Optimized Tone Mapping of HDR Images via HVS Model-Based 2D Histogram Equalization

Nam Hoang Nguyen†, Tu Van Vo†, Younghoon Jeong‡, Youngsu Moon‡, and Chul Lee†
†Department of Computer Engineering, Pukyong National University, Busan, Korea
‡Samsung Electronics Co., Ltd., Suwon, Korea
E-mails: {namnguyen, tuvovan}@pukyong.ac.kr†, {yh1004.jeong, mys66}@samsung.com‡, chullee@pknu.ac.kr†

***Abstract*—We propose an optimized human visual system (HVS) response model-based tone-mapping algorithm to display high dynamic range (HDR) image content on standard dynamic range (SDR) displays. We first measure the HVS response differences using a 2D histogram when an HDR image is displayed on an HDR device and its tone-mapped image on an SDR device. Then, we formulate an optimization problem to minimize the difference. By efficiently solving the optimization problem with a closed-form solution, we obtain an optimal tone-mapping curve. Experiment results demonstrate that the proposed algorithm provides higher image quality than a conventional algorithm in the HDR10 standard.**

I. Introduction

The market for ultra-high-definition (UHD) televisions has been rapidly expanding due to UHD broadcasting and the recent expansion of high-definition content produced. UHD broadcasting provides more realistic visual experiences by extending image resolution, bit depth, and color gamut [1]. Also, UHD broadcasting supports high dynamic range (HDR) image content, presenting a wider dynamic range of luminance levels than conventional standard dynamic range (SDR) content [2]. Because an HDR image contains the actual scene colors instead of those to be displayed on devices, it can present a look similar to that experienced through the human visual system (HVS) [2].

The most common HDR transport and broadcasting format is HDR10 Media Profile (HDR10) [3], which is currently used in UHD broadcasting and online streaming services, *e.g.*, Amazon and Netflix. HDR10 image content can represent a luminance of up to 4,000 cd/m^2 (nits) using 10-bit color depth [1], thereby providing higher image quality with more color, brighter highlights, and improved details. However, an SDR display can present contents with lower dynamic range than that defined in HDR10. Moreover, because the peak luminance represented by each device is different, devices that support HDR10 contents display them differently. Therefore, when HDR10 content is displayed on an SDR device or when the maximum luminance values of HDR content and the display are different, a process of converting the dynamic range of the input HDR content, called tone mapping, is required.

This work was supported in part by Samsung Electronics and in part by the National Research Foundation of Korea grant funded by the Korean Government (MSIT) (No. NRF-2016R1C1B2010319).

Many tone-mapping algorithms have been proposed, and they can be classified into global or local algorithms according to their methods of deriving mapping functions [2]. A global tone-mapping algorithm uses a single mapping function to convert all pixels in an entire image. For example, Drago *et al.* [4] employed a logarithmic function to adaptively reduce the dynamic range imitating the visual response of HVS. Gommelet *et al.* [5] proposed a global tone-mapping algorithm by formulating an optimization problem to minimize the spatial gradient-based distortion between an HDR image and its tone-mapped image. Mantiuk *et al.* [6] developed another global tone-mapping algorithm that minimizes the perceptual contrast distortion, given the display characteristics. Whereas global tone-mapping algorithms are computationally efficient in general, they may fail to preserve local details in an input HDR image, especially in regions of high contrast.

By comparison, local tone-mapping algorithms derive the mapping for each pixel by considering its local neighborhood. The zone system [7] and the retinex theory [8] have been employed to develop local tone-mapping algorithms that preserve details in the input HDR image. Recently, in [9], a gradient domain guided image filter was employed by incorporating an edge-aware constrain to better preserve local details. In [10], Ma *et al.* developed a local tone-mapping algorithm that iteratively updates the tone-mapped image to improve its structural fidelity and naturalness. Local tone-mapping algorithms provide tone-mapped results with fine details but often require high computational complexity.

Although conventional tone-mapping algorithms have attempted to provide high-quality results [4]–[10], little effort has been made to consider the characteristics of display devices on which the tone-mapped images are presented. The display-adaptive tone mapping in [6] is computationally inefficient because the optimization problem is solved iteratively. To address these problems, in this work, we propose an efficient HVS model-based tone-mapping algorithm that preserves the perceptual quality of the tone-mapped image. Specifically, we first formulate an optimization problem that maximizes the perceptual responses between an input HDR image and its tone-mapped image using the 2D histogram. Then, we obtain an optimal tone mapping curve (TMC) by efficiently solving the optimization problem with a closed-form solution. Experimental results demonstrate that the proposed tone-mapping algorithm preserves the perceptual similarities between the input HDR image and the tone-mapped image on real displays.

The rest of this paper is organized as follows. Section II briefly reviews related work. Section III describes the proposed tone-mapping algorithm. Section IV provides experimental results. Finally, Section V concludes the paper.

II. Related Work

In this section, we briefly review related works, upon which the proposed algorithm is based.

A. 2D Histogram Equalization

We employ the 2D histogram technique [11]–[13] to consider the local details in an HDR image in the HDR10 format. We are given an input image $\mathbf{X} = \{x(i,j) | 1 \leq i \leq H, 1 \leq j \leq W\}$ of size $H \times W$, where $x(i,j) \in [1, K]$. Let $h_x(m,n)$ represent the number of pairs of neighboring pixels with value m and n in image $\mathbf{X}$. Then, the 2D histogram of $\mathbf{X}$ is defined as

$$\mathbf{H}_x = \{h_x(m,n) | 1 \leq m \leq K, 1 \leq n \leq K\}. \tag{1}$$

The contrast of the image is enhanced using the histogram equalization. Specifically, an optimization problem is formulated to minimize the distance between the output 2D histogram $\mathbf{H}$ and the uniformly distributed 2D histogram $\mathbf{H}_u$, given by

$$\mathbf{H}_t = \arg\min_{\mathbf{H}} \|\mathbf{H} - \mathbf{H}_u\|_F^2 \tag{2}$$

where $\|\mathbf{Y}\|_F = \left(\sum_{ij} |Y_{ij}|^2\right)^{1/2}$ denotes the Frobenius norm of a matrix. Then, from the optimal 2D histogram $\mathbf{H}_t$ in (2), the transformation function that maps input pixel intensities to output intensities is derived using the histogram matching scheme [11].

B. HDR10 Standards

The major improvement in quality of videos is via the use of HDR. The HDR10 standard has been recently adopted [3] to facilitate transport and broadcasting of HDR videos. HDR10 supports a luminance of up to 4,000 nits using 10-bit color depth with a nonlinear transfer function for perceptually unnoticeable quantization [14], [15], which is called electro-optical transfer function (EOTF), and the wide-gamut BT.2020 color space [1]. It also uses static metadata to send color calibration data of the mastering display in the encoded stream [16].

C. HVS Response Model

We measure the perceptual response of the HVS to an image luminance using the threshold-versus-intensity (TVI) function, defined as the visibility threshold luminance ΔL at a background luminance L, given by $\Delta L = \text{TVI}(L)$. The response $r(L)$ is scaled in just noticeable difference (JND) units and specifies the HVS response under different luminance levels [17]. In practice, given the lowest luminance value under consideration L_0, we compute the luminance value $L_1 = L_0 + \text{TVI}(L_0)$, where the response is $r(L_1) = 1$. Then, the perceptual response to a luminance value can be obtained by the recursive procedure, given by

$$L_i = L_{i-1} + \text{TVI}(L_{i-1}) \text{ and } r(L_i) = i. \tag{3}$$

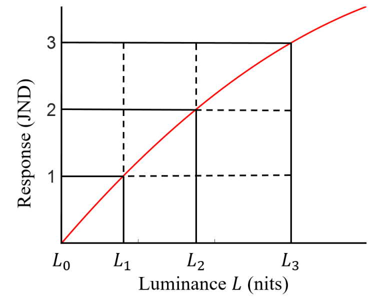


Fig. 1. Illustration of the luminance to HVS response mapping function $r(L)$.

Fig. 1 illustrates an example HVS response curve. In this work, we employ the TVI function developed in [18].

III. Proposed Algorithm

A. HVS Response Preserving Optimization

We develop a tone-mapping algorithm that preserves the perceptual responses between the input HDR image in the HDR10 format and its tone-mapped image displayed on a low-luminance display. To this end, we first quantify the response differences between two luminance levels, recognized by the HVS, of an image using the EOTF [14] and the HVS response model in (3). Specifically, we define the HVS response differences $R_i(m,n)$ and $R_o(m,n)$ between two luminance values on the reference and target displays, respectively, for two pixel values m and n.

Because the HVS is more sensitive to the luminance difference between neighboring pixels than the absolute luminance values [19], [20], we only consider luminance differences between neighboring pixels in this work. Let $D_i(x,y)$ and $D_o(x,y)$ denote the HVS response differences between two pixels at locations x and $y \in \mathcal{N}(x)$, where $\mathcal{N}(x)$ is a set of neighboring pixels of x, in the HDR and its tone-mapped images, respectively. Then, we minimize the total response differences for all pixels in the image as

$$\sum_{x,y \in \mathcal{N}(x)} (D_i(x,y) - D_o(x,y))^2. \tag{4}$$

We solve the problem in (4) using the 2D histogram equalization technique. Specifically, let $\mathbf{H}_i$ and $\mathbf{H}_o$ denote the 2D histograms of the input HDR image in the HDR10 format and its tone-mapped image, respectively. Then, we can obtain the 2D histogram of the tone-mapped image $\mathbf{H}_o$, which minimizes the cost in (4), by

$$\underset{\mathbf{H}_o}{\text{minimize}} \ \|\mathbf{H}_i \circ \mathbf{R}_i - \mathbf{H}_o \circ \mathbf{R}_o\|_F^2 \tag{5}$$

where the (m, n)th elements of the HVS response matrices $\mathbf{R}_i$ and $\mathbf{R}_o$ are the HVS response differences $R_i(m, n)$ and $R_o(m, n)$, respectively.

B. Luminance-Preserving Optimization

The optimization problem in (5) considers only the HVS response differences between neighboring pixels. However, the solution may cause visual differences between the input HDR image and the tone-mapped image, since the absolute luminance values are not considered. Therefore, to consider the absolute luminance, we employ a guide TMC that can preserve the absolute luminance values in the HDR image during the tone mapping procedure.

We first obtain the guide image $\mathbf{I}_g$ by applying the guide TMC to the input HDR image. Then, the guide 2D histogram $\mathbf{H}_g$ is computed from the guide image $\mathbf{I}_g$. The output 2D histogram $\mathbf{H}_o$ should be close to the guide 2D histogram $\mathbf{H}_g$ to maintain the luminance values of the tone-mapped image. Note that, according to the HVS response model [18], the HVS is more sensitive at the dark region than at the bright region, since only a small change of luminance causes a significant change in the response. Thus, we define a weight matrix $\mathbf{W}$, in which the elements corresponding to dark regions have higher values to better preserve small luminance values. Then, the luminance-preserving optimization can be formulated as

$$\underset{\mathbf{H}_o}{\text{minimize}} \ \|\mathbf{W} \circ (\mathbf{H}_o - \mathbf{H}_g)\|_F^2. \tag{6}$$

We now have two objectives: the 2D histogram $\mathbf{H}_o$ should minimize the HVS response difference in (5); and it should minimize the luminance difference in (6). To achieve both objectives simultaneously, we formulate the optimization

$$\underset{\mathbf{H}_o}{\text{minimize}} \ \alpha\|\mathbf{H}_i \circ \mathbf{R}_i - \mathbf{H}_o \circ \mathbf{R}_o\|_F^2 + (1-\alpha)\|\mathbf{W} \circ (\mathbf{H}_o - \mathbf{H}_g)\|_F^2 \tag{7}$$

where parameter α controls the relative importance between HVS response preservation and luminance preservation.

C. Smoothness Constraint

Peak values in the 2D histogram may cause steep slopes in the final TMC. Consequently, it may result in quality degradation in the tone-mapped images, *e.g.*, contour artifacts. To address such problems, we add a smoothness constraint to the optimization problem as in [11], given by

$$\underset{\mathbf{H}_o}{\text{minimize}} \ \|\mathbf{H}_o \mathbf{D}\|_F^2 \tag{8}$$

where $\mathbf{D}$ is a bidiagonal difference matrix

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{9}$$

D. Final Optimization

The final optimization problem, consisting of the HVS response preserving term in (5), luminance-preserving term in (6), and smoothness term in (8), can be formulated as

$$\underset{\mathbf{H}_o}{\text{minimize}} \ \alpha\|\mathbf{H}_i \circ \mathbf{R}_i - \mathbf{H}_o \circ \mathbf{R}_o\|_F^2 + (1-\alpha)\|\mathbf{W} \circ (\mathbf{H}_o - \mathbf{H}_g)\|_F^2 + \beta\|\mathbf{H}_o\mathbf{D}\|_F^2 \tag{10}$$

where parameter β controls the level of smoothness. Note that, since the optimization problem in (10) is an unconstrained quadratic problem of a single variable $\mathbf{H}_o$, we can obtain a closed-form solution efficiently.

Then, we obtain the optimal TMC from the input 2D histograms $\mathbf{H}_i$ and the output 2D histogram $\mathbf{H}_o$ using a histogram matching technique [12]. Specifically, we first normalize $\mathbf{H}_x$ as

$$h_x(m, n) = \frac{h_x(m, n)}{\sum_{i=1}^{K}\sum_{i=1}^{K} h_x(i, j)} \tag{11}$$

where $K = 1024$ for a 10-bit image. Next, the cumulative density function (CDF) of $\mathbf{H}_x$ is computed by

$$P_x = \left\{ P_x(m) = \sum_{i=1}^{m}\sum_{j=1}^{m} h_x(i, j) | m = 1, 2, ..., K \right\}. \tag{12}$$

We compute the CDFs P_i and P_o for $\mathbf{H}_i$ and $\mathbf{H}_o$, respectively, using (11) and (12). Finally, the input pixel value m is mapped to the output pixel value m' by

$$m' = \underset{i \in \{1,2,...,K\}}{\arg\min} \ |P_i(m) - P_o(i)|. \tag{13}$$

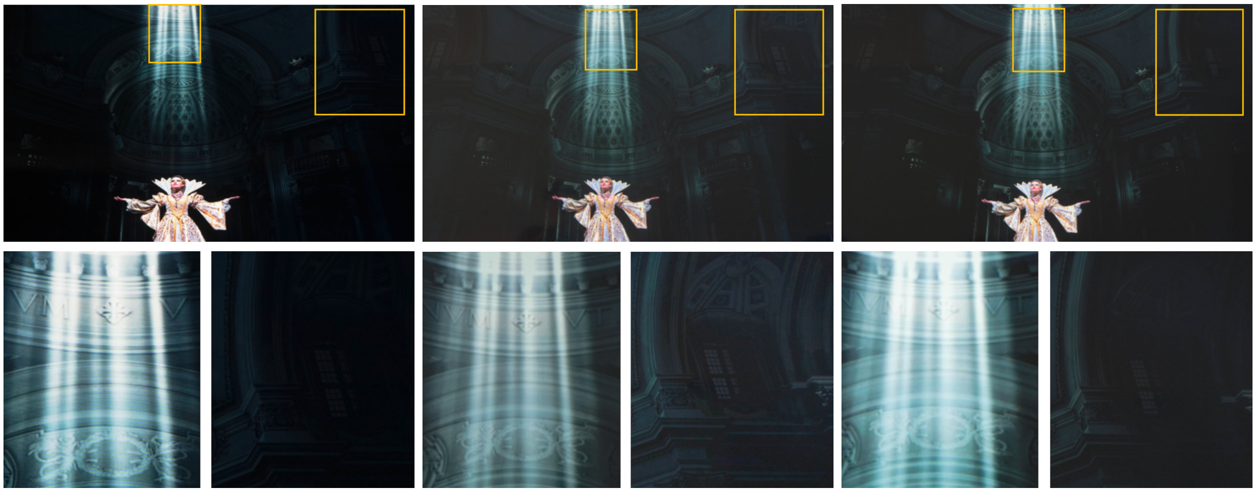
IV. Experimental Results

We evaluate the performance of the proposed tone-mapping algorithm subjectively using two real displays. We compare the performance of the proposed algorithm with that of the HDR10 standard in [14], [16]. Each element in the weight matrix $\mathbf{W}$ in (6) is defined as

$$W(m, n) = 1023 - \max(m, n) \tag{14}$$

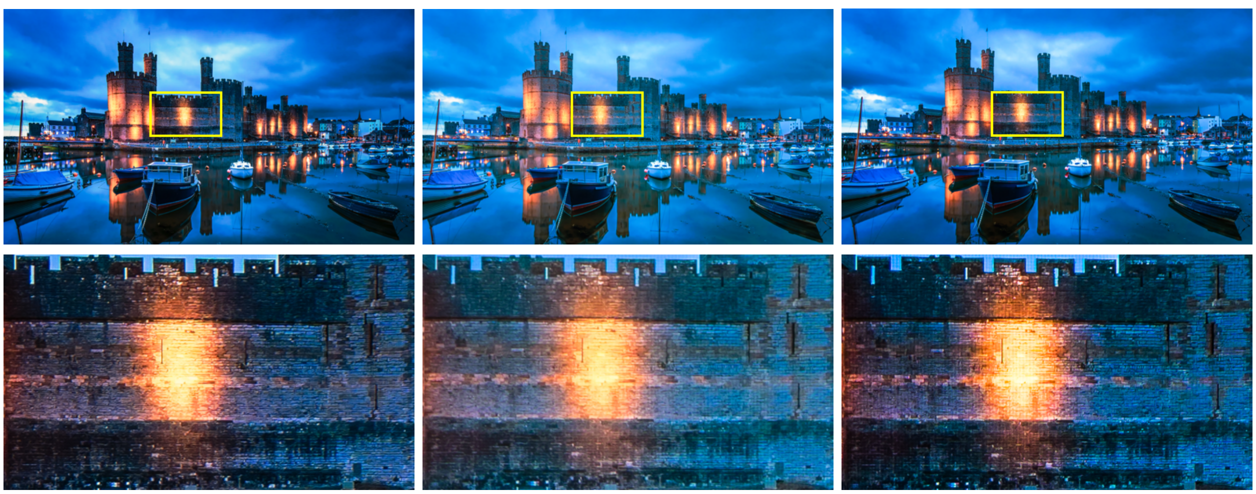
to assign larger values to the pixels in dark regions. The parameters α and β in (10) are fixed to 0.2 and 0.5, respectively. The evaluation used two displays: a television with 1,000-nits peak luminance value as a reference HDR display and another television with 500-nits peak luminance as a target SDR display. The HDR display showed the original image and the SDR display showed the tone-mapped image, which was converted by each algorithm.

Figs. 2 and 3 compare the tone mapping results on the display devices, which were taken using a digital camera, and their detailed parts on the *Queen* and *Castle* images, respectively. The HDR10 standard [14], [16] in Figs. 2(b) and 3(b) alters the overall contrast in the images providing brighter results at the dark regions and darker results at the bright regions compared to the references in Figs. 2(a)



(a) (b) (c)

Fig. 2. Comparison of the displayed results of the *Queen* image. (a) The reference image on the HDR display. The results on the SDR display tone-mapped by (b) the HDR10 standard [14], [16] and (c) the proposed algorithm. The magnified parts are shown in the bottom row.



(a) (b) (c)

Fig. 3. Comparison of the displayed results of the *Castle* image. (a) The reference image on the HDR display. The results on the SDR display tone-mapped by (b) the HDR10 standard [14], [16] and (c) the proposed algorithm. The magnified parts are shown in the bottom row.

and 3(a). This is because the HDR10 standard uses a single static TMC for all images, Thus, it is not adaptive to the characteristics of the input images. On the contrary, we see that the proposed algorithm in Figs. 2(c) and 3(c) produces faithful results without contrast alternation, which is perceptually more similar to those on the reference HDR display in Figs. 2(a) and 3(a), since it employs the 2D histogram to take the characteristics of the input image into account in the tone mapping procedure. Thus, the proposed tone-mapping algorithms is adaptive to the characteristics of the input image.

V. Conclusions

We have proposed an optimized HVS model-based tone-mapping algorithm for displaying HDR10 image content on conventional SDR devices. We first quantified the HVS response to an image when it was displayed on a device using the HVS model-based 2D histogram. Then, we formulated an optimization problem to minimize the HVS response difference between an input HDR and its tone-mapped image. Finally, we obtained the optimal TMC by efficiently solving the optimization problem. Experimental results demonstrated that the proposed tone-mapping algorithm provides higher image qualities than the HDR10 standard on real displays.

References

[1] *Parameter Values for Ultra-High Definition Television Systems for Production and International Programme Exchange*, ITU-R BT.2020-2, 2015.

[2] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd ed. Morgan Kaufmann Publishers, 2010.

[3] *Common File Format & Media Formats Specification Version 2.1*, Digital Entertainment Content Ecosystem (DECE), 2015.

[4] F. Drago, K. Myszkowski, T. Annen, N. Chiba, P. Brunet, and D. Fellner, "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 419–426, Sep. 2003.

[5] D. Gommelet, A. Roumy, C. Guillemot, M. Ropert, and J. L. Tanou, "Gradient-based tone mapping for rate-distortion optimized backward-compatible high dynamic range compression," *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5936–5949, Dec. 2017.

[6] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 68:1–68:10, Aug. 2008.

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

[8] L. Meylan and S. Susstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2820–2830, Sep. 2006.

[9] F. Kou, W. Chen, C. Wen, and Z. Li, "Gradient domain guided image filtering," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4528–4539, Nov. 2015.

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

[11] T. Celik and T. Tjahjadi, "Contextual and variational contrast enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3431–3441, Dec. 2011.

[12] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognit.*, vol. 45, no. 10, pp. 3810–3824, Oct. 2012.

[13] C. Lee, C. Lee, and C. S. Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5372–5384, Dec. 2013.

[14] *High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays*, SMPTE ST 2084:2014, 2014.

[15] *Image Parameter Values for High Dynamic Range Television for Use in Production and International Programme Exchange*, ITU-R BT.2100-1, 2017.

[16] *Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images*, SMPTE ST 2086:2014, 2014.

[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] H. R. Blackwell, "Contrast thresholds of the human eye," *J. Opt. Soc. Am.*, vol. 36, no. 11, pp. 624–643, Nov. 1946.

[19] B. Wandell, *Foundations of Vision*. Sinauer Associates, 1995.

[20] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, Jul. 2002.