HIGH DYNAMIC RANGE VERSUS STANDARD DYNAMIC RANGE COMPRESSION EFFICIENCY

Ronan Boitard[⋆] *Mahsa T. Pourazad*[⋆†] *Panos Nasiopoulos*[⋆]

[⋆] University of British Columbia, Vancouver, Canada
[†]TELUS Communications Inc., Vancouver, Canada

ABSTRACT

High Dynamic Range (HDR) image and video technology aims at conveying the full range of perceptible shadow and highlight details with sufficient precision. HDR is regarded by many experts as the next evolution in digital media. However, industrial broadcasters have concerns regarding the bandwidth overhead that this new technology entails. While many consider that broadcasting HDR content would increase bandwidth requirements by around 20%, this number is based on studies where, in addition to the SDR main stream, HDR-related side information is conveyed. A recent subjective evaluation reported that encoding HDR video content in a single layer might require less bandwidth than its associated SDR version. Similar results were discussed in the MPEG ad-hoc group on High Dynamic Range and Wide Color Gamut. In this article, we explain how having more information can result in lower bandwidth requirements. To this end, we describe several limitations of the human vision system that, when exploited, optimize the HDR distribution pipeline for a human observer. Our theoretical assumption about the higher efficiency of HDR is backed up by a statistical analysis of pixel distribution in real images. The Spatial Index objective metric also reconfirms our assumption.

***Index Terms*—** High Dynamic Range, Video Coding, Perceptual Encoding, HDR10

I. INTRODUCTION

High Dynamic Range (HDR) image and video technology can capture, represent and display much more light and color information than the Standard Dynamic Range (SDR) technology, and thus greatly enhance the viewers Quality of Experience (QoE) [1]. This technology is regarded as the next evolution in digital media. However, broadcasting companies have concerns regarding the increased bit-rates that sending HDR information would entail.

In recent consumer electronic tradeshows such as CES, NAB and IBC, it was claimed that encoding HDR content requires around 20% bit-rate overhead compared to its SDR counterpart. These reports are based on studies where, in addition to the SDR main stream, HDR-related side information is conveyed [2]. When considering a single layer scheme, be it HDR or SDR, a recent study reported that encoding HDR content natively, for example using HDR10 [3], results in lower bit-rates than the SDR version [4]. These results are intuitively contradictory, more information using less bandwidth. In this article, we explain in detail why a pipeline such as HDR10 can be more efficient than broadcasting the SDR version.

The rest of this paper is organized as follows. Section II introduces the concept of color pixels encoding. Section III outlines the difference between the traditional SDR and the emerging HDR pipelines. Section IV provides a statistical analysis of both HDR and SDR content to illustrate that more visual information does not mean higher entropy signal. Finally, Section V concludes this article.

II. COLOR PIXEL ENCODING

Color pixels are traditionally represented using integer code values whose distribution is optimized for human observers. Thus, images are encoded in a perceptually linear domain with two goals in mind:

- removing information that would be invisible after decoding (visual noise),
- optimizing the limited bit-depth to minimize the visual loss due to quantization.

In the case of SDR, perceptual encoding is performed by an inverse electro optic transfer function called gamma encoding (ITU Recommendation BT.1886) [5], whose design was based on psychophysical studies for luminance values ranging from 1 to 100 cd/m^2 (capacity of the CRT display technology), representing a contrast of 1:100 [6]. Although existing monitors achieve lower black and higher white levels than this range, the BT.1886 recommendation has been for a long time the only available standard to encode SDR content. Note that at the acquisition stage, SDR cameras have a proprietary perceptual encoding function that differs from the BT.1886 [5]. This curve usually corresponds to a sigmoid-shape (in camera jargon, a tail for shadows and a knee for highlights) and aims at preserving some information in highlights and shadows [7].

HDR imagery matches and can even surpass the human vision system limitations [1]. In the case of HDR, pixels correspond to absolute light intensity (measured in cd/m^2) represented by floating point values that can cover a wide

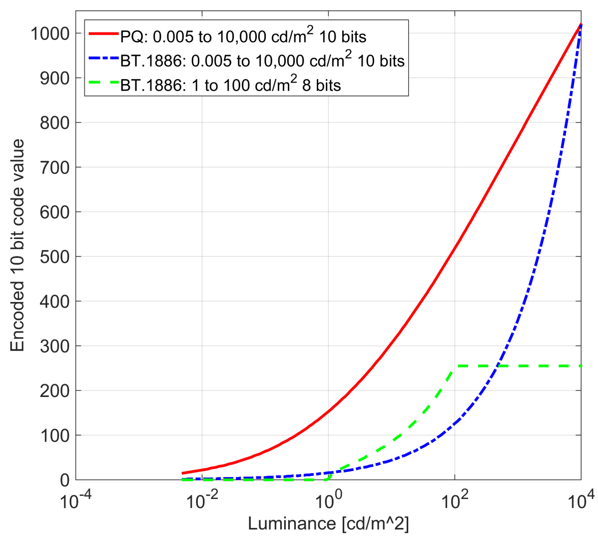

Fig. 1: Luminance values encoded with the Perceptual Quantizer (PQ) [9] and Gamma encoding (BT.1886) [5].

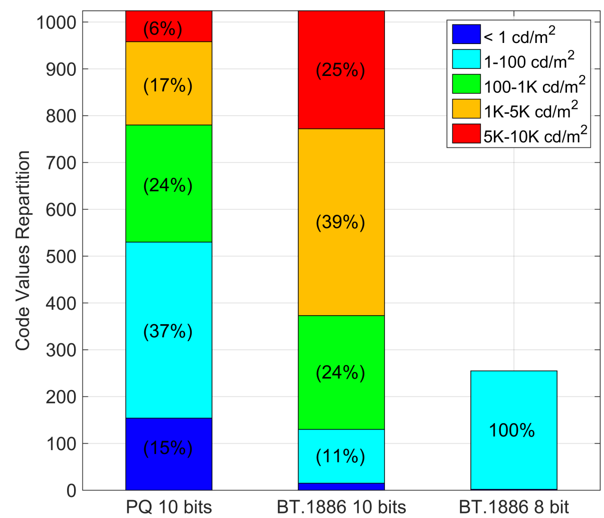

Fig. 2: Distribution of code values depending on the chosen perceptual encoding.

range of luminance (much larger than the capacity of a CRT display). Since image and video processing is devised to process integer code values, some encoding is needed to convert HDR floating point values to the integer format. However, using the BT.1886 function is not ideal since it was designed for luminance values ranging from 1 to 100 cd/m^2. To address this issue, SMPTE recently standardized another perceptual encoding [8], known as Perceptual Quantizer (PQ) [9]. PQ was derived using the peak sensitivities of a Contrast Sensitivity Function (CSF) model [10] with the objective of predicting the width of the quantization step that would not introduce any visible artifacts at different luminance levels. PQ has shown to be the most efficient encoding approach, requiring no more than 11 bits to represent gray patches without any visual loss [11] (natural images are reported to require no more than 10 bits [9]).

Fig. 1 shows 10 bit encoded luminance code values using PQ and BT.1886 for luminance values ranging from 0.005 to 10,000 cd/m^2. The BT.1886 as designed originally on 8 bits for luminance values ranging from 1 to 100 cd/m^2 is also plotted. Fig. 2 plots, for all three encodings, the distribution of these code values in 5 approximate luminance ranges: shadows (0 to 1 cd/m^2), typical (1 to 100 cd/m^2), white (100 to 1,000 cd/m^2), highlights (1,000 to 5,000 cd/m^2) and super highlights (5,000 to 10,000 cd/m^2) [12]. Using the BT.1886 with 10 bits over the entire luminance range results in 88% of the code values being assigned to luminance values ranging from 100 to 10,000 cd/m^2. The typical SDR range in this case is then left with only 11% of the code values (112 code values against 255 in SDR). Note also that all the information in dark areas (≤ 1 cd/m^2) is encoded using only 10 code values. In contrast, PQ assigns 52% of its range for luminance values ranging from 0.005 to 100, using slightly more code values than what the traditional BT.1886 uses for the 1 to 100 cd/m^2 range, while keeping 48% of its code words for the high luminance range.

Optimizing the limited bit-depth is also achieved by restricting the scope of color values (color gamut) that can be represented. In high-definition television technology, the scope of supported color values is described by the ITU-R Recommendation BT.709 [13]. With the introduction of the ultra high-definition television, a larger color gamut was needed, which is described by the ITU-R Recommendation BT.2020 [14]. BT.2020 uses 10 or 12 bits to cover 75.8% of the full visible gamut. For comparison purposes, the BT.709 uses 8 bits to cover approximately 35.9% of this gamut.

When compressing video content, both recommendations describe the luminance/chrominance decomposition to convert R'G'B' values into Y'C_bC_r values. A study reported that 10 bits are sufficient to represent the BT.2020 gamut using the Y'C_bC_r representation [11]. The Y'C_bC_r decomposition includes a luminance dependent scaling so that only a small portion of the bit-depth is used at low luminance intensity while the full range is used at high luminance intensity. Fig. 3 illustrates this scaling. Note how, at the black level (left), every R'G'B' combination is gathered around the white point. This means that when quantized on 10 bits, chroma values will be in the range 480-638, say 17% of the available code values. At the white level, most of the chroma plane is used. The explanation for this scaling and its impact on compression efficiency can be found in [15].

III. HDR/SDR ACQUISITION, REPRESENTATION AND REPRODUCTION

As mentioned above, high dynamic range images require a different perceptual encoding than BT.1886. However, most images or videos may not possess such a wide range of

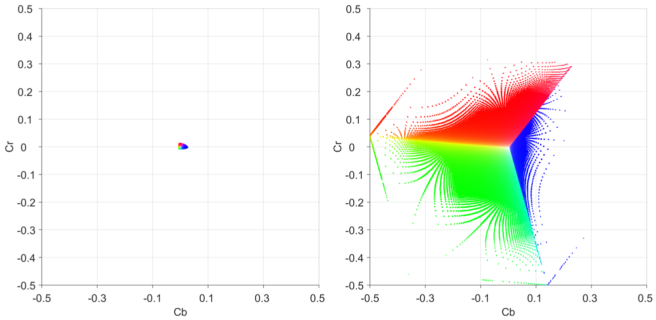

Fig. 3: Projection on the Y'$C_b$$C_r$ plane of all R'G'B'-10 bits combinations of the BT.2020 color space. Left: reference black luminance intensity (Y'=0). Right: reference white luminance intensity (Y'=1).

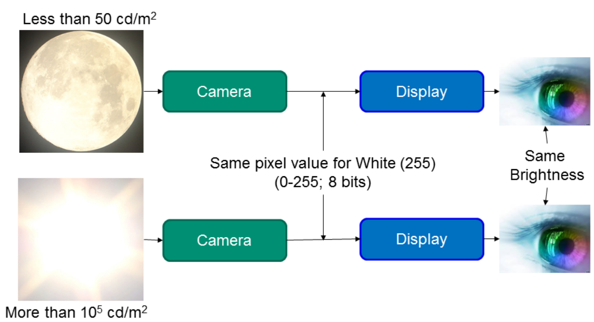

Fig. 4: Simplified SDR pipeline.

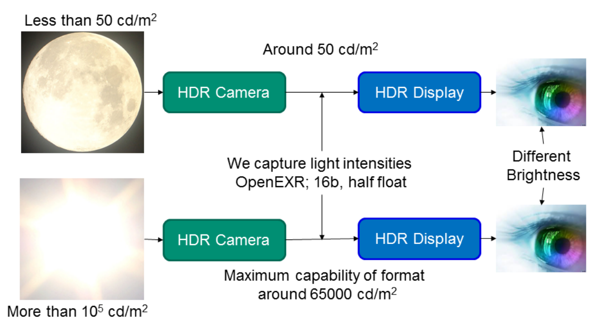

Fig. 5: Simplified HDR pipeline.

luminance, stretching from 0.005 to 10,000 cd/m^2. That is where the main difference between HDR and SDR lies. Given two scenes with different ambient illumination (night and daylight), the SDR pipeline will try to optimize the distribution of pixels to its limited dynamic range to preserve as much information as possible. On the other hand, the HDR pipeline does not need to drop any information as it can keep all the ambient illumination information between those two scenes.

This effect is illustrated in Figs. 4 and 5. For the SDR pipeline, the moon, which corresponds to a low luminance level, is mapped to the same white value (255 with 8 bits) as the sun, thus both objects end up having similar

	HDR 10 bits		SDR 8 bits	
Sequence	SI	I-Frame	SI	I-Frame
FireEater2	30	176 K bits	35	358 K bits
Market3	196	2.162 M bits	289	9.884 M bits

Table I: Spatial Index (SI) [16] and bitrate necessary to compress HDR 10 bits and SDR 8 bits images using Intra-Coding.

brightness when shown on a SDR display (Fig. 4). In the HDR pipeline, the two different ambient illuminations can be represented simultaneously, allowing us to preserve the difference of brightness at the display stage (Fig. 5). In a nutshell, the difference is that HDR pixels correspond to absolute light intensity (directly related to cd/m^2), while SDR pixels are integer code values whose light intensity is relative to the camera exposure (at the acquisition stage) and the capabilities of the used display (at the reproduction stage). This difference between absolute and relative light is the reason for the higher HDR compression efficiency.

IV. IMAGE STATISTICS AND COMPRESSION EFFICIENCY

In HDR images, by applying perceptual encoding on absolute luminance values, we ensure that visual noise is minimal, thus resulting in using an optimal amount of code values to represent all the visual information. However, in the SDR case, relative values do not correspond to the same luminance level at the acquisition and reproduction stage, and that prevents perceptual encoding from being attuned to the human visual system response to light. For this reason, perceptual encoding was based on a typical display model (average capabilities of CRT displays).

We illustrate this difference in interpreting light by plotting the luminance histogram of two scenes with different ambient illumination in Figs. 6 and 7. Fig. 6 is a night scene with a dynamic range too wide to be captured with an SDR camera, thus luminance values above 300 cd/m^2 are clipped. In addition, the impact of coarse quantization at low luminance levels can be observed, as the distribution of consecutive intervals (bins) is altered in the SDR 8 bits content (fails to represent values of the original image shown in black). Thus, the SDR version has less visual information, especially in highlights, compared to the HDR 10 bits (shown with blue circles). Despite this loss of information, the Spatial Index (SI) [16], a metric used to assess the coding complexity of a scene, of the SDR content is higher than that of the HDR content (see Table I). Furthermore, when coding the associated sequence in Intra-Only mode using HEVC, the amount of bits required to code this image is doubled even though the SDR version uses only 8 bits rather than 10.

Regarding the daylight scene shown in Fig. 7, the SDR version can almost cover the entire range, only clipping some highlights and losing details in shadows due to quantization.

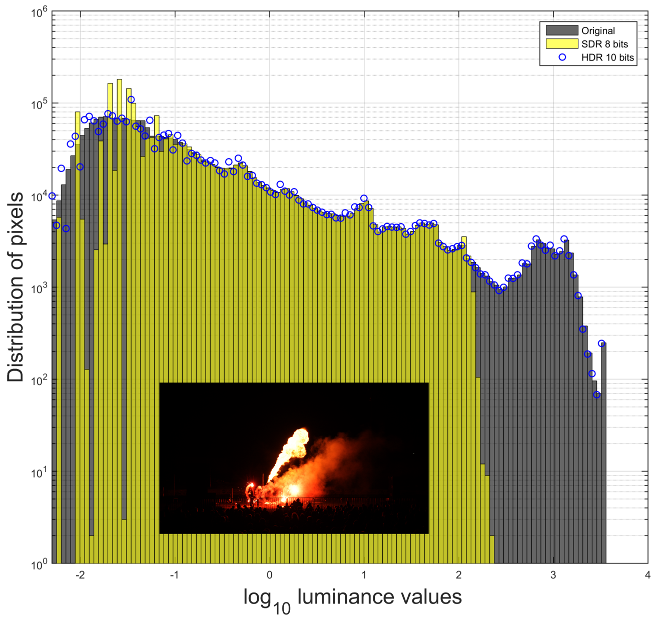

Fig. 6: Histogram comparison of original luminance, HDR 10 bits and SDR 8 bits for a night scene (FireEater2 [17]).

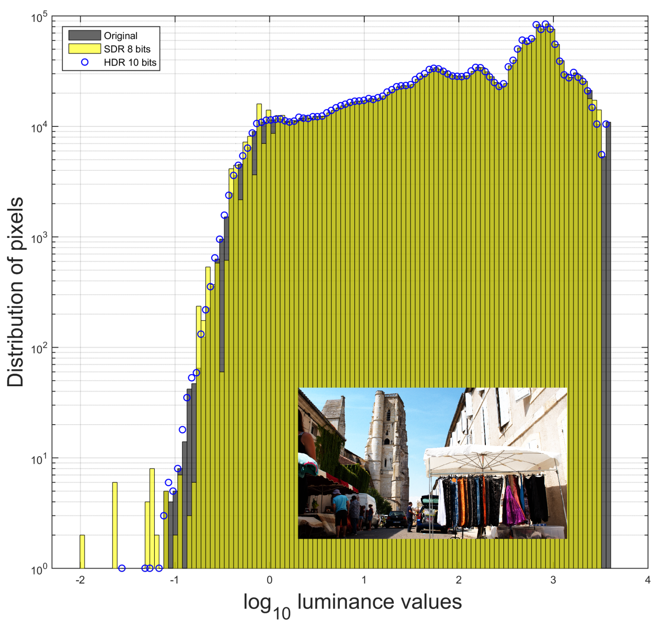

Fig. 7: Histogram comparison of original luminance, HDR 10 bits and SDR 8 bits for a daylight scene (Market3 [17]).

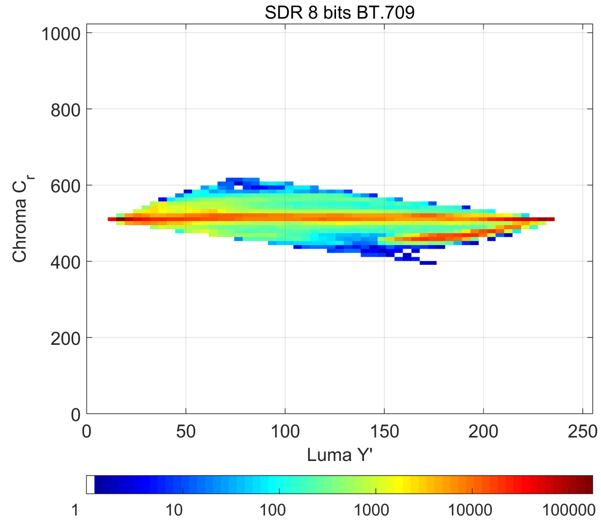

Fig. 8: Luma Y' versus chroma C_r using the SDR 8 bits BT.709 representation [13] for the first frame of the Market3 sequence [17]. C_r values have been centered around 512 for easier comparison.

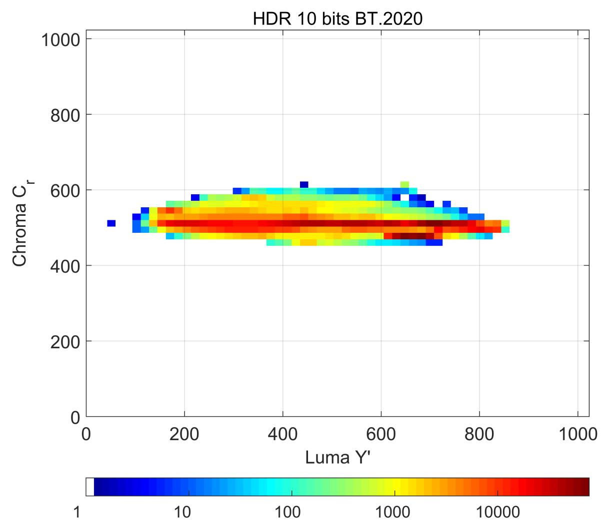

Fig. 9: Luma Y' versus chroma C_r using the HDR 10 bits BT.2020 representation [14] for the first frame of the Market3 sequence [17].

However, similarly to the night scene, the SI is much higher in the SDR case compared to the HDR, and the required bitrate for Intra coding is 5 times more for SDR 8 bits compared to HDR 10 bits. Indeed, although HDR content is encoded on 10 bits, not all code values are used since the PQ mapping is fixed and independent of the content.

Another aspect of importance is the representation of color. Figs. 8 and 9 plot the distribution of C_r pixels per luma value for both the SDR 8 bits and HDR 10 bits representation. This figure shows that SDR 8 bits with the BT.709 representation requires a slightly wider range to represent color information compared to the HDR 10 bits using the BT.2020. Since the luminance/chrominance representation is the same in both cases (Y'$C_b$$C_r$), using 10 bits instead of 8 should results in more code values being used. However, HDR is recommended to be used only with the BT.2020 color gamut or higher. Using a wider color gamut allows to compensate for the additional 2 bits (from 8 to 10), resulting in a similar distribution of chroma code values which should result in a similar compression efficiency. Note that the SDR 8 bits chroma code values, in Fig. 8, have been centered around 512 for easier comparison but are still quantized on 8 bits ([385, 640] instead of [0, 255]).

By using fewer code values than the SDR in some ranges, PQ removes visual noise (high frequency signal) before the compression stage, thus preventing video coding techniques preserving this information. Regarding the color representation, the use of a different container (BT.2020 instead of BT.709) allows a smooth transition to 10 bits for representing the same color values. Note that the results reported in [4] show that using HDR10 encoding [3] not only requires less bandwidth compared to SDR 8 or 10 bits, but it also achieves a higher mean opinion score in subjective experiments.

V. CONCLUSION

In this article, we described the reasons why distributing HDR content can be more efficient than its SDR counterpart in terms of compression efficiency. We illustrated that traditional perceptual encoding such as the BT.1886 cannot accurately represent higher dynamic range since its quantization is too coarse in dark areas. For the dynamic range currently considered in digital video as HDR (0.005 to 10,000 cd/m^2), the proposed HDR10 pipeline showed to provide good results [4].

We also showed that having less information and a smaller bit-depth does not result in higher compression efficiency. For the past 10 years, the focus has been on improving codec standards and only thanks to the emergence of HDR imagery this trend has shifted toward improving color pixel representation. Statistical analysis of content encoded using the Perceptual Quantizer (PQ) is a good indication that further improvements in compression efficiency can be achieved by choosing the right representation (perceptual encoding and color representation) for the video signal.

VI. REFERENCES

[1] R. Boitard, M. T. Pourazad, P. Nasiopoulos, and J. Slevinsky, "Demystifying high-dynamic-range technology: A new evolution in digital media.," *IEEE Consumer Electronics Magazine*, vol. 4, no. 4, pp. 72–86, oct 2015.

[2] D. Touzé, S. Lasserre, Y. Olivier, R. Boitard, and E. François, "Hdr video coding based on local ldr quantization," in *HDRi2014 - Second International Conference and SME Workshop on HDR imaging (2014),*, 2014.

[3] D. Le Gall et al., "High dynamic range with hevc main10," in *ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11*, Warsaw, PL, 2015.

[4] M. Azimi, R. Boitard, B. Oztas, S. Ploumis, H. R. Tohidypour, M. Pouzarad, and P. Nasiopoulos, "Compression efficiency of hdr / ldr content," in *7th International Workshop on Quality of Multimedia Experience (QoMEX)*, 2015.

[5] International Telecommunication Union, "Reference electro-optical transfer function for flat panel displays used in hdtv studio production bt series broadcasting service," in *Recommendation ITU-R BT.1886*, 2011.

[6] T. Borer, "Bbc research & development white paper whp283," Tech. Rep. July, BBC Research Department Report, 2014.

[7] J. Petit and R. K. Mantiuk, "Assessment of video tone-mapping : Are cameras s-shaped tone-curves good enough?," *Journal of Visual Communication and Image Representation*, vol. 24, pp. 1020–1030, 2013.

[8] Society of Motion Picture & Television Engineers, "High dynamic range electro- optical transfer function of mastering reference displays table," in *SMPTE ST 2084*. 2014, pp. 1–14, SMPTE.

[9] S. Miller, M. Nezamabadi, and S. Daly, "Perceptual signal coding for more efficient usage of bit codes," *SMPTE Motion Imaging Journal*, vol. 122, no. 4, pp. 52–59, May 2013.

[10] P. G. J. Barten, "Physical model for the contrast sensitivity of the human eye," in *SPIE 1666, Human Vision, Visual Processing, and Digital Display III,*, B. E. Rogowitz, Ed., Aug. 1992, pp. 57–72.

[11] R. Boitard, R. K. Mantiuk, and T. Pouli, "Evaluation of color encodings for high dynamic range pixels," in *Proc. SPIE 9394, Human Vision and Electronic Imaging XX*, 2015, pp. 1–9.

[12] M. Nezamabadi, S. Miller, S. Daly, and R. Atkins, "Color signal encoding for high dynamic range and wide color gamut based on human perception," in *Proc. SPIE 9015, Color Imaging XIX*, 2014, pp. 1–12.

[13] International Telecommunication Union, "Parameter values for the hdtv standards for production and international programme exchange," in *Recommendation ITU-R BT.709-3*, 1998.

[14] International Telecommunication Union, "Parameter values for ultra-high definition television systems for production and international programme exchange," in *Recommendation ITU-R BT.2020*, 2012.

[15] R. Boitard, M. T. Pourazad, and P. Nasiopoulos, "Chroma scaling for high dynamic range video compression," in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

[16] H. Yu and S. Winkler, "Image complexity and spatial information," *2013 5th International Workshop on Quality of Multimedia Experience, QoMEX 2013 - Proceedings*, pp. 12–17, 2013.

[17] D. Touzé and E. Francois, "Description of new version of hdr class a and a' sequences," in *ISO/IEC JTC1/SC29/WG11 MPEG2014/M35477*, Geneva, Switzerland, 2015, IEEE.