Two-Layer Lossless Coding of HDR Images Specialized for Radiance Format

Kai Yang*, Taizo Suzuki*, and Taichi Yoshida†
* University of Tsukuba, Ibaraki, Japan
E-mail: yang@wmp.cs.tsukuba.ac.jp, taizo@cs.tsukuba.ac.jp
† The University of Electro-Communications, Tokyo, Japan
E-mail: t-yoshida@uec.ac.jp

***Abstract*—Two-layer coding of high dynamic range (HDR) images, which consists of a base layer for encoding the tone-mapped low dynamic range (LDR) image and an enhancement layer for encoding the residual signals between LDR and HDR images, is a constructive solution as a coding framework that can simultaneously provide the LDR and HDR images. This study proposes a two-layer coding that greatly improves the compression performance while satisfying the reversibility for the Radiance format of HDR images. Specifically, in the enhancement layer, the image is decomposed into mantissa and exponential information by utilizing the integer conversion (RGBE converter) of the Radiance format and the luminance and chrominance components of the residual signals are sparsified by the proposed exponent-adjusted mantissa prediction. The experimental results of HDR image coding show that the proposed method improved the average bitrate by 1.41 bpp compared with the best existing method thanks to the specialized scheme to the Radiance format.**

I. Introduction

High dynamic range (HDR) image, which has a high contrast ratio between the maximum luminance and the minimum luminance of the image, has been widely used in many fields such as photography, computer graphics, and filming [1], [2]. However, since most existing devices cannot display HDR images accurately, a technology called tone mapping [3–8] that maps an HDR image to a low dynamic range (LDR) image with 24-bit full color information is essential. On the other hand, inverse tone mapping, which transforms an LDR image into an HDR image, also plays an important role in two-layer coding of HDR images described later.

The Radiance and OpenEXR formats [9], [10] that represent floating point numbers have been mainly used to encode HDR images. To compress the HDR images more efficiently, Xu et al. proposed a logarithmic method that encodes an HDR image with the JPEG 2000 after logarithmically converting it from floating-point numbers to an integers [11]. However, it does not guarantee true reversibility of HDR images due to calculation loss. Also, the Radiance format, OpenEXR format, and logarithmic method require a new tone mapping architecture to provide LDR images for most existing devices.

Recently, two-layer coding of HDR images, already employed by the international standard JPEG XT, is becoming mainstream [12–15]. It consists of a base layer for encoding the tone-mapped LDR image and an enhancement layer for encoding the residual signals between the HDR and LDR images and has two properties as follows:

- Image selectivity: LDR images can be provided to general users while providing HDR images to high-end users such as photographers, designers, and medical staff.
- Backward compatibility: Arbitrary general coders such as JPEG and JPEG 2000 can be directly used in the coder parts with almost no modification.

In addition to the above properties, the reversibility of HDR images can also be realized when both the integer conversion and coder in the enhancement layer are reversible.

As previous studies on two-layer coding, Mai et al. proposed a tone mapping optimization method to compress an HDR image efficiently at the expense of quality of the tone-mapped LDR image [12]. On the other hand, several methods, which allow any tone mapping and inverse tone mapping, have been introduced [13–15]. Watanabe et al. employed the histogram packing (HP) method [16] in the enhancement layer based on the framework of JPEG XT (referred as "basic method") and the lossy logarithmic method [13]. Iwahashi et al. presented a reversible integer conversion instead of the lossy logarithmic method to guarantee reversibility [14]. In addition, Yoshida et al. compressed the energy of the residual signals efficiently by the HP method and the adaptive inverse tone mapping (global approximation and gradation prediction) in the enhancement layer while guaranteeing reversibility based on the integer conversion of JPEG XR [15]. However, in the Yoshida's method, the performance on the Radiance format was not as good as on the OpenEXR format and the gradation prediction requires high computational cost on the decoder.

This study proposes an efficient two-layer coding of HDR images that guarantees reversibility and has better compression performance for the Radiance format. Specifically, in the enhancement layer, the image is decomposed into mantissa and exponential information by utilizing the integer conversion (RGBE converter) of the Radiance format and the luminance and chrominance components of the residual signals are sparsified by the proposed exponent-adjusted mantissa prediction. Of course, the computational cost on the decoder is not high unlike the gradation prediction in the Yoshida's method. The experimental results show that the proposed method achieves higher compression performance than the conventional ones

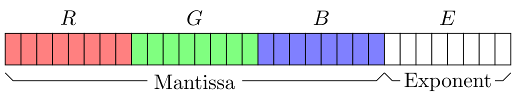

Fig. 1. Structure of the Radiance format.

without any loss in the entire process.

Notation: When A is an image (a component) and p (or q, whose total number is a quarter of one of p) is a pixel index of A, $A(p)$ (or $A(q)$) is the p-th (or q-th) pixel value of A. $\lceil\cdot\rceil$, $\lfloor\cdot\rfloor$, and $|\cdot|_2$ are the ceil function, floor function, and L_2-norm, respectively.

II. Preparation

A. Radiance Format

The Radiance format (extension ".hdr") [9] is a format of total 32-bit per pixel (bpp) representation, where 8 bits are assigned to each mantissa and exponential information (Fig. 1). By eliminating the sign bit and sharing the exponent part, each channel can be encoded efficiently. Its integer and inverse conversions (RGBE and rgb converters) are defined as follows:

$$E(p) = \lceil \log_2(\max(m(p)) + 128) \rceil, \tag{1}$$

$$M(p) = \left\lfloor \frac{256 \times m(p)}{2^{E(p)-128}} \right\rfloor, \tag{2}$$

$$m(p) = \frac{M(p) + 0.5}{256} 2^{E(p)-128}, \tag{3}$$

where E, $M \in \{R, G, B\}$, and $m \in \{r, g, b\}$ are the exponent value, mantissa (integer color) value, and the floating point color value, respectively. The conversed information E and M are further losslessly compressed from 32 bpp by the run length coding.

B. Encoder of JPEG 2000

JPEG 2000 is one of the most efficient image compression standards. At the encoder, RGB values are first decorrelated to YCbCr space by a reversible color transform (RCT) as follows:

$$Y = \left\lfloor \frac{R + 2G + B}{4} \right\rfloor, \tag{4}$$

$$C_\mathrm{b} = B - G, \tag{5}$$

$$C_\mathrm{r} = R - G, \tag{6}$$

where Y, C_b, and C_r denote the luminance and chrominance components. Then, each component is separately decorrelated to LL component (low frequency component in both horizontal and vertical directions), HL component (high frequency component in the vertical direction), LH component (high frequency component in the horizontal direction), and HH component (remaining high frequency components) by the 5/3-tap discrete wavelet transform (5/3-DWT). After that, they are encoded by EBCOT. Note that the preprocessing (RCT and 5/3-DWT) affects the compression performance.

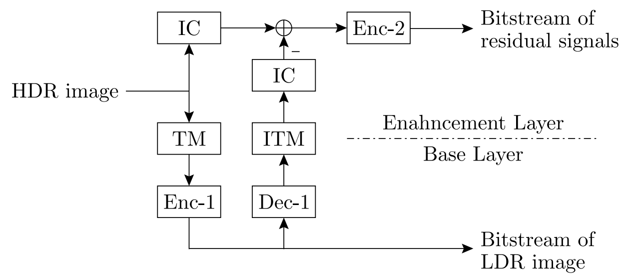

Fig. 2. Encoder of the basic encoder: IC, TM, ITM, Enc-1, Dec-1, and Enc-2 are integer converter, tone mapping, inverse tone mapping, lossy encoder for the tone-mapped LDR image, decoder related to Enc-1, and lossless encoder for the residual signals, respectively.

C. Basic Method of Two-Layer Coding of HDR Images

Two-layer coding of HDR images consists of the base and enhancement layers. In the base layer, an HDR image is converted to an LDR image by tone mapping and encoded by a general lossy encoder such as JPEG. In the enhancement layer, the predicted HDR image is generated by decoding and inverse tone mapping from the tone-mapped LDR image encoded in the base layer and the residual signals between the logarithmically-converted HDR and the predicted HDR images are encoded by a general lossless encoder such as JPEG 2000 lossless mode. Let "basic method" be one that introduced the logarithmic conversion [11] to integer converters in the JPEG XT-based framework as shown in the Fig. 2. Unfortunately, in such two-layer coding, reversibility cannot be guaranteed because the calculation loss occurs according to the lossy logarithmic conversion. Its entire compression performance is even worse than the conventional one-layer coding such as the Radiance format, OpenEXR format, and logarithmic method.

III. Two-Layer Lossless Coding of HDR Images Specialized for Radiance Format

In this section, a two-layer coding that greatly improves the compression performance while satisfying the reversibility for the Radiance format of HDR images is proposed. The proposed method has two strategies, RGBE converter for the integer converter and exponent-adjusted mantissa prediction. Fig. 3 shows the encoder of the proposed method.

A. RGBE Converter for Integer Converter

Unlike the basic method, the proposed method is based on the Radiance format and uses the RGBE converter in (1) and (2) as integer converters. The use of the RGBE converter contributes to reduce the amount of information from 16 bits to 8 bits for each color. In addition, E tends to approach a certain value, indicating relatively small amount of information. The rgb converter in (3) is also used as the inverse conversion.

B. Exponent-adjusted Mantissa Prediction

1) Basic Strategy: The compression performance is further improved by incorporating a mantissa prediction based on the decomposition with the RGBE converter as shown in Fig. 3. As shown in (2), E has a very large effect on M. Therefore,

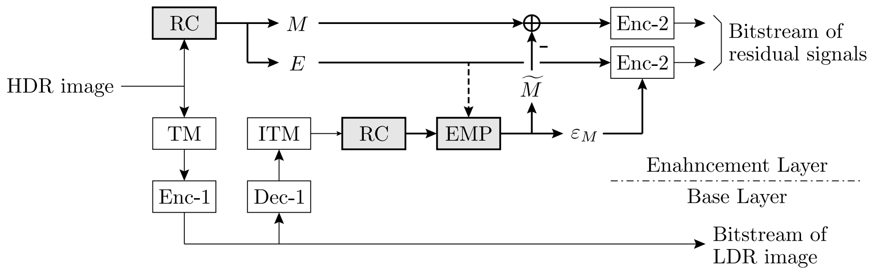

Fig. 3. Encoder of the proposed method: RC, EMP, TM, ITM, Enc-1, Dec-1, and Enc-2 are RGBE converter, exponent-adjusted mantissa prediction, tone mapping, inverse tone mapping, encoder for the tone-mapped LDR image, decoder for LDR image, and lossless encoder for the residual signals, respectively.

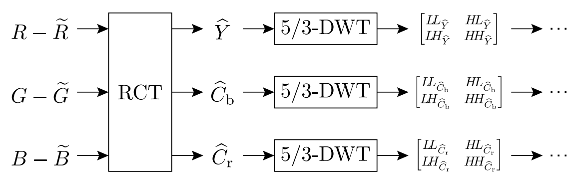

Fig. 4. Preprocessing in encoder of JPEG 2000 lossless mode in the enhancement layer.

for better compression performance, we consider predicting M in the enhancement layer by adjusting E. First, the predicted mantissa information $\widetilde{M} \in \{\widetilde{R}, \widetilde{G}, \widetilde{B}\}$ in the enhancement layer is expressed as follows:

$$\widetilde{M}(p) = \left\lfloor \frac{256 \times \widetilde{m}(p)}{2^{\widetilde{E}_M(p)-128}} \right\rfloor, \tag{7}$$

where $\widetilde{m} \in \{r', g', b'\}$ and $\widetilde{E}_M \in \{\widetilde{E}_R, \widetilde{E}_G, \widetilde{E}_B\}$ are the mantissa and exponential information obtained with the inverse tone mapping of LDR images in the base layer. By using the unknown value $\varepsilon_M \in \{\varepsilon_R, \varepsilon_G, \varepsilon_B\}$ that produces the well-predicted mantissa information, (7) is rewritten by

$$\widetilde{M}(p) = \left\lfloor \frac{256 \times \widetilde{m}(p)}{2^{E(p)+\varepsilon_M-128}} \right\rfloor. \tag{8}$$

ε_M, which has only 32 bits, is transmitted as side information.

2) Adjustment Considering Encoder: We apply JPEG 2000 lossless mode to a coder in the enhancement layer in accordance with the previous studies [13–15].[1] Since the preprocessing greatly affects the compression performance as described in Section II-B, we adjust ε_M by considering those transforms. Fig. 4 shows the preprocessing in the encoder of JPEG 2000 lossless mode in the enhancement layer.

In the enhancement layer, the luminance and chrominance components $\widehat{Y}$, $\widehat{C}_\mathrm{b}$, and $\widehat{C}_\mathrm{r}$ of the residual signals are obtained as follows:

$$\widehat{Y} = \left\lfloor \frac{(R-\widetilde{R}) + 2(G-\widetilde{G}) + (B-\widetilde{B})}{4} \right\rfloor, \tag{9}$$

$$\widehat{C}_\mathrm{b} = (B-\widetilde{B}) - (G-\widetilde{G}), \tag{10}$$

$$\widehat{C}_\mathrm{r} = (R-\widetilde{R}) - (G-\widetilde{G}). \tag{11}$$

[1]The other coders can also be applied.

Here, we sparsify the chrominance components by adjusting the exponential information in (8) via reducing the L_2 norm of $\widehat{C}_\mathrm{b}$ and $\widehat{C}_\mathrm{r}$ represented as follows:

$$|\widehat{C}_\mathrm{b}|_2 = \sqrt{\sum_p |(B(p)-\widetilde{B}(p)) - (G(p)-\widetilde{G}(p))|^2}, \tag{12}$$

$$|\widehat{C}_\mathrm{r}|_2 = \sqrt{\sum_p |(R(p)-\widetilde{R}(p)) - (G(p)-\widetilde{G}(p))|^2}. \tag{13}$$

Similar to the above chrominance components, the L_2 norm of the HH components are represented as follows:

$$|HH_{\widehat{Y}}|_2 = \sqrt{\sum_q |HH_{\widehat{Y}}(q)|^2}, \tag{14}$$

$$|HH_{\widehat{C}_\mathrm{b}}|_2 = \sqrt{\sum_q |HH_{\widehat{C}_\mathrm{b}}(q)|^2}, \tag{15}$$

$$|HH_{\widehat{C}_\mathrm{r}}|_2 = \sqrt{\sum_q |HH_{\widehat{C}_\mathrm{r}}(q)|^2}, \tag{16}$$

where $HH_{\widehat{Y}}$, $HH_{\widehat{C}_\mathrm{b}}$, and $HH_{\widehat{C}_\mathrm{r}}$ denote the HH components of $\widehat{Y}$, $\widehat{C}_\mathrm{b}$, and $\widehat{C}_\mathrm{r}$, respectively. Consequently, ε_M is calculated as follows:

$$\begin{aligned}\arg\min_{\varepsilon_M}(&w(1)|\widehat{C}_\mathrm{r}|_2 + w(2)|\widehat{C}_\mathrm{b}|_2 \\ &+ w(3)|HH_{\widehat{Y}}|_2 + w(4)|HH_{\widehat{C}_\mathrm{r}}|_2 + w(5)|HH_{\widehat{C}_\mathrm{b}}|_2),\end{aligned} \tag{17}$$

where $w(k)$s are weight factors.

The bitstreams of the tone-mapped LDR image, ε_M, E, and residual signals are stored and transmitted to the decoder. By applying the rgb conversion in (3) to the RGBE components obtained with the inverse processing of the transmitted bitstreams, the original HDR image is completely restored.

IV. Experimental Results

For the test images, we used 36 full-color HDR images (Radiance format) cropped from the database in [17] to a size of 1024×1536. In order to measure the compression performance, the bitrate was used as objective index. It indicates that the smaller bitrate is the better compression performance. For each method, we applied the Reinhard's method [3] for the tone mapping, the Huo's method [18] for the inverse tone mapping, JPEG (quality factor $Q = 85$) for

TABLE I
LOSSLESS CODING RESULTS (BITRATE [BPP]).

Image	Radiance	Basic	[13]	[15]	Prop.1	Prop.2
img1	27.13	29.90	27.09	19.95	20.25	**17.18**
img2	26.34	26.73	23.79	15.55	17.74	**14.65**
img3	27.04	34.16	32.01	18.05	18.83	**16.68**
img4	23.89	28.23	25.38	16.91	13.64	**12.81**
img5	26.67	30.99	28.23	20.36	20.74	**18.35**
img6	23.80	25.47	22.70	15.30	15.10	**12.87**
img7	24.00	27.18	24.24	12.27	13.08	**11.34**
img8	24.21	27.79	24.89	11.26	11.89	**10.33**
img9	26.75	27.52	24.58	18.44	17.69	**15.44**
img10	23.37	28.42	25.56	13.64	13.71	**11.69**
img11	26.55	28.14	25.23	18.49	18.75	**16.17**
img12	24.95	30.43	27.54	12.63	13.97	**11.83**
img13	25.05	30.43	27.68	13.10	14.31	**12.25**
img14	24.59	30.50	27.34	13.35	14.92	**12.45**
img15	25.57	30.84	28.05	14.00	16.24	**13.70**
img16	26.18	32.74	29.98	16.56	18.34	**15.50**
img17	25.18	31.63	28.82	13.90	16.17	**13.58**
img18	27.32	31.33	28.59	18.81	20.84	**17.81**
img19	24.96	31.79	28.93	14.21	15.34	**13.38**
img20	25.15	24.34	21.44	14.15	14.31	**12.33**
img21	24.78	26.48	23.71	14.16	15.32	**13.05**
img22	24.98	29.71	27.41	14.65	15.74	**12.95**
img23	24.63	28.23	25.37	11.76	13.11	**11.04**
img24	25.17	29.68	26.84	13.48	14.52	**12.11**
img25	25.12	30.80	27.78	13.54	15.30	**12.44**
img26	25.79	29.01	26.07	19.55	18.66	**15.81**
img27	25.57	27.30	24.46	17.37	16.92	**14.22**
img28	25.89	28.92	26.14	16.36	17.56	**15.02**
img29	25.20	30.25	27.50	15.65	16.85	**14.28**
img30	24.32	29.77	26.92	**11.88**	13.77	12.81
img31	25.70	26.09	23.06	15.38	16.36	**14.30**
img32	25.79	26.82	23.96	17.05	17.13	**14.50**
img33	24.50	30.58	27.66	12.78	14.37	**12.42**
img34	24.58	29.50	26.74	12.76	15.19	**12.40**
img35	25.20	29.44	26.50	11.96	13.96	**11.65**
img36	24.85	24.87	22.28	14.15	15.05	**13.03**
Average	25.30	29.06	26.23	15.09	15.99	**13.68**

the base layer coder, and JPEG 2000 lossless mode for the enhancement layer coder. Let Prop. 1 and 2 be the proposed methods with only the RGBE converter for integer converter and with both of the RGBE converter for integer converter and the exponent-adjusted mantissa prediction, respectively. The proposed methods were compared with the basic method, Watanabe's method [13], and Yoshida's method [15] in lossless coding.

The experimental results are shown in Table I. Note that the basic method and Watanabe's method are not exactly lossless due to calculation loss in the logarithmic method. The proposed methods not only guarantee reversibility but also outperformed the conventional methods for almost all images. Especially, the Prop. 2 improved the average bitrate by 1.41 bpp compared with Yoshida's method, which is the best existing method. Since the Prop. 2 always showed better performance compared with the Prop. 1, it is clear that the exponent-adjusted mantissa prediction had a significant effect on the compression performance improvement.

V. CONCLUSION

This study proposed a two-layer coding of HDR images specialized for the Radiance format with the RGBE converter for the integer converter and the exponent-adjusted mantissa prediction. The proposed method almost always provided the better compression performance than conventional methods while guaranteeing reversibility and improved the average bitrate by 1.41 bpp compared with the best existing method.

REFERENCES

[1] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*, Morgan Kaufmann, 2010.

[2] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging*, AK Peters/CRC Press, 2017.

[3] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, July 2002.

[4] E. Reinhard, "Parameter estimation for photographic tone reproduction," *J. Graph. Tools*, vol. 7, no. 1, pp. 45–51, Nov. 2002.

[5] B. Gu, W. Li, M. Zhu, and M. Wang, "Local edge-preserving multiscale decomposition for high dynamic range image tone mapping," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 70–79, Aug. 2012.

[6] Z. Li and J. Zheng, "Visual-salience-based tone mapping for high dynamic range images," *IEEE Trans. Ind. Electron.*, vol. 61, no. 12, pp. 7076–7082, Mar. 2014.

[7] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image compression by optimizing tone mapped image quality index," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3086–3097, May 2015.

[8] S. Paris, S.W. Hasinoff, and J. Kautz, "Local Laplacian filters: Edge-aware image processing with a Laplacian pyramid," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 68, Mar. 2015.

[9] G. Ward, *Graphics Gems II*, chapter Real Pixels, pp. 80–85, 1991.

[10] R. Bogart, F. Kainz, and D. Hess, "OpenEXR image file format," in *Proc. ACM SIGGRAPH*, San Diego, CA, Jan. 2003.

[11] R. Xu, S.N. Pattanaik, and C.E. Hughes, "High-dynamic-range still-image encoding in JPEG 2000," *IEEE Comput. Graph. Appl.*, vol. 25, no. 6, pp. 57–64, Nov. 2005.

[12] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. Ward, and W. Heidrich, "Optimizing a tone curve for backward-compatible high dynamic range image and video compression," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1558–1571, Dec. 2010.

[13] O. Watanabe, H. Kobayashi, and H. Kiya, "Lossless two-layer coding using histogram packing technique for HDR images," in *Proc. ISCAS*, Florence, Italy, May 2018, pp. 1–5.

[14] M. Iwahashi, T. Yoshida, N.B. Mokhtar, and H. Kiya, "Bit-depth scalable lossless coding for high dynamic range images," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 22, pp. 1–15, Mar. 2015.

[15] T. Yoshida, M. Iwahashi, and H. Kiya, "Two-layer lossless coding for high dynamic range images based on range compression and adaptive inverse tone-mapping," *IEICE Trans. Fundamentals*, vol. 101, no. 1, pp. 259–266, Jan. 2018.

[16] A.J. Pinho, "An online preprocessing technique for improving the lossless compression of images with sparse histograms," *IEEE Signal Process. Lett.*, vol. 9, no. 1, pp. 5–7, Feb. 2002.

[17] H. Nemoto, P. Korshunov, P. Hanhart, and T. Ebrahimi, "Visual attention in LDR and HDR images," in *Proc. VPQM*, Chandler, AZ, Feb. 2015, pp. 1–6.

[18] Y. Huo, F. Yang, L. Dong, and V. Brost, "Physiological inverse tone mapping based on retina response," *The Visual Computer*, vol. 30, no. 5, pp. 507–517, Sep. 2014.