

# TWO-DIMENSIONAL WIENER FILTERS FOR ERROR RESILIENT TIME DOMAIN LAPPED TRANSFORM

Jie Liang<sup>1</sup>, Xin Li<sup>2</sup>, Guoqian Sun<sup>1</sup>, and Trac D. Tran<sup>3</sup>

<sup>1</sup>School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6 Canada

<sup>2</sup>Lane Department of CSEE, West Virginia University, Morgantown, WV 26506, USA

<sup>3</sup>ECE Department, The Johns Hopkins University, Baltimore, MD, 21218, USA

## ABSTRACT

This paper presents the design of two-dimensional Wiener filters for error resilient time domain lapped transform. Two solutions are discussed, and a multi-pass approach is also proposed to make the algorithm adaptive to input statistics. Design examples and image coding experiments show that the adaptive 2-D Wiener filters provide significant improvement over the existing 1-D Wiener filtering method.

## 1. INTRODUCTION

Error concealment is an important technique to mitigate the effect of transmission error in multimedia communications. Among the error concealment algorithms that have been proposed [1], some methods such as the reversible variable length coding [2] introduce error resilience at the encoder. Some of them estimate the lost data at the decoder by virtue of interpolation or 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 provides a powerful and flexible platform [5].

The conventional lapped transform applies a postfilter after the DCT to remove the remaining redundancy between neighboring blocks and improve the coding efficiency of the DCT. In [6], it was realized that the lapped transform can also be designed to introduce redundancy to the transformed coefficients, thereby enhancing the error resilience of the system. In [7], this approach is used to generate multiple description coding (MDC), where the transformed image is split into various subimages, and each subimage is entropy coded independently to obtain one description. A maximum smoothness recovery design criterion is used in [8] to improve the error concealment quality, at the cost of increased complexity.

Recently, a new family of lapped transform named the time-domain lapped transform (TDLT) [9] is developed by applying a prefilter at block boundaries before the DCT, making it more compatible to existing DCT-based infrastructures. The application of the TDLT in error concealment shows that more flexibilities and better performance can be achieved [10]. However, the mean average reconstruction method proposed in [6] is still used in [10], where the lost coefficient blocks are simply estimated by averaging its neighboring blocks.

In [11], we show that the estimation error of the lost TDLT coefficients can be minimized by applying the Wiener filter, which

yields up to 80% error reduction compared to the mean reconstruction method in [10]. More than 4 dB improvement can be achieved in image coding experiments.

Two issues still exist in [11]. First of all, the design of the Wiener filter is based on one-dimensional (1-D) signal model, and the result is applied to two-dimensional (2-D) signal via the classic separate approach, *i.e.*, each row of the lost block is first estimated from its horizontal neighbors. A vertical estimate is then obtained for each column, and the average of the two estimates is used as the final result. This simple treatment does not exploit the 2-D geometric structures of the input and can create some artifacts, especially near the edges. In this paper, we generalize the approach in [11] and derive the 2-D Wiener filter for the estimation of lost transformed coefficients. Another limitation of [11] is that it uses a fixed Wiener filter. In this paper, we will show how to introduce adaptivity to this framework.

The implementation of TDLT for image coding is illustrated in Fig. 1 (a), which only shows the operations required to obtain the transformed coefficients of one block of size  $M \times M$  (in the center of the figure). The encoder first applies an  $M \times M$  prefilter at all block boundaries, first in each row and then in each column, or vice versa. After that, the 2-D DCT is applied to each block. As a result, each block of TDLT coefficients is a function of the input samples in a  $2M \times 2M$  region that spans nine neighboring blocks. The inverse of the two steps are employed at the decoder.

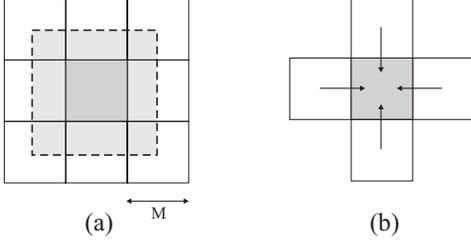
If some coefficient blocks are lost during transmission, the decoder can first apply the inverse DCT to all correctly received blocks, then estimate each of the lost block from its neighboring blocks, before applying the postfiltering. In the 1-D model considered in [6, 10, 11], only two neighboring blocks are required in the estimation. In the 2-D case, it can be seen from Fig. 1 that the estimation error in each block propagates to nine blocks (more precisely a  $2M \times 2M$  region). Therefore in this paper we limit our attention to the estimation of the lost block from the eight neighboring blocks. To reduce the complexity, we consider two special cases that involve different amount of neighboring information.

## 2. 2-D ESTIMATION FROM 4-CONNECTION NEIGHBORS

In this section, we consider the estimation of a lost block from its four immediate neighboring blocks, as shown in Fig. 1 (b). The estimation is performed after applying the inverse DCT to all received blocks. Suppose  $\mathbf{S}_0$  is the inverse DCT result of a lost TDLT coefficient block of size  $M \times M$ , and  $\mathbf{S}_i$ ,  $i = 1, \dots, 4$  the inverse DCT of the top, bottom, left, and right neighbor of  $\mathbf{S}_0$ , respectively, as shown in Fig. 1 (b). If we use  $\tilde{\mathbf{S}}_i$ ,  $i = 0, \dots, 4$  to represent the  $M^2 \times 1$  vector obtained by stacking the  $M$  columns of block  $\mathbf{S}_i$  together, all

---

This work is supported by NSERC Discovery Grant, SFU Presidential Research Grant, NSF CAREER Grant CCR-0093262, and NSF Grant CCF-0430869. Author emails: jiel@sfu.ca, xin.li@ieee.org, guoqians@sfu.ca, trac@jhu.edu.



**Fig. 1.** (a) 2-D implementation of TDLT. All gray areas are affected by the pre/post-filtering, and the DCT/IDCT is applied to the dark gray area. (b) Estimation of the lost block from four neighbors.

coefficients of the 4-connection neighbors can be put into a  $4M^2 \times 1$  vector as

$$\tilde{\mathbf{S}}_{4C} \triangleq [\tilde{\mathbf{S}}_1^T, \tilde{\mathbf{S}}_2^T, \tilde{\mathbf{S}}_3^T, \tilde{\mathbf{S}}_4^T]^T. \quad (1)$$

Our first objective is to find an  $M^2 \times 4M^2$  linear filter  $\mathbf{H}_1$  such that  $\hat{\tilde{\mathbf{S}}}_0 = \mathbf{H}_1 \tilde{\mathbf{S}}_{4C}$  is the optimal linear estimate of  $\tilde{\mathbf{S}}_0$  in terms of MSE. By the orthogonality principle, the optimal prediction error is uncorrelated with the observation, *i.e.*,

$$E\{(\mathbf{H}_1 \tilde{\mathbf{S}}_{4C} - \tilde{\mathbf{S}}_0)(\mathbf{H}_1 \tilde{\mathbf{S}}_{4C} - \tilde{\mathbf{S}}_0)^T\} = \mathbf{0}. \quad (2)$$

The optimal solution is the 2-D Wiener filter

$$\mathbf{H}_1 = \mathbf{R}_{\tilde{\mathbf{S}}_0 \tilde{\mathbf{S}}_{4C}} \mathbf{R}_{\tilde{\mathbf{S}}_{4C} \tilde{\mathbf{S}}_{4C}}^{-1}. \quad (3)$$

The correlation matrices involved above can be obtained as the follows. As shown in Fig. 1 (a), each input block  $\mathbf{S}_i$  to the DCT is a function of  $2M \times 2M$  input samples  $\mathbf{X}_i$ . Let  $\mathbf{P}$  be the  $M \times M$  prefilter, and  $\mathbf{P}_0, \mathbf{P}_1$  its first  $M/2$  rows and the second  $M/2$  rows, respectively, the relationship between  $\mathbf{X}_i$  and  $\mathbf{S}_i$  is thus

$$\mathbf{S}_i = \mathbf{P}_{12} \mathbf{X}_i \mathbf{P}_{12}^T, \quad (4)$$

where  $\mathbf{P}_{12}$  is an  $M \times 2M$  matrix given by

$$\mathbf{P}_{12} \triangleq \text{diag}\{\mathbf{P}_1, \mathbf{P}_0\}. \quad (5)$$

Denote  $\otimes$  the Kronecker product, (4) can be turned into the 1-D expression

$$\tilde{\mathbf{S}}_i = (\mathbf{P}_{12} \otimes \mathbf{P}_{12}) \tilde{\mathbf{X}}_i \triangleq \tilde{\mathbf{P}}_{12} \tilde{\mathbf{X}}_i, \quad (6)$$

where  $\tilde{\mathbf{X}}_i$  is obtained from  $\mathbf{X}_i$  by stacking its columns. By the definition of  $\tilde{\mathbf{S}}_{4C}$  in (1), we have

$$\begin{aligned} \tilde{\mathbf{S}}_{4C} &= \text{diag}\{\tilde{\mathbf{P}}_{12}, \tilde{\mathbf{P}}_{12}, \tilde{\mathbf{P}}_{12}, \tilde{\mathbf{P}}_{12}\} [\tilde{\mathbf{X}}_1^T, \tilde{\mathbf{X}}_2^T, \tilde{\mathbf{X}}_3^T, \tilde{\mathbf{X}}_4^T]^T \\ &\triangleq \tilde{\mathbf{P}}_{12,4} \tilde{\mathbf{X}}_{4C}. \end{aligned} \quad (7)$$

From (6) and (7), we get

$$\begin{aligned} \mathbf{R}_{\tilde{\mathbf{S}}_0 \tilde{\mathbf{S}}_{4C}} &= \tilde{\mathbf{P}}_{12} \mathbf{R}_{\tilde{\mathbf{X}}_0 \tilde{\mathbf{X}}_{4C}} \tilde{\mathbf{P}}_{12,4}^T, \\ \mathbf{R}_{\tilde{\mathbf{S}}_{4C} \tilde{\mathbf{S}}_{4C}} &= \tilde{\mathbf{P}}_{12,4} \mathbf{R}_{\tilde{\mathbf{X}}_{4C} \tilde{\mathbf{X}}_{4C}} \tilde{\mathbf{P}}_{12,4}^T. \end{aligned} \quad (8)$$

These matrices can be obtained once the 2-D correlation matrix of the input is known. Further discussions on the input statistics are given in Sec. 5.

### 3. FINAL MSE AND JOINT OPTIMIZATION

Our ultimate goal is to reduce the reconstruction error after the post-filtering, rather than after the Wiener filtering of the lost blocks. When quantization noise is ignored, the final reconstruction error in error concealment is solely caused by the estimation error  $\mathbf{v}_0 = \mathbf{S}_0 - \hat{\mathbf{S}}_0$ . After post-filtering, the  $M \times M$  estimation error propagates to a  $2M \times 2M$  region. The reconstruction error in this region is given by

$$\mathbf{e} = \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_0 \end{bmatrix} \mathbf{v}_0 \begin{bmatrix} \mathbf{T}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_0 \end{bmatrix}^T \triangleq \mathbf{T}_{21} \mathbf{v}_0 \mathbf{T}_{21}^T, \quad (9)$$

where  $\mathbf{T}_0$  and  $\mathbf{T}_1$  are the left half and the right half of the post-filter  $\mathbf{T}$ , respectively, *i.e.*,  $\mathbf{T} = [\mathbf{T}_0, \mathbf{T}_1]$ .

The equation above can be turned into 1-D by the Kronecker product as

$$\tilde{\mathbf{e}} = (\mathbf{T}_{21} \otimes \mathbf{T}_{21}) \tilde{\mathbf{v}}_0 \triangleq \tilde{\mathbf{T}}_{21} \tilde{\mathbf{v}}_0. \quad (10)$$

Thus

$$\mathbf{R}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}} = \tilde{\mathbf{T}}_{21} \mathbf{R}_{\tilde{\mathbf{v}}_0 \tilde{\mathbf{v}}_0} \tilde{\mathbf{T}}_{21}^T. \quad (11)$$

Each diagonal entry of  $\mathbf{R}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}$  corresponds to the MSE of one pixel in the  $2M \times 2M$  region. We can then extract all diagonal entries from  $\mathbf{R}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}$  and regroup them into a  $2M \times 2M$  matrix, which represents the distribution of the final reconstruction error in the  $2M \times 2M$  region. We define the overall MSE of the  $2M \times 2M$  region as

$$\mathcal{E} = 1/(4M^2) \text{trace}\{\mathbf{R}_{\tilde{\mathbf{e}}\tilde{\mathbf{e}}}\}. \quad (12)$$

An optimization program can be set up to obtain the TDLT pre- and post-filters so that the system achieves different tradeoffs between the error resilience and the compression capability. The objective function of the optimization is defined to be:

$$J = G_{TC} - \alpha \mathcal{E}, \quad (13)$$

where  $G_{TC}$  is the coding gain of the TDLT when there is no transmission error. Maximizing the objective function with different value of  $\alpha$  will lead to different solutions. Other criteria can also be included in the objective function. For example, we can explicitly control the distribution of the error in the  $2M \times 2M$  region in order to obtain smooth transition from healthy blocks to concealed blocks.

### 4. 2-D ESTIMATION FROM 8-CONNECTION NEIGHBORS

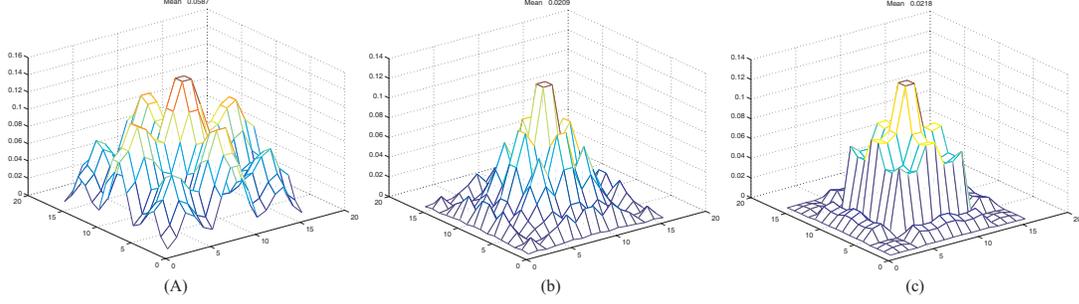
The 2-D Wiener filter in Sec. 2 is based on coefficients in the four immediate neighboring blocks. In this section, we consider 2-D Wiener filter that involves information in all eight neighboring blocks. To reduce the complexity, we investigate the Wiener filter based on  $M/2$  layers of coefficients around the lost block, as shown by the light gray area in Fig. 1 (a). The number of neighboring coefficients in this region is  $3M^2$ , which is less than the  $4M^2$  used in the last section.

Denote  $\tilde{\mathbf{S}}_{8C}$  the vector obtained by stacking all of these  $3M^2$  neighboring coefficients, it is straightforward to show that the 2-D Wiener filter for the estimation of the lost vector  $\tilde{\mathbf{S}}_0$  is

$$\mathbf{H}_2 = \mathbf{R}_{\tilde{\mathbf{S}}_0 \tilde{\mathbf{S}}_{8C}} \mathbf{R}_{\tilde{\mathbf{S}}_{8C} \tilde{\mathbf{S}}_{8C}}^{-1}. \quad (14)$$

To find the expression of  $\mathbf{R}_{\tilde{\mathbf{S}}_0 \tilde{\mathbf{S}}_{8C}}$  and  $\mathbf{R}_{\tilde{\mathbf{S}}_{8C} \tilde{\mathbf{S}}_{8C}}$ , we first write the  $2M \times 2M$  prefiltered region as

$$\mathbf{S}_{2M} = \text{diag}\{\mathbf{P}, \mathbf{P}\} \mathbf{X}_{2M} \text{diag}\{\mathbf{P}, \mathbf{P}\}^T \triangleq \mathbf{P}_{2M} \mathbf{X}_{2M} \mathbf{P}_{2M}^T. \quad (15)$$



**Fig. 2.** 2-D error distributions of various error resilient TDLTs with similar coding gains; (a) by 1-D Wiener filter (MSE: 0.0587); (b) by 4-connection 2-D Wiener filter (MSE: 0.0209); (c) by 8-connection 2-D Wiener filter (MSE: 0.0218).

With the help of the Kronecker product, this can be converted into

$$\tilde{\mathbf{S}}_{2M} = \tilde{\mathbf{P}}_{2M} \tilde{\mathbf{X}}_{2M}. \quad (16)$$

Therefore

$$\mathbf{R}_{\tilde{\mathbf{S}}_{2M} \tilde{\mathbf{S}}_{2M}} = \tilde{\mathbf{P}}_{2M} \mathbf{R}_{\tilde{\mathbf{X}}_{2M} \tilde{\mathbf{X}}_{2M}} \tilde{\mathbf{P}}_{2M}^T, \quad (17)$$

which can be obtained once the prefilter and the input statistics are known.  $\mathbf{R}_{\tilde{\mathbf{S}}_0 \tilde{\mathbf{S}}_{8C}}$  and  $\mathbf{R}_{\tilde{\mathbf{S}}_{8C} \tilde{\mathbf{S}}_{8C}}$  in (14) are simply submatrices of  $\mathbf{R}_{\tilde{\mathbf{S}}_{2M} \tilde{\mathbf{S}}_{2M}}$ .

## 5. 2-D STATISTICAL MODELS AND ADAPTIVE FILTERING

Both (8) and (17) requires the knowledge of the 2-D auto-correlation matrix of the input. We have tested with two commonly used models for natural images. The first one is the separable AR(1) model, where

$$r_{xx}(h, k) = E\{x(m, n)x(m+h, n+k)\} = \sigma_x^2 \rho_r^{|h|} \rho_c^{|k|}, \quad (18)$$

where  $\rho_r$  and  $\rho_c$  are row and column correlation coefficients, respectively.

Another model is the isotropic model, where

$$r_{xx}(h, k) = \sigma_x^2 \rho^{\sqrt{h^2+k^2}}. \quad (19)$$

This is in general a more accurate representation of natural images than the separable model, and we will use this model in the following designs and experiments. The value of  $\rho$  is chosen to be 0.95.

The Wiener filter based on the model in (19) works well for smooth images. However, its performance may not be satisfactory for images with more high frequency components. In these cases, a simple multi-pass method can be applied. We first use the optimized 2-D Wiener filter and the inverse TDLT to obtain the first reconstructed image. This image is then used to estimate the auto-correlation matrix, which allows us to compute an improved Wiener filter according to (3) or (14). This filter is thus applied to the input data to get a second-pass reconstruction image. The procedure can be repeated to get multi-pass results. Our experiments show that the best results is usually obtained in three passes.

## 6. DESIGN AND ERROR CONCEALMENT EXAMPLES

Fig. 2 (a) shows the error distribution given by the 1-D Wiener filter in [11] when the prefilter is optimized for both coding efficiency and error resilience. The coding gain is reduced from 9.61 dB to 8.41

dB, but the MSE is reduced by almost 70% (from 0.18 to 0.06). Fig. 2 (b) shows the result of the 4-connection 2-D Wiener filter at similar compression efficiency, which further reduces the MSE to 0.02. Fig. 2 (c) is the error distribution of the 8-connection 2-D Wiener filter (14). Its MSE is also 0.02.

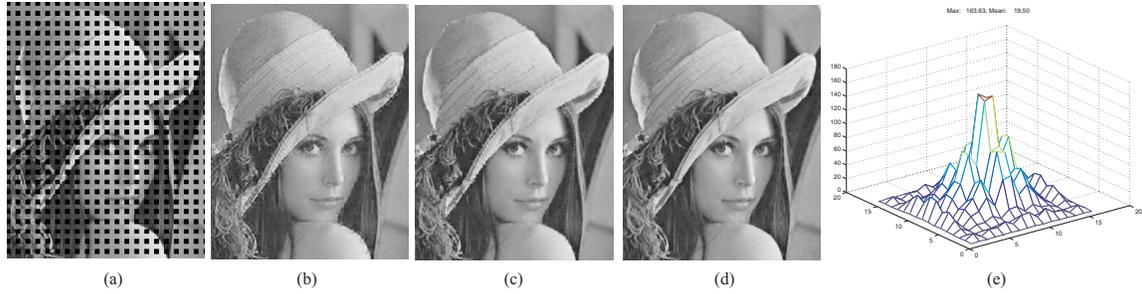
In Fig. 3, we compare the error concealment performances of the mean reconstruction method in [10], the 1-D Wiener filter in [11], and the 4-connection 2-D Wiener filter proposed in this paper. The forward transforms have similar coding gains in these cases. A 25% regular loss scenario is considered. Such loss pattern can happen in block splitting-based multiple description coding [7]. The PSNR given by the 1-D Wiener filter is 3.8 dB higher than the mean reconstruction method. However, due to the separate estimations in horizontal and vertical directions and the final average operation, some artifacts can be noticed, especially near the edges. In contrast, the 2-D method yields another 1.9 dB improvement, and the visual artifacts are also reduced significantly. The measured error distribution around each lost block is given in Fig. 3 (e), whose shape agrees very well with the theoretical distribution in Fig. 2 (b).

Fig. 4 demonstrates the performance of the multi-pass filtering approach for image Barbara, as discussed in Sec. 5. The 8-connection filter is used. This image is dominated by textures; therefore (19) is no longer suitable. This explains the artifacts in the first-pass result in Fig. 4 (c). It is also slightly worse than the 1-D Wiener filter result in Fig. 4 (b). However, after estimating the auto-correlation matrix from the reconstructed image, re-computing the Wiener filter, and repeating the filtering, the next two passes achieve 5.7 dB and 6.2 dB improvements, respectively. In fact, we can only get 0.05 dB higher even with the correlation matrix estimated from the original uncorrupted image. This shows that the multi-pass approach is very effective in boosting the performance of error concealment in the lapped transform framework.

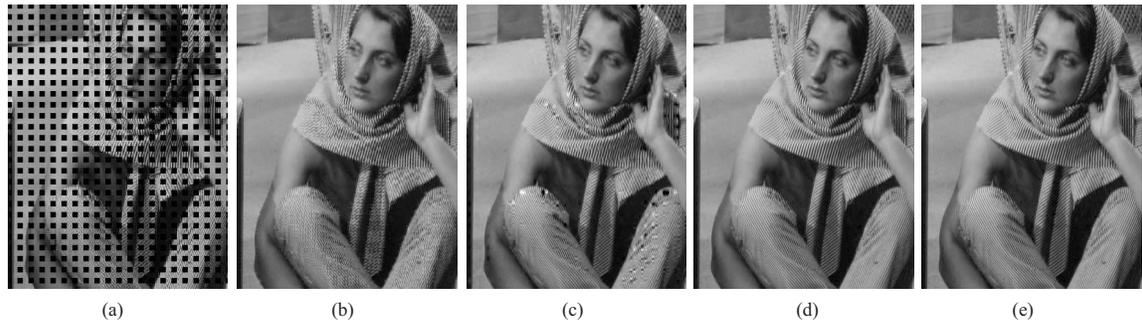
## 7. CONCLUSIONS

Two-dimensional Wiener filter for the error resilient lapped transform is developed. Together with multi-pass adaptive filtering, the 2-D solution can yield more than 5.5 dB improvement over the 1-D Wiener filter in certain cases. The derivation in this paper can be modified to obtain 2-D Wiener filters with less than four available neighboring blocks. These filters are needed at the boundary of images. They are also necessary inside the image when more than two descriptions are used in MDC or in random block loss scenarios.

In the 1-D case [11], the size of the filter is  $M \times 2M$ , and the filter is applied to each row and each column. Ignoring the possible fast implementation, this requires roughly  $M \times 2M \times 2M$



**Fig. 3.** Portions of error concealment results of different methods; (a) Loss pattern; (b) by filter P2 in [10] with mean reconstruction method: 29.86 dB; (c) by the 1-D Wiener filter P21 in [11]: 33.61 dB; (d) by the 4-connection 2-D Wiener filter: 35.50 dB; (e) Measured error distribution of (d).



**Fig. 4.** Portions of error concealment results of different methods; (a) Loss pattern; (b) by the 1-D Wiener filter P21 in [11]: 28.73 dB; (c) The first pass result with the 8-connection 2-D filter: 28.05 dB; (d) The second pass result with the 8-connection 2-D filter: 33.72 dB; (e) The third pass result with the 8-connection 2-D filter: 34.25 dB.

multiplication-addition operations, or  $4M$  operations for each coefficient in the lost block. The size of the 2-D Wiener filters in (3) and (14) are  $M^2 \times 4M^2$  and  $M^2 \times 3M^2$ , respectively, thus their complexities are approximately  $4M^2$  and  $3M^2$  operations for each coefficient, which are acceptable for small block size.

When the multi-pass approach is used, the estimation of the auto-correlation matrix could be quite time-consuming. How to reduce the complexity of the adaptive algorithm without too much performance tradeoff is our ongoing research. Finally, the current adaptive algorithm still uses the same Wiener filter at every lost block within each pass. This can be improved by adapting the filter to local statistics, as in [12]. However, the complexity will be further increased. How to implement it efficiently is another future topic.

## 8. 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, "Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 9 (6), pp. 895–908, Sept. 1999.
- [8] —, "Lapped orthogonal transform designed for error resilient image coding," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 12(9), pp. 752–764, Sept. 2002.
- [9] 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.
- [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] J. Liang, C. Tu, T. D. Tran, and L. Gan, "Wiener filtering for generalized error resilient time domain lapped transform," *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Proc.*, vol. 2, pp. 181–184, Mar. 2005.
- [12] X. Li and M. Orchard, "Novel sequential error concealment using orientation adaptive interpolation," *IEEE Trans. Circ. Syst. for Video Tech.*, vol. 12 (10), pp. 857–864, Oct. 2002.