IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 21, 2019, accepted August 12, 2019, date of publication August 23, 2019, date of current version September 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2937211

Novel Tone Mapping Method via Macro-Micro Modeling of Human Visual System

DISHENG MIAO[1,2], ZHONGJIE ZHU[1,2], YONGQIANG BAI[1], GANGYI JIANG[3], AND ZHIYONG DUAN[2]

[1]Ningbo Key Laboratory of DSP, Zhejiang Wanli University, Ningbo 315000, China
[2]Physical Engineering College, Zhengzhou University, Zhengzhou 450000, China
[3]Institute of Technology, Ningbo University, Ningbo 315211, China

Corresponding author: Zhongjie Zhu (zhongjiezhu@yeah.net)

This work was supported in part by the National Natural Science Foundations of China under Grant 61671412, in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LY19F010002, in part by the Natural Science Foundation of Ningbo, China, under Grant 2018A610053, in part by the Ningbo Municipal Project of Science and Technology beneficial to People under Grant 2017C50011, and in part by the Innovation and Consulting Project from Ninghai Power Supply Company, State Grid Corporation of Zhejiang, China.

ABSTRACT Tone mapping(TM) aims to adapt high dynamic range (HDR) images to conventional displays with visual information preserved. In this paper, a novel TM method based on macro-micro modeling is proposed, which can address the common problems in existing TM methods, such as exposure imbalance and halo artifact. From a microscopic perspective, multi-layer decomposition and reconstruction are applied to model the properties of brightness, structure, and detail for HDR images, and then different strategies are adopted for each layer by the human visual system (HVS) to reduce the overall brightness contrast and retain as much scene information. From a macroscopic perspective, scene content-based global operator is designed to adaptively adjust the scene brightness so that it is consistent with the subjective perception of human eyes. Both the micro and macro models are processed in parallel, which can ensure the integrity and subjective consistency of scene information. Experiments with numerous HDR images and TM methods are conducted and the results show that the proposed method achieves visually compelling results with little exposure imbalance and halo artifact, and is superior to the current state-of-the-art TM methods in both subjective and objective evaluations.

INDEX TERMS Tone mapping, macro-micro modeling, high dynamic range, human visual system.

I. INTRODUCTION

Dynamic range is the logarithm of the ratio of the maximum to the minimum luminance for a digital image [1]. In reality, the real scene with wide dynamic range can be effectively perceived by human eyes [2]–[4]. However, existing low dynamic range (LDR) images adopt 8 bits/color/pixel to represent a limited range of luminance, and often include overexposed or underexposed phenomena, resulting in some loss of scene information. To address this deficiency, multi-exposure image fusion methods, which fusion of LDR images that are taken from the same scene at different exposure levels, are often applied to display more information in a limited dynamic range [5]–[7]. However, multi-exposure image fusion methods are still unable to retain the complete information. Therefore, high dynamic range (HDR) imaging technology employs floating-point numbers to represent a wide range of luminance, which can retain the fidelity of real scene accurately. But unfortunately, these HDR display devices are difficult to popularize because of their cost and technical problems [8], [9]. Hence, how to visualize HDR images effectively on existing display devices has become a problem that must be solved in practical applications. To tackle the problem, tone mapping(TM), which essentially functions as intensities mapping of an HDR image to the target display range, has been developed [10], [11].

Up to now, numerous tone mapping operators(TMOs) have emerged and can be roughly classified into two categories: Global TMOs and Local TMOs. For the former, all pixel values in the HDR image are mapped with the same mapping function, so these are relatively simple, fast, efficient and can maintain good overall masking effect [12].

The associate editor coordinating the review of this article and approving it for publication was Tianhua Xu.

FIGURE 1. Examples of global TMO and local TMO, (a) (b) global TMO, (c) local TMO.

Larson *et al.* employed a histogram adjustment technique by incorporating human contrast sensitivity, spatial acuity, and color sensitivity [13]. Drago *et al.* utilized a logarithmic bias power function to adaptively compress luminance values and preserve details and contrast [14]. Mantiuk *et al.* minimized visible contrast distortions according to their visibility predicted by the model of the HVS [15]. Jung *et al.* proposed a limit curve based on perceptual quantization transfer function to adjust the degree of contrast enhancement [16]. Khan *et al.* employed histogram of luminance to construct a lookup table for TM [17]. Lee *et al.* introduced a new asymmetric sigmoid curve based on the model of the retinal response curve [12]. These global TMOs effectively reduce the overall luminance contrast, but also are easy to result in the loss of local information by a single function, especially in the bright or dark regions as shown in Fig. 1(a) and (b). In contrast, local TMOs also take local neighborhood features into account, which are flexibly maintain more details and contrast for each region. Reinhard *et al.* utilized the technique of photographic practice and central surround function to deal with the local visual information [18]. Li *et al.* used a symmetrical analysis-synthesis filter bank for local gain control and luminance compression [19]. Shan *et al.* developed overlapping window-based linear functions and constructed the guidance map via local statistical information for TM [20]. Furthermore, Durand *et al.* presented a two-scale decomposition method to model the Retinex theory for better TM performance [21]. Farbam *et al.* proposed an edge-preserving smoothing operator based on the weighted least squares optimization framework for multi-scale decompositions [22]. Barai *et al.* proposed a salience guide edge-preserving operator which uses saliency region information of an HDR image as input to decompose the image [23]. For these local TMOs, the introduction of local information enhances the scene details, but to a certain extent, it also causes the imbalance of global scene brightness and halo phenomenon as shown in Fig. 1(c).

Towards filling these gaps of details lost, unbalance lighting contrast and halo phenomenon, we propose a TM method based on macro-micro model of human visual system(HVS). Firstly, by expanding the human eyes perception, the proposed method models the visual cognitive mechanism from micro and macro perspectives respectively. And then, two models are processed in parallel to obtain the desired LDR image. In brief, this method has three major contributions as follows.

(1) Both microscopic model and macroscopic model for TM are designed. This is inspired by the evidence in visual physiology, i.e., the HVS first rapidly and unconsciously produces a global perception, and then gradually focuses on specific local areas for the perception of image quality [24], [25]. Meanwhile, both the micro and macro models are processed in parallel to ensure the integrity and subjective consistency of scene information.

(2) From a microscopic perspective, multi-layer decomposition and reconstruction are applied to model visual information of brightness, structure, and detail for HDR images. This is inspired by the discoveries in intrinsic decomposition, that is hierarchical representation using a series of visual content descriptions is coincident with the perceptual-cognitive process of visual signals [26], [27]. And then different strategies are adopted for each layer by the property of HVS, to reduce the overall brightness contrast and retain as much scene information.

(3) From a macroscopic perspective, scene content-based global operator is designed to adaptively adjust the scene brightness so that it is consistent with the subjective perception of human eyes. This is motivated by the fact that an image with high quality must consider the interaction and conversion between human subjective perception and device objective display [28].

To validate the performance of the proposed method, we compare results with both the classical and state of the art TM methods. The results show that the proposed method achieves visually compelling results with little exposure imbalance and halo artifact, and is superior to the current state-of-the-art TM methods in both subjective and objective evaluations.

The rest of this paper is organized as follows. Section II describes the motivation and methodology of the proposed TM method in detail. And then, the relevant experimental results and comparative analysis are presented in Section III. Finally, Section IV concludes this paper.

II. MOTIVATION AND METHODOLOGY

As shown in Fig. 2, we design both micro model and macro model and combine them in parallel to ensure the integrity and subjective consistency of scene information for TM process. On the one hand, the microscopic model is described by multi-layer decomposition. Three layers are extracted to portray visual information of HDR images, and then different strategies are adopted for each layer by the property of HVS. On the other hand, the macroscopic model is described by a global operator, which is utilized to adaptively adjust the brightness based on the scene content of HDR image. Finally, two intermediate results are fused to obtain the desired LDR image.

A. PREPROCESSING

Compared with LDR images, the biggest feature of HDR images is the expansion of luminance dynamic range for real scene, and the color information remains basically

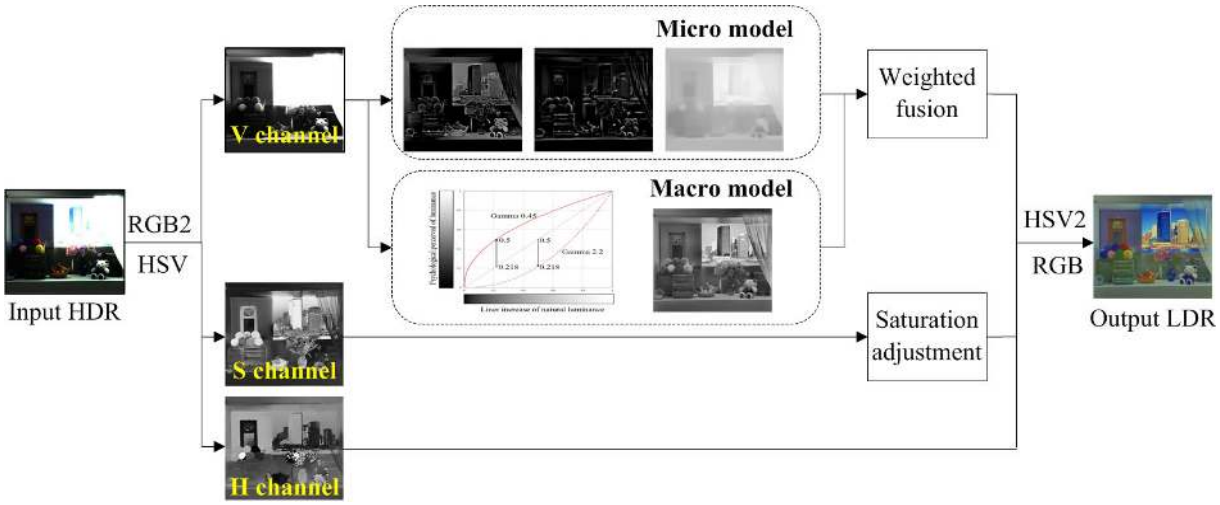


FIGURE 2. Flowchart of the proposed macro-micro tone mapping method.

unchanged. Meanwhile, relevant psychophysical studies have shown that the HSV color space is more consistent with the perception of human eyes than the RGB space [29]. Hence, the HDR image is firstly transformed into HSV color space with three channel, and the Value channel is selected for subsequent operations. The channels of RGB space are normalized first with Eq. (1), and then each channel of HSV space can be obtained with Eq. (2)-(4) as follows [30].

$$\begin{aligned} R' &= R/sum(R,G,B) \\ G' &= G/sum(R,G,B) \\ B' &= B/sum(R,G,B) \end{aligned} \tag{1}$$

$$H = \begin{cases} \dfrac{G'-B'}{\max(R',G',B')-\min(R',G',B')} \cdot 60°, \\ \qquad R' = \max(R',G',B') \\ \left(2+\dfrac{B'-R'}{\max(R',G',B')-\min(R',G',B')}\right) \cdot 60°, \\ \qquad G' = \max(R',G',B') \\ \left(4+\dfrac{B'-R'}{\max(R',G',B')-\min(R',G',B')}\right) \cdot 60° \\ \qquad B' = \max(R',G',B') \end{cases} \tag{2}$$

$$S = \frac{\max(R',G',B')-\min(R',G',B')}{\max(R',G',B')} \tag{3}$$

$$V = \max(R',G',B') \tag{4}$$

where R, G, B are the values of each channel in RGB space, respectively. Similarly, H, S, V are the values of each channel in HSV space, respectively.

B. MICRO MODEL

Due to the limitation of dynamic range, the LDR images after TM cannot preserve all the visual information and local structure of the original HDR images. Hence, the micro model is performed to extract useful features to describe visual information and local structure. Here, we resort to multi-layer decomposition and reconstruction as the basis

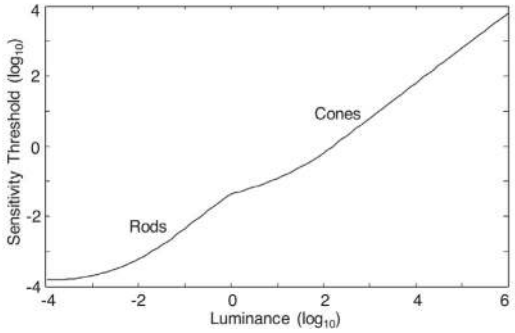


FIGURE 3. Threshold versus intensity curve giving sensitivity of the HVS for each dynamic range of luminance.

for local feature extraction, which is known as an effective technique to describe signals by decomposing into some layers with different visual perception characteristics. Previous studies have shown strong evidence that the receptive fields of photoreceptor cells can be characterized as being spatially, brightly and structurally sensitive, which closely correspond to the characteristic of each layer used in multi-layer decomposition [31].

Here, the original HDR image is decomposed into three layers, that is brightness layer, structure layer, and detail layer. And then different strategies are adopted for each layer by the property of HVS. For the brightness layer, we reduce the overall brightness contrast to match the display luminance range by the recent study in [32] and [33]. Kim *et al.* modeled the impact of ambient conditions on the perceived luminance and chrominance, and used it to improve visibility of display of mobile devices in outdoor [32]. And Kundu *et al.* developed similar psychophysical studies to determine just noticeable difference(JND) perceived by human eyes at a given adaptation level as shown in Fig. 3 [33]. Note that if luminance is in night vision range, the rods in the eye play a more important role, whereas in light vision range, cones mechanism comes into play. In the intermediate vision conditions, both cones

and rods work together. For the structure layer, we keep it unchanged to retain the integrity of the scene, as human vision is very sensitive to structure information [31], which has been demonstrated in recent subjective and objective studies [34]. For the detail layer, we enhance it to enrich the visual effect of LDR images after TM. These three layers are interrelated, restricted and promoted mutually, and together constitute the validation of micro model. And the specific procedure is described as follows.

As a non-iterative filtering based on local nonlinearity, bilateral filter is an effective method for multi-layer decomposition and is utilized in this paper [21], [35]. And the bilateral filter can be defined as:

$$BF(I) = \frac{\sum_{p\in\Omega} f(p-b)g(I_p - I_b)I_p}{\sum_{p\in\Omega} f(p-b)g(I_p - I_b)} \tag{5}$$

where I is the input image, *BF(I)* denotes the output image for I, Ω neighbors domain window, f is a Gaussian function in the spatial domain, g is a Gaussian function in the intensity domain, I_p and I_b are values of pixel p and b, respectively.

For different scenes, the dynamic range fluctuates significantly. To avoid its influence, a logarithm transform is used to non-linearly normalize the luminance channel as:

$$V_{nor} = \frac{\log(V_{HDR} + \delta) - \log(V_{\min} + \delta)}{\log(V_{\max} + \delta) - \log(V_{\min} + \delta)} \tag{6}$$

where V_{HDR} is luminance value of HDR images, δ is minimal value to avoid zero value. Here, we set $\delta = 0.00001$. This step simulates the response of human eyes to luminance, and initially reduces the dynamic range.

A piecewise invariant detail layer and a piecewise smooth base layer can be obtained by applying the bilateral filter to V_{nor}. First scale decomposition:

$$V_b = BF(V_{nor}) \tag{7}$$

$$V_D = V_{nor} - V_b \tag{8}$$

where V_D is the detail layer. After the first-level decomposition, the detail information is retained in the detail layer (V_D), and the main structure information is transferred to the bottom layer (V_b). Second-level decomposition, the bilateral filter is applied to V_b

$$V_B = BF(V_b) \tag{9}$$

$$V_S = V_b - V_B \tag{10}$$

where V_S is the structure layer which saves the structure information of the image, V_B is the brightness layer which included local mean luminance. To sum up, two-level decomposition schemes produces V_D, V_S,V_B. The relationship between them is given as:

$$V = \sum (V_D, V_S, V_B) \tag{11}$$

We can get three layers, including detail layer (V_D), structure layer (V_S) and brightness layer (V_B). And then different strategies are adopted for each layer by the property of HVS. Finally, the three decomposed layers are fused to obtain the luminance layer:

$$V_{micro} = V_S + \alpha_B \cdot (V_B)^{\gamma} + \beta_D \cdot \left[\operatorname{sign} V_D \left(\frac{|V_D|}{\max(|V_D|)} \right)^{\theta} \cdot \max(|V_D|) \right] \tag{12}$$

where α_B and β_D are weight coefficients for brightness layer (V_B) and detail layer (V_D). Through a large number of experiments, we get the right value of α_B and β_D. The value of α_B is not too high or too low to ensure the image brightness in the appropriate range. Similarly, images with high values of β_D will be distorted, and images with too low values of β_D will lose detail. Obviously, for the V_B, the gamma function is used to compress dynamic range. For the V_S, we keep it unchanged to retain the integrity of the scene. For the V_D, we use a nonlinear stretch function to enhance its details.

C. MACRO MODEL

For the visual cognitive mechanism, the HVS first rapidly and unconsciously produces a global perception of the entire image, and then gradually focuses on specific local areas [24], [25]. From a macroscopic perspective, a common sense is that high-quality images should consider the interaction and conversion between human subjective perception and device objective display. Although it is difficult to give a precise and rigorous definition of the interaction and conversion, we can exploit the existing theoretical analysis and the relevant results of subjective experiments to characterize and optimize them.

Intensity masking relates to the lower sensitivity exhibited by the HVS in darker and brighter areas. For LDR images, Weber-Fechner's law, which states that the minimum perceivable visual stimulus difference increases with background luminance, describes this masking phenomenon. Subsequently, relevant scholars have conducted in-depth research on this issue and point out that the overall effect of intensity masking on the distortion sensitivity of the HVS is a U-shaped curve that defines the maximum amount of distortion tolerated for a given average intensity value for HDR images [36]. Corresponding intensity-dependent quantization(IDQ) profile for HDR image and LDR image is shown in Fig. 4. Obviously, a higher visibility threshold occurs in either very dark or very bright regions in an image, and a lower occurs in regions with medium luminance (Weber-Fechner's) regions, which is consistent with the human subjective perception.

Besides this luminance sensitivity of human eyes, the property of device objective display also must be taken into account. Studies have shown that the corresponding relationship has the Gamma curve correlation between objective display brightness and subjective perception brightness as shown in Fig. 5 [28]. The horizontal direction shows a linear increase in natural brightness. And the vertical direction shows psychological perceived of luminance. Psychological perceived of luminance are not exactly the same as the linear increase in natural brightness. As we can see from the Fig. 5,

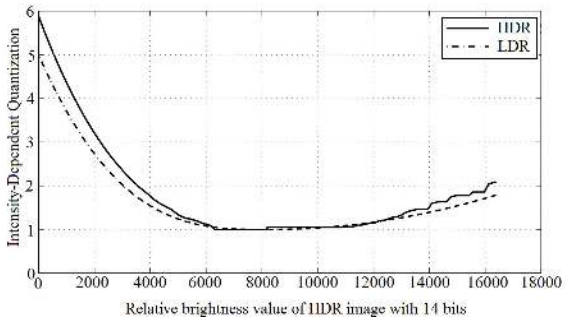


FIGURE 4. Intensity-dependent quantization profile for HDR image and LDR image [36].

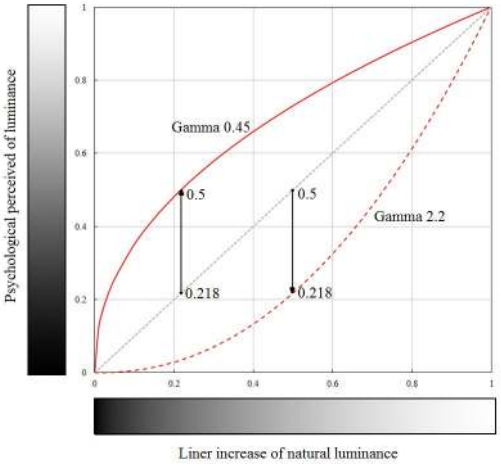


FIGURE 5. Gamma curve correlation between objective display brightness and subjective perception brightness.

the objective display brightness is closer to subjective perception brightness after gamma correction.

Hence, we utilize these two characteristics to adjust the global luminance for better visual fidelity as follows.

$$V_{macro} = \lambda \cdot V_{HDR}^{\gamma} \tag{13}$$

where V_{HDR} is the luminance value of the HDR image, V_{macro} is the luminance value after compression with the macro model. The setting of parameter λ and γ directly affects the actual performance, and will be discussed subsequently. To improve the generalization ability for the dynamic range of brightness, we combine parameter setting with image content information. Here, parameter λ is used to control the overall luminance of the image. Through a large number of experimental data analysis, we get the empirical formula as follows:

$$\lambda = \alpha_1 \cdot V_{mean}^{\beta_1} \tag{14}$$

where V_{mean} is average luminance of HDR image. α_1 and β_1 are empirical coefficients.

Parameter γ is important for image adjustment. Due to the variation of brightness intensity, it is necessary to select an appropriate γ value according to the actual scene. we set the value of parameter γ by analyzing the proportion of low-value pixels.

$$\gamma = \alpha_2 \cdot \frac{\sum p_i}{\sum p} + \beta_2 \tag{15}$$

where p is the number of image pixel, p_i is the number of pixel which below the average luminance. α_2 and β_2 are empirical coefficient. Through data analysis and fitting, we set $\alpha_1 = 0.5414$, $\beta_1 = -0.142$ and $\alpha_2 = -0.5$, $\beta_2 = 0.59$ in our method.

D. FUSION AND RECONSTRUCTION

After parallel processing of the two models, we can obtain corresponding intermediate results with different characteristics. For micro model, it enriches the local visual information but maybe result in halo artifact. In contrast, it avoids halo artifact but losses a part of detail information for macro model. Therefore, we fuse these results to achieve the expecting effect as follows.

$$V_{LDR} = \omega_1 \cdot V_{macro} + \omega_2 \cdot V_{micro} \tag{16}$$

where V_{LDR} represents the luminance channel, ω_1 and ω_2 represents the fusion coefficient. Here, we set $\omega_1 = 0.5$ and $\omega_2 = 0.5$ based on experience. Through a large number of experiments, such settings of ω can be better compatible with the advantages of the macro and micro models. It not only preserves the details of the image, but also maintains the global brightness balance, and can effectively remove the halo.

In order to prevent over-saturation and make the image more natural, we have adjusted the saturation of the appropriate.

$$S_{LDR} = \rho \cdot S_{HDR} \tag{17}$$

where S_{LDR} denotes the Saturation channel of final TM LDR image. S_{HDR} denotes the Saturation channel of HDR image. Parameter ρ is used to control the saturation and is set in the range of 0.8 and 1 with experience. In our method, we set $\rho = 0.9$, it will ensure that the output image is more natural.

With the Hue channel of HDR image, S_{LDR}, and V_{LDR}, the TM LDR image can be obtained in HSV space. Finally, the output LDR image is obtained by color space conversion.

III. EXPERIMENT RESULTS AND ANALYSIS

To verify the effectiveness of the proposed TM method, an extensive set of experiments were carried out on the MATLAB R2016a platform and Win7 operating system. Firstly, the performance with respect to subjective and objective effects are validated by numerical HDR images as shown in table 1 [37], [38]. And then, the comparison with the state-of-the-art TM methods is presented in detail, including "Khan *et al.*" [17], "Reinhard *et al.*" [18], "Durand *et al.*" [21], "Li *et al.*" [39], "Liang *et al.*" [40], "Bansal *et al.*" [41]. The source codes of these comparison methods are publicly available in the "HDR-Toolbox" [42] or provided in the authors' homepages, and we use the default

FIGURE 6. Local comparison results with different TM methods for image *"Design Center"*.

TABLE 1. List of the test HDR images.

No.	Name	*D*	Size	No.	Name	*D*	Size
1	Belg.	4.7	769×1025	11	Yose.	4.2	1000×664
2	Tint.	2.5	1000×750	12	Atri.	6.6	1016×760
3	Pill.	2.7	612×920	13	Cath.	4.7	1023×767
4	Desk	5.9	852×1136	14	Syna.	2.6	769×1025
5	Kitc.	4.8	664×1000	15	Rose.	7.4	480×720
6	Bott.	4.9	688×912	16	Tree	6.1	906×928
7	Loun.	6.3	664×1000	17	Lake	6.0	2353×4183
8	Atri.	4.5	1016×760	18	Bar	5.2	2412×4288
9	Design.	4.4	1000×656	19	Flam.	7.4	2848×4288
10	Desk(2)	7.0	874×644	20	River	3.6	1332×2000

D represents dynamic rang.

parameters of those codes which were optimized by the authors.

A. SUBJECTIVE ASSESSMENT

An important task of TM is to preserve all the visual information and local structure of the original HDR images. In order to make the results more convincing, different scenarios and different dynamic ranges are selected to verify our method, including indoor light scene (as shown in Fig. 6 with dynamic range is 5.9), indoor dark scene (as shown in Fig. 7 with dynamic range is 4.4), outdoor light scene (as shown in Fig. 8 with dynamic range is 5.2) and outdoor dark scene(as shows in Fig. 9 with dynamic range is 3.6). These scenes typically contain a lot of details and bright/dark areas, which can well verify the effectiveness of our method.

It can be seen from Fig. 6 and 7, the whole structure and a lot of details of the image are preserved through micro model of our method. In terms of structural retention, our method could retain structure more complete. Other methods exist problems of losing structure information (e. g. Fig. 6 (b), (c), (e), (f) and Fig. 7 (b), (c), (g)). In terms of detail retention, our method retains more details which shows in the zoom-in of Fig. 6 and Fig. 7. In Fig. 6(e) and 6(g), the details of windows are lost in the result of Li and Khan. In Fig. 7(d), a distinct halo phenomenon around the pen is appeared in Li's result and we avoided this problem. In Fig. 7(e), Liang's result lost details. In Fig. 7(f), the luminance of Bansal's result is too high to conform to the actual situation, for example, the drawer region should be darker.

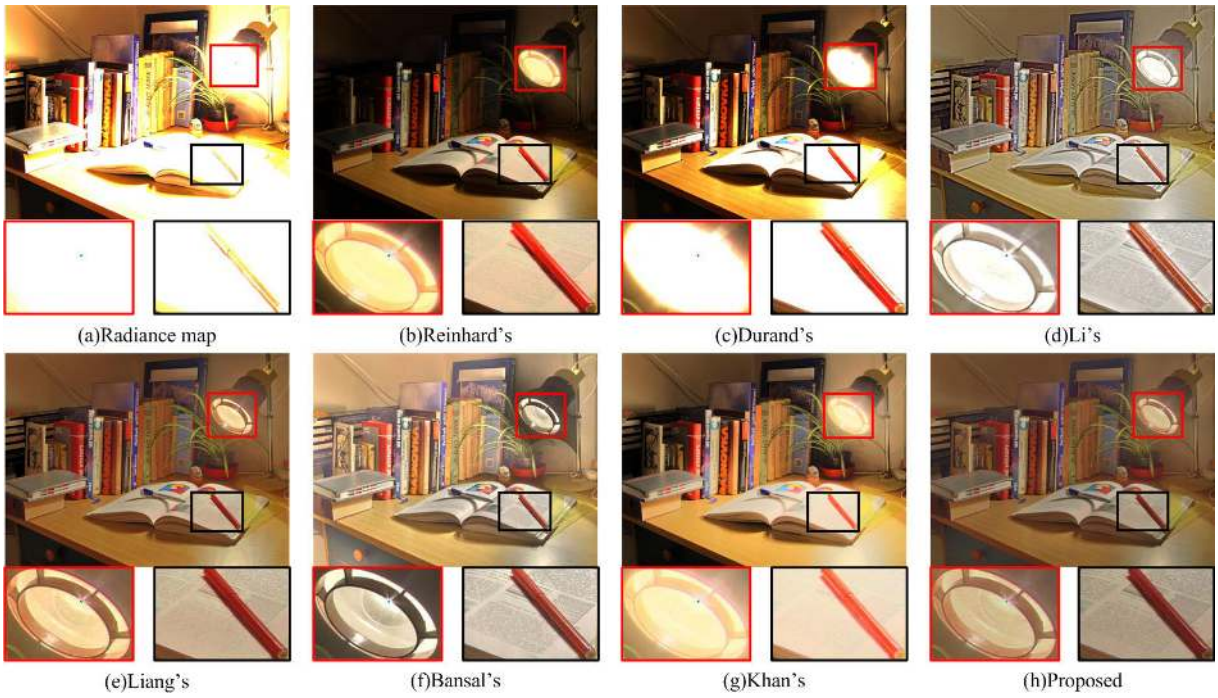


FIGURE 7. Local comparison results with different TM methods for image "D*esk*".

In addition to maintaining structure and detail, the overall effect of the TM should be consistent with the perception of the human eyes. Through the macro model of our method, the luminance of TM result will be in a suitable range which not too bright or too dark. Fig. 8 and 9 show the comparison of TM results on two outdoor scenes. The results of we proposed are close to real perception. For example, the white clouds are grayish(in Fig. 8) and sky is dim(in Fig. 9) in our result, and it is more consistent with reality scene and the human perception.

To sum up, our method mimics the area where normal human eyes see images after natural adaptation, achieving the purpose of TM.

B. OBJECTIVE ASSESSMENT

In order to avoid the influence of subjective preferences and environment, we also used information entropy, variance and tone mapping quality index (TMQI) [43] to objectively evaluate images. The computation of image variance is given as follows:

$$V = \sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left[f(i,j) - \frac{1}{M\cdot N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} f(i,j)\right]^2 \tag{18}$$

where *f(i,j)* is the matrix corresponding to image pixel. M and N are the length and width of image. Variance represents the contrast of an image. The greater variance, the greater contrast of an image. The computation of information entropy is given as follows:

$$H = -\sum_{i=0}^{255} p_i \cdot \log_2(p_i) \tag{19}$$

where p_i refers to the proportion of pixels in the image with gray value i. Information entropy is used to describe how much information an image contains. The bigger information entropy, the more information the image contains and the richer the details of the image.

TMQI is a widely accepted method of tone mapping evaluation. This method first evaluates the structural fidelity and naturalness of tone mapping images. The naturalness index provides useful information regarding the correlations between image naturalness and different image attributes, and can be computed as [43]:

$$N = P_m \cdot P_d / K \tag{20}$$

where K is a normalization factor, and P_m and P_d are Gaussian and Beta probability density functions respectively.

For structural similarity, TMQI first calculates the local similarities between corresponding patches x and y of HDR and LDR image pairs as [43]:

$$S = \frac{2\sigma_x \cdot \sigma_y + C_1}{\sigma_x^2 + \sigma_y^2 + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x \cdot \sigma_y + C_2} \tag{21}$$

where σ_x, σ_y and $\sigma_{x,y}$ are the local standard deviations and cross correlation between the corresponding HDR and LDR patches, and C_1 and C_2 are positive stabilizing constants.

FIGURE 8. Global comparison results with different TM methods for image "*Bar Harbor Sunrise*".

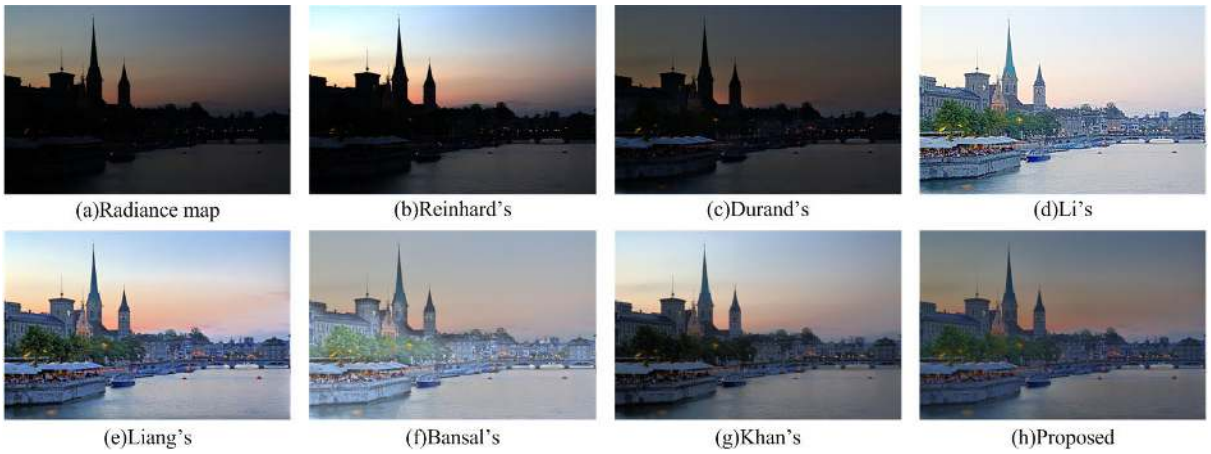


FIGURE 9. Global comparison results with different TM methods for image "*River*".

Then the power function is used to adjust the two indicators, and the average value is taken to get the final score [43].

$$Q = a \cdot S^{\alpha} + (1 - a) \cdot N^{\beta} \tag{22}$$

where a adjusts the relative importance of the two components, and α and β determine their respective sensitivities. The higher of TMQI value, the better the quality of tone mapping image, and vice versa.

For the images shown in the paper, we compare our method with the six above-mentioned algorithms using TMQI naturalness, structured similarity, and final score, in Tables 2, 3, and 4 respectively. In each table, the winner algorithms score is shown in bold font. In terms of naturalness score (Table 2), our method can rank the top two in three images. In terms of structured similarity score (Table 3), our method produces best results for all images. In terms of final score (Table 4), our method is also excellent. There are two images ranked first and others ranked second.

In order to make sure that our method works effective, twenty HDR images in a high dynamic database are randomly selected for testing. The HDR image information is shown in Table 1. In order to show the comparison results of twenty images more intuitively, we use the scatter plot to show the naturalness score, structural similarity score and final score of TMQI, which shows in Fig. 10, 11 and 12. Obviously, the method we proposed get the highest score in most images.

TABLE 2. Naturalness score of TMQI for the test images.

Method	Desk	Design.	Bar.	River
Reinhard's	0.692	0.010	0.029	0.001
Durand's	0.032	**0.438**	0.001	0.025
Li's	0.488	0.027	0.059	0.103
Liang's	0.277	0.040	**0.121**	0.062
Bansal's	**0.980**	0.133	0.116	0.069
Khan's	0.405	0.009	**0.209**	**0.168**
Proposed	**0.670**	**0.339**	0.120	**0.107**

Fig. 13 shows the average score of TMQI including naturalness, structural similarity and final score. More specifically, the average score of test images is shown in Table 5. In the table, the two highest scorers are shown with bold font.

We can see that our method rank first from Fig.13 and Table 5. And the score of our method has improved in natural-

TABLE 3. Structural similarity score of TMQI for the test images.

Method	Desk	Design.	Bar.	River
Reinhard's	0.815	0.604	0.784	0.488
Durand's	0.817	0.697	0.617	**0.651**
Li's	**0.826**	0.100	0.782	0.527
Liang's	0.812	0.731	**0.785**	0.532
Bansal's	0.810	**0.856**	0.769	0.557
Khan's	0.799	0.417	0.748	0.628
Proposed	**0.862**	**0.876**	**0.847**	**0.651**

TABLE 4. Final score of TMQI for the test images.

Method	Desk	Design.	Bar.	River
Reinhard's	0.906	0.688	0.760	0.645
Durand's	0.771	**0.829**	0.693	0.718
Li's	0.875	0.412	0.770	0.699
Liang's	0.832	0.748	0.789	0.689
Bansal's	**0.947**	0.812	0.783	0.700
Khan's	0.853	0.614	**0.799**	**0.752**
Proposed	**0.915**	**0.862**	**0.806**	**0.744**

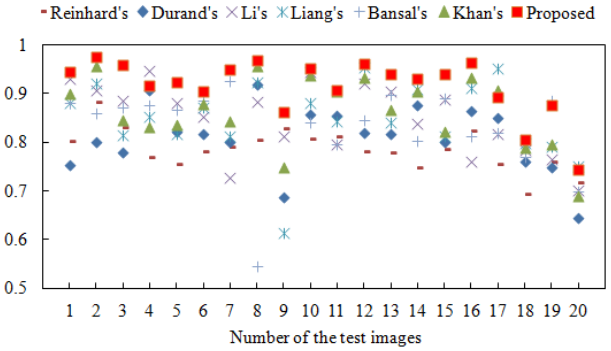


FIGURE 10. Naturalness score of TMQI for the test images.

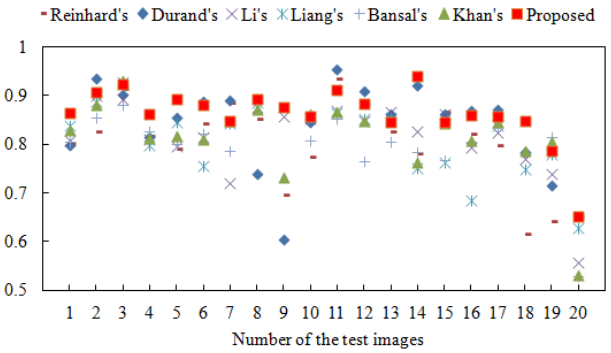


FIGURE 11. Structural similarity score of TMQI for the test images.

ness, structural similarity and final score. The improvement of structural similarity indicates that our micro model is effective. Meanwhile, the role of macro model is reflected in the improvement of naturalness. Furthermore, both the micro and

TABLE 5. Average score of different objective evaluation for test images.

Method	Variance	Entrop	S	N	Q
Reinhard's	33.9	6.77	0.801	0.167	0.786
Durand's	57.8	6.78	**0.825**	0.199	0.809
Li's	45.0	**7.16**	0.814	0.395	0.845
Liang's	42.9	7.12	0.791	0.435	**0.848**
Bansal's	38.5	6.89	0.735	0.356	0.809
Khan's	**64.4**	6.86	0.817	**0.474**	0.864
Proposed	**62.4**	**7.82**	**0.861**	**0.694**	**0.915**

S, N, Q represent structural similarity score, naturalness score, final score of TMQI.

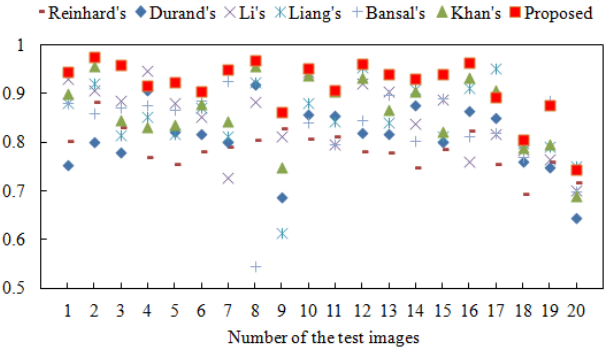


FIGURE 12. Final score of TMQI for test images.

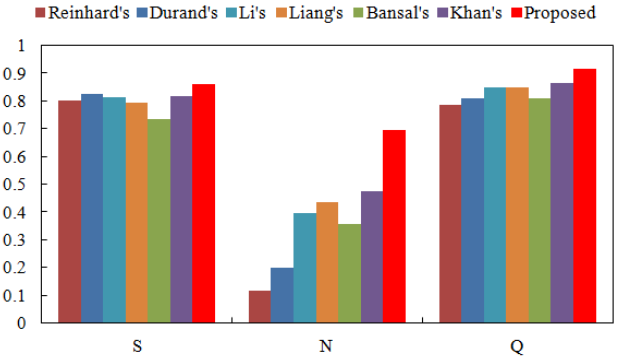


FIGURE 13. Average score of TMQI for test images, S is structural similarity score, N is Naturalness score, Q is final score.

macro models are processed in parallel to ensure the integrity and subjective consistency of scene information, so that final score of our method could be improved. Moreover, in the evaluation of information entropy and variance, the score of we proposed method also ranks high.

Furthermore, different scenarios and dynamic ranges images have been selected in the test to verify our method, including indoor light scene (as shown in Fig. 6 with dynamic range is 5.9), indoor dark scene (as shown in Fig. 7 with dynamic range is 4.4), outdoor light scene (as shown in Fig. 8 with dynamic range is 5.2) and outdoor dark scene(as shown in Fig. 9 with dynamic range is 3.6). Due to space limitations, other test images only show objective evaluation scores. As we can see from the results, our method performs

excellent in different scenarios and dynamic ranges. Besides, the size of the image does not affect the mapping effect, but only the operating efficiency, the bigger the image, the longer it takes.

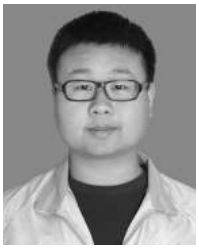
IV. CONCLUSION

In this paper, a novel tone mapping(TM) method based on macro-micro modeling is proposed, to visualize high dynamic range(HDR) images effectively on existing display devices. On the one hand, the micro model is designed with multi-layer decomposition technology to characterize the properties of brightness, structure, and detail for HDR images. And then different strategies are adopted for each layer by the human visual system(HVS) to reduce the overall brightness contrast and retain as much scene information. On the other hand, the macro model is designed with a scene content-based global operator to adaptively adjust the scene brightness. Finally, the desired TM image are fused with the results of the two models. Experiments results show that the proposed method achieves visually compelling results with little exposure imbalance and halo artifact, and is superior to the current state-of-the-art TM methods in both subjective and objective evaluations.

REFERENCES

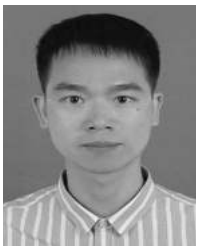
[1] J. M. Dicarlo and B. A. Wandell, "Rendering high dynamic range images," *Proc. SPIE*, pp. 392–402, May 2000.
[2] G. Yue, C. Hou, K. Gu, S. Mao, and W. Zhang, "Biologically inspired blind quality assessment of tone-mapped images," *IEEE Trans. Ind. Electron.*, vol. 65, no. 3, pp. 2525–2536, Mar. 2018.
[3] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, "Perception-based tone reproduction," in *High Dynamic Range Imaging: Acquisition, Display, Image-Based Lighting*, 2nd ed. San Mateo, CA, USA: Morgan Kaufmann, 2010, ch. 7, pp. 237–243.
[4] V. Hulusic, K. Debattista, G. Valenzise, and F. Dufaux, "A model of perceived dynamic range for HDR images," *Signal Process. Image Commun.*, vol. 51, pp. 26–39, Feb. 2017.
[5] Y. Li, Y. Sun, M. Zheng, X. Huang, G. Qi, H. Hu, and Z. Zhu, "A novel multi-exposure image fusion method based on adaptive patch structure," *Entropy*, vol. 20, no. 12, p. 935, 2018.
[6] Y. Li, M. H. Zheng Hu, and H. Wang, "A novel multi-exposure image fusion approach based on parameter dynamic selection," in *Proc. Chin. Intell. Syst. Conf.*, 2019, pp. 411-419.
[7] H. Fan, D. Zhou, R. Nie, and C. Yu, "A color multi-exposure image fusion approach using structural patch decomposition," *IEEE Access*, vol. 6, pp. 42877–42885, 2018.
[8] C. A. Parraga and X. Otazu, "Which tone-mapping operator is the best? A comparative study of perceptual quality," *J. Opt. Soc. Amer. A*, vol. 35, no. 4, pp. 626–638, Apr. 2018.
[9] 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.
[10] E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. Vis. Comput. Graph.*, vol. 11, no. 1, pp. 13–24, Jan./Feb. 2005.
[11] A. Artusi, T. Pouli, F. Banterle, A.O. Akyüz, "Automatic saturation correction for dynamic range management algorithms," *Signal Process. Image Commun.*, vol. 63, pp. 100–112, Apr. 2018.
[12] D.-H. Lee, M. Fan, S.-W. Kim, M.-C. Kang, and S.-J. Ko, "High dynamic range image tone mapping based on asymmetric model of retinal adaptation," *Signal Process. Image Commun.*, vol. 68, pp. 120–128, Oct. 2018.
[13] G. W. Larson, H. Rushmeier, and C. Piatko, "A visibility matching tone reproduction operator for high dynamic range scenes," *IEEE Trans. Vis. Comput. Graph.*, vol. 3, no. 4, pp. 291–306, Dec. 1997.
[14] 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.
[15] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–10, 2008.
[16] C. Jung and K. Xu, "Naturalness-preserved tone mapping in images based on perceptual quantization," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 2403–2407.
[17] I. R. Khan, S. Rahardja, M. M. Khan, M. M. Movania, and F. Abed, "A tone-mapping technique based on histogram using a sensitivity model of the human visual system," *IEEE Trans. Ind. Electron.*, vol. 65, no. 4, pp. 3469–3479, Apr. 2018.
[18] 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.
[19] Y. Li, L. Sharan, and E. H. Adelson, "Compressing and companding high dynamic range images with subband architectures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 836–844, Jul. 2005.
[20] Q. Shan, J. Jia, and M. S. Brown, "Globally optimized linear windowed tone mapping," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 4, pp. 663–675, Aug. 2010.
[21] 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, Jul. 2002.
[22] 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.
[23] N. R. Barai, M. Kyan, and D. Androutsos, "Human visual system inspired saliency guided edge preserving tone-mapping for high dynamic range imaging," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 1017–1021.
[24] Q. Lu, W. Zhou, and H. Li, "A no-reference image sharpness metric based on structural information using sparse representation," *Inf. Sci.*, vol. 369, pp. 334–346, Nov. 2016.
[25] S. H. Bae and M. Kim, "A novel image quality assessment With globally and locally consilient visual quality perception," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2392–2406, May 2016.
[26] S. Bi, X. Han, and Y. Yu, "An L_1 image transform for edge-preserving smoothing and scene-level intrinsic decomposition," *ACM Trans. Graph.*, vol. 34, no. 4, pp. 1–78, Jul. 2015.
[27] J. Chang, R. Cabezas, and J. W. Fisher, "Bayesian nonparametric intrinsic image decomposition," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 704–719.
[28] K. Yang, H. Li, H. Kuang, C. Li, and Y. Li, "An adaptive method for image dynamic range adjustment," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 29, no. 3, pp. 640–652, Mar. 2018.
[29] M. Manchanda and R. Sharma, "Fusion of visible and infrared images in HSV color space," in *Proc. 3rd IEEE Int. Conf. Comput. Intell. Commun. Technol. (CICT)*, Jul. 2017, pp. 1–6.
[30] A. Ford and A. Roberts, *Color Space Conversions*. London, U.K.: Westminster Univ., 1998.
[31] X. S. Zhang and Y. J. Li, "A retina inspired model for high dynamic range image rendering," in *Proc. Int. Conf. Brain Inspired Cogn. Syst.*, Nov. 2016, pp. 68–79.
[32] D. C. Kim, J. H. Yoo, W. H. Choe, and Y. H. Ha, "Visibility enhancement of mobile device through luminance and chrominance compensation upon hue," in *IEEE Trans. Ind. Electron.*, vol. 64, no. 4, pp. 3039–3047, Dec. 2016.
[33] M. K. Kundu and S. K. Pal, "Thresholding for edge detection using human psychovisual phenomena," *Pattern Recognit. Lett.*, vol. 4, no. 6, pp. 433–441, 1986.
[34] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," *IEEE Trans. Multimedia*, vol. 18, no. 3, pp. 432–443, Mar. 2016.
[35] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.
[36] Y. Zhang, M. Naccari, D. Agrafiotis, M. Mrak, and D. R. Bull, "High dynamic range video compression exploiting luminance masking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 950–964, May 2015.
[37] *High Dynamic Range Image Examples*. Accessed: Jun. 1, 2018. [Online]. Available: http://www.anyhere.com/gward/hdrenc/pages/originals.html?tdsourcetag=s_pctim_aiomsg
[38] *The HDR Photographic Survey*. Accessed: Jun. 1, 2018. [Online]. Available: http://rit-mcsl.org/fairchild//HDRPS/HDRthumbs.html

[39] H. Li, X. Jia, and L. Zhang, "Clustering based content and color adaptive tone mapping," *Comput. Vis. Image Understand.*, vol. 168, pp. 37–49, Mar. 2018.
[40] Z. T. Liang, J. Xu, D. Zhang, Z. S. Cao, and L. Zhang, "A hybrid l1-l0 layer decomposition model for tone mapping," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 4758–4766.
[41] N. Bansal and S. Raman, "Regularized tone mapping using edge preserving filters," in *Proc. 21st Nat. Conf. Commun. (NCC)*, May 2015, pp.1-6.
[42] F. Banterle, A. Artusi, K. Debattista, and A. Chalmers, "A brief overview of the MATLAB HDR toolbox," in *Advanced High Dynamic Range Imaging: Theory and Practice*, 1st ed. Natick, MA, USA: A K Peters, 2011, pp. 233–248.
[43] 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.



DISHENG MIAO received the B.Sc. degree from the China University of Petroleum, in 2015. He is currently pursuing the M.Eng. degree in instrument and meter engineering with Zhengzhou University, under the supervision of Prof. Z. J. Zhu. His research interests include tone mapping, visual attention, image segmentation, neural network, and deep learning.

ZHONGJIE ZHU received the Ph.D. degree in electronics science and technology from Zhejiang University, China, in 2004. He is currently a Professor with the Faculty of Electronics and Information Engineering, Zhejiang Wanli University, China. His research interests mainly include video compression and communication, image analysis and understanding, watermarking and information hiding, and 3D image signal processing.



YONGQIANG BAI received the B.S. and M.S. degrees from ZhengZhou University, China, in 2006 and 2009, respectively, and the Ph.D. degree from Ningbo University, China, in 2019. He is currently a Researcher with Zhejiang Wanli University, China. His research interests mainly include data hiding and image processing.

GANGYI JIANG received the M.S. degree from Hangzhou University, China, in 1992, and the Ph.D. degree from Ajou University, South Korea, in 2000. He is currently a Professor with the Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include digital video compression and communication, multi-view video coding, image-based rendering, and image processing.

ZHIYONG DUAN received the Ph.D. degree in microelectronics and solid state electronics from Shanghai Jiaotong University, China, in 2005. He is currently a Professor with the Physical Engineering College, Zhengzhou University. His research interests include nanoimprint lithography, MEMS/NEMS, signal processing, and so on.