Evaluation of the Encoding Accuracy of the PQ based HDR Content Delivery Formats

Asif Siddiq[†], Ishtiaq Rasool Khan[*] and Jameel Ahmed[†]
[†]Riphah International University, Islamabad, Pakistan
E-mail: asif.siddik@gmail.com
E-mail: jameel.ahmed@riphah.edu.pk
[*]University of Jeddah, Jeddah, Saudi Arabia
E-mail: irkhan@uj.edu.sa

***Abstract*—Recently, new formats like HDR10 and Dolby Vision have emerged and become the popular choice for delivering HDR content. HDR10 uses a bit-depth of 10 to quantize the data after applying a non-linear transformation operator known as the perceptual quantizer (PQ). It has been known that at this precision of just 10 bits, banding artifacts can be observed in the lower luminance regions. However, in general, HDR10 has good visual quality on HDR displays and is widely recognized by the display makers and content developers. This paper evaluates the accuracy of the content encoded in HDR10 format, especially when the dynamic range is high. It is observed that the accuracy of HDR10 is low when there are extremely dark and bright regions in the image. Therefore while it is the right choice for HDR display, other formats should be used to preserve the original dynamic range for applications such as machine-based analysis and tone-mapping for LDR display. We also propose a dual-layer codec using companding, which resolves the known drawbacks of two-layer backward compatible formats. The proposed format does not require metadata for the reconstruction of the HDR image, and therefore its decoding function has a low computational cost. Moreover, its encoding accuracy is higher than HDR10 and the existing state of the art popular formats like JPEG-XT and HDR-MPEG.**

Keywords: HDR10, Dolby Vision, Dual Layer Codec, Non-linear Quantization, Perceptual Quantizer

I. INTRODUCTION

High dynamic range (HDR) images capture all regions of the scene with high precision and therefore show the dark and bright areas with much higher accuracy than the standard low dynamic range (LDR) images. The high precision of HDR pixel intensity values requires a higher bit-depth for encoding as the traditional formats of 24 bits per pixel do not have enough distinct quantization levels to store the HDR data accurately. HDR formats such as 96-bit TIFF, 32-bit LogLuv, 48-bit OpenEXR, 32-bit RGBE, and 32-bit XYZE have been proposed to encode the raw pixel values. These formats encode the entire dynamic range of the image and are generally used as the reference, however their size is very large, and they are not backward compatible, i.e., they cannot be displayed on the traditional displays or processed using the existing codecs.

A solution to the problem came in the form of backward-compatible two-layer formats. For encoding in these formats, the HDR image is transformed into an LDR image through a tone-mapping process and encoded in a standard 24-bit format. Additional information describing the tone-mapping process, which can be used to predict the HDR image, i.e., to reverse the tone-mapping operation, is also encoded. Finally, the residual data, i.e., the predicted HDR image's difference from the ground truth, is encoded in one or more additional channels. The base layer, i.e., the 24-bit LDR image, is for the traditional displays and applications, while the enhancement layer, i.e., the residual data and the metadata giving information of the tone-mapping operation, are used by HDR-enabled displays and applications. Different two-layer formats differ in the process of creating the enhancement layer. JPEG-XT, the extended JPEG standard, implements some of them in its various profiles. Two-layer formats could be handy during the transition period when both LDR and HDR displays are in use; however, they have the drawbacks of increased decoder complexity due to the use of metadata and relatively poor overall compression.

To resolve the issues of the backward-compatible two-layer designs, some formats relaxed the restriction of backward compatibility, some of them increased the bit-depth of the base layer to 10 or 12 bits and dropped the enhancement layer, while some others dropped both the enhancement layer and the metadata. These approaches are particularly effective for the encoding of HDR video. The most famous recent formats, HDR10 and Dolby Vision, do not use the enhancement layer. Perceptual quantizer (PQ) is a new electro-optical transfer function (EOTF) used to quantize the content replacing the standard gamma curve. The term PQ10 refers to the content quantized linearly using a bit-depth 10 after applying PQ transformation. HDR10 uses PQ10 encoding with some static metadata, i.e., the metadata does not change for the entire scene or the video. Dolby vision uses PQ10 (or PQ12 using 12 bits) but allows more flexibility by using dynamic metadata changing with every frame. As the use of static metadata is restricting, a variant of HDR10 called HDR10+ is gaining popularity, which allows using dynamic metadata.

HDR10 is a simple format compatible with almost all HDR displays. HDR10 is a display-referred format, which means that the encoded values represent the actual display luminance values. Except for some adjustments in dark and bright pixels, the encoded values can be rendered on display without processing. The format has been widely studied and accepted for the delivery and visualization of HDR content and is undoubtedly the favorite choice for HDR display. For LDR display, the format is not backward compatible and requires tone-mapping. Here dual-layer formats become more useful as they contain the LDR version in the base layer and can display it without any additional processing.

As for accuracy, to the best of our knowledge, there have not been any comprehensive studies yet comparing the single layer HDR10 (and Dolby Vision) with two-layer formats. This is true that HDR10 is intended only for the delivery of content for viewing, but in the applications beyond viewing where contents are processed by the machines (such as in scene analysis, medical diagnosis, and autonomous vehicle control, etc.), a higher level of accuracy is desirable. This work intends to fill this gap and presents a detailed comparison of the accuracy of HDR10 and different variants of two-layer formats with reference to the ground truth raw images. The aim is to determine if the precision of 10 bits (or 12 bit) per channel in single layer structures is sufficient for preserving the high dynamic range of natural scenes. In addition, we present a new two-layer format, which addresses some drawbacks of the existing two-layer formats, and include that in the said comparative evaluations. Unlike existing two-layer formats, the proposed format does not require metadata. It has a very efficient decoding operation, making it a preferred choice for real-time applications on devices that have lower computational and/or battery power.

II. Existing Techniques of HDR Encoding

This section presents some of the existing techniques used for encoding the HDR images and video. Traditional single-layer raw formats, backward and non-backward compatible two-layer formats, and the new single-layer formats that apply PQ or HLG transformation before quantizing the content to the fixed bit-width are briefly reviewed.

A. Traditional Raw HDR formats

Finer details captured in HDR images require higher number of bits per pixel to retain precision. Floating point format is therefore used to encode raw HDR data loselessly with little or no compression. TIFF encodes uncompressed data using 96 bits per pixel [1]. LogLuv is implemented as a part of public TIFF library and is a more compact format having 24 and 32 bits per pixel [2], [3]. Industrial Light and Magic (ILM) introduced OpenEXR format that encodes data as 16 bits per channel in “half” floating-point type reserving 1 sign bit, 10 bits for the mantissa, and 5 bits for the exponent [4]. OpenEXR supports 96 and 72 bits per pixel data types but 48 bits per pixel (16 bits per channel) is the most widely used type. Radiance picture format known as RBBE encodes three floating-point channels of the HDR data into four 8-bit integer channels [5]. The fourth channel is the common exponent of the rest three. Covering the entire visible gamut, CIE variant XYZE format has slightly lesser resolution than RGBE. Lossless options in JPEG2000 [25] and JPEG-XT [16], [17] can also be used to encode the raw HDR data.

B. Two-layer backward compatible formats

The basic idea behind two-layer formats is to prepare a tone-mapped version of the HDR image and store it as the base layer while keeping some additional details in the enhancement layer [6]. This ensures backward compatibility with the existing LDR applications and displays. JPEG-HDR presented by G. Ward [7] encodes the tone-mapped image in standard JPEG format in the base layer, and the luminance ratio of the HDR and LDR versions in the enhancement layer as another JPEG image in the wrappers provided in JPEG. HDR-MPEG format by Mantiuk et al. [8] uses inverse tone-mapping functions to reduce the dynamic range of the enhancement layer for improved coding efficiency. The inverse function is generated in the form of a look-up table (LUT) by taking the average of all HDR values that map to an LDR value. The LUT is used to reconstruct the HDR image, and the error in this step is saved in the enhancement layer.

Many variants using the above concept have been proposed [9-19]. Hill function was used in [9],[10] to smooth the LUT to reduce the enhancement layer's entropy by removing the high-frequency noise. Using the same idea, piecewise linear functions [11, 12] and cubic splines [14] have been used to remove the noise and improve the coding efficiency. Some works used non-linear quantization intervals in the enhancement layer to improve the coding efficiency [20]. JPEG-XT, the extended standard by JPEG, includes some of the two-layer formats for saving HDR images in JPEG format [13].

Some single-layer lossless formats were proposed in the past, but they did not get much popularity [21-23]. On the other hand, some works presented lossless two-layer formats by encoding the first or both layers losslessly using a format like lossless JPEG2000 [15].

C. Two-layer non-backward compatible formats

In the two-layer non-backward-compatible format proposed by Su et al. [24], the base layer is formed using any quantization method, not necessarily using a tone-mapping algorithm. In other words, backward compatibility is not a design constraint in this format. The enhancement layer is encoded after applying a non-linear quantization algorithm. Parameters used in generating both layers are encoded along with the data for reconstruction of the dynamic range at the decoder. Mai et al. [25] proposed a dual-layer structure in which the based layer is obtained by tone-mapping, but the tone-mapping curve is optimized for coding efficiency while visual quality is not the only or the significant decision parameter. Dufaux et al. [26] proposed a signal splitting and recombination at the encoder and decoder side, respectively. The input image is split into the most significant bits (MSB) and the least significant bits (LSB), forming two layers for HDR transmission. A preprocessing step is used in [27, 28], where the input HDR image is split into two images using a non-linear mapping function – a monochromatic modulation picture and a residual picture. The modulation picture consists of relatively low-frequency components of the input image, and a residual picture represents the remaining relatively high-frequency components. The residual picture can be used for backward compatibility or to improve the compression ratio. For the latter case representing a non-backward-compatible scenario, the split images after going through some color space and signal format changes are encoded and transmitted to the decoder. At the decoder end, an inverse transformation and recombination of signals from the two layers are used to reconstruct the HDR image.

D. HDR10, Dolby Vision, HLG and other single layer formats

The Optical to Electrical Transfer Function (OETF) determines how the scene/image captured by a camera is coded. In contrast, Electrical to Optical Transfer Function (EOTF) determines how the coded information will be displayed on the screen. HDR is characterized by high luminance and wider color gamut (WCG). The Gamma curve, which has been used traditionally as the EOTF for rendering the content on the LDR displays is no longer capable of reproducing the content on the HDR displays. Dolby developed a new non-linear EOTF curve called the PQ, which has been widely accepted and standardized by SMPTE ST-2084 [29] and ITU-R BT.2100 [30]. PQ has been designed to cover luminance values from 0.001 to 10,000 nits (cd/m^2). Depending on the distribution pipeline, PQ can be coded in 10 bits (PQ10) or 12 bits (PQ12). The commonly used HDR10 uses PQ10 while Dolby Vision, which is Dolby's suite for HDR technologies, offers both 10-bit and 12-bit modes. Encoding with 10 bits shows some banding issues as few codes in low luminance lie above the Barten's ramp [31], while 12-bit mode lies below the Barten's ramp for all code and produces no banding artifacts.

Besides having a lower bit-depth, HDR10 uses only static metadata, which remains fixed for the entire scene. The static metadata includes the Maximum Content Light Level (MaxCLL), which represents the brightest pixel in the whole video stream, and the Maximum Frame-Average Light Level (MaxFALL), which is the maximum of the average frame values of the entire video stream [32]. In addition, characteristics of the mastering display (on which the creative intent was established) specified in SMPTE ST 2086 [33] (including display primaries, the white point, and the display luminance range) are included. Dolby Vision includes additional dynamic metadata that can change with every frame and is based on the SMPTE ST 2094-10 [34]. Dolby Vision allows YCbCr and ICpCt signal formats. When YCbCr is used with a bit-depth of 10, the stream becomes fully compatible with HDR10 players who can simply ignore the supplementary enhancement information (SEI) dynamic metadata. The metadata in HDR10 and Dolby Vision is contained in the SEI messages of the High-Efficiency Video Coding (HEVC) format [35].

Some other similar formats for HDR content delivery have been proposed recently. Philips has designed its own HDR delivery system using a single layer along with metadata [36]. It utilizes the 10-bit HEVC codec, along with Philips HDR EOTF and display tuning, to optimize the image's peak luminance. Technicolor HDR supports HDR content production and delivery to HDR and legacy SDR displays [37]. The technologies and standards used in this workflow facilitate an open approach, including a single layer SDR/HDR HEVC encoding, the MPEG standardized Color Remapping Information metadata (CRI) for HDR to SDR conversion, a Parameterized Electro-Optical Transfer Function (P-EOTF), and SHVC.

Hybrid Log-Gamma (HLG) is another transfer function developed for camera capture (OETF) jointly developed by BBC and NHK Japan specified in ARIB STD-B67 [13]. Later ITU-R BT.2100 [30] defined display EOTF and OOTF as well. The HLG curve maintains compatibility with the LDR displays using the standard gamma curve for the lower luminance values and the Log curve for the higher luminance values. The term HLG10 is coined for the combination of HLG, BT.2100 wide gamut colorimetry, and quantization at bit-depth 10. Unlike PQ based display referred systems, HLG is a scene-referred format. This means that the pixel values are proportional to the scene intensities and are adjusted depending on the display capabilities and the viewing environment.

III. NON-LINEAR QUANTIZATION USING COMPANDING

Quantization is inherently a lossy process. Lower signal values suffer more percentage loss as compared to higher signal values in uniform quantization. One way to reduce this effect is to increase the number of the quantization intervals, but it increases the number of bits, thus increasing the data. A more preferred solution is to use non-linear quantization. The underlying idea in non-linear quantization of images is generally to utilize more quantization intervals for darker pixels where details are easier to be lost compared to the brighter pixels. Gamma correction and PQ are both examples of this. Display algorithms, such as tone-mapping operators, use other strategies to form the non-linear transformations for better visualization [38-44].

Companding, a short form of the term "compression expanding", is a non-linear technique that reduces the dynamic range of a signal by nonlinearly raising the amplitude of the weak level signals at the coder side and bringing them back to original values at the decoder side. Companding has been standardized by ITU-T G.711 and is widely used for compression and expanding audio (voice and unvoiced) signals. Besides being a non-linear process, companding does not require metadata to reconstruct the signal at the decoder end, as it is defined as a closed-form function. The compression and expansion functions are given below in equations (1) and (2), respectively.

$$y = F(x) = \mathrm{sgn}(x)\frac{ln(1+\mu|x|)}{ln(1+\mu)}, \quad -1 \le x \le 1 \qquad (1)$$

$$x = F^{-1}(y) = \mathrm{sgn}(y)\frac{(1+\mu)^{|y|}-1}{\mu}, \quad -1 \le y \le 1 \qquad (2)$$

where μ is a constant that determines the slope of the transformation curve. Performance analysis of few two layers codec has been done in [45].

We reshape the luminance channel only using equation (1), while chrominance channels are left unchanged. All three channels are then linearly quantized, each using 8 bits per pixel. The quantized luminance is transformed back to HDR using (2), and the difference between the original and the reconstructed values is again quantized linearly using 8 bits per pixel and saved in the enhancement layer. The structure is not different from the existing two-layer formats, but there is no metadata requirement, making the decoding operation simpler. A disadvantage is that the first layer is not suitable for viewing on the LDR displays; however, the requirement of backward compatibility has already been dropped in many two-layer formats for improved efficiency and accuracy, as described in detail in section II-C. A detailed evaluation of the proposed format will be presented in the next section.

IV. Experimental Evaluations

In this section, we evaluate the accuracy of the HDR10 format compared to some existing two-layer formats. The two-layer formats used in this study include JPEG-XT Profile C [7], HDR-MPEG [8], Khan 2016 [20], and the companding based scheme explained in section III. The original images in RGBE formats are taken as the reference to evaluate the results of these algorithms. The experiments are conducted using 20 images from the HDR datasets on the accompanying disk of [46]. The images having a higher value of the dynamic range were chosen from the dataset to evaluate the systems' performance in relatively challenging cases. Note that the results shown for HDR10 also refer to the performance of the 10-bit version of Dolby Vision, as both formats use the same encoding and decoding pipelines and differ only in the type of metadata. In the results shown here for HDR10, we assume that the required metadata for reconstruction of the HDR content is available.

For the evaluations, the following five objective metrics were used: mean square error (MSE), peak signal to noise ratio (PSNR), signal to noise ratio (SNR), objective mean opinion score of quality in HDR-VDP2 metric [47], and a PSNR metric defined for HDR images [48]. Graphs in Fig. 1 to Fig. 5 show the average performance of formats on the test dataset.

Fig. 1 shows the values of the MSE in the reconstructed images of the formats being evaluated. A smaller value is preferred. The proposed format has the lowest value improving 57% from the second-best HDR-MPEG. HDR10 has the second largest MSE value after JPEG-XT. The proposed format performs 12X better than HDR10.

Fig. 2 shows the average PSNR scores. An improvement of 5.05% in PSNR is observed in the proposed system compared to the second-best Khan 2016. The second-lowest value observed is again for HDR10. The proposed codec outperforms HDR10 by 27%.

Fig. 3 shows the average SNR scores of all the formats. The proposed codec again outperformed the remaining. There is an improvement of 4.69% over the second-best Khan 2016. HDR10 again gets the second-lowest rank and stands 25% behind the proposed codec.

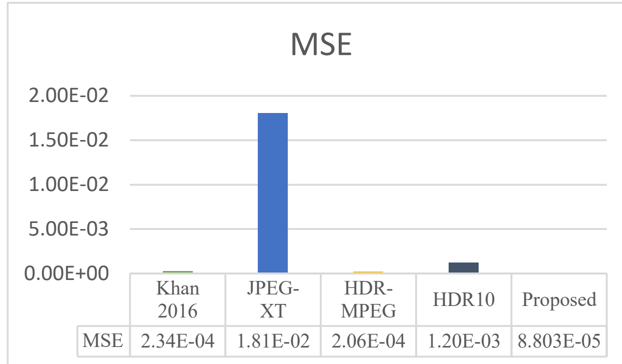

Fig. 1 Mean Square Error

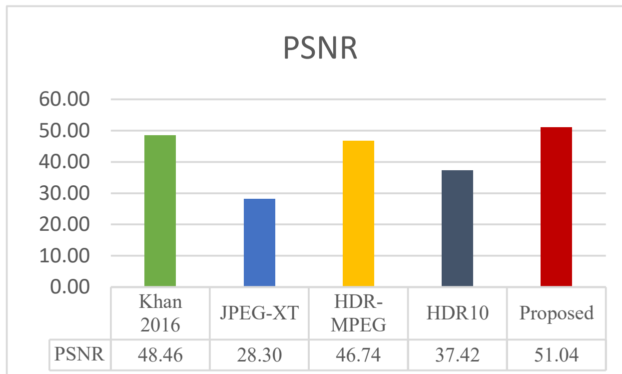

Fig. 2 Peak Signal to Noise Ratio

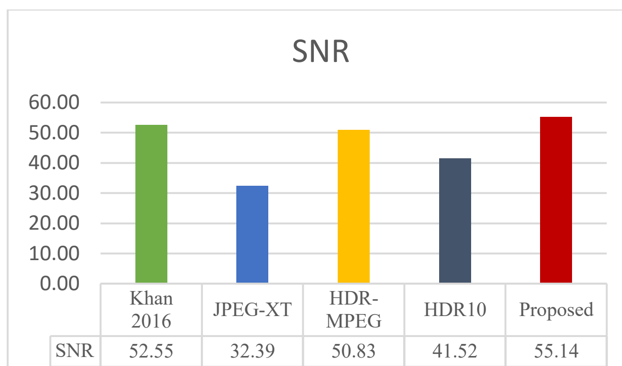

Fig 3 Signal to Noise Ratio

Fig. 4 shows the mean opinion score (MOS) obtained by the HDR-VDP2 metric which is determined considering the characteristics of the HVS [47]. The proposed format has the best value with an improvement of 4.94% from the second-best Khan 2016, while HDR10 remains the second lowest again trailing by the proposed codec by 8%.

The results above clearly show that the proposed system outperforms the other systems by a fair margin. It can be argued that PSNR and SNR metrics were originally defined for LDR images and therefore their values are not very meaningful for HDR images. Another argument could be that HDR10 uses 30 bits per pixel while other formats in the comparison use 32 bits.

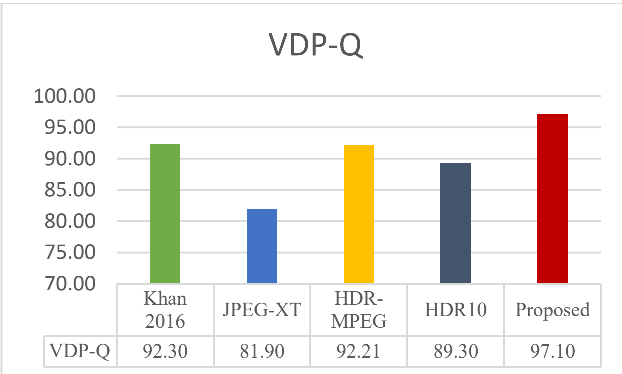

Fig. 4 Objective Mean Opinion Score by HDR-VDP2

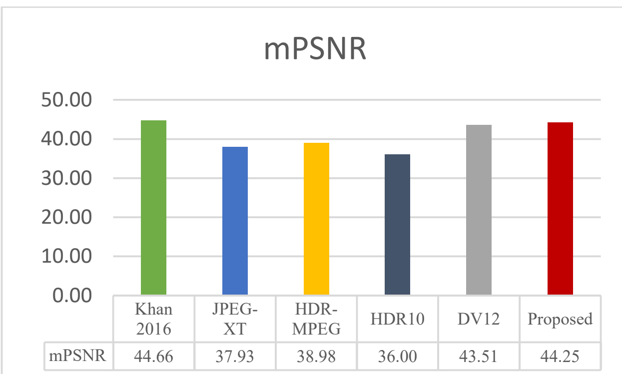

Fig. 5 HDR Peak Signal to Noise Ratio

Therefore. we conducted another set of experiments in which a PSNR metric especially designed for the HDR images, called mPSNR and implemented in the HDR Toolbox by Benterle [48] was used. Moreover, we included the 12-bit version of Dolby Vision format also in this comparison, which uses 36 bits per pixel. The results are shown in Fig. 5.

It can be noted that even for mPSNR both HDR10 and 12-bit Dolby Vision have lower scores when compared with the proposed codec. In fact, HDR10 scored the lowest for this metric. The score of the proposed format is approximately equal to Khan 2016, which is around 12% higher than JPEG-XT and HDR-MPEG, 19% higher than HDR10, and 2% higher than 10.

In the above experiments, we used lossless encoding on the final quantized outputs of each format (except JPEG-XT) to compare the encoding algorithm's performance alone without considering other sources of error caused by lossy encodings. JPEG-XT is inherently a lossy format but allows a lossless mode. However, in lossless mode, the image size becomes very large, even larger than the reference HDR image size. Therefore, for these experiments above, we tuned the parameters such that its performance became comparable to the other two-layer formats. The quality of the base layer was set to 100 (the maximum) while that of the enhancement layer was set to 98.

V. Discussion and Conclusion

The paper addressed two important topics related to the encoding of HDR images and videos. Two-layer structures that got much interest from researchers over the past years lost their popularity due to higher decoder complexity and poor compression. More efficient formats like HDR10 and Dolby vision have become de facto industry standards for HDR content delivery. In this paper, it was shown that despite having acceptable display quality HDR10 is not as good as 2-layer formats. This means that if the contents are intended to be analyzed by machines (instead of viewing by humans) in applications like scene analysis, medical diagnostics, autonomous vehicle control, astronomical imaging, etc., then the contents encoded in HDR10 format might not be the right choice. Moreover, for the LDR display, where HDR contents need to be tone-mapped, the original contents should be utilized instead of those reconstructed from the HDR10 format. This is to say that the existing raw and two-layer formats that have much higher accuracy compared to HDR10 are still relevant and find their applications. Some comparisons of the accuracy of HDR10 and some two-layer formats were presented, and the latter outperformed the former, except for JPEG-XT, by fairly a wide margin in every objective metric used in the evaluations.

The second issue addressed in this paper is related to the existing two-layer formats. Using the metadata requires handling an additional set of values and this, along with the prediction module, adds to the complexity of the decoding function. These factors restricted the application of these formats and were addressed in this paper. A new two-layer structure was presented, which uses a close form function for prediction and does not require metadata, which reduces the computational complexity. Moreover, the accuracy of the proposed format is better than the existing two-layer formats. It can be expected that through such developments, the two-layer formats might regain their popularity, given that their accuracy is much higher than the single-layer HDR10 and Dolby Vision.

VI. Acknowledgement

This project was funded by Lockheed Martin and King Abdulaziz City for Science and Technology (KACST). The authors, therefore, acknowledge with thanks their technical and financial support.

References

[1] A. D. Association, "TIFF" revision 6.0, Jun. 3, 1992," *Adobe Systems Incorporated, Mountain View, CA*.

[2] G. W. Larson, "Overcoming gamut and dynamic range limitations in digital images," in *Color and imaging conference*, 1998, vol. 1998, no. 1, pp. 214-219: Society for Imaging Science and Technology.

[3] G. Ward-Larson, " LogLuv Encoding for Full-Gamut, High-Dynamic Range Images," *Journal of Graphics Tools,* vol. 3, pp. 15-31, 2012.

[4] R. Bogart, F. Kainz, and D. Hess, "OpenEXR image file format," *SIGGRAPH , Sketches & Applications* . 2003.

[5] G. Ward, *Real pixels* (Graphics Gems II). San Francisco C.A Morgan Kaufmann, 1991, pp. 80-83.

[6] K. E. Spaulding, R. L. Joshi, and G. J. Woolfe, "Using a residual image formed from a clipped limited color gamut digital image to represent an extended color gamut digital image," USA, 2001.

[7] G. Ward and M. Simmons, "JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG," presented at the ACM SIGGRAPH 2006 Courses, 2006.

[8] R. Mantiuk, A. Efremov, K. Myszkowski, and H.-P. Seidel, "Backward compatible high dynamic range MPEG video compression," in *ACM SIGGRAPH 2006 Papers*, 2006, pp. 713-723.

[9] M. Okuda and N. Adami, "Two-layer coding algorithm for high dynamic range images based on luminance compensation," *Journal of Visual Communication Image Representation,* vol. 18, no. 5, pp. 377-386, 2007.

[10] M. Okuda and N. Adami, "JPEG compatible RAW image coding based on polynomial tone mapping model," *IEICE transactions on fundamentals of electronics, communications computer sciences,* vol. 91, no. 10, pp. 2928-2933, 2008.

[11] K. Ishtiaq Rasool, "Two layer scheme for encoding of high dynamic range images," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008, pp. 1169-1172.

[12] I. R. Khan, "HDR image encoding using reconstruction functions based on piecewise linear approximations," *Multimedia Systems,* vol. 21, no. 6, pp. 615-624, 2015.

[13] T. Richter, "On the standardization of the JPEG XT image compression," in *2013 Picture Coding Symposium (PCS)*, 2013, pp. 37-40.

[14] I. R. Khan, "Effect of smooth inverse tone-mapping functions on performance of two-layer high dynamic range encoding schemes," *Journal of Electronic Imaging,* vol. 24, no. 1, p. 013024, 2015.

[15] M. Iwahashi and H. Kiya, "Two layer lossless coding of HDR images," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 1340-1344.

[16] T. Fujiki, N. Adami, T. Jinno, and M. Okuda, "High dynamic range image compression using base map coding," in *Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference*, 2012, pp. 1-4.

[17] L. Jian, H. Firas, and C. Joan, "Embedding high dynamic range tone mapping in JPEG compression," in *Proc.SPIE*, 2013, vol. 8655.

[18] K. Pavel and E. Touradj, "Context-dependent JPEG backward-compatible high-dynamic range image compression," *Optical Engineering,* vol. 52, no. 10, pp. 1-12, 8/1 2013.

[19] I. R. Khan, "A new quantization scheme for HDR two-layer encoding schemes," in *ACM SIGGRAPH 2014 Posters*, 2014, pp. 1-1.

[20] I. R. Khan, "A nonlinear quantization scheme for two-layer hdr image encoding," *Signal, Image Video Processing,* vol. 10, no. 5, pp. 921-926, 2016.

[21] D. S. Taubman and M. W. Marcellin, "JPEG2000," *Kluwer International Series in Engineering Computer Science,* 2002.

[22] R. Xu, S. N. Pattanaik, and C. E. Hughes, "High-dynamic-range still-image encoding in JPEG 2000," *IEEE Computer Graphics and Applications,* vol. 25, no. 6, pp. 57-64, 2005.

[23] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H.-P. Seidel, "Perception-motivated high dynamic range video encoding," *ACM Transactions on Graphics,* vol. 23, no. 3, pp. 733-741, 2004.

[24] G.-M. Su, S. Qu, S. N. Hulyalkar, T. Chen, W. C. Gish, and H. Koepfer, "Layer decomposition in hierarchical VDR coding," 2016.

[25] 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 transactions on image processing,* vol. 20, no. 6, pp. 1558-1571, 2010.

[26] F. Dufaux, P. Le Callet, R. Mantiuk, and M. Mrak, Elsevier, Ed. *High dynamic range video: from acquisition, to display and applications*. UK: Academic Press, 2016.

[27] S. Lasserre, F. Le Leannec, P. Lopez, Y. Olivier, D. Touze, and E. Francois, "High Dynamic Range Video Coding," in *Joint Collaborative Team on Video Coding (JCT-VC), 16th Meeting*, San Jose, CA, 2014, pp. 1-8.

[28] F. Le Leannec *et al.*, "Modulation channel information SEI message," in *Document JCTVC-R0139 (m33776), 18th JCT-VC Meeting,*, Sapporo, Japan, 2014.

[29] *High dynamic range electro-optical transfer function of mastering reference displays*, 2014.

[30] ITU-R. (2017). *Image parameter values for high dynamic range television for use in production and international programme exchange*. Available: http://www.itu.int/rec/R-REC-BT.2100

[31] G. J. B. Peter, "Formula for the contrast sensitivity of the human eye," in *Proc.SPIE*, 2003, vol. 5294.

[32] [MaxCLL]:. (2002, July, 30). *CTA 861-G A DTV Profile for Uncompressed High Speed Digital Interfaces*. Available: http://www.techstreet.com/standards/cta-861-g?product_id=1934129

[33] SMPTE, "[SMPTE ST 2086] Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images," *The Society of Motion Picture Television Engineers Journal,* 2014.

[34] SMPTE, "ST 2094-10: 2016:" Dynamic Metadata for Color Volume Transform-Application# 1," vol. 21, pp. 1-15, June, 13 2016.

[35] "ISO/IEC: Doc. ISO/IEC 23008-2:2015 Information Technology—High Efficiency Coding and Media Delivery in Heterogeneous Environments—Part 2: High Efficiency Video Coding," 2013.

[36] R. Goris, R. Brondijk, and R. van der Vleuten, "Philips response to CfE for HDR and WCG," presented at the M-36266,ISO/IEC JTC1/SC29/WG11, 112th MPEG meeting, Warsaw, Poland, September, 2015.

[37] S. Lasserre, E. François, F. Le Léannec, and D. Touzé, "Single-layer HDR video coding with SDR backward compatibility," in *Applications of Digital Image Processing XXXIX; 997108*, San Diego, California, United States, 2016, vol. 9971: International Society for Optics and Photonics.

[38] I. R. Khan, W. Aziz, and S. Shim, "Tone-Mapping Using Perceptual-Quantizer and Image Histogram," *IEEE Access,* vol. 8, pp. 31350-31358, 2020.

[39] A. Almaghthawi, F. Bourennani, and I. R. Khan, "Differential Evolution-based Approach for Tone-Mapping of High Dynamic Range Images," *International Journal of*

Advanced Computer Science and Applications, vol. 11, pp. 341-347, 2020.
[40] A. Mehmood, I. R. Khan, H. Dawood, and H. Dawood, "Enhancement of CT images for visualization," in *ACM SIGGRAPH 2019 Posters*, 2019, pp. 1-2.
[41] J. Han, I. R. Khan, and S. Rahardja, "Lighting condition adaptive tone mapping method," in *ACM SIGGRAPH 2018 Posters*, 2018, pp. 1-2.
[42] I. R. Khan, S. Rahardja, M. M. Khan, M. M. Movania, and F. Abed, "A tone-mapping technique based on histogram using a sensitivity model of the human visual system," *IEEE Transactions on Industrial Electronics,* vol. 65, no. 4, pp. 3469-3479, 2017.
[43] I. Khan, Z. Huang, F. Farbiz, and C. Manders, "HDR image tone mapping using histogram adjustment adapted to human visual system," in *7th International Conference on Information, Communications and Signal Processing (ICICS)*, Macau, China, 2009, pp. 1-5: IEEE.
[44] S. Rahardja *et al.*, "Eye HDR: Gaze-adaptive system for displaying high-dynamic-range images," in *ACM SIGGRAPH ASIA 2009 Art Gallery & Emerging Technologies: Adaptation*, 2009, pp. 68-68.
[45] A. Siddiq, J. Ahmed, and I. R. Khan, "A Companding Based Two-Layer codec for HDR images," presented at the ACM SIGGRAPH 2020 Posters, 2020.
[46] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High dynamic range imaging: acquisition, display, and image-based lighting*, 2nd ed. Burlington, MA 01803, USA: Morgan Kaufmann, 2010.
[47] 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 Transactions on graphics,* vol. 30, no. 4, pp. 1-14, 2011.
[48] F. Banterle. (July 29). *Hdr toolbox for matlab*. Available: https://github.com/banterle/HDR_Toolbox.