Hierarchical Prediction and Context Adaptive Coding for Lossless Color Image Compression

Seyun Kim, *Student Member, IEEE*, and Nam Ik Cho, *Senior Member, IEEE*

***Abstract*— This paper presents a new lossless color image compression algorithm, based on the hierarchical prediction and context-adaptive arithmetic coding. For the lossless compression of an RGB image, it is first decorrelated by a reversible color transform and then *Y* component is encoded by a conventional lossless grayscale image compression method. For encoding the chrominance images, we develop a hierarchical scheme that enables the use of upper, left, and lower pixels for the pixel prediction, whereas the conventional raster scan prediction methods use upper and left pixels. An appropriate context model for the prediction error is also defined and the arithmetic coding is applied to the error signal corresponding to each context. For several sets of images, it is shown that the proposed method further reduces the bit rates compared with JPEG2000 and JPEG-XR.**

***Index Terms*— Lossless color image compression, reversible color transform, hierarchical prediction, context adaptive arithmetic coding.**

I. Introduction

Digital images are usually encoded by lossy compression methods due to their large memory or bandwidth requirements. The lossy compression methods achieve high compression ratio at the cost of image quality degradation. However, there are many cases where the loss of information or artifacts due to compression needs to be avoided, such as medical, prepress, scientific and artistic images. As cameras and display systems are going high quality and as the cost of memory is lowered, we may also wish to keep our precious and artistic photos free from compression artifacts. Hence efficient lossless compression will become more and more important, although the lossy compressed images are usually satisfactory in many cases.

Along with the standardization or independently, many lossless image compression algorithms have been proposed. Among a variety of algorithms, the most widely used ones may be Lossless JPEG [1], JPEG-LS [2], LOCO-I [3], CALIC [4], JPEG2000 [5] (lossless mode) and JPEG XR [6]. The LOCO-I and CALIC were developed in the process of JPEG standardization, where most ideas in LOCO-I are accepted for the JPEG-LS standard although the CALIC provides better compression performance at the cost of more computations. For the compression of color images, the color components are first decorrelated by a color transform, and each of the transformed components is independently compressed by the above referenced methods. For example, the RGB to YC_bC_r transform [7] may be the most frequently used one for the lossy compression of color image and video. However, in the case of lossless compression, most color transforms cannot be used due to their uninvertibility with integer arithmetic. Hence an invertible version of color transform, the reversible color transform (RCT) was defined and used in JPEG2000 [5]. There have also been much research for finding better RCTs [8]–[10], among which we adopt a transform proposed in [9] because it approximates the conventional YC_bC_r transform very well.

The purpose of this paper is to develop a hierarchical prediction scheme, while most of existing prediction methods in lossless compression are based on the raster scan prediction which is sometimes inefficient in the high frequency region. The "hierarchical" prediction for the compression was already proposed in [11], but only pixel interpolation is used here. In this paper, we design an edge directed predictor and context adaptive model for this hierarchical scheme. To be specific, we propose a method that can use lower row pixels as well as the upper and left pixels for the prediction of a pixel to be encoded. For the compression of color images, the RGB is first transformed to YC_uC_v by an RCT mentioned above [9], and Y channel is encoded by a conventional grayscale image compression algorithm. In the case of chrominance channels (C_u and C_v), the signal variation is generally much smaller than that of RGB, but still large near the edges. For more accurate prediction of these signals, and also for accurate modeling of prediction errors, we use the hierarchical scheme: the chrominance image is decomposed into two subimages; i.e. a set of even numbered rows and a set of odd numbered rows respectively. Once the even row subimage X_e is encoded, we can use all the pixels in X_e for the prediction of a pixel in the odd row subimage X_o. In addition, since the statistical properties of two subimages are not much different, the pdf of prediction errors of a subimage can be accurately modeled from the other one, which contributes to better context modeling for arithmetic coding. Experiments on various kinds of images are performed, and it is shown that the proposed

Manuscript received September 11, 2013; revised November 18, 2013; accepted November 21, 2013. Date of publication December 3, 2013; date of current version December 12, 2013. This work was supported in part by Samsung Electronics and in part by the Ministry of Science, ICT and Future Planning, Korea, through the Information Technology Research Center support Program supervised by the National IT Industry Promotion Agency under Grant NIPA-2013-0301-13-4005. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Ali Bilgin.

S. Kim is with Samsung Electronics, Yongin 151-744, Korea (e-mail: light4u@ispl.snu.ac.kr).

N. I. Cho is with the Department of Electrical and Computer Engineering, INMC, Seoul National University, Seoul 446-711, Korea (e-mail: nicho@snu.ac.kr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2293428

Algorithm 1 Calculation of $dir(i, j)$

if $|x_o(i,j) - \hat{x_h}(i,j)| + T_1 < |x_o(i,j) - \hat{x_v}(i,j)|$ **then**
  $dir(i,j) \leftarrow H$
else
  $dir(i,j) \leftarrow V$
end if

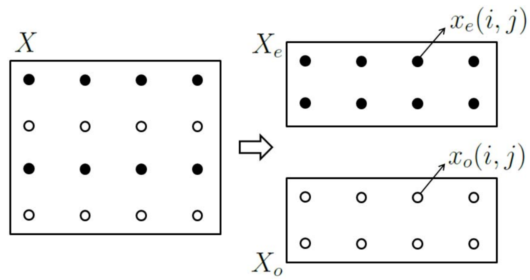


Fig. 1. Input image and its decomposition.

method provides higher coding gain than JPEG2000 and JPEG-XR in many cases.

II. Hierarchical Decomposition and Pixel Prediction

The chrominance channels C_u and C_v resulting from the RCT usually have different statistics from Y, and also different from the original color planes R, G, and B. In the chrominance channels, the overall signal variation is suppressed by the color transform, but the variation is still large near the object boundaries. Hence, the prediction errors in a chrominance channel are much reduced in a smooth region, but remain relatively large near the edge or within a texture region.

For the efficient lossless compression, it is important to accurately estimate the pdf of prediction error for better context modeling, along with the accurate prediction. For this, we propose a hierarchical decomposition scheme as depicted in Fig. 1, which shows that pixels in an input image X is separated into two subimages: an even subimage X_e and an odd subimage X_o. Then, X_e is encoded first and is used to predict the pixels in X_o. In addition, X_e is also used to estimate the statistics of prediction errors of X_o. In actual implementation, X_e is decomposed once more as will be explained later.

For the compression of X_o pixels using X_e, directional prediction is employed to avoid large prediction errors near the edges. For each pixel $x_o(i, j)$ in X_o, the horizontal predictor $\hat{x_h}(i, j)$ and vertical predictor $\hat{x_v}(i, j)$ are defined as

$$\begin{aligned} \hat{x_h}(i, j) &= x_o(i, j-1) \\ \hat{x_v}(i, j) &= round\left(\frac{x_e(i, j) + x_e(i+1, j)}{2}\right), \end{aligned} \tag{1}$$

and one of them is selected as a predictor for $x_o(i, j)$. With these two possible predictors, the most common approach to encoding is "mode selection," where better predictor for each pixel is selected and the mode (horizontal or vertical) is also transmitted as side information. However, the vertical predictor is more often correct than the horizontal one when the predictors are defined as (1) because upper and lower pixels are used for the "vertical" whereas just a left pixel is used for the "horizontal." The horizontal predictor is more accurate only when there is a strong horizontal edge. For example, the frequency of selecting horizontal predictor is just 0.03% $\sim$ 1.45% for the images in Kodak set [13] which is one of the image sets used in the experiments. Hence, the vertical predictor is used for most pixels, and mode selection is used only when the pixel seems to be on a strong horizontal edge.

Algorithm 2 Calculation of $\hat{x_o}(i, j)$

if $dir(i-1,j) = H$ or $dir(i,j-1) = H$ **then**
  Calculate $dir(i,j)$ by Algorithm 1
  Encode $dir(i,j)$
  if $dir(i,j) = H$ **then**
    $\hat{x_o}(i,j) \leftarrow \hat{x_h}(i,j)$
  else
    $\hat{x_o}(i,j) \leftarrow \hat{x_v}(i,j)$
  end if
else
  $\hat{x_o}(i,j) \leftarrow \hat{x_v}(i,j)$
  Calculate $dir(i,j)$ by Algorithm 1
end if

For implementing this idea, we define a variable for the direction of edge at each pixel $dir(i, j)$, which is given either H or V. Actually, it is given H only when the horizontal edge is strong, and given V for the rest. Deciding $dir(i, j)$ is summarized in Algorithm 1, where it can be seen that the direction is given H only when $|x_o(i, j) - \hat{x}_h(i, j)|$ is much smaller than $|x_o(i, j) - \hat{x_v}(i, j)|$ by adding a constant T_1 to the former when comparing them.

Based on the directions of pixels, the overall prediction scheme is summarized in Algorithm 2. It can be seen that the mode selection is tried when more than one of $dir(i-1, j)$ or $dir(i, j-1)$ are H, and the vertical prediction is performed for the rest.

III. Proposed Coding Scheme

In this section, we explain the overall process of image compression, including the new encoding scheme. An input RGB color image is transformed into YC_uC_v color space by an RCT. The luminance image Y is encoded by any of lossless grayscale image coders, such as CALIC, JPEG-LS, or JPEG2000 lossless. The chrominance images C_u and C_v are encoded using the method described in Section II. To be specific, a chrominance image $X^{(0)} \in \{C_u, C_v\}$ is decomposed row by row into an even subimage $X_e^{(1)}$ and an odd subimage $X_o^{(1)}$ as shown in Fig. 2. The subimage $X_o^{(1)}$ is predicted and encoded using $X_e^{(1)}$, as described in Section II. The subimage $X_e^{(1)}$ can be further decomposed column by column into the even subimage $X_e^{(2)}$ and the odd subimage $X_o^{(2)}$ as shown in the last figure of Fig. 2, where the subimage $X_o^{(2)}$ is compressed using $X_e^{(2)}$.

In the predictive lossless compression, efficient encoding of the prediction error $e(i, j) = x_o(i, j) - \hat{x_o}(i, j)$ plays an important role. Although the proposed prediction method usually generates small prediction errors owing to the RCT and

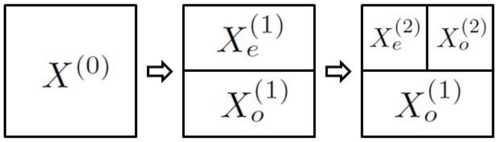
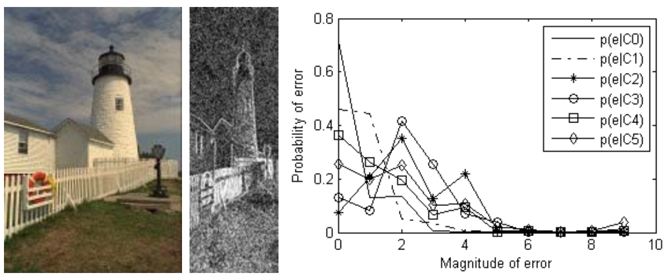


Fig. 2. Illustration of hierarchical decomposition.

TABLE I

AVERAGE OF COMPRESSED BIT RATES (bpp) FOR 24 KODAK IMAGES

	BPP
JPEG2000	9.5353
JPEG2000 with RCT [9]	9.4586
JPEG-XR	10.9214
JPEG-XR with RCT [9]	10.8521
Proposed	8.8587



(a) (b) (c)

Fig. 3. An example of context and pdf of error depending on the context. (a) Input image. (b) Context. (c) Conditional pdf.

the sophisticated prediction scheme, there are still relatively large errors near the edge or texture region, which degrades the compression performance. For the efficient compression, the statistics of symbols (prediction errors) should well be described by an appropriate model and/or parameters. We model the prediction error as a random variable with pdf $P(e|C_n)$, where C_n is the coding context that reflects the magnitude of edges and textures. Specifically, C_n is the level of quantization steps of pixel activity $\sigma(i, j)$ defined as

$$\sigma(i, j) = |x_e(i, j) - x_e(i + 1, j)|. \tag{2}$$

Note that the local activity and its quantization steps are calculated with the pixels in X_e, because all the pixels of X_e are available and its statistical property would be almost the same as that of X_o. The local activity is quantized into K steps such that C_n represents the step

$$q_{n-1} \leq \sigma(i, j) < q_n \tag{3}$$

for $n = 1, \ldots, K$ with $q_0 = 0$ and $q_K = \infty$. The length of quantization steps is determined such that each step includes the same number of elements (local activities). For each context, a generic adaptive arithmetic coder [12] is used to encode the prediction error. For illustration, Fig. 3 shows an input image, the local activity of a subimage (context), and $P(e|C_n)$ for several C_n. It describes the statistical property of prediction error very well, in that the error magnitude is large when the local activity is strong. Hence the proposed model can be effective for the compression with arithmetic coding.

TABLE II

COMPRESSED BIT RATES (bpp) FOR THE MEDICAL IMAGES

	Size	JPEG2000	JPEG-XR	Proposed
PET1	256 × 256	6.7390	8.0839	**5.6453**
PET2	256 × 256	7.3403	8.5533	**6.1598**
PET3	256 × 256	7.0232	8.4425	**5.8768**
Eye1	3216 × 2136	5.7498	7.4635	**4.6208**
Eye2	3216 × 2136	5.4467	7.3490	**4.3350**
Eyeground	1600 × 1216	3.2763	5.6944	**2.9656**
Endoscope1	603 × 552	7.3532	8.6395	**7.0451**
Endoscope2	568 × 506	5.1304	7.2928	**4.8968**
Avg.		6.0074	7.6899	5.1932

TABLE III

COMPRESSED BIT RATES (bpp) FOR THE COMMERCIAL DIGITAL CAMERA IMAGES

	Size	JPEG2000	JPEG-XR	Proposed
Ceiling	4288 × 2848	7.5571	8.8331	**7.2080**
Locks	4288 × 2848	7.4574	8.8296	**7.1623**
Flamingo	4288 × 2848	7.0366	8.2698	**6.6371**
Berry	4288 × 2848	7.2468	8.6646	**6.8917**
Sunset	4288 × 2848	6.3586	7.9263	**5.9700**
Flower	4032 × 3024	6.4141	8.1298	**6.0655**
Park	4032 × 3024	5.8977	7.6534	**5.5622**
Fireworks	4032 × 3024	5.7797	7.4469	**5.2855**
Avg.		6.7185	8.2192	6.3478

IV. EXPERIMENTAL RESULTS

As stated in the introduction, the state-of-the-art lossless compression method may be the CALIC [4], which shows higher coding gain than the JPEG-LS (or LOCO-I) [2], [3] at the cost of higher computational complexity. For the compression of color image, the JPEG2000 and JPEG-XR [6] lossless provide better coding gain than the independent encoding of each channel by CALIC and also than the encoding by CALIC after RCT. Hence we compare the proposed method with JPEG2000 and JPEG-XR. The excutables for our encoder/decoder and all the images used in the experiments are publicly available at our website [18].

We first apply the algorithm on Kodak image set [13], which is widely used for the test of lossless compression [14]–[16] and demosaicking [17]. In all the experiments, the parameter T_1 in Algorithm 2 and number of contexts K are set to 3 and 6 respectively. The luminance images and decomposed highest level images $X_e^{(2)}$ in Fig. 2 are encoded by JPEG2000 lossless. Experiments are summarized in Table I, which shows that the proposed method performs better than the compared methods.

It should also be noted that different color transforms are used in each of the methods stated above. Hence, for fair comparison, we also perform experiments with the same RCT defined in [9], the results of which are denoted as "JPEG2000 with RCT [9]" and "JPEG-XR with RCT [9]" in Table I. It can be seen that the recent RCT improves the coding gain though not significant. On the average, the proposed algorithm improves 7.10% and 18.89% over JPEG2000 and JPEG-XR respectively.

The proposed method is also tested on medical images in Fig. 4 and compared with JPEG2000 and JPEG-XR in Table II. The test medical images are positron emission tomography

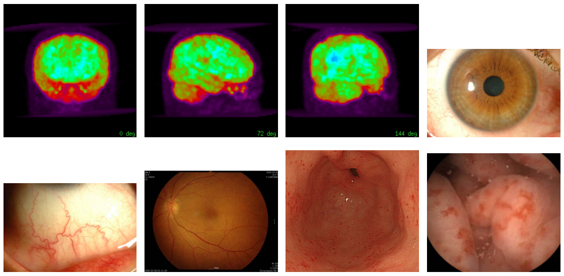


Fig. 4. The medical images.

Fig. 5. The digital camera images.

TABLE IV

COMPRESSED BIT RATES (bpp) FOR CLASSIC IMAGES

	Size	CALIC	JPEG2000	JPEG-XR	Proposed
Lena	512 × 512	**13.1787**	13.5848	14.0942	13.6461
Peppers	512 × 512	**13.8661**	14.8000	15.3245	15.2102
Mandrill	512 × 512	18.1511	**18.0939**	18.2553	18.5305
Barbara	640 × 512	14.9567	**11.1612**	12.1408	11.4575
Avg.		15.0392	14.4100	14.9537	14.7111

(PET) images for human brain, digital camera images for eyes and eyeground, and endoscope images for human intestine, which are generally smooth and hence less bits are generated when compared with the case of Kodak images. On the average, the proposed algorithm produced 13.55% less bits than JPEG2000 lossless. In addition, experiments for images from commercial digital cameras (shown in Fig. 5) are also conducted, and the results are compared in Table III. The first five images are captured with NIKON D90, and the rest are captured with OLYMPUS E-P1. On the average, the proposed algorithm produces 5.52% less bits than JPEG2000 lossless.

It is also noted that the proposed method does not always perform best for every set of images. The proposed hierarchical encoding scheme sometimes works better and sometimes worse than the conventional methods, depending on image sets and also depending on the channels (Y, C_u, and C_v). It is also true for every compression algorithms, i.e. the coding gain of compression algorithms differ on different set of images. For example, on the set of classical test images such as Lena, Peppers, and Mandrill, even the channel independent CALIC sometimes performs better than JPEG2000 and our algorithm, as shown in Table IV.

Finally, the CPU times taken by the above stated methods are measured for 24 Kodak images, and their averages are summarized in Table V. It shows that the JPEG2000 spends CPU time about 2 times more than JPEG-XR. Since our method employs JPEG2000 and needs additional steps for hierarchical prediction and context modeling, it needs slightly more computation time than the JPEG2000.

TABLE V

COMPARISON OF CPU TIMES (SECONDS) ON A PC WITH INTEL CORE-I5 2.67 GHZ CPU

	JPEG2000	JPEG-XR	Proposed
Endocing Time	0.8125	0.3491	0.8835
Decoding Time	0.6945	0.4617	0.7908

V. CONCLUSION

We have proposed a lossless color image compression method based on a hierarchical prediction scheme and context-adaptive arithmetic coding. For the compression of an RGB image, it is first transformed into YC_uC_v color space using an RCT. After the color transformation, the luminance channel Y is compressed by a conventional lossless image coder. Pixels in chrominance channels are predicted by the hierarchical decomposition and directional prediction. Finally, an appropriate context modeling of prediction residuals is introduced and arithmetic coding is applied. The proposed method and several conventional methods have been tested on the Kodak image set, some medical images, and digital camera images, and it is shown that average bit rate reductions over JPEG2000 for these sets are shown to be 7.10%, 13.55%, and 5.52% respectively.

REFERENCES

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

[2] *Information Technology—Lossless and Near-Lossless Compression of Continuous-Tone Still Images (JPEG-LS)*, ISO/IEC Standard 14495-1, 1999.

[3] M. Weinberger, G. Seroussi, and G. Sapiro, “The LOCO-I lossless image compression algorithm: Principles and standardization into JPEG-LS,” *IEEE Trans. Image Process.*, vol. 9, no. 8, pp. 1309–1324, Aug. 2000.

[4] X. Wu and N. Memon, “Context-based, adaptive, lossless image coding,” *IEEE Trans. Commun.*, vol. 45, no. 4, pp. 437–444, Apr. 1997.

[5] *Information Technology—JPEG 2000 Image Coding System—Part 1: Core Coding System*, INCITS/ISO/IEC Standard 15444-1, 2000.

[6] *ITU-T and ISO/IEC, JPEG XR Image Coding System—Part 2: Image Coding Specification*, ISO/IEC Standard 29199-2, 2011.

[7] G. Sullivan, “Approximate theoretical analysis of RGB to YCbCr to RGB conversion error,” ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6 document JVT-I017, 2003.

[8] H. S. Malvar, G. J. Sullivan, and S. Srinivasan, “Lifting-based reversible color transformations for image compression,” *Proc. SPIE*, vol. 707307, pp. 707307-1–707307-10, Aug. 2008.

[9] S. Pei and J. Ding, “Improved reversible integer-to-integer color transforms,” in *Proc. 16th IEEE ICIP*, Nov. 2009, pp. 473–476.

[10] T. Strutz, “Adaptive selection of colour transformations for reversible image compression,” in *Proc. 20th Eur. IEEE Signal Process. Conf.*, Aug. 2012, pp. 1204–1208.

[11] P. Roos, M. A. Viergever, M. C. A. van Dijke, and J. H. Peters, “Reversible intraframe compression of medical images,” *IEEE Trans. Med. Imag.*, vol. 7, no. 4, pp. 328–336, Dec. 1988.

[12] A. Said, “Arithmetic coding,” in *Lossless Compression Handbook*, K. Sayood, Ed. San Diego, CA, USA: Academic, 2003.

[13] (1991). *Images from KODAK Photo CD Photo Sampler* [Online]. Available: http://www.site.uottawa.ca/~edubois/demosaicking

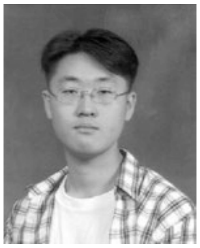
[14] Z. Mai, P. Nasiopoulos, and R. Ward, “A wavelet-based intra-prediction lossless image compression scheme,” in *Proc. Int. Conf. Consum. Electron.*, Jan. 2009, pp. 1–2.

[15] H. S. Malvar and G. J. Sullivan, “Progressive-to-lossless compression of color-filter-array images using macropixel spectral-spatial transformation,” in *Proc. DCC*, Apr. 2012, pp. 3–12.

[16] N. Zhang and X. Wu, “Lossless compression of color mosaic images,” *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1379–1388, Jun. 2006.

[17] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, “Color plane interpolation using alternating projections,” *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 997–1013, Sep. 2002.

[18] (Dec. 3, 2013) [Online]. Available: http://ispl.snu.ac.kr/light4u/project/LCIC



Seyun Kim received the B.S. and Ph.D. degrees from Seoul National University, Seoul, Korea, in 2002 and 2013, respectively. He is currently a Senior Engineer with Samsung Electronics, Yongin, Korea. His research interests include image processing, image compression, and computer vision.

Nam Ik Cho received the B.S., M.S., and Ph.D. degrees in control and instrumentation engineering from Seoul National University, Seoul, Korea, in 1986, 1988, and 1992, respectively. From 1991 to 1993, he was a Research Associate with the Engineering Research Center for Advanced Control and Instrumentation, Seoul National University. From 1994 to 1998, he was with the University of Seoul, Seoul, as an Assistant Professor of electrical engineering. He joined the Department of Electrical and Computer Engineering, Seoul National University, in 1999, where he is currently a Professor. His research interests include image processing, adaptive filtering, and computer vision.