IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 25, 2021, accepted September 3, 2021, date of publication September 10, 2021, date of current version September 21, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3112046

Human Visual System Model-Based Optimized Tone Mapping of High Dynamic Range Images

NAM HOANG NGUYEN[1], TU VAN VO[2], (Student Member, IEEE), AND CHUL LEE[1], (Member, IEEE)

[1]Department of Multimedia Engineering, Dongguk University, Seoul 04620, South Korea
[2]Department of Computer Engineering, Pukyong National University, Busan 48513, South Korea

Corresponding author: Chul Lee (chullee@dongguk.ac.kr)

This work was supported in part by Samsung Electronics, and in part by the National Research Foundation of Korea (NRF) Grant by the Korean Government (MSIT) under Grant NRF-2019R1A2C4069806.

ABSTRACT High dynamic range (HDR) image and video technology can provide significant picture quality improvement over the standard dynamic range (SDR). However, when HDR content is represented on an SDR display, dynamic range compression may result in image quality deterioration. To address this problem, we propose an optimized human visual system (HVS) response model-based tone-mapping algorithm to preserve the perceptual responses between the HDR image and its tone-mapped image. First, we measure the HVS response differences using a 2D histogram when an HDR image is displayed on an HDR device and when its tone-mapped image is displayed on an SDR device. Then, we formulate an optimization problem to minimize the differences. By efficiently solving the optimization problem, we obtain an optimal tone-mapping curve. Experimental results on actual displays demonstrate that the proposed algorithm provides superior image quality compared with conventional algorithms in terms of both subjective and objective evaluations.

INDEX TERMS High dynamic range, HDR10, tone-mapping, human visual system, contrast sensitivity, optimization.

I. INTRODUCTION

The ultra-high-definition (UHD) television standard, an advanced form of the high-definition television standard, can provide users with improved visual quality by offering increased resolution. Currently, due to the recent expansion of UHD video content and broadcasting, the market for UHD television has been rapidly expanding. UHD broadcasting produces video content with higher quality by improving the resolution, bit depth, and color gamut [1]. In particular, UHD broadcasting supports high dynamic range (HDR) image content, which can represent a wider dynamic range of luminance levels than standard dynamic range (SDR) content [2]. Due to the higher dynamic range capability, HDR content can present a look similar to that experienced by the human eye through the human visual system (HVS) [2].

The HDR pipeline consists of three main steps: content generation, storage, and visualization [3]. First, a camera is used to capture HDR images, either by capturing and then synthesizing multiple exposures, *e.g.*, [4], [5], or by capturing a single exposure using an HDR sensor, *e.g.*, [6], [7]. Because of its higher dynamic range, HDR content requires considerable storage space and transmission bandwidth compared to SDR content. Therefore, to make HDR content manageable, various techniques have been developed to store, process, and distribute HDR images. Finally, an HDR-compatible display can be used to represent HDR content, bearing in mind that dynamic range compression should be employed to adapt HDR content for representation on SDR displays. Using HDR imaging, it is possible to capture, store, manipulate, transmit, and display images that more accurately represent real-world scenes. Extensive effort has been focused on utilizing the advancements of HDR content, resulting in the adoption of HDR transport and broadcasting formats.

The most common HDR transport and broadcasting format is HDR10 Media Profile (HDR10) [8], which is currently used in UHD transport and broadcasting services due to the widespread support by display manufacturers and service providers. HDR10 is designed to support up to 4,000 cd/m^2 (nits) [9], thereby providing higher perceptible contrast

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

between the brightest and darkest regions of an image, which better represents what the HVS can see. The high peak brightness also preserves the details in highlighted and shadows areas that are lost when using conventional SDR format. In addition, while the color depth for SDR content is typically 8-bit, HDR10 content uses 10-bit color depth [9]. This advancement allows the production of more realistic images, with smooth transitions from dark to bright regions and less contour artifacts. However, an SDR display can present content with a lower dynamic range than that defined in HDR10. Moreover, because the peak luminance represented by each device is different, devices that support HDR10 content display such content differently. Therefore, when HDR10 content is displayed on an SDR device or when the maximum luminance values of the HDR content and the display are different, a process of converting the dynamic range of the input HDR content, called tone mapping [2], is required.

Many tone-mapping algorithms have been proposed, and they can be classified as global or local algorithms according to their methods of deriving transformation functions [2]. In a global tone-mapping algorithm, a single monotonic function is employed to convert all pixels in an entire image. For example, Drago *et al.* [10] employed a logarithmic function to reduce the dynamic range by imitating the visual response of the HVS. Gommelet *et al.* [11] formulated an optimization problem to minimize the spatial gradient-based distortion between an HDR image and its tone-mapped image. Mantiuk *et al.* [12] developed another global tone-mapping algorithm that minimizes the perceivable contrast distortion, given the display characteristics. Recently, Ok and Lee [13] proposed a tone-mapping algorithm to reproduce the details in dark and bright regions by maximizing the tone-mapping quality metric of a tone-mapped image. Yang *et al.* [14] used two gamma functions to adjust the luminance of the dark and bright regions and then combined the two adjusted luminance maps to improve the details and structural fidelity of the tone-mapped image. Song and Cosman [15] considered the display contrast and human visual adaptation to construct the tone-mapping function, thereby enhancing the luminance and visibility of details. Nguyen *et al.* [16] proposed an HVS response model-based tone-mapping algorithm to minimize the visual response difference between an HDR image and its tone-mapped image. Although these global tone-mapping algorithms are computationally efficient in general, they may fail to preserve the 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 to take into account neighboring pixel intensities [2]. For example, Reinhard *et al.* [17] proposed a local tone-mapping algorithm by adopting the Zone System that was used to predict the mapping method between the actual luminance and a set of print zones. Meylan and Süsstrunk [18] employed a center-surround retinex model to preserve high-contrast edges and reduce artifacts. In [19], a gradient domain guided image filter was employed by incorporating an edge-aware constraint to preserve local details more effectively. Ma *et al.* [20] developed a tone-mapping algorithm based on a tone-mapped image quality index to improve the structural fidelity and naturalness of the tone-mapped image. Recently, Mezeni and Saranovac [21] employed detail and base layer decomposition to preserve and enhance the details while preserving global contrast in the tone-mapped images. Shu and Wu [22] employed locally adaptive rank-constrained optimal tone mapping to reduce artifacts such as halos and double edges. Local tone-mapping algorithms provide tone-mapped results with fine details but often show high computational complexity.

Recently, inspired by the success of deep learning in various image processing tasks, several deep learning-based tone mapping techniques that use convolutional neural networks (CNNs) or generative adversarial networks (GANs) have been developed [23]–[27]. For example, Zhang *et al*. [23] developed a tone-mapping network by employing the improved Wasserstein GAN and constructed a dataset of HDR images and the corresponding tone-mapped images, manually adjusted by photographers. Montulet and Briassouli [24] proposed an end-to-end tone-mapping algorithm based on deep convolutional GANs using a dataset of HDR images and their tone-mapped images obtained by experts. Rana *et al*. [25] improved the quality of a tone-mapped image by employing a multiscale conditional GAN to alleviate the problems of the conventional GAN-based tone-mapping algorithms [23], [24], such as blurring, tiling patterns, and saturation artifacts. Panetta *et al.* [26] further improved the performance of the GAN-based tone-mapping algorithm by developing an attention-guided generator architecture. Kim *et al.* [27] proposed a detail-preserving tone-mapping algorithm that consists of two CNNs; one restores the details in the input HDR image, and the other compresses the dynamic range. However, the main disadvantage of the learning-based approaches is that their performance is highly dependent on the training datasets.

Although conventional tone-mapping algorithms have attempted to obtain high-quality results [10]–[27], little effort has been made to consider the characteristics of the display devices on which the tone-mapped images are presented. The display-adaptive tone-mapping algorithm [12] is computationally inefficient, because the computation of the HVS response to the input HDR image requires high computational complexity. Meanwhile, the luminance enhancement tone-mapping algorithm [15] considers only the image characteristics based on the pixel value distribution, without the textual details. The HVS response model-based tone-mapping algorithm [16] uses a guided tone-mapping curve (TMC) to preserve the absolute luminance that may cause detail loss in the bright regions. We also note that, although recent deep learning-based algorithms have shown significant tone-mapping performance improvements [23]–[27], they too have been developed without considering the characteristics of the display devices.

To address the aforementioned limitations, we develop an efficient HVS response model-based tone-mapping algorithm that preserves the perceptual quality of tone-mapped images in this work. Specifically, we first formulate an optimization problem that minimizes the perceptual response differences between an input HDR image in HDR10 format and its tone-mapped image using a 2D histogram. Then, we obtain an optimal TMC by efficiently solving the optimization problem. Experimental results demonstrate that the proposed tone-mapping algorithm provides higher quality tone-mapped images than other state-of-the-art algorithms, by preserving the perceptual similarities between the input HDR images and the corresponding tone-mapped images on actual displays.[1]

The remainder of this paper is organized as follows. Section II briefly reviews the relevant background. Section III describes the proposed tone-mapping algorithm. Section IV presents the experimental results. Finally, Section V concludes the paper.

II. BACKGROUND

In this section, we provide the background, on which the proposed algorithm is based, including 2D histogram equalization, the HDR10 standards, the HVS response model, and relative contrast measurement.

A. 2D HISTOGRAM EQUALIZATION

We employ 2D histogram equalization [28]–[30] to consider the local details in an HDR image. A 2D histogram $\mathbf{H}_x$ is constructed using a mutual relationship between each pixel and its neighboring pixels in the image. Assume that an input image $\mathbf{X} = \{x(i,j)|1 \le i \le H, 1 \le j \le W\}$ of size $H \times W$ is given, where $x(i,j) \in [1, K]$. In this work, we consider a 10-bit image, thus $K = 1,024$. Let $H_x(m,n)$ denote 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

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

The contrast of an image is enhanced via the histogram equalization technique. Specifically, an optimization problem can be formulated to minimize the distance between the 2D histogram $\mathbf{H}$ of the output image and the uniformly distributed 2D histogram $\mathbf{H}_u$, which is 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 the input pixel intensities to the output intensities can be derived using the histogram matching scheme [28].

[1] Preliminary results of this work have been presented in part in [16]. In this paper, we develop a new algorithm to preserve the HVS model-based global contrast and provide rigorous solutions to the optimization. Furthermore, more comprehensive experiments are included, which show the effectiveness of the proposed algorithm, including objective quality assessment and computational complexity evaluation.

B. HDR10 STANDARDS

Significant video quality improvement is primarily achieved via the use of HDR. HDR10 has recently been adopted to facilitate transport and broadcasting of HDR videos [8]. The HDR10 standard supports a luminance of up to 4,000 units with a nonlinear transfer function for perceptually unnoticeable quantization to generate an efficient mapping from the pixel values to the absolute luminance values [31], [32], which is called an electro-optical transfer function (EOTF). However, to improve the display quality of the HDR10 content further, content mastering by the experts from the studio is required. Specifically, the content mastering process involves selecting a viewing environment and examining the quality of the HDR content after performing adjustments. The purpose of mastering is to improve the color reproduction to reproduce the creators' intent more exactly [33]. In addition, HDR10 uses the wide-gamut BT.2020 color space [9] with 10-bit color depth, which covers 75.8% of the CIE 1931 color space [34]. While an 8-bit RGB image only has 256 levels per channel, corresponding to 16.7 million colors, a 10-bit image supports 1,024 levels per channel and can display up to 1.07 billion colors. Because of the wide color gamut capability, HDR10 provides better color shades and smoother transitions between light and dark areas. HDR10 also uses static metadata to transport the color calibration data of the mastering display, which are encoded as supplemental enhancement information (SEI) messages within the video stream [33]. Moreover, the high-efficiency video coding (HEVC) Main 10 profile [35] also supports HDR10 encoding, which improves the compression ratio, bandwidth usage, and image quality.

In this work, as we consider HDR images in the HDR10 format, we use the standard EOTF to accurately compute the actual luminance of each pixel when an HDR10 image is presented on a specific display. According to standards [31], [32], the EOTF is defined with a high luminance range from 0 to 10,000 nits. Specifically, the linear luminance value L, which corresponds to the optical output of a display device, can be obtained from the normalized pixel intensity N as

$$L = \left(\frac{\max[(N^{1/m_2} - R_1), 0]}{R_2 - R_3 N^{1/m_2}}\right)^{1/m_1} \tag{3}$$

where the model parameters are $m_1 = 0.1593$, $m_2 = 78.8437$, $R_1 = 0.8359$, $R_2 = 18.8515$, and $R_3 = 18.6875$. Then, the absolute output luminance value of the display, L_d, in nits is defined as

$$L_d = 10,000L. \tag{4}$$

C. HVS RESPONSE MODEL

We measure the perceptual response of the HVS to an image using the contrast sensitivity, which is related to the ability to distinguish an object from its background [36]. According to the Weber-Fechner law, the minimum detectable contrast is constant regardless of luminance [37]. However, many

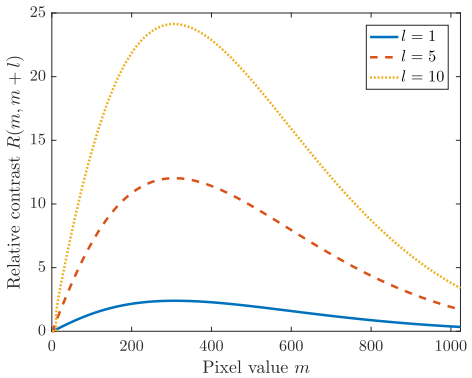


FIGURE 1. Relative contrast of pixel pair $(m, m+l)$ for $l = 1$, $l = 5$, and $l = 10$.

psychophysical experiments have shown that the contrast sensitivity varies based on other factors such as the viewing angle, screen size, and spatial frequency of the image [36]. Thus, contrast sensitivity functions (CSFs) were proposed to accurately measure the sensitivity of the human vision at a variety of spatial frequencies. In this work, we consider the contrast sensitivity of an image using the CSF from [38], since all aforementioned conditions that affect the contrast sensitivity are taken into account. Specifically, the CSF is defined as a function of luminance L and spatial frequency u as

$$S(L,u) = \frac{e^{-2\pi^2\sigma^2u^2}/k}{\sqrt{\frac{2}{T}\left(\frac{1}{X_0^2}+\frac{1}{X_{\max}^2}+\frac{u^2}{N_{\max}^2}\right)\left(\frac{1}{\eta pE}+\frac{\Phi_0}{1-e^{-(u/u_0)^2}}\right)}} \tag{5}$$

where $\sigma = \sqrt{\sigma_0^2 + (R_{ab}d)^2}$, $d = 5 - 3\tanh\left(0.4\log\left(\frac{LX_0^2}{40^2}\right)\right)$, and $E = \frac{\pi d^2}{4}L(1-(d/9.7)^2+(d/12.4)^4)$ [38]. The values of the constants used in the model are $k = 3.0$, $\sigma_0 = 0.5$ arc min, $R_{ab} = 0.08$ arc min/mm, $T = 0.1$ sec, $X_{\max} = 12°$, $N_{\max} = 15$ cycles, $\eta = 0.03$, $\Phi_0 = 3 \times 10^{-8}$ sec deg^2, $u_0 = 7$ cycles/deg, $p = 1.2 \times 10^6$ photons/sec/deg^2/Td, and $X_0 = 40°$ [38]. For each value of L, the peak sensitivity is obtained by

$$S_{\max}(L) = \max_u S(L,u). \tag{6}$$

Then, the minimum detectable contrast for each luminance level is formulated using the peak sensitivity [39]

$$M(L) = \frac{1}{S_{\max}(L)} \times 2 \times \frac{1}{1.27} \tag{7}$$

where the values 2 and $\frac{1}{1.27}$ are for the conversions from modulation to contrast and sinusoidal to rectangular waves, respectively [39].

D. RELATIVE CONTRAST

We quantify the HVS responses of two luminance values in an HDR image by employing the relative contrast [15]. Specifically, the relative contrast is measured in just noticeable difference (JND) units and used to quantify the HVS response between two luminance values. Specifically, given two pixel values m and $m+l$, we first measure the contrast between their luminance values as

$$C(m, m+l) = 2 \times \frac{L(m+l) - L(m)}{L(m+l) + L(m)} \tag{8}$$

where $L(.)$ denotes the EOTF of the display. Then, the relative contrast is defined as the ratio between the contrast in (8) and the minimum detectable contrast in (7) as

$$R(m, m+l) = \frac{C(m, m+l)}{M(L(m))}. \tag{9}$$

Figure 1 shows examples of the relative contrast for pixel value pairs $(m, m+l)$ with different values of l. As illustrated, the relative contrast increases in proportion to the value of l, since the contrast in (8) gets larger when the luminance differences between the two pixel values increase. In this work, we consider the relative contrast between two consecutive pixel values using $l = 1$. With increasing l, the relative contrast between pixel pairs with larger luminance differences is taken into account.

III. PROPOSED ALGORITHM

The reduced dynamic range caused by tone mapping results in perceptual image quality deterioration. Our objective is to develop a tone-mapping algorithm that preserves the perceptual responses between the input HDR image in HDR10 format and its tone-mapped image presented on an SDR display. To this end, we first measure the contextual information of an input HDR image and the perceptual quality when it is represented on an HDR display using a 2D histogram and the HVS response model. Then, we formulate an optimization problem to minimize the perceptual differences between the input and tone-mapped images in terms of local and global relative contrast. Finally, the optimal TMC is obtained by solving the optimization problem. We describe these procedures in the following sections.

A. LOCAL CONTRAST-PRESERVING OPTIMIZATION

Because the HVS is more sensitive to the luminance difference between neighboring pixels than the absolute luminance values [40], [41], we consider the local contrast between pixels and their neighbors in this work. Specifically, let $R_{\text{ref}}(x, y)$ and $R_{\text{tgt}}(x, y)$ be the relative contrast in (9) between two luminance values at locations x and $y \in \mathcal{N}(x)$, where $\mathcal{N}(x)$ denotes a set of neighboring pixels of x, in the reference HDR display and the target SDR display, respectively. Then, to preserve the local contrast of the tone-mapped image on the SDR display, we minimize the total relative contrast differences for all pixels in the image, defined as

$$\sum_{x, y \in \mathcal{N}(x)} (R_{\text{ref}}(x, y) - R_{\text{tgt}}(x, y))^2. \tag{10}$$

To minimize the cost in (10), we first take into account the HVS responses of neighboring pixels. Specifically, the HVS response of an image is quantified by the relative contrast using the EOTF [31], [32] and the HVS response model [38].

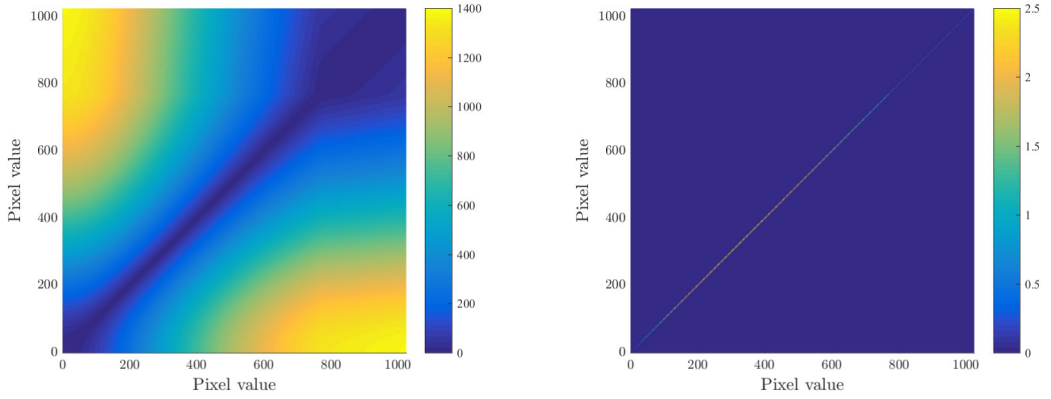


FIGURE 2. Examples of (a) local contrast matrix $\mathbf{R}_l$ and (b) global contrast matrix $\mathbf{R}_g$.

Let $T(.)$ be the tone-mapping function that converts an input pixel value into an output pixel value. We then define the relative contrast $R_l(m, n)$ and $R_{l,T}(m, n)$ between the actual luminance of two pixel values m and n on the reference display and the corresponding pixel values $T(m)$ and $T(n)$ on the target displays, respectively, *i.e.*,

$$R_l(m, n) = R(m, n), \tag{11}$$

$$R_{l,T}(m, n) = R(T(m), T(n)). \tag{12}$$

In this work, we obtain the optimal TMC T using 2D histogram equalization. Let $\mathbf{H}$ and $\mathbf{H}_T$ denote the 2D histograms of the input HDR image, in the HDR10 format, and its tone-mapped image, obtained using T. We note that, although a 2D histogram can efficiently consider the textual information of an image, the local details in a high-resolution image may not be taken into account faithfully using a small neighborhood size. Furthermore, constructing a 2D histogram using a larger neighborhood window is time-consuming. To address this issue, we adopt a multiscale 2D histogram approach to consider the contextual information of the input image more effectively. More specifically, a Gaussian pyramid, $\mathbf{I}_0, \mathbf{I}_1, \ldots, \mathbf{I}_N$, is constructed from the HDR input image $\mathbf{I}$, where $\mathbf{I}_0$ is the full resolution image and $\mathbf{I}_N$ is the coarsest image. Then, we define the input 2D histogram as the combination of the corresponding 2D histograms $\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_N$ of these scaled images with their weights as

$$\mathbf{H} = \omega_0\mathbf{H}_0 + \omega_1\mathbf{H}_1 + \cdots + \omega_N\mathbf{H}_N \tag{13}$$

where ω_i is the ith weight for the 2D histogram of corresponding scaled image.

Next, we quantify the HVS response between every two pixel values by exploiting display characteristics and then formulate an optimization problem using a 2D histogram. Specifically, two local contrast matrices $\mathbf{R}_l$ and $\mathbf{R}_{l,T}$ are constructed for the reference and target displays, where the (m, n)th elements are the relative contrast $R_l(m, n)$ and $R_{l,T}(m, n)$ in (11) and (12), respectively. An example of the local contrast matrix $\mathbf{R}_l$ for the reference display is illustrated in Figure 2(a). Then, the optimal 2D histogram $\mathbf{H}_T$ and TMC T, which minimize the cost in (10), can be obtained by

$$\underset{\mathbf{H}_T, T}{\text{minimize}} \ \|\mathbf{H} \circ \mathbf{R}_l - \mathbf{H}_T \circ \mathbf{R}_{l,T}\|_F^2 \tag{14}$$

where $\circ$ denotes the element-wise product of two matrices.

(a) (b)

FIGURE 3. (a) Reference images on the HDR display and (b) tone-mapped images obtained using the local contrast-preserving optimization.

B. GLOBAL CONTRAST-PRESERVING OPTIMIZATION

The optimization problem in (14) considers only the relative contrast between the luminance of neighboring pixels. However, the solution may cause visual differences between the input HDR image and the tone-mapped image, since the global contrast is not considered. Figure 3 compares reference images with tone-mapped images obtained using the local contrast-preserving optimization. The output tone-mapped images contain over-enhanced artifacts, *e.g.*, the details of the liquid regions in the *Paint* image in the first row and the mountain regions in the *Temples* image in the second row. We address this issue by maintaining the global contrast of every pixel in the tone-mapped image. Specifically, given a pixel value pair $(m(x), m(x) + l)$ in the reference display, where $m(x)$ denotes the pixel value at location x, by applying T, we obtain the corresponding pair $(T(m(x)), T(m(x) + l))$ in the target display. We attempt to preserve the global relative

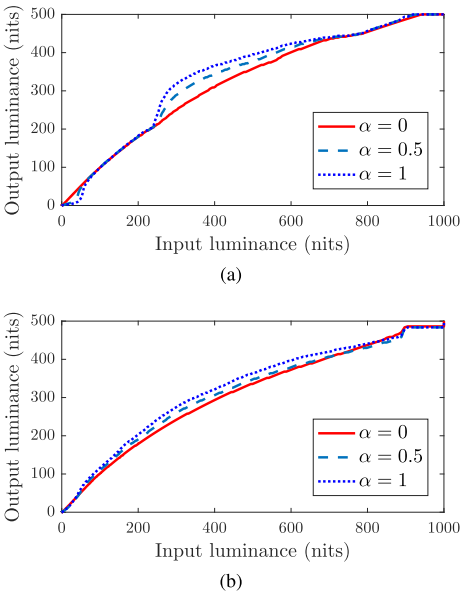


FIGURE 4. Comparison of the TMCs on the (a) *Motor* image and (b) *Flower* image obtained using different values of α.

contrast in the tone-mapped image by minimizing

$$\sum_x (R(m(x), m(x)+l) - R(T(m(x)), T(m(x)+l)))^2. \quad (15)$$

To minimize the cost in (15), we first define two global contrast matrices $\mathbf{R}_g$ and $\mathbf{R}_{g,T}$, where the $(m, m+l)$th elements are the relative contrast $R(m, m+l)$ and $R(T(m), T(m+l))$ in the reference and target displays, respectively, as

$$R_g(m, m+l) = R(m, m+l), \quad (16)$$

$$R_{g,T}(m, m+l) = R(T(m), T(m+l)). \quad (17)$$

Figure 2(b) shows an example of the global contrast matrix $\mathbf{R}_g$ for $l = 1$. The target 2D histogram $\mathbf{H}_T$ of the tone-mapped image and the optimal T, which minimize the cost in (15), can be obtained using the 2D histogram technique as

$$\underset{\mathbf{H}_T, T}{\text{minimize}} \ \|\mathbf{H} \circ \mathbf{R}_g - \mathbf{H}_T \circ \mathbf{R}_{g,T}\|_F^2. \quad (18)$$

There are now two objectives: that the output 2D histogram $\mathbf{H}_T$ should minimize the local contrast differences between neighboring pixels in (14) and that it should minimize the global contrast differences between pixel value pairs in (18). To achieve both objectives simultaneously, we formulate the optimization

$$\underset{\mathbf{H}_T, T}{\text{minimize}} \ \alpha\|\mathbf{H} \circ \mathbf{R}_l - \mathbf{H}_T \circ \mathbf{R}_{l,T}\|_F^2 + (1-\alpha)\|\mathbf{H} \circ \mathbf{R}_g - \mathbf{H}_T \circ \mathbf{R}_{g,T}\|_F^2 \quad (19)$$

where the parameter α controls the relative importance between local and global contrast preservation. When $\alpha = 0$, T is obtained using only the global contrast-preserving term. As α increases, T, which minimizes the cost in (19), gets closer to the TMC obtained using the local contrast-preserving optimization. Examples of TMCs generated using different values of α are shown in Figure 4, which exemplifies how α in (19) controls the relative importance of local and global contrast preservation. We will discuss the impacts of different values of α on the tone-mapped images in Section IV.

C. SMOOTHNESS CONSTRAINT

Peak values in the 2D histogram may cause steep slopes in the final TMC. Consequently, quality degradation in the tone-mapped images, *e.g.*, contour artifacts, may occur. To address such problems, we add a smoothness constraint to the optimization problem as in [28], given by

$$\underset{\mathbf{H}_T, T}{\text{minimize}} \ \|\mathbf{H}_T \mathbf{D}\|_F^2 \quad (20)$$

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

$$\mathbf{D} = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (21)$$

D. FINAL OPTIMIZATION

The final optimization problem, consisting of the local contrast-preserving term in (14), global contrast-preserving term in (18), and smoothness term in (20), can be formulated as

$$\underset{\mathbf{H}_T, T}{\text{minimize}} \ \alpha\|\mathbf{H} \circ \mathbf{R}_l - \mathbf{H}_T \circ \mathbf{R}_{l,T}\|_F^2 + (1-\alpha)\|\mathbf{H} \circ \mathbf{R}_g - \mathbf{H}_T \circ \mathbf{R}_{g,T}\|_F^2 + \beta\|\mathbf{H}_T \mathbf{D}\|_F^2 \quad (22)$$

where the parameter β controls the level of smoothness. As β increases, the algorithm attempts to further reduce the high peak in the output 2D histogram by minimizing the deviations between its components.

E. SOLUTION TO THE OPTIMIZATION

The final optimization problem in (22) is a joint optimization of $\mathbf{H}_T$ and T. However, since $\mathbf{H}_T$ is a function of T, joint optimization over both $\mathbf{H}_T$ and T is intractable in practice. In this work, we employ the alternating direction method of multipliers (ADMM) [42], in which each variable is optimized separately, while fixing the other one. Specifically, we iteratively update $\mathbf{H}_T$ and T individually. These subproblems are described below.

Updating $\mathbf{H}_T$*:* In the first step, given the estimate of $T^{(k)}$, we update $\mathbf{H}_T$ as

$$\mathbf{H}_T^{(k+1)} = \arg\min_{\mathbf{H}_T} \alpha\|\mathbf{H} \circ \mathbf{R}_l - \mathbf{H}_T \circ \mathbf{R}_{l,T^{(k)}}\|_F^2 + (1-\alpha)\|\mathbf{H} \circ \mathbf{R}_g - \mathbf{H}_T \circ \mathbf{R}_{g,T^{(k)}}\|_F^2 + \beta\|\mathbf{H}_T \mathbf{D}\|_F^2. \quad (23)$$

Because this subproblem is an unconstrained quadratic problem of a single variable $\mathbf{H}_T$, we can obtain a closed-form solution efficiently. In the Appendix, we derive

the closed-form solution to the subproblem in (23), which is given by

$$\mathbf{H}_T^{(k+1)} = \mathrm{vec}^{-1}\Big(\big(\mathrm{diag}(\mathrm{vec}(\mathbf{A})) + (\mathbf{B}^T \otimes \mathbf{I})\big)^{-1}\mathrm{vec}(\mathbf{C})\Big) \tag{24}$$

where $\mathbf{A} = 2(\alpha\mathbf{R}_{l,T^{(k)}} \circ \mathbf{R}_{l,T^{(k)}} + (1-\alpha)\mathbf{R}_{g,T^{(k)}} \circ \mathbf{R}_{g,T^{(k)}})$, $\mathbf{B} = 2\beta\mathbf{D}\mathbf{D}^T$, $\mathbf{C} = \alpha\mathbf{R}_{l,T^{(k)}} \circ (\mathbf{H} \circ \mathbf{R}_l + \mathbf{H}^T \circ \mathbf{R}_l^T) + (1-\alpha)\mathbf{R}_{g,T^{(k)}} \circ (\mathbf{H} \circ \mathbf{R}_g + \mathbf{H}^T \circ \mathbf{R}_g^T)$, and $\mathbf{I}$ is the identity matrix. Also, $\mathrm{vec}(\cdot)$ is the vectorization operator, $\otimes$ is the Kronecker product, and vec^{-1} is the inverse of the vectorization operator defined as

$$\mathrm{vec}^{-1}(\mathrm{vec}(\mathbf{Y})) = \mathbf{Y}, \text{ for } \mathbf{Y} \in \mathbb{R}^{K\times K}. \tag{25}$$

Updating T: Next, given $\mathbf{H}_T^{(k+1)}$, we estimate $T^{(k+1)}$ using the histogram matching technique [29]. Specifically, we first normalize a 2D histogram $\mathbf{H}_x$ as

$$H_x(m,n) = \frac{H_x(m,n)}{\sum_{i=1}^{K}\sum_{i=1}^{K} H_x(i,j)} \tag{26}$$

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

$$P_x = \left\{P_x(m) = \sum_{i=1}^{m}\sum_{j=1}^{m} H_x(i,j) | m = 1,2,\ldots,K\right\}. \tag{27}$$

We compute the CDFs P and $P_T^{(k+1)}$ for $\mathbf{H}$ and $\mathbf{H}_T^{(k+1)}$, respectively, using (26) and (27). Finally, we obtain the TMC $T^{(k+1)}(m)$ for the input pixel value m by

$$T^{(k+1)}(m) = \underset{i\in\{1,2,\ldots,K\}}{\arg\min} |P(m) - P_T^{(k+1)}(i)|. \tag{28}$$

The optimization variables $\mathbf{H}_T$ and T are iteratively updated until convergence. Specifically, we define the convergence rate at the kth iteration as

$$\xi^{(k)} = \frac{\|T^{(k+1)} - T^{(k)}\|}{\|T^{(k)}\|}, \tag{29}$$

and run the iteration until $\xi^{(k)} < 10^{-3}$ or the maximal number of iterations is reached. The complete algorithm to solve the final optimization in (22) is summarized in Algorithm 1.

Algorithm 1 Optimization for Solving (22)

Input: $\mathbf{H}$, α, β

1: **Initialization:** T = identity mapping, $k = 1$
2: **while** not converged **do**
3: Construct $\mathbf{R}_{l,T^{(k)}}$ and $\mathbf{R}_{g,T^{(k)}}$ via (12) and (17)
4: Update $\mathbf{H}_T^{(k+1)}$ via (24)
5: Update $T^{(k+1)}$ via (28)
6: $k = k + 1$
7: **end while**

Output: Optimal TMC $T^{(k)}$

IV. EXPERIMENTAL RESULTS

We evaluate the performance of the proposed tone-mapping algorithm in comparison with those of Mantiuk *et al.*'s display adaptive tone-mapping algorithm [12], Song and Cosman's content-dependence tone-mapping algorithm [15], and Nguyen *et al.*'s HVS response model-based tone-mapping algorithm [16]. Note that these algorithms were developed by exploiting the display characteristics.

A. EXPERIMENTAL SETTINGS

We selected test images with resolutions of 3840 × 2160 that were extracted from HDR10 video sequences. As luminance, contrast, colorfulness, and naturalness are important factors that influence the display quality of an image [43], [44], we first classified the images into different groups according to their characteristics. Then, to maximize the diversity of the image characteristics, test images were selected randomly from each group. In the experiments, we used 54 images with different characteristics.

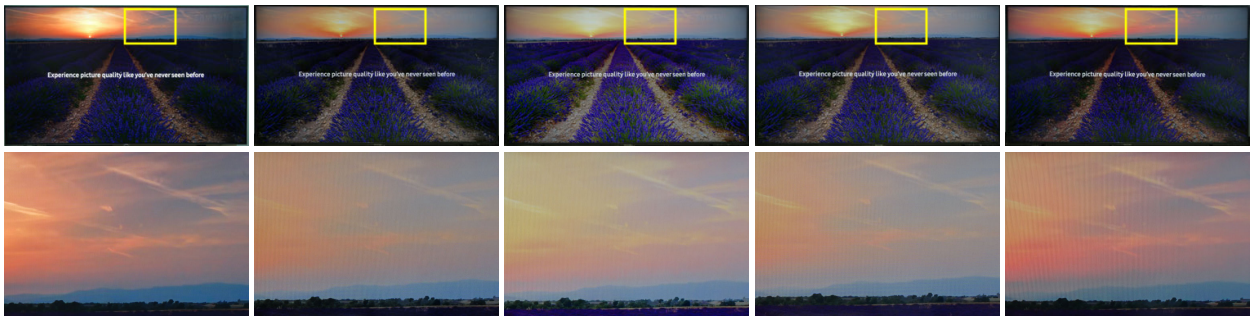
The 2D histogram of the input image was created by considering a neighborhood with a 3 × 3 window, and α and β in (22) are fixed to 0.2 and 0.5, respectively, in all experiments, unless specified otherwise. Additionally, l in (16) and (17) is set to 1; thus, we consider the relative contrast between every two consecutive pixel values. For the conventional algorithms, we used the publicly available implementation of Mantiuk *et al.*'s algorithm [12] from their webpage and incorporated the EOTF as a lookup table into their algorithm. We implemented Song and Cosman's algorithm [15] with 10-bit bit-depth and used the EOTF as the display model. The source code for Nguyen *et al.*'s algorithm [16] was provided by the authors.

Two displays were used for the evaluation: a television with a peak luminance of 1,000 nits was utilized as the reference HDR display and another television with a peak luminance of 500 nits was employed as the target SDR display. Both televisions use the same BT.2020 color space [9]. The HDR display presented the original images, and the SDR display showed the tone-mapped images, which were converted by each algorithm. All the tests were conducted in moderate room illumination.

B. SUBJECTIVE EVALUATION

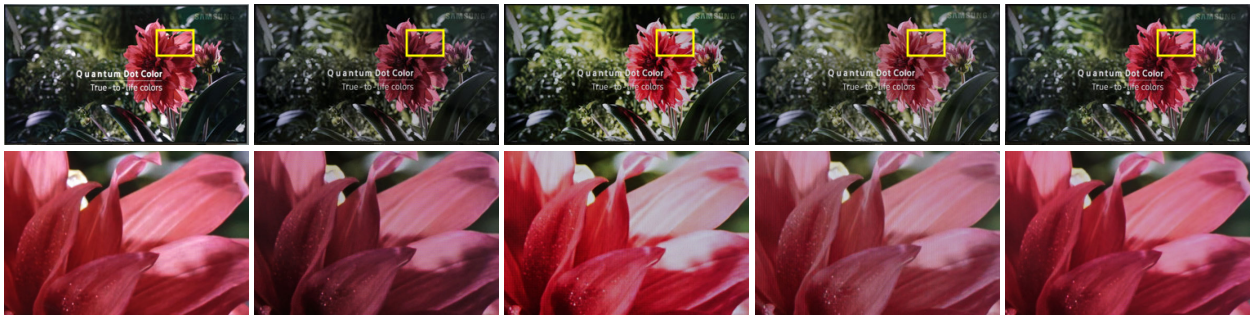
We first evaluate the performances of the tone-mapping algorithms by performing a subjective test. Figures 5–8 compare the tone-mapping results on the display devices, which were taken using a digital camera, as well as the detailed parts of the *Lavender*, *Flower*, *BlueHouse*, and *Glass* images. Note that, since Figures 5–8 present screen-captured images, moiré-pattern artifacts appear due to interference between the color filter array of the camera and the screen's subpixel layout. Therefore, we compare only the global and local contrast and colors in the images qualitatively, instead of comparing the amount of visible moiré artifacts.

Mantiuk *et al.*'s algorithm [12] reduces the contrast in the sky regions of the *Lavender* image in Figure 5(b). Their algorithm also darkens the *Flower* and *BlueHouse* images, as shown in Figures 6–7(b), respectively. Moreover,



(a) (b) (c) (d) (e)

FIGURE 5. Comparison of the displayed results of the *Lavender* image. (a) Reference image on the HDR display. Results on the SDR display that are tone-mapped by (b) Mantiuk *et al.*'s algorithm [12], (c) Song and Cosman's algorithm [15], (d) Nguyen *et al.*'s algorithm [16], and (e) the proposed algorithm. The magnified parts are shown in the bottom row.



(a) (b) (c) (d) (e)

FIGURE 6. Comparison of the displayed results of the *Flower* image. (a) Reference image on the HDR display. Results on the SDR display that are tone-mapped by (b) Mantiuk *et al.*'s algorithm [12], (c) Song and Cosman's algorithm [15], (d) Nguyen *et al.*'s algorithm [16], and (e) the proposed algorithm. The magnified parts are shown in the bottom row.

(a) (b) (c) (d) (e)

FIGURE 7. Comparison of the displayed results of the *BlueHouse* image. (a) Reference image on the HDR display. Results on the SDR display that are tone-mapped by (b) Mantiuk *et al.*'s algorithm [12], (c) Song and Cosman's algorithm [15], (d) Nguyen *et al.*'s algorithm [16], and (e) the proposed algorithm. The magnified parts are shown in the bottom row.

the results obtained for the *Glass* image in Figure 8(b) are significantly brighter than the reference image in Figure 8(a) since the luminance is extremely enhanced, resulting in detail loss in the bright regions, such as the flame and glass. This degradation occurs because Mantiuk *et al.*'s algorithm exploits the Laplacian pyramid to represent the image contrast, which is not effective since the total local contrast of every pixel with its neighbors is not thoroughly considered. Song and Cosman's algorithm [15] alters the overall contrast in the tone-mapped images, resulting in quality degradation. For example, the sky region of the *Lavender* image in Figure 5(c) loses a considerable amount of contrast. The brightness of the petal and glass regions in the *Flower* and *Glass* images in Figures 6(c) and 8(c), respectively, are increased noticeably, thus the details in these regions are lost. In addition, there are over-enhancement artifacts in the wall and stair regions of the *BlueHouse* image in Figure 7(c). This is because Song and Cosman's algorithm only exploits a 1D histogram that cannot consider the spatial information in the input image. Nguyen *et al.*'s algorithm [16] in Figure 5(d) also reduces the contrast in the sky regions of the *Lavender* image. Their algorithm yields output results with perceptual detail losses in the bright regions of the tone-mapped images, *e.g.*, the petal in the *Flower* image, the wall in the *BlueHouse* image, and the glass in the *Glass* image in Figures 6–8(d). This loss occurs because their algorithm attempts to preserve the absolute luminance value using a guided TMC without taking into account the global contrast of the tone-mapped image. On the contrary, the proposed algorithm in Figures 5–8(e) produces the best results without contrast

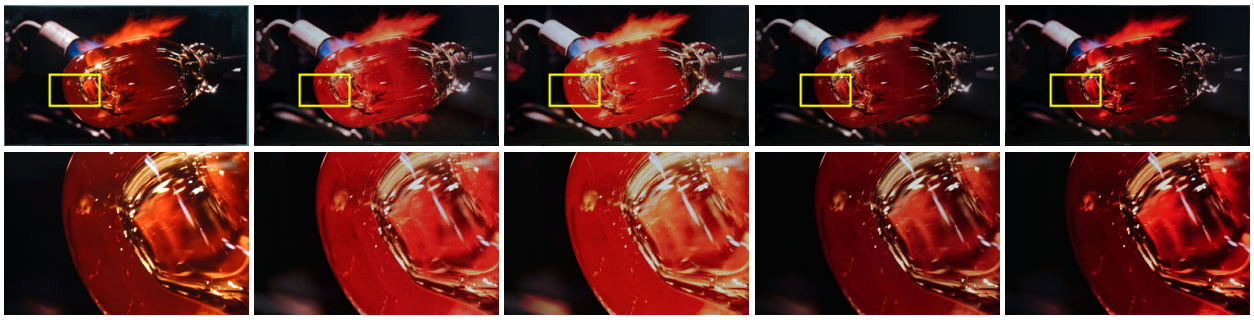


FIGURE 8. Comparison of the displayed results of the *Glass* image. (a) Reference image on the HDR display. Results on the SDR display that are tone-mapped by (b) Mantiuk *et al.*'s algorithm [12], (c) Song and Cosman's algorithm [15], (d) Nguyen *et al.*'s algorithm [16], and (e) the proposed algorithm. The magnified parts are shown in the bottom row.

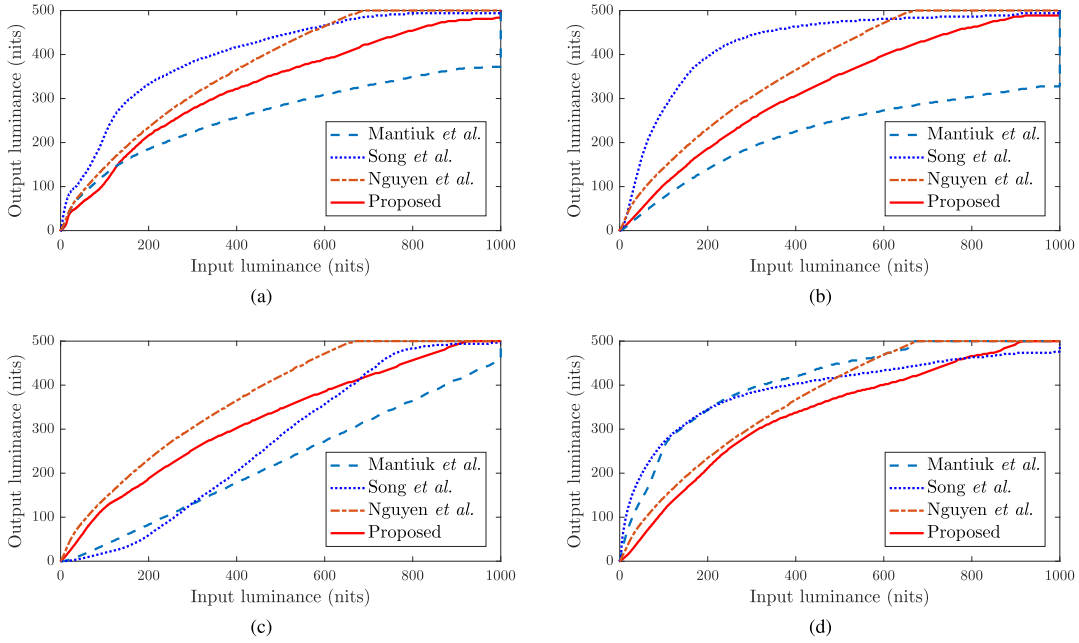


FIGURE 9. Comparison of the TMCs obtained for the (a) *Lavender*, (b) *Flower*, (c) *BlueHouse*, and (d) *Glass* images using different tone-mapping algorithms.

alternation and visible artifacts, which are perceptually more similar to those on the reference HDR display in Figures 5–8(a), since both the local and global relative contrast of the output images are considered adaptively to the HVS.

Figure 9 compares the TMCs that yield the output images in Figures 5–8. Mantiuk *et al.*'s algorithm produces TMCs that reduce the contrast and darken the output images, as shown in the *Lavender*, *Flower*, and *BlueHouse* images. Moreover, the large saturated region in the TMC obtained by their algorithm causes detail loss in the bright regions in the *Glass* image, since the total local contrast of every pixel is not considered. Song and Cosman's algorithm enhances the brightness of the output images significantly, resulting in saturated regions and detail loss in the *Lavender*, *Flower*, and *Glass* images, while darkening the output *BlueHouse* image in lower luminance regions. This saturation and detail loss occurs because their algorithm does not take into account the textual details of the image. Nguyen *et al.*'s algorithm produces similar TMCs for the images with different characteristics, which indicates that their algorithm is less adaptive since the guided TMC in their algorithm is image characteristic-independent. In contrast, the TMCs obtained by the proposed algorithm produce results adaptive to the image characteristics thereby minimizing the saturated region.

We also conducted a subjective evaluation of the proposed algorithm against the HDR10 standard [31], [33]. 13 subjects participated in this evaluation, and they rated the quality of each converted image on the SDR display in comparison with the original image on the HDR display on a five-point Likert scale [45] in terms of both bright preservation and local contrast. Table 1 compares the average scores on 13 randomly selected test images. The proposed algorithm gets significantly higher scores than the HDR10 standard on both criteria. This is because the proposed algorithm takes the HVS model-based global and local contrast of the input image into account by employing the 2D histogram and exploiting the characteristics of the display devices on which the tone-mapped images are presented.

TABLE 1. Mean subjective evaluation scores on the brightness preservation and the local contrast of the HDR10 standard [31], [33] and the proposed algorithm. Boldface numbers denote the highest scores for each criterion.

	HDR10	Proposed
Brightness	3.00	**4.01**
Local contrast	3.00	**4.02**

C. OBJECTIVE EVALUATION

In addition to the subjective evaluation, we also compare the results of the tone-mapping algorithms using four objective quality metrics: the tone-mapped image quality index (TMQI) [46], feature similarity index for tone-mapped images (FSITM) [47], visual difference predictor for HDR images (HDR-VDP) [48], [49], and dynamic range independent image quality assessment metric (DRI-IQM) [50]. To this end, we first convert the pixel values of the original and tone-mapped images into the absolute luminance maps using the standard EOTF [30], [31], and then compute the objective metrics between the luminance maps.

Table 2 lists the average values of the first three metrics obtained by the different algorithms from 54 test images. For each metric, the best results are indicated in bold. First, TMQI quantifies structural fidelity and naturalness of a tone-mapped image using a statistical naturalness model built upon various types of natural scenes. A high TMQI score demonstrates that the tone-mapped image has good structural details and is natural-looking. In this test, Song and Cosman's algorithm and Nguyen *et al.*'s algorithm yield the highest scores, whereas the proposed algorithm ranks third. However, the differences in TMQI are insignificant among these algorithms, which all yield comparable results according to the TMQI scores. Second, FSITM evaluates the similarity of a tone-mapped image based on the local phase similarity with respect to the HDR image. A high FSITM score indicates a high degree of structural similarity between the images in the pair. Mantiuk *et al.*'s algorithm and the proposed algorithm produces the highest scores since they attempt to minimize the contrast distortion between the HDR image and its tone-mapped image. Third, HDR-VDP, which is derived from the contrast sensitivity, estimates the probability with which a human observer can detect differences between two images in a pair. HDR-VDP provides higher scores for the image pairs that have lower probabilities of visual detection of differences. The proposed algorithm outperforms all the conventional algorithms in terms of the HDR-VDP since the perceptual differences between the input HDR image and tone-mapped image are minimized effectively.

Figure 10 compares the results of the DRI-IQM [50] assessment for the test images in Figures 5–8. DRI-IQM predicts the perceptual differences between the reference and query images and generates a distortion map that shows the loss, amplification, and reversal of visible contrast in the query image. As can be seen from the results, the conventional algorithms produce tone-mapped images with contrast alteration compared to the reference images. For example,

TABLE 2. Objective evaluation results of the tone-mapping algorithms using three metrics: TMQI [46], FSITM [47], and HDR-VDP [48], [49]. Boldface numbers denote the highest scores for each metric.

	TMQI	FSITM	HDR-VDP
Mantiuk *et al.* [12]	0.77	**0.95**	70.66
Song and Cosman [15]	**0.79**	0.91	70.28
Nguyen *et al.* [16]	**0.79**	0.94	75.12
Proposed	0.78	**0.95**	**76.45**

the results show that Mantiuk *et al.*'s algorithm causes contrast loss in the bright regions and amplifies the contrast along the edges of the *Glass* image since the local contrast is not taken into account effectively. Song and Cosman's algorithm experiences the same problem in the *Glass* image. Further, the results for the *Flower* image exhibit notable contrast loss. This loss occurs because their algorithm uses a 1D histogram that cannot consider the local detail of an image. Moreover, the tone-mapped images obtained by Nguyen *et al.*'s algorithm also exhibit visible contrast loss, *e.g.*, the bright regions of the *Lavender* and *BlueHouse* images, since the guided TMC that is used to preserve the absolute luminance does not consider the global contrast of the tone-mapped image. On the contrary, the proposed algorithm provides images with less contrast distortion by minimizing both the local and global contrast differences with respect to the reference HDR images.

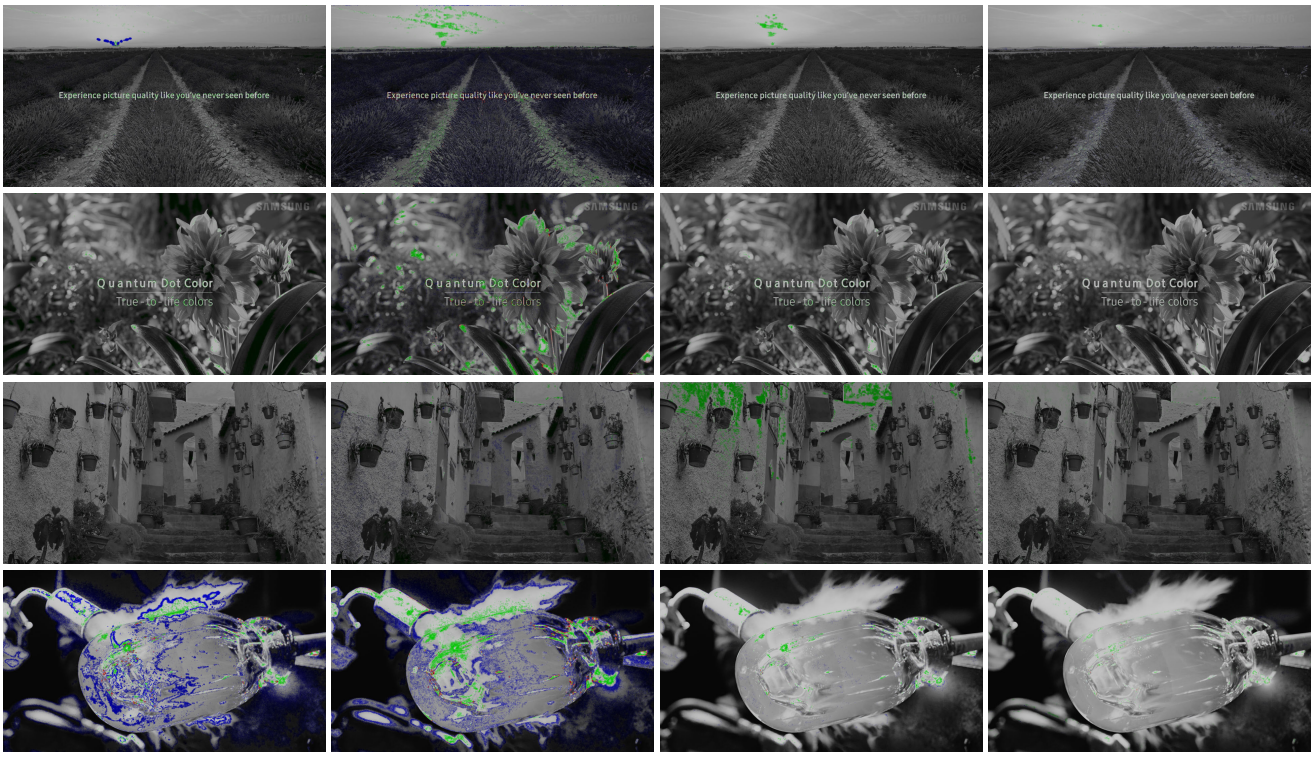
D. IMPACTS OF α ON TONE MAPPING

As discussed in Section III, the proposed algorithm performs local and global contrast preservation simultaneously, and the parameter α in (22) controls the relative importance of these types of preservation. In Figure 11, we show the tone-mapped *Motor* and *Flower* images obtained using the proposed algorithm with three different values of α. When $\alpha = 0$, only the global contrast-preserving term is included in the proposed algorithm. As α increases, the proposed algorithm further enhances the local contrast of the tone-mapped images. For example, the local details of the dust regions in the *Motor* image and the petals regions in the *Flower* image are increased noticeably in Figures 11(a)–(c), in proportion to the increase in α. Therefore, α can be selected adaptively depending on the application or user preferences.

E. COMPUTATION TIMES

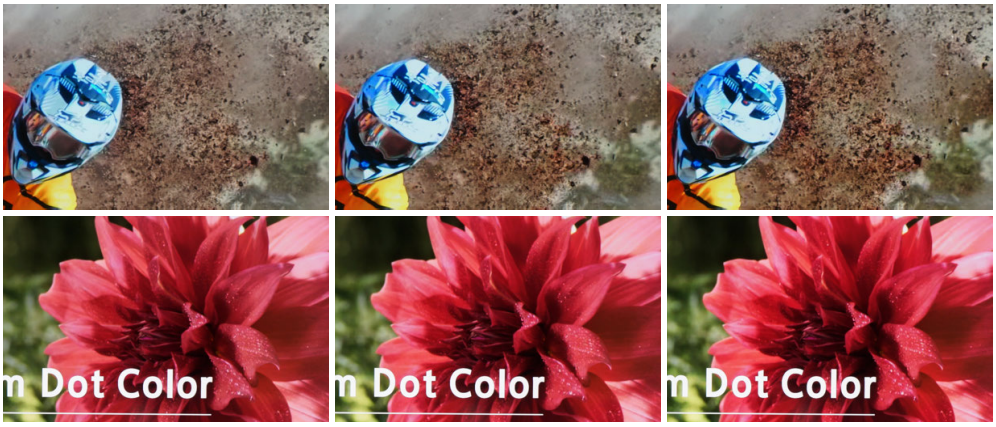
Table 3 compares the average execution times obtained by applying the different tone-mapping algorithms to ten images of resolution 3840×2160. In this test, we use a PC with a 3.6 GHz CPU and 32 GB RAM. Except for Mantiuk *et al.*'s algorithm, which is developed using C++, all the other tone-mapping algorithms were implemented using MATLAB without code optimization.

Song and Cosman's algorithm is the most efficient in terms of computation time due to its simplicity. However, it fails to preserve the perceptual quality on the target SDR display. Aside from this algorithm, the proposed algorithm is the most efficient. It should also be noticed that the proposed



(a) (b) (c) (d)

FIGURE 10. DRI-IQM assessment on the *Lavender* (top row), *Flower* (second row), *BlueHouse* (third row), and *Glass* (bottom row) images. The colormap represents the predicted visible distortions between reference HDR images and tone-mapped images obtained by (a) Mantiuk *et al.*'s algorithm [12], (b) Song and Cosman's algorithm [15], (c) Nguyen *et al.*'s algorithm [16], and (d) the proposed algorithm. The colors green, blue, and red indicate loss of visible contrast, amplification of invisible contrast, and reversal of visible contrast, respectively.



(a) (b) (c)

FIGURE 11. Tone-mapped *Motor* (top row) and *Flower* (bottom row) images for various values of α. (a) $\alpha = 0$, (b) $\alpha = 0.5$, and (c) $\alpha = 1$.

TABLE 3. The average computation times in seconds of Mantiuk *et al.*'s algorithm [12], Song and Cosman's algorithm [15], Nguyen *et al.*'s algorithm [16], and the proposed algorithm for the test images.

Mantiuk *et al.*	Song and Cosman	Nguyen *et al.*	Proposed
16.2	3.0	14.9	13.8

algorithm is implemented in MATLAB without code optimization; the computation performance could be further enhanced by using low-level languages such as C/C++. The use of parallel computing could also reduce the execution time.

V. CONCLUSION

We have proposed an optimized HVS model-based tone-mapping algorithm for displaying HDR10 image content on conventional SDR devices. We first quantified the perceptual quality of an image when it was displayed on a device using a HVS model-based 2D histogram. Then, we formulated an optimization problem to minimize the perceptual quality difference between an input HDR and its tone-mapped image in terms of local and global contrast. Finally, we obtained the optimal TMC by solving the optimization problem using the ADMM approach. Experimental results demonstrated

that the proposed tone-mapping algorithm provides a higher image quality than other conventional algorithms on actual displays. One important direction for future work is to more accurately quantify the perceptual contrast by considering ambient illumination [15], [51], [52].

APPENDIX
DERIVATION OF THE CLOSED-FORM SOLUTION $\mathbf{H}_T^{(k+1)}$ IN (24)

In (23), we define $f(\mathbf{H}_T)$ as

$$\begin{aligned} f(\mathbf{H}_T) = {} & \alpha\|\mathbf{H}\circ\mathbf{R}_l - \mathbf{H}_T\circ\mathbf{R}_{l,T}\|_F^2 \\ & + (1-\alpha)\|\mathbf{H}\circ\mathbf{R}_g - \mathbf{H}_T\circ\mathbf{R}_{g,T}\|_F^2 \\ & + \beta\|\mathbf{H}_T\mathbf{D}\|_F^2. \end{aligned} \tag{30}$$

Using the trace of a matrix $\mathrm{tr}(\cdot)$, $f(\mathbf{H}_T)$ can be rewritten as

$$\begin{aligned} & f(\mathbf{H}_T) \\ & = \alpha\mathrm{tr}\Big((\mathbf{H}\circ\mathbf{R}_l - \mathbf{H}_T\circ\mathbf{R}_{l,T})(\mathbf{H}\circ\mathbf{R}_l - \mathbf{H}_T\circ\mathbf{R}_{l,T})^T\Big) \\ & \quad + (1-\alpha)\mathrm{tr}\big((\mathbf{H}\circ\mathbf{G}_l - \mathbf{H}_T\circ\mathbf{R}_{g,T}) \\ & \quad \times(\mathbf{H}\circ\mathbf{R}_g - \mathbf{H}_T\circ\mathbf{R}_{g,T})^T\big) + \beta\mathrm{tr}\big(\mathbf{H}_T\mathbf{D}(\mathbf{H}_T\mathbf{D})^T\big). \end{aligned} \tag{31}$$

Then, the target 2D histogram $\mathbf{H}_T$ is obtained by solving

$$\nabla_{\mathbf{H}_T} f(\mathbf{H}_T) = \mathbf{0} \tag{32}$$

where $\nabla_{\mathbf{H}_T} f(\mathbf{H}_T)$ is the Jacobian matrix, and $\mathbf{0}$ denotes the matrix with all zero elements. The Jacobian matrix is obtained by [53]

$$\begin{aligned} & \nabla_{\mathbf{H}_T} f(\mathbf{H}_T) \\ & = \alpha\mathbf{R}_{l,T}\circ(2\mathbf{H}_T\circ\mathbf{R}_{l,T} - \mathbf{H}\circ\mathbf{R}_L - \mathbf{H}^T\circ\mathbf{R}_l^T) \\ & \quad + 2(1-\alpha)\mathbf{R}_{g,T}\circ(2\mathbf{H}_T\circ\mathbf{R}_{g,T} - \mathbf{H}\circ\mathbf{R}_g - \mathbf{H}^T\circ\mathbf{R}_g^T) \\ & \quad + 2\beta\mathbf{H}_T\mathbf{D}\mathbf{D}^T \\ & = 2(\alpha\mathbf{R}_{l,T}\circ\mathbf{R}_{l,T} + (1-\alpha)\mathbf{R}_{g,T}\circ\mathbf{R}_{g,T})\circ\mathbf{H}_T \\ & \quad - \alpha\mathbf{R}_{l,T}\circ(\mathbf{H}\circ\mathbf{R}_L + \mathbf{H}^T\circ\mathbf{R}_l^T) \\ & \quad - (1-\alpha)\mathbf{R}_{g,T}\circ(\mathbf{H}\circ\mathbf{R}_G + \mathbf{H}^T\circ\mathbf{R}_g^T) + 2\beta\mathbf{H}_T\mathbf{D}\mathbf{D}^T \\ & = \mathbf{A}\circ\mathbf{H}_T + \mathbf{H}_T\mathbf{B} - \mathbf{C} \end{aligned} \tag{33}$$

where $\mathbf{A} = 2(\alpha\mathbf{R}_{l,T}\circ\mathbf{R}_{l,T} + (1-\alpha)\mathbf{R}_{g,T}\circ\mathbf{R}_{g,T})$, $\mathbf{B} = 2\beta\mathbf{D}\mathbf{D}^T$, and $\mathbf{C} = \alpha\mathbf{R}_{l,T}\circ(\mathbf{H}\circ\mathbf{R}_l + \mathbf{H}^T\circ\mathbf{R}_l^T) + (1-\alpha)\mathbf{R}_{g,T}\circ(\mathbf{H}\circ\mathbf{R}_g + \mathbf{H}^T\circ\mathbf{R}_g^T)$.

From (33), the problem in (32) can be written as

$$\mathbf{C} = \mathbf{A}\circ\mathbf{H}_T + \mathbf{H}_T\mathbf{B}. \tag{34}$$

Using the vectorization operation $\mathrm{vec}(\cdot)$, the problem in (34) is expanded as

$$\begin{aligned} \mathrm{vec}(\mathbf{C}) & = \mathrm{vec}(\mathbf{A}\circ\mathbf{H}_T) + \mathrm{vec}(\mathbf{H}_T\mathbf{B}) \\ & = \mathrm{vec}(\mathbf{A})\circ\mathrm{vec}(\mathbf{H}_T) + \mathrm{vec}(\mathbf{I}\mathbf{H}_T\mathbf{B}) \\ & = \mathrm{diag}(\mathrm{vec}(\mathbf{A}))\mathrm{vec}(\mathbf{H}_T) + (\mathbf{B}^T\otimes\mathbf{I})\mathrm{vec}(\mathbf{H}_T) \\ & = \big(\mathrm{diag}(\mathrm{vec}(\mathbf{A})) + (\mathbf{B}^T\otimes\mathbf{I})\big)\mathrm{vec}(\mathbf{H}_T) \end{aligned} \tag{35}$$

where $\mathbf{I}$ and $\otimes$ denote the identity matrix and the Kronecker product, respectively. Thus, we have

$$\mathrm{vec}(\mathbf{H}_T) = \big(\mathrm{diag}(\mathrm{vec}(\mathbf{A})) + (\mathbf{B}^T\otimes\mathbf{I})\big)^{-1}\mathrm{vec}(\mathbf{C}). \tag{36}$$

Finally, let vec^{-1} denote the inverse of the vectorization operator, then the closed-form solution $\mathbf{H}_T$ is obtained by

$$\mathbf{H}_T = \mathrm{vec}^{-1}\Big(\big(\mathrm{diag}(\mathrm{vec}(\mathbf{A})) + (\mathbf{B}^T\otimes\mathbf{I})\big)^{-1}\mathrm{vec}(\mathbf{C})\Big). \tag{37}$$

REFERENCES

[1] *Ultra High Definition Television—Image Parameter Values for Program Production*, document SMPTE ST 2036-1:2014, 2014.
[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. San Mateo, CA, USA: Morgan Kaufmann, 2010.
[3] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*, 1st ed. Natick, MA, USA: A. K. Peters, 2011.
[4] S. Mann and R. W. Picard, "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures," in *Proc. 46th Ann. IS&T Conf.*, May 1995, pp. 442–448.
[5] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH*, Aug. 1997, pp. 369–378.
[6] E. Reuss and L. Borg, "Camera raw for high dynamic range workflows," *SMPTE Motion Imag. J.*, vol. 125, no. 5, pp. 28–34, Jul. 2016.
[7] J. Keinert, M. Schöberl, M. Ziegler, F. Zilly, and S. Foessel, "High-dynamic range video cameras based on single shot non-regular sampling," *SMPTE Motion Imag. J.*, vol. 123, no. 8, pp. 49–54, Nov. 2014.
[8] *Common File Format & Media Formats Specification Version 2.1*, Digit. Entertainment Content Ecosyst. (DECE), Little Rock, AR, USA, 2015.
[9] *Parameter Values for Ultra-High Definition Television Systems for Production and International Programme Exchange*, ITU-R BT.2020-2, 2015.
[10] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 419–426, Sep. 2003.
[11] 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.
[12] 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.
[13] J. Ok and C. Lee, "HDR tone mapping algorithm based on difference compression with adaptive reference values," *J. Vis. Commun. Image Represent.*, vol. 43, pp. 61–76, Feb. 2017.
[14] K.-F. Yang, H. Li, H. Kuang, C.-Y. Li, and Y.-J. Li, "An adaptive method for image dynamic range adjustment," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 3, pp. 640–652, Mar. 2019.
[15] Q. Song and P. C. Cosman, "Luminance enhancement and detail preservation of images and videos adapted to ambient illumination," *IEEE Trans. Image Process.*, vol. 27, no. 10, pp. 4901–4915, Oct. 2018.
[16] N. H. Nguyen, T. Van Vo, Y. Jeong, Y. Moon, and C. Lee, "Optimized tone mapping of HDR images via HVS model-based 2D histogram equalization," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2018, pp. 700–704.
[17] 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.
[18] L. Meylan and S. Süsstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2820–2830, Sep. 2006.
[19] 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.
[20] 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.
[21] E. Mezeni and L. V. Saranovac, "Enhanced local tone mapping for detail preserving reproduction of high dynamic range images," *J. Vis. Commun. Image Represent.*, vol. 53, pp. 122–133, May 2018.

[22] X. Shu and X. Wu, "Locally adaptive rank-constrained optimal tone mapping," *ACM Trans. Graph.*, vol. 37, no. 3, pp. 38:1–38:10, Jul. 2018.
[23] N. Zhang, C. Wang, Y. Zhao, and R. Wang, "Deep tone mapping network in HSV color space," in *Proc. IEEE Vis. Commun. Image Process. (VCIP)*, Dec. 2019, pp. 1–4.
[24] R. Montulet and A. Briassouli, "Deep learning for robust end-to-end tone mapping," in *Proc. Brit. Mach. Vis. Conf. (BMVC)*, Sep. 2019, pp. 1–10.
[25] A. Rana, P. Singh, G. Valenzise, F. Dufaux, N. Komodakis, and A. Smolic, "Deep tone mapping operator for high dynamic range images," *IEEE Trans. Image Process.*, vol. 29, pp. 1285–1298, 2020.
[26] K. Panetta, L. Kezebou, V. Oludare, S. Agaian, and Z. Xia, "TMO-Net: A parameter-free tone mapping operator using generative adversarial network, and performance benchmarking on large scale HDR dataset," *IEEE Access*, vol. 9, pp. 39500–39517, 2021.
[27] H.-Y. Kim, S. Park, Y.-G. Shin, S.-W. Jung, and S.-J. Ko, "Detail restoration and tone mapping networks for X-ray security inspection," *IEEE Access*, vol. 8, pp. 197473–197483, 2020.
[28] T. Celik and T. Tjahjadi, "Contextual and variational contrast enhancement," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3431–3441, Dec. 2011.
[29] T. Celik, "Two-dimensional histogram equalization and contrast enhancement," *Pattern Recognit.*, vol. 45, no. 10, pp. 3810–3824, Oct. 2012.
[30] 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.
[31] *High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays*, SMPTE ST 2084:2014, 2014.
[32] *Image Parameter Values for High Dynamic Range Television for Use in Production and International Programme Exchange*, ITU-R BT.2100-1, 2017.
[33] *Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images*, SMPTE ST 2086:2014, 2014.
[34] T. Smith and J. Guild, "The C.I.E. Colorimetric standards and their use," *Trans. Opt. Soc.*, vol. 33, no. 3, pp. 73–134, Jan. 1931.
[35] *High Efficiency Video Coding*, ITU-T H.265, 2018.
[36] P. G. J. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality*. Eindhoven, The Netherlands: Technische Universiteit Eindhoven, 1999.
[37] A. K. Jain, *Fundamentals of Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, 1989.
[38] P. G. J. Barten, "Formula for the contrast sensitivity of the human eye," in *Proc. SPIE*, vol. 5294, pp. 231–238, Dec. 2003.
[39] *The Present State of Ultra-high Definition Television*, SMPTE ST 2246-6:2017, 2017.
[40] B. Wandell, *Foundations of Vision*. Sunderland, MA, USA: Sinauer Associates, 1995.
[41] 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.
[42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.
[43] L. Krasula, K. Fliegel, P. L. Callet, and M. Klíma, "Objective evaluation of naturalness, contrast, and colorfulness of tone-mapped images," in *Proc. SPIE, Appl. Digital Image Process.*, vol. 9217, Aug. 2014, pp. 600–609.
[44] M. Čadík, M. Wimmer, L. Neumann, and A. Artusi, "Evaluation of HDR tone mapping methods using essential perceptual attributes," *Comput. Graph.*, vol. 32, no. 3, pp. 330–349, Jun. 2008.
[45] R. Likert, "A technique for the measurement of attitudes," *Arch. Psychol.*, vol. 26, no. 40, pp. 5–55, Jun. 1932.
[46] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb. 2013.
[47] H. Z. Nafchi, A. Shahkolaei, R. F. Moghaddam, and M. Cheriet, "FSITM: A feature similarity index for tone-mapped images," *IEEE Signal Process. Lett.*, vol. 22, no. 8, pp. 1026–1029, Aug. 2015.
[48] R. Mantiuk, K. J. Kim, A. G. Rempel, and W. Heidrich, "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.*, vol. 30, no. 4, pp. 1–14, Jul. 2011.
[49] M. Narwaria, R. K. Mantiuk, M. P. D. Silva, and P. L. Callet, "HDR-VDP-2.2: A calibrated method for objective quality prediction of high-dynamic range and standard images," *J. Electron. Imag.*, vol. 24, no. 1, Jan. 2015, Art. no. 010501.
[50] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 69:1–69:10, Aug. 2008.
[51] Y. Zhang, D. Agrafiotis, M. Naccari, M. Mrak, and D. R. Bull, "Visual masking phenomena with high dynamic range content," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2013, pp. 2284–2288.
[52] Y. Zhang, D. Agrafiotis, and D. R. Bull, "High dynamic range content calibration for accurate acquisition and display," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4703–4707.
[53] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Lyngby, Denmark: Technical Univ. of Denmark, 2008.

NAM HOANG NGUYEN received the B.S. degree in electronics and telecommunication from University of Science, Vietnam National University, Ho Chi Minh City, Vietnam, in 2012, and the M.S. degree in computer engineering from Pukyong National University, Busan, South Korea, in 2019.

In 2019, he was a Researcher with the Department of Multimedia Engineering, Dongguk University, Seoul, South Korea. He is currently a Machine Learning Engineer with Cybercore Co. Ltd., Ho Chi Minh City, where he is developing computer vision algorithms. His current research interests include image processing and computer vision, especially high dynamic range imaging and deep learning-based computer vision.

TU VAN VO (Student Member, IEEE) received the B.S. degree in electrical engineering from Hanoi University of Science and Technology, Hanoi, Vietnam, in 2016, and the M.S. degree in computer engineering from Pukyong National University, Busan, South Korea, in 2019.

He is currently with InhandPlus Inc., Seoul, South Korea, where he is developing computer vision algorithms for medication behavior analysis. His current research interests include image processing and computer vision, especially high dynamic range imaging and deep learning-based computer vision.

CHUL LEE (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2003, 2008, and 2013, respectively.

He was with Biospace Inc., Seoul, from 2002 to 2006, where he was involved in the development of medical equipment. From 2013 to 2014, he was a Postdoctoral Scholar with the Department of Electrical Engineering, Pennsylvania State University, University Park, PA, USA. From 2014 to 2015, he was a Research Scientist with the Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong. From 2015 to 2019, he was an Assistant Professor with the Department of Computer Engineering, Pukyong National University, Busan, South Korea. In March 2019, he joined the Department of Multimedia Engineering, Dongguk University, Seoul, where he is currently an Assistant Professor. His current research interests include image processing and computational imaging with an emphasis on restoration and high dynamic range imaging.

Dr. Lee received the Best Paper Award from the *Journal of Visual Communication and Image Representation*, in 2014. He is currently an Editorial Board Member of the *Journal of Visual Communication and Image Representation*.