

Multi-Scale Exposure Fusion Based on Multi-Visual Feature Measurement and Detail Enhancement Representation

Yong Yang^{ID}, Senior Member, IEEE, Danjie Zhang^{ID}, Weiguo Wan^{ID}, and Shuying Huang^{ID}, Member, IEEE

Abstract—Multiexposure image fusion (MEF) is used to generate a high-quality image from a series of images with different exposure levels. The multiscale-based MEF method achieves better fusion performance than the single-scale-based method because it can better preserve the global contrast information. However, this method still has the problem of the loss of details in the fused results. To solve this problem, a novel multiscale exposure image fusion method based on multivisual feature measurement and detail enhancement representation in the intensity-hue-saturation (IHS) color space is proposed. First, three visual features of the source multiexposure images, namely contrast, saturation, and exposure, are measured and are then adopted to construct the initial weight maps with adaptive weighting coefficients. Second, to optimize the initial weight maps and obtain the middleweight maps, a decision map construction method is proposed by comparing the pixel values of the detail maps of the intensity components, which can enhance the representation of detail information. Third, guided filtering is applied to eliminate the noise in the middleweight maps to obtain the final weight maps, which improves the visual effect of the fused image. Finally, image pyramid decomposition and reconstruction are performed on the source images and the final weight maps to achieve the final fused image. Numerical experimental results indicate that the proposed method outperforms state-of-the-art methods in terms of subjective visual and quantitative evaluations.

Index Terms—Detail enhancement representation, image pyramid, multiexposure image fusion (MEF), visual feature measurement.

Manuscript received February 18, 2022; revised April 24, 2022; accepted May 5, 2022. Date of publication May 23, 2022; date of current version May 31, 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 61862030 and Grant 62072218, in part by the Project of the Education Department of Jiangxi Province under Grant GJJ200541, in part by the Postdoctoral Research Projects of Jiangxi Province under Grant 2020KY44, and in part by the 15th Student Research Project of Jiangxi University of Finance and Economics (JUFE) under Grant 20200614185004404. The Associate Editor coordinating the review process was Dr. Susanna Spinsante. (*Corresponding authors:* Weiguo Wan; Shuying Huang.)

Yong Yang is with the School of Computer Science and Technology, Tiangong University, Tianjin 300387, China, and also with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: greatyangy@126.com).

Danjie Zhang is with the School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: lyq_zhangdj@163.com).

Weiguo Wan is with the School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China (e-mail: wanwgplus@163.com).

Shuying Huang is with the School of Software, Tiangong University, Tianjin 300387, China (e-mail: shuyinghuang2010@126.com).

Digital Object Identifier 10.1109/TIM.2022.3176881

I. INTRODUCTION

THE dynamic range of images captured by current digital cameras is much smaller than that of natural scenes. Hence, information loss often occurs in under- or overexposed areas in daily photography, resulting in poor imaging results [1]. The high dynamic range (HDR) imaging method can effectively solve this problem by merging a series of low-dynamic-range (LDR) images of the same scene with different exposure levels [2]. There are usually two HDR imaging methods: the first aims to reconstruct the HDR content through the camera response function (CRF) and then reproduce the tone in the luminance domain [3]. However, the exposure time is unknown in most cases, and different devices have different CRFs [4], increasing the complexity of the operation process. The second method aims to fuse the multiexposure images with advanced image fusion methods to obtain a single HDR image [5]–[7]. By comparison, as a method that can directly generate high-quality and well-exposed images, multiexposure image fusion (MEF) has been widely used in HDR imaging for mobile devices.

With the development of MEF techniques, a large number of MEF methods have been proposed. These methods can be roughly divided into two categories [8]: traditional MEF methods and deep learning-based MEF methods. Traditional MEF methods obtain the final fusion image by weighted averaging of the input source images. In these methods, the structural and visual information of each input image is utilized to obtain the weight maps, which can provide spatial proportionality for these input images [9]. Therefore, how to obtain suitable weight maps is the core issue of MEF. Traditional methods are usually divided into spatial domain- and transform domain-based methods. In spatial domain-based methods, the pixel-wise MEF method works on pixels, which mainly obtains the fusion weight maps by extracting the local area features of the source images. Liu and Wang [10] proposed an MEF algorithm based on the dense scale-invariant feature transform (SIFT) [11], which constructs the three measurement factors of local contrast, exposure quality, and spatial consistency of each source image. Afterward, Hayat and Imran [12] presented a ghost-free MEF method based on the dense SIFT and guided filter, which uses three features to estimate the initial weights. Raman and Chaudhuri [13] proposed using bilateral filtering to extract image details and then construct the weight factors based on the extracted details for the fusion of the source

images. Lee *et al.* [14] proposed an MEF method based on adaptive weights, which defines two weight functions to reflect the relative pixel intensity between the source images and their global gradients. Li and Kang [15] proposed using local contrast, brightness, and color dissimilarity to estimate the fusion weight maps and used an edge-preserving recursive filter to refine the weight maps to guide the fusion process. The above pixel-wise MEF methods can be easily implemented; however, noise and halo artifacts are usually generated in the fusion results [16].

Unlike pixel-wise MEF methods, patch-wise MEF methods first decompose the source images into patches and then process these patches to obtain the fused image. Ma and Wang [17] first decomposed each overlapped patch into three conceptually relatively independent components, signal strength, signal structure, and mean intensity and then processed the three components to obtain the fusion result. On the basis of [17], Ma *et al.* [2] proposed an MEF algorithm based on structural patch decomposition (SPD), which can better solve the artifact phenomenon in the fused result. In order to reduce the running time, Li *et al.* [18] proposed a fast multiscale SPD method, which generates better fusion results and achieves a faster running time. However, there are still details lost in the fusion image. To solve this problem, the edge retention factor was used by Li *et al.* [1] to preserve more details in fusion results. Compared with pixel-wise MEF methods, the weight maps of the patch-wise MEF method produce less noise. However, patch-wise MEF methods may produce obvious visible artifacts at the target boundaries [8].

Considering that the fusion performance of existing MEF methods is affected by a single scale, many researchers have proposed using two-scale or multiscale fusion methods in the transform domain to improve the fusion effects [8]. Wang *et al.* [6] proposed a MEF algorithm in the Luminance, Chrominance, Chroma (YUV) color space, which can solve the global optimization problem with new parameters and retain the details in the brightest and darkest regions of the source images. Li *et al.* [19] proposed a two-scale image fusion method based on guided filtering [20]. Although the computational efficiency is higher, it produces visible halo artifacts in the fusion result. Mertens *et al.* [5] proposed a multiscale fusion method, which first constructs the weight maps and then constructs the multiscale weight maps and exposure images by the Gaussian pyramid [21] and Laplacian pyramid [22], respectively. However, this method suffers from edge detail degradation owing to the Gaussian filtering operation. To overcome this limitation, some other edge-preserving smoothing filtering methods have been adopted. Li *et al.* [23] proposed a multiscale exposure fusion method, which uses weight-guided filtering (WGF) [24] to smooth the weight maps and uses the detail extraction module to refine the image details. Kou *et al.* [25], [26] proposed to use gradient domain-based guided filtering (GGF) [27] to preserve image edges and alleviate the halo phenomenon in the fusion process. Although the above multiscale exposure fusion methods can better retain the global contrast information than single-scale exposure fusion methods, they still have the problem of losing the details in the process of multiscale decomposition and reconstruction [6].

In recent years, many deep learning-based MEF methods have been proposed. Prabhakar *et al.* [28] first proposed a convolution neural network (CNN)-based MEF method, which does not require ground truth as a reference during the training process. Another unsupervised MEF method based on the generative adversarial network (GAN) was proposed by Yang *et al.* [29] to better retain the contents of the exposure image pairs in the fused image. Li and Zhang [30] proposed a CNN-based MEF method that can be applied to both static and dynamic scenes by extracting features of multiple source images. Xu *et al.* [31] proposed a GAN-based MEF method, which utilizes the self-attention mechanism to achieve attention-driven and long-range dependency. Both these methods need reference images in the training process; so, a large number of data samples and a lot of time are needed for training. In general, compared with traditional methods, deep learning-based methods have been more favored by researchers in recent years. However, deep learning-based methods require extensive samples and take a long time for network training.

Based on the above analysis, this article proposes a new multiscale exposure fusion method based on multivisual feature measurement and detail enhancement representation in the intensity-hue-saturation (IHS) color space instead of the RGB color space. To better construct the initial weight maps, the three visual features of contrast, saturation, and exposure are extracted, and among these, the exposure feature is obtained by a Gaussian model with adaptive parameters. To solve the issue of the loss of detail information in the fusion process, a decision map construction method is proposed to enhance the representation of detail information. The experimental results show that the proposed method achieves a better fusion effect in both subjective and objective results. The main contributions of the proposed MEF method are as follows.

- 1) A novel multiscale exposure fusion framework is proposed based on multivisual feature measurement and detail enhancement representation, which can achieve better fusion results with more precise details and natural color.
- 2) An initial weight map calculation method is designed based on the measured visual features and their adaptive weighting coefficients, which can retain more brightness and color information from the source images.
- 3) To enhance the detail information representation, a decision map construction method is presented by comparing the pixel values of the detail maps of the intensity components.
- 4) The results of extensive experiments prove that our method can improve the visual effect of the fused images and can achieve higher evaluation indexes compared with state-of-the-art fusion approaches.

The remainder of this article is organized as follows. In Section II, the related works of guided filter and pyramid image decomposition-based MEF are briefly introduced. Section III introduces the proposed MEF method in detail. In Section IV, extensive experiments and the corresponding analysis are presented. The conclusion is given in Section V.

II. RELATED WORK

A. Guided Filter

He *et al.* [20] proposed the guided filter, which uses the local linear transformation of the guided image to realize the filtering of the input image. The guided image can be served by the input image or by a different image. The guided filter has better edge-retention performance and can effectively solve the gradient inversion problem. It is also known as one of the fastest edge-retention filters. At present, the guided filter has been successfully applied to the fields of computer vision and image processing. The local linear relationship between the guided and the filtered output Q in the window ω_p centered at a pixel p is defined as follows [20]:

$$Q_i = a_p L_i + b_p \quad \forall i \in \omega_p \quad (1)$$

where a_p and b_p are constants in ω_p of size $(2r+1) \times (2r+1)$, and r is the radius of the guided filter, which can be calculated by minimizing the squared difference between the input image X_i and the output image Q_i as follows:

$$E(a_p, b_p) = \sum_{i \in \omega_p} ((a_p L_i + b_p - X_i) + \varepsilon a_p^2) \quad (2)$$

where ε is a regularization term that determines the blur degree of the smoothing. Then, a_p and b_p can be directly obtained by linear regression [32]

$$a_p = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_p} (L_i X_i - \mu_p \bar{X}_p)}{\sigma_p^2 + \varepsilon} \quad (3)$$

$$b_p = \bar{X}_p - a_p \mu_p \quad (4)$$

where μ_p and σ_p^2 are the mean and variance of L in ω_p , respectively; $|\omega|$ is the number of pixels in ω_p ; and \bar{X}_p is the mean of X_i in ω_p . After obtaining a_p and b_p , the filtered output Q_i can be calculated as follows:

$$Q_i = \bar{a}_i L_i + \bar{b}_i \quad (5)$$

where $\bar{a}_i = (1/|\omega|) \sum_{i \in \omega_p} a_p$ and $\bar{b}_i = (1/|\omega|) \sum_{i \in \omega_p} b_p$ are the average of a_p and b_p in ω_p , respectively. For more information on a guided filter, see [20]. In this article, the guided filter is defined as follows:

$$Q = GF(X, L) \quad (6)$$

where $GF(\cdot)$ denotes the guided filtering operator, X represents the input image, L is the guided image, and Q is the output filtered image.

B. Pyramid Image Decomposition-Based MEF

A Gaussian pyramid [21] is composed of the decomposed images obtained by performing Gaussian filtering and down-sampling on the source image. Specifically, the source image is taken as the basis, and then the convolution operation is performed on it. The upper layer of the Gaussian pyramid is obtained by down-sampling the image after Gaussian filtering. Finally, the upper layer is taken as the input, and the filtering and down-sampling operations are repeatedly conducted to obtain the next-highest layer, and the Gaussian pyramid is

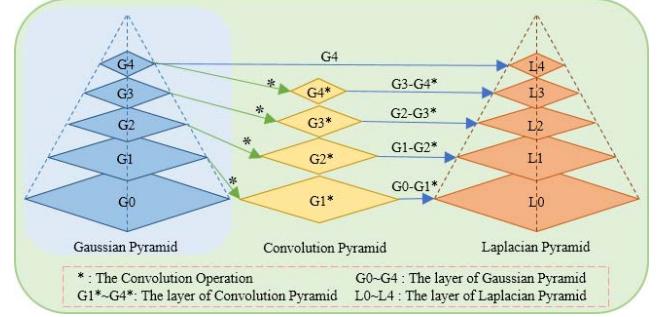


Fig. 1. Construction process of Gaussian and Laplacian Pyramid.

obtained by iterations. The Laplacian pyramid [22] is constructed on the basis of the Gaussian pyramid, which can make up for the high-frequency details lost in the Gaussian pyramid operation. After obtaining the Gaussian pyramid, the upper layer after filtering is subtracted from the bottom layer of the Gaussian pyramid to obtain the bottom layer of the Laplacian pyramid. By iterating this process, the Laplacian pyramid is obtained. The construction processes of the Gaussian pyramid and Laplacian pyramid are shown in Fig. 1. The source image can be reconstructed by merging each layer of the Laplacian pyramid from the top to the bottom.

Pyramid image decomposition has been utilized in the MEF field [5], [6]. It can avoid the problems of traditional MEF methods, such as undesirable halos around edges, obvious splicing traces, and poor color information [22]. The pyramid image decomposition-based MEF consists of the following steps: 1) the weight maps are calculated by measuring the visual features of the source images; 2) the source images are decomposed into Laplacian pyramids, and the weight maps are decomposed into Gaussian pyramids; 3) the fused pyramid map can be obtained by combining the Laplacian and Gaussian pyramids; and 4) the fusion image is reconstructed by the inverse pyramid operation on the fused pyramid map.

C. Image Quality Evaluation Indexes for MEF

The MEF-structural similarity (SSIM) index (MEF-SSIM) [33] is based on the image patch consistency measure and the multiscale SSIM [34] framework. It describes an image patch x from three conceptually different components, namely luminance, contrast, and structure. The luminance \hat{l} of the fused image patch \hat{x} is determined by the mean value of the patch x . The contrast \hat{c} of the fused image patch is determined by the highest contrast of all source image patches. And the structure \hat{s} of the fused image patch is computed by a weighted average of the input structure vectors. The fused image patch \hat{x} can be obtained by combining \hat{l} , \hat{c} , and \hat{s}

$$\hat{x} = \hat{c} \cdot \hat{s} + \hat{l}. \quad (7)$$

Hence, the overall structural quality measure of the fused image y can be calculated as follows:

$$S_q^F = \frac{1}{Q} \sum_{q=1}^Q \left(\frac{2\sigma_{\hat{x}y} + C}{\sigma_{\hat{x}}^2 + \sigma_y^2 + C} \right) \quad (8)$$

where q is the spatial patch index and Q is the total number of image patches. $\sigma_{\hat{x}}^2$ and σ_y^2 denote the local variances of

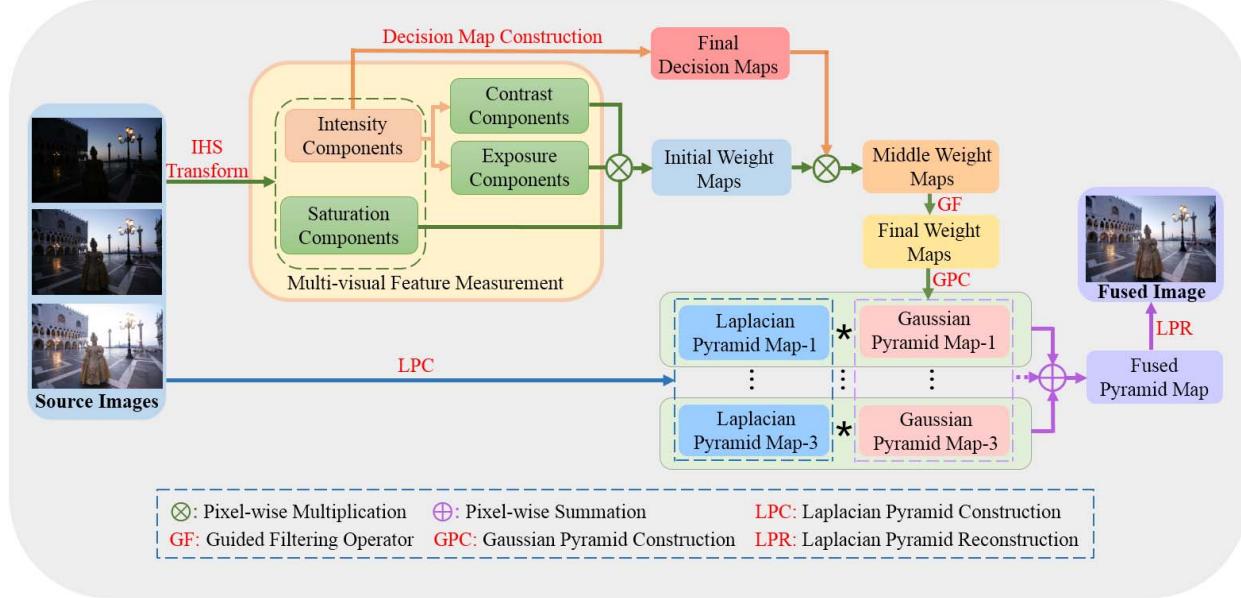


Fig. 2. Flowchart of the proposed MEF method.

\hat{x} and y , respectively, and $\sigma_{\hat{x}y}$ denotes the local covariance between \hat{x} and y . C is a small positive constant for stabilizing. The MEF-SSIM is defined as follows:

$$S^F = \prod_{l=1}^L [S_q^F]^{\omega_l} \quad (9)$$

where L is the number of scales and ω_l is the weight for l th scale.

The visual information fidelity for fusion (VIFF) is a multiresolution image fusion metric that can measure the visual information fidelity (VIF) between source images and fused image. The larger the VIFF value, the better the fusion effect. The calculation process of VIFF is as follows. First, the source images and fused image are filtered and divided into blocks. Then, the visual information is extracted by using a VIF model from each pair of the source and fused image blocks. Next, the VIFF of each subband is calculated based on the extracted effective visual information of each block. Finally, the overall VIFF value is obtained by weighting the VIFF of each subband. For detail information of VIFF, refer to [35].

III. PROPOSED METHOD

In this section, the proposed method is introduced in detail. First, the input source images are converted from the RGB color space to the IHS color space; then, three visual features of the source images, namely contrast, saturation, and exposure, are extracted. Next, the initial weight maps are constructed through these visual features and their weighting coefficients, which are adaptively determined by the mean luminance of the intensity components. Then, the decision map construction method is proposed to enhance the representation of detail information to obtain the middleweight maps. In addition, the guided filter is adopted to refine the middleweight maps to obtain the final weight maps with less noise. Finally, the pyramid image decomposition and

reconstruction are conducted to achieve the fusion result. The flowchart of the proposed method is illustrated in Fig. 2.

A. Multivisual Feature Measurement

Assume that the input multiexposure image is M_k , $k = 1, \dots, N$, where N is the total number of input images. It is found that the color description in the IHS space is more intuitive [36], as the color components are independent of each other. Hence, we convert the source images from the RGB to the IHS color space. Because contrast, saturation, and exposure are three important visual features for measuring useful information of input images, we propose extracting these features to construct the initial weight maps.

1) *Contrast*: The contrast can well reflect the detail features of each exposed image [37]; so, we introduce the contrast for detail measurement. The contrast is calculated by the convolution operation on the intensity component I_k with the Laplacian operator $LP = [0 \ 1 \ 0, 1 \ -4 \ 1, 0 \ 1 \ 0]$. I_k is the intensity component of the k th input image, which can be obtained as follows:

$$I_k(x, y) = \frac{1}{3}[R_k(x, y) + G_k(x, y) + B_k(x, y)] \quad (10)$$

where R_k , G_k , and B_k represent the three channels of the k th input image.

Therefore, the contrast for each pixel location (x, y) in the intensity component of k th source image is obtained as follows:

$$C_k(x, y) = |LP * I_k(x, y)| \quad (11)$$

where $*$ means the convolution operation, and $C_k(x, y)$ is the value at the pixel position (x, y) in the k th contrast component. The contrast can better represent the edge details of the source image, as shown in Fig. 3(a).

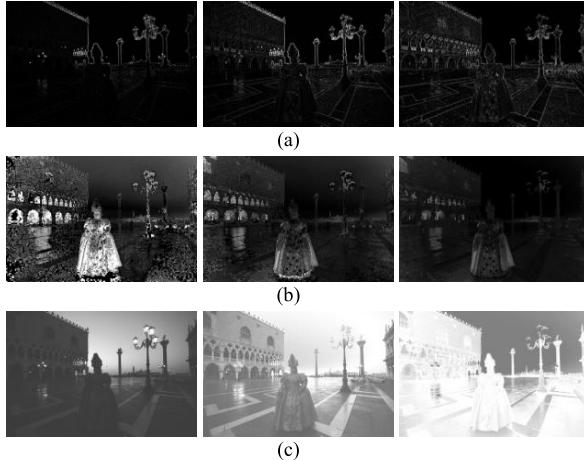


Fig. 3. Visualization results of multivisual features measurement. (a) Contrast components. (b) Saturation components. (c) Exposure components.

2) *Saturation*: It is difficult to obtain useful information from an overly dark or overly bright area in under- or over-exposed images. However, saturation is an important feature that highlights image color information, which is helpful to measure the image quality in the process of constructing the initial weight map [5]. The visualization results of saturation are shown in Fig. 3(b). The saturation can be obtained as follows:

$$S_k(x, y) = 1 - \frac{[\min(R_k(x, y), G_k(x, y), B_k(x, y))]}{[R_k(x, y) + G_k(x, y) + B_k(x, y)]/3} \quad (12)$$

where $S_k(x, y)$ is the saturation component of the k th input image.

3) *Exposure*: The exposure reflects the brightness characteristic of each exposed image. The key information of the image can be obtained in the well-exposed area of the image sequence with different exposure levels. In the existing methods, the exposure feature is obtained by assigning reasonable weights to all pixels through the Gaussian model [5]

$$E_k(x, y) = \exp\left\{-\frac{[I_k(x, y) - \mu]^2}{2\sigma^2}\right\} \quad (13)$$

where σ and μ are two parameters of the Gaussian kernel, which are set to 0.2 and 0.5, respectively. In this case, the image details are mainly concentrated around the normalized pixel value of 0.5 so that the details cannot be fully preserved in the brighter areas of the low-exposed image or the darker areas of the over-exposed image. To preserve more details in darker areas of over-exposed images and brighter areas of under-exposed images, greater weights should be assigned to these areas. Based on the above analysis, an adaptive μ calculated from the mean luminance of the intensity component is designed to solve the mentioned problem. The optimal μ value of each input image is adaptively determined according to different average brightness values. The calculation method is defined as follows:

$$\mu'_k = 1 - \bar{I}_k \quad (14)$$

where \bar{I}_k is the mean luminance of I_k . Combining (13) and (14), the new exposure calculation formula is as

follows:

$$E_k(x, y) = \exp\left\{-\frac{[I_k(x, y) - \mu'_k]^2}{2\sigma^2}\right\} \quad (15)$$

where the value of σ is set to 0.5 through experiments in this article. The visualization of the exposure features is shown in Fig. 3(c), which reflects the brightness information of the source images.

B. Construction of the Initial Weight Maps

By combining the above three visual features, the initial weight maps with complementary information can be obtained, which highlights the significant information of the source images. In this work, the initial weight maps are calculated by linearly multiplying these visual features and their weighting coefficients, which are adaptively obtained. With the extraction of the contrast $C_k(x, y)$, saturation $S_k(x, y)$, and exposure $E_k(x, y)$ of each exposed image, the initial weight map W_k is obtained as follows:

$$W_k(x, y) = [C_k(x, y)]^{\omega_C} \cdot [S_k(x, y)]^{\omega_S} \cdot [E_k(x, y)]^{\omega_E} \quad (16)$$

where ω_C , ω_S , and ω_E are the weighting coefficients of the contrast, saturation, and exposure feature measurement factors, respectively.

In the existing methods, these weighting coefficients are usually set to the default value of 1, which will cause the loss of color information during the fusion process and the darkening of the fused image. Therefore, different weighting coefficients should be set for different exposure image sequences to obtain better fused images. Higher or lower contrast will result in the loss of edge information, and the final fused image will be brighter or darker, resulting in a poor visual effect. Considering the relationship between contrast and brightness, the weighting coefficient of the contrast component is defined as follows:

$$\omega_C = 2^{\text{mod}((\lfloor I_k \rfloor + 1), 2)} \quad (17)$$

where $\text{mod}(\cdot)$ represents the mod function and $\lfloor \cdot \rfloor$ denotes a ceiling function. With this weighting coefficient, more edge information will be preserved in the contrast map.

In addition, a higher or lower exposure will also result in brighter or darker fusion results, which cannot well preserve the brightness and color information of the source images, thereby reducing the quality of the fusion images. To address these problems, exposure and saturation maps are justified to balance brightness and color information with the weighting coefficients, which can be, respectively, defined as follows:

$$\omega_E = 2^{\text{mod}(\bar{I}_k, 2)}/0.5 \quad (18)$$

$$\omega_S = 2^{\text{mod}(\bar{I}_k, 2)}/10. \quad (19)$$

C. Decision Map and Guided Filtering for Weight Map Optimization

1) *Decision Map Construction*: The generated initial weight maps may contain noises; thus, they should be refined before

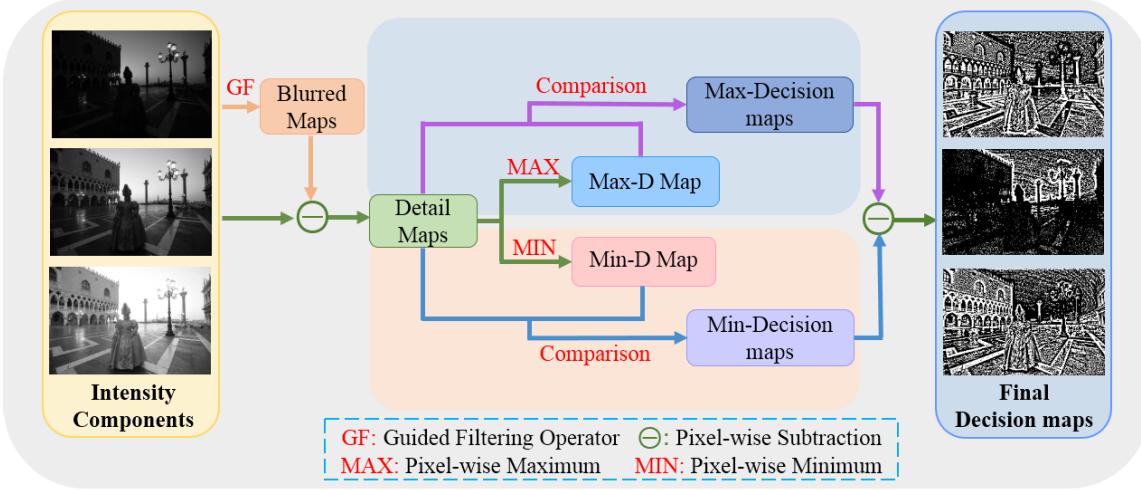


Fig. 4. Flowchart of the proposed decision map construction.

participating in the fusion [12]. To solve this problem, a decision map construction method is proposed to refine the initial weight maps to improve the details of the fusion images and decrease the noise, and the construction process is shown in Fig. 4. Assume that an input intensity component is I_k [see Fig. 5(a)]. To enhance the detail representation, the detail maps of the source images are first extracted through the following two steps: 1) guided filtering is conducted to filter I_k with itself serving as the guided image, and the corresponding blurred maps are obtained, as shown in Fig. 5(b); and 2) the detail map M_k^d is obtained by subtracting the blurred map from the intensity component I_k , as shown in Fig. 5(c), which can be defined as follows:

$$M_k^d = I_k - GF(I_k, I_k), \quad k = 1, \dots, N. \quad (20)$$

Then, according to the rules of the pixel-wise maximum and minimum, we can obtain the maximum detail map and minimum detail map by taking these rules on N detail maps. Next, each value of all the detail maps and the value of the maximum detail map at the same position are compared to obtain the max-decision maps. Assuming that the value of the k th detail map at position (x, y) is the largest, the value of the k th max-decision map at the same position is set as 1, and the values of other max-decision maps at the same position are set as 0. Thus, the formula for the max-decision map M_k^{ad} is expressed as follows:

$$M_k^{ad}(x, y) = \begin{cases} 1, & \text{if } M_k^d(x, y) = \max(M_1^d(x, y), \dots, M_N^d(x, y)) \\ 0, & \text{otherwise} \end{cases} \quad (21)$$

where $M_k^d(x, y)$ is the value of the detail map at position (x, y) in the k th detail map. This way, the details contained in the max-decision maps are the most abundant, as shown in Fig. 5(d). Similarly, each pixel value of all detail maps and the minimum detail map at the same location are compared to

generate the min-decision map M_k^{id} [see Fig. 5(e)] as follows:

$$M_k^{id}(x, y) = \begin{cases} 1, & \text{if } M_k^d(x, y) = \min(M_1^d(x, y), \dots, M_N^d(x, y)) \\ 0, & \text{otherwise.} \end{cases} \quad (22)$$

The final decision map M_k^{ad} is obtained by subtracting the min-decision map from the max-decision map, which can be defined as follows:

$$M_k^{fd}(x, y) = M_k^{ad}(x, y) - M_k^{id}(x, y) \quad (23)$$

where $M_k^{fd}(x, y)$ is the k th final decision map. Compared with M_k^{ad} , the final decision maps can enhance the detail representation of the source images. The proposed decision map construction process is presented in Algorithm 1.

With the obtained final decision maps, the initial weight maps can be refined to generate the middleweight maps by pixel-wise multiplication of the initial weight maps and the final decision maps, which is defined as follows:

$$W_k^d = W_k(x, y) \cdot |M_k^{fd}(x, y)| \quad (24)$$

where $|\cdot|$ denotes the absolute value operation. The generated middleweight maps highlight the detail information of the source images, which helps in the later fusion process of the multiexposure images.

2) *Final Weight Map Construction*: The middleweight maps can be used to generate the fused results with more details, but noise and artifacts will be introduced into the fusion result at the same time. To tackle these problems, guided filtering is adopted to achieve the final weight maps, which is constructed as follows:

$$W_k^f = GF(W_k^d, W_k^d) \quad (25)$$

where W_k^f is the k th final weight map. To ensure that the sum of the weights in the final weight maps is 1, the weight maps are normalized as follows:

$$\overline{W}_k^f = \left[\sum_{k=1}^N W_k^f \right]^{-1} W_k^f \quad (26)$$

where \overline{W}_k^f is the k th normalized weight map of the source images.

D. Image Pyramid Decomposition and Reconstruction

If weight fusion is directly performed on the input image sequences and weight maps, the fused result will have the problems of edge information loss and an inconsistent target connection [38]. The multiresolution blending algorithm with pyramidal image decomposition can avoid the above problems [37]. Therefore, we refer to the Gaussian pyramid to decompose the final weight maps, and the Laplacian pyramid to decompose the source images. The Gaussian pyramid contains a series of low-pass filters, and the specific construction process is as follows:

$$G^{l+1} = \text{Down}(G^l \otimes G_{5 \times 5}) \quad (27)$$

where $\text{Down}(\cdot)$ denotes down-sampling, \otimes denotes the filtering operator, l is the l th layer of the pyramid, and $G_{5 \times 5}$ represents the Gaussian kernel, which is used to smooth the bottom layers of the pyramid.

In the process of constructing the Gaussian pyramid, the processed images need to be convolved and down-sampled operations, which will result in the loss of the high-frequency information in the image. The Laplacian pyramid can overcome this shortcoming and retain the high-frequency information in the image. The Laplacian pyramid is defined as follows:

$$L^l = G^l - \text{Up}(G^{l+1}) \otimes G_{5 \times 5} \quad (28)$$

where $\text{Up}(\cdot)$ denotes up-sampling. The total number of decomposition layers is T , which is defined as follows:

$$T = \lfloor \log_2(\min(W, H)) \rfloor \quad (29)$$

where W and H represent the width and height of the source image, respectively. Then, the layers in different pyramid levels are fused by weight blending of the Laplacian pyramid of the source image and the Gaussian pyramid of the normalized final weight map. The process is performed as follows:

$$L\{F\}^l = \sum_{k=1}^N \left[G\{\overline{W}^f\}_k^l \times L\{M\}_k^l \right] \quad (30)$$

where $L\{F\}^l$ is the l th layer of the fused pyramid. Finally, the fused image F is obtained by reconstructing the fused pyramid $L\{F\}^l$. The procedure of our proposed fusion method is given in Algorithm 2.

IV. EXPERIMENTAL RESULTS

To validate the effectiveness of the proposed MEF method, all the experiments are conducted on 30 LDR sequences with different exposure levels from the datasets of [39]–[41]. Two commonly used metrics, the MEF SSIM index (MEF-SSIM) and the VIFF, are adopted to evaluate the quality of fused images. The MEF-SSIM is a quantitative evaluation metric that measures the SSIM between the source exposure images and the fused image. The VIFF reflects the degree to which the visual information of the source images is preserved in the

TABLE I
PERFORMANCE COMPARISON OF FIXED AND ADAPTIVE PARAMETER μ

Parameter Setting	MEF-SSIM	VIFF
$\mu = 0.5$	0.9796	0.8509
Adaptive μ	0.9799	0.8642

fused image. The higher the MEF-SSIM and VIFF values, the better the performance of the fusion method.

In our experiments, we first conducted a series of ablation studies to verify the effectiveness of the proposed method. Then, the qualitative and quantitative results of our method against state-of-the-art MEF methods are proposed. Finally, the running time of the different methods is analyzed. In all the objective evaluation results, the best results are marked in bold.

A. Ablation Studies

In this section, we first verify the performance of the two parameters μ and σ of the Gaussian function in the exposure measurement of the proposed method. Then, we analyze the effect of weighting coefficients, including ω_C , ω_S , and ω_E on the quality of the fused image. Next, we verify the effectiveness of different decision maps for the fusion results. Finally, we discuss the parameters in the guided filtering.

1) *Effect of Parameters μ and σ :* First, we fix the parameter σ as 0.2 to compare the effectiveness of the fixed parameter and the proposed adaptive parameter μ . The results of the comparison experiments are presented in Table I, which shows that our proposed adaptive μ achieves better values on both MEF-SSIM and VIFF than the fixed μ .

In addition, to verify the effectiveness of parameter σ , we perform six groups of experiments on different σ values, and the results are presented in Table II. It can be found from the table that when parameter $\sigma = 0.5$, the MEF-SSIM is the optimal value, but the VIFF is not optimal; and when parameter $\sigma = 0.3$, the VIFF obtains the optimal value, but the MEF-SSIM is not optimal. As MEF-SSIM is the main evaluation metric in the MEF field, we finally set $\sigma = 0.5$ as the default value.

2) *Effect of Weighting Coefficients:* To evaluate the effectiveness of the fixed weighting coefficients and the proposed adaptive weighting coefficients for the construction of the initial weight maps, a comparison is conducted. The fixed weighting coefficients are set as the same as in [5] and [6], namely $\omega_C = \omega_S = \omega_E = 1$. Fig. 6 shows the examples of fusion results of the fixed and adaptive weighting coefficients intuitively. The fusion results of the adaptive weighting coefficients show a better fusion performance than that of the fixed weighting coefficients. For example, the results of fixed weighting coefficients [Fig. 6(a)] look dark as a whole (see the sky areas and red boxes), and the visual effect is poor. Moreover, the objective evaluation is performed on the 30 LDR sequences to compare the effectiveness of the fixed and adaptive weighting coefficients, and the results are presented in Table III. From the table, it can be seen that the adaptive weighting coefficients can achieve higher MEF-SSIM

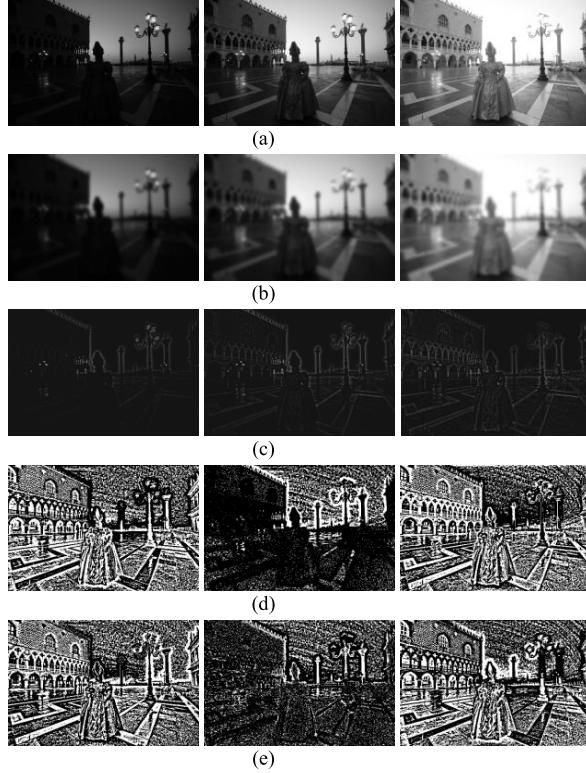


Fig. 5. Visualization results of decision map construction. (a) Intensity components. (b) Blurred maps. (c) Detail maps. (d) Max-decision maps. (e) Min-decision maps.

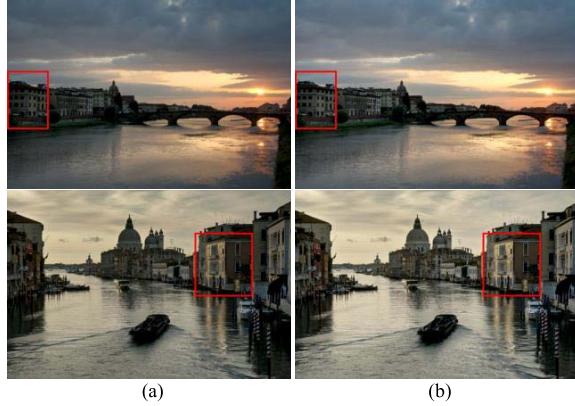


Fig. 6. Fused results of proposed method by different weighting coefficients. (a) Results of fixed weighting coefficients. (b) Results of adaptive weighting coefficients.

TABLE II
PERFORMANCE COMPARISON OF PARAMETER σ

Parameter Setting	MEF-SSIM	VIFF
$\sigma = 0.1$	0.9779	0.8579
$\sigma = 0.2$	0.9799	0.8642
$\sigma = 0.3$	0.9809	0.8660
$\sigma = 0.4$	0.9815	0.8648
$\sigma = 0.5$	0.9818	0.8608
$\sigma = 0.6$	0.9815	0.8550

and VIFF values, which indicate that the proposed adaptive weighting coefficients are superior to the fixed weighting coefficients.

TABLE III
PERFORMANCE COMPARISON OF FIXED WEIGHTING COEFFICIENTS AND ADAPTIVE WEIGHTING COEFFICIENTS

Weighting Coefficients Setting	MEF-SSIM	VIFF
Fixed	0.9775	0.8321
Adaptive	0.9818	0.8608

TABLE IV
OBJECTIVE EVALUATION RESULTS OF DIFFERENT DECISION MAPS

Max-Decision Maps	Min-Decision Maps	Final Decision Maps	MEF-SSIM	VIFF
×	×	×	0.9789	0.8328
√	×	×	0.9809	0.8602
×	√	×	0.9817	0.8585
×	×	√	0.9818	0.8608

TABLE V
PERFORMANCE COMPARISON OF THE REGULARIZATION PARAMETER ε AND THE RADIUS r IN THE GUIDED FILTERING

Parameters	$r=1$	$r=2$	$r=4$	$r=8$
$\varepsilon=1e-4$	0.9571/0.7479	0.9572/0.7542	0.9569/0.7618	0.9565/0.7679
$\varepsilon=1e-2$	0.9774/0.8270	0.9762/0.8227	0.9748/0.8177	0.9728/0.8088
$\varepsilon=1e-1$	0.9809/0.8484	0.9813/0.8527	0.9814/0.8560	0.9810/0.8553
$\varepsilon=1$	0.9812/0.8507	0.9817/0.8562	0.9818 /0.8608	0.9814/ 0.8612

3) *Effect of the Decision Maps:* In this part, different decision maps including max-decision maps, min-decision maps, and the final decision maps for the performance of the fused image are compared. The objective evaluation results are presented in Table IV. From the table, it can be seen that the results without the decision maps obtain the lowest MEF-SSIM and VIFF values, which reflects the effectiveness of the decision maps. In addition, the final decision maps achieve the highest MEF-SSIM and VIFF values, which indicates that the proposed final decision maps are better than the max- and min-decision maps. Thus, the use of final decision maps can further improve the quality of the fusion image.

4) *Effect of Parameters in the Guided Filtering:* In this part, we use a series of experiments to verify the effect of the regularization parameter ε and radius r in the guided filtering on the fusion performance. The average MEF-SSIM and VIFF values under different parameter settings are presented in Table V. From the table, it can be seen that when $\varepsilon = 1$ and $r = 4$, the MEF-SSIM obtains the optimal value, and when $\varepsilon = 1$ and $r = 8$, the VIFF obtains the optimal value. But with the increase of radius r , the fused image will be over-smoothed and cannot preserve more fine details. Therefore, we finally set $\varepsilon = 1$ and $r = 4$ as the default values of the guided filtering in this work.

B. Comparison With Existing MEF Methods

1) *Qualitative Comparison:* In this section, the proposed method is compared with eight state-of-the-art MEF methods,

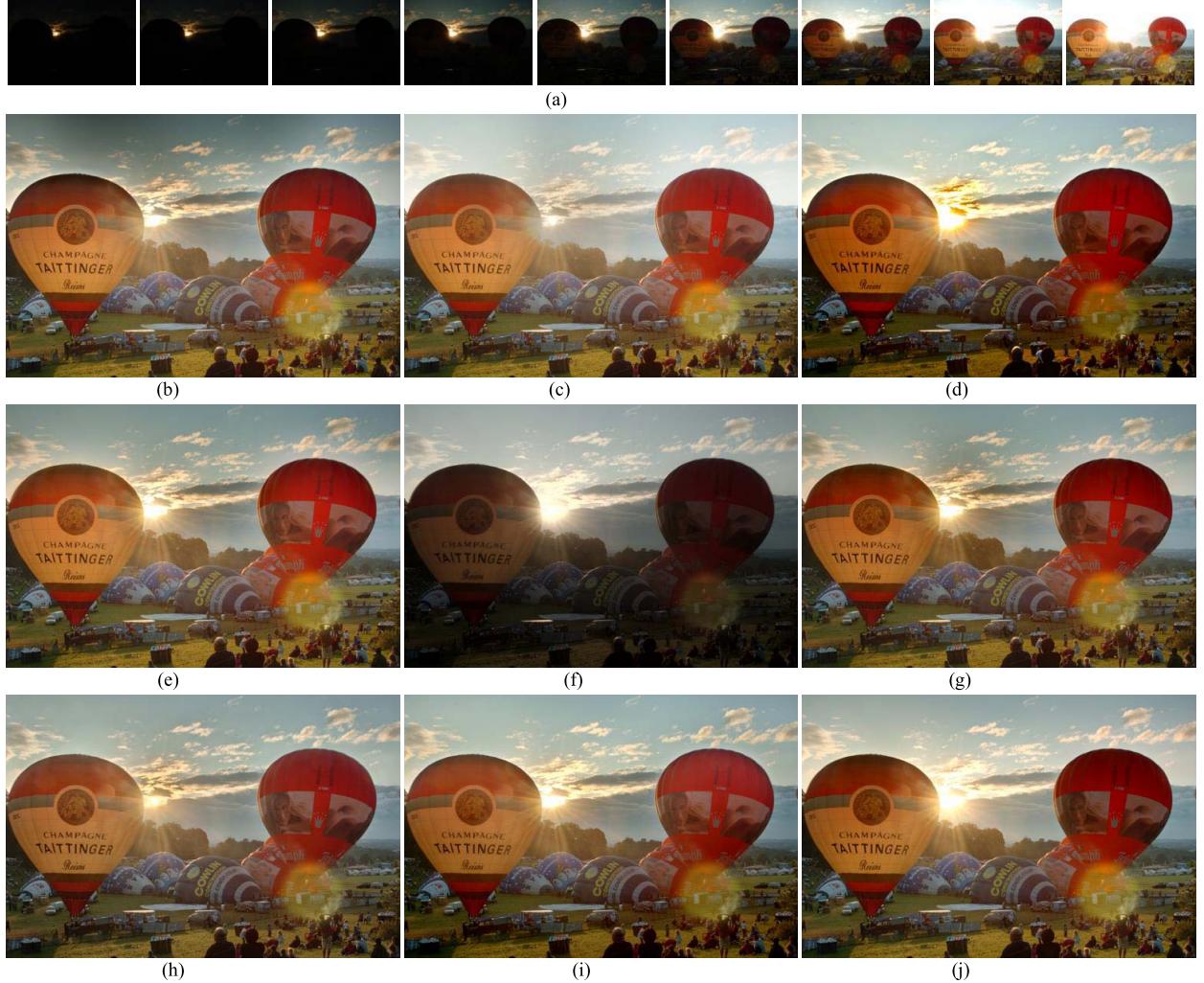


Fig. 7. Experimental results of fusing the multiexposed image “Balloons.” (a) Source image sequence. (b)–(j) Fusion results of nine algorithms. (b) Li13 [19]. (c) Liu15 [10]. (d) Ma17 [2]. (e) Yang18 [42]. (f) Li18 [30]. (g) Wang20 [6]. (h) Ma20 [43]. (i) Li21 [1]. (j) Ours.

which include six traditional methods (i.e., Li13 [19], Liu15 [10], Ma17 [2], Yang18 [42], Wang20 [6], and Li21 [1]) and two deep learning methods (i.e., Li18 [30] and Ma20 [43]). For a fair comparison, all fusion results are generated from the source codes that are publicly available online, with the default parameters suggested by the authors. In order to verify the effectiveness of the proposed method, three typical examples are illustrated in Figs. 7–9, respectively.

The comparison of eight MEF methods with our method on a typical “Balloons” sequence is shown in Fig. 7. As can be seen from Fig. 7, the result generated by Li18 has the problem of low brightness and low contrast. The fusion result obtained by Li13 has inconsistent brightness artifacts in the sky regions. The fusion results of Liu15 and Ma20 obviously have different degrees of over-brightness, and the sun in the results looks tiny, which makes the scenes uncoordinated. Furthermore, the result of Ma20 suffers from loss of color information. The results of Ma17, Yang18, and Wang20 are affected by the halo artifacts in the sun region. In addition, the result generated by Yang18 has low contrast and nonuniform brightness distortions in the cloud area. The fusion result of Li21 looks natural and retains more details, but produces a visible halo around the

balloons. In contrast, our method can better preserve the color and luminance information and avoid the halo artifacts.

Fig. 8 shows the fusion results of different methods on the “ChineseGarden” sequence. In Fig. 8, it is obvious that the fused image of Yang18 has uneven brightness, which leads to a poor visual effect. The fused result generated by Li18 improves the global uniformity, but the overall contrast of the image decreases, resulting in a loss of details, as shown in the red rectangle. The result of Wang20 preserves more details, but the brightness of the entire image is inconsistent. The result of Liu15 has better brightness in the house area, but the sky area is slightly darkened. It is clear that the fused results of Li13, Ma17, Ma20, and Li21 suffer from different degrees of artifacts, which can be seen from the magnified blue rectangles. Compared with the results of the above MEF methods, our result has more details, better brightness, and richer color information, which is more suitable for visual perception.

In Fig. 9, we compare the eight methods with our method on the “Mask” sequence. The results of Yang18 and Li18 have insufficiently low contrast of the entire images, resulting in darker images overall compared with the results of other

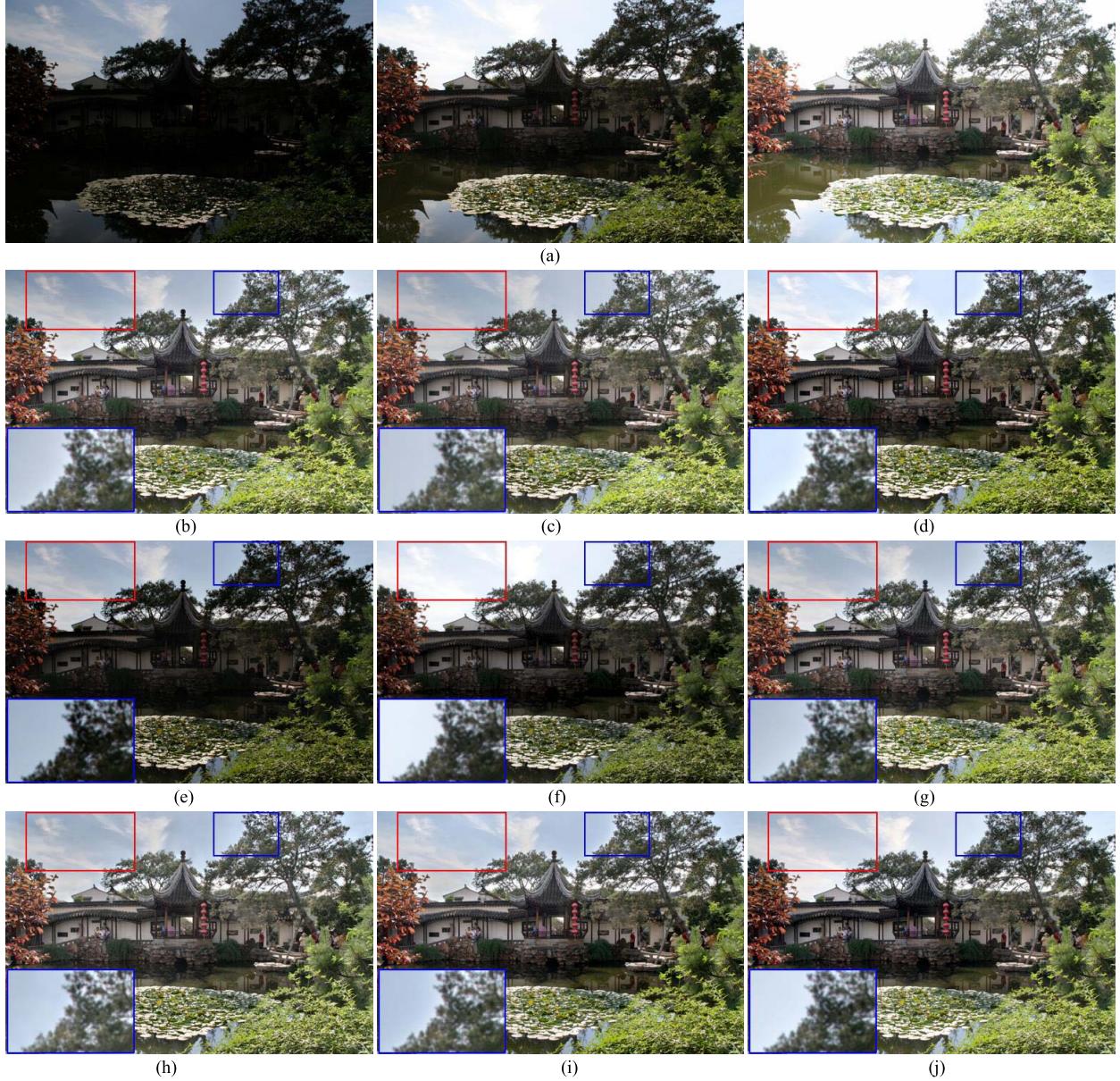


Fig. 8. Experimental results of fusing the multiexposed image “ChineseGarden.” (a) Source image sequence. (b)–(j) Fusion results of nine algorithms. (b) Li13 [19]. (c) Liu15 [10]. (d) Ma17 [2]. (e) Yang18 [42]. (f) Li18 [30]. (g) Wang20 [6]. (h) Ma20 [43]. (i) Li21 [1]. (j) Ours.

methods. The results obtained by Li13 and Liu obviously have inconsistent brightness in the sky areas. The fusion result of Ma17 clearly shows inconsistent brightness in the entire image. Wang20, Ma20, and Li21 perform well in preserving the details, but their results have some artifacts around the top of the house, leading to poor visual effects. In contrast, our method obtains a better fusion result with uniform brightness and less halo artifacts and color distortion.

Through the previous comparison of subjective visual effects, it is shown that our proposed method can better integrate the detail information, color information, and brightness information of the source images with different exposure levels, and the obtained fusion results are better than those of other fusion methods in terms of visual performance.

2) *Quantitative Comparison*: In addition to subjective analysis, we also conducted a quantitative analysis of the

abovementioned comparative methods on the 30 LDR sequences including the source images of Figs. 7–9. We objectively evaluated the performance of these MEF methods through two evaluation metrics: MEF-SSIM and VIFF. Table VI shows the objective assessment results of the proposed method and other comparison methods. In the table, the first three rows correspond to the evaluation results of Figs. 7–9, from which it can be seen that the MEF-SSIM and VIFF values of our method are the highest. This indicates that the objective evaluation is consistent with the analysis of the previous subjective experiments. Besides, on the other images, the two indicators of our method are optimal in most image sequences. In addition, the average score for all the testing sequences is reported as well in the last row. It is found that the average results of our proposed method rank first in both MEF-SSIM and VIFF indicators, which is better



Fig. 9. Experimental results of fusing the multiexposed image “Mask.” (a) Source image sequence. (b)–(j) Fusion results of nine algorithms. (b) Li13 [19]. (c) Liu15 [10]. (d) Ma17 [2]. (e) Yang18 [42]. (f) Li18 [30]. (g) Wang20 [6]. (h) Ma20 [43]. (i) Li21 [1]. (j) Ours.

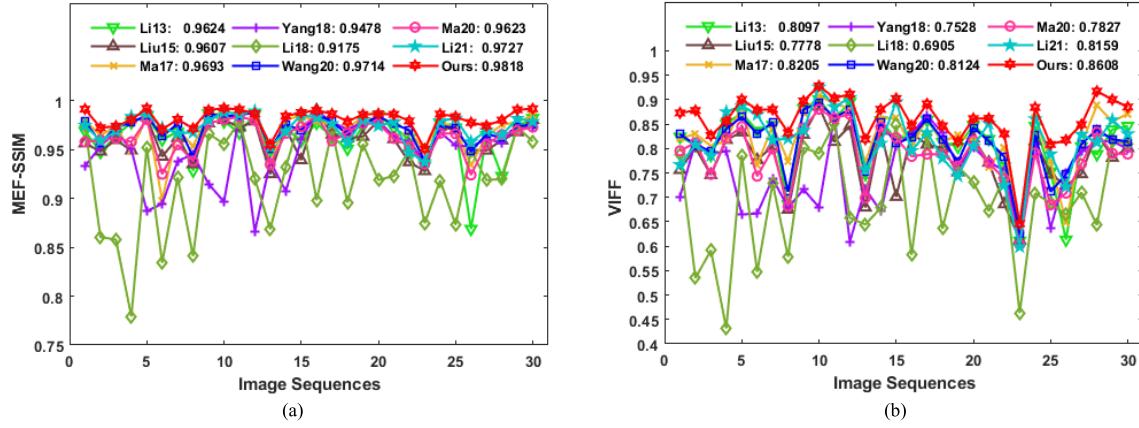


Fig. 10. Quantitative results of different methods on 30 image sequences. (a) MEF-SSIM. (b) VIFF.

than the results of other methods. Among other methods, the multiscale-based methods Li21 and Wang20 outperform the patch-wise method Ma17 on the metric MEF-SSIM,

but their performance on the metric VIFF is lower than that of Ma17. The two metrics of the deep learning-based methods Ma20 and Li18 are inferior to those of the pixel-wise

TABLE VI
PERFORMANCE COMPARISON OF THE PROPOSED METHOD WITH THE EXISTING MEF ALGORITHMS USING MEF-SSIM AND VIFF METRICS

Source Images	Li13 [19]	Liu15 [10]	Ma17 [2]	Yang18 [42]	Li18 [30]	Wang20 [6]	Ma20 [43]	Li21 [1]	Ours
Balloons	0.9476/0.8073	0.9482/0.8014	0.9650/0.8301	0.9534/0.8133	0.8605/0.5347	0.9572/0.8075	0.9600/0.8119	0.9590/0.8080	0.9724/0.8778
ChineseGarden	0.9844/0.8730	0.9835/0.8333	0.9853/0.8893	0.8872/0.6652	0.9526/0.7864	0.9854/0.8646	0.9809/0.8446	0.9870/0.8871	0.9924/0.9906
Mask	0.9788/0.9005	0.9805/0.8863	0.9877/0.9074	0.8967/0.6797	0.9560/0.7900	0.9838/0.8931	0.9808/0.8803	0.9885/0.9255	0.9925/0.9281
Arno	0.9691/0.8241	0.9565/0.7564	0.9796/0.8203	0.9332/0.7015	0.9740/0.7884	0.9790/0.8293	0.9579/0.7952	0.9749/0.7667	0.9915/0.8729
BelgiumHouse	0.9639/0.7869	0.9600/0.7459	0.9728/0.8048	0.9684/0.7885	0.8580/0.5917	0.9726/0.7941	0.9618/0.7471	0.9690/0.7857	0.9740/0.8267
Cave	0.9780/0.8508	0.9493/0.8179	0.9484/0.8160	0.9504/0.7947	0.7792/0.4317	0.9775/0.8402	0.9579/0.8212	0.9839/0.8750	0.9805/0.8562
House	0.9604/0.8362	0.9429/0.7765	0.8981/0.7668	0.8945/0.6667	0.8342/0.5474	0.9636/0.8300	0.9249/0.7433	0.9722/0.8694	0.9708/0.8791
Kluki	0.9682/0.8481	0.9621/0.8051	0.9705/0.8468	0.9384/0.7384	0.9216/0.7287	0.9764/0.8540	0.9550/0.7992	0.9667/0.8169	0.9810/0.8800
Lamp	0.9285/0.6853	0.9353/0.6750	0.9538/0.7742	0.9453/0.6786	0.8410/0.5784	0.9424/0.7115	0.9383/0.6820	0.9676/0.8204	0.9717/0.8323
Laurenziana	0.9872/0.8821	0.9771/0.8286	0.9854/0.8728	0.9149/0.7161	0.9661/0.8015	0.9874/0.8799	0.977/0.8418	0.9802/0.8376	0.9905/0.8972
Office	0.9674/0.8599	0.9729/0.8138	0.9877/0.8922	0.9855/0.8668	0.9785/0.8507	0.9856/0.8639	0.9837/0.8621	0.9858/0.8849	0.9911/0.9032
Tower	0.9864/0.8952	0.9790/0.8461	0.9861/0.8815	0.8661/0.6091	0.9205/0.6578	0.9877/0.8797	0.9799/0.8702	0.9896/0.9040	0.9868/0.9111
Studio	0.9430/0.7330	0.9250/0.6793	0.9335/0.7719	0.9379/0.7179	0.8686/0.6445	0.9505/0.7532	0.9351/0.6988	0.9526/0.7568	0.9551/0.8190
Window	0.9713/0.8575	0.9717/0.8364	0.9816/0.8617	0.9074/0.6712	0.9316/0.6782	0.9758/0.8527	0.9667/0.8441	0.9687/0.8104	0.9843/0.8801
Cove	0.9593/0.8187	0.9393/0.7012	0.9839/0.8626	0.9599/0.8053	0.9854/0.8480	0.9699/0.8107	0.9733/0.8218	0.9857/0.8955	0.9880/0.9023
Frontier	0.9773/0.8271	0.9903/0.8306	0.9796/0.8034	0.9895/0.8311	0.8983/0.5820	0.9864/0.8226	0.9733/0.8218	0.9828/0.7953	0.9903/0.8470
Extinguisher	0.9628/0.8583	0.9654/0.8082	0.9750/0.8623	0.9783/0.8674	0.9696/0.8283	0.9787/0.8608	0.9589/0.7884	0.9766/0.8325	0.9866/0.8914
Willy Desk	0.9509/0.8134	0.9618/0.8007	0.9717/0.8129	0.9708/0.8321	0.8959/0.6358	0.9652/0.8186	0.9672/0.7881	0.9572/0.7802	0.9789/0.8462
HorseshoeLake	0.9741/0.8037	0.9634/0.7514	0.9841/0.8283	0.9779/0.7619	0.9557/0.7633	0.9764/0.7728	0.9768/0.7640	0.9807/0.7437	0.9860/0.8138
Exhibition	0.9852/0.8477	0.9795/0.8210	0.9797/0.8153	0.9786/0.8024	0.9189/0.7304	0.9840/0.8430	0.9765/0.8163	0.9743/0.8051	0.9867/0.8612
Archway	0.9722/0.8195	0.9608/0.7712	0.9636/0.7617	0.9637/0.7727	0.9232/0.6720	0.9765/0.8151	0.9615/0.7709	0.9783/0.8512	0.9861/0.8617
Glass	0.9618/0.7634	0.9373/0.6861	0.9680/0.8028	0.9577/0.7546	0.9497/0.7413	0.9698/0.7830	0.9535/0.7375	0.9473/0.7248	0.9794/0.8299
TableLamp	0.9336/0.6108	0.9281/0.6122	0.9498/0.6223	0.9377/0.6271	0.8747/0.4631	0.9386/0.6241	0.9353/0.6034	0.9374/0.5987	0.9507/0.6459
Wall Drug	0.9769/0.8458	0.9736/0.8154	0.9796/0.8627	0.9679/0.8113	0.9185/0.7085	0.9743/0.8273	0.9668/0.7906	0.9793/0.8605	0.9864/0.8840
Door	0.9726/0.7527	0.9699/0.7472	0.9710/0.7728	0.9543/0.6366	0.8731/0.6877	0.9724/0.7113	0.9663/0.6846	0.9821/0.7886	0.9838/0.8083
Candle	0.8688/0.6133	0.9498/0.7243	0.9340/0.6487	0.9532/0.7505	0.9489/0.6647	0.9480/0.7475	0.9241/0.7086	0.9580/0.7214	0.9777/0.8174
Cars	0.9647/0.8044	0.9492/0.7478	0.9636/0.8087	0.9608/0.7943	0.9188/0.7099	0.9658/0.8091	0.9554/0.7685	0.9695/0.8290	0.9745/0.8489
Rock	0.9230/0.7882	0.9595/0.8333	0.9755/0.8884	0.9572/0.8150	0.9200/0.6433	0.9623/0.8386	0.9654/0.8332	0.9650/0.8126	0.9803/0.9172
Exploratorium	0.9734/0.8389	0.9679/0.7812	0.9782/0.8551	0.9748/0.8148	0.9731/0.8171	0.9736/0.8187	0.9695/0.7898	0.9819/0.8584	0.9908/0.9002
Cemetery	0.9824/0.8463	0.9805/0.8011	0.9851/0.8713	0.9733/0.7990	0.9584/0.8094	0.9761/0.8144	0.9730/0.7885	0.9788/0.8298	0.9919/0.8849
Average	0.9624/0.8097	0.9607/0.7778	0.9693/0.8205	0.9478/0.7528	0.9175/0.6905	0.9714/0.8124	0.9623/0.7827	0.9727/0.8159	0.9818/0.8608

Algorithm 1 Pseudocode of the Proposed Decision Map Construction Method

```

Input: Intensity component  $I_k$ ,  $k = 1, 2, \dots, N$ .
Output: Final decision map  $M_k^{fd}(x, y)$ .
1: for all intensity components ( $I_k$ ) do
2:   Decompose the intensity components  $I_k$  using Eq. (20) to obtain the detail map  $M_k^d$ .
3:   for all pixels at same position ( $x, y$ ) in ( $I_k$ ) do
4:     Construct  $M_k^{ad}(x, y)$ ,  $M_k^{id}(x, y)$  by Eqs. (21)-(22).
5:     Construct  $M_k^{fd}(x, y)$  by Eq. (23).
6:   end for
7: end for

```

method Li13 and the multiscale-based method Yang18, respectively.

In order to intuitively show that our method is superior to other advanced fusion methods, Fig. 10 shows the values

of MEF-SSIM and VIFF for various methods on 30 image sequences by using the line chart. As can be seen from Fig. 10, the values of these two metrics of our method are the largest in most image sequences. It demonstrates that the proposed method can better preserve useful information from source images into the fused image compared to other methods. On the whole, our method achieves a better objective evaluation performance than the comparison methods.

C. Running Time Comparison With State-of-the-Art MEF Methods

To compare the computational efficiency of the different MEF methods, we calculated their average running time on the 30 multiexposure image sequences, as shown in Table VII. In our experiments, the deep learning-based method Ma20 was tested on NVIDIA RTX3060 GPU and AMD Ryzen R5-5600X CPU, and other test methods were implemented in MATLAB R2018a on a computer with a 3.3-GHz CPU

Algorithm 2 Pseudocode of the Proposed MEF Method

Input: Source images M_k , $k = 1, 2, \dots, N$.
Output: Fused image F .

```

1: for all source images ( $M_k$ ) do
2:   for all pixels at same position ( $x, y$ ) in ( $M_k$ ) do
3:     Construct  $I_k(x, y)$ ,  $S_k(x, y)$ ,  $C_k(x, y)$ ,  $E_k(x, y)$  and  $W_k(x, y)$ 
       by Eqs. (10)-(16).
4:   end for
5: end for
6: Construct the final decision map  $M_k^{fd}(x, y)$  using
   Algorithm 1.
7: for all source images ( $M_k$ ) do
8:   for all pixels at same position ( $x, y$ ) in ( $M_k$ ) do
9:     Construct  $W_k^d(x, y)$  and  $W_k^f(x, y)$  by Eqs. (24)-(25).
10:    Construct the fused pyramids  $L\{F\}^l$  by Eq. (30).
11:   end for
12: end for
13: Compute fused image  $F$  by using Laplacian pyramid
   reconstruction on  $L\{F\}^l$ 
```

TABLE VII
AVERAGE RUNNING TIME OF DIFFERENT MEF METHODS

Methods	Time/s
Li13 [19]	1.09
Liu15 [10]	2.16
Ma17 [2]	5.73
Yang18 [42]	2.92
Li18 [30]	5.32
Wang20 [6]	0.73
Ma20 [43]	2.84
Li21 [1]	0.19
Ours	1.91

and 16-GB RAM. As shown in Table VII, since Ma17 is a patch-based method, it is more time-consuming compared to other MEF methods. Besides, the deep learning-based method Li18 takes around five seconds to fuse one image sequence, which is also slow. Li21 has the highest computation efficiency owing to the use of multiscale SPD, but the quality of its generated images is poorer than that of our results. In contrast, our method can achieve acceptable efficiency. Therefore, the proposed method achieves the optimal balance between computational efficiency and fusion performance.

V. CONCLUSION

In this article, we have proposed a novel multiscale exposure image fusion method based on multivisual feature measurement and detail enhancement representation, which is used to fuse image sequences with different exposure levels in the IHS color space. First, the initial weight maps are constructed according to the measured multivisual features of the source images and the adaptive weighting coefficients of these features. Second, the decision map construction method is designed to optimize the initial weight maps to obtain the middle weight maps through comparing the pixel values of the

detail maps of the intensity components. Third, the final weight maps are acquired by using the guided filtering to remove the noise in the middle weight maps. Finally, the image pyramid decomposition and reconstruction algorithm are employed for generating the fused image with better image details and less halo artifacts and color distortion. The experimental results indicate that the proposed method is superior to state-of-the-art MEF methods. In the future, we plan to further improve the computational efficiency of the proposed method.

REFERENCES

- [1] H. Li, T. Chan, X. Qi, and W. Xie, "Detail-preserving multi-exposure fusion with edge-preserving structural patch decomposition," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 31, no. 11, pp. 4293–4304, Nov. 2021.
- [2] K. Ma, H. Li, H. Yong, Z. Wang, D. Meng, and L. Zhang, "Robust multi-exposure image fusion: A structural patch decomposition approach," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, May 2017.
- [3] 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.
- [4] J. Wang, H. Liu, and N. He, "Exposure fusion based on sparse representation using approximate K-SVD," *Neurocomputing*, vol. 135, pp. 145–154, Jul. 2014.
- [5] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: A simple and practical alternative to high dynamic range photography," *Comput. Graph. Forum*, vol. 28, no. 1, pp. 161–171, 2009.
- [6] Q. Wang, W. Chen, X. Wu, and Z. Li, "Detail-enhanced multi-scale exposure fusion in YUV color space," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 8, pp. 2418–2429, Aug. 2020.
- [7] M. Nejati, M. Karimi, S. M. R. Soroushmehr, N. Karimi, S. Samavi, and K. Najarian, "Fast exposure fusion using exposedness function," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 2234–2238.
- [8] X. Zhang, "Benchmarking and comparing multi-exposure image fusion algorithms," *Inf. Fusion*, vol. 74, pp. 111–131, Oct. 2021.
- [9] Y. Que, Y. Yang, and H. J. Lee, "Exposure measurement and fusion via adaptive multiscale edge-preserving smoothing," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4663–4674, Dec. 2019.
- [10] Y. Liu and Z. Wang, "Dense SIFT for ghost-free multi-exposure fusion," *J. Vis. Commun. Image Represent.*, vol. 31, pp. 208–224, Aug. 2015.
- [11] C. Liu, J. Yuen, and A. Torralba, "SIFT flow: Dense correspondence across scenes and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 978–994, May 2011.
- [12] N. Hayat and M. Imran, "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter," *J. Vis. Commun. Image Represent.*, vol. 62, pp. 295–308, Jul. 2019.
- [13] S. Raman and S. Chaudhuri, "Bilateral filter based compositing for variable exposure photography," in *Proc. Eurographics*, 2009, pp. 1–4.
- [14] S.-H. Lee, J. S. Park, and N. I. Cho, "A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1737–1741.
- [15] 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, May 2012.
- [16] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and A. C. Bovik, "Single-scale fusion: An effective approach to merging images," *IEEE Trans. Image Process.*, vol. 26, no. 1, pp. 65–78, Jan. 2017.
- [17] K. Ma and Z. Wang, "Multi-exposure image fusion: A patch-wise approach," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 1717–1721.
- [18] H. Li, K. Ma, H. Yong, and L. Zhang, "Fast multi-scale structural patch decomposition for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 5805–5816, 2020.
- [19] S. Li, X. Kang, and J. Hu, "Image fusion with guided filtering," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2864–2875, Jul. 2013.
- [20] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
- [21] A. Toet, "Image fusion by a ratio of low-pass pyramid," *Pattern Recognit. Lett.*, vol. 9, no. 4, pp. 245–253, 1989.
- [22] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.

- [23] Z. Li, Z. Wei, C. Wen, and J. Zheng, "Detail-enhanced multi-scale exposure fusion," *IEEE Trans. Image Process.*, vol. 26, no. 3, pp. 1243–1252, Mar. 2017.
- [24] Z. Li, J. Zheng, Z. Zhu, W. Yao, and S. Wu, "Weighted guided image filtering," *IEEE Trans. Image Process.*, vol. 24, no. 1, pp. 120–129, Jan. 2015.
- [25] F. Kou, Z. Li, C. Wen, and W. Chen, "Multi-scale exposure fusion via gradient domain guided image filtering," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Jul. 2017, pp. 1105–1110.
- [26] F. Kou, Z. Li, C. Wen, and W. Chen, "Edge-preserving smoothing pyramid based multi-scale exposure fusion," *J. Vis. Commun. Image Represent.*, vol. 53, pp. 235–244, May 2018.
- [27] F. Kou, W. Chen, C. Wen, and Z. Li, "Gradient domain guided image filtering," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 4528–4539, Nov. 2015.
- [28] K. R. Prabhakar, V. S. Srikanth, and R. V. Babu, "DeepFuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 4724–4732.
- [29] Z. Yang, Y. Chen, Z. Le, and Y. Ma, "GANFuse: A novel multi-exposure image fusion method based on generative adversarial networks," *Neural Comput. Appl.*, vol. 33, no. 11, pp. 6133–6145, Jun. 2021.
- [30] H. Li and L. Zhang, "Multi-exposure fusion with CNN features," in *Proc. 25th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2018, pp. 1723–1727.
- [31] H. Xu, J. Ma, and X.-P. Zhang, "MEF-GAN: Multi-exposure image fusion via generative adversarial networks," *IEEE Trans. Image Process.*, vol. 29, pp. 7203–7216, 2020.
- [32] N. R. Draper and H. Smith, *Applied Regression Analysis*. New York, NY, USA: Wiley, 1981.
- [33] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3345–3356, Nov. 2015.
- [34] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [35] Y. Han, Y. Cai, Y. Cao, and X. Xu, "A new image fusion performance metric based on visual information fidelity," *Inf. Fusion*, vol. 14, pp. 127–135, Apr. 2013.
- [36] C.-L. Chien and W.-H. Tsai, "Image fusion with no gamut problem by improved nonlinear IHS transforms for remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 651–663, Jan. 2014.
- [37] Y. Liu, D. Wang, J. Zhang, and X. Chen, "A fast fusion method for multi-exposure image in YUV color space," in *Proc. IEEE 3rd Adv. Inf. Technol., Electron. Autom. Control Conf. (IAEAC)*, Oct. 2018, pp. 1685–1689.
- [38] Y. Yang, J. Wu, S. Huang, and P. Lin, "Multiexposure estimation and fusion based on a sparsity exposure dictionary," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4753–4767, Jul. 2020.
- [39] Waterloo IVC Multi-Exposure Fusion Image Database. Accessed: Apr. 2020. [Online]. Available: <http://ivc.uwaterloo.ca/database/MEF/MEF-Database.php>
- [40] The HDR Photographic Survey. Accessed: Jul. 2020. [Online]. Available: <http://rit-mcsrl.org/fairchild/HDR.html>
- [41] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2049–2062, Apr. 2018.
- [42] Y. Yang, W. Cao, S. Wu, and Z. Li, "Multi-scale fusion of two large-exposure-ratio images," *IEEE Signal Process. Lett.*, vol. 25, no. 12, pp. 1885–1889, Dec. 2018.
- [43] K. Ma, Z. Duannmu, H. Zhu, Y. Fang, and Z. Wang, "Deep guided learning for fast multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 2808–2819, 2020.



Yong Yang (Senior Member, IEEE) received the Ph.D. degree from Xi'an Jiaotong University, Xi'an, China, in 2005.

From 2009 to 2010, he was a Post-Doctoral Research Fellow with Chonbuk National University, Jeonju, South Korea. He is currently a Distinguished Professor with the School of Computer Science and Technology, Tiangong University, Tianjin, China. His current research interests include image fusion, image super-resolution reconstruction, medical image processing and analysis, and pattern recognition.

Dr. Yang is an Associate Editor of the IEEE ACCESS and an Editor of the KSII Transactions on Internet and Information Systems.



Danjeie Zhang received the B.S. degree in computer science and technology from Ningxia Normal University, Ningxia, China, in 2020. She is currently pursuing the M.S. degree in computer science and technology with the Jiangxi University of Finance and Economics, Nanchang, China.

Her current research interests include image fusion and multiexposure image processing and analysis.



Weiguo Wan received the B.S. degree in mathematics and applied mathematics from Jiangxi Normal University, Nanchang, China, in 2014, and the Ph.D. degree in computer science and engineering from Jeonbuk National University, Jeonju, South Korea, in 2020.

He is currently a Lecturer with the School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang. His current research interests include computer vision, deep learning, face sketch synthesis and sensing image fusion.



Shuying Huang (Member, IEEE) received the Ph.D. degree in computer application technology from the Ocean University of China, Qingdao, China, in 2013.

She is currently a Professor with the School of Software, Tiangong University, Tianjin, China. Her current research interests include image and signal processing, and pattern recognition.