proc.*, vol. 4(4), pp. 470–477, Apr. 1995.

[4] J. W. Park and S. U. Lee, "Recovery of corrupted image data based on the NURBS interpolation," *IEEE Trans. Image Proc.*, vol. 9(7), pp. 1003–1008, Oct. 1999.

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

[6] S. S. Hemami, "Reconstruction-optimized lapped transforms for robust image transmission," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 6 (2), pp. 168–181, Apr. 1996.

[7] D. Chung and Y. Wang, "Lapped orthogonal transform designed for error resilient image coding," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 12(9), pp. 752–764, Sept. 2002.

[8] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-processing," *IEEE Trans. Signal Proc.*, vol. 51 (6), pp. 1557–1571, Jun. 2003.

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

[10] C. Tu, T. D. Tran, and J. Liang, "Error resilient pre-/post-filtering for DCT-based block coding systems," *IEEE Trans. Image Proc.*, to appear.

[11] C. Tu and T. D. Tran, "Context based entropy coding of block transform coefficients for image compression," *IEEE Trans. Image Proc.*, vol. 11(11), pp. 1271–1283, Nov. 2002.