Received November 19, 2017, accepted January 4, 2018, date of publication January 15, 2018, date of current version February 28, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2793228

Improvement of JPEG XT Floating-Point HDR Image Coding Using Region Adaptive Prediction

OH-JIN KWON[1], SEUNGCHEOL CHOI[1], AND DONGKYOO SHIN[2]
[1]Department of Electrical Engineering, Sejong University, Seoul 05006, South Korea
[2]Department of Computer Engineering, Sejong University, Seoul 05006, South Korea

Corresponding author: Seungcheol Choi (choisc@sju.ac.kr)

This work was supported by the Standard Technology Development and Spread Program of KATS/KEIT [10069128, Developing Standard Platform and Library for JPEG Privacy-Security].

ABSTRACT Displays capable of representing high-dynamic range (HDR) images were recently released in the digital consumer electronics market, and consumers have become increasingly interested in HDR images. However, the majority of today's digital imaging devices are still low dynamic range (LDR) using an 8-bit representation for each RGB color components; hence, a backward-compatible HDR image format with an existing LDR-based imaging environment is required. Accordingly, JPEG XT, a new HDR image coding standard providing JPEG backward compatibility was established by the JPEG committee based on the market needs. JPEG XT consists of a base and a residual layer. Both layers are independently encoded by a legacy JPEG encoder. Most importantly, the base layer contains a tone-mapped LDR version of the HDR image for compliance with a legacy JPEG decoder. To date, three profiles have been defined for JPEG XT to encode HDR images, which are represented using floating-point values. This paper proposes a new image-coding scheme that uses a region adaptive prediction method with modified current specifications on JPEG XT Part 7. The proposed method adaptively predicts the HDR pixel value using different prediction information for each block based on the ratio between the LDR and HDR blocks, which is invariant to the tone-mapping operator, while the existing JPEG XT profiles globally predict the HDR values using an inverse transformation of the LDR values. Experimental results using various sample images show the superiority of the proposed method in terms of objective and visual evaluations.

INDEX TERMS JPEG XT, HDR from LDR predictor (HLP), high-dynamic range image, image coding.

I. INTRODUCTION

High-dynamic range (HDR) imaging has recently been widely used in various applications. Images taken by typical digital cameras with limited dynamic range capabilities usually suffer from lack of details in the under- and over-exposed areas. The most commonly used method in capturing HDR images, which solves this problem, is to take a sequence of low-dynamic range (LDR) images of the same scene with varying exposures and combine them into a single image [1]. Several formats have been established in the HDR imaging industry using a floating point for each color channel, as listed in [2, Table 1]. These image file formats are lossless for the subsequent applications: storage, transmission, or display. However, lossy formats have been also required in the industry, where the storage and transmission costs are critically considered.

As for lossy image coding, the JPEG standard (ISO/IEC 10918-1) has been widely used and has established a matured tool chain in the digital imaging industry, but it offers no support for HDR images.

TABLE 1. HDR image formats.

Format	Pixel encoding	Compression type	Bits/pixel
HDR	RGBE	Run-length	32
	XYZE	Run-length	32
EXR	Half RGB	Wavelet, ZIP	48
TIFF	IEEE RGB	None	96
	LogLuv24	None	24
	LogLuv32	Run-length	32

Although the newer JPEG image-coding standards, such as JPEG 2000 and JPEG XR, support HDR images, their usage has been limited to a few niche markets, such as digital cinema and medical imaging. The JPEG working group (WG)

TABLE 2. Description of the JPET XT standard [5].

	Title	Description
Part 1	Core coding system	Definition of the core coding technology, which is the legacy JPEG specifications
Part 2	Coding of high-dynamic range images	Definition of the text-based legacy syntax for HDR photography
Part 3	Box file format	Definition of an extensible box-based format extending legacy JPEG format
Part 4	Conformance testing	Definition of the methodology to verity that the standard is meeting the normative requirements
Part 5	Reference software	Reference software of JPEG XT subparts 6–9
Part 6	Intermediate-dynamic range (IDR) integer coding	Definition of a coding for integer samples between 9 and 16 bits per sample
Part 7	HDR floating-point coding	Definition of a coding for HDR image (i.e., image consisting of floating-point samples)
Part 8	Lossless and near-lossless coding	Definition of a lossless and near-lossless coding for IDR and HDR images
Part 9	Alpha channel coding	Definition of an extension to support opacity information on the other subparts of this specification

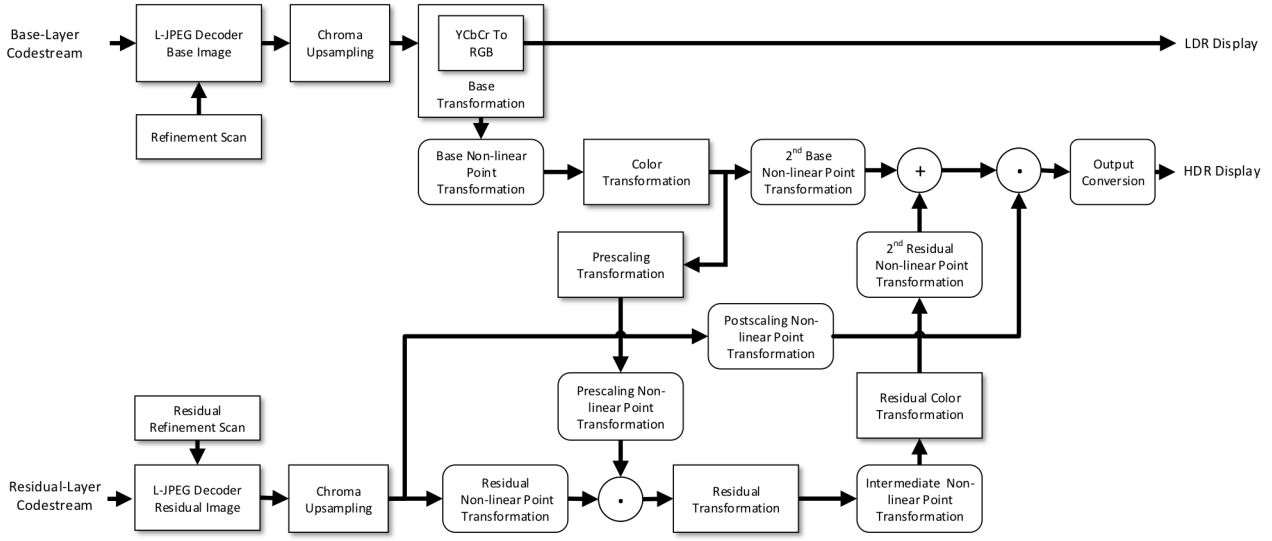


FIGURE 1. High-level overview of the decoding process of JPEG XT Part 7 [4].

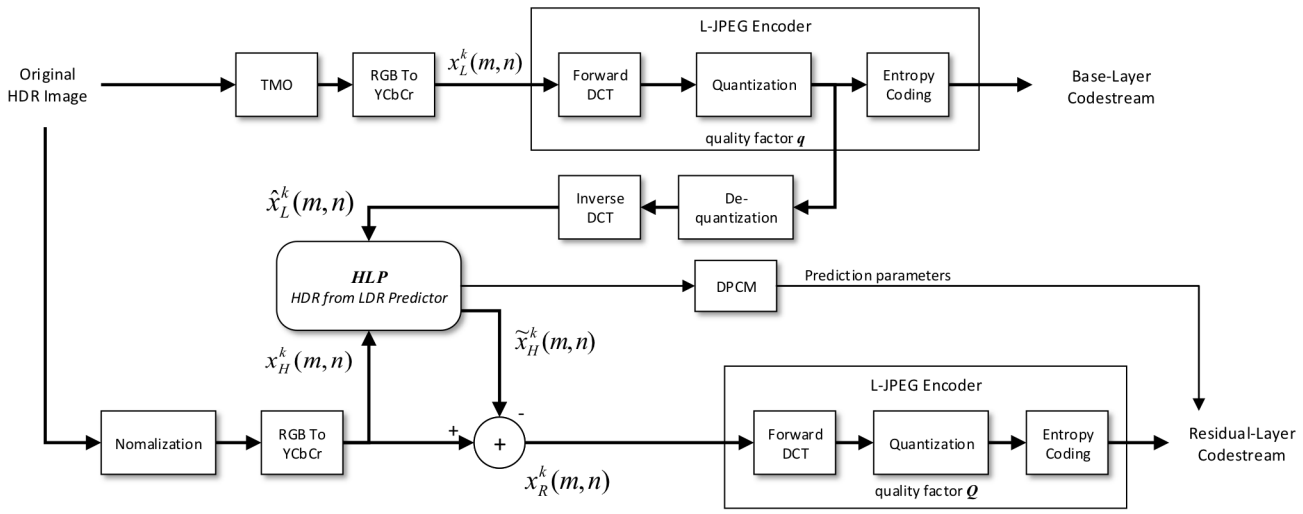
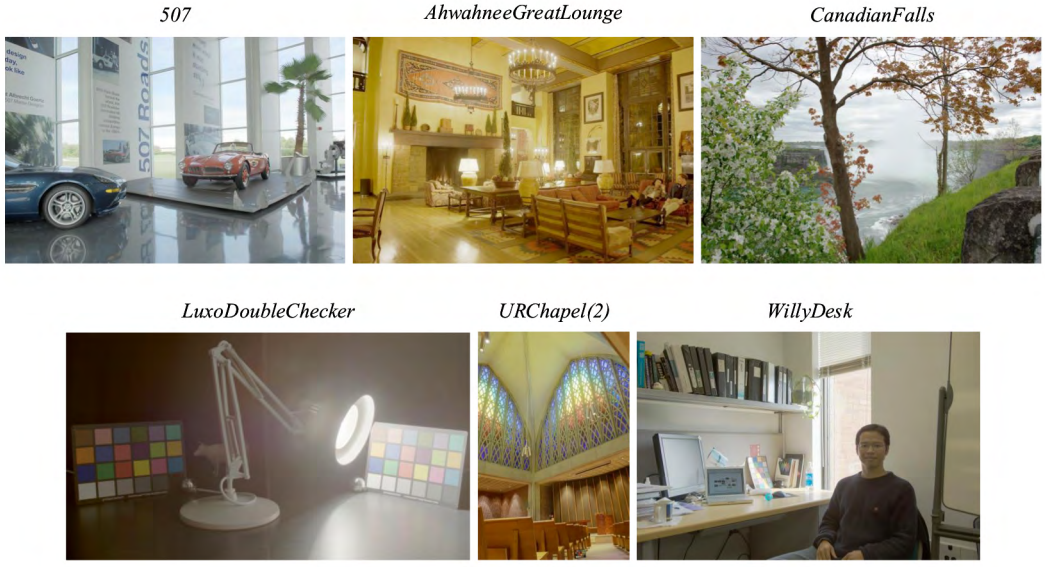


FIGURE 2. High-level overview of the encoding process of the proposed scheme.

recognized that the main reason for this is because they are not backward-compatible with the legacy JPEG (L-JPEG) standard; hence, they recently established a new HDR image-coding standard, called JPEG XT (ISO/IEC 18477) that is backward-compatible with the L-JPEG-based imaging environment [3].

(a)



(b)

FIGURE 3. Sample images [8] used to verify our assumption: LDR images (a) uniformly tone-mapped and (b) tone-mapped by [9].

Table 2 lists the nine parts composing the JPEG XT specification. Part 7 [4] specifies the floating-point-valued HDR image coding, which is the main concern of this study. To date, three profiles have been defined for lossy HDR image coding in JPEG XT Part 7. They encode an HDR image into a codestream consisting of two layers: base and residual layers, each of which can be decoded by an L-JPEG decoder. The base layer contains a tone-mapped LDR version of the input HDR image, whereas the residual layer provides additional information to reconstruct the HDR image from the LDR version. The choice of the tone-mapping operator (TMO) is open to the user. The tone-mapped LDR image encoded in the base layer is displayed when a JPEG XT file is decoded by the existing L-JPEG format-compliant device. Meanwhile, a JPEG XT-compliant decoder can reconstruct the HDR image using the information from the residual layer.

This study shows that the performance of the existing profiles of JPEG XT Part 7 is highly dependent on the choice of TMOs. Furthermore, we propose a scheme that improves the robustness of JPEG XT Part 7 to various TMOs with a modified current specification. A novel method that block-wisely predicts the input HDR image from its tone-mapped LDR version is proposed herein based on a relationship between both LDR and HDR images. We design the predictor to be locally adaptive to the differently tone-mapped regions of the LDR image such that it may compensate for the effect of various TMOs for generating the residual layer.

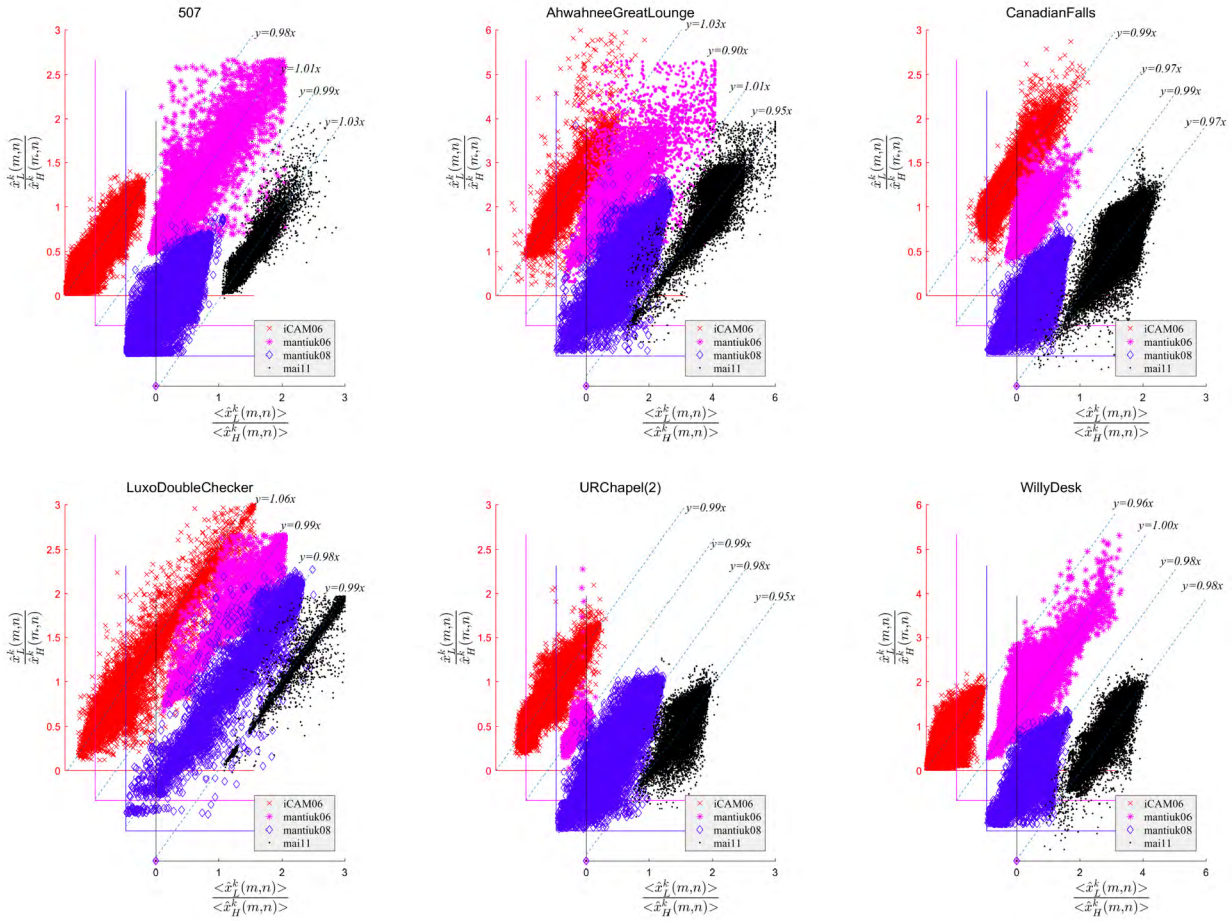


FIGURE 4. Plots representing the correlation between $\frac{\langle \hat{x}_L^k(m,n) \rangle}{\langle x_H^k(m,n) \rangle}$ and $\frac{\hat{x}_L^k(m,n)}{x_H^k(m,n)}$ on a luminance component. Each plot explicates the relationship using a linear equation.

The remainder of this paper is organized as follows: Section II provides the overview of JPEG XT Part 7; Section III explains the proposed HDR image-coding scheme; Section IV presents the experimental results showing the superiority of the proposed scheme; and Section V concludes this paper.

II. OVERVIEW OF THE JPEG XT PART 7 PROFILES

Figure 1 shows the high-level overview of the decoding process of JPEG XT Part 7 [4]. The JPEG XT standard does not define a normative encoding process, but only defines the decoding process. Therefore, any encoding process can be used when its encoded file is decoded by the normative decoding process. As shown in Fig. 1, the decoding process of JPEG XT is defined as a flow of functional blocks, and each block performs various types of processing: refinement scan, chroma upsampling, color transformation, non-linear-point transformation, scaling transformation, etc. [6], [7]. All the functional blocks are not used in the real implementation of JPEG XT Part 7. JPEG XT defines three different types of implementation, namely, Profiles A, B, and C, for lossy HDR image coding. Although one more profile, called Profile D, is defined in JPEG XT Part 7 [4], this profile is not for coding HDR images and does not assume the residual layer. The profile is provided only to simply extend the dynamic range of the traditional LDR image encoded by the L-JPEG. The choice of the profile is open to users. Each profile is differentiated by how it specifies a subset of the functional blocks for reconstructing the output HDR image.

The first step of the decoding process in the three profiles of JPEG XT is to apply the L-JPEG decoder for both base- and residual-layer codestreams. The base- and residual-layer images are then represented by the spatial domain pixels. Profile A basically reconstructs the HDR image by pixel-wise multiplying a scale factor signaled in the residual layer to the base image, while Profile B reconstructs the HDR image as the pixel-wise quotient of the base image and a divisor image delivered by the residual layer. Profile C reconstructs the HDR image by employing a pixel-wise sum to merge base and residual images. Compared to Profiles A and B, Profile C can be characterized as the only one profile which

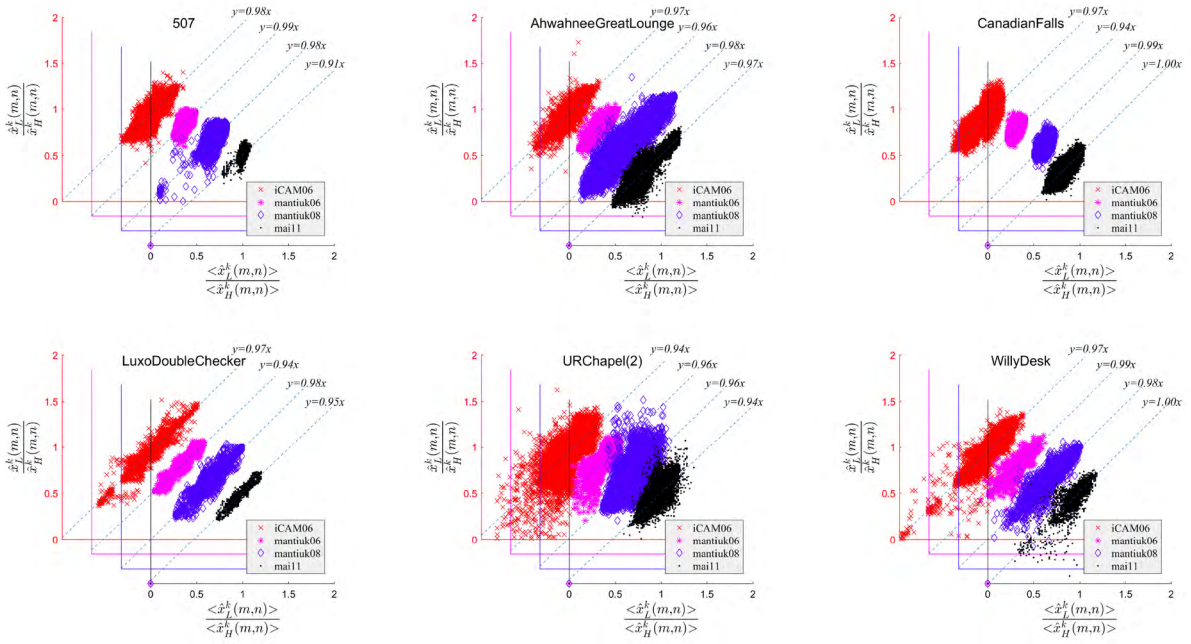


FIGURE 5. Plots representing the correlation between $\frac{<\hat{x}_L^k(m,n)>}{<x_H^k(m,n)>}$ and $\frac{\hat{x}_L^k(m,n)}{x_H^k(m,n)}$ on chrominance components: Cb and Cr. Each plot explicates the relationship using a linear equation.

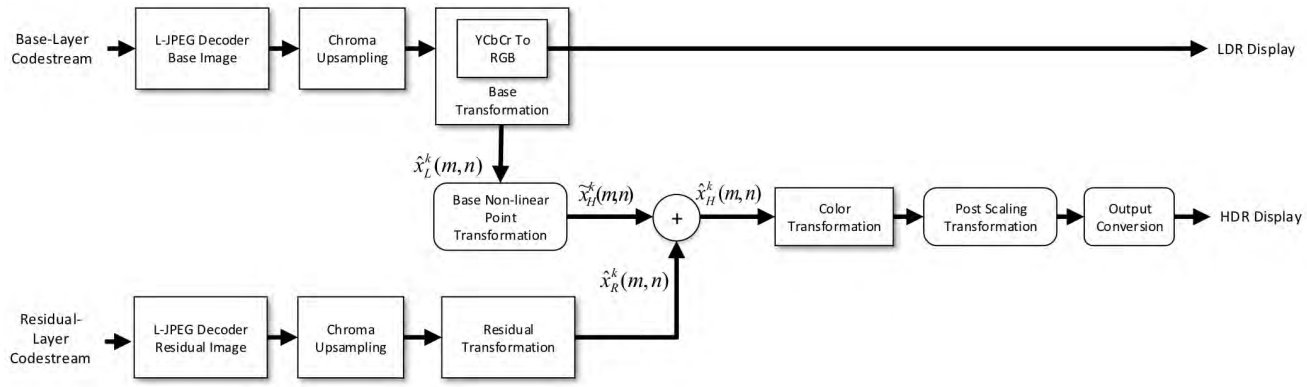


FIGURE 6. High-level overview of the decoding process of the proposed scheme.

uses the refinement scan functional block for the extension of the coding precision in the DCT domain. In JPEG XT, the base layer contains the LDR image, which results from L-JPEG encoding the tone-mapped LDR version of the original HDR image. The choice of TMO in the encoding process is assumed to be open to users. Therefore, an operation compensating or estimating the effect of TMO used in the encoding process is demanded before merging base and residual images into the output HDR image in the decoding process. JPEG XT Part 7 basically implements this operation using the non-linear transformation functional block in Fig. 1. For this operation, both Profiles A and B use pixel-wise inverse gamma correction procedures, while Profile C uses a pixel-wise global inverse tone-mapping procedure in addition to gamma correction [3], [6], [7].

Note that all the functional blocks defined in the current JPEG XT are limited to performing pixel-wise operations. However, the behaviors of state-of-the-art TMOs include global, local, segmentation-based, and frequency/gradient-based operations [1], [2], which cannot be compensated by simple pixel-wise operations. Therefore, we claim that all the existing profiles of JPEG XT inevitably show highly dependent performances on the choice of TMOs. We will show our experimental results on this problem in Section IV.

III. PROPOSED SCHEME

A. HDR FROM THE LDR PREDICTOR

In the encoding process of JPEG XT, an input HDR image is tone-mapped by a TMO. The resulting LDR version

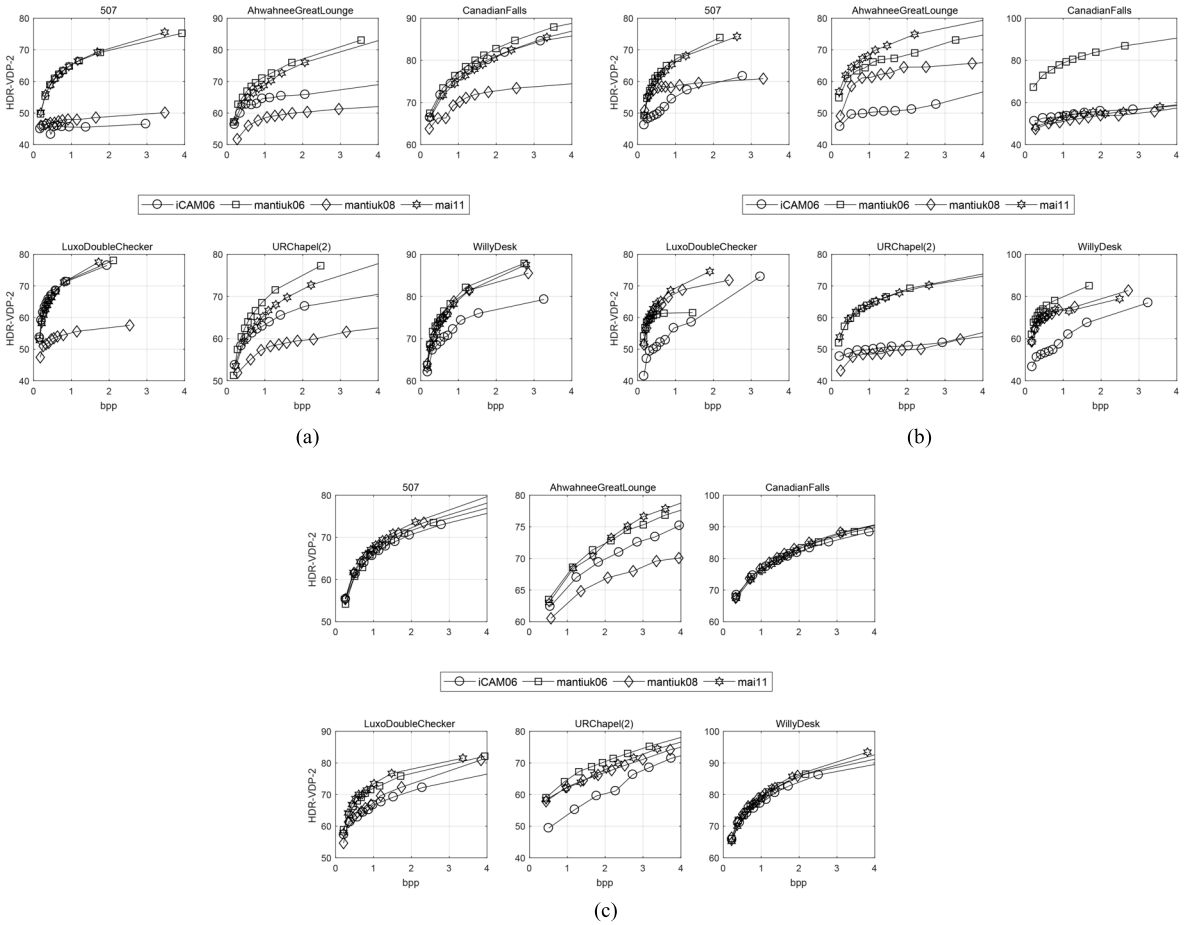


FIGURE 7. HDR-VDP-2 performance of Profiles A, B, and C on the four TMOs. Bpp is an acronym for bits per pixel. (a) Profile A. (b) Profile B. (c) Profile C.

(base image) of the input HDR image is compressed by the L-JPEG encoder, and the base-layer codestream is constructed to provide L-JPEG backward compatibility. As described in Section II, each profile generates the residual image in its own way by comparing the input HDR image and the image, which is the L-JPEG-decoded image from the base-layer codestream. The residual layer codestream is obtained by L-JPEG encoding the residual image [4]. The L-JPEG assumes that the image is represented in the YCbCr color space and block-wisely performs the encoding (8×8 size). We denote herein a pixel at (m, n) in the kth block of the input HDR image and the base image by $x_H^k(m, n)$ and $x_L^k(m, n)$, respectively. Figure 2 denotes the encoded version of the base image by $\hat{x}_L^k(m, n)$.

The key feature of the proposed scheme is to predict $x_H^k(m, n)$ based on $\hat{x}_L^k(m, n)$. We call this the HDR from the LDR predictor (HLP) (Fig. 2). We need to find a relationship between $x_H^k(m, n)$ and $\hat{x}_L^k(m, n)$, which is invariant to different TMOs, to design an encoding scheme that is robust to different TMOs. We designed the HLP based on the assumption that the behavior of a TMO on $x_H^k(m, n)$ is almost similar to that on $< x_H^k(m, n) >$, where $< x_H^k(m, n) >$ is the mean value of the kth block, including a pixel at (m, n). This assumption can be represented by the following equation:

$$\frac{\hat{x}_L^k(m, n)}{x_H^k(m, n)} \approx \frac{< \hat{x}_L^k(m, n) >}{< x_H^k(m, n) >}, \quad (m, n) \in \{1 \le m \le 8, 1 \le n \le 8\}. \tag{1}$$

We performed a test to verify our assumption for each of the Y, Cb, and Cr components of a pixel and describe our observation results using six sample HDR images. Sample HDR images were tone-mapped by two TMOs (i.e., uniform TMO and TMO proposed by Mai *et al.* [9]) for display purpose. Figures 3(a) and (b) depict the results. Note that our observation results for the other images provided the same conclusion in terms of designing the HLP. We also describe our observation results using various TMOs. Four TMOs were selected based on a thorough review of previous researches [10]–[15]. These TMOs are frequently used in HDR imaging research. We denoted these herein as iCAM06 [16], mantiuk06 [17], mantiuk08 [18], and mai11 [9]. The quality

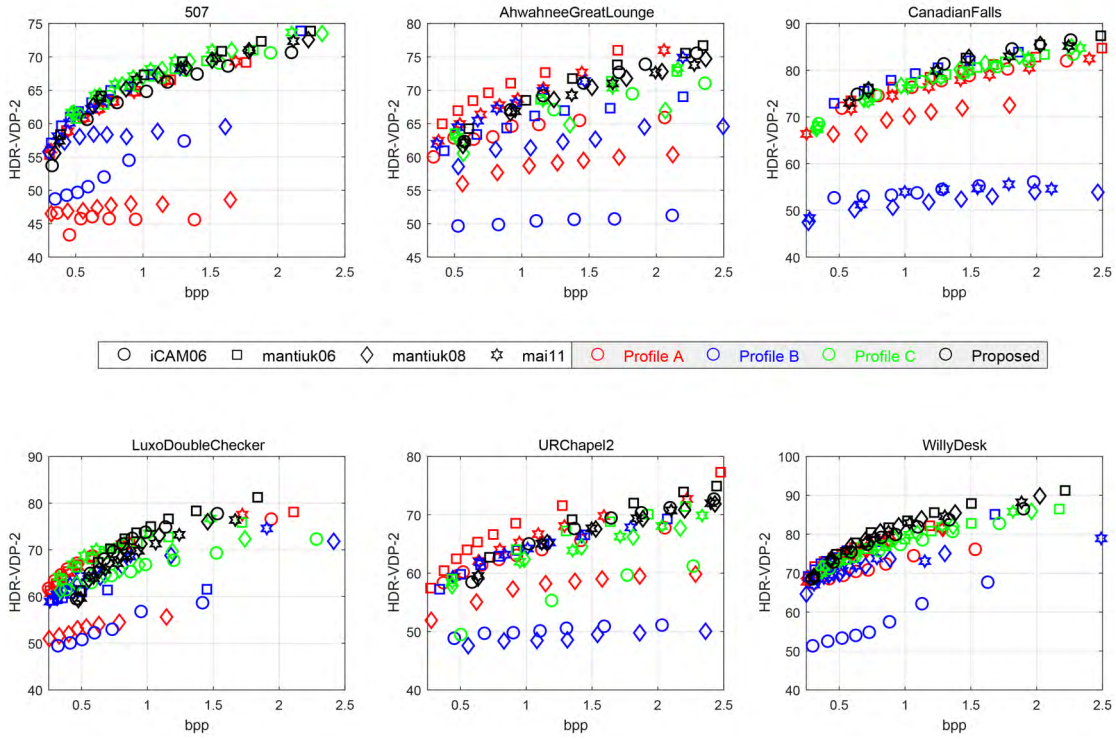


FIGURE 8. HDR-VDP-2 performance of the compared methods.

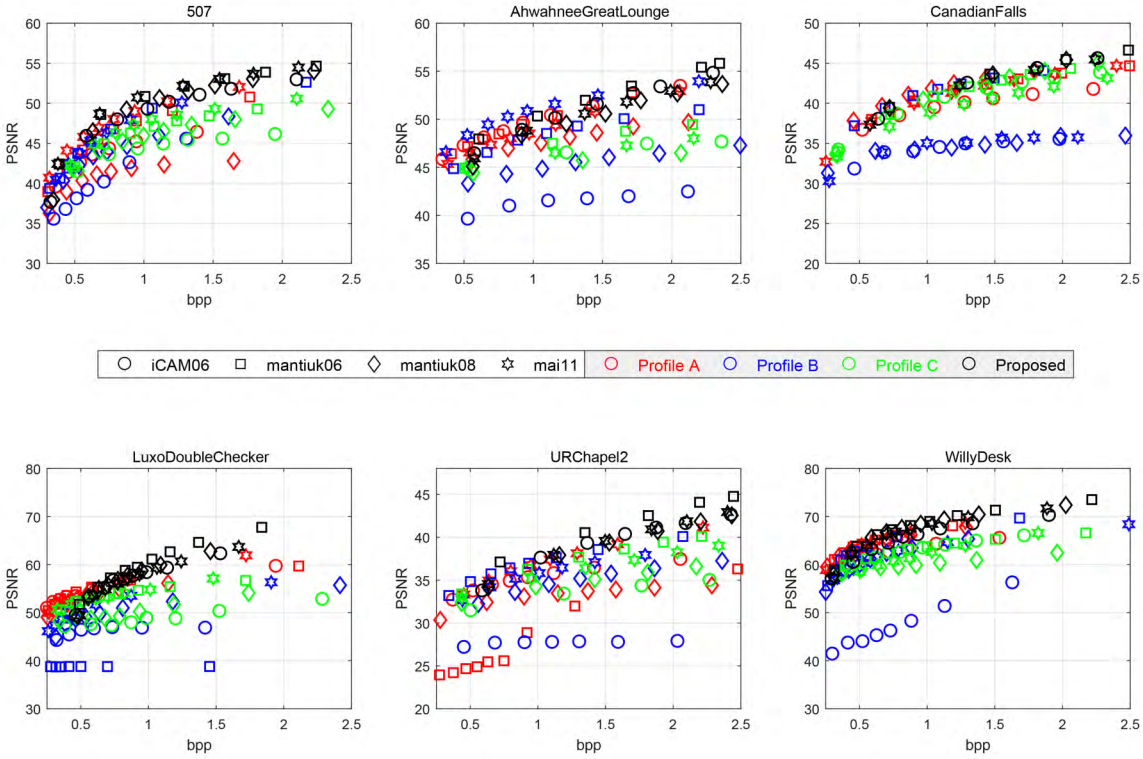


FIGURE 9. PSNR performance of the compared methods.

factor q (Fig. 2) for the L-JPEG encoding of the base image was set to 95, such that the observation provides better results. We also observed the distribution given by different quality factors and conclude that the effect of a different quality factor was negligible in the design of the HLP.

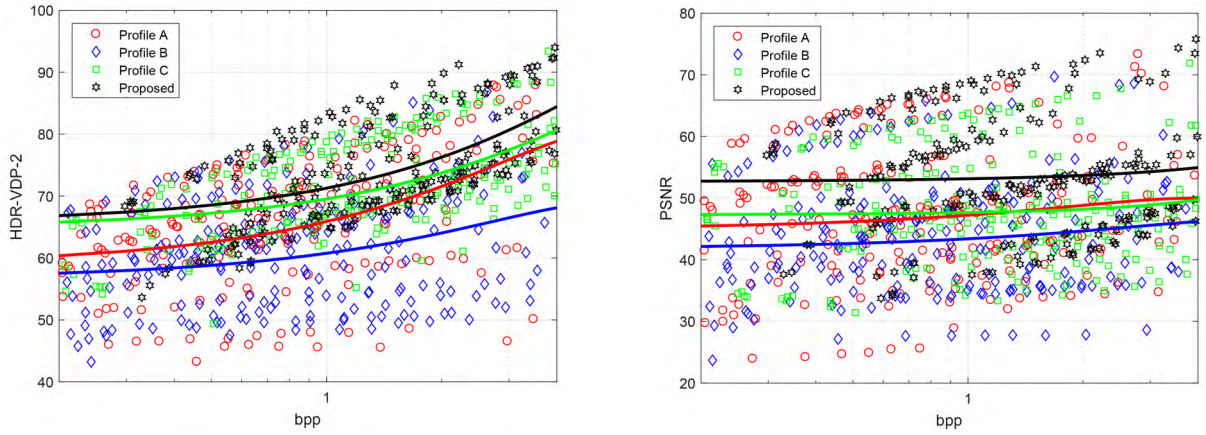
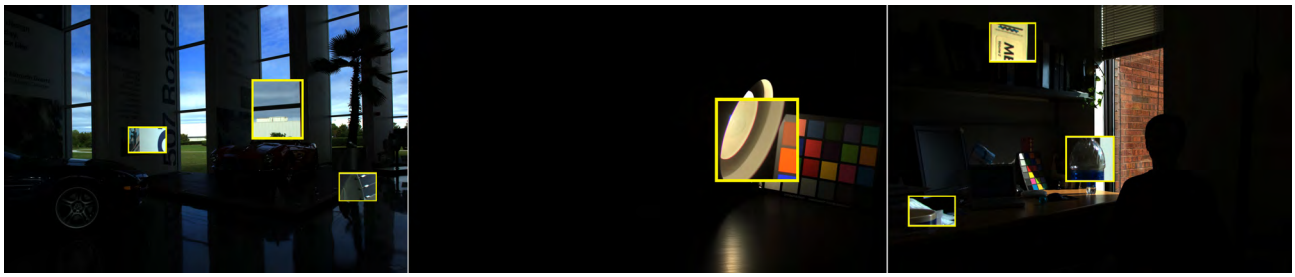


FIGURE 10. Comparison of the HDR-VDP-2 and PSNR performances of the compared methods with best-fitting curves.



(a) (b) (c)

FIGURE 11. Selected parts of the sample images for the visual tests: (a) 507, (b) LuxoDoubleChecker, and (c) WillyDesk. The selected parts are uniformly tone-mapped and enlarged for display.

Figure 4 shows the relationship between two variables: $\frac{<\hat{x}_L^k(m,n)>}{<x_H^k(m,n)>}$ and $\frac{\hat{x}_L^k(m,n)}{x_H^k(m,n)}$ on a luminance component. Each plot shows the distribution for the four TMOs. The distribution explicated that the variables were highly correlated and can be approximated by a linear function that minimizes the mean squared error between variables. As shown in Fig. 4, the slopes of all cases of four TMOs combined with six images were close to 1, such that our assumption represented in (1) was clarified for a luminance component even if a different TMO induced a different shape of the distribution.

Figure 5 illustrates the distribution on chrominance components (Cb and Cr components together). Although the shape of the distribution differs when different TMOs are used, the high correlation between the variables is similar to that of the luminance component. From this observation, we may conclude that our assumption of (1) is also useful for chrominance components. Our HLP for each of the Y, Cb, and Cr color components is given by the following equation:

$$\tilde{x}_H^k(m,n) = \alpha(k)\hat{x}_L^k(m,n), \tag{2}$$

with

$$\alpha(k) = \frac{< x_H^k(m,n) >}{< \hat{x}_L^k(m,n) >}, \quad (m,n) \in \{1 \le m \le 8, 1 \le n \le 8\}. \tag{3}$$

The originality of the proposed scheme is that the HDR values are block-wisely predicted by the ratio of the mean value of the HDR block to the corresponding LDR block. Meanwhile, the JPEG XT profiles globally predict the HDR values using a pixel-wise inverse transformation on the LDR image.

B. PROPOSED ENCODING PROCESS

Figure 2 depicts the procedure of the proposed encoding scheme. An original HDR image is tone-mapped to integer values for the base layer. The tone-mapped LDR version of the HDR image is L-JPEG encoded as performed in the existing JPEG XT profiles.

For the residual layer, each RGB color component of the original HDR image is first normalized into a floating-point value within the 0.0 to 255.0 range. The range of the original color values for each RGB color component is embedded as an overhead information into the residual layer codestream. The *HLP* function block performs the proposed prediction denoted by (2) in the YCbCr color space. The prediction coefficient $\alpha(k)$ is encoded by the differential pulse-code modulation method, as similarly performed for the DC coefficients of DCT in the L-JPEG encoder [19], which quantizes the difference between $\alpha(k)$ and $\alpha(k-1)$ using 8 bits, where the resulting quantized difference is also lossless encoded by an arithmetic coder [19]. Finally, the error image denoted by

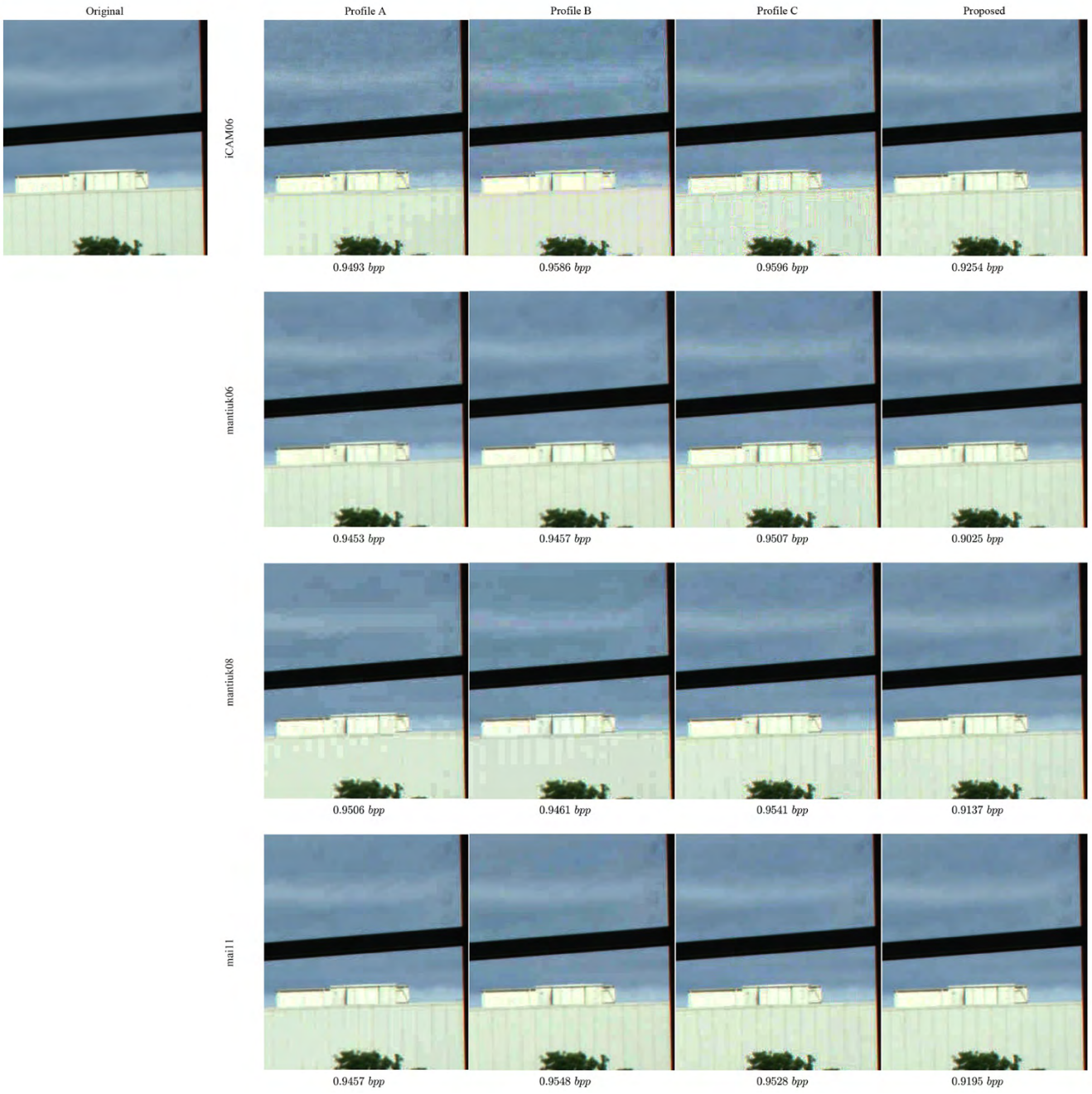


FIGURE 12. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

$x_R^k(m, n)$ in Fig. 2, which is the difference between $x_H^k(m, n)$ and its predicted version $\tilde{x}_H^k(m, n)$, is encoded using an L-JPEG encoder.

The syntax of the proposed file format was compliant with the box-based file format defined in JPEG XT Part 3, except for a newly defined function block to support the proposed block-wise prediction of the HDR image.

C. PROPOSED DECODING PROCESS

The decoding process of the proposed scheme can simply be implemented by reversing the encoding process. Figure 6 shows the overview of the decoding process of the proposed scheme. The decoder first parses the base- and residual-layer codestreams when a decoder that implements the proposed scheme loads the JPEG XT file. An L-JPEG decoder then decodes the base and residual images. The base-layer codestream is a complete L-JPEG file containing an LDR image; thus, no more processing is required to display the LDR image on the LDR devices. The decoder reconstructs the HDR image according to the function blocks in the residual layer once the proposed decoder is configured to display the HDR image on the HDR display.

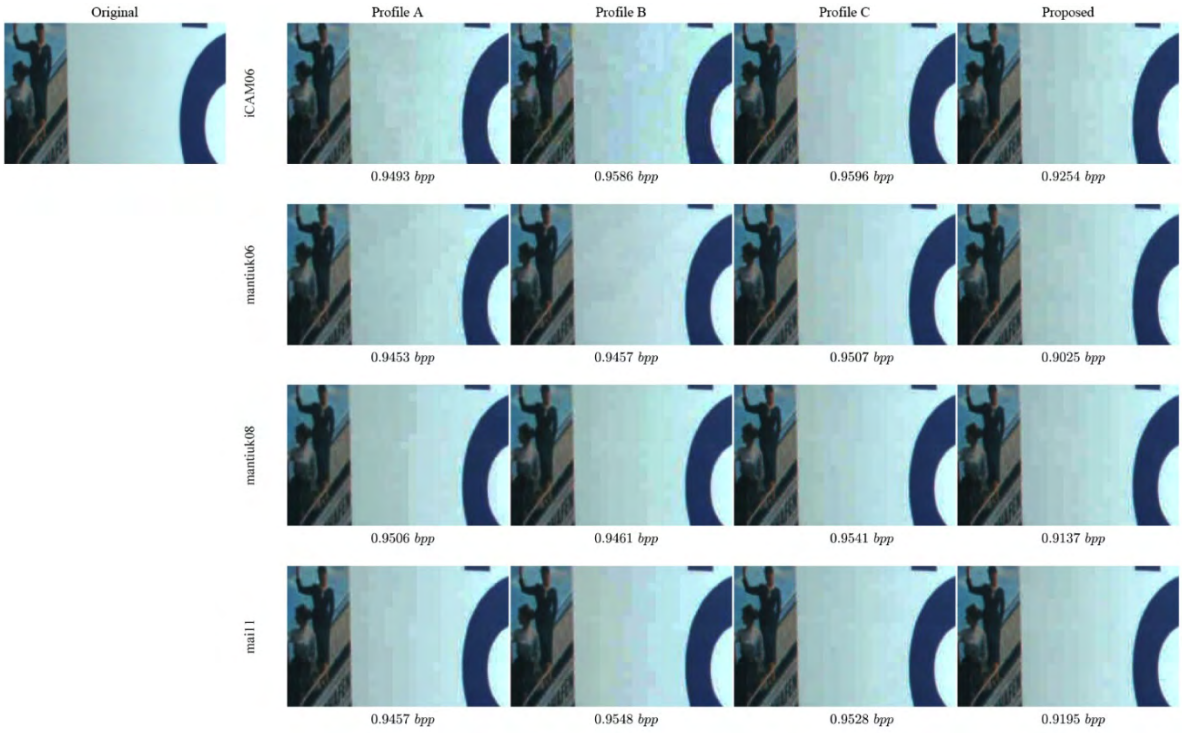


FIGURE 13. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

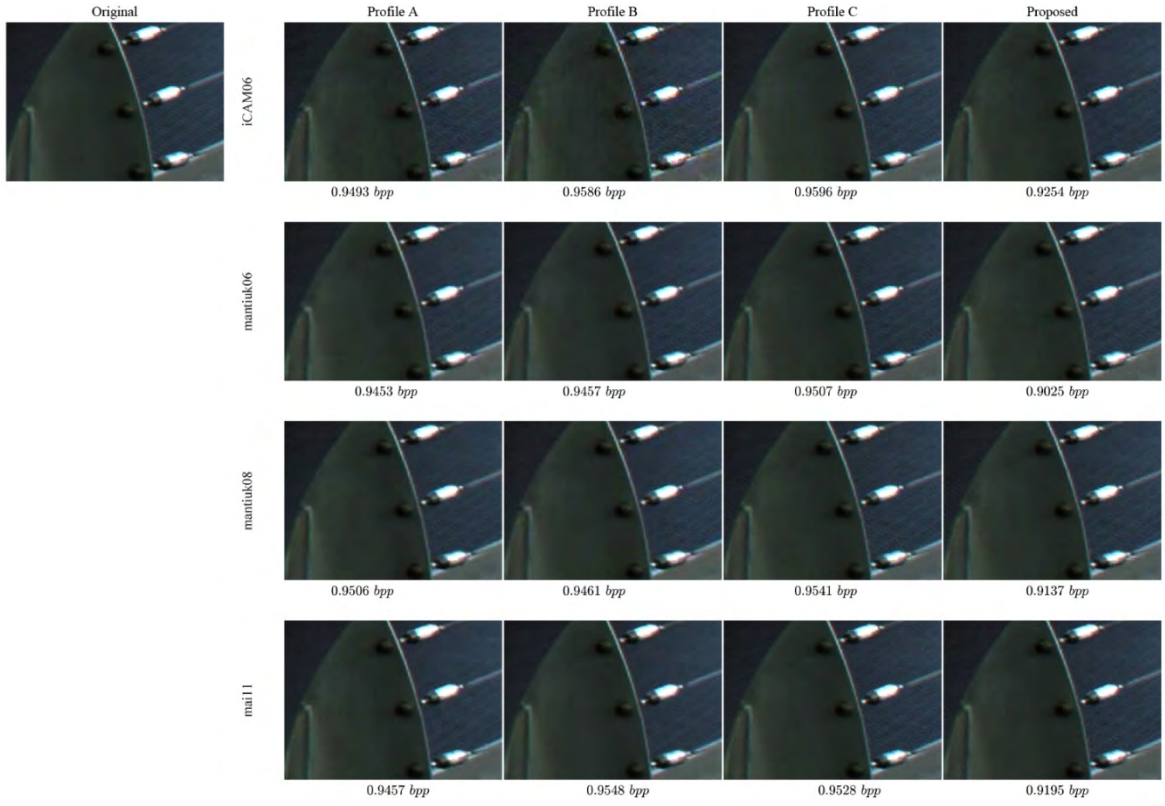


FIGURE 14. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

A function block denoted by *Base Non-linear Point Transformation* in Fig. 6 should be applied to each of the components from the base layer in the YCbCr color space. Details of the point transformation were provided in [4].

The process of the proposed *HLP* in the proposed scheme is defined by the *Base Non-linear Point Transformation* function block. The output of this process is the predicted HDR image. The proposed decoder then reconstructs a normalized

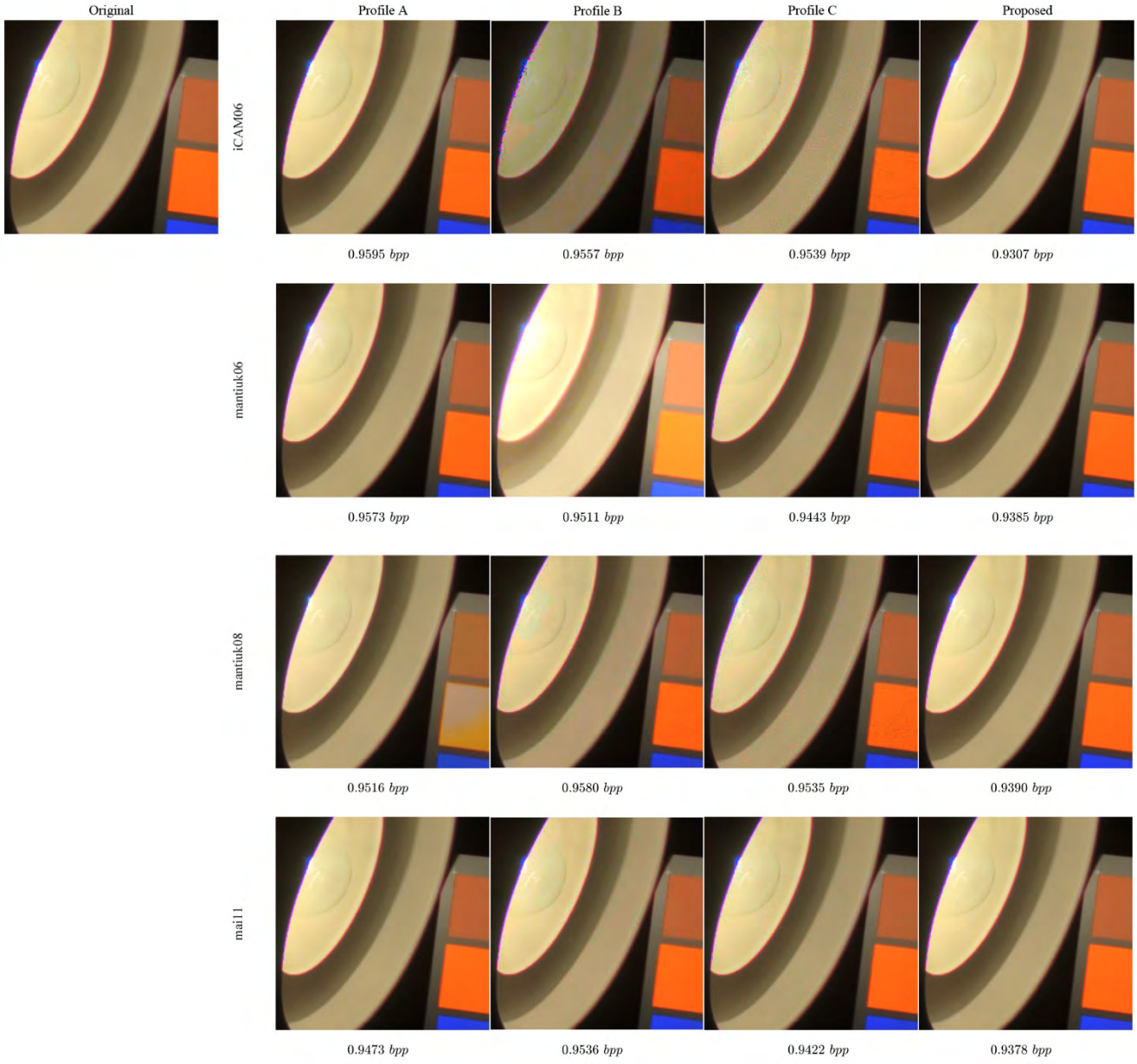


FIGURE 15. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

HDR image denoted by $\hat{x}_H^k(m,n)$ by adding the predicted HDR image to the decoded error image denoted by $\hat{x}_R^k(m,n)$. Finally, the HDR image is reconstructed by scaling, which is defined by the *Post Scaling Transformation* function block, to the range of values corresponding to the original image.

As shown in Figs. 1 and 6, the proposed decoder can be configured as a subset of the full configuration of JPEG XT Part 7 like existing Profiles A, B, and C. However, two function blocks that signal the prediction parameters and the scaling process must be defined anew to predict the HDR value and restore the original range of the HDR image. First, the *Base Non-linear Point Transformation* must be modified to support block-wise processing because the current specification only defines the pixel-wise process. Second, the proposed *Post Scaling Transformation* need to be newly defined. These modifications can be defined within the current syntax of the specification, such that the proposed scheme can be implemented by minor modifications of the current specification.

IV. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETUP

We compared the performance of the proposed coding scheme with that of the JPEG XT profiles. The evaluations were performed using the same images and TMOs used for our observation in Section IIIA. We also performed

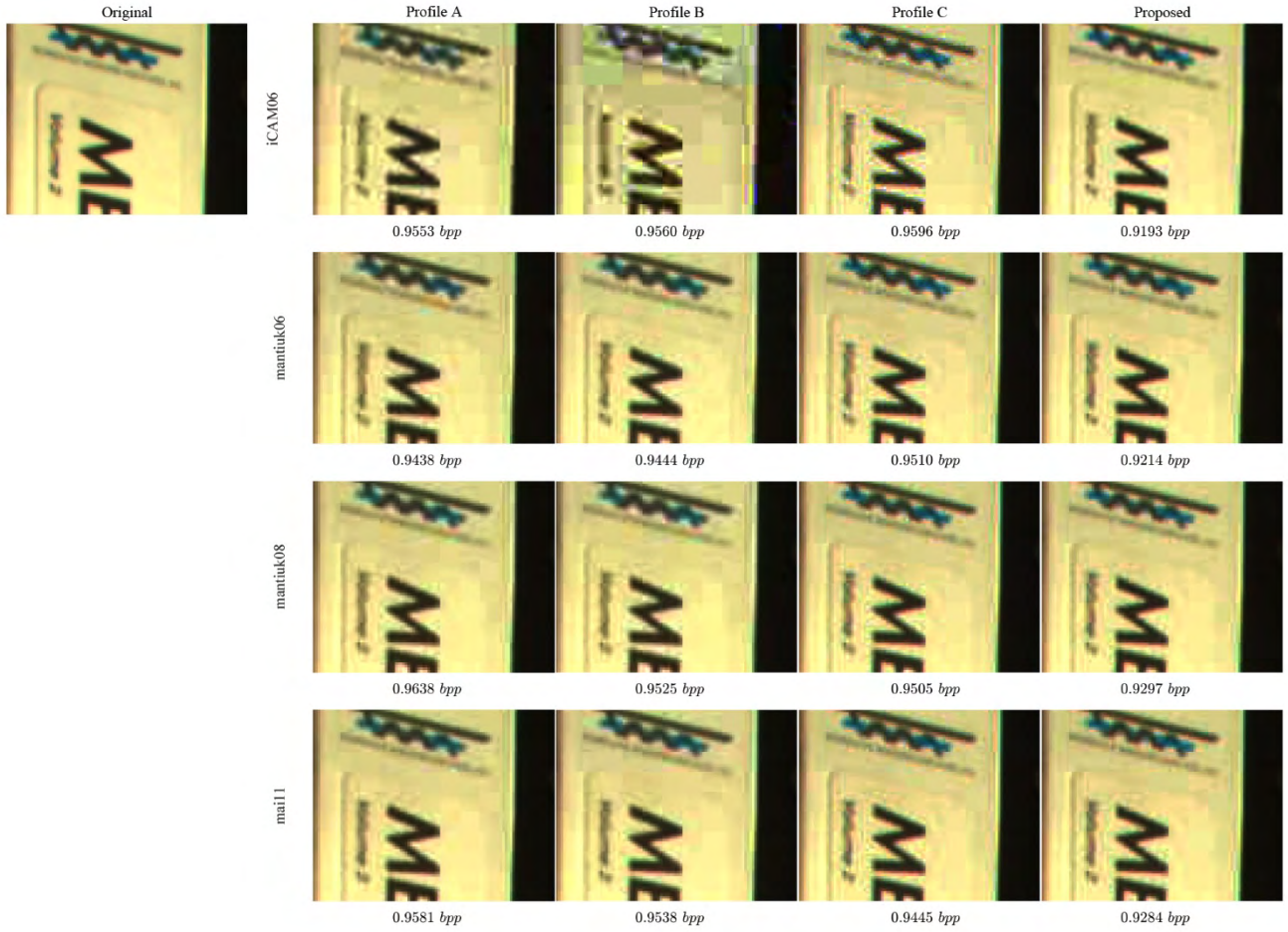


FIGURE 16. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

our experiment on other sample images and arrived at the same conclusion in terms of the performance comparisons. We examined the image quality metric by changing the bit rate via the quality factor of the compared methods. For the implementation of Profiles A, B, and C, we used the reference software provided by the JPEG WG [5] with their recommended parameter settings.

B. ANALYZING THE TMO INFLUENCE ON THE JPEG XT PROFILES

Subjective evaluation methods can generally provide the best results to evaluate the performance of the video and image codec. However, in most cases, objective evaluation methods are used because of their cost efficiency. We evaluated the performance of the existing profiles using the HDR-VDP-2 objective metric [20], which has recently been considered as the best-performing objective quality metric for an HDR image by many researchers [7]. HDR-VDP-2 is a perceptual metric specifically designed for HDR images. The metric evaluates the quality of the decoded images with a quality index between 0 and 100, where 100 is achieved when no visible difference from the original image is observed. The HDR-VDP-2 metric requires several parameters to consider the visualization conditions. We used the predefined default value for the parameters for the experiments.

Figure 7 shows the TMO influence on the JPEG XT profiles. Profile C exhibited the most stable HDR-VDP-2 performance among the three profiles. Profiles A and B exhibited a relatively random behavior, which was highly dependent on the applied TMO for the LDR image of the base layer. However, in the test cases of the *AhwahneeGreatLounge* and *URChapel(2)* images, Profile C also exhibited an inconsistent HDR-VDP-2 performance for different TMOs. We confirmed from these experiments that the existing JPEG XT profiles exhibited an inconsistent performance according to the TMO used for generating the base image (i.e., the choice of TMO influences the compression performance for the same HDR image).

C. PERFORMANCE EVALUATIONS

We used two image quality metrics for the objective evaluation of the proposed scheme compared to the existing JPEG XT profiles. We chose the PSNR [21] in addition to the HDR-VDP-2. The PSNR was chosen because it is a metric

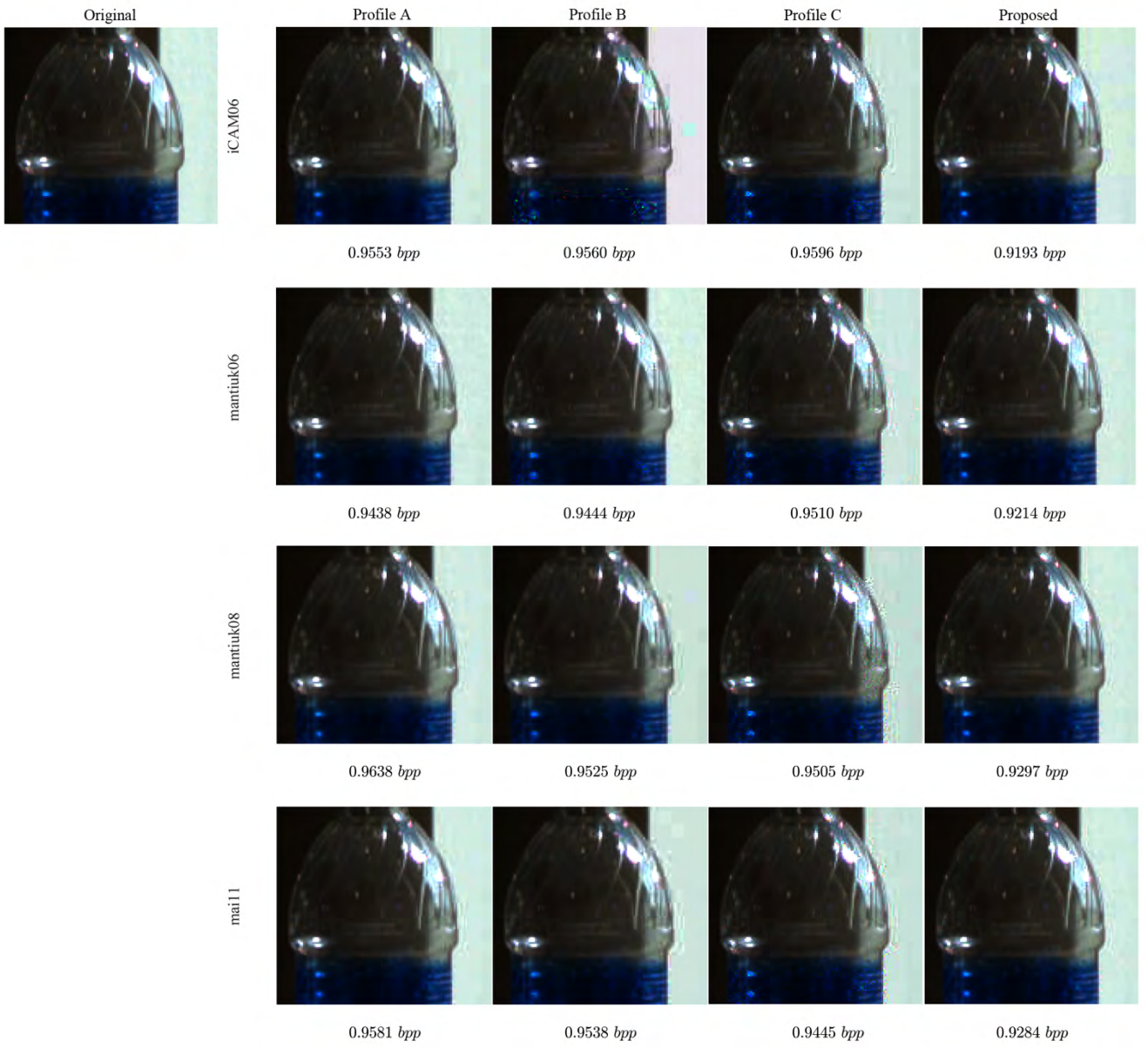


FIGURE 17. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

that has been traditionally used in the history of image compression researches. Only the luminance component was employed for the HDR-VDP-2 calculations, and the programs provided by the respective authors were used. The PSNR metric used in this paper supports color images, and its definition can be found in [22].

Figures 8–10 show the comparison results. Figures 8 and 9 depict the results of the HDR-VDP-2 and PSNR performance, respectively, for each sample image. Figure 10 illustrates those results for all the sample images together in a graph. Each plot depicts the results between 0.25 and 2.5 bits per pixel (bpp) to observe the influence of the TMO selection on the compression performance. The results of the objective evaluation may be summarized as follows:

1) Interestingly, the general aspect of the results of the HDR-VDP-2 and PSNR performance evaluation was quite similar when we compared Figs. 8 and 9.
2) Compared to the existing profiles, the proposed method showed a consistent performance for different TMOs.
3) In terms of both the HDR-VDP-2 and PSNR performances averaged over different TMOs, the proposed method outperforms the existing profiles for the all sample images as shown in Figs. 8 and 9, respectively.
4) The result of the objective comparison averaged over all the sample images confirmed that the proposed method, which adaptively predicted the HDR value for each block, was superior to the existing profiles (Fig. 10).

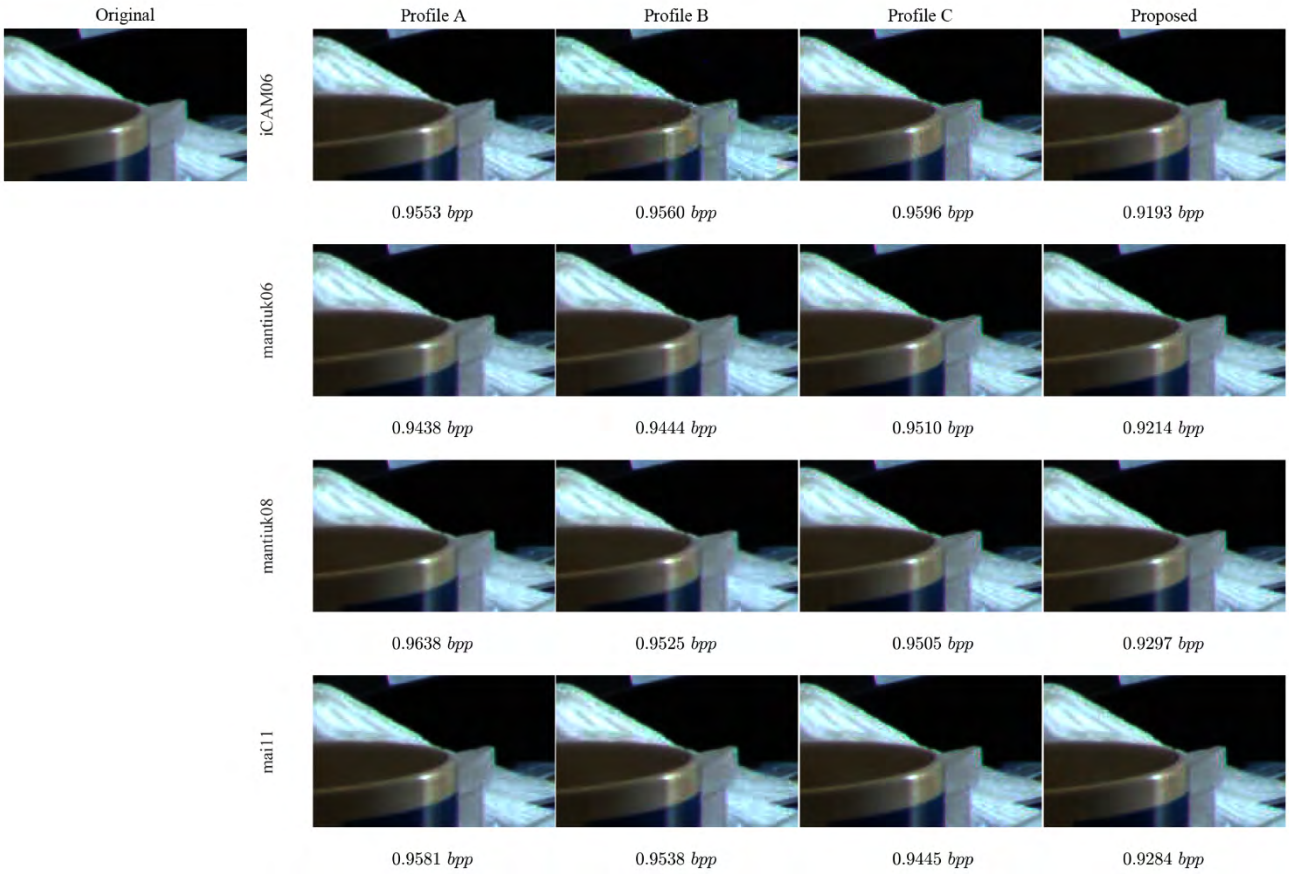


FIGURE 18. Coded image of the compared methods with four different TMOs at approximately 0.95 bpp.

For a visual comparison, the proposed method was visually compared to the JPEG XT profiles. We selected several parts of interest in the coded images (Fig. 11). These parts were selected, such that they include various features (e.g., dark and bright regions with text, texture, and/or flat area). Figures 12–18 show the resulting images coded at approximately 0.95 bpp for four TMOs. We set the bit rate of the proposed method to be the lowest among the compared methods. The regions of the reconstructed HDR image were uniformly tone-mapped and zoomed in to visualize the differences. The overall visual quality of the reconstructed HDR image tended to decline when iCAM06 TMO was used. This phenomenon prominently occurred in Profiles A, B, and C.

The visual comparison showed that the existing JPEG XT profiles showed an inconsistent visual quality according to the TMO used in most comparisons. However, the proposed method visually showed an insignificant difference in the seven resulting images. Most importantly, the proposed method objectively and visually exhibited better performances while consuming fewer bits than the existing JPEG XT profiles. The proposed method exhibited substantially less blurring and blocky artifacts in edges and flat regions and a better texture rendition.

V. CONCLUSION

We describe herein a new scheme for a JPEG backward-compatible HDR image coding scheme that includes a base-layer, which is a tone-mapped LDR version of the HDR image. JPEG XT is a specification that allows any arbitrary (possibly local) TMOs, but only defines the pixel-wisely global inverse transformation. For this reason, JPEG XT shows an inconsistent performance according to the applied TMO for generating the base image. Our method can adapt well to the applied TMO because of the local prediction method based on the observation of a correlation between the HDR and LDR values. Thus, the proposed method enables the control of the rate distortion, which is not dependent on the TMO used for the LDR. The robustness of the prediction is further enhanced by applying scaling to the prediction parameters of the block.

We compared the performance of our method against three existing JPEG XT profiles. We also performed objective tests to evaluate the quality of the coded images. For the objective tests, we compared the results using two objective image quality metrics: PSNR and HDR-VDP-2. Our experiments showed significant coding gains on the residual layer compared with the JPEG XT profiles vulnerable to the local tone-mapping because of the TMO-robustness of the block-wise

local prediction. We demonstrated the improvement of our method using six sample images combined with four TMOs.

Tone mapping is a major technique of HDR imaging, and TMO works with various attributes. Therefore, providing robustness to the TMO might be an important tool in enhancing the reliability of HDR image coding.

REFERENCES

[1] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*. Boca Raton, FL, USA: CRC Press, 2011.

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

[3] T. Richter, T. Bruylants, P. Schelkens, and T. Ebrahimi, "The JPEG XT suite of standards: Status and future plans," *Proc. SPIE*, vol. 9599, p. 95990T, Sep. 2015.

[4] *Information Technology: Scalable Compression and Coding of Continuous-Tone Still Images HDR Floating-Point Coding*, document ISO/IEC 18477-7:2016, ISO/IEC, 2016.

[5] (2017). *JPEG Official Site*. [Online]. Available: http://www.jpeg.org

[6] T. Richter, A. Artusi, and T. Ebrahimi, "JPEG XT: A new family of JPEG backward-compatible standards," *IEEE Trans. Multimedia*, vol. 23, no. 3, pp. 80–88, Jul./Sep. 2016.

[7] A. Artusi *et al.*, "JPEG XT: A compression standard for HDR and WCG images [standards in a nutshell]," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 118–124, Mar. 2016.

[8] M. D. Fairchild, "The HDR photographic survey," in *Proc. Color Imag. Conf.*, 2007, vol. 2007. no. 1, pp. 233–238.

[9] Z. Mai, H. Mansour, R. Mantiuk, P. Nasiopoulos, R. K. 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, Jun. 2011.

[10] G. Valenzise, F. de Simone, P. Lauga, and F. Dufaux, "Performance evaluation of objective quality metrics for HDR image compression," *Proc. SPIE*, vol. 9217, p. 92170C, Sep. 2014.

[11] C. Mantel, S. C. Ferchiu, and S. Forchhammer, "Comparing subjective and objective quality assessment of HDR images compressed with JPEG-XT," in *Proc. IEEE 16th Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2014, pp. 1–6.

[12] P. Hanhart, M. V. Bernardo, P. Korshunov, M. Pereira, A. M. G. Pinheiro, and T. Ebrahimi, "HDR image compression: A new challenge for objective quality metrics," in *Proc. 6th Int. Workshop Quality Multimedia Exper. (QoMEX)*, Sep. 2014, pp. 159–164.

[13] A. Pinheiro *et al.*, "Performance evaluation of the emerging JPEG XT image compression standard," in *Proc. IEEE 16th Int. Workshop Multimedia Signal Process. (MMSP)*, Sep. 2014, pp. 1–6.

[14] S. Choi, O. Kwon, D. Jang, and S. Choi, "Performance evaluation of JPEG XT standard for high dynamic range image coding," in *Proc. World Congr. Eng.*, vol. 1. 2015, pp. 1–5.

[15] R. K. Mantiuk, T. Richter, and A. Artusi, "Fine-tuning JPEG-XT compression performance using large-scale objective quality testing," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2152–2156.

[16] J. Kuang, G. M. Johnson, and M. D. Fairchild, "iCAM06: A refined image appearance model for HDR image rendering," *J. Vis. Commun. Image Represent.*, vol. 18, no. 5, pp. 406–414, 2007.

[17] R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "A perceptual framework for contrast processing of high dynamic range images," *ACM Trans. Appl. Perception*, vol. 3, no. 3, pp. 286–308, Jul. 2006.

[18] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 68.

[19] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*. Norwell, MA, USA: Springer, 1992.

[20] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, 2011, Art. no. 40.

[21] D. Salomon, *Data Compression: The Complete Reference*. New York, NY, USA: Springer, 2004.

[22] S. Choi, O.-J. Kwon, J. Lee, and Y. Kim, "A JPEG backward-compatible image coding scheme for high dynamic range images," *Digit. Signal Process.*, vol. 67, pp. 1–16, Aug. 2017.

OH-JIN KWON received the M.S. degree in electrical engineering from the University of Southern California at Los Angeles, Los Angeles, in 1991, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park, in 1994. He was a Researcher with the Agency for Defense Development of Korea, from 1984 to 1989, and the Head of the Media Laboratory with Samsung SDS Co., Ltd., Seoul, South Korea, from 1995 to 1999. Since 1999, he has been a Faculty Member with Sejong University, Seoul, South Korea, where he is currently an Associate Professor. His research interests include image and video fusion, coding, watermarking, analyzing, and processing.

SEUNGCHEOL CHOI received the B.S. and M.S. degrees in computer science and the Ph.D. degree in electronics engineering from Sejong University, Seoul, South Korea, in 1998, 2001, and 2017, respectively. He was a Researcher with Galaxia Communications from 2001 to 2013. He has been a Post-Doctoral Researcher with Sejong University since 2017. His research interests include image and video coding, high-dynamic range imaging, image processing, image fusion, and JPEG.

DONGKYOO SHIN received the B.S. degree from Seoul National University, South Korea, in 1986, the M.S. degree from the Illinois Institute of Technology, Chicago, Illinois, in 1992, and the Ph.D. degree from Texas A&M University, College Station, Texas, in 1997, all in computer science. He is currently a Professor with the Department of Computer Engineering, Sejong University, South Korea. He was with the Korea Institute of Defense Analyses from 1986 to 1991, where he developed the database application software. He was with the Multimedia Research Institute of Hyundai Electronics Co., South Korea, as a Principal Researcher, from 1997 to 1998. His research interests include natural user interaction, information security, bio-signal data processing, and ubiquitous computing.