Perception-Based Histogram Equalization for Tone Mapping Applications

Stelios Ploumis[1], Ronan Boitard [1], Mahsa T. Pourazad [1, 2], Panos Nasiopoulos [1]

[1]*University of British Columbia, Vancouver, Canada*
[2] *TELUS Communications Inc., Vancouver, Canada*

Abstract*—Due to the ever increasing commercial availability of High Dynamic Range (HDR) content and displays, backward compatibility of HDR content with Standard Dynamic Range displays is currently a topic of high importance. Over the years, a significant amount of Tone Mapping Operators (TMOs) have been proposed to adapt HDR content to the restricted capabilities of SDR displays. Among them, the Histogram Equalization (HE) is considered to provide good results for a wide set of images. However, the naïve application of HE results either in banding artifacts or noise amplification when the HDR image has large unified areas (i.e. sky). In order to differentiate relevant information from noise in a uniform background, or in dark areas, the authors proposed a ceiling function. Their method results in noise-free but dim images. In this paper we propose a novel ceiling function which is based on the Perceptual Quantizer (PQ) function. Our method uses as threshold the number of code-words that PQ assigns on a luminance range in the original HDR image and the corresponding number of code-words in the resulting SDR image. We limit the number of code-words on SDR to be equal or less than the HDR. The saved code-words during the ceiling operation are redistributed to increase the contrast as well as the brightness of the final image. Results shows that provided SDR images are noise-free and brighter than the one obtained with prior HE operators. Finally since the proposed method is a Global TMO, it is thereby of low complexity and suitable for real time applications.

Keywords: HDR, Tone Mapping, Histogram Equalization, Perceptual Encoding

INTRODUCTION

High Dynamic Range (HDR) technology has recently gained a lot of interest in both academia and industry. HDR can capture and reproduce a wide range of luminance values, very close to what the human eyes can perceive in real world scene [1]. Such a technology contrasts with Standard Dynamic Range (SDR) technology which is limited to a small portion of this range (0.1-100 cd/m^2). Due to the significant improvement in the Quality of Experience (QoE), HDR is expected to replace SDR in the near future. However, during the transition from SDR to HDR, most displays will only have SDR capabilities. Thus, the ability to display HDR content on legacy SDR displays is highly desirable. This may be achieved through a process known as tone mapping, which maps the large amount of HDR information to the limited SDR range. Considering that legacy displays have a much more limited dynamic range than HDR ones, it is impossible for all the tonal levels of HDR content to be reproduced on SDR displays. Naïve linear mapping from HDR to SDR do not preserve image contrast, brightness and details. Thus tone mapping aids in sophisticated decision-making regarding which tonal levels should be preserved and which dismissed. Tone mapping results in SDR content with much less information and lower visual quality than the original HDR, yet still offering a better viewing experience than similar content captured and displayed in SDR (see Figure 1).

Over the years, several Tone Mapping Operators (TMOs) have been proposed. They can be distinguished into two categories: local and global operators. Local operators [2, 3, 4] tone map each pixel based on its spatial neighborhood. These operators usually deal well with edges and preserve most of the visual information. Their main drawbacks are the appearance of halo artifacts around edges and their high computational cost, which makes them unlikely candidates for real time applications. Global operators [5, 6, 7] use statistics of an HDR image to compute a monotonously increasing tone map curve for the whole image. While these operators have low computational cost, they usually fail to deliver a desirable SDR image when dealing with dark scenes or highly contrasted HDR images.

In this work, we mainly focus on global TMOs because of their ability to be used in real time applications, like TV broadcasting and video streaming. Among the state-of-the-art global TMOs, Ward et al. operator [5] is known to well preserve the contrast of original HDR images while delivering natural looking results [8, 9, 10]. This TMO is based on Histogram Equalization (HE), which efficiently redistributes tonal levels to

Figure 1. A comparison between a SDR image (top) and a HDR tone mapped image (bottom). [Images courtesy of Zicong (Jack) Mai.]

preserve information where the pixel density is the highest in the intensity domain. This algorithm works well when applied to most HDR images since it preserves contrast in large areas with low luminance variations. However, this algorithm cannot detect when a luminance range is too small to contain details worth preserving, for example large background areas such as sky or night scenes. In those cases, it will emphasize noise rather than visual information.

In this paper, we propose a Histogram Equalization (HE) based TMO that can differentiate relevant information from noise in a uniform background or in dark areas. Our proposed scheme first compares the number of distinguishable tonal levels in the HDR luminance image with those in the candidate tone mapped SDR luminance image. Wherever this number of tonal levels is higher in the SDR luminance than in the HDR one, we modify the mapping curve. Our technique then redistributes saved tonal levels to increase details in highlights and/or to achieve SDR images with higher overall contrast.

The rest of this paper is structured as follows. Section II provides an overview of the Histogram Equalization (HE) and Perceptual Quantizer (PQ). Section III describes our proposed tone mapping method. Section IV presents some results while finally Section V concludes the paper.

Overview of Historgam Equalization and Perceptual Quantizer

Perceptual Quantizer (PQ)

The Perceptual Quantizer (PQ) is an inverse Electro Optical Transfer Function (EOTF), which optimizes the distribution of light intensities with respect to the human visual system properties. Since the human eye does not perceive and interpret the light in a linear way, an inverse EOTF aims at transforming physically linear values into perceptually linear values. A transform domain is defined as perceptually linear if a variation of intensity, at any brightness level, corresponds to the same variation in perception.

In digital imagery, pixels are traditionally represented using integer code-words whose distribution is optimized for human observer. Perceptual linearity is very useful to encode the highest amount of information on the lowest amount of code-words. The current dominant inverse EOFT is the BT.1886 [12] or Gamma encoding function. BT.1886 was designed for the luminance values ranging from 0.1 to 100 cd/m^2. However HDR technology can represent a much larger luminance range and commercial TVs are expected to reproduce luminance intensities ranging from 0.005 to 4,000 cd/m^2 in the near future. That is why SMPTE standardized, in 2012, the SMPTE ST 2084 inverse EOTF, also known as Perceptual Quantizer (PQ). PQ was derived using the peak sensitivities values of Barten's Contrast Sensitivity Function (CSF) [13]. The Barten's CSF models the human vision contrast detection threshold for different spatial frequencies, background luminance and viewing angle. Figure 2 plots the Barten Ramp (dashed line) which delimitates invisible contrast steps (green area) from visible ones (pink area). PQ was designed to be as close as possible to the Barten Ramp using the smallest possible bit-depth. Figure 2 also plots the PQ using 12 bits, BT.1886 15 bits and log 13 bits encoding. Observe that PQ, which uses the lowest bit-depth, remains almost always at the same distance from the Barten Ramp for the entire range of 0.005 to 10,000 cd/m^2.This proximity to the Barten Ramp means that PQ, using 12 bits, assigns two successive code-words to two different luminance levels if and only if the difference between them is just below the just visual noticeable threshold. The BT.1886 using 15 bits fends off the Barten Ramp for luminance values higher than 10 cd/m^2. This means that bright areas will be

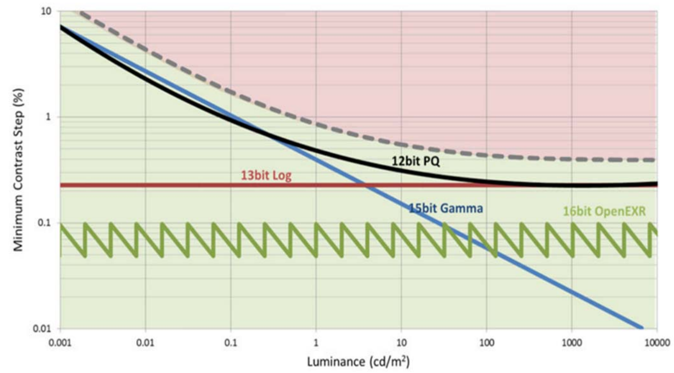
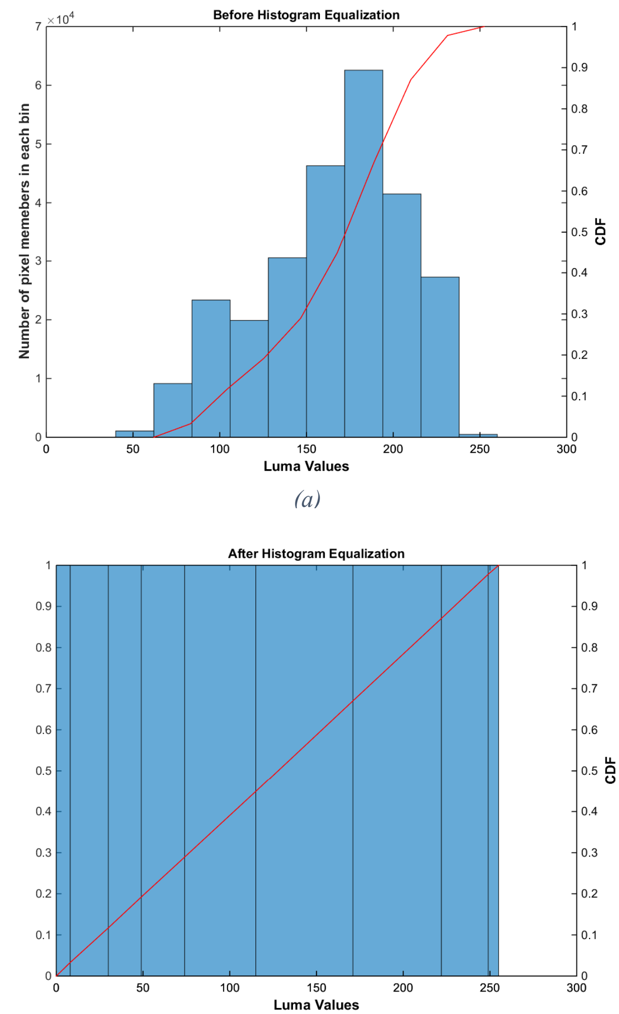


Figure 2. Barten CSF (pink area), PQ curve (black), BT.1886 curve (blue) log curve (red) [Image courtesy of Dolby Laboratories]



(b)

Figure 3. Before (a) and after (b) the application of HE on an image

oversampled, using more code-words than necessary or in other words encoding visual noise.

Histogram Equalization and Ceiling

Histogram Equalization (HE) is a widely used technique in image processing, aiming mainly at increasing the global contrast of an image. Usually in user-generated images, most of the pixel's intensities are gathered over a relatively small portion of the histogram, usually on its center (see Figure 3a). HE consists in spreading intensities over the available range, thus increasing the global contrast of the image. Figure 3 illustrates the application of an HE. The cumulative distribution function is first computed (red curve). This function is used to expand bins associated with many pixels and shrink bins with smaller pixel density (see Figure 3b). In a nutshell, Histogram Equalization efficiently redistributes tonal levels to preserve information where the pixels' density is the highest in the intensity domain. However, such an approach cannot differentiate relevant information from noise in a uniform background or in dark areas.

When Ward et al. [5] adapted the HE to tone mapping, they proposed to build the histogram of the luminance values in the log domain instead of the linear one since the response of the human vision system to luminance variations is approximately log linear when photoreceptors are fully responsive [14]. This can also be seen in Figure 2, where the 13 bit log encoding is almost parallel to Barten's Ramp for high luminance values. The cumulative distribution function is then used to map the HDR log luminance values to the new SDR range. The authors assume as destination SDR range the luminance values between 1 – 100 cd/m^2. Afterwards they transfer the values back to the luminance domain and normalize them to the range [0, 1] which are then gamma encoded using the BT.1886 and quantized on the targeted bit-depth. This TMO handles well most natural HDR images but struggles to provide pleasant reproduction for images with relatively large areas belonging to a small luminance range, such as a sky or a night scene. Indeed, in those cases, a large amount of pixels gathered in a small luminance range (which we will denote as a bin of the histogram) will be map to a large proportion of the destination range since the slope of the cumulative distribution function will be very steep. This results in exaggeration of contrast in spatial areas that belongs to those bins. Furthermore, only a relatively small proportion of the destination range will be left for the rest of the histogram bins to be map, resulting to significant loss of information. Finally, especially for scenes with large dark areas (i.e., a night scene), the exaggeration of contrast amplifies noise.

To overcome these issues Ward et al. [5] proposed two ceiling functions to prevent large bins to get assigned most of the available dynamic range. Such ceiling functions mainly aims at reducing the tonal levels that are assigned to the large bins described previously. In the first ceiling function, the authors proposed a linear ceiling which limits the maximum density of pixels in a single bin independently of the luminance levels that this bin represents. However as we described before, humans do not interpret light in a linear way, thus this ceiling function it results in noisy images when it deals with dark scenes. The second ceiling function is based on the Blackwell CSF [15]. In this method, the authors applied different limits on the bins' density depending on their luminance level. In their implementation, they relied only on five different luminance levels. This method manages to reduce visible noise from dark scenes but resultes in dimmer images.

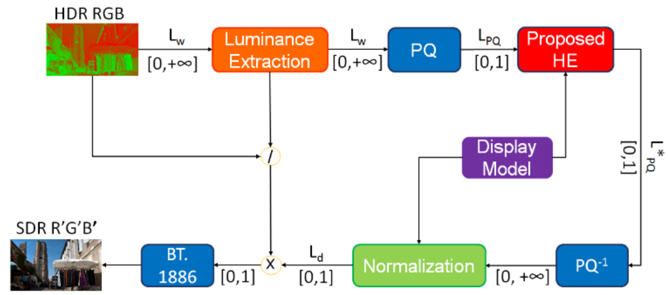


Figure 4. Flowchart of the proposed method

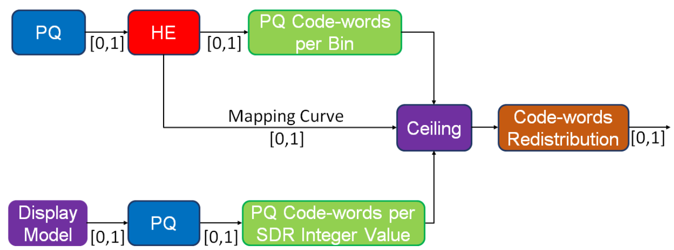


Figure 5. Proposed Histogram Equalization

PROPOSED TONE-MAPPING METHOD

Ward et al. [5] TMO performs well for most natural images but fails when HDR images are too dark or have large uniform areas. This is because their method does not redistribute code-words saved during the ceiling process. Furthermore, their method does not take into account the capabilities of the targeted display. With the variety of peak luminance and bit-depth that commercial displays can achieve, the number of different tonal levels achievable can greatly vary. In this paper, we propose an alternative ceiling function based on the PQ to constrain the histogram equalization. We also propose an algorithm to redistribute the saved code-words. Figure 4 illustrates the workflow of our method. First, we compute the histogram of the HDR luminance channel in the PQ domain. The PQ domain allows to overcome a problem inherent to the log space, that is to say its lack of perceptual uniformity for low luminance values. The size of histogram bins corresponds to the luminance range that code-words represent when quantizing a luminance range from 0 to 10,000 cd/m^2 on 10 bits. We use 10 instead of 12 bits because 12 bits corresponds to the most conservative case (maximum peak of contrast sensitivity). These low contrast would be masked by edges and color changes and as a matter of fact Kunkel et al. reported in [16] that 10 bits was sufficient for natural images. Furthermore, using 10 bits in the PQ space allows direct compatibility with the HDR10 format [17] which is the most likely format to be used by broadcasters.

Our method performs the HE, as detailed in [5], to map the HDR luminance values ranging from 0 to 10,000 cd/m2 to the targeting range (from 0 to 1). As described before, such a naïve application of the HE results in noise amplification or banding artifacts when dealing with dark or uniform images. Thus we propose to supervise the HE process using a ceiling function. The proposed ceiling function will ensure that the amount of visible information in the tone mapped image is not greater than in the HDR one. However, since HDR values correspond to physical light intensities while SDR values correspond to relative light intensities that depend on the used TV, we need to transform both these values to the same domain in order to compare them. For this purpose, we chose the PQ domain since it is currently the most perceptually uniform as explained in Figure 2 and Section II.

Converting HDR values to the PQ domain is straightforward by using the PQ transfer function. However, to convert SDR integer values to physical luminance values, one needs to assume the capabilities of the display on which the content will be reproduced. That is why, our ceiling function will take as input parameters the capabilities of the targeted display (black level, white level, EOTF and bit-depth). Using such a display model and the PQ function, we can convert SDR integer values to the expected light intensity emitted by the targeted display and then convert them to the PQ domain. We thus have two different luminance ranges projected in the PQ domain which will be used for the ceiling process: the HDR luminance values and the mapped SDR luminance values. To differentiate visible from invisible information, we also quantize the PQ domain using 10 bits.

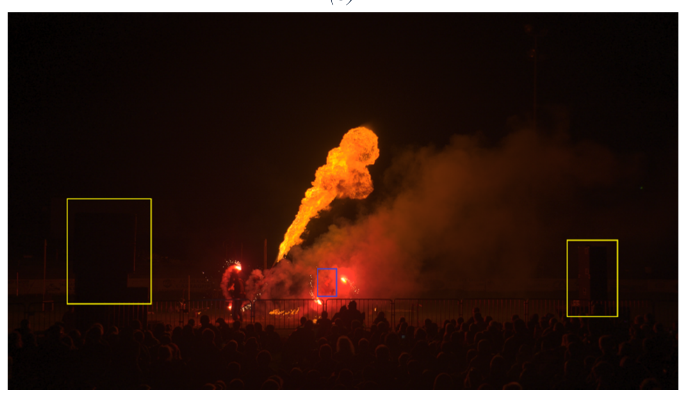
We chose PQ 10 bits since it approximately assigns two successive code-words if and only if the visual difference between two luminance levels is detectable by human eye. For our proposed tone mapping, it means that if a bin of the HDR histogram is represented using more than one code-word when mapped to the destination range, then more information is visible which may result in noise amplification or banding artifacts (denoted as Overload bin). On the other hand, if it is represented using only one code-word, then the same amount of visual information is reproduced (denoted as Full bin). Finally, if less than a code-word is used, then we have loss of visual information (denoted as Available bin). Such a loss is acceptable since tone mapping is about selecting which information should be preserved and which dismissed. In a nutshell, the goal of the ceiling is to detect which bins are overloaded.

The ceiling process starts by comparing independently the PQ code-words associated with each bin of the HDR histogram and the projection of its range in the mapped range. If the bin is in an Overload state, we reduce the slope of the mapping curve for this particular bin. The ceiling process is applied to all overloaded bins.

The next step is to redistribute the saved dynamic range obtained during the ceiling process. We propose to redistribute the saved dynamic range to the largest Available bins. When all existing Available bins are Full, we use the remaining available code-words to increase the slope at gathering of HDR bins with very low pixel density. This redistribution increases contrast between gatherings of pixels with similar luminance, in other word contrast between different object of the scene (global contrast). Furthermore we can also redistribute these code-words to the bins located at the beginning of the histogram in

(a)

(b)



(c)

Figure 6. Comparison between Ward et al. linear ceiling (a), perceptual ceiling (b) and the proposed method (c)

order to increase the global brightness of the image and thus preventing too dim images. The whole process of the proposed Histogram Equalization is detailed in Figure 5.

The last stages of our operator consist in perceptually decoding the mapped luminance in the PQ domain L_{PQ} using the inverse PQ and inverting the display model to obtain the mapped luminance L_d (Normalization step). Finally, RGB channels are scaled using original (L_w) and tone mapped (L_d) luminance channels and RGB channels are then perceptually encoded using the BT.1886 and quantized on the targeted bit-depth.

RESULTS

In this section we evaluate our proposed method by comparing it with the TMO presented by Ward et al. [5]. As we described before, the naïve use of the HE as a TMO results either in banding or/and noise amplification when the original HDR image has large unified areas (i.e. sky). Thus a ceiling method is needed to restrict the HE as well as to redistribute the tonal values.

Figure 6a, depicts the resulting SDR image using the linear ceiling presented by Ward et al. [5]. We observe that the image suffers from noise, which is explained by the fact that the same threshold is applied on the size of each bin regardless of its luminance level. The second ceiling method proposed by Ward et al. is perceptually based and efficiently remove noise as illustrated in Figure 6b. However the overall brightness of the image is dimmer and objects like the the speaker and the audience are hardly visible. Figure 6c presents the results obtained using the proposed method. Note how we manage to eliminate the noise while keeping most of the visual information such as the speakers or advertisement on the railings. Finally, details on the bright part of the image, like the flame, the smoke and the person in front of the flame are also preserved.

Furthermore we compare the two methods (perceptual Ward and the proposed one) using as input a compressed HDR image. Since our method aims at preventing large uniform area to be emphasized, it should also prevent common compression artifacts such as blocking to be emphasized. Figure 7a shows the result obtained using the perceptual ceiling method of Ward et al. [5]. The image suffers from blocking artifacts in the cloudy sky. Figure 7a1 zooms on the sky area to better illustrates the blocking. On contrast, our method handles better the compressed images, thus blocking artifacts are less visible (see Figures 7b, 7b1).

CONCLUSIONS AND FUTURE WORK

In this article, we proposed a method to broaden the range of images that HE TMOs can handle. Since previous approaches [5] struggled with large background areas such as sky or night scenes, we propose a perceptually-based ceiling function to prevent noise amplification and/or banding artifacts. To increase local and global contrast as well as global brightness, we also proposed an algorithm to redistribute saved dynamic range during the ceiling process. Finally, our method adapts to the targeted display capabilities and is thus also suitable for tone mapping HDR content to HDR displays with different peak luminance.

Results show that the proposed TMO efficiently reduces noise in dark scenes as well as preserves the overall brightness of the scene. Furthermore, it handles well the compression artifacts, an important feature for broadcast applications where the content is transmitted compressed. Finally, the complexity of the proposed TMO is relatively low, making it suitable for real time applications.

In the future, we plan to extend our algorithm to video sequences. To achieve that we need to ensure the temporal coherency between successive frames in order to prevent flickering [18].

(a)

(a1)

(b1)

(b)

Figure 7. Comparison on compressed image between Ward et al. perceptual ceiling (a), and it artifacts (a1) and the proposed method (b) and its artifacts (b1).

REFERENCES

[1] Boitard R, Pourazad MT, Nasiopoulos P, Slevinsky J. Demystifying High-Dynamic-Range Technology: A new evolution in digital media. Consumer Electronics Magazine, IEEE. 2015 Oct;4(4):72-86.

[2] Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J. (2002, July). Photographic tone reproduction for digital images. In ACM Transactions on Graphics (TOG) (Vol. 21, No. 3, pp. 267-276). ACM.

[3] Durand, F., & Dorsey, J. (2002, July). Fast bilateral filtering for the display of high-dynamic-range images. In ACM transactions on graphics (TOG) (Vol. 21, No. 3, pp. 257-266). ACM.

[4] Pattanaik S, Yee H. Adaptive gain control for high dynamic range image display. InProceedings of the 18th spring conference on Computer graphics 2002 Apr 24 (pp. 83-87). ACM.

[5] Larson GW, Rushmeier H, Piatko C. A visibility matching tone reproduction operator for high dynamic range scenes. Visualization and Computer Graphics, IEEE Transactions on. 1997 Oct;3(4):291-306.

[6] Pattanaik, Sumanta N., et al. "Time-dependent visual adaptation for fast realistic image display." Proceedings of the 27th annual conference on Computer graphics and interactive techniques. ACM Press/Addison-Wesley Publishing Co., 2000.

[7] Drago F, Myszkowski K, Annen T, Chiba N. Adaptive logarithmic mapping for displaying high contrast scenes. InComputer Graphics Forum 2003 Sep 1 (Vol. 22, No. 3, pp. 419-426). Blackwell Publishing, Inc.

[8] Čadík M, Wimmer M, Neumann L, Artusi A. Evaluation of HDR tone mapping methods using essential perceptual attributes. Computers & Graphics. 2008 Jun 30;32(3):330-49.

[9] Yoshida A, Blanz V, Myszkowski K, Seidel HP. Perceptual evaluation of tone mapping operators with real-world scenes. InElectronic Imaging 2005 2005 Mar 18 (pp. 192-203). International Society for Optics and Photonics.

[10] Ledda P, Chalmers A, Troscianko T, Seetzen H. Evaluation of tone mapping operators using a high dynamic range display. ACM Transactions on Graphics (TOG). 2005 Jul 1;24(3):640-8.

[11] Miller S, Nezamabadi M, Daly S. Perceptual signal coding for more efficient usage of bit codes. Motion Imaging Journal, SMPTE. 2013 May;122(4):52-9.

[12] Recommendation, I. T. U. R. B. T. (2011). Reference electro-optical transfer function for flat panel displays used in HDTV studio production.

[13] Barten PG. Contrast sensitivity of the human eye and its effects on image quality. SPIE press; 1999.

[14] Pattanaik SN, Ferwerda JA, Fairchild MD, Greenberg DP. A multiscale model of adaptation and spatial vision for realistic image display. InProceedings of the 25th annual conference on Computer graphics and interactive techniques 1998 Jul 24 (pp. 287-298). ACM.

[15] Blackwell HR. Contrast thresholds of the human eye. JOSA. 1946 Nov 1;36(11):624-43.

[16] Kunkel T, Ward G, Lee B, Daly S. HDR and wide gamut appearance-based color encoding and its quantification. InPicture Coding Symposium (PCS), 2013 2013 Dec 8 (pp. 357-360). IEEE..

[17] D. Le Gall, A. Tourapis, M. Raulet, W. Wan, D. Malas, Y. Syed, J. Parkins, J. Boyce, J. Helman, C. Fogg, B. Mandel, G. Reitmeier, A. Duenas, D. Rusanovskyy, M. Budagavi, E. Alshina, and T. Suzuki, “High Dynamic Range with HEVC Main10,” in ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, 2015.

[18] Boitard R, Cozot R, Thoreau D, Bouatouch K. Survey of temporal brightness artifacts in video tone mapping. In HDRi2014-Second International Conference and SME Workshop on HDR imaging 2014..