# OPTIMAL BLOCK BOUNDARY PRE/POST-FILTERING FOR WAVELET-BASED IMAGE AND VIDEO COMPRESSION

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

[1]School of Engineering Science, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada
[2]Windows Digital Media Division, Microsoft Corporation, Redmond, WA, 98052, USA
[3]ECE Department, The Johns Hopkins University, Baltimore, MD, 21218, USA

**ABSTRACT**

This paper presents a pre/post-filtering method to reduce the reconstruction errors near block boundaries in wavelet-based image and video compression. It can be used effectively to mitigate the tiling artifact in JPEG 2000 and the jittering artifact in 3-D wavelet-based video compression. In this method, a short prefilter is applied across the boundaries of image tiles or video frame groups before wavelet compression, and a post-filter is applied at the same place after wavelet reconstruction. The optimal pre/post-filter is obtained by formulating and solving the corresponding rate-distortion optimization problem. A low-complexity structure is then proposed to approximate the optimal solution. The performance of the proposed method is demonstrated by both image and video coding examples.

## 1. INTRODUCTION

It is well known that DCT-based image coding systems exhibit annoying blocking artifacts at low bit rates. Wavelet transform (WT) is one of the solutions to this problem. Blocking artifact can be removed satisfactorily when WT is applied to the whole image. In addition to applications in image coding, various 3-D wavelet-based methods have been proposed to video coding [1, 2]. These approaches inherit the scalability of wavelet and are promising for video delivery over heterogeneous networks.

However, due to the constraints of CPU and memory on different platforms, block-based WT has to be adopted in certain circumstances. In image coding, this requires that a large image to be compressed by dividing it into some smaller pieces and encoding each part separately. In JPEG2000, each piece is called a *tile*. As a result, blocking artifacts known as *tiling artifacts* can be observed at low bit rates.

Block-based approach is more imperative in 3-D wavelet-based video coding. In this case, the video sequence is usually divided into groups of frames (GOF), and each group is compressed by 3-D wavelet separately. This unavoidably leads to serious degradation of quality at group boundaries - a phenomenon known as the *jittering artifact* [3].

Some methods have been proposed to reduce the tiling artifacts in wavelet-based image compression, including, for example, the method of projection onto scaling functions [4], the point-symmetric extension [5], and the odd-tile length method [6, 7]. To reduce the jittering artifact in 3-D wavelet video coding, a pipeline implementation is proposed in [3] to achieve the WT of the whole video sequence without too much buffering requirement, thus eliminating the boundary artifacts.

In this paper, we propose a pre/post-filtering approach, where a pre-processing operator is applied at tile boundaries or GOF boundaries before the wavelet compression. At the decoder side, a post-processing is performed at the same place after inverse WT. The optimal pre/post-filter is found by formulating a rate-distortion problem for the boundary filter bank. The optimal solution is then approximated by a low-cost structure. Various design examples are presented. Experimental results show that the proposed scheme leads to significantly PSNR improvement at tile/GOF boundaries in both image and video coding, and the visual quality is also enhanced satisfactorily.

## 2. PROBLEM FORMULATION

Fig. 1 (a) illustrates an example of the proposed wavelet-based image/video compression with $2K$-point pre-filter $\mathbf{P}$ and post-filter $\mathbf{P}^{-1}$ operating at tile/GOF boundaries, where $x_{0,n}$'s and $x_{1,n}$'s ($n = 0, \ldots, N-1$) are input samples of two neighboring blocks of size $N$. The WT is implemented through a number of lifting and scaling steps, with symmetric extension employed at block boundary. For simplicity purpose, only one-level of WT is considered in the design, but the result can be applied satisfactorily to systems with multiple levels of WT.

Our first objective is to find the optimal pre- and post-filters in the rate-distortion sense. To this end, we first identify the number of reconstructed pixels that are affected by the given pre- and post-filters, as labeled in Fig. 1 (a). Starting from these reconstructed pixels and tracing back, we can find all wavelet coefficients and input samples that contribute to these reconstructed samples. As a result, we can obtain the boundary filter bank model as given in Fig. 1 (b). The objective is to find the optimal pre- and post-filters such that the mean-squared error (MSE) of the reconstructed boundary vectors $\hat{\mathbf{y}}_0$ and $\hat{\mathbf{y}}_1$ is minimized. The solution can be obtained as follows.

First, the boundary WT coefficients can be written as

$$\mathbf{u} = \begin{bmatrix} \mathbf{u}_0 \\ \mathbf{u}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{F}_0 & 0 \\ 0 & \mathbf{F}_1 \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N_0'} & 0 & 0 \\ 0 & \mathbf{P} & 0 \\ 0 & 0 & \mathbf{I}_{N_1'} \end{bmatrix} \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{x}_1 \end{bmatrix} \triangleq \mathbf{F}\,\mathbf{x}, \quad (1)$$

where $\mathbf{F}_0$ and $\mathbf{F}_1$ represent the boundary forward WTs at the two neighboring blocks, the sizes of $\mathbf{x}_i$, $\mathbf{F}_i$ and $\mathbf{u}_i$ are $N_i \times 1$, $M_i \times N_i$, and $M_i \times 1$, respectively, $N_0^{'} = N_0 - k$, $N_1^{'} = N_1 - K$, and $\mathbf{I}_n$ is the $n \times n$ identity matrix.

This work was supported by NSF CAREER Grant CCR-0093262. E-mails: jiel@sfu.ca, chentu@microsoft.com, trac@jhu.edu.

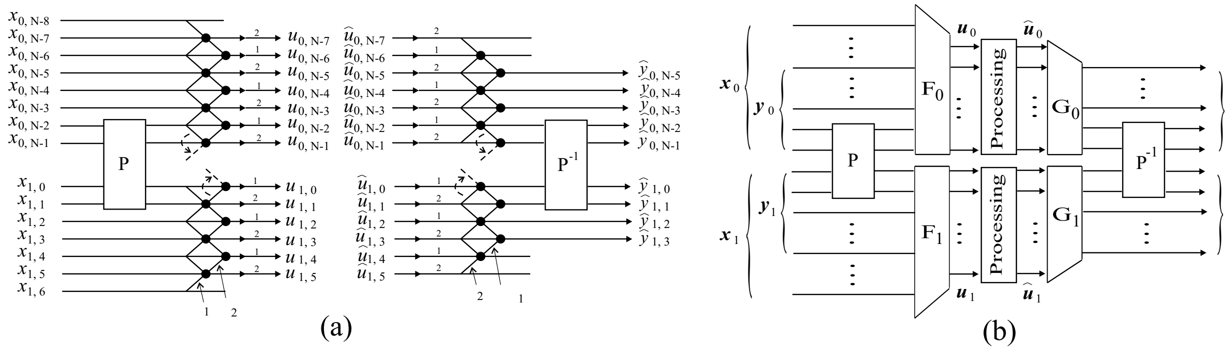



**Fig. 1**. (a) An example of wavelet-based image/video compression with pre- and post-processing at block boundaries. (b) General model for block boundary filter bank.

Let $\mathbf{q}_i$ be the quantization noise applied to $\mathbf{u}_i$, and $\mathbf{e}_i = \hat{\mathbf{y}}_i - \mathbf{y}_i$ be the boundary reconstruction errors of two blocks, we have

$$\mathbf{e} = \begin{bmatrix} \mathbf{e}_0 \\ \mathbf{e}_1 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{L_0'} & 0 & 0 \\ 0 & \mathbf{P}^{-1} & 0 \\ 0 & 0 & \mathbf{I}_{L_1'} \end{bmatrix} \begin{bmatrix} \mathbf{G}_0 & 0 \\ 0 & \mathbf{G}_1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_0 \\ \mathbf{q}_1 \end{bmatrix} \triangleq \mathbf{G}\,\mathbf{q}, \quad (2)$$

where $\mathbf{G}_i$ represents the boundary inverse WTs, the sizes of $\mathbf{q}_i$, $\mathbf{G}_i$, and $\mathbf{y}_i$ are $M_i \times 1$, $L_i \times M_i$, and $L_i \times 1$, respectively, $L_0^{'} = L_0 - K$, and $L_1^{'} = L_1 - K$.

The mean squared error (MSE) of boundary samples can thus be expressed as

$$\mathcal{E} = \text{tr}\{E\{\mathbf{e}\,\mathbf{e}^T\}\} = \text{tr}\{\mathbf{G}\,\mathbf{R}_{\mathbf{q}\,\mathbf{q}}\,\mathbf{G}^T\} = \text{tr}\{\mathbf{R}_{\mathbf{q}\,\mathbf{q}}\,\mathbf{G}^T\,\mathbf{G}\}, \quad (3)$$

where tr$\{\mathbf{A}\}$ is the trace of matrix $\mathbf{A}$. If subband noises are uncorrelated, $\mathcal{E}$ can be written as

$$\mathcal{E} = \sum_{k=0}^{M_0-1} \sigma^2_{q_{0,\,k}} \parallel \mathbf{G}_k \parallel_2^2 + \sum_{k=0}^{M_1-1} \sigma^2_{q_{1,\,k}} \parallel \mathbf{G}_{M_0+k} \parallel_2^2, \quad (4)$$

where $\sigma^2_{q_{i,\,k}}$ is the variance of the $k$-th entry in $\mathbf{q}_i$, and $\parallel \mathbf{G}_k \parallel_2^2$ is the norm of the $k$-th column of $\mathbf{G}$.

Define $\mathbf{R}_{\mathbf{x}\,\mathbf{x}}$ to be the autocorrelation matrix of the input, we have $\mathbf{R}_{\mathbf{u}\,\mathbf{u}} = \mathbf{F}\,\mathbf{R}_{\mathbf{x}\,\mathbf{x}}\,\mathbf{F}^T$. Let

$$\sigma^2_{u_{0,\,k}} = \mathbf{R}_{\mathbf{u}\,\mathbf{u}}(k,\,k), \quad \sigma^2_{u_{1,\,k}} = \mathbf{R}_{\mathbf{u}\,\mathbf{u}}(M_0+k,\,M_0+k), \quad (5)$$

the variance of the quantization noise can be modeled by

$$\sigma^2_{q_{i,\,k}} = C\,2^{-2b_{i,\,k}}\,\sigma^2_{u_{i,\,k}}, \quad (6)$$

where $C$ is a constant, and $b_{i,\,k}$ is the number of bits allocated to the $k$-th channel in $\mathbf{u}_i$.

Since the two blocks are quantized separately, the optimal bit allocation problem can be formulated as

$$\mathcal{E}_{min} = \min_{\mathbf{b}_0,\,\mathbf{b}_1} \Big( \sum_{k=0}^{M_0-1} C\,2^{-2b_{0,\,k}}\,\sigma^2_{u_{0,\,k}} \parallel \mathbf{G}_k \parallel_2^2 + \sum_{k=0}^{M_1-1} C\,2^{-2b_{1,\,k}}\,\sigma^2_{u_{1,\,k}} \parallel \mathbf{G}_{M_0+k} \parallel_2^2 \Big), \quad (7)$$

$$\text{subject to} \quad \frac{1}{M_0} \sum_{k=0}^{M_0-1} b_{0,\,k} = \bar{b}_0, \quad \frac{1}{M_1} \sum_{k=0}^{M_1-1} b_{1,\,k} = \bar{b}_1.$$

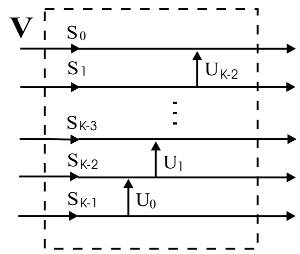



**Fig. 2**. Fast structure for matrix $\mathbf{V}$ in pre-filter.

where $\bar{b}_0$ and $\bar{b}_1$ are the targeted average bit rates of the two blocks.

This is a standard Lagrangian problem, and the solution can be found to be

$$\mathcal{E}_{min} = M_0\,2^{-2\bar{b}_0}\,\beta_0^2 + M_1\,2^{-2\bar{b}_1}\,\beta_1^2, \quad (8)$$

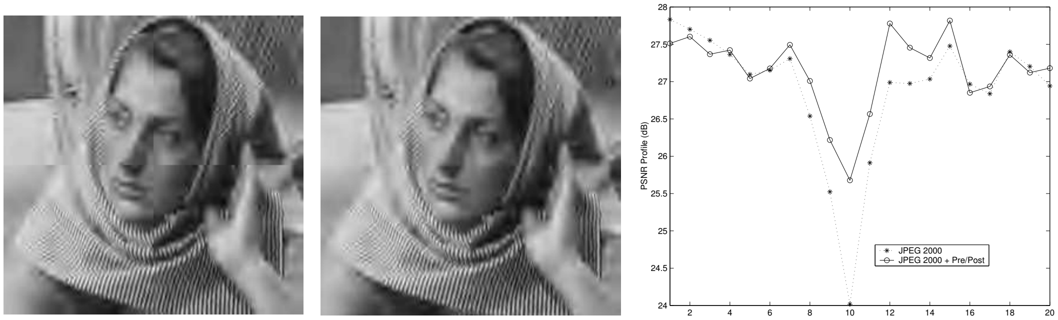
where

$$\beta_0^2 = \left( \prod_{k=0}^{M_0-1} \sigma^2_{u_{0,\,k}} \parallel \mathbf{G}_k \parallel_2^2 \right)^{1/M_0}, \qquad \beta_1^2 = \left( \prod_{k=0}^{M_1-1} \sigma^2_{u_{1,\,k}} \parallel \mathbf{G}_{M_0+k} \parallel_2^2 \right)^{1/M_1}. \quad (9)$$

Similar to standard subband coding problems, we can define the coding gain of the boundary filter bank with respect to the PCM scheme, where each sample is compressed directly to the bit rate of $\bar{b}_i$ in the two blocks. Therefore the coding gain of the pre/post-filter can be written as

$$\gamma = \frac{\mathcal{E}_{PCM}}{\mathcal{E}_{min}} = \frac{\left(L_0\,2^{-2\bar{b}_0} + L_1\,2^{-2\bar{b}_1}\right)\sigma_x^2}{M_0\,2^{-2\bar{b}_0}\,\beta_0^2 + M_1\,2^{-2\bar{b}_1}\,\beta_1^2} = \frac{(R_b\,L_0 + L_1)\,\sigma_x^2}{R_b\,M_0\,\beta_0^2 + M_1\,\beta_1^2}, \quad (10)$$

where $R_b = 2^{-2\,(\bar{b}_0 - \bar{b}_1)}$. In this paper, $R_b$ is fixed to be 1. The optimal pre- and post-processing operators can be found by setting up an optimization program to maximize the coding gain as defined above.




(a) (b) ( c )

**Fig. 3**. An image coding example: (a) Decoded result with JPEG 2000 only. (b) Decoded result with JPEG 2000 and 8-point pre- and post-filtering. (c) Horizontal PSNR across tile boundaries.

**Table 1**. Optimal parameters and fast approximations of the pre-filter in Fig. 2 for 9/7 Wavelet and AR(1) input with $\rho = 0.95$.

| $\mathbf{P}$ Rows | $S_0$ | $S_1$ | $S_2$ | $S_3$ | $U_0$ | $U_1$ | $U_2$ | Gain (dB) |
|---|---|---|---|---|---|---|---|---|
| 0 | - | - | - | - | - | - | - | 9.28 |
| 2 | 7/4 | - | - | - | - | - | - | 9.37 |
| | 2 | - | - | - | - | - | - | 9.36 |
| 4 | 1.63 | 1.30 | - | - | 0.55 | - | - | 9.41 |
| | 3/2 | 5/4 | - | - | 1/2 | - | - | 9.40 |
| | 2 | 1 | - | - | 1/2 | - | - | 9.36 |
| 6 | 1.65 | 1.33 | 1.17 | - | 0.17 | 0.63 | - | 9.42 |
| | 3/2 | 5/4 | 9/8 | - | 1/8 | 1/2 | - | 9.41 |
| | 2 | 1 | 1 | - | 1/4 | 1/2 | - | 9.38 |
| 8 | 1.66 | 1.35 | 1.19 | 1.13 | 0.10 | 0.21 | 0.63 | 9.43 |
| | 5/3 | 4/3 | 6/5 | 9/8 | 1/8 | 1/4 | 5/8 | 9.43 |
| | 2 | 1 | 1 | 1 | 1/8 | 1/4 | 1/2 | 9.38 |

## 3. EFFICIENT STRUCTURE

In addition to coding gain, some other properties are desired for a filter bank to be used in image and video compression. In the pre/post-processing of the WT, these requirements impose further constraints on the structure of the pre/post-filters. In this paper, we propose the following structure:

$$\mathbf{P} = \frac{1}{2}\begin{bmatrix} \mathbf{I} & \mathbf{J} \\ \mathbf{J} & -\mathbf{I} \end{bmatrix}\begin{bmatrix} \mathbf{I} & 0 \\ 0 & \mathbf{V} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{J} \\ \mathbf{J} & -\mathbf{I} \end{bmatrix}, \qquad (11)$$

where $\mathbf{J}$ is the reversal identity matrix. This structure is identical to that of the pre-/post-filters in the time-domain lapped transform [8]. All freedoms in this structure lie in the $K \times K$ matrix $\mathbf{V}$, which can be optimized for coding gain. Since $\mathbf{V}$ is applied to the lower part of the butterfly output, the pre-filter has no effect when the input is constant. Therefore the superior energy compaction capability of WT for a constant input is preserved. Moreover, the prefilter satisfies $\mathbf{P} = \mathbf{JPJ}$, *i.e.*, it provides identical boundary processing to both blocks.

Optimization results show that when $K$ is small, the optimal matrix $\mathbf{V}$ can be approximated very well by the fast structure in Fig. 2, which involves $K$ scaling factors and $K - 1$ lifting steps. Some optimal parameters for the fast pre-filter are listed in Table 1. Various rational approximations that allow fast implementation are also presented. Notice that those with integer scaling factors enable lossless compression, if lossless WT is used.

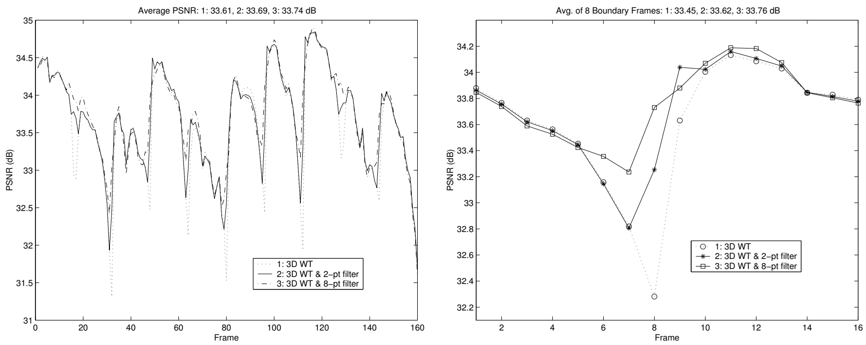
## 4. IMAGE/VIDEO CODING EXPERIMENTS

In this section, we demonstrates the performance of the pre- and post-filtering in the removal of tiling artifact in JPEG 2000 and the jittering artifact in 3D wavelet video coding.

An image coding example is given in Fig. 3 with the $512 \times 512$ gray-scale image Barbara. The tile size is chosen to be $128 \times 128$ and 5-level 9/7 WT is applied. Fig. 3 (a) shows a portion of the JPEG2000 decoded result with Kakadu version 3.4 [9] at 0.2 bits/pixel. Tiling artifact can be clearly observed. Compared with this, the result is much more pleasant in Fig. 3 (b), where the second 8-point pre- and post-filter in Table 1 is employed. Fig. 3 (c) shows the average horizontal PSNRs of the two decoded images across all rows and all tile boundaries, where each PSNR value is obtained from all co-located boundary pixels. The first 10 samples belong to the left tiles, whereas the next 10 samples belong to the right tiles. It can be seen that with the help of pre- and post-processing, the PSNR for the last pixel of the left tile is improved by more than 1.5 dB. Other boundary pixels are also improved, leading to mitigated tiling artifacts.

We next apply the fast pre-/post-processing solutions to reduce the jittering artifact in 3-D wavelet video coding. Since the GOF size is much smaller than the tile size in image compression, the percentage of pre/post-processed data is much larger in the 3-D video coding than in the image compression. As a result, the pre/post-filtering not only reduces the jittering artifact, but also improves the PSNR performance, as shown in Table 2. In Fig. 4, the coding results of the first 160 frames of the QCIF sequence Akiyo with different pre- and post-filtering are displayed. The GOF size is chosen to be 16, and 3-level 9/7 WT is applied to each dimension. 3-D SPIHT is then used to encode the wavelet coefficients [2], and the compression ratio is 120 : 1. The lossless $2 \times 2$ filter and lossless $8 \times 8$ filter in Table 1 are used respectively.

Fig. 4 (a) plots the PSNR of all frames in the three scenarios, and the average result across all GOP boundaries is plotted in Fig. 4 (b), where frames 1-8 belong to one GOP and frames 9-16 belong to the next GOP. It can be observed from Fig. 4 that the original 3-D wavelet-based video coding experiences serious PSNR drop at GOF boundaries, whereas the transition at GOF boundaries become much smoother when pre/post-filtering is applied.




(a) (b)

**Fig. 4**. A 3-D wavelet video coding example. (a) PSNR of all frames. (b) Average PSNR across all group boundaries.

**Table 2**. Video coding results (PSNR in $dB$) with 9/7 wavelet-based 3-D SPIHT and various pre/post-filtering

| Comp. Ratio | SPIHT Only | SPIHT & lossless 2-pt Filter | SPIHT & lossless 8-pt Filter | SPIHT & lossy 4-pt Filter | SPIHT & lossy 6-pt Filter | SPIHT & lossy 8-pt Filter |
|---|---|---|---|---|---|---|
| Akiyo (QCIF, 288 frames) | | | | | | |
| 30 : 1 | 41.23 | 41.35 | 41.36 | 41.47 | 41.55 | 41.44 |
| 60 : 1 | 38.25 | 38.39 | 38.45 | 38.48 | 38.53 | 38.48 |
| 120 : 1 | 33.62 | 33.70 | 33.74 | 33.75 | 33.78 | 33.75 |
| Claire, QCIF, 480 frames) | | | | | | |
| 30 : 1 | 41.93 | 42.09 | 42.07 | 42.24 | 42.31 | 42.18 |
| 60 : 1 | 39.37 | 39.56 | 39.63 | 39.66 | 39.72 | 38.67 |
| 120 : 1 | 35.33 | 35.47 | 35.56 | 35.56 | 35.60 | 35.61 |
| Foreman (QCIF, 400 frames) | | | | | | |
| 30 : 1 | 32.06 | 32.18 | 32.18 | 32.25 | 32.27 | 32.19 |
| 60 : 1 | 29.31 | 29.41 | 29.45 | 29.48 | 29.51 | 29.46 |
| 120 : 1 | 26.81 | 26.91 | 26.96 | 26.97 | 26.99 | 26.99 |

## 5. CONCLUSION

In this paper, we present a block boundary pre-/post-processing approach for wavelet-based image and video coding. It can be used to reduce the tiling artifact in JPEG 2000 and the jittering artifact in 3-D wavelet video coding. The optimal solution is obtained by solving a rate-distortion problem. A low-complexity structure that preserves the lossy-to-lossless compression of the wavelet is proposed, based on the result in time-domain lapped transform. The effectiveness of the proposed scheme is demonstrated by both image and video coding examples. As a final note, although the proposed framework requires both pre-filter and post-filter, experimental results show that at low bit rate, post-processing alone is actually effective enough to reduce the boundary artifact.

## 6. REFERENCES

[1] D. Taubman and A. Zakhor, "Multirate 3-D subband coding of video," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 572–589, Sept. 1994.

[2] B.-J. Kim, Z. Xiong, and W. A. Pearman, "Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT)," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 10, no. 12, pp. 1365–1374, Dec. 2000.

[3] J. Xu, Z. Xiong, S. Li, and Y.-Q. Zhang, "Memory-constrained 3-D wavelet transform for video coding without boundary effects," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 12, no. 9, pp. 812–818, Sept. 2002.

[4] K. Berkner and E. Schwartz, "Removal of tile artifacts using projection onto scaling functions for JPEG 2000," *Proc. of ICIP 2002*, vol. 1, pp. 373–376, Rochester, NY, Sept. 2002.

[5] I. Kharitonenko, X. Zhang, and S. Twelves, "A wavelet transform with point-symmetric extension at tile boundaries," *IEEE Trans. Image Process.*, vol. 11, no. 12, pp. 1357–1364, Dec. 2002.

[6] J. Wei, M. Pickering, M. Frater, and J. Arnold, "A new method for reducing boundary artifacts in blocking-based wavelet image compression," *Proc. of SPIE Visual Comm. and Image Process.*, pp. 1290–1295, 2000.

[7] J. Wei, M. Pickering, M. Frater, J. Boman, J. Arnold, and W. Zeng, "Boundary artifact reduction using odd tile length and the low pass first convention (OTLPF)," *Proc. of SPIE Conf. on Appl. of Dig. Image Process.*, Jul. 2001.

[8] 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.

[9] D. Taubman, "Kakadu version 3.4," *available at http://www.kakadusoftware.com*, 2001.