

Detail-preserving underexposed image enhancement via optimal weighted multi-exposure fusion

Shiguang Liu, *Member, IEEE*, Yu Zhang

Abstract—Underexposed image enhancement aims at revealing hidden details that are barely noticeable in underexposed images due to low light conditions. Previous work may inevitably wash out some weak edges and lose details when handling several underexposed images. To deal with these problems, this paper presents a detail-preserving underexposed image enhancement method based on a new optimal weighted multi-exposure fusion mechanism. Providing an input underexposed image, we propose a novel multi-exposure image enhancement method which can generate a multi-exposure image sequence. However, none of these images are good enough, as images with high exposure have good brightness and color information, whereas sharp details are better preserved in the images with lower exposure. In order to preserve details and enhance the blurred edges, we propose to solve an energy function to compute the optimal weight of the three measurements: local contrast, saturation, and exposedness. Then a weighted multi-exposed fusion method is used to generate the final image. Since the proposed approach is computationally light-weight, it is possible to implement it on mobile devices such as smart phones and compact cameras. Various experiment results validate our new method.

Index Terms—Underexposed image enhancement, detail preserving, weighted multi-exposure fusion.

I. INTRODUCTION

IMAGE enhancement has become a crucial step in image processing. Images taken by consumer digital cameras usually lack details in the under-exposed areas if the camera has an improper setting or the light condition is poor [1] [2]. Due to low visibility, the image details are lost, and colors in images are washed out. To improve the visual appearance of the underexposed images, image enhancement technique emerges. Image enhancement can be widely used in color correction, contrast enhancement, and noise reduction, etc. Underexposed image enhancement aims at revealing hidden details that can be barely noticed in images which are captured in low-light-level photographic environment.

Researchers have investigated the problems of image exposure lightness for years, and have proposed many underexposed image enhancement methods. For example, histogram equalization [3]–[6] can better distribute the intensity values over the histogram in the underexposed images. Unfortunately, it would produce unrealistic effects. Relying on adjusting a single tone mapping curve to adapt the exposure to regions

This work was supported in part by the Natural Science Foundation of China under grant nos. 61672375 and 61170118.

Shiguang Liu was with School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, P.R. China (e-mail: lsg@tju.edu.cn).

Yu Zhang was with School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, P.R. China (e-mail: zhangyu_tju@tju.edu.cn).

throughout the process is also a common method. However, it is actually difficult for users to determine the optimal curve that best suits different input images by tuning two or more parameters. Moreover, even with an optimal curve, it may also fail to ensure that all regions are well-exposed. Single tone mapping curve may induce disturbing visual artifacts, such as distorted appearance. Recently, Yuan and Sun [6] have presented an automatic exposure correction method that is capable of estimating the desired exposure for different regions. This method demonstrates successful results in correcting exposure of some static images. However, their method did not consider denoising and noise would become noticeable after lighten the dark areas.

Different from tone mapping based methods, image fusion methods combine well-exposed regions together from an image sequence with different exposures. Most previous exposure fusion approaches have the following workflow: based on some defined well-exposedness measures [7]–[9], weighting maps are determined for each of the input images. With these weighting maps all images are fused to the final version. However, the main drawback of the image fusion schemes is that they require multiple images with different exposure information, they do not consider camera motion during the exposure time and therefore produces ghost artifacts [10], [11]. To deal with this problem, Im et al. [12] proposed a method that only use a single underexposed image as the input and removed ghost artifacts. They generated three LDR images by local histogram stretching, removed noise by combining high-frequency regions from a noisy LDR image and flat regions from the result of averaging filter image and finally fused LDR images to generate HDR images. However, because of the default weight parameters in fusion process, they can not preserve good intensity and color information while preserving sharp details, which is also a problem in the general image fusion approach. This means that they can not preserve color information and details at the same time. The other approaches do not select the best input image parts and combine them to one image. Instead, they directly formulate an energy whose minimizer gives the optimal result. They used variational method to automatically combining an exposure-bracketed pair of images within a single image that reflects the desired properties of each one [13]–[15]. Bertalmó et al. [15] introduced an energy functional consisting of two terms, one measuring the difference in edge information with the short-exposure image and the other measuring the local color difference with a warped version of the long-exposure image. However, it requires to impose a smoothness constraint on the resulting image for optimizing the energy immediately for the

image itself.

To address the above problems, we combine the two different image fusion schemes to preserve details, color and brightness at the same time. The problem lies in that the weight value of contrast, exposedness and saturation should change according to the contrast values, the exposedness values and the saturation values respectively to better represent the exposed regions, but the default weight can not achieve this goal. Different from using default weight, we proposed to solve an energy function to compute the optimal weight of the three measurements: local contrast, saturation, and exposedness. In this case, the three measures will have the best weights and thus preserve details, color and brightness at the same time. Therefore, the colors of the output image match closely to those images with high exposure, while retaining details from the images with low exposure.

In this paper, we explore the advantages of the two different image fusion schemes together to generate the fusion result which can therefore better preserve details. Instead of requiring multiple images with different exposure information as the input, our input is only a single underexposed image. Firstly, we construct multi-exposure image sequence containing different exposed versions of each underexposed image using a series of tone mapping curves. Guided by some visual perception quality measures which can represent the desirable visual appearance, we adaptively locate locally best exposed regions from all multi-exposure image sequences and then seamlessly integrate them into a well-exposed image through some weight values. These weight values can be solved by formulating as an energy function so that the colors of the output image match closely to those of images with high exposure, while retaining details from the images with low exposure. The main contributions of this paper can be summarized as follows:

- A novel detail-preserving underexposed image enhancement method is developed, which can better preserve details in images than state-of-the-art methods.
- We propose to solve weight values of three measurements by formulating an energy function so that the colors of the output image match closely to those images with high exposure, while retaining details from the images with low exposure.
- We design an optimal weighted multi-exposure enhancement mechanism to integrate visually best exposed regions derived from different tone mapping curves into a single well-exposed image.

II. RELATED WORK

Underexposed image enhancement is one of the essential research issues for image processing [16]. One of the simplest, efficient and most widely used techniques for image enhancement is histogram equalization [3]–[6]. Generally speaking, histogram equalization evens out the original intensity histogram as much as possible to increases the contrasts in image scene. Pizer et al. [3] redistributed the intensities using multiple adaptive histogram equalization techniques over distinct small regions. Although this method can improve local contrast and bring out more details, it has a tendency to overly

amplify noise in relatively homogeneous regions. To limit the amplification, Zuiderveld et al. [4] presented the contrast limited adaptive histogram equalization by setting a limit on the derivative of the slope of the transformation function, but this method requires accurate parameter setting, and may produce ghosting artifacts.

Tone mapping is another common dynamic range tuning technique [17]–[20], which attracts broad interests recently. Many different tone mapping operators have been suggested with different advantages and disadvantages. Here, we simply review current tone mapping operators in two categories, namely global operators and local operators. Global operators apply a spatially uniform remapping of intensity to compress the dynamic range [21]–[23]. The main advantage of these methods is high processing efficiency. However, these methods may sometimes fail to produce pleasing result. Local tone mapping operators apply a spatially varying remapping [20] [24]. The mapping changes for different regions in the image. It often yields more pleasing images, however, the result may still look unnatural. The operators employ very different techniques to compress the dynamic range: from bilateral filtering [25], which decomposes the image into edge-aware low and high frequency components, to compression in the gradient domain. The following two local operators are related to our method. Bennett and McMillan [26] presented to comprise an adaptive spatio-temporal accumulation filter for reducing noise and a scale adaptive LDR tone mapping approach for making previously unwatchable details visually noticeable. Although this approach is capable of improving the visibility of underexposed videos, it may output videos with visual artifacts, such as uneven exposure, and flickering artifacts, etc. The method by Li et al. [24] used a pyramidal image decomposition, and attenuated the coefficients at each level to compress the dynamic range. Our method is also pyramid-based, but it works on the coefficients of the different exposures instead of those of an in-between HDR image. We aim at creating pleasing images and try to reproduce as much detail and color as possible.

Image fusion techniques have been widely applied in the fields such as image correction [27], and video enhancement [28], etc. Image fusion based methods [18] [29] aimed to combine relevant information from multiple images taken from the same scene and then produce a fused image, which is more informative than each one. Most previous exposure fusion approaches have the following workflow: based on specific quality measures, they first determine for each pixel or region of the input images how much it should contribute to the final composite, then they fuse these pixels or regions to the resulting overall well-exposed image. In this mechanism, several input images should be prepared before image fusion. Hsieh et al. [30] used a linear function to fuse the input image and histogram equalization enhanced image to get a fused image. Pei et al. [31] generated two images, a histogram equalization enhanced image and a sharpened image using Laplacian operator, and then fused their discrete wavelet transform (DWT) coefficients to get a fused image with higher contrast and sharpness. However, due to the camera motion during the exposure time they will produce ghost artifacts

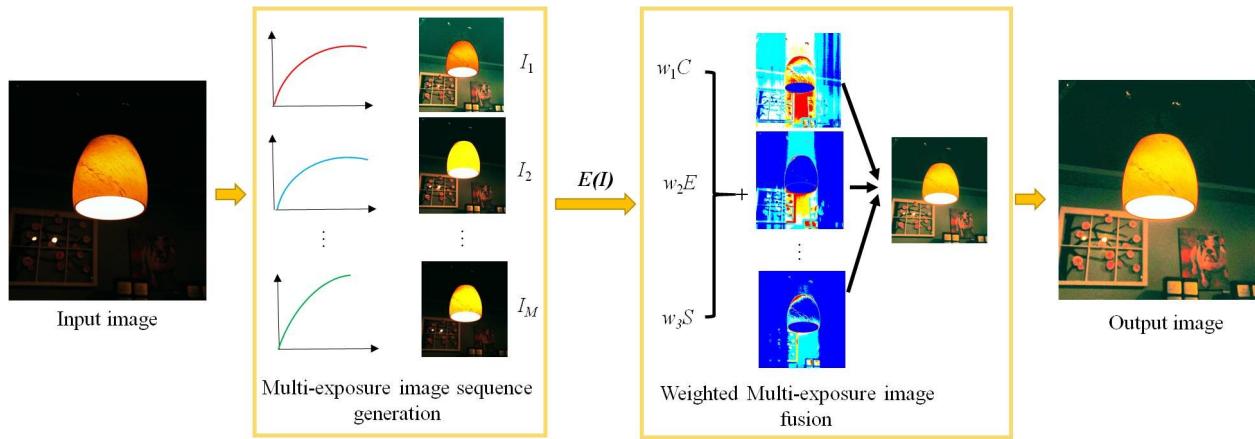


Fig. 1: The proposed novel framework of our optimal weighted underexposed image enhancement method.

[10], [11]. Instead of requiring multiple images with different exposure information as the input, some methods only use a single underexposed image [12], which remove the ghost artifacts and preserve edges. In contrast to the above approaches, the other approaches do not specify the quality of the input images first and then fuse them later on. Instead, they directly design the quality of the fusion result. In this regard, Raman and Chaudhuri [13] formulated an energy function whose minimizer gives the fused composite. However, optimizing the energy immediately for the image itself has two drawbacks: First, it restricts the possible model assumptions. Second, it requires to impose a smoothness constraint on the resulting image. This is not intuitive and may lead to over-smoothed results. A more suitable idea of Kotwal and Chaudhuri [14] is to express the composite image as a weighted average of the input images and to optimize for these weights. This still allows to directly model assumptions on the fusion result, but additionally opens the possibility to impose a smoothness constraint on the weight maps and not on the image itself.

III. OUR METHOD

Let I be the input underexposed image. For image I , the underexposed image enhancement can be formulated as a mapping function which can map the underexposed image I to a well-exposed image O . The great challenge lies in exploring a suitable mapping function that can restore all underexposed regions without introducing uneven exposure, and other visual artifacts.

Fig. 1 shows the workflow of our method. Firstly, we perform a multi-exposure image sequence generation by using a series of tone mapping curves to generate several images (I_1, I_2, \dots, I_M) with different exposure. Then, we formulate an energy function $E(I)$ to get several weight values (w_1, w_2, w_3) so that the colors of the output image match closely to those of images with high exposure, while retaining details from the images with low exposure. Finally, we fuse the image sequences with the weighted three measurements (C, S, E) and integrate the best exposed regions through the weight values into a well-exposed frame through four features.

A. Multi-Exposure Image Sequence Generation

Given an input underexposed image, we first remap it using a series of tone mapping curves to generate a multi-exposure image sequence which contains different exposed versions of the original image. Here, we adopt a LDR tone mapping operator to adjust different exposure images. The mapping function can be defined as:

$$f(I, \alpha) = I + \frac{I}{\sqrt{I^2 + \alpha^2}}(I_{\max} - I) \quad (1)$$

where $I \in [0, 1]$ denotes the input normalized intensity, and I_{\max} denotes the upper bound of the intensity range. α is the factor that controls the intensity climb. A small α corresponds to a large intensity increase, and there is not much increase in intensity for a large α .

For intensity, we manipulate the tone of the input image by mapping the intensity channel only and keeping the color channels unchanged. The intensity image is set to $I = (20R + 40G + B)/61$, where R, G and B refer to the RGB channels, respectively. Chrominance ratios are set to $(\rho_R, \rho_G, \rho_B) = (R, G, B)/I$. We then apply the mapping function on I to get mapped intensity I' . Finally, we multiply I' by the chrominance ratios (ρ_R, ρ_G, ρ_B) to obtain the output RGB channels.

In our work, we bound $\alpha \in [\alpha_{\min}, \alpha_{\max}]$, and the number of α is set to M . Within the range, we can obtain a series of mapping functions $\{f(I, \alpha_m)\}_{m=1}^M$ where

$$\alpha_m = \alpha_{\min} + (\alpha_{\max} - \alpha_{\min})[(m - 1)/(M - 1)] \quad (2)$$

After applying the series of mapping functions $\{f(I, \alpha_m)\}_{m=1}^M$ on each frame, we will get M multi-exposure image sequences. In our experiments, similarly to Zhang et al.'s method [26], α_{\min} and α_{\max} are set to 0.1 and 1.0, and M is set to 6. Fig. 2 shows the multi-exposure image sequence.

Then we locate all the best exposed regions based on some quantized visual quality measures. Inspired by Mertens et al.'s method [7], we choose three measures, namely local contrast, saturation, and exposedness. For local contrast C , we apply Laplacian filter on the input multi-exposure image sequence, and select the absolute value of the filtering result as the local

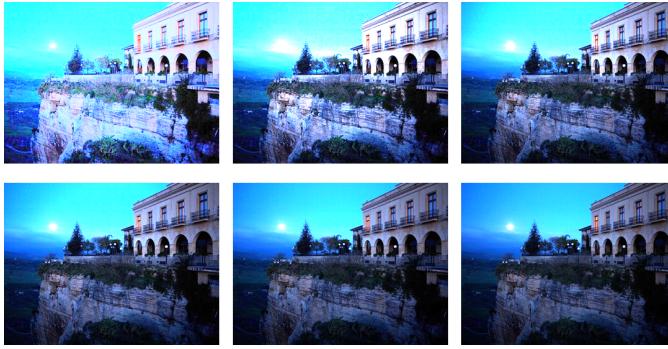


Fig. 2: Multi-exposure image sequence generation. From left to right, top to bottom are the images with α from 6 to 1. It is obvious that the higher value α is, the higher luminance and exposedness of the image will be.

contrast. The filtering result with high values means edges and textures. Then, saturation measure S can be estimated by the standard deviation with R , G , and B channels. Finally, in order to judge how much a pixel is exposed, we need to keep the intensities that are not near 0 (underexposed) or one (over-exposed). We weigh each intensity based on how close it is to 0.5 by computing the Gaussian-modeled distance as follows:

$$Ep = \exp \left\{ -\frac{(I_p - 0.5)^2}{2\sigma_e^2} \right\} \quad (3)$$

where I_p denotes the intensity of the pixel p , and σ_e is set to 0.25 in our experiments.

By combining these three measurements, we can determine the visual friendliness of a pixel. For an input image, we define the visual weight map as $W = C^{w_1} \times S^{w_2} \times E^{w_3}$, where C , E , and S denote quantized contrast, exposedness, and saturation, respectively. w_1 , w_2 and w_3 are exponents that control the contribution of each measure. If an exponent is set to zero, the corresponding measure will not be taken into account.

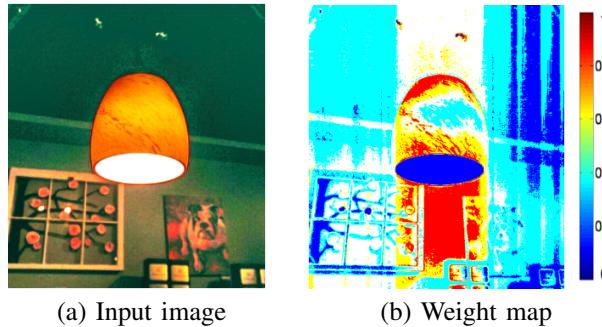


Fig. 3: Locating the best exposed regions. (a) is one frame from the multi-exposure image sequence, and (b) is the weight map. The higher weight value indicates that the pixel is in better exposure status.

Fig. 3 illustrates how to locate the best exposed regions. We first compute the weight map for each input image. Since higher visual friendliness indicates that the pixel is well exposed, we can then identify the best exposed regions based on the weight map. In this case, we visualize each normalized

weight map according to the rightmost color bar in Fig. 3(b). Finally, we integrate all the best regions of the multi-exposed image sequences into an intermedia well-exposed image.

By default, in many methods these three exponents are set to 1.0. However, default weight values may wash out some weak edges when handling severely underexposed images. In the multi-exposure image sequence, the three measurements cannot represent accurate visual quality measures. For example, images with low exposure contain sharp objects details, but low brightness and poor colors, therefore, the contrast values are high, while the exposedness values and the saturation values are low, and so the weight value of exposedness and saturation should be higher than the others simultaneously. Directly combining these three measurement by the default cannot better represent the exposed regions. To solve this problem, we propose to solve an optimal problem to obtain the weight values of the three measurements to preserve the edges in the result.

B. Optimal Weight Values Estimation

Our goal is to fuse the three measurements together to integrate them into a visual weight map and then fuse them with each multi-exposure image sequence to get the final output image. Basically, we propose to compute the visual weight map as:

$$W = C^{w_1} \times S^{w_2} \times E^{w_3} \quad (4)$$

where $w_i (i = 1, 2, 3)$ denotes the weight of each measurement. We constrain these weights to be non-negative and to sum up to 1, i.e. $w_i \geq 0$ and $\sum_{i=1}^3 w_i = 1$. This provides a close attachment to the input data that prevents undesirable effects such as color shifts or halos.

In the multi-exposed image sequence, none of these images are good enough: images with large α preserve good intensity and color information in the sense that the colors from the actual scene are retained, but may result in blurred details and noises. Images with small α preserve sharp details, but these images are usually dark and noisy. In this section, we propose an optimal approach for automatically combining three measurements of each multi-exposure images to reflect the optimal properties of each one. Specifically, they are optimal in the sense that the resulting weight values w_i is optimal for the following energy function:

$$\begin{aligned} E(w_1, w_2, w_3) = & -w_1 \int_{\Omega} g_{\sigma}(x, y) \cdot \Psi_{\lambda}(I(x) - I(y)) dx dy \\ & -w_2 \int_{\Omega} ((I_{Cb}(x) - 0.5)^2 + (I_{Cr}(x) - 0.5)^2) dx \\ & + w_3 \int_{\Omega} (E(x) - 0.5)^2 dx \end{aligned} \quad (5)$$

subject to

$$w_i \geq 0 \quad \text{and} \quad \sum_{i=1}^3 w_i = 1 \quad (6)$$

where $x = (x_1, x_2)$ denotes the position on the rectangular image domain. The first term is the contrast term, which penalizes uniform images much more than images with a high local contrast. Intuitively speaking, the energy favors solutions that differ a lot from pixel to pixel. Here, the locality is introduced

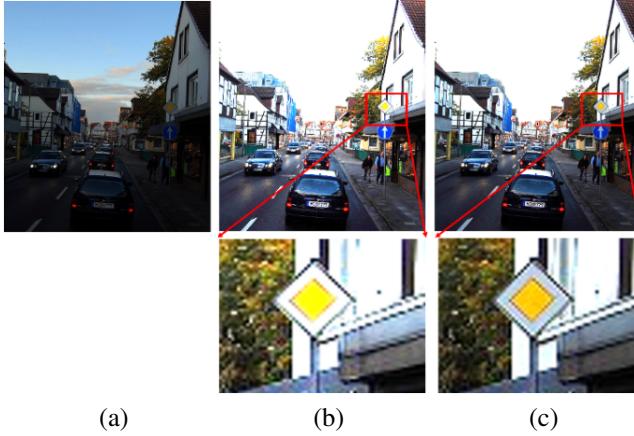


Fig. 4: Comparison between weight enhancement and without weight enhancement. (a) the original underexposed image. (b) straightforward default weight values yields unpleasing blurry edges. (c) using the enhancement weight, the edges are more clear, and the details are preserved well.

by the Gaussian weighting $g_\sigma(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(\frac{-|x-y|^2}{2\sigma^2}\right)$. Furthermore, $\Psi_\lambda(s) = \sqrt{s^2 + \lambda^2}$ is a function with a sigmoid-shaped derivative that shows connections to the nonlinear response of the visual system [30] [31]. The parameter λ allows to tune this nonlinear behavior. The second term in our energy functional is the saturation term. Here we transform the input images from the *RGB* to the *YCbCr* color space, I_{Cb} and I_{Cr} denote the *Cb* channel and *Cr* channel of I . This term favors values far away from grey values, and thus fuses images with vivid color. The third term is the exposure term, it favors the pixels with middle exposed values.

Minimizing Equation (5) is a standard linear programming problem, and we apply the simplex algorithm to solve it. After we get the weight values through minimizing the energy functional, we integrate them with the contrast, exposedness, and saturation and get the final visual weight maps. Fig. 4 shows the comparison results between weight enhancement and without enhancement. It can be seen that the default weight would leads to blurry edges, while using the weight in Equation (5) can well preserve details.

C. Multi-Exposure image Fusion

To implement the fusion between multi-exposure images I_t and I_t^M , we compose I_t and I_t^M into Lapacian pyramids $\{L[I_t]\}$ and $\{L[I_t^M]\}$. Meanwhile, corresponding Guassian pyramids $\{G[W_t]\}$ and $\{G[W_t^M]\}$ are constructed, where W_t and W_t^M are normalized weight map of I_t and I_t^M , respectively. By mixing the coefficients between the Laplacian pyramid and Guassian pyramid at each level independently, we can construct the output Laplacian pyramid $L[I_t']$. We define the l -th level of the output Laplacian pyramid as follows:

$$L_l[I_t']_p = G_l[W_t]_p L_l[I_t]_p \quad (7)$$

where p denotes the pixel coordinate. After all coefficients of the output Laplacian pyramid have been computed, we collapse the output pyramid to get the fusion result P_t . As

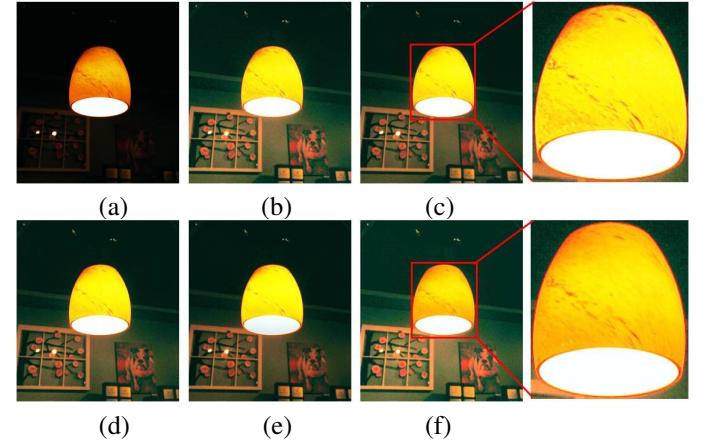


Fig. 5: Influence of different weight values. From left to right, top to bottom: the input underexposed image, weight values (1.0, 1.0, 1.0), weight values (1.0, 0, 1.0), weight values (0, 1.0, 1.0), weight values (1.0, 1.0, 0), weight values solved by our method.

demonstrated in Fig. 4(c), under our optimized weighted multi-exposure image fusion framework, we achieve the desired visually pleasing result with the multiscale fusion.

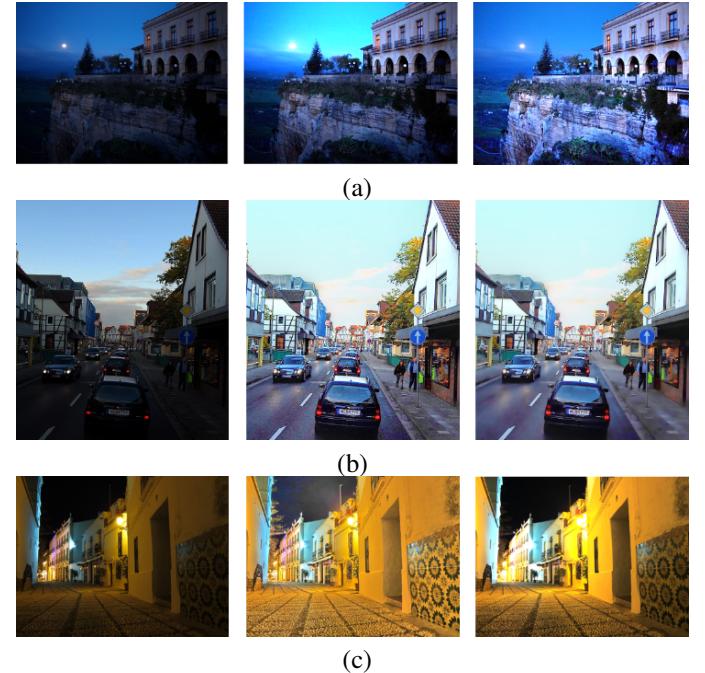


Fig. 6: The result comparison among different tone mapping methods. From left to right: the original input video frame which is underexposed, the result of Mertens et al.'s method [7], the result of our method.

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

A. Subjective Assessments

First, we illustrate the optimality of our model parameters. To this end, we depict several fused images for varying weight settings in Fig. 5. Fig. 5(a) is the original underexposed image,

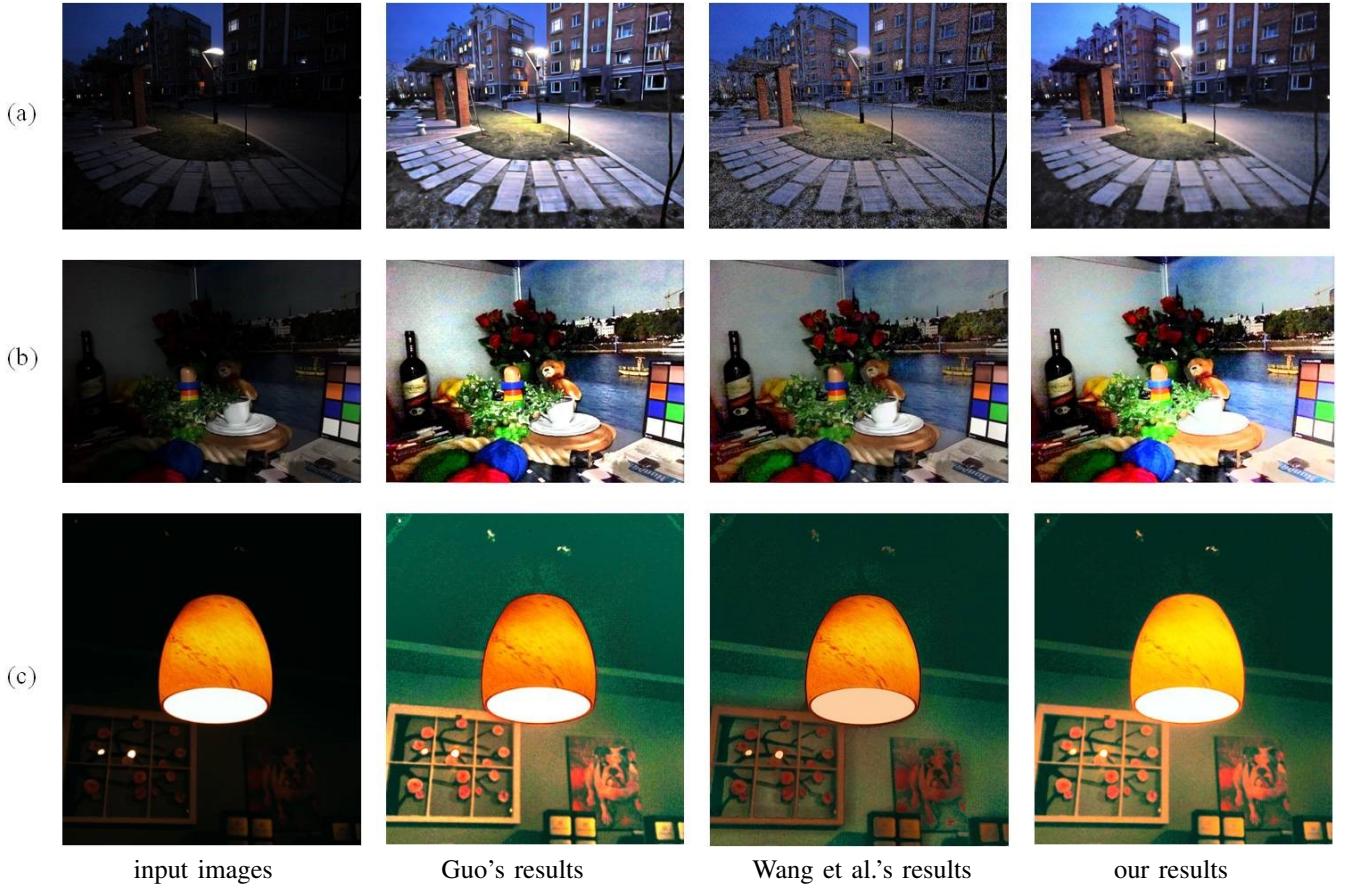


Fig. 7: The result comparison among different image enhancement methods. The first column shows the original images. The second column and the third column show Guo's results [32] and Wang et al.'s results [33], respectively. The last column shows our results. The enhancement result produced by Guo's method suffer from some artifacts, such as the over exposure in (a), the noise in the region of the wall of (b), and the noise around the lamp in (c). In contrast, our results can keep even exposure and show better visual appearance.

and we test different influence of weight values. The default weight values (1.0, 1.0, 1.0) suffer from blurred edges, and the other weight combinations suffer from different levels of overexposed regions. Our result can better preserve the details and can better reveal the exposure regions. To verify our weighted multi-exposure image fusion method can better preserve details than the other image fusion methods we compare the results of our methods with Mertens et al.'s method [7]. Fig. 6 shows the comparisons between our methods and Mertens et al.'s method [7]. Their results look either less vivid, or have lower global contrast than ours.

We now compare the performance of our algorithm with state-of-the-art approaches. Fig. 7 shows the comparisons between our methods and Guo's method [32] and Wang et al.'s method [33], respectively. The input images show poor-exposed scenarios. It can be observed that Guo's enhancement results suffer from different levels of underexposure and overexposure in some regions, while our method brings more visible details into poor-exposed areas while preserving the original appearance in well-exposed areas. Wang et al.'s results show poor luminance in the image scenes. In contrast, the results by our method show the best visual appearance. In addition, comparisons with other multiple exposed fusion

methods (Im et al.'s method [10], Bilcu et al.'s method [11] and Im et al.'s method [12]) are shown in Fig. 8. It can be observed that our method retain details from the images with low exposure and the colors of the output image match closely to those images with high exposure, while enhancing the blurred edges.

B. Objective Assessments

As subjective assessment depends on human visual system, it is hard to find an objective measure that is in accordance with the subjective assessment. Objective assessment is often used to explain some important characteristics of the image. We assess the detail enhancement through two objective fusion quality metrics: Q_0 [34] and $Q^{AB/F}$ [35]. Meanwhile, we assess the naturalness preservation through the proposed lightness-order-error (LOE) measure of Wang et al. method [33]. The quantitative LOE measure is based on the lightness order error between the original image and the enhanced version, the smaller the LOE value is, the better Naturalness the enhanced image will be kept. The index Q_0 evaluates the success of edge information transferred from the source images to the fused image. The index $Q^{AB/F}$ is designed

through modeling image distortion as a combination of three factors: loss of correlation, luminance distortion, and contrast distortion. The larger the Q_0 and $Q^{AB/F}$ value are, the better the fusion results will be. Tables I, II, and III give the three measurements for the images in Fig. 7. It can be concluded that the proposed image fusion method can better preserve the naturalness of the enhanced image (see the performance of LOE), it can also better preserve the edge information of source images (see the performance of Q_0) and introduces less color, contrast, and image quality distortions (see the performance of $Q^{AB/F}$). At the same time, we also calculate three quality metrics for images shown in Fig. 8 to better prove the advantages talked above. As can be seen from Table IV, our method gets lower LOE compared to all three methods, and gets higher Q_0 and $Q^{AB/F}$ than Im et al.'s method [10] and Bilcu et al.'s method [11] while performing no much difference compared to Im et al.'s method [12].

TABLE I
QUANTITATIVE ASSESSMENT OF
DIFFERENT ENHANCEMENT METHODS OF LOE

| Images | Guo [32] | Wang et al. [33] | Our method |
|-----------|----------|---------------------|------------|
| Fig. 7(a) | 33.1112 | 22.1236 | 15.3816 |
| Fig. 7(b) | 60.0947 | 50.8764 | 39.0381 |
| Fig. 7(c) | 38.9067 | 52.4568 | 22.9647 |

TABLE II
QUANTITATIVE ASSESSMENT OF
DIFFERENT ENHANCEMENT METHODS OF Q_0

| Images | Guo [32] | Wang et al. [33] | Our method |
|-----------|----------|---------------------|------------|
| Fig. 7(a) | 0.1371 | 0.1134 | 0.1465 |
| Fig. 7(b) | 0.1055 | 0.0831 | 0.1327 |
| Fig. 7(c) | 0.0544 | 0.0389 | 0.0991 |

TABLE III
QUANTITATIVE ASSESSMENT OF
DIFFERENT ENHANCEMENT METHODS OF $Q^{AB/F}$ FOR FIG. 7

| Images | Guo [32] | Wang et al. [33] | Our method |
|-----------|----------|---------------------|------------|
| Fig. 7(a) | 0.543 | 0.435 | 0.570 |
| Fig. 7(b) | 0.510 | 0.562 | 0.614 |
| Fig. 7(c) | 0.592 | 0.318 | 0.676 |

TABLE IV
QUANTITATIVE ASSESSMENT OF DIFFERENT
IMAGE FUSION METHODS OF LOE, Q_0 AND $Q^{AB/F}$ FOR FIG. 8

| method | LOE | Q_0 | $Q^{AB/F}$ |
|----------------------------|--------|--------|------------|
| Im et al.'s method [10] | 1.6125 | 0.5586 | 0.1651 |
| Bilcu et al.'s method [11] | 1.0744 | 0.8544 | 0.2906 |
| Im et al.'s method [12] | 0.7417 | 0.8611 | 0.4819 |
| our method | 0.6141 | 0.8576 | 0.4920 |

C. Performance Evaluation

We tested our new underexposed image enhancement approach on both personal computer (PC) and mobile devices. For mobile device, we constructed an application with a simple

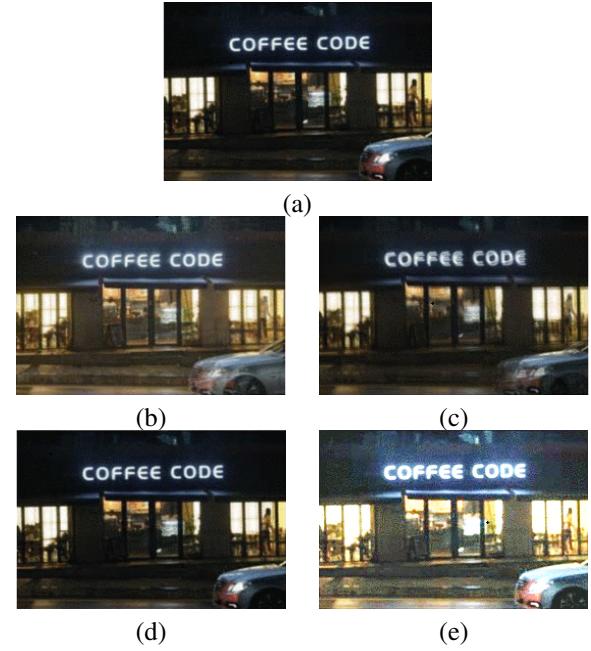


Fig. 8: The result comparison among different image fusion methods. (a) the original input underexposed image, (b) the result of Im et al.'s method [10], (c) the result of Bilcu et al.'s method [11], (d) the result of Im et al.'s method [12], and (e) the result of our method.

user interface, with the function of reading image, processing image and showing execution time, etc. Figure 9 shows the interface of our application. The size of the tested images ranges from 230x254 to 1920x1280.

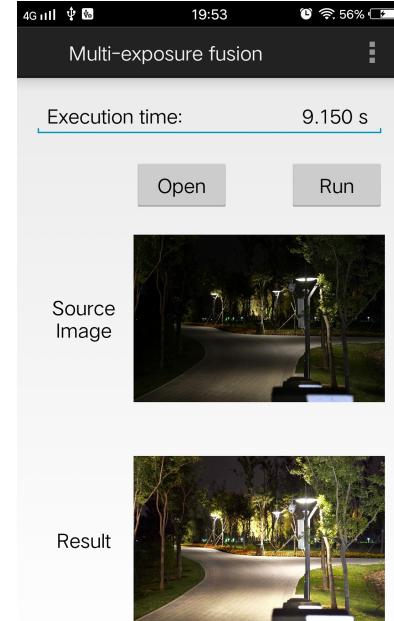


Fig. 9: The user interface of our developed application on mobile device.

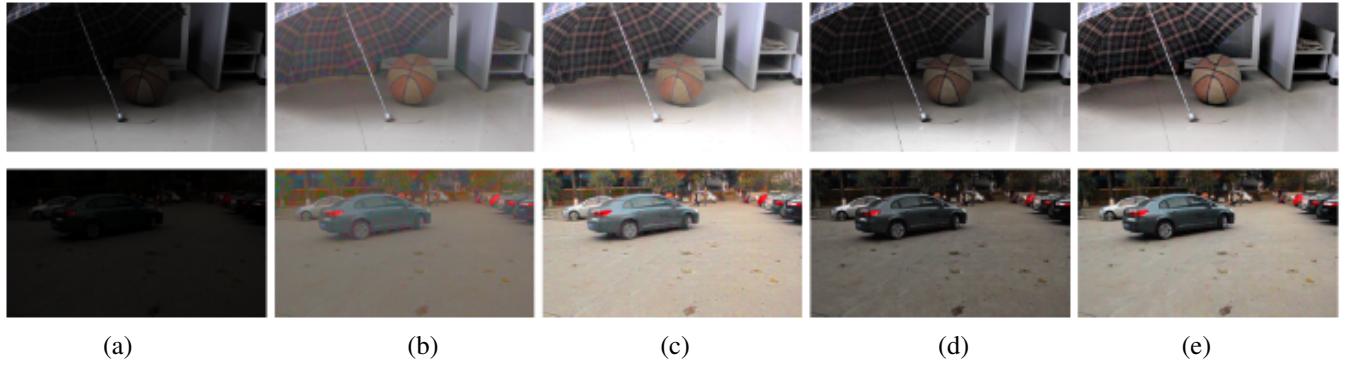


Fig. 10: The result comparison among different tone mapping methods. (a) the original input underexposed video frame, (b) the result of Gamma curves based method, (c) the result of Bennett and Mcmillan's method [26], (d) the result of Yuan and Sun's method [6], and (e) the result of our method.

TABLE V
CONSUMING TIME OF DIFFERENT METHODS FOR PROCESSING DIFFERENT IMAGES

| Images | Mertens et al.'s method [7] | Park et al.'s method [36] | Our method |
|-----------|-----------------------------|---------------------------|------------|
| Fig. 6(a) | 5.3214s | 3.3720s | 6.9746s |
| Fig. 6(b) | 6.6746s | 4.9742s | 6.9461s |
| Fig. 6(c) | 6.3197s | 7.9256s | 8.3455s |
| Fig. 7(a) | 3.3642s | 3.0365s | 5.3267s |
| Fig. 7(b) | 4.7642s | 4.6412s | 5.0944s |
| Fig. 7(c) | 5.3115s | 5.9664s | 6.3464s |

TABLE VI
CONSUMING TIME OF DIFFERENT METHODS FOR PROCESSING DIFFERENT IMAGES ON DIFFERENT DEVICES

| Image Size | method | Device A | Device B | Device C |
|------------|--------------------|----------|----------|----------|
| 230x254 | our method | 0.915s | 0.898s | 0.534s |
| | Mertens et al. [7] | 0.798s | 0.678s | 0.463s |
| | Park et al. [36] | 0.563s | 0.487s | 0.332s |
| 604x343 | our method | 1.450s | 1.432s | 0.895s |
| | Mertens et al. [7] | 1.392s | 1.363s | 0.873s |
| | Park et al. [36] | 1.378s | 1.358s | 0.850s |
| 886x470 | our method | 2.412s | 2.423s | 1.804s |
| | Mertens et al. [7] | 2.122s | 2.016s | 1.515s |
| | Park et al. [36] | 1.912s | 1.787s | 1.278s |
| 1280x720 | our method | 4.095s | 3.938s | 3.472s |
| | Mertens et al. [7] | 3.425s | 3.232s | 2.836s |
| | Park et al. [36] | 2.908s | 2.656s | 2.132s |
| 1920x1280 | our method | 9.982s | 9.150s | 8.991s |
| | Mertens et al. [7] | 7.897s | 7.518s | 7.135s |
| | Park et al. [36] | 7.082s | 6.734s | 6.173s |

First, we compared execution time of our approach with other methods. This part of experiment was performed on a PC equipped with 4 GB RAM, and a graphics card with 2 GB memory. Table V shows the execution times of processing images in Fig. 6 in the manuscript. As can be seen from this table, the computational time of enhancing a general image using our weighted multi-exposure fusion is similar to that of the state-of-the-art efficient methods including Mertens et al.'s method [7] and Park et al.'s method [36]. Mertens et al.'s method [7] is the basic image enhancement method which is based on image fusion, and Park et al.'s method [36] is the state-of-the-art image enhancement method which can be

applied in Consumer Electronics due to its high efficiency. The size of most of our input underexposed images in this experiment is 1080x1920, and the execution time for enhancing an image is 6 seconds on average, which is close to the other two methods. On the other hand, our method can keep even exposure and show better visual appearance than the others. Thus, the proposed method is an efficient underexposed image enhancement method which can produce great enhancement quality.

In practice, we tested our method on different mobile devices with different configurations using images with different sizes. We chose three mobile devices to test our approach: device A with 1.8 GHz CPU and 4 GB RAM, device B with 2.1 GHz CPU and 6 GB RAM, and device C with 2.2 GHz CPU and 6 GB RAM. For every device, we tested five images with different sizes. In addition, an application was built to read an image and implemented our approach to generate final results, in which execution timings was calculated to show the efficiency of our approach on mobile devices. Table VI collects the execution time of processing images with different sizes on different devices. As can be seen from this table, the execution time for enhancing an image with size of 1920x1280 is less than 10 seconds on average. Given an image with size of 886x470, the execution time is less than 2.5 seconds. On the other hand, we also compared with Mertens et al.'s method [7] and Park et al.'s method [36] on the same devices and showed the experiment result in Table VI. The result shows that the execution time of our method is close to the other two methods on mobile devices too. This indicates that our approach can be used on mobile device efficiently like the other two methods. Note that our program was not optimized and there is much room for improvement in efficiency. In summary, our approach can be used for most consumer digital cameras to serve as the image processing module on mobile devices, such as smart phones, pads, etc.

D. Expanded to Underexposed Video Enhancement

Our enhancement can also be expanded to underexposed video enhancement. In Fig. 10, we compare our method with different video enhancement methods. To validate the performance of tone mapping carefully evaluate both the

proposed tone mapping operator Equation (1), we first sample a family of tone mapping curves from the gamma curve model. Then, we replace all the tone mapping curves with the sampled gamma curves and finally receive the gamma curves based results shown in Fig. 10 (b). Compared with the gamma curves, our tone mapping curves have better performance in revealing details and noise control. The tone mapping component of [32] leads to clear ghosting artifacts around the basketball. Moreover, their tone mapped frames exhibit visually unpleasant appearance, such as the whitish floor and severely degraded textures, as shown in Fig. 10(c). In contrast, our method produces visually pleasing result without introducing those visual artifacts. In Fig. 10(d), we can see that both our method and [6] successfully avoid inducing distortion appearance and ghosting artifacts. However, our method has better performance in improving the visibility of severely underexposed regions, and tends to produce results with better overall appearance, such as more reasonable brightness and vivid colors.

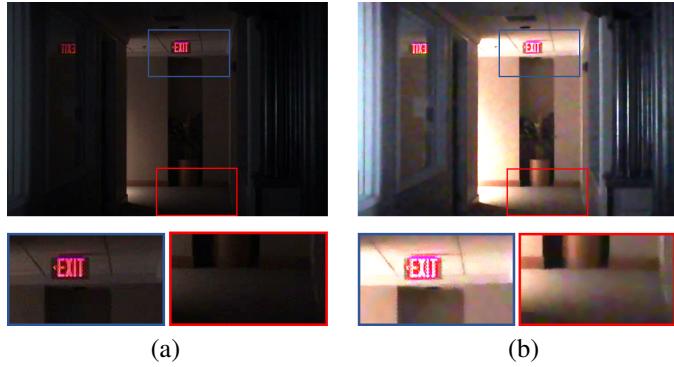


Fig. 11: Limitations of our method. (a) one video frame from the nighttime video. (b) our enhancement result.

E. Limitations

When handling severely underexposed images, our method may fail to produce a satisfactory result. In Fig. 11, for instance, the input image is one frame from a nighttime video. Dark area in the red rectangle actually contains weak color differences. Under our image enhancement framework, these weak color differences will be first treated as details to be amplified, and then appear in the form of severe noise interference. Since the following filtering operation fails to completely remove these noises, our final result still suffers some noticeable noise artifacts, as illustrated in the rectangles. We will further explore this issue in the future work.

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a novel detail-preserving underexposed image enhancement method based on optimal weighted multi-exposure fusion mechanism. Firstly, we constructed multi-exposure image sequence containing different exposed versions of each underexposed image using a series of tone mapping curves. Guided by some visual perception quality measures which can represent the desirable visual

appearance, we adaptively located locally best exposed regions from all multi-exposure image sequences and then seamlessly integrated them into a well-exposed image through some weight values. These weight values could be solved by formulated as an energy function so that the colors of the output image match closely to those of images with high exposure, while retaining details from the images with low exposure. Finally, we fused the image sequences and integrate the best exposed regions into a well-exposed image. Various experiment results validated that our new method can preserve the details in the underexposed images.

However, our method also has some limitations. When handling severely underexposed images, our method may fail to produce a satisfactory result, because the bilateral filter may wash out some weak edges. Furthermore, we will investigate how to integrate the exposure enhancement and the noise reduction into a unified framework. We will also introduce more visual measures to better imitate human vision.

REFERENCES

- [1] A. T. Celebi, R. Duvar and O. Urhan, "Fuzzy fusion based high dynamic range imaging using adaptive histogram separation," *IEEE Trans. Consum. Electron.*, vol. 61, no. 1, pp. 119-127, 2015.
- [2] D. H. Lee, Y. J. Yoon, S. J. Kang and S. J. Ko, "Correction of the overexposed region in digital color image," *IEEE Trans. Consum. Electron.*, vol. 60, no. 2, pp. 173-178, 2014.
- [3] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vision Graph. Image Process.*, vol. 39, no. 3, pp. 355-368, 1987.
- [4] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems*, 1994, pp. 474-485.
- [5] G. Cao, Y. Zhao and R. Ni, "Anti-forensics of contrast enhancement in digital images," in *Proc. ACM Multimedia and Security Workshop.*, 2010, pp. 25-34.
- [6] L. Yuan and J. Sun, "Automatic exposure correction of consumer photographs," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2012, pp. 1-9.
- [7] T. Mertens, J. Kautz and F. Van Reeth, "A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum.*, vol. 28, no. 1, pp. 161-171, 2009.
- [8] S. Li and X. Kang, "Fast multi-exposure image fusion with median filter and recursive filter," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 626-632, 2012.
- [9] S. Raman and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proc. Eurographics*, 2009, pp. 369-378.
- [10] J. Im, S. Lee, and J. Paik, "Improved elastic registration for removing ghost artifacts in high dynamic imaging," *IEEE Trans. Consumer Electronics.*, vol. 57, no. 2, pp. 932-935, 2011.
- [11] R. Bilcu, A. Burian, A. Kuutila, and M. Vehvilainen, "High dynamic range imaging on mobile devices," in *Proc. IEEE Conf. Electronics, Circuits and Systems.*, 2008, pp. 1312-1315.
- [12] J. Im, J. Jeon, M. H. Hayes and J. Paik, "Single image-based ghost-free high dynamic range imaging using local histogram stretching and spatially-adaptive denoising," *IEEE Trans. Consum. Electron.*, vol. 57, no. 4, pp. 1478-1484, 2011.
- [13] S. Raman and S. Chaudhuri, "A matte-less, variational approach to automatic scene compositing," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, 2010, pp. 574-579.
- [14] K. Kotwal and S. Chaudhuri, "An optimization-based approach to fusion of multiexposure, low dynamic range images," in *Proc. Int. Conf. Inf. Fus. (FUSION)*, 2008, pp. 1942-1948.
- [15] M. Bertalmio and S. Levine, "Variational approach for the fusion of exposure bracketed pairs," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 712-723, 2013.
- [16] E. J. Chae, E. S. Lee, W. S. Kang, H. J. Cheong and J. Paik, "Spatially adaptive antialiasing for enhancement of mobile imaging system using combined wavelet-fourier transform," *IEEE Trans. Consum. Electron.*, vol. 59, no. 1, 862-868, 2013.

- [17] X. Kang and S. Li, "Fast multi-exposure image fusion with median filter and recursive filter," *IEEE Trans. Consum. Electron.*, vol. 58, no. 2, pp. 626-632, 2012.
- [18] S. Daly, R. Mantiuk and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1-10, 2008.
- [19] M. S. Brown, Q. Shan and J. Y. Jia, "Globally optimized linear windowed tone mapping," *IEEE Trans. Vis. Comput. Graph.*, vol. 16, no. 4, pp. 663-675, 2010.
- [20] D. Lischinski, R. Fattal and M. Werman, "Gradient domain high dynamic range compression," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 249-256, 2002.
- [21] R. C. Gonzalez and R. E. Woods, *Digital image processing*. 2nd ed., China: Electronic Industry Press, 2007, pp. 133-149.
- [22] K. Devlin and E. Reinhard, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. Vis. Comput. Graph.*, vol. 11, no. 1, pp. 13-24, 2005.
- [23] T. Annen F. Drago, K. Myszkowski and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum.*, vol. 22, no. 3, pp. 26-34, 2003.
- [24] Y. Li, L. Sharan and E. H, "Compressing and companding high dynamic range images with subband architectures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 836-844, 2005.
- [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] E. P. Bennett and L. Mcmillan, "Video enhancement using per-pixel virtual exposures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 845-852, 2005.
- [27] W. Li, X. Dong and L. Yuan, "Temporally consistent region-based video exposure correction," in *Proc. IEEE Int. Conf. Multimedia Expo. (ICME)*, 2015, pp. 1-6.
- [28] L. Zhang, C. X. Xiao, Q. Zhang and Y. Nie, "Underexposed video enhancement via perception-driven progressive fusion," *IEEE Trans. Vis. Comput. Graph.*, vol. 22, no. 6, pp. 1773-1785, 2016.
- [29] W. Q. Gu, X. Y. Fang, J. G. Liu and Y. W. Tang, "A method to improve the image enhancement result based on image fusion," in *Proc. Int. Conf. Mul. Technol. (ICMT)*, 2011, pp. 55-58.
- [30] C. H. Hsieh, B. C. Chen, C. M. Lin and Q. Zhao, "Detail aware contrast enhancement with linear image fusion," in *Proc. Int. Symp. on Aware Computing*, 2010, pp. 1-5.
- [31] L. Pei, Y. Zhao and H. Luo, "Application of wavelet-based image fusion in image enhancement," in *Proc. Int. Congress on Image and Signal Processing (CISP)*, 2010, pp. 649-653.
- [32] X. Guo, "LIME: a method for low-light image enhancement," in *Proc. ACM Conf. Mul. (ACM MM)*, 2016, pp. 87-91.
- [33] S. Wang, J. Zheng and H. M. Hu, "Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3538-3548, 2013.
- [34] Z. Wang and A. C. Bovik, "A universal image quality index," *IEEE Signal Proc. Let.*, vol. 9, no. 3, pp. 81-84, 2002.
- [35] C. S Xydeas and V. Petrovic, "Objective image fusion performance measure," *Military Technical Courier*, vol. 36, no. 4, pp. 308-309, 2000.
- [36] S. Park, B. Moon, S. Ko, S. Yu, "Low-light image enhancement using variational optimization-based Retinex model," *IEEE Trans. Consum. Electron.*, vol. 63, no. 2, pp. 70-71, 2017.



Yu Zhang received the B.E. degree and the M.E. degree from School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, P.R. China in 2015 and 2017, respectively. Her research interests include image processing and video enhancement.



Shiguang Liu (M'16) received the Ph.D. degree from the State Key Laboratory of CAD & CG, Zhejiang University, Hangzhou, P.R. China in 2007. He is currently a Professor with School of Computer Science and Technology, College of Intelligence and Computing, Tianjin University, Tianjin 300350, P.R. China. His research interests include image/video editing, computer graphics, visualization, and virtual reality.