

COLOR IMAGE CODING BY USING INTER-COLOR CORRELATION

Xing San ^{*}, Hua Cai [†], and Jiang Li [†]

^{*} Dept. of Electronic Eng. and Information Science, University of Science and Technology of China

[†] Media Communication Group, Microsoft Research Asia, Beijing, China

ABSTRACT

Inter-color correlation between the luminance component and chrominance components has been utilized for color image coding for years. However, the correlation has not been clearly analyzed. In this paper, we analyze the inter-color correlation and answer two questions related to color image coding: (1) what kind of inter-color correlation exists in color images after the discrete wavelet transform?; and, (2) how strong is it? This analysis helps us to find a most suitable inter-color context and eventually leads to a new embedded color image codec. By using the discovered inter-color context, significant performance improvement can be achieved when encoding chrominance components.

Index Terms— Color image coding, image coding, inter-color correlation

1. INTRODUCTION

One key problem in image compression is to analyze the statistical characteristics of source data for better entropy coding. After devoting many efforts in this area in the last decade, researchers have come up with several elegant ways for modeling the statistical characteristics. For example, in the embedded zerotree wavelet (EZW) coder [1], coefficients after discrete wavelet transform (DWT) can be represented by a zerotree structure. In the set partitioning in hierarchical trees (SPIHT) coder [2], a so-called spatial orientation tree (SOT) was adopted. And in JPEG2000 [3][4], the values of neighboring pixels are used as context when modeling the statistics of a current pixel.

The above methods were designed and optimized for monochrome images. However, most digital images are colorized. When coding color images, each component is usually coded independently without any inter-color prediction [4][5]. Although being extensions of algorithms for coding monochrome images can simplify coding system design, coding efficiency will be sacrificed since the correlation among different color components is not exploited. In fact, one can often easily recognize the shapes of the chrominance components from that of the luminance component when they are separately displayed. The phenomenon indicates that most chrominance changes in a real scene are accompanied by a luminance change [6]. Having noticed the inter-color correlation among different color components, some researchers began placing their attention on this field. For instance, color-EZW (CEZW) [7] and color-SPIHT (C-SPIHT) [8] were proposed to exploit underlying inter-color correlation by expanding the existing zerotree or SOT structure across the spectral planes. By using the inter-color correlation, these methods obtain a better performance than the traditional coding schemes.

^{*}The work presented in this paper was carried out at Microsoft Research Asia, Beijing, China.



Fig. 1. The five typical color images used for analysis.

Though there exists the inter-color correlation, it does not always mean that a significant chrominance change will be accompanied by a significant luminance change at the same pixel position. Moreover, the correspondence becomes even more random after DWT. Generally speaking, the significant wavelet coefficients generated by the chrominance changes may be found within a neighboring region of the luminance coefficients. In such a situation, the simple *one-one mapping* adopted in CEZW and CSPIHT is not always effective.

This paper studies two questions related to color image coding: (1) what kind of inter-color correlation exists in color images after DWT?; and, (2) how strong is the inter-color correlation? To answer these questions, we extensively analyzed the correlation in Section 2 of this paper. This analysis helps us to find the most suitable inter-color context and eventually leads to a new color image codec. This new codec achieves a significant performance improvement when coding the chrominance components.

The rest of the paper is organized as follows. Section 2 analyzes the inter-color correlation in detail. Based on this analysis, we present our new embedded color image codec in Section 3. The experimental results are reported and discussed in Section 4. Finally, some conclusions are drawn in Section 5.

2. INTER-COLOR CORRELATION ANALYSIS

As we mentioned in Section 1, in most of natural color images a strong relationship exists between a large spatial change in the chrominance signal and a large spatial change in the luminance signal. Hence, after DWT, the relationship in spatial changes usually results in the correspondence in the position of large wavelet coefficients in high-frequency subbands.

To gain a better understanding of the inter-color correlation, we present some numerical analysis in this section. We choose five typical color images (Fig. 1), *Lena*, *Baboon*, *Barbara*, *Goldhill*, and *Peppers*, for analysis. The RGB channels of each image are first decorrelated by means of either irreversible color transform (ICT) [3] or reversible color transform (RCT) [3] to generate the luminance and chrominance components. Each component is then transformed by the Daubechies 9/7-tap filters [9]. In the DWT domain, we analyze the emergence probabilities of the large chrominance coefficients at different positions compared with the large luminance coefficients

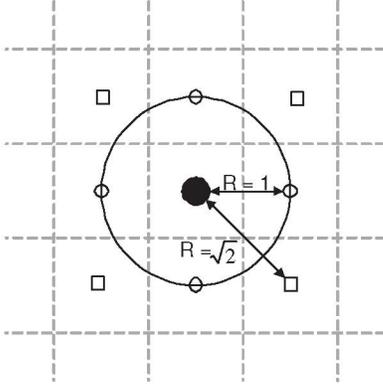


Fig. 2. The definition of R .

(where a coefficient is regarded as large if it is not less than a given threshold). The average emergence probabilities of the five images are measured using the following formula:

$$p(R) = N(R)/Sum(R) \quad (1)$$

Here, R denotes the distance between the position of a chrominance coefficient and that of the nearest large luminance coefficient. $Sum(R)$ denotes the number of chrominance coefficients of which the distance between a certain chrominance coefficient and the corresponding nearest large luminance coefficients is R . $N(R)$ denotes the number of large chrominance coefficients among the $Sum(R)$ chrominance coefficients. Hence $N(R) \leq Sum(R)$. As shown in Fig. 2, $R = 0$ means large luminance coefficient appears exactly at the same position of the chrominance coefficient, i.e., the *one-one mapping* in CEZW and CSPIHT. $R = 1$ means the nearest large luminance coefficients appear at the horizontal or vertical neighboring positions. And $R = \sqrt{2}$ means the nearest large luminance coefficients appear at the neighboring corner positions.

From the statistical results listed in Table 1, several important conclusions can be drawn. First, it is true that after DWT a large chrominance coefficient is often accompanied by large luminance coefficients around the same location. And the correlation increases when the comparison threshold decreases. Second, the inter-color correlation after RCT is usually larger than that after ICT, because ICT provides a more efficient color decorrelation. Third, the emergence probability at $R = 0$, i.e., the *one-one mapping*, has the largest value. That is the reason that CEZW and CSPIHT use this kind of correlation to construct the new zerotree and SOT across the spectral plane and obtain a satisfactory performance. On the other hand, the correlation at $R = 1$ and $R = \sqrt{2}$ is also not weak, which indicates that further performance improvement is still achievable given the good results attained by CEZW and CSPIHT. And fourth, the emergence probability drops quickly when the distance R increases. As can be seen from Table 1, the correlation at $R > \sqrt{2}$ becomes very weak. Therefore, taking the correlation at $R > \sqrt{2}$ into account could not bring significant improvement and thus might be simply ignored in practical design.

We then further compare the conditional entropy with and without inter-color prediction. As shown in Fig. 3, in *method-A*, we choose nine context events that exploit only the correlation in the same chrominance component. Among them, c_1 is the value of the corresponding coefficient in the lower resolution and the same bit-plane; $c_2 \sim c_9$ are the values of the eight neighboring coefficients in the more significant bit-plane. And in *method-B*, we add another

Table 1. Emergence probability $p(R)$ at different thresholds T .

(a) reversible color transform					
T	Comp.	$R = 0$	$R = 1$	$R = \sqrt{2}$	$R > \sqrt{2}$
32	U	0.2422	0.1061	0.0941	0.0246
	V	0.3366	0.2165	0.2416	0.0786
16	U	0.4271	0.3062	0.2342	0.0285
	V	0.5464	0.4319	0.3452	0.0379
8	U	0.5897	0.3999	0.1960	0.0545
	V	0.6445	0.4244	0.1792	0.0547
4	U	0.6889	0.4425	0.2574	0.0815
	V	0.6870	0.4283	0.2520	0.0961

(b) irreversible color transform					
T	Comp.	$R = 0$	$R = 1$	$R = \sqrt{2}$	$R > \sqrt{2}$
32	U	0.1351	0.0720	0.0824	0.0141
	V	0.0881	0.0104	0.0045	0.0006
16	U	0.3534	0.2318	0.1451	0.0078
	V	0.2170	0.0746	0.0421	0.0022
8	U	0.4779	0.2391	0.0443	0.0036
	V	0.3880	0.1598	0.0332	0.0039
4	U	0.5235	0.2065	0.0555	0.0145
	V	0.5026	0.1930	0.0547	0.0143

four context events that are from the current bit-plane of the luminance component. As can be seen from Fig. 4, c_{10} is the significance of the luminance coefficient at the same position (i.e. $R = 0$); c_{11} and c_{12} are of respective significance of the vertical two neighboring luminance coefficients and horizontal two neighboring luminance coefficients (i.e. $R = 1$); and c_{13} is the significance of the four neighboring corner luminance coefficients (i.e. $R = \sqrt{2}$). Note that each of the context events has two states: *zero* and *one*. If the corresponding coefficient of the context event is significant, the state is set to *one*, otherwise it is set to *zero*. In addition, for $c_{11} \sim c_{13}$, if any one of the corresponding coefficients is significant, the state of that context event is set to *one*.

We choose the bit-plane of threshold 16 of the image *Lena* after ICT and DWT for the analysis. The conditional entropy of the component U and V are shown respectively in Fig. 5 and Fig. 6, which is obtained by calculating the conditional entropy of the n most important context events. It is clear from these figures that the conditional entropy of *method-B* drops faster than that of *method-A*. The faster downtrend indicates that the context events of inter-color correlation are even stronger than some context events from the neighboring samples of the same component.

3. IMPLEMENTATION FOR EMBEDDED COLOR IMAGE CODING

From the analysis given in Section 2, it is clear that the corresponding 3×3 neighboring luminance coefficients can be used as context when coding a chrominance coefficient. In this section, we discuss how to design an embedded color image codec that exploits the inter-color correlation.

In our scheme, the original RGB image is first transformed to YUV components using either ICT or RCT. Then, DWT is individually performed on each component. After that, the transformed coefficients are separated into a series of bit-planes. Starting from the most significant bit-plane, three coding passes, i.e., Zero Coding, Sign Coding, and Magnitude Refinement are applied for each

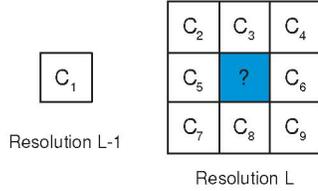


Fig. 3. The 9 context events used in *method-A*.

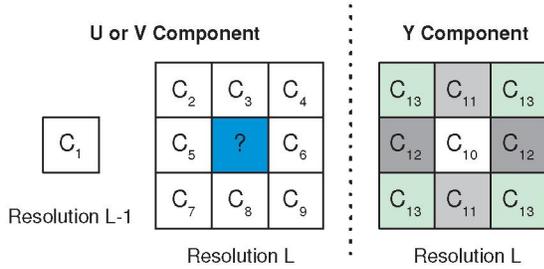


Fig. 4. The 13 context events used in *method-B*.

bit-plane with a MQ arithmetic coding engine [10].

1) *Zero Coding*: The zero coding pass encodes those coefficients, which become significant at the current bit-plane. It is the most important pass among the three passes for an embedded wavelet encoder. Considering the sensitivity of visual systems on the luminance component, the coding priority of luminance should be higher than chrominance in a same bit-plane. So a coding order of Y, U, and V is adopted within each bit-plane. In such a situation, when we encode the U or V component, the positions of the significant coefficients of the Y component in the current bit-plane are known. Hence, it is possible to exploit the inter-color correlation. When encoding the Y component, we choose the nine context events used in *method-A* described before for context-based arithmetic coding. And we choose the thirteen context events used in *method-B* for encoding the U and V components.

2) *Sign Coding*: Once a coefficient is found significant in the zero coding pass, its sign value should be coded in this pass. Here, we choose the same fixed contexts as that of JPEG2000 [3].

3) *Magnitude Refinement*: This pass refines the coefficients that are found significant in the more significant bit-plane(s). We use only two conditioning states for each component: a one state is set if the coefficient just changed to be significant in the nearest more significant bit-plane; otherwise the state is set to zero.

4. EXPERIMENTAL RESULTS

We choose three standard 24-bit color images, *Lena* (512×512), *Peppers* (512×512), and *Monarch* (768×512), as test images. After the color transform, the YUV components of each image is carried out a six-stage dyadic wavelet decomposition using the Daubechies 9/7-tap filters [9]. The wavelet coefficients are then encoded using our newly embedded color image codec.

We first evaluate the performance when using the inter-color relation to encode the chrominance components. As a comparison bench-mark, the performance without inter-color prediction (i.e. using the nine context events only) is also tested. From the results listed in Table 3, we can conclude that about 5% ~ 15% bit rate reduction

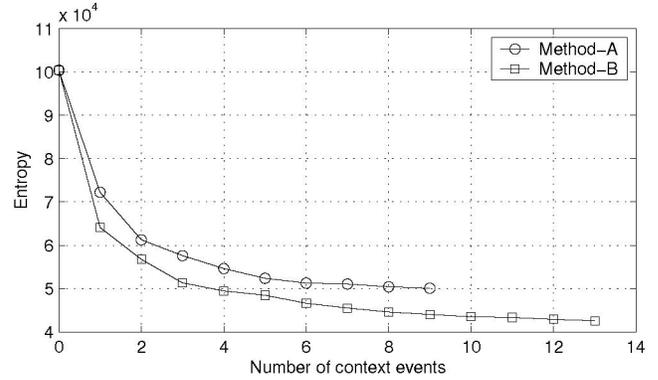


Fig. 5. Conditional entropy of U component.

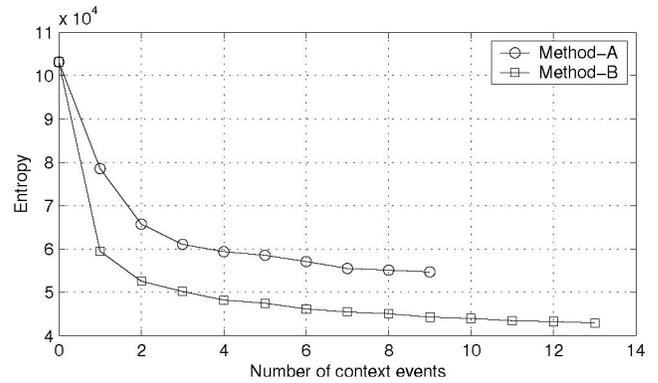


Fig. 6. Conditional entropy of V component.

on average is achievable by using the inter-color correlation.

We also compare the performance of the proposed color image codec with that of the Jasper implementation of JPEG2000 [11]. In the comparison, images were encoded and decoded by the Jasper codec in *real* mode with the same bit rate of our output bitstream. And, ICT color transform is used in our codec. To ensure a common basis of comparison, the decoded RGB image is measured with the following PSNR formula:

$$PSNR = 10 \log_{10} \frac{255^2 \times 3}{[MSE(R) + MSE(G) + MSE(B)]} \quad (2)$$

As shown in Table 2, the proposed codec achieves a higher performance than JPEG2000.

5. CONCLUSIONS AND DISCUSSIONS

This paper analyzes the inter-color correlation of color images and identifies the most important inter-color contexts between the luminance component and the chrominance components. Based on the analysis, a new color image coding scheme is proposed and evaluated. Significant bit rate reduction can be achieved after employing the inter-color correlation.

In addition to conventional color images, the discovered inter-color correlation can also be applied for multi-spectral images. This is one of our likely future research goals.

Table 2. Performance evaluation of the new color image codec compared with Jasper. The bitstream is truncated at bit-plane of threshold 16, 8, and 4, respectively. In the table, *Length* denotes the total number of bytes after truncation (including all of the overhead bits).

Image	Length (Bytes)	PSNR(dB)	Jasper PSNR(dB)
Lena	7138	30.51	30.17
	15981	32.99	32.85
	38610	35.48	35.43
Peppers	9081	29.64	29.18
	18989	31.82	31.54
	50932	34.05	33.91
Monarch	14370	31.19	30.78
	26940	35.09	34.76
	46210	38.94	38.75

6. REFERENCES

- [1] J. M. Shapiro, "Embedded image coding using zerotrees of wavelets coefficients," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3445–3462, Dec. 1993.
- [2] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, pp. 243–250, June 1996.
- [3] Majid Rabbani and Rajan Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, pp. 3–48, 2002.
- [4] ISO/IEC 15444-1, "Information technology—JPEG2000—Image coding system—Part 1: Core coding system," 2000.
- [5] G. K. Wallace, "The JPEG still picture compression standard," *Comm. of the ACM*, vol. 34, pp. 30–44, Apr. 1991.
- [6] J. O. Limb and C. B. Rubinstein, "Plateau coding of the chrominance component of color picture signals," *IEEE Transactions on Communications*, vol. 22, pp. 812–820, June 1974.
- [7] K. Shen and E. J. Delp, "Color image compression using an embedded rate scalable approach," in *Proc. of IEEE International Conference on Image Processing*, Santa Barbara, 1997.
- [8] A. A. Kassim and W. S. Lee, "Embedded color image coding using SPIHT with partially linked spatial orientation trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 203–206, Feb. 2003.
- [9] Marc Antonini, Michel Barlaud, Pierre Mathieu, and Ingrid Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, pp. 205–220, Apr. 1992.
- [10] ISO/IEC 14492, "Information technology – coded representation of picture and audio information – lossy/lossless coding of bi-level images (JBIG 2)," *ISO/IEC/JTC 1/SC 29/WG 1*, 1999.
- [11] M. Adams and F. Kossentini, "Jasper: A software-based JPEG-2000 codec implementation," *Proc. IEEE Int. Conf. Image Processing*, vol. 2, pp. 53–56, Sept. 2000.

Table 3. Coding length of chrominance components. The bitstream is truncated at bit-plane of threshold 16, 8, and 4. The *Length-1* and *Length-2* respectively denote the length of truncated output bitstream with the nine context events and the thirteen context events.

(a) Coding length of U component with ICT color transform

Image	Length-1 (Bytes)	Length-2 (Bytes)	Gain
Lena	686	635	7.43 %
	1787	1618	9.46 %
	5341	4986	6.65 %
Peppers	1306	1182	9.94 %
	2908	2589	10.97 %
	8975	8320	7.30 %
Monarch	1205	1149	4.65 %
	2294	2173	5.27 %
	4175	3963	5.08 %

(b) Coding length of V component with ICT color transform

Image	Length-1 (Bytes)	Length-2 (Bytes)	Gain
Lena	688	638	7.27 %
	1953	1721	11.88 %
	5773	5038	12.73 %
Peppers	1613	1387	14.01 %
	3448	2974	13.75 %
	10361	9594	7.40 %
Monarch	1280	1208	2.63 %
	2506	2359	5.87 %
	4594	4353	5.25 %

(c) Coding length of U component with RCT color transform

Image	Length-1 (Bytes)	Length-2 (Bytes)	Gain
Lena	1956	1656	15.33 %
	5612	4979	11.28 %
	18530	17372	6.25 %
Peppers	3990	3604	9.67 %
	11861	11263	5.04 %
	31907	30357	4.86 %
Monarch	2671	2575	3.59 %
	4908	4702	4.20 %
	8893	8568	3.65 %

(d) Coding length of V component with RCT color transform

Image	Length-1 (Bytes)	Length-2 (Bytes)	Gain
Lena	1825	1514	17.04 %
	5407	4828	10.71 %
	18703	181010	3.71 %
Peppers	3202	2646	17.36 %
	9996	9046	9.50 %
	31223	30103	3.59 %
Monarch	2458	2148	12.61 %
	4434	3912	11.78 %
	8017	7284	9.14 %