EMBEDDED COLOR IMAGE CODING USING CONTEXT-MODELED WAVELET DIFFERENCE REDUCTION

Yufei Yuan and Mrinal K. Mandal

Department of Electrical & Computer Engineering
University of Alberta, Edmonton AB, Canada T6G 2V4
Email: {yuan, mandal}@ece.ualberta.ca

ABSTRACT

A novel embedded wavelet coding algorithm for color image and video compression is proposed. A color image is first represented in YUV color space. The discrete wavelet transform of each of the three color components is then calculated separately. The luminance component is compressed with context-modeled wavelet difference reduction (CM-WDR). The chrominance components are compressed using the contextual information obtained during the coding of the luminance component. Our experimental results show that the proposed algorithm performs better than two latest wavelet-based benchmark schemes, namely, CEZW and JPEG-2000, in terms of color SNR and component PSNR.

1. INTRODUCTION

Most wavelet-based image coding schemes assume that the input is a grayscale (or monochrome) image. However, color images are almost always involved for real-world applications. A simple and direct extension of these schemes to color image coding is to encode the components of a color image as three independent grayscale images. This strategy is adopted by the latest still image coding standard JPEG-2000 [1]. In the JPEG-2000 coding standard, a color image, such as RGB, is first decorrelated using YCbCr or RCT transform before the wavelet transform. After the color transform stage, the decorrelated components are treated independently as grayscale images.

However, there are two limitations with this simple strategy. Firstly, statistical dependency present among the decorrelated components is not exploited. It has long been observed that in color images the locations of significant spatial changes in the chrominance components generally coincide with significant spatial changes in the luminance component [2]. Hence, the prediction of a chrominance sample can be made efficiently using the previously transmitted chrominance and luminance samples, and the present luminance sample. Secondly, explicit rate allocation among the color components is needed, which complicates the rate control process.

In order to utilize the statistical dependency among decorrelated components and retain the embeddedness and implicit bit allocation of wavelet coders [3, 4], some researchers have proposed algorithms for embedded color image and video coding system. Extension of EZW [3] to color image coding has been proposed as CEZW [5], where zerotree structure is modified to accommodate chrominance components. In CEZW, each node in the chrominance component has two parent nodes, one in the same chrominance component and the other in the luminance component. Hence, for each zerotree root in luminance pyramid, all of its children, including those in chrominance components, are examined. In the extension of SPIHT [4] to color image coding [6], the construction process of initial lists for the three components are interleaved. However, this extension of SPIHT to color coding does not address the exploitation of the statistical dependency among the color components.

A novel wavelet coding algorithm, named color wavelet difference reduction (CWDR), for encoding color image is proposed in this paper. The CWDR first transforms a color image into YUV space, and then applies discrete wavelet transform (DWT) to the Y, U, and V components. The encoding of Y component is done by the context-modeled wavelet difference reduction (CM-WDR) [7]. The encoding of U and V components is based on the contextual information formed during the encoding of Y component. Our experimental results showed that in terms of color SNR and component PSNR, CWDR performs better than CEZW, which also exploits statistical dependency, and JPEG-2000, which is the state-of-the-art image coding standard.

The paper is organized as follows. Section 2 briefly introduces the CM-WDR employed to encode the Y component. Section 3 then describes the CWDR algorithm, which is followed by experimental results in Section 4. Conclusive remarks are given in Section 5.

0-7803-8484-9/04/$20.00 ©2004 IEEE

2. CONTEXT-MODELED WAVELET DIFFERENCE REDUCTION

CM-WDR is a very simple and efficient embedded wavelet image coding algorithm. In CM-WDR, the basic techniques of progressive bit-plane coding and successive scalar quantization are inherited from the EZW [3]. However, the zerotree structure and explicit entropy coding, *i.e.*, arithmetic coding, are not used in CM-WDR.

In CM-WDR, each node within a bit-plane p, *i.e.*, the p-th bit of the binary representation of the magnitude of a wavelet coefficient with the least-significant bit (LSB) labeled the 0-th bit, is classified as either being a significant symbol or being a refinement symbol. The sequence of significant symbols is encoded with an adaptive run-length coder (RLC). The adaptivity of the RLC results from the contextual information formed by encoding the bit-planes that are higher than p. Specifically, the position of a significant symbol within the sequence is determined by its contextual information formed by the encoding of its parent and neighbors in higher bit-planes. In other words, CM-WDR tries to improve the lossless compression ratio of the significant symbol sequence, which is a binary sequence, by clustering the nodes that are more likely to generate stop symbol (bit "1") and placing them in the front of the sequence.

3. COLOR WAVELET DIFFERENCE REDUCTION

The proposed CWDR is based on the same context modeling method adopted in CM-WDR. However, the problem of rate allocation persists before a method is found to assign bit budget among the three components. Since peak signal-to-noise ratio (PSNR) is a popular image fidelity criterion, we also use it in this paper. The perceptual weighting of different color components is beyond the scope of this paper, uniform weighting is assumed in this paper. To retain the feature of embeddedness, we progressively interleave the bit-streams for the bit-planes of the three components as shown in Fig. 1. The order of chrominance components is determined by their most significant bit-planes. The chrominance component with higher most significant bit-plane is placed before the other. If the most significant bit-planes of both chrominance components are equally high, U is placed before V.

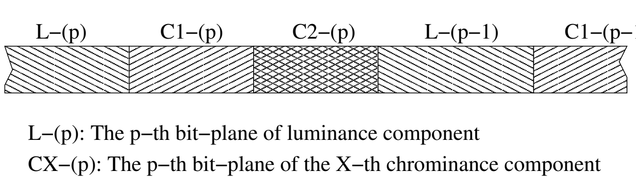

Fig. 1. The composition of the output bit-stream.

Each bit-plane of the luminance component (Y) is encoded using CM-WDR. During the encoding of the p-th bit-plane of Y, contextual information for a node is formed by examining its parent node and neighboring nodes in the higher bit-planes. However, for bit-planes of chrominance components (U and V), the contextual information for a node is formed by examining the corresponding node at the same coordinates of the same bit-plane of Y. This simplifies the context-modeling process and exploits the latest updated information, *i.e.*, p-th bit-plane of Y is used jointly with bit-planes higher than p. This is also the reason that we prefer interleaving the three components at the bit-plane level over interleaving at the coefficient level [6]. If the interleaving occurs at the coefficient level, the latest information about p-th bit-plane of Y can not be used to form contextual information for encoding p-th bit-plane of U and V.

The efficacy of the context-modeling method is demonstrated in Fig. 2. The white pels in Fig. 2(a) and Fig. 2(c) represent the significant coefficients of the U and the V component at the second bit-plane (with threshold value of 4). White pels in Fig. 2(b) and Fig. 2(d) represent those of the U and the V component whose significance fails to be predicted by using the contextual information from Y. It is clear that the simple context modeling approach for chrominance components are very efficient.

The pseudo-code for CWDR is listed as follows:

```
begin
  for i := Y to V do
      P_i := dwt_2d(i);
      B_{i,max} := floor(log_2(max(|P_i|)) od;
  B_max := max({B_{i,max}});
  if B_{U,max} < B_{V,max}
    then encoding_order := Y → V → U
     else encoding_order := Y → U → V fi;
  T := pow(2, B_max); b := B_max;
  while r < R_budget do
        for i := Y to V do
            if b > B_{i,max} then continue fi;
            if i ≠ Y
              then update_UV_scanning_order(P_i, b) fi;
            sorting_pass(P_i, T);
            refinement_pass(P_i, T);
            if i = Y
              then update_scanning_order(P_i, b) fi od;
        T := T/2; b := b − 1 od;
end
```

Note that the sorting pass and refinement pass of CWDR are the same as in CM-WDR. The major difference is in the scanning order updating procedure for chrominance components. When updating the scanning order for U (or V), the nodes whose corresponding nodes in Y are significant are scanned first. Since this simple approach is very efficient, the parent-children relationship and neighborhood re-

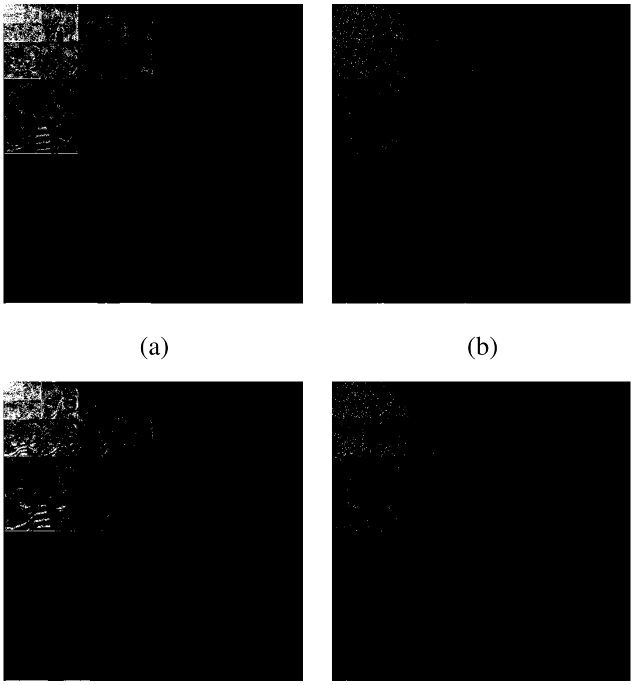

Fig. 2. Prediction efficacy of the context-model used on *Girls* at a decomposition level of 6: (a) the significant coefficients (white pels) at the 2nd bit-pane of U, (b) the significant coefficients in (a) that fail to be identified by context-modeling, (c) the significant coefficients at the 2nd bit-pane of V, (b) the significant coefficients in (c) that fail to be identified by context-modeling.

lationship are not used for U and V, and the complexity of context-modeling is decreased.

4. EXPERIMENTAL RESULTS

In our experiments, three RGB color images of size 512×512, *Girls*, *Lena*, and *Peppers*, have been used [1]. The CEZW and the JPEG-2000 are used as benchmark schemes. Instead of implementing CEZW, we directly downloaded the decoded copies of *Girls* from the authors' public FTP directory [2]. For JPEG-2000, we have used the Kakadu implementation of the standard [8]. To evaluate the fidelity of the decoded images, both the color SNR that is based on the 1964 CIE formula [9] and the PSNR for each component are calculated. The color of pels is represented in CIE modified UCS color space (U^*, V^*, W^*). The color SNR is calculated as follows:

$$SNR = 10 \log_{10} \frac{(\Delta s_0)^2}{(\Delta s)^2} \tag{1}$$

where $(\Delta s)^2$ is the mean square distance between the original image and the decoded image, $(\Delta s_0)^2$ is the mean square distance between the original image and its mean color:

$$\begin{aligned}
(\Delta s)^2 &= \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (\Delta u^*_{ij})^2 + (\Delta v^*_{ij})^2 + (\Delta w^*_{ij})^2 \\
(\Delta s_0)^2 &= \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (u^*_{ij} - u_0)^2 + (v^*_{ij} - v_0)^2 + \\
&\quad (w^*_{ij} - w_0)^2 \\
u_0 &= \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} u^*_{ij} \\
v_0 &= \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} v^*_{ij} \\
w_0 &= \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} w^*_{ij}
\end{aligned} \tag{2}$$

where M and N are the width and the height of the image, respectively.

As shown in Table 1, the color SNR gap between CEZW and CWDR is so amazing that CWDR at 0.5 bpp is better than CEZW at 1.5 bpp. However, that gap between JPEG-2000 and CWDR is much smaller. In the only case (at 1.0 bpp) that CWDR performs worse than JPEG-2000 in terms of color SNR, we can see that the PSNR of Y component of CWDR is 1.55 dB higher than that of JPEG-2000. This result actually favors CWDR because it shows that with CWDR, the mean square error (MSE) of Y lowers from 13 to 9, while the MSEs of U and V suffer a minor increase from 1.5 to 2, and 2.7 to 3, respectively.

Since only the results for *Girls* are available for CEZW, we compared JPEG-2000 and CWDR on other two images. We can see that CWDR always performs better than JPEG-2000 in terms of color SNR. The performance gap increases with the increase of bit-rate. The authors' opinion is that when bit-rate becomes higher, more coefficients become significant and the context model is more efficient for U and V.

Fig. 3 plots the decoded *Peppers* at 0.25 bpp, *i.e.*, a compression ratio of 96:1. The subjective quality of the two images are very similar.

5. CONCLUSION

A novel wavelet embedded coding algorithm has been proposed for color images. It encodes the luminance compo-

[1] http://links.uwaterloo.ca/colorset.base.html

[2] ftp://skynet.ecn.purdue.edu/pub/dist/delp/icip97-coding.

Table 1. Color SNR (dB) and component PSNR (dB) vs. bit rate for CEZW, JPEG-2000, and CWDR on *Girls*.

Bit-rate (bpp)		SNR	Y-PSNR	U-PSNR	V-PSNR
0.5	CEZW	12.19	33.90	39.03	36.97
	JPEG2K	15.37	34.00	43.90	41.70
	CWDR	16.04	34.64	44.61	42.58
1.0	CEZW	15.17	37.38	41.88	39.64
	JPEG2K	18.20	37.09	46.54	43.92
	CWDR	17.46	38.64	44.61	43.15
1.5	CEZW	15.59	40.70	41.88	39.64
	JPEG2K	19.89	38.89	48.08	45.45
	CWDR	20.08	41.20	47.17	45.66

Table 2. Color SNR (dB) and component PSNR (dB) vs. bit rate for JPEG-2000 and CWDR on *Lena*.

Bit-rate (bpp)		SNR	Y-PSNR	U-PSNR	V-PSNR
0.25	JPEG2K	11.43	32.83	33.43	31.13
	CWDR	11.89	32.49	34.54	31.39
0.5	JPEG2K	13.14	34.81	35.12	32.86
	CWDR	13.49	35.69	35.34	33.15
1.0	JPEG2K	14.64	37.54	36.39	33.87
	CWDR	15.00	38.24	36.53	34.95

nent with the CM-WDR algorithm. For the two chrominance components, a simple and efficient context-modeling approach is employed to exploit the statistical dependency among the color components. No explicit rate allocation among color components is needed. Embeddedness and precise rate control are retained. The proposed algorithm performs better than CEZW and JPEG-2000 in terms of color SNR and PSNR.

6. REFERENCES

[1] M. W. Marcellin, M. J. Gormish, A. Bilgin, and M. P. Boliek, "Overview of JPEG-2000," in *Proc. of IEEE DCC*, 2000, pp. 523–541.

[2] A. N. Netravali and C. B. Rubinsein, "Luminance adaptive coding of chrominance signals," *IEEE Trans. Comm.*, vol. 27, no. 4, pp. 703–710, April 1979.

[3] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Acoustics, Speech and Signal Proc.*, vol. 41, no. 12, pp. 3445–3462, December 1993.

Table 3. Color SNR (dB) and component PSNR (dB) vs. bit rate for JPEG-2000 and CWDR on *Peppers*.

Bit-rate (bpp)		SNR	Y-PSNR	U-PSNR	V-PSNR
0.25	JPEG2K	14.97	32.96	33.56	33.26
	CWDR	15.23	32.48	33.28	33.52
0.5	JPEG2K	18.45	35.61	37.79	35.69
	CWDR	18.96	36.14	37.10	36.56
1.0	JPEG2K	21.70	38.52	40.80	38.07
	CWDR	22.41	39.55	40.79	39.29

(a) JPEG-2000 at 0.25 bpp (b) CWDR at 0.25 bpp

Fig. 3. The decoding output of the *Peppers* at compression ratio of 96:1.

[4] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits and Sys. for Video Tech.*, vol. 6, no. 3, pp. 243–250, June 1996.

[5] K. Shen and E. J. Delp, "Color image compression using an embedded rate scalable approach," in *Proc. of IEEE ICIP*, 1997, vol. 3, pp. 34–37.

[6] B.-J. Kim and W. A. Pearlman, "Low-delay embedded 3-D wavelet color video coding with SPIHT," in *Proc. SPIE Visual Comm. Image Proc.*, 1998, vol. 3309, pp. 955–964.

[7] Y. Yuan and M. K. Mandal, "Context-modeled wavelet difference reduction coding based on fractional bit-plane partitioning," in *Proc. of IEEE ICIP*, 2003, vol. 2, pp. 251–254.

[8] D. Taubman, "Kakadu software, Ver. 4.0.3," http://www.kakadusoftware.com/.

[9] A. K. Jain, *Fundamentals of digital image processing*, Prentice Hall, Englewood Cliffs, NJ, 1989.