Optimal tile boundary artifact removal with CREW

Edward L. Schwartz, Kathrin Berkner, and Michael J. Gormish
Ricoh Silicon Valley, Inc.
2882 Sand Hill Road, Suite 115, Menlo Park, CA 94025
email: {schwartz,berkner,gormish,crew}@crc.ricoh.com, www.crc.ricoh.com/CREW/

Abstract: For practical image compression systems, handling images divided into independent tiles in the pixel and transform domain is important. However, tiling introduces artifacts when wavelet coefficients are quantized. The compression scheme CREW (Compression with Reversible Embedded Wavelets) is a system that can perform wavelet transforms on independent tiles. This paper presents a "detiling" solution for a family of wavelet systems similar to that used in CREW. The solution eliminates the blocking artifacts and computes smooth approximations.

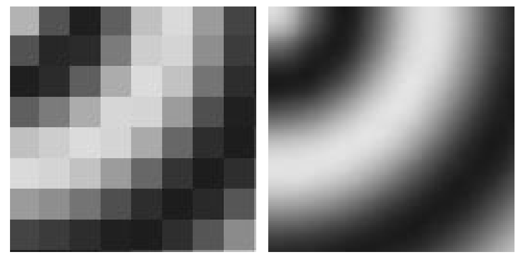

Fig. 1 Blocking artifacts in a DCT (left) and wavelet-based tile-compressed (right) image

1 Introduction

For practical image compression systems, handling images divided into independent tiles in the pixel as well as in the transform coefficient domain is important. A partition into simple rectangular tiles is the most convenient method for implementation. Tiling allows rectangular regions-of-interests (ROI), as either an encoder or a decoder option, and aids low memory operation and parallel processing. However, naive decoding for independent tiles in a wavelet based lossy compression system can create visible tile boundary artifacts. These artifacts can occur regardless of the filter chosen.

A well known property of the wavelet transform that makes it an attractive tool for image compression is that using only parts of the transform coefficients provides smooth reconstruction. This property occurs because wavelets form an unconditional basis in a wide range of smoothness spaces (see e.g. [4]). As a consequence, lossy full-frame wavelet-based image compression generally causes some ringing around edges in the decompressed image, but not really "sharp" or "peaky" artifacts. In order to illustrate the general advantage regarding smoothness properties of wavelet transforms over a Discrete Cosine Transform (DCT), the following example is considered. Let the starting set of pixels be the same for both methods and consider two approximate reconstructions of this set of pixels, one using a full-frame 4-level 2-10 wavelet system (defined below) and one using a 16×16 DCT. These transforms are chosen such that the DCT DC basis vector and the lowpass wavelet filter at level four both average over 16×16 blocks. The other coefficients are quantized to zero, making irrelevant the differences between the DCT AC basis vectors and the wavelet highpass synthesis filter. The difference in the applied synthesis filters results in a globally smooth wavelet reconstruction and a blocky DCT reconstruction (Figure 1). In light of this example, it is clear that quantized wavelet transform coefficients can introduce artifacts only at boundaries of tiles (see Figure 3).

This paper focuses on wavelet filters with simple low pass analysis filters such as the CREW (Compression with Reversible Embedded Wavelets) 2-10 filter [1]. CREW allows to compress independent tiles of an image and uses an integer-to-integer transform with reversible wavelets that allows lossy as well as lossless compression. A detiling solution for the 2-10 filter and other two-tap analysis filters is presented in this paper. For those filters the proposed solution is shown to be optimal for "smooth" images where "smooth" is defined as all quantized full-frame highpass coefficients at the tile boundaries are zero. The proposed method is based only on the properties of the chosen wavelet system. Therefore, the image is not required to fit any particular model.

This paper is organized as follows. Section 2 gives a short review on the tile-boundary artifact problem in wavelet-based compressed images. Section 3 presents a detiling solution via the novel *pseudo full-frame inverse transform*. This transform is defined for the family of two-tap analysis filters. Experiments in Section 4 show effective results applying the presented detiling solution.

2 Artifacts in wavelet transform coefficients on tiles

In a lossy compressed version of an image, computed using wavelet transforms on independent tiles of the original image, a decompressed tile itself looks smooth, but the transition to the next tile can become a sharp edge-type boundary. These tile-boundary artifacts occur because transform coefficients used for quantization are computed from disjoint sets of pixels. No matter what filter or what extension rule is used at the boundaries (symmetric extension, replication, anti-symmetric extension), in general, it will not match the actual image on the other side of the boundary and cause artifacts for some images as seen in Figure 3 and 4. In detail, the boundary artifact problem is demonstrated in a numerical example in Figure 2 where a one-dimensional linear signal is transformed. Comparing the full-frame- with the tile-coefficients clarifies the introduced artifact at the boundary between the two tiles. So called “tile-boundary artifacts” are often viewed as “unacceptable,” in particular at low bitrates. This problem occurs only in lossy, but not lossless decoders.

Significant work has been done in the field of block-artifact removal in DCT-compressed images. One approach to removal of tile boundary artifacts is to apply a postprocessing step and simply smooth the decompressed image at the boundaries with a lowpass filter. However, in order to provide a successful solution this approach requires an image model and does not work for arbitrary images. In general, smoothing at boundaries will introduces new artifacts. Therefore, detiling solutions are desirable that are incorporated into the decoder and use information from various quantized wavelet coefficients. One approach that matches this criterion for the DCT is the one that described in Section K.8 of the JPEG standard [5]. This method works on the quantized coefficients and uses them to compute a polynomial fit through the given coefficients. Since the polynomial model is not naturally associated with a DCT this approach requires rather complex calculations. Another well-known approach is the *Projection Onto Convex Sets* [7]. This method needs a model that cannot be specified in the transform domain. Therefore, it requires an iterative scheme switching back and force between the transform and the data domain causing high computational cost.

A way to avoid block artifacts in wavelet-based compression is presented in [3]. In that approach the authors compute wavelet coefficients of overlapping tiles. The overlap-size depends on the maximal level of decomposition in the wavelet tree. Storing wavelet coefficients computed from overlap regions is equivalent to storing selected coefficients from the full-frame decomposition. The higher the level of decomposition the more full-frame wavelet coefficients have to be stored. “Line-based” or “rolling buffer” methods have similar storage requirements. These methods complicate random access and parallel processing.

Pixels	1	2	3	4	5	6	7	8	9	10
Full Frame Low Pass	1		3		5		7		9	
Full Frame High Pass	0		0		0		0		0	
Tiled Pixels	1	2	3	4	5	6	6	5	4	3
Low Pass	1		3		5		5		3	
High Pass	0		0		-1					
Quantized High Pass	0		0		0					
Reconstructed PIxels	1	2	3	4	5	5				

Fig. 2 Numerical example for appearance of tile boundary artifacts introduced by quantization.

3 Detiling solution for two-tap lowpass analysis filters

In this section a detiling solution is presented that is incorporated in the decoder and makes use of the smoothness properties of wavelet systems.

The goal is to obtain a decompressed image that is smooth at the tile boundaries. In order to achieve this goal, at each scale J a smooth target image is defined. This consists of the full-frame coefficients at scale J modified such that all highpass coefficients, that would have been effected by boundary corrections in the tile transform, are set to zero. Due to the nature of the wavelet transform, this image has the same degree of smoothness as the synthesis scaling function which is determined by the synthesis lowpass filter. If the analysis filter is a two-tap filter, the lowpass coefficients of the independent tiles and those of the full frame are identical given an even number of samples in each tile at each level. The detail coefficients from full-frame and tile transform, however, differ. These observations are, of course, only true for analysis lowpass filters of length two. This class of filters contains both the 2-10 filter and the S+P transform filter derived in [6].

A correction of the tile highpass coefficients at a scale J using the neighboring lowpass coefficients at the same scale is obtained introducing the novel *pseudo-full-frame inverse transform* W_{pseudo}. This transform is defined for biorthogonal wavelet transforms that use a two-tap analysis filter. At each level of decomposition, it transforms the tile-image to the smooth target image. In the following paragraph one step of this inverse transform in one dimension is explained.

The full-frame wavelet decomposition of the target image is given by lowpass coefficients s^J and highpass coefficients d^j for j=1,...,J. These coefficients are associated with the lowpass operator H and the highpass operator G, i.e. $s^J = Hs^{J-1}$ and $d^J = Gs^{J-1}$, followed by downsampling. In contrast, the tile wavelet decomposition is given by coefficients $\tilde{s}^J$ and $\tilde{d}^j$, j=1,..., J, computed using the operators $\tilde{H}$ and $\tilde{G}$, i.e. $\tilde{s}^J = \tilde{H}\tilde{s}^{J-1}$ and $\tilde{d}^J = \tilde{G}\tilde{s}^{J-1}$. Both operator schemes are applied to all coefficients. The difference between the two schemes is that $\tilde{H}$ and $\tilde{G}$ include the operations at the tile boundaries whereas H and G include boundary operations only at the image boundary. To compute the inverse transforms the synthesis operators H^*, $\tilde{H}^*$, G^* and $\tilde{G}^*$ are necessary. One step of the inverse transforms is performed as

$$s^{J-1} = H^*s^J + G^*d^J \quad , \quad \tilde{s}^{J-1} = \tilde{H}^*\tilde{s}^J + \tilde{G}^*\tilde{d}^J \quad .$$

Since the analysis lowpass filter is a two-tap filter the synthesis highpass filter is also a two-tap filter, and it follows that

$$\tilde{s}^J = s^J \text{ and } \tilde{G}^* = G^* . \qquad (1)$$

It is important to notice that this equality is valid only for wavelet systems with a two-tap analysis lowpass filter.

The goal is to substitute the coefficients $\tilde{d}^j$ by new coefficients $\hat{d}^j$ such that applying the inverse transform given by the operators $\tilde{H}^*$ and $\tilde{G}^*$ to the coefficients ($\tilde{s}^J$, $\hat{d}^J$) yields the same coefficients as obtained from the full-frame inverse transform of (s^J,d^J) using the operators H^* and G^*, i.e.

$$\tilde{H}^*\tilde{s}^J + \tilde{G}^*\hat{d}^J = H^*s^J + G^*d^J . \qquad (2)$$

Because of Equation (1) the above condition reduces to

$$\tilde{G}^*\hat{d}^J = G^*d^J . \qquad (3)$$

Due to the choice of the target image, the components of $\tilde{G}^*\hat{d}^J$ and G^*d^J are the same except of those that were effected by the boundary corrections of the forwards tile transform. For those coefficients d_{tb} the condition $d_{tb} = 0$ (tb = "tile boundary") has to be satisfied. Equation (3) has a unique solution. It becomes clear that the modifications, that have to be made in order to satisfy condition (3), differ from scale to scale since the target image is scale-dependent.

In order to demonstrate the pseudo-full-frame inverse transform the example of the 2-10 wavelet transform used in CREW the implementation of this integer-to-integer transform is reviewed first (see [1]). Given the data x_n the forward transform W is computed as follows:

$s_n = \lfloor(x_{2n}+x_{2n+1})/2\rfloor$ (lowpass coefficients)
$p_n = \lfloor(3s_{n-2} - 22\, s_{n-1} + 22\, s_{n+1} - 3\, s_{n+2} + 32) / 64\rfloor$
$d_n = x_{2n}-x_{2n+1} + p_n$ (highpass coefficients)

This implementation using the intermediate ("lifting") step p can be done for every biorthogonal wavelet system as it shown in [2]. From the computed coefficients the data samples x_n are obtained via the inverse transform W*:

$$x_{2n} = s_n + \lfloor(d_n-p_n+1)/2\rfloor, \quad x_{2n+1} = s_n - \lfloor(d_n-p_n)/2\rfloor.$$

Using these notations, the exact calculations for one step of the pseudo-full-frame inverse transform for the 2-10 filter are:

- Compute the p-portion of the inverse transform W* as a function of s coefficients, denoted by p_f.
- Compute the p-portion of the inverse transform $\tilde{W}^*$ as a function of $\tilde{s}$ coefficients, denoted by p_t.
- Set $\hat{d}$ = -(p_f-p_t) in the reconstruction step of the wavelet transform.
- Clip $\hat{d}$ to be consistent with the quantization.

The extension to the two-dimensional case is given by the following procedure where the standard notation SS, SD, DS, DD is chosen for the lowpass and highpass coefficients.

```
for (scale=maxscale; scale > 0; scale--)
  save DS and DD coefficients effected by boundary
  reconstruct SD coefficients smooth accross boundary
  for each tile
    vertical inverse transform on tile
  reconstruct D coefficients smooth accross boundary
  for each tile
    horizontal inverse transform on tile
```

Now the problem occurs how to make the modified detail coefficients consistent with the actual quantization. Making an SD coefficient consistent with the quantization is easy, since the quantization is known. First the minimum and maximum allowed values are determined from the quantized SD value, given a specified number of unknown bits or a scalar quantization factor. In a second step the minimal and maximal values are used to clip the modified wavelet coefficients -(P_f-P_t). Handling the horizontal transform is slightly more complicated because the quantization of DS and DD coefficients need to be propagated through the vertical transform. In detail, this works as follows:

compute the pseudo-full-frame inverse transform of the lowpass coeffients to obtain new lowpass coefficients (t=top,b=bottom):

$$\hat{S}^T = (\hat{S}^{(t)}, \hat{S}^{(b)}) = W_{pseudo}(\tilde{SS},\tilde{SD})$$

compute the tile-inverse transform of the highpass coefficients:

$$(\tilde{D}^{(t)}, \tilde{D}^{(b)}) = \tilde{W}^*(\tilde{DS},\tilde{DD}) = \tilde{H}^*[\tilde{DS}] + \tilde{G}^*[\tilde{DD}]$$

compute P-portions

$$P^{(t)} = -(P_f^{(t)} - P_t^{(t)}) \quad , \quad P^{(b)} = -(P_f^{(b)} - P_t^{(b)})$$

$$P^{(vert)} = D\tilde{D} - (\tilde{D}^{(t)} - \tilde{D}^{(b)})$$

$$P_{DS} = \lfloor (P^{(t)} + P^{(b)})/2 \rfloor, \; P_{DD} = P^{(t)} - P^{(b)} + P^{(vert)}$$

clip P_{DS} to the quantization bounds of DS.
clip P_{DD} to the quantization bounds of DD.

compute tile-inverse transform

$$\hat{D}^T = (\hat{D}^{(t)}, \hat{D}^{(b)}) = \tilde{W}^*(P_{DS}, P_{DD})$$

apply inverse transform $\tilde{W}^*$ in horizontal direction to $(\hat{S}, \hat{D})$ to obtain a smooth approximation of the lowpass coefficients at the next finer resolution.

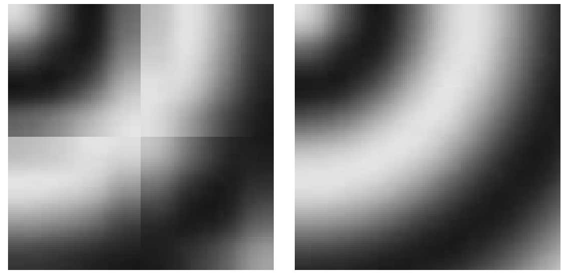

Fig. 3 Decompressed image with artifacts (left), detiling solution (right).

4 Experimental results

In this section some experimental results using the 2-10 filter are discussed. In a first example the property of the presented detiling solution of being optimal for smooth images is demonstrated. In Figure 3 the algorithm is applied to a synthetic image that does not contain edges and yields zero highpass coefficients. For this example reducing the blocking artifacts by computing smooth approximations with the proposed method yields exactly the original image. The example of the real-world image *Boats* from the JPEG test set is shown in Figure 4. For a fixed (very high) quantization level the result shows the blocking artifacts introduced by the wavelet transform on tiles and the effect of “smoothing across edges” in the detiling solution when applying the pseudo-full-frame inverse transform. For the two tile-decompressed images the compressed data is the same.

5 Conclusions

In this paper a detiling solution for artifact reduction in decompressed images using wavelet transform on independent tiles is presented. This solution uses smoothness properties of wavelet systems and is provided by the novel pseudo full-frame inverse transform. This transform is introduced for length-two filters and is explicitly presented for the 2-10 filter used in CREW. Experimental results successfully show the removal of blocking artifacts in the decompressed images. For encoding, lossless decoding, and decoding when artifacts are not objectionable, there is no cost for having the ability to do detiling. The proposed technology has been presented to ISO/IEC JTC 1/SC 29/WG 1 for use with JPEG 2000.

6 Bibliography

[1] M. Boliek, E. L. Schwartz, M. J. Gormish, and A. Keith,“Decoding compression with reversible imbedded wavelets (CREW) codestreams”, in *Journal of Electronic Imaging*, vol. 7, no. 3, pp. 402-409,1998.

[2] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, “Wavelet transforms that map integers to integers,” *Appl. Comput. Harmon. Anal.*, vol. 5, no. 3, pp. 332–369, 1998.

[3] I. K. Eom, Y. S. Kim, and J. H. Kim, “A block wavelet transform for sub-band image coding/decoding,” in *SPIE Electronic Imaging*, vol. 2669, (San Jose, California), pp. 169–77, January 1996.

[4] Y. Meyer, *Wavelets and operators*. Cambridge University Press, 1992.

[5] W. B. Pennebaker and J. L. Mitchell, *JPEG - Still Image and Data Compression Standard*. Van Nostrand Reinhold, New York, 1993.

[6] A. Said and W. A. Pearlman, “An image multiresolution representation for lossless and lossy image compression,” *IEEE Trans. Image Process.*, vol. 5, no. 9, pp. 1303–1310, 1996.

[7] A. Zakhor, “Iterative procedure for reduction of blocking effects in transform coding,” *IEEE Trans. Circ. Sys.*, vol. 2, no. 1, pp. 91–95, 1992.

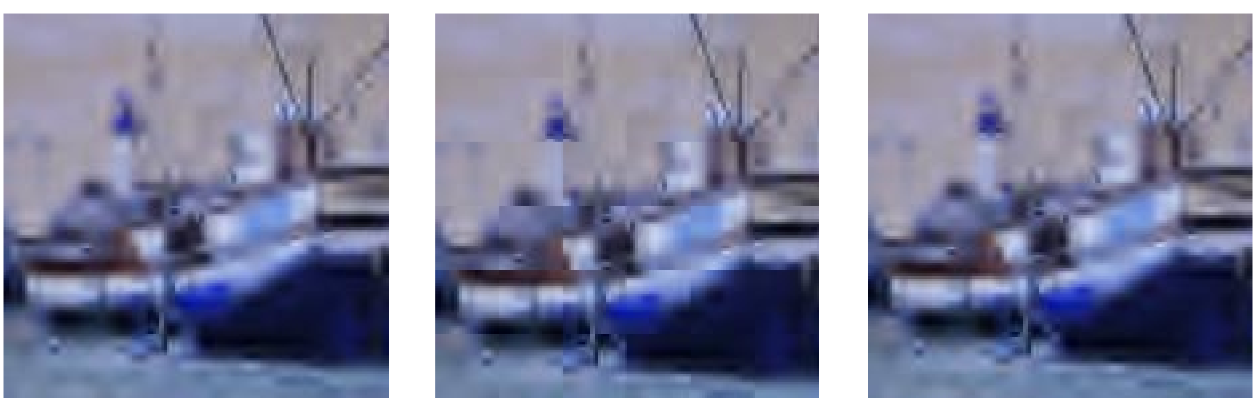

Fig. 4 Full-frame decompressed *Boats* image (left), tile-decompressed image without (middle) and with (right) detiling.