IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 6, 2018, accepted June 8, 2018, date of publication June 13, 2018, date of current version July 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2846772

Non-Iterative Tone Mapping With High Efficiency and Robustness

GYUJIN BAE[1], (Student Member, IEEE), CHAN YOUNG JANG[2], (Student Member, IEEE), SUNG IN CHO[3], (Member, IEEE), AND YOUNG HWAN KIM[1], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Pohang University of Science and Technology, Pohang 37673, South Korea
[2]Department of System LSI, Samsung Electronics Co. Ltd., Hwaseong 18448, South Korea
[3]Department of Electronic Engineering, College of Information and Communication Engineering, Daegu University, Gyeongsan 38453, South Korea

Corresponding author: Young Hwan Kim (youngk@postech.ac.kr)

This work was supported by the Ministry of Science and ICT, South Korea, through the Institute for Information and Communications Technology Promotion, the ICT Consilience Creative Program, under Grant IITP-2018-2011-1-00783.

ABSTRACT This paper proposes an efficient approach for tone mapping, which provides a high perceptual image quality for diverse scenes. Most existing methods, optimizing images for the perceptual model, use an iterative process and this process is time consuming. To solve this problem, we proposed a new layer-based non-iterative approach to finding an optimal detail layer for generating a tone-mapped image. The proposed method consists of the following three steps. First, an image is decomposed into a base layer and a detail layer to separate the illumination and detail components. Next, the base layer is globally compressed by applying the statistical naturalness model based on the statistics of the luminance and contrast in the natural scenes. The detail layer is locally optimized based on the structure fidelity measure, representing the degree of local structural detail preservation. Finally, the proposed method constructs the final tone-mapped image by combining the resultant layers. The performance evaluation reveals that the proposed method outperforms the benchmarking methods for almost all the benchmarking test images. Specifically, the proposed method improves an average tone mapping quality index-II (TMQI-II), a feature similarity index for tone-mapped images (FSITM), and a high-dynamic range-visible difference predictor (HDR-VDP)-2.2 by up to 0.651 (223.4%), 0.088 (11.5%), and 10.371 (25.2%), respectively, compared with the benchmarking methods, whereas it improves the processing speed by over 2611 times. Furthermore, the proposed method decreases the standard deviations of TMQI-II, FSITM, and HDR-VDP-2.2, and processing time by up to 81.4%, 18.9%, 12.6%, and 99.9%, respectively, when compared with the benchmarking methods.

INDEX TERMS Tone mapping, high-dynamic range compression, guided filter, statistical model, structure fidelity.

I. INTRODUCTION

The luminance of a real-world scene ranges from 10^{-6} cd/m^2 (extreme darkness) to 10^8 cd/m^2 (bright sunshine). Hence, a standard 8-bit display device is unable to display the high-dynamic range (HDR) images from the real world with satisfactory viewing quality. To overcome this problem, several techniques to generate the HDR-like low-dynamic range (LDR) images from the real-world scenes have been proposed. The HDR-like LDR images are expected to reconcile smoothly the HDR of luminance levels in natural scenes with the LDR of standard display devices.

One approach to generate the HDR-like LDR images is to acquire HDR images from natural scenes and convert them into LDR images. The process of converting HDR images to LDR images with naturalness is called tone mapping. Another popular approach to generate the HDR-like LDR images from the real world is the multi-exposure image fusion [1]–[4]. This approach captures multi-exposure LDR images and directly converts them into a high-quality LDR image without obtaining an HDR image. While both approaches are promising, this paper focuses on the tone mapping methods of the first approach.

Generally, the HDR images are acquired using the camera with HDR sensors or by combining multiple LDR images taken with different exposure levels [5], [6]. Then, the process of tone mapping is applied to the HDR images. Tone mapping methods are grouped into two categories: global tone mapping (spatially-uniform tone mapping) and local

tone mapping (spatially-varying tone mapping). Global tone mapping methods [7]–[12] compress the dynamic range of the entire input image using a point-to-point matching function based on logarithmic, gamma, or sigmoidal functions. These methods have a low computational complexity and preserve the intensity orders of the original scenes; thus, avoiding halo artifacts [11]. However, these methods have a limitation in preserving the details of an HDR image [13].

Numerous local tone mapping methods have been proposed to preserve the detail and to enhance the local contrast [14]–[32]. The local tone mapping method uses a different transformation function adaptively to the local characteristics of the image. This local processing is a better imitation of a human visual system, which locally adapts to each part of a scene to form a percept where all details are visible [24]. Retinex-based methods are one of the prominent tone mapping methods [14]–[24]. These methods decompose an image into an illumination image and a reflectance image. The illumination image is assumed to include the low-frequency components, whereas the reflectance image corresponds to the high-frequency components. The illumination image is estimated by using a Gaussian filter, and the reflectance image is obtained by dividing the input image by the illumination image. This produces severe halo artifacts in the resulting images [23]. To avoid the halo artifacts, Meylan and Susstrunk [24] proposed using an edge-preserving filter. In contrast with an isotropic filter, this filter considered the directionality of the edges, and thus, it effectively removed the halo artifacts.

Edge-preserving filtering has become an important factor in filter design to avoid halo artifacts. Durand and Dorsey [25] introduced a layer-based method to display HDR images. They decomposed an image into a base layer and a detail layer using a bilateral filter. A base layer, containing low-frequency components, is typically computed by applying an edge-preserving smoothing operator to the image. The detail layer is then defined as the difference between the original image and the base layer. Next, each layer is multiplied by some scaling factors, and recombined into a single image to obtain the final image. Farbman *et al.* [26] proposed the global optimization-based edge-preserving smoothing technique to decompose the luminance of an HDR image into a base layer and a detail layer. A global optimization-based approach often yields excellent image quality, but it requires a high computational cost. He *et al.* [27] proposed a guided filter that could transfer the structure of a guidance image to the filter output. Moreover, the guided filter had a naturally fast and non-approximate linear time algorithm, regardless of the kernel size and intensity range. However, such methods often produce over-exaggerated and unnatural visual effects. This is because they contain one or more parameters whose optimal values are often image-dependent. Thus, it is difficult to guarantee a consistent image quality for various images.

To solve this problem, numerous optimization methods have been studied for tone mapping. Generally, there are two major research topics on optimization. The first topic involves developing the image quality assessment model as an objective function. The second topic is concerned about efficiently obtaining optimized images for the objective function. Shan *et al.* [28] proposed a window-based tone mapping method that performed local linear adjustments on small overlapping windows over an entire HDR image. Then, the parameter values for the linear adjustments were determined by solving the optimization problem to preserve the local contrast. This method preserves the local structures, including the sharp edges. However, it often suffers from over-enhancement and produces unnatural results. Yeganech and Wang [29] developed the tone mapping quality index (TMQI) to assess the quality of tone-mapped images. The associated model was based on combining a structural similarity index measure (SSIM)-motivated structural fidelity measure with a statistical naturalness model. The TMQI clearly provided a better assessment result for tone-mapped images than the well-known image quality assessment metrics [30]. However, the TMQI metric has some limitations. The structural fidelity measure often incorrectly predicts the contrast strength in a homogeneous area owing to its oversensitivity to noise. The statistical naturalness model does not work well on images with various image contents. This is because the model does not consider that each image has a different luminance level to appear natural depending on its content. To overcome these limitations, Ma *et al.* [31] proposed an improved TMQI, namely, TMQI-II. For the structural fidelity measure, a contrast viability model adapted to the local luminance levels was proposed to solve the problems caused by noise. For the statistical naturalness model, an image-dependent statistical naturalness model was suggested to better quantify the naturalness of the tone-mapped images. In addition, an iterative tone mapping method based on TMQI-II was proposed. In each iteration, it alternately improved the structural fidelity and statistical naturalness of the resulting image. Therefore, the method produced good quality tone-mapped images. However, the method had a high computational complexity. Moreover, because the number of iterative operations required for each input image was different, the processing time for each input image was extremely variable. Lappra *et al.* [32] proposed the perceptually optimized image rendering method. It first constructs a constrained optimization problem, which optimizes the rendered image to minimize perceptual difference from the original HDR image. Then, it uses the Normalized Laplacian Pyramid (NLP) [33] to quantify the perceptual difference. Next, it finds iteratively an optimized image using a gradient-based method. The Lappra's method has the advantage of minimizing the loss of contrast between the tone-mapped image and the input HDR image. However, it sometimes induces spatial distortion caused by over-enhanced noise. Furthermore, it has a high computational complexity due to the iterative process.

As described above, numerous studies have been conducted to develop a tone mapping method that provides

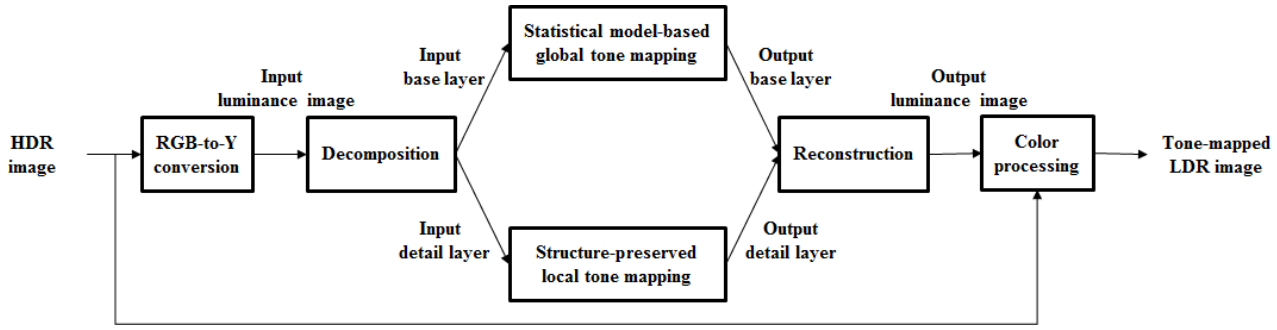


FIGURE 1. Overall block diagram of the proposed tone mapping method.

a high perceptual image quality. These particularly focused on the research to provide consistent results for all types of HDR images. Although the methods were able to provide a high image quality, the high computational complexity of the methods was the bottleneck in real-time applications.

In this paper, we propose an efficient tone mapping algorithm that provides a high perceptual image quality with a low computational complexity consistently for scenes with diverse characteristics. In the proposed method, to reduce the computational complexity, perceptual models are applied to the illumination and detail components independently. For the preservation of the scene naturalness, the proposed method globally compresses the illumination components using an image-dependent statistical naturalness model. In addition, the proposed method locally enhances the detail components, considering the structural similarity with an HDR image for the preservation of the structure.

The remainder of this paper is organized as follows. In Section II, we describe the proposed algorithm. In Section III, we present the experimental results and evaluate the performance of the proposed method. Finally, we conclude this study in Section IV.

II. PROPOSED METHOD

The block diagram of the proposed method is illustrated in Fig. 1. First, to decompose an image into illumination, detail, and color components, the proposed method converts the input RGB image into a luminance image using the RGB-to-YCbCr color conversion [34]. The computed luminance image is decomposed into a base layer and a detail layer. Next, to obtain a structure-preserved image without artifacts, the base layer is globally compressed based on an image-dependent statistical naturalness model. The detail layer is locally optimized by balancing the naturalness and contrast preservation. The resultant layers are recombined to yield the final luminance result. Finally, the final tone-mapped image is obtained by applying color processing that preserves the ratio between the color channels.

A. DECOMPOSITION AND RECONSTRUCTION

The decomposition step is to divide an image into a base layer and a detail layer, representing the illumination and detail image, respectively. The base layer consists of low-frequency components, thereby determining the overall brightness of a scene. In contrast, the detail layer is composed of high-frequency components such as detail components. The base layer is obtained using an edge-preserving filter, and the detail layer is computed by subtracting the base layer from the input luminance image in the logarithmic scale. As an edge-preserving filter, we use the guided filter proposed in [27], which provides a good edge-preserving property and fast filtering speed. The decomposition equation is defined as follows:

$$\begin{aligned} g_{in}(x,y) &= \log(L_{in}(x,y)), \\ \log(B_{in}(x,y)) &= G(g_{in}(x,y)), \\ \log(D_{in}(x,y)) &= g_{in}(x,y) - \log(B_{in}(x,y)), \end{aligned} \tag{1}$$

where, $L_{in}(x, y)$ is the luminance value of an input luminance image at location (x, y), $B_{in}(x, y)$ and $D_{in}(x, y)$ represent the luminance values of the input base and detail layer at location (x, y), respectively, $g_{in}(x, y)$ denotes the log-luminance value of the input image at location (x, y), and G denotes the edge-preserving smoothing operator.

The reconstruction process is the inverse process of the decomposition process. The luminance image is easily obtained by multiplying the pixel values of the base and detail layers.

B. STATISTICAL MODEL-BASED GLOBAL TONE MAPPING

Global tone mapping is performed considering the characteristics of the base layer. The base layer contains the low-frequency components. In contrast with the existing tone mapping methods in which the base layer is simply multiplied using a scale factor [25]–[27], the proposed method uses the following image-dependent statistical model [31] to enrich the naturalness of a scene. First, we calculate the log-average luminance, L_{avg}, of the scene as

$$L_{avg} = \exp\left(\frac{1}{N}\sum_{x,y}\log\varepsilon + L_{in}(x,y)\right), \tag{2}$$

where N is the total number of pixels in the image and ε is a positive constant to avoid instability. Next, the luminance

image is computed by

$$B^s(x, y) = \frac{h}{L_{avg}} B_{in}(x, y), \tag{3}$$

where h is a constant and its value is selected as 0.12, that is the value used in [31]. Finally, the output base layer is obtained by compressing the computed image. The output base layer, B_{out} is defined as

$$B_{out}(x, y) = \left(\frac{B^s(x, y)}{1 + B^s(x, y)} \right)^{\lambda}, \tag{4}$$

where the value of λ is selected as 2.2 [31]. The model is developed considering the correlation between the naturalness and average luminance. In addition, the model is proposed based on the logarithmic function. The use of a logarithmic function assumes that most of the structural detail in the HDR image is in a low dynamic range, and thus, it is reasonable to boost the lower luminance levels while compressing the higher luminance levels. Fig. 2(b) shows the output base layer. The base layer preserves the natural appearance well. However, the high luminance values are further compressed. This inevitably causes detail loss in the high-luminance areas. The detail loss can typically be captured by performing a local tone mapping in the detail layer.

(a) (b)

FIGURE 2. (a) input base layer (B_{in}) and (b) output base layer (B_{out}).

C. STRUCTURE-PRESERVED LOCAL TONE MAPPING

The detail layer is assumed to only contain high-frequency components. However, the detail layer actually contains some low-frequency components as well as high-frequency components because the decomposition process is a non-trivial problem. Thus, the proposed method generates a new detail layer considering both the local contrast and naturalness information. Specifically, in the smooth region, the proposed method should preserve the naturalness of a scene, whereas in the detail region, it should preserve the local contrast.

To obtain the optimal detail layer, we suppose that output detail layer D_{out} is linearly dependent on input detail layer D_{in} in a local patch because the pixels are highly correlated locally. The linear equation to obtain the optimal detail layer is as follows:

$$D_{out}(x, y) = k_{opt}(x, y) \cdot D_{in}(x, y), \tag{5}$$

where $k_{opt}(x, y)$ is the optimal coefficient value at location (x, y). It is calculated considering the characteristics of the luminance value in the local patch.

Three processes are performed to generate the detail layer that preserves the local structures without artifacts. First, a natural-looking detail layer is created that does not contain unnatural artifacts such as over-enhanced noise. The natural-looking detail layer is obtained based on the following two considerations. First, the luminance image which is globally compressed using the statistical naturalness model has high naturalness [31]. Second, the output luminance image obtained by combining the base and detail layers should be looked natural. The natural-looking detail layer is calculated as follows:

$$\begin{aligned} L_N(x, y) &= \left(\frac{L^S(x, y)}{1 + L^S(x, y)} \right)^{\lambda}, \quad where, \\ L^S(x, y) &= \frac{h}{L_{avg}} \cdot L_{in}(x, y), \\ D_N(x, y) &= \frac{L_N(x, y)}{B_{out}(x, y)}, \\ k_N(x, y) &= \frac{D_N(x, y)}{D_{in}(x, y)}, \end{aligned} \tag{6}$$

where $L_N(x, y)$ is the luminance value of a natural-looking luminance image. The same formula was used to obtain the output base layer B_{out} using Eq. (4). $D_N(x, y)$ is the luminance value of a natural-looking detail layer at location (x, y) and $k_N(x, y)$ is the coefficient value of the detail layer being natural at location (x, y). As shown in Fig. 3, D_N exhibits a natural characteristic in a homogeneous area such as the walls and wood board in the middle part of the image. In contrast, the detail regions with high luminance values have a large loss of the details.

Second, a contrast-preserved detail layer is generated by considering the contrast similarity with an HDR image. The proposed method uses the contrast sensitivity function, which is widely used to analyze the contrast similarity of two different dynamic range images (LDR, HDR) [29], [31], [35], [36]. The function is defined as follows:

$$\begin{aligned} \widetilde{\sigma}(x, y) = F(v|\theta, \tau) &= \frac{1}{\theta\sqrt{2\pi}} \int_{-\infty}^{v} \exp\left[-\frac{(t - \tau)^2}{2\theta^2} \right] dt, \\ v &= \frac{\sigma(x, y)}{\mu(x, y)}, \end{aligned} \tag{7}$$

where $\widetilde{\sigma}(x, y)$ is the signal strength at location (x, y); τ is a threshold determined by the contrast sensitivity function [29], [31]; τ_H and τ_L are the threshold values in the HDR and LDR images, respectively, and are set as 0.06 and 2.6303, respectively; θ is the standard deviation of the normal distribution and is set as $\tau/3$ based on Crozier's law [35], [36]; $\mu(x, y)$ and $\sigma(x, y)$ denote the local mean and local standard deviation calculated in a patch centered at (x, y), respectively. Then, the patch size is set as 5×5. We can distinguish between the significant and insignificant signal strengths by passing local standard deviation σ through a nonlinear mapping. The nonlinear mapping is designed so that a significant signal strength is mapped to 1 and an insignificant signal

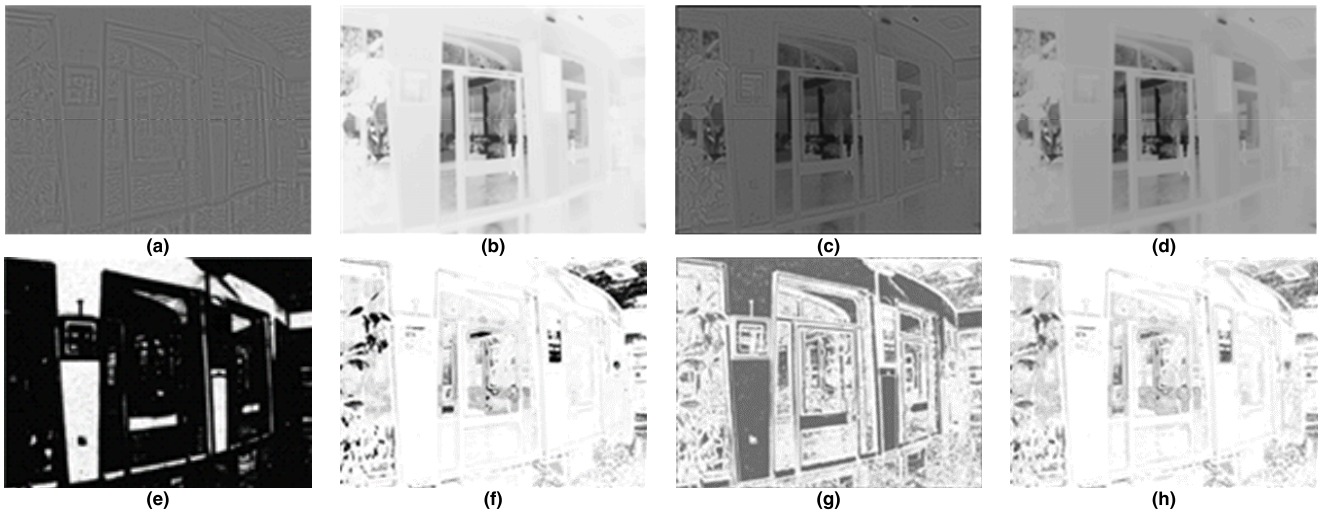


FIGURE 3. (a) input detail layer (D_{in}), (b) natural-looking detail layer (D_N), (c) normalized contrast-preserved detail layer (D_C^N), (d) output detail layer (D_{out}), (e) structure fidelity map of D_{in}, (f) structure fidelity map of D_N, (g) structure fidelity map of D_C^N, and (h) structure fidelity map of D_{out}.

strength is mapped to 0, with a smooth transition in-between. Furthermore, we can assume that the signal strength values of the tone-mapped LDR image within the patch are approximately the same as those of the HDR image. The mathematical condition to preserve the local contrast is described as follows:

$$\begin{aligned} &\widetilde{\sigma}_H(x,y) = \widetilde{\sigma}_{L_C}(x,y) \\ &where, \\ &L_C(x,y) = B_{out}(x,y)\cdot D_C(x,y), \\ &D_C(x,y) = k_C(x,y)\cdot D_{in}(x,y), \end{aligned} \tag{8}$$

where $L_C(x, y)$ and $D_C(x, y)$ are the luminance values of the contrast-preserved luminance image and contrast-preserved detail layer at location (x, y), respectively; $\widetilde{\sigma}_H(x, y)$ and $\widetilde{\sigma}_{L_C}(x, y)$ are the signal strengths of the HDR and LDR images, respectively. According to this condition, the standard deviation of the LDR image σ_{L_C} to preserve the local contrast should satisfy the following equation:

$$\sigma_{L_C}(x,y) = F^{-1}(\widetilde{\sigma}_H(x,y)|\theta_L,\tau_L). \tag{9}$$

Eq. (8) is rewritten as follows:

$$\begin{aligned} &\sigma^2_{L_C}(x,y) \\ &= \sigma_{A_2B_2}(x,y) + \mu_{A_2}(x,y)\cdot\mu_{B_2}(x,y) \\ &\quad - \left(\sigma_{A_1B_1}(x,y) + \mu_{A_1}(x,y)\cdot\mu_{B_1}(x,y)\right)^2 \\ &= k_C^2(x,y)\cdot\sigma_{A_2C_2}(x,y) + k_C^2(x,y)\cdot\mu_{A_2}(x,y)\cdot\mu_{C_2}(x,y) \\ &\quad - \left(k_C(x,y)\cdot\sigma_{A_1C_1}(x,y) + k_C(x,y)\cdot\mu_{A_1}(x,y)\cdot\mu_{C_1}(x,y)\right)^2 \\ &= k_C^2(x,y)\cdot\begin{pmatrix}\sigma_{A_2C_2}(x,y) + \mu_{A_2}(x,y)\cdot\mu_{C_2}(x,y) \\ -\left(\sigma_{A_1C_1}(x,y) + \mu_{A_1}(x,y)\cdot\mu_{C_1}(x,y)\right)^2\end{pmatrix}, \end{aligned}$$

where,

$$\begin{aligned} A_1 &= B_{out}, \quad A_2 = B_{out}^2, \\ B_1 &= k_C D_{in}, \quad B_2 = k_C^2 D_{in}^2, \\ C_1 &= D_{in}, \quad C_2 = D_{in}^2, \end{aligned} \tag{10}$$

where $\mu_A(x, y)$ and $\sigma_{AB}(x, y)$ denote the local mean and covariance between two corresponding patches centered at (x, y), respectively. Using (8)–(10), we can calculate the k_C value as follows:

$$k_C(x,y) = \sqrt{\frac{\sigma^2_{L_C}(x,y)}{\begin{pmatrix}\sigma_{A_2C_2}(x,y) + \mu_{A_2}(x,y)\cdot\mu_{C_2}(x,y) \\ -\left(\sigma_{A_1C_1}(x,y) + \mu_{A_1}(x,y)\cdot\mu_{C_1}(x,y)\right)^2\end{pmatrix}}}, \tag{11}$$

However, D_C is an over-enhanced image because it does not consider the brightness. Therefore, to maintain the local average brightness, we normalize the k_C values using the local mean value of each pixel.

$$\begin{aligned} k_C^N(x,y) &= k_C(x,y)\cdot\frac{\mu_N^{local}(x,y)}{\mu_C^{local}(x,y)}, \\ D_C^N(x,y) &= k_C^N(x,y)\cdot D_{in}(x,y), \\ L_C^N(x,y) &= B_{out}(x,y)\cdot D_C^N(x,y), \end{aligned} \tag{12}$$

where $\mu_N^{local}(x, y)$ and $\mu_C^{local}(x, y)$ represent the local mean of the D_N and D_C values calculated in a patch centered at (x, y), respectively; $D_C^N(x, y)$ and $L_C^N(x, y)$ is the luminance value of a normalized contrast-preserved detail layer and a normalized contrast-preserved luminance layer at location (x, y), respectively.

Finally, we compute the optimal detail layer by combining the natural-looking detail layer, D_N, and the normalized contrast-preserved detail layer, D_C^N, via an optimization process. To combine the two detail layers, we construct a constrained optimization problem, given in Eq. (13), and use the structure fidelity measure [31]. The measure indicates the degree of structure preservation. The higher measure value means that the structure is preserved better. Fig. 3 shows the examples of four detail layers, D_{in}, D_N, D_C^N, and D_{out}, and the structure fidelity maps corresponding to the detail layers. Each structure fidelity map is obtained through the

comparison of the input luminance image and the reconstructed luminance image generated by combining each detail layer and output base layer. In the structure fidelity map, the brighter part indicates the higher structural fidelity. In D_N, the structure is kept well in homogeneous regions. However, the details may be lost in the texture regions due to the limitation of global tone mapping. On the other hand, in case of D_C^N, the structure is preserved well in the texture regions. However, noise signals in the homogeneous regions are too much enhanced due to oversensitivity to noise. Since D_N and D_C^N are complementary, we can obtain an optimized structure-preserved detail layer by using the two detail layers. The constrained optimized problem is defined as follows:

$$\min_{k_{opt}} E(x,y) = \min_{k_{opt}} \begin{pmatrix} w_N(x,y)\cdot\left(k_{opt}(x,y)-k_N(x,y)\right)^2 \\ +w_C^N(x,y)\cdot\left(k_{opt}(x,y)-k_C^N(x,y)\right)^2 \end{pmatrix},$$

subject to :

$$\begin{aligned} &D_{out}(x,y) \geq D_{out}(x',y'), \quad \text{if } D_{in}(x,y) \geq D_{in}(x',y'), \\ &D_{out}(x,y) < D_{out}(x',y'), \quad \text{if } D_{in}(x,y) < D_{in}(x',y'), \end{aligned}$$

where,

$$\begin{aligned} w_N(x,y) &= S_{local}(L_N(x,y), L_{in}(x,y)), \\ w_C^N(x,y) &= S_{local}(L_C^N(x,y), L_{in}(x,y)), \\ S_{local}(X,Y) &= \frac{2\tilde{\sigma}_X(x,y)\tilde{\sigma}_Y(x,y)+C_1}{\tilde{\sigma}_X^2(x,y)+\tilde{\sigma}_Y^2(x,y)+C_1} \\ &\quad \cdot \frac{\sigma_{XY}(x,y)+C_2}{\sigma_X(x,y)\sigma_Y(x,y)+C_2}, \\ &\quad (x',y') \in \Omega(x,y), \end{aligned} \tag{13}$$

where $\Omega(x,y)$ is a 5×5 patch and w_N and w_C^N are weighting factors. The $w_N(x,y)$ and $w_C^N(x,y)$ values are set as the local structure fidelity values [31] of the D_N and D_C^N images at location (x,y), respectively. $\sigma_X(x,y)$, $\sigma_Y(x,y)$, and $\sigma_{XY}(x,y)$ denote the local standard deviations and covariance between the two corresponding patches centered at (x,y), respectively. C_1 and C_2 are two small positive constants to avoid instability. The objective function is expressed as the weighted sum of squared errors, which represents the loss of structure information. Therefore, minimizing an objective function implies minimizing the loss of structure information. Because the objective function is a convex function, it has a minimum value at the pole. Thus, we compute k_{opt}, the optimal coefficient, by taking the partial derivatives of function E with respect to k_{opt} and setting them to zeros:

$$\begin{aligned} \frac{\partial E(x,y)}{\partial k_{opt}(x,y)} &= 2\cdot w_N(x,y)\cdot\left(k_{opt}(x,y)-k_N(x,y)\right) \\ &\quad +2\cdot w_C(x,y)\cdot\left(k_{opt}(x,y)-k_C^N(x,y)\right)=0, \\ k_{opt}(x,y) &= \frac{w_N(x,y)k_N(x,y)+w_C(x,y)k_C^N(x,y)}{w_N(x,y)+w_C(x,y)}. \end{aligned} \tag{14}$$

The k_{opt} value is determined with emphasis on preserving structure; the objective function increases the weight of the natural-looking detail layer (D_N) if it has a high degree of structure preservation. Otherwise, the objective function gives more weight to the normalized contrast-preserved detail layer (D_C^N).

Next, the k_{opt} values are updated to satisfy the constraint, which ensures the basic monotonic principle. The violation of the monotonic principle can cause halo artifacts [11]. To ensure the principle, the output detail layer should preserve the intensity orders of the input detail layer. Therefore, the boundary values are calculated using the values of the neighboring pixels in the patch. The low boundary value (k_L) is set as the k value of the neighboring pixel having the highest value in the output detail layer using the k_{opt} calculated using Eq. (14) among the neighboring pixels having a value smaller than the value of the center pixel in the input detail layer. Inversely, the upper boundary value (k_U) is set as the k value of the neighboring pixel having the lowest value in the output detail layer among the neighboring pixels having a value higher than the value of the center pixel in the input detail layer.

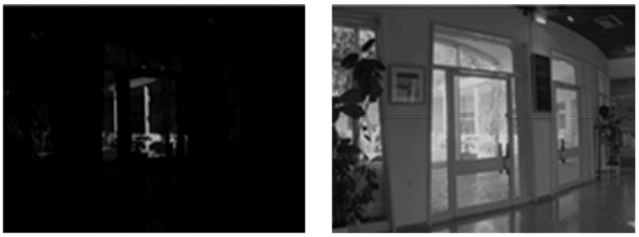
$$k_{opt}(x,y) = \begin{cases} k_{opt}(x,y), & k_L \leq k_{opt}(x,y) \leq k_U \\ k_U(x,y), & k_{opt}(x,y) \geq k_U \\ k_L(x,y), & k_{opt}(x,y) \leq k_L. \end{cases} \tag{15}$$

The output detail layer obtained using the optimal k values is shown in Fig. 3(d).

The output luminance image can be obtained by combining the output base layer and detail layer. Output luminance image L_{out} is defined as follows:

$$L_{out}(x,y) = B_{out}(x,y)\cdot D_{out}(x,y). \tag{16}$$

The output luminance image is shown in Fig. 4.



(a) (b)

FIGURE 4. (a) input luminance image (L_{in}) and (b) output luminance image (L_{out}).

D. COLOR PROCESSING

The color image can be reconstructed by preserving the ratio between the red, green, and blue channels. The equation we use is the same as in [37]:

$$C_{out}(x,y) = \left(\frac{C_{in}(x,y)}{L_{in}(x,y)}\right)^s \cdot L_{out}(x,y), \tag{17}$$

where $C_{in}(x,y)$ and $C_{out}(x,y)$ represent the RGB values of the input and output RGB images at location (x,y), respectively, and $L_{in}(x,y)$ and $L_{out}(x,y)$ denote the luminance values before and after the HDR compression at location

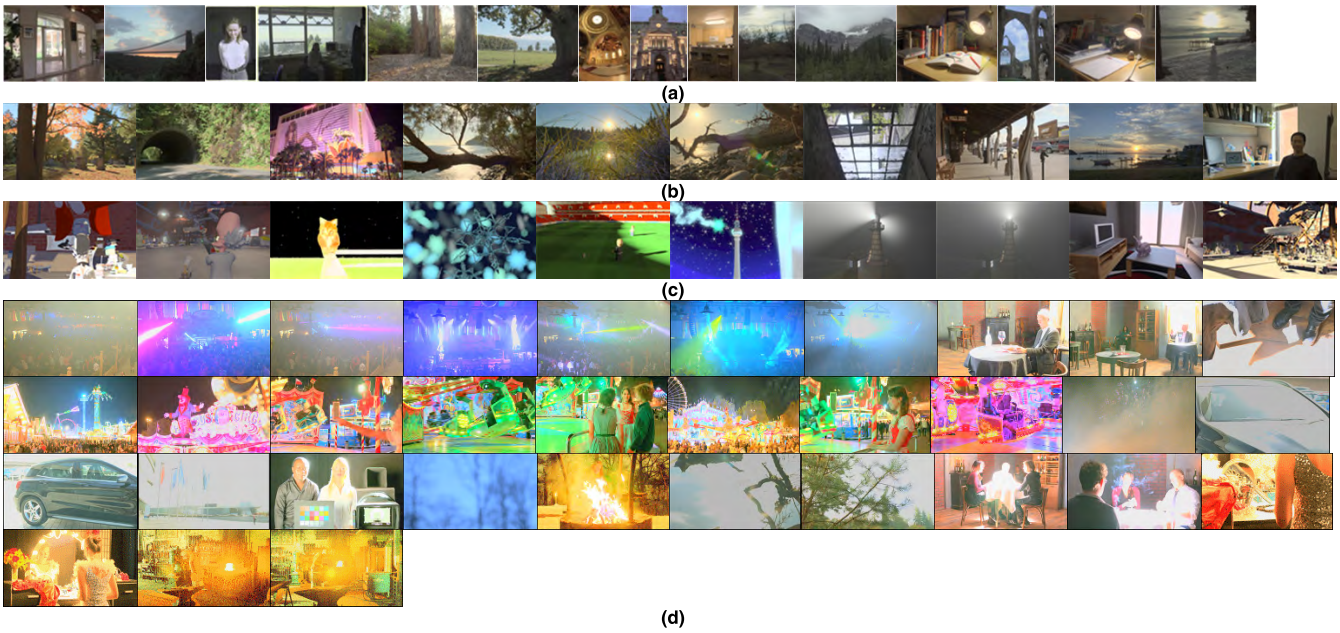


FIGURE 5. HDR test images used in the experiment: (a) basic image set, (b) TMIQD (natural scenes), and (c) TMIQD (synthetic scenes), and (d) Froehlich's database.

(x, y), respectively. This function transforms the relationship of the colors from level L_{in} to level L_{out}. Exponent s is the saturation parameter with the value of 0.4 [37].

III. EXPERIMENTAL RESULTS

A. EXPERIMENTAL SETTING

To evaluate the performance of the proposed method, three image sets were used for the experiments as shown in Fig. 5: basic image set used by Ma *et al.* [31], Tone-mapped image quality database (TMIQD) [38], and Froehlich's database [39]. The basic image set consists of 15 images, and the TMIQD consists of 10 natural scenes and 10 synthetic scenes. Froehlich's database consists of 15 video sequences (Beerfest light show, Bistro, Carousel fireworks, Car close shot, Car full shot, Car long shot, Fireplace, Fishing close shot, Fishing long shot, HDR test image, Poker full shot, Poker travelling slow motion, Showgirl, Smith hammering, and Smith welding). We extracted 33 meaningful frames from the video sequences and performed the experiments for performance evaluation. The image sets contain indoor and outdoor scenes, human and static objects, day and night views, as well as natural and synthetic scenes. Various types of images were used to demonstrate the generality and robustness of the proposed method.

For the benchmarking methods, Meylan's [24], Farbman's [26], Shan's [28], Ma's [31], Laparra's [32], and Shin's [40] methods were used. Among these, Meylan's and Farbman's methods are the most representative retinex-based and layer-based methods, respectively. Shan's [28], Ma's [31], Laparra's [32], and Shin's [40] methods are perceptual model-based methods. In addition, since Laparra's method performs only luminance processing without color processing, we used Mantiuk's method [37] for the post-processing of Laparra's method.

In all the benchmarks, various parameters were optimized to produce the best performance for an unbiased evaluation. The convergence tolerance of Ma's method was set as 0.0001.

B. PERFORMANCE EVALUATION OF PROPOSED METHOD

Various experiments were conducted to evaluate the performances of the proposed and benchmarking tone mapping algorithms.

First, the image quality of the benchmarking and proposed methods was evaluated using the TMQI-II, the feature similarity index for tone-mapped images (FSITM) [41], the high dynamic range-visible difference predictor (HDR-VDP)-2.2 [42], and the dynamic range difference metric (DRIM) [30]. The FSITM is based on the local phase similarity of the original HDR image and converted LDR image. The HDR-VDP-2.2 can predict whether the difference between two images is visible to a human observer or not. This takes into account several aspects of the human visual system such as the diagonal display size, display resolution, and viewing distance. The diagonal display size, display resolution, and viewing distance were set as 24 inch, 1920 × 1200 pixels, and 0.5 meters. The DRIM visualizes three kinds of contrast variations of contrast loss, contrast amplification, and contrast reversal when comparing image pairs. The metric uses color coding to represent the contrast distortion error for each pixel: no error (gray), contrast loss (green), contrast amplification (blue), and contrast reversal (red). Contrast loss indicates that the contrast becomes invisible in the tone-mapped image. Contrast amplification indicates that the contrast becomes

TABLE 1. TMQI-II results of the proposed and benchmarking methods.

Image set		Meylan [24]	Farbman [26]	Shan [28]	Ma [31]	Laparra [32]	Shin [40]	Proposed method
Basic image set	Avg.	0.754	0.502	0.509	0.908	0.457	0.258	**0.949**
	Std.	0.183	0.161	0.180	0.175	0.150	0.108	**0.029**
	Min.	0.431	0.313	0.225	0.474	0.273	0.123	**0.882**
TMIQD (Natural scenes)	Avg.	0.737	0.599	0.454	**0.971**	0.356	0.487	0.936
	Std.	0.203	0.258	0.071	**0.011**	0.110	0.206	0.022
	Min.	0.838	0.560	0.427	**0.953**	0.198	0.268	0.903
TMIQD (Synthetic scenes)	Avg.	0.570	0.333	0.560	0.863	0.251	0.435	**0.904**
	Std.	0.241	0.130	0.214	0.213	0.132	0.215	**0.068**
	Min.	0.531	0.386	0.540	0.437	0.056	0.267	**0.900**
Froehlich's database	Avg.	0.393	0.498	0.567	0.749	0.209	0.531	**0.954**
	Std.	0.088	0.229	0.204	0.208	0.079	0.126	**0.042**
	Min.	0.228	0.244	0.347	0.405	0.043	0.122	**0.789**
Total	Avg.	0.564	0.533	0.584	0.833	0.292	0.448	**0.943**
	Std.	0.229	0.241	0.200	0.202	0.149	0.181	**0.045**
	Min.	0.228	0.101	0.225	0.405	0.043	0.123	**0.757**

TABLE 2. FSITM results of the proposed and benchmarking methods.

Image set		Meylan [24]	Farbman [26]	Shan [28]	Ma [31]	Laparra [32]	Shin [40]	Proposed method
Basic image set	Avg.	0.892	0.846	0.840	0.896	0.776	0.788	**0.898**
	Std.	0.033	0.036	0.041	0.047	0.034	0.050	**0.032**
	Min.	0.807	0.777	0.759	0.808	0.728	0.670	**0.823**
TMIQD (Natural scenes)	Avg.	0.839	0.826	0.780	0.857	0.780	0.864	**0.867**
	Std.	0.029	**0.011**	0.037	0.043	0.037	0.017	0.035
	Min.	0.771	0.810	0.702	0.746	0.713	**0.842**	0.777
TMIQD (Synthetic scenes)	Avg.	0.803	0.801	0.772	0.818	0.753	0.824	**0.827**
	Std.	0.073	**0.034**	0.076	0.071	0.039	0.052	0.066
	Min.	0.703	0.736	0.660	0.716	0.709	0.730	**0.740**
Froehlich's database	Avg.	0.821	0.769	0.839	**0.850**	0.765	0.842	0.844
	Std.	0.041	0.079	0.058	0.050	**0.040**	0.048	0.051
	Min.	0.737	0.582	0.740	0.751	0.667	**0.765**	0.747
Total	Avg.	0.837	0.799	0.821	0.852	0.768	0.831	**0.856**
	Std.	0.053	0.067	0.062	0.054	**0.036**	0.052	0.053
	Min.	0.703	0.582	0.660	0.708	0.667	0.671	**0.740**

TABLE 3. HDR-VDP-2.2 results of the proposed and benchmarking methods.

Image set		Meylan [24]	Farbman [26]	Shan [28]	Ma [31]	Laparra [32]	Shin [40]	Proposed method
Basic image set	Avg.	73.193	71.505	71.583	73.927	68.721	73.591	**73.950**
	Std.	**5.953**	6.376	6.268	6.250	6.672	6.912	6.019
	Min.	**64.754**	62.566	61.741	62.554	59.073	62.362	64.458
TMIQD (Natural scenes)	Avg.	41.219	36.953	40.846	**42.989**	32.725	38.690	42.611
	Std.	2.653	2.170	3.136	3.994	**1.240**	2.122	4.190
	Min.	**38.298**	34.899	37.292	37.826	31.532	34.783	37.273
TMIQD (Synthetic scenes)	Avg.	38.720	34.471	38.480	40.729	31.780	34.841	**41.139**
	Std.	2.002	1.093	1.100	2.950	**0.426**	1.839	2.975
	Min.	**36.176**	32.539	36.082	35.295	31.178	32.428	36.080
Froehlich's database	Avg.	47.190	39.410	46.688	46.990	34.134	44.723	**47.294**
	Std.	4.343	3.839	6.370	5.837	**2.183**	6.877	5.664
	Min.	35.540	33.426	37.836	37.54	31.349	33.441	**36.301**
Total	Avg.	50.803	45.403	50.114	51.423	41.210	48.751	**51.580**
	Std.	13.021	14.556	**12.983**	13.262	15.051	14.857	13.152
	Min.	35.540	32.539	36.062	35.295	31.968	32.428	**36.080**

visible while the contrast is invisible in the input HDR image. Contrast reversal means that the contrast has a different polarity between the tone-mapped image and input HDR image. The less colored regions and lighter color intensity represent the better tone-mapped image quality. Note that DRIM assigns only one type of error to each pixel, which is the type of a predominant error at that location.

Tables I, II, and III present the performance comparison of the proposed tone mapping method and six benchmarking methods. The performance is compared using the

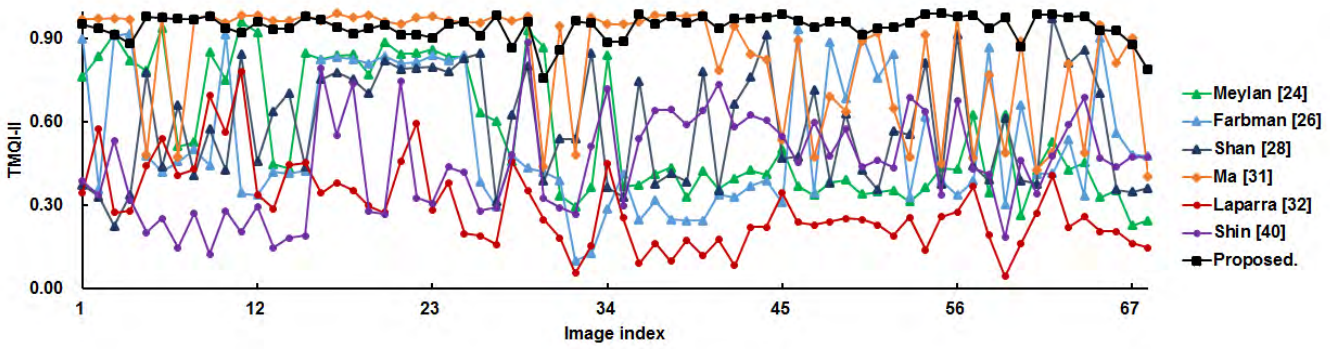


FIGURE 6. Comparison of TMQI-II with the different tone mapping methods.

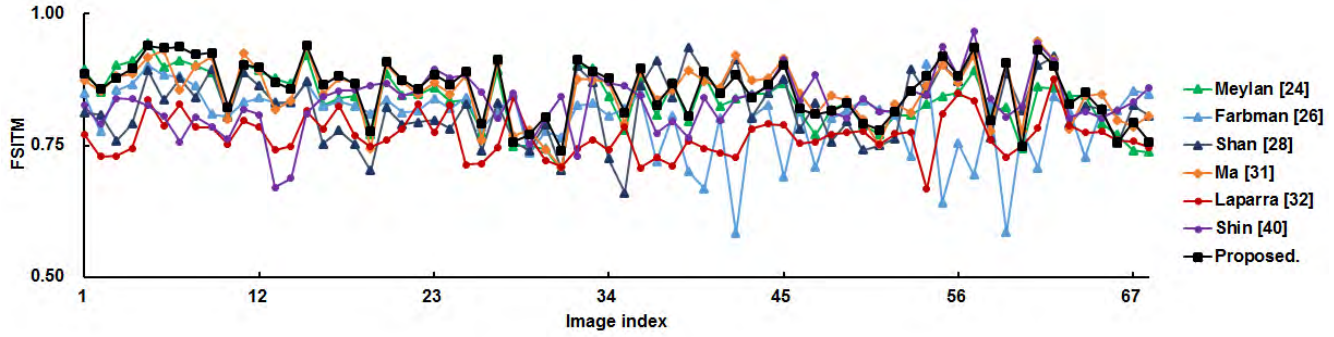


FIGURE 7. Comparison of FSITM with the different tone mapping methods.

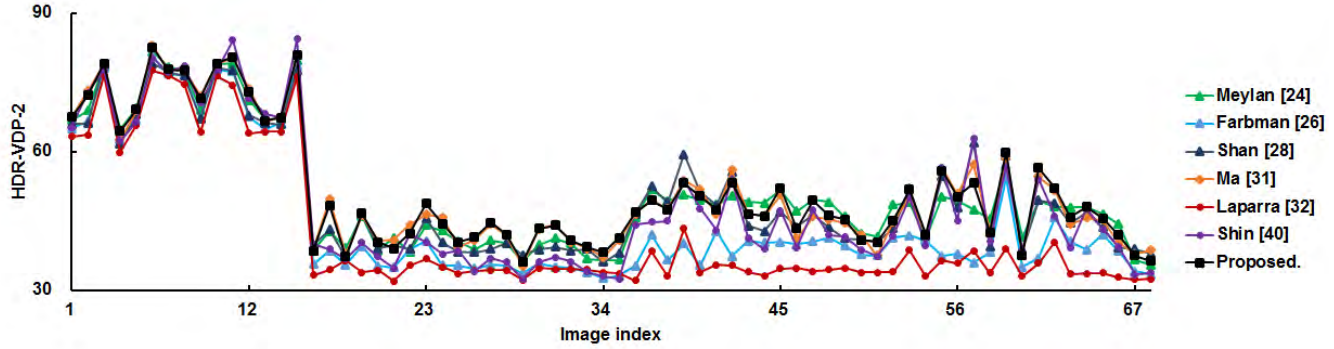


FIGURE 8. Comparison of HDR-VDP-2.2 with the different tone mapping methods.

average (avg.), standard deviation (std.), and the minimum (min.) of TMQI-II, FSITM, and HDR-VDP-2.2 values obtained for the three data sets, Basic image set, TMIQD (Naturalness and Synthetic scenes), and Froehlich's database. Figs. 6, 7, and 8 compare the seven methods in terms of the same evaluation metrics for each of the test images. In Table I, we find that the proposed method yields the higher TMQI-II on the average value, compared to the benchmarking methods. Specifically, the average TMQI-II value of proposed method is as much as 0.378 (67.0%), 0.410 (76.8%), 0.358 (61.3%), 0.109 (13.1%), 0.651 (223.4%), and 0.495 (110.4%) higher than those of Meylan's method, Farbman's method, Shan's method, Ma's method, Laparra's method, and Shin's method. Furthermore, the proposed method reduces the standard deviation of TMQI-II by 0.155 (77.6%) $\sim$ 0.195 (81.4%) compared with the benchmarking methods. With respect to the minimum TMQI-II value, the proposed method yields the best performance among the seven methods. In Fig. 6, the proposed method yields the highest TMQI-II values for more images than the benchmarking methods. From the viewpoint of TMQI-II, the proposed method is regarded to preserve naturalness and structure information better than the benchmarking methods generally. Tables II and III illustrate that the proposed method has the higher average and higher minimum values of FSITM and HDR-VDP-2.2, when compared with the benchmarking

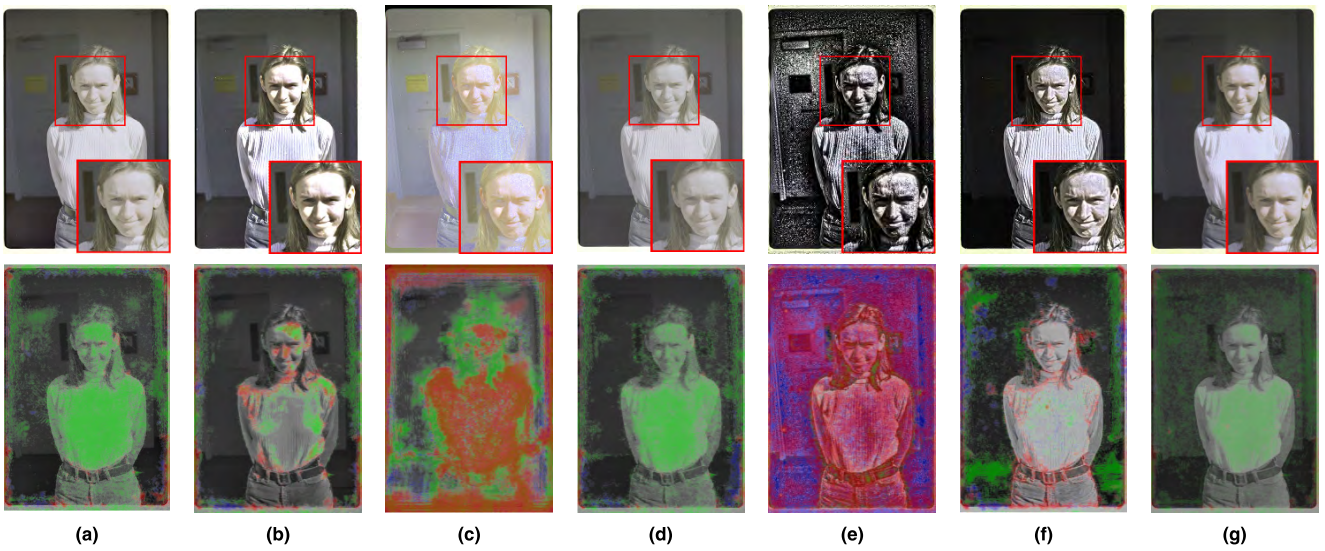


FIGURE 9. Human image example. Tone mapping results (top row) and distortion maps (bottom row). Distortion maps are generated using DRIM [30]. (a) Meylan's method, (b) Farbman's method, (c) Shan's method, (d) Ma' method, (e) Laparra's method, (f) Shin's method, and (g) the proposed method.

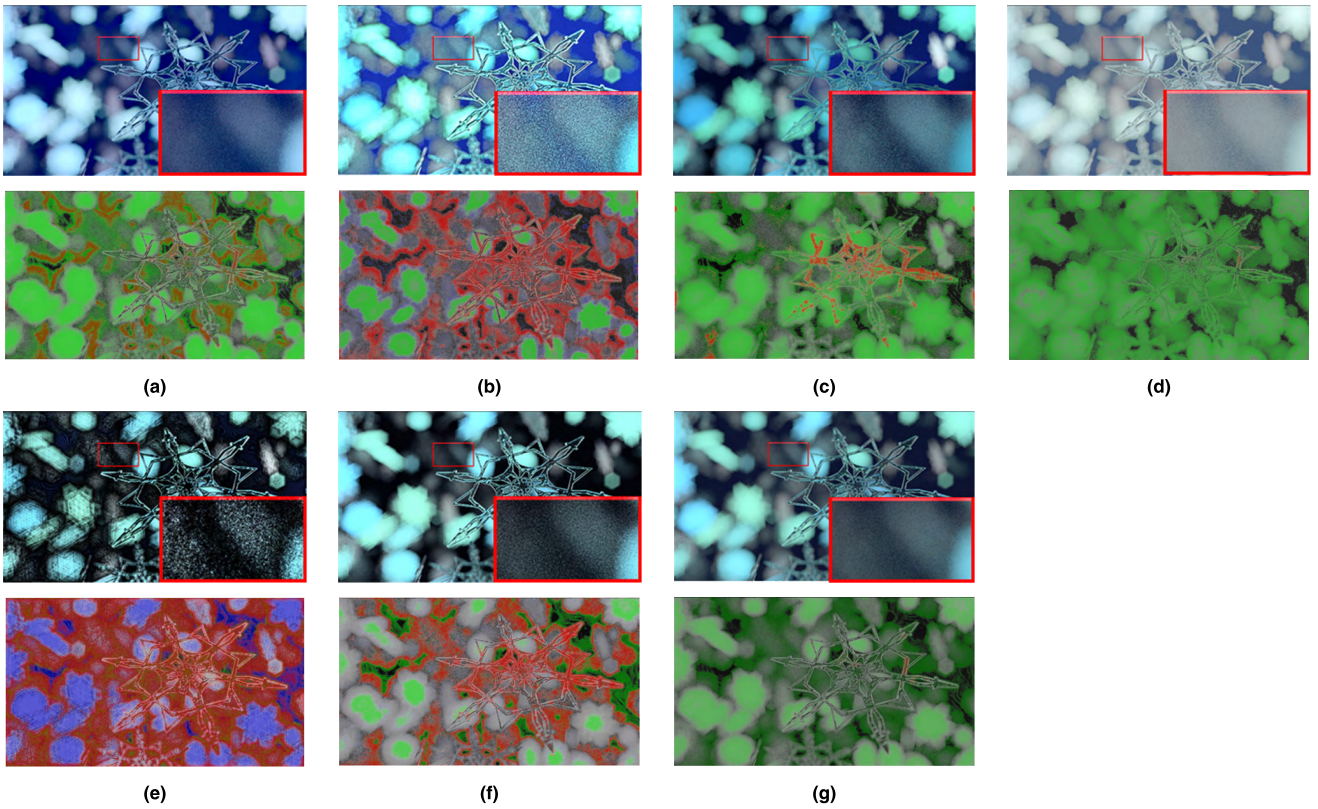


FIGURE 10. Example of a synthetic scene. Tone mapping results (top row) and distortion maps (bottom row). Distortion maps are generated using DRIM [30]. (a) Meylan's method, (b) Farbman's method, (c) Shan's method, (d) Ma' method, (e) Laparra's method, (f) Shin's method, and (g) the proposed method.

methods. In addition, the proposed method provides the lower standard deviation values of FSITM and HDR-VDP-2.2 than Ma's method. Ma's method has the highest average values of FSITM and HDR-VDP-2.2 among the benchmarking methods. In Figs. 7 and 8, the proposed method yields slightly higher FSITM and HDR-VDP-2.2 values compared to Ma's method. This means that the proposed method produces the tone-mapped image which is more similar to the input image than the benchmarking methods.

Figs. 9, 10, and 11 present the resulting images of the proposed and benchmarking methods and the corresponding distortion maps. Based on Figs. 9 and 10, we can say that

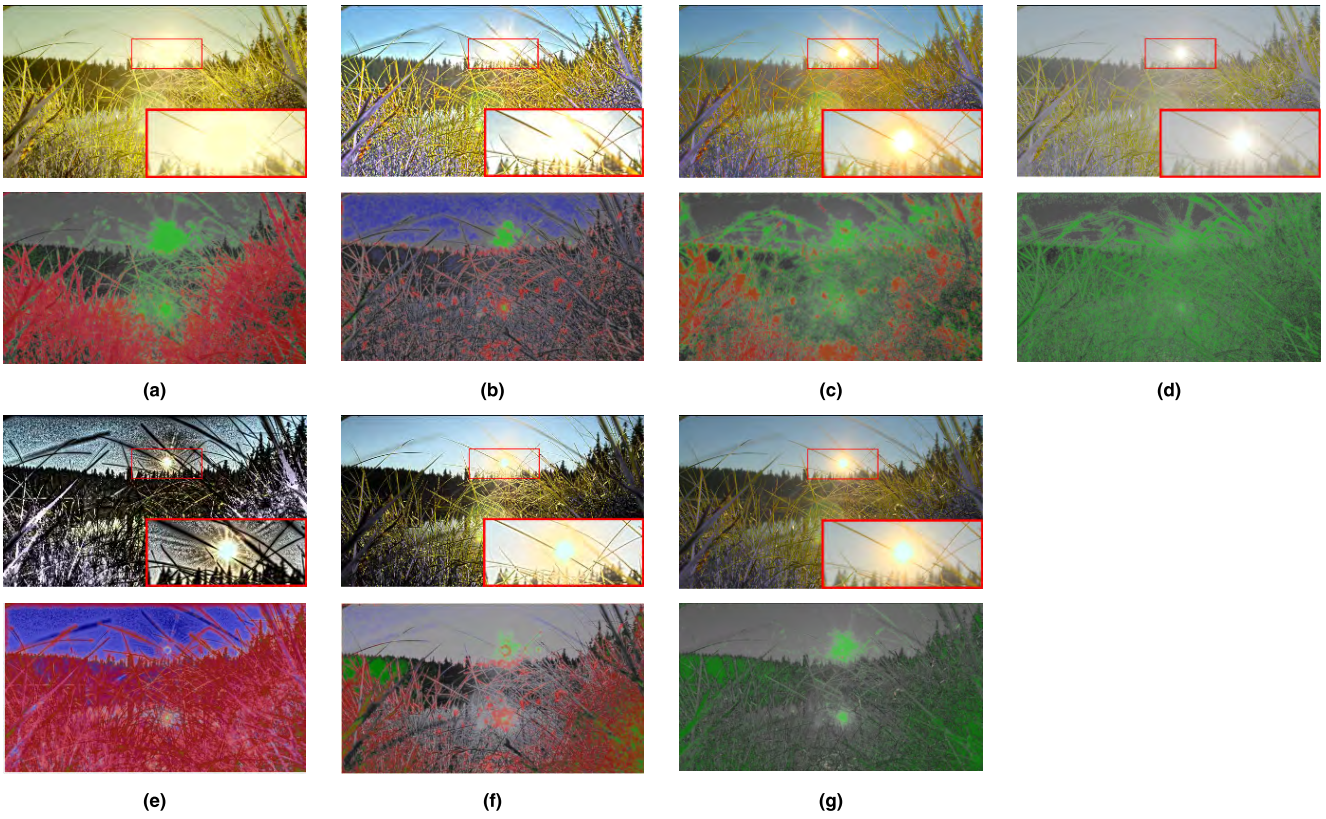


FIGURE 11. Example of a natural scene. Tone mapping results (top row) and distortion maps (bottom row). Distortion maps are generated using DRIM [30]. (a) Meylan's method, (b) Farbman's method, (c) Shan's method, (d) Ma' method, (e) Laparra's method, (f) Shin's method, and (g) the proposed method.

both Farbman's and Shan's methods cause contrast reversal in detail regions, thereby leading to an unnatural appearance. Farbman's method is good at preserving the details of the different scales, but it produces over-exaggerated details. Furthermore, it strongly boosts the noise in the smooth regions as shown in the red box in Figs. 9(b) and 10(b). According to Figs. 9(c) and 10(c), Shan's method commonly leads to a high global contrast. However, the method often loses the local details of the objects. Therefore, it degrades the visibility of the output image. In comparison, the proposed method does not have these problems, as is evident from the red boxes in Figs. 9(g) and 10(g). Fig. 11 presents images with bright regions. Meylan's method does not cause halo artifacts. However, it results in a contrast loss in the bright areas of the scene, as shown in Fig. 11(a). The images created by Ma's method look natural. However, it produces a low contrast compared with the other methods, as depicted in Fig. 11(d). Therefore, the perceived image quality decreases owing to the low contrast. This is because the required convergence tolerance varies depending on the type of image. Laparra's method induces contrast amplification or contrast reversal in most regions (Figs. 9(e), 10(e), and 11(e)). The images produced by Laparra's method appear unnatural. Shin's method does not provide a consistent quality for various images. In Fig. 9(f), it causes contrast loss in the dark areas of the scene due to strong contrast compression. Conversely, in Fig. 11(f), it causes contrast reversal in detail regions. Figs. 9(g), 10(g), and 11(g) are the resulting images of the proposed method. The proposed method causes almost no contrast amplification and reversal; the overall appearance of the images looks natural without artifacts. Furthermore, the structure information is relatively well preserved in the images.

TABLE 4. Average TMQI-II, FSTIM, and HDR-VDP-2.2 for case 1, case 2, and the proposed method.

	Case 1*	Case 2**	Proposed method
Avg. TMQI-II	0.932	0.719	**0.943**
Avg. FSTIM	0.823	0.738	**0.856**
Avg. HDR-VDP-2.2	50.49	47.66	**51.58**

* Case 1 denotes the case that the proposed method considers the naturalness only in the detail layer.
** Case 2 denotes the case that the proposed method considers the contrast only in the detail layer.

Second, we evaluated the effect of each of two steps, related to naturalness and contrast preservation of the detail layer, respectively. The experiment was conducted using all datasets. Table IV shows the quality of the tone mapping process for each step. Case 1 denotes the case that the

TABLE 5. Average ranking scores made by 10 subjects for each test image set.

Image set	Meylan [24]	Farbman [26]	Shan [28]	Ma [31]	Laparra [32]	Shin [40]	Proposed method
Basic image set	2.63	3.47	4.79	3.30	6.69	6.58	**2.31**
TMIQD (Natural scenes)	3.37	3.70	5.14	2.94	6.74	4.28	**1.81**
TMIQD (Synthetic scenes)	3.09	3.66	4.00	3.64	5.89	4.93	**2.74**
Froehlich's database	4.01	4.65	3.78	3.91	6.76	2.52	**2.37**

TABLE 6. Average, standard deviation, and maximum C_T for the proposed and benchmark methods and average and standard deviation C_T rate of the proposed method.

	Meylan [24]	Farbman [26]	Shan [28]	Ma [31]	Laparra [32]	Shin [40]	Proposed method
Avg. C_T (μs)	16.95	0.95	3.92	1592.77	131.62	164.62	**0.61**
Std. C_T (μs)	8.98	0.20	0.21	66.84	26.07	29.52	**0.06**
Max. C_T (μs)	24.17	1.01	4.41	3129.72	178.20	256.38	**0.69**
Avg. C_T rate	27.79	1.56	6.43	2611.10	215.78	269.87	**1**
Std. C_T rate	149.67	3.34	3.47	1113.98	434.61	492.05	**1**

TABLE 7. Average TMQI-II, FSTIM, HDR-VDP-2.2, and C_T at different patch sizes.

	Patch size						
	3×3	5×5	7×7	9×9	11×11	13×13	15×15
Avg. TMQI-II	0.941	**0.943**	**0.943**	**0.943**	**0.943**	**0.943**	**0.943**
Avg. FSTIM	0.855	0.856	0.856	0.856	**0.858**	**0.858**	**0.858**
Avg. HDR-VDP-2.2	51.551	51.580	51.582	51.585	51.587	**51.588**	**51.588**
Avg. C_T (μs)	**0.608**	0.612	0.684	0.764	0.889	1.032	1.190

proposed method considers the naturalness only and does not consider the contrast in the detail layer. On the contrary, Case 2 is the case that the proposed method considers the contrast only and does not consider the naturalness in the detail layer. In another words, Case 1 and Case 2 use D_N and D_C^N instead of D_{out}, respectively. Since the input detail layer contains both low-frequency components and high-frequency components, the proposed method considering both naturalness and contrast shows the best image quality.

Third, we performed a subjective test after inviting 10 subjects to rank tone-mapped images, generated by the proposed and six benchmarking methods. As seven methods were used for the test, each subject ranked the seven methods from one to seven after appreciating the tone-mapped images. All of the 68 images presented in Fig. 5 were used as test images. Table V compares the average subjective rankings of the seven methods for each image set. In the table, the lower ranking score corresponds to the better quality. The table indicates that the proposed method obtained the best average rank for every data set among the tone-mapping methods used for the test.

Fourth, we evaluated the computation time per pixel (C_T). For the C_T comparison, the proposed and all the methods were executed using MATLAB on a PC with an Intel I5-4970 3.40 GHz processor. The comparison results for the computation time are listed in Table VI. The proposed method considerably reduces the average, maximum, and standard deviation of C_T by up to 1592.16 μs (99.9%), 3129.03 μs (99.9%), and 66.78 μs (99.9 %), respectively, compared with the benchmarking methods. This is because the proposed method is performed without any iteration process.

Fifth, we evaluated the performance of the proposed method with various patch sizes. The experiment was conducted using all datasets. For the experiment of the performance sensitivity to patch sizes, all parameter values except the patch size were fixed. Table VII shows the average TMQI-II, FSTIM, HDR-VDP-2.2 values, and computation time per pixel (C_T) for different patch sizes. As the patch size increases, the computation time also increases, but the accuracy remains almost the same. In the proposed method, we set the patch size to 5×5.

Finally, the overall performance of the proposed and benchmarking methods was evaluated using the image quality and computation time. Fig. 12 shows the average TMQI-II and computation time. The proposed method not only provides the most outstanding image quality but also has the fastest operation speed compared with those of the benchmarking methods. Fig. 13 shows another superiority of the proposed method. The proposed method provides the least standard deviation of image quality and computation time compared to those of the benchmark methods. The experimental results verify that compared with the benchmarking methods, the proposed method consistently yields the most outstanding image quality with a low computational cost.

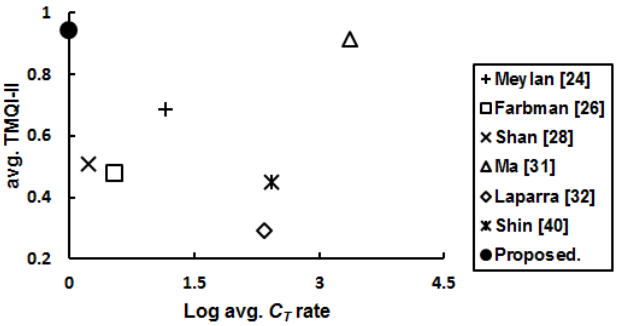


FIGURE 12. Comparison of the avg. TMQI-II and log avg. C_T rate obtained with the proposed and benchmarking methods.

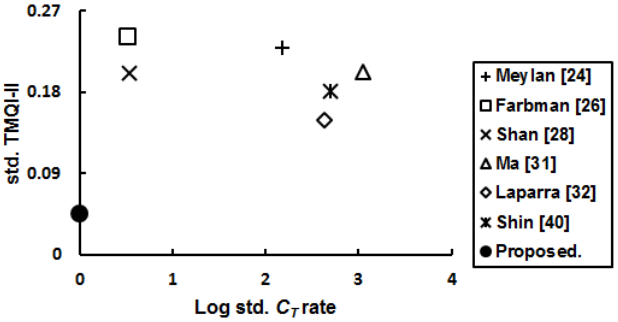


FIGURE 13. Comparison of the std. TMQI-II and log std. C_T rate obtained with the proposed and benchmarking methods.

IV. CONCLUSION

This paper proposed an efficient tone mapping method that provides a high perceptual image quality with a low computation complexity. The feature of the proposed method is that it finds an optimal detail layer and generates a tone-mapped image without using the iterative process. The proposed method first decomposed an image into illumination components and detail components. Next, the proposed method independently performed the compression of the illumination components and the enhancement of the detail components using perceptual models. Finally, the output image was generated by combining the components. The experiments revealed that the proposed method improved the average TMQI, FSITM, and HDR-VDP-2.2 by up to 0.651 (223.4%), 0.088 (11.5%), and 10.371 (25.2%), respectively, and decreased the execution time by up to 99.9% compared with the benchmarking methods. Furthermore, the proposed method improved the standard deviations of TMQI-II, FSITM, HDR-VDP-2.2, and processing time by up to 81.4%, 18.9%, 12.6%, and 99.9%, respectively, compared with the benchmarking methods.

REFERENCES

[1] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, 2009.
[2] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, and L. Zhang, "Robust multi-exposure image fusion: A structural patch decomposition approach," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, May 2017.
[3] Z. G. Li, J. H. Zheng, and S. Rahardja, "Detail-enhanced exposure fusion," *IEEE Trans. Image Process.*, vol. 21, no. 11, pp. 4672–4676, Nov. 2012.
[4] K. Ma, Z. Duanmu, H. Yeganeh, and Z. Wang, "Multi-exposure image fusion by optimizing a structural similarity index," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 60–72, Mar. 2018.
[5] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. SIGGRAPH*, Los Angeles, CA, USA, Aug. 1997, pp. 369–378.
[6] E. Reinhard, G. Ward, S. Pattanaik, and P. E. Debevec, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. San Mateo, CA, USA: Morgan Kaufmann, 2005.
[7] J. Tumblin and H. Rushmeier, "Tone reproduction for realistic images," *IEEE Comput. Graph. Appl.*, vol. 13, no. 6, pp. 42–48, Nov. 1993.
[8] G. Ward, "A contrast-based scalefactor for luminance display," in *Graphics Gems IV*. New York, NY, USA: Academic, 1994, pp. 415–421.
[9] J. A. Ferwerda, S. N. Pattanaik, P. Shirley, and D. P. Greenberg, "A model of visual adaptation for realistic image synthesis," in *Proc. SIGGRAPH*, 1996, pp. 249–258.
[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, 2003.
[11] G. Qiu, J. Guan, J. Duan, and M. Chen, "Tone mapping for HDR image using optimization a new closed form solution," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2006, pp. 996–999.
[12] J. Duan, G. Qiu, and G. M. D. Finlayson, "Learning to display high dynamic range images," *Pattern Recognit.*, vol. 40, no. 10, pp. 2641–2655, 2007.
[13] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 249–256, 2002.
[14] E. H. Land, "The retinex," *Amer. Sci.*, vol. 52, no. 2, pp. 247–264, 1964.
[15] E. H. Land and J. J. McCann, "Lightness and retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.
[16] B. K. P. Horn, "Determining lightness from an image," *Comput. Graph. Image Process.*, vol. 3, no. 4, pp. 277–299, 1974.
[17] A. Hurlbert, "Formal connections between lightness algorithms," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 3, no. 10, pp. 1684–1693, 1986.
[18] D. H. Brainard and B. A. Wandell, "Analysis of the retinex theory of color vision," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 3, no. 10, pp. 1651–1661, 1986.
[19] A. Rizzi, C. Gatta, and D. Marini, "From retinex to automatic color equalization: Issues in developing a new algorithm for unsupervised color equalization," *J. Electron. Imag.*, vol. 13, no. 1, pp. 15–28, 2004.
[20] R. Kimmel, M. Elad, D. Shaked, R. Keshet, and I. Sobel, "A variational framework for retinex," *Int. J. Comput. Vis.*, vol. 52, no. 1, pp. 7–23, 2003.
[21] M. Elad, R. Kimmel, D. Shaked, and R. Keshet, "Reduced complexity retinex algorithm via the variational approach," *J. Vis. Commun. Image Represent.*, vol. 14, no. 1, pp. 369–388, 2003.
[22] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.
[23] M. Elad, "Retinex by two bilateral filters," in *Proc. 5th Int. Conf. Scale Space PDE Methods Comput. Vis.*, 2005, pp. 217–229.
[24] 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.
[25] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 257–266, 2002.
[26] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249–256, 2008.
[27] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
[28] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone mapping," *IEEE Trans. Vis. Comput. Graphics*, vol. 16, no. 4, pp. 663–675, Jul./Aug. 2010.
[29] 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.
[30] T. O. Aydin, R. Mantiuk, K. Myszkowski, and H.-P. Seidel, "Dynamic range independent image quality assessment," *ACM Trans. Graph.*, vol. 27, no. 3, 2008, Art. no. 69.

[31] 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.
[32] V. Laparra, A. Berardino, J. Ballé, and E. P. Simoncelli, "Perceptually optimized image rendering," *J. Opt. Soc. Amer. A, Opt. Image Sci.*, vol. 34, no. 9, pp. 1511–1525, 2017.
[33] V. Laparra, J. Ballé, A. Berardino, and E. P. Simoncelli, "Perceptual image quality assessment using a normalized Laplacian pyramid," *Electron. Imag.*, vol. 16, pp. 1–6, 2016.
[34] P.-S. Tsai, C.-K. Liang, T.-H. Huang, and H. H. Chen, "Image enhancement for backlight-scaled TFT-LCD displays," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 4, pp. 574–583, Apr. 2009.
[35] W. J. Crozier, "On the variability of critical illumination for flicker fusion and intensity discrimination," *J. Gen. Physiol.*, vol. 19, no. 3, pp. 503–522, 1936.
[36] P. G. J. Barten, "Contrast sensitivity of the human eye and its effects on image quality," *J. Electron. Imag.*, vol. 15, no. 3, 2006.
[37] R. Mantiuk, A. Tomaszewska, and W. Heidrich, "Color correction for tone mapping," *Comput. Graph. Forum*, vol. 28, no. 2, pp. 193–202, 2009.
[38] [Online]. Available: http://mmtg.fel.cvut.cz/tmiqd-database/
[39] J. Froehlich, S. Grandinetti, B. Eberhardt, S. Walter, A. Schilling, and H. Brendel, "Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays," *Proc. SPIE*, vol. 9023, p. 90230X, Mar. 2014.
[40] S. Shin, K. Kong, J. Lee, W.-J. Song, K. J. Kwon, and S. G. Kim, "Tone mapping to minimize visual sensation distortion considering brightness and local band-limited contrast," *J. Soc. Inf. Display*, vol. 25, no. 10, pp. 621–635, 2017.
[41] 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.
[42] M. Narwaria, R. Mantiuk, M. P. Da Silva, and P. Le 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, p. 010501, 2015.

GYUJIN BAE (S'13) received the B.S. degree in electronics engineering from Kyungpook National University, South Korea, in 2013, and the M.S. degree in electrical and electronic engineering from the Pohang University of Science and Technology, South Korea, in 2015, where he is currently pursuing the Ph.D. degree. His current research interests include image processing and computer vision.

CHAN YOUNG JANG (S'11) received the B.S. degree in electronic engineering from Pusan National University, Busan, South Korea, in 2011, and the Ph.D. degree in electrical and computer engineering from the Pohang University of Science and Technology, South Korea, in 2017. He is currently a Senior Engineer with Samsung Electronics Co. Ltd., Hwaseong, South Korea. His current research interests include image processing algorithms, image analysis, and computer vision.

SUNG IN CHO (S'10–M'15) received the B.S. degree in electronic engineering from Sogang University, South Korea, in 2010, and the Ph.D. degree in electrical and computer engineering from the Pohang University of Science and Technology, South Korea, in 2015. From 2015 to 2017, he was a Senior Researcher with LG Display, South Korea. He is currently an Assistant Professor in electronic engineering with Daegu University, South Korea. His current research interests include image analysis and enhancement, video processing, multimedia signal processing, and circuit design for LCD and OLED systems.

YOUNG HWAN KIM (S'86–M'89–SM'14) received the B.E. degree in electronics from Kyungpook National University, South Korea, in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of California at Berkeley, Berkeley, CA, USA, in 1985 and 1988, respectively. From 1977 to 1982, he was with the Agency for Defense Development, South Korea, where he was involved in various military research projects, including the development of autopilot guidance and control systems. From 1983 to 1988, he was a Post-Graduate Researcher with the Electronic Research Laboratory, University of California at Berkeley, where he was involved in developing VLSI CAD programs. He is currently a Professor with the Division of Electronic and Computer Engineering, Pohang University of Science and Technology, South Korea. His research interests include plasma and liquid crystal display systems, multimedia circuit design, MPSoC and GPGPU system design for display and computer vision applications, statistical analysis and design technology for deep-submicrometer semiconductor devices, and power noise analysis. He has served as a Committee Member and the General Chair of various Korean domestic and international technical conferences, including the International SoC Design Conference, the IEEE ISCAS 2012, and the IEEE APCCAS 2016. He has served as an Editor for the *Journal of the Institute of Electronics Engineers of Korea*.