Multi-Focus Image Fusion With Point Detection Filter and Superpixel-Based Consistency Verification

QIANG CHEN🆔, BIN YANG🆔, (Member, IEEE), YUEHUA LI, AND LIHUI PANG

College of Electric Engineering, University of South China, Hengyang 421001, China

Corresponding author: Bin Yang (yangbin01420@163.com)

ABSTRACT An accurate and efficient measurement of pixel's sharpness is a critical factor in most multi-focus image fusion methods. In our practice, we found that the focused regions get more blurred than the defocus regions when the multi-focus images are blurred digitally. Based on this observation, a novel multi-focus image fusion method is presented in this paper. In the given fusion scheme, focused regions detection is achieved by point detection filter and Gaussian filter, which has been certified more effective than other frequently used image clarity measures. Moreover, unlike the other commonly used consistency verification, we propose a superpixel-based consistency verification (SCV) method by integrating the image superpixels to improve the fusion performance. Image superpixels can perceptually represent meaningful image local features. Two datasets of multi-focus images are used to conduct experiments. Experimental results demonstrate that the proposed method can be competitive with or even outperform the state-of-the-art fusion methods in terms of both subjective visual perception and objective evaluation metrics.

INDEX TERMS Clarity measure, consistency verification, image superpixel, multi-focus image fusion, point detection filter.

I. INTRODUCTION

With the development of electronics and computer technology, multiple imaging sensors have been synergistically used in various image processing systems. Although there are many kinds of sensors, most optical cameras cannot capture a single sharper image with all related objects in focus, due to optical lens' limitation of depth-of-focus [1]. Multi-focus image fusion is an effective technology to solve this problem. In this technology, we firstly shoot multiple images of the same scene with different focus settings. Then the source images are merged to get a composite image in which all objects in the scene are in focus. The composite image also known as the fused image which is more suitable for machine or human perception [2]. These years, multi-focus image fusion technique has been successfully applied in many

image processing fields such as digital imaging, microscopic imaging, machine vision and so on [3]–[6].

During the past two decades, a great variety of multi-focus image fusion algorithms have been proposed and these algorithms can be roughly divided into two categories: transform domain algorithms and spatial domain algorithms [7]. Transform domain algorithms share a 'decomposition-fusion-reconstruction' framework. It decomposes the input images into their multi-scale domain firstly. Then the transformed coefficients are fused by certain fusion rules. Finally, the fused image is reconstructed from the integrated coefficients. Pyramid transform [8], wavelet transform [9], contourlet transform [10], curvelet transform [11], shearlet transform [12] are commonly selected as the decomposition method. Transform domain algorithms are usually treated as general image fusion methods which can be used to achieve large range of fusion tasks such as multi-focus, visual and infrared image, remote sensing and so on. More recently,

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

sparse representation (SR) also has recently attracted significant attention in image fusion [13]–[15]. In these methods, the source image is represented with sparse coefficients using an over-complete dictionary. Then, the coefficients are combined with the fusion rules. Finally, the fused image is reconstructed from the combined sparse coefficients and the dictionary. Apart from the selection of transform methods, the fusion rules designed for merging transformed coefficients in high- or low-frequency domain also play a critical role in these methods, and many researches have also been taken in this direction [16], [17].

Spatial domain-based fusion methods also attract numerous researches since human eyes perceive images in spatial domain. Averaging the pixel values of all source images directly is the simplest algorithm in pixel-based image fusion methods, but it leads to blurring effect, contrast reducing and original image information loss. Aiming to make full use of the spatial context information, Li *et al.* first divide the source images into uniform blocks and maximize the spatial frequency (SF) [18] in each block [19]. However, the fusion result may produce undesirable block artifacts on object boundaries. To mitigate these artifacts, many region based image fusion methods have been presented. The region based approaches first segment the source images into regions rather than blocks by using image segmentation techniques such as normalized cut [20], watershed transform [21], and enhance linear spectral clustering [7]. Then the divided regions with different source images are fused according to their clarity, such as SF, energy of Laplacian of the image (EOL) [22], sum-modified-Laplacian (SML) [23] and FSWM [24]. Obviously, image segmentation techniques as well as clarity measures have great influence on the fusion results.

Recently, several state-of-the-art spatial domain fusion methods have been proposed, including multi-scale weighted gradient [25], guided filtering [26], dense SIFT [2]. The Dense SIFT methods employed the local feature descriptors dense SIFT as the activity level measurement and to match the mis-registered pixels between multiple source images to improve the fusion quality in object boundaries [2]. The multi-scale weighted gradient-based image fusion method reconstructed the fused image by making its gradient as close as possible to the magnitude of the merged gradient rather than employing a decision map [25]. Li *et al.* first employed guided filter to optimize the weighted coefficients of base layers and detail layers, and obtained satisfactory fusion results [26]. The utilization of guided filter alleviates the misalignment of decision map with object boundaries. Similarly, Qiu *et al.* proposed a novel focus region detection method based on guided filter and mean filter, guided filter is not only used to solve the problem of initial decision map misalignment with object boundaries, but also applied to refine rough focus maps in focus region detection [27]. Although these newly presented algorithms could improve the visual qualities of the fusion images, most of them involve

relatively complex procedures and might produce artifacts around the object boundaries due to the inaccurate fusion decision maps. In addition, to overcome issues in handcraft methods, recently CNN-based image fusion approaches have been applied to image fusion [28], [29]. Liu *et al.* first introduced CNN to fuse multi-focus image. They formulated multi-focus image as a classification task and used their model to classify image patches as focused or defocused to obtain the focus map [28]. However, the use of empirical risk minimization and max pooling limit their fusion performance. Thus, Du *et al.* [29] presented a deep support value convolutional neural network (DSVCNN) by replacing the empirical risk minimization and max pooling layer by structural risk minimization and standard convolutional layers. Due to the strong capability in feature extraction as well as in data representation, the CNN based methods also achieve state-of-the-art performances. However, these CNN-based methods are time-consuming and require a lot of datasets.

In most multifocus image fusion methods, the consistency verification is usually used as a post-processing technique to remove the misjudgment in the final fused results. Majority filter is the earliest consistency verification based on the idea that neighboring coefficients should maintain a very consistent focus judgment [30]. For mathematical morphology-based fusion methods, the focus misjudgment in initial focus segmentation map is seen as noise or undesirable artifacts which can be easily removed by morphology operators [31]. However, the nonlinear mathematical morphology has the risk of removing the details in the fused images if the morphology operator is designed irrationally. In the guided filtering-based image fusion method [26], although the simple Laplacian coefficients are used as saliency measure, the quality of fused result has a great improved due to the guided filtering-based consistency verification. A reasonable consistency verification strategy can evidently improve the performance of most fusion scheme.

In this paper, a novel multi-focus image fusion method with point detection filter-based focus measure and superpixel-based consistency verification is proposed. In experimental practice, we find that if an image is digitally blurred with a blur kernel, the local quality of the focus areas will degrade more evidently than that of defocus areas. Based on this observation, we can separate the focused areas of different image by comparing their image quality degradation between the raw image and the corresponding blurred image. We presented a new approach to evaluate the pixel's sharpness by convoluting the input image with a point detection filter. This approach is effective and easy to implement. The performance evaluation of the proposed clarity measure is demonstrated in Section II-A. Comparing the quality changes of the source images, the focused and defocused areas are separated using the focus map which is further optimized by SCV. Different from the region based image
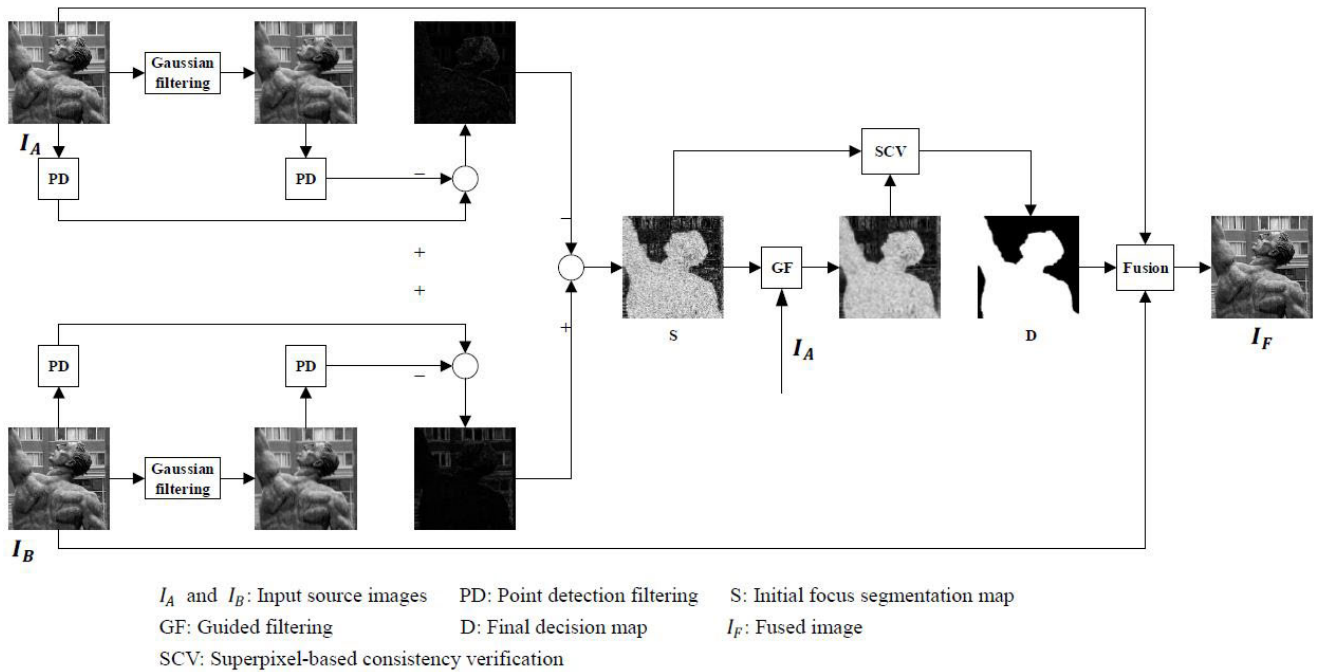
I_A and I_B: Input source images   PD: Point detection filtering   S: Initial focus segmentation map
GF: Guided filtering   D: Final decision map I_F: Fused image
SCV: Superpixel-based consistency verification

FIGURE 1. The schematic diagram of the proposed fusion algorithm.

fusion methods, the superpixel algorithm simple linear iterative clustering (SLIC) [32] is used for consistency verification. Finally, the fused image is constructed by integrating the focus regions in each source image according to the refined focus map. The effectiveness of our method is validated on two datasets under three objective quality metrics. The results are compared with five state-of-the-art multi-focus image fusion methods. Experimental results show that the proposed algorithm is competitive with these methods, in both terms of subjective visual perception and quantitative evaluations.

To sum up, this paper has following contributions:

(1) This paper presents a novel measure on image pixel's sharpness. The strategy is inspired by the found that image quality of focus areas is degraded significantly more than that of blurred areas when an image is digitally blurred. Furthermore, the present sharpness evaluate strategy can be easily implemented using Gaussian filtering and point detection filter.

(2) Inspired by the success of consistency verification in multi-focus image fusion, this paper proposed an extension of guide filter based consistency verification to integrate the image superpixels, which demonstrates better performance in visual perception.

(3) To demonstrate the efficacy of fusion strategy and show its superiority over other state-of-the-art alternatives, experiments on two datasets of multi-focus images are conducted. The results demonstrate the superiority of the proposed method in both subjective visual perception and objective evaluation metrics.

All frequently used symbols in this paper are listed in Table 1.

TABLE 1. List of all frequently used symbols.

Symbol	Definition
I_A, I_B	input source images
I_F	fused image
M_q	the Gaussian blurred image of I_A and I_B
T_q	activity level map of I_A and I_B
W	the point detection filter
H	the 5×5 Gaussian kernel
S	initial segmentation map
D	finial decision map
σ	the standard deviation in Gaussian kernel
r	the block size in guided filtering
K	the number of superpixels k in SLIC

II. MULTI-FOCUS IMAGE FUSION FRAMEWORK

The schematic diagram of the proposed fusion algorithm is illustrated in Fig. 1, in which we can see that there are three main steps in the proposed fusion method. In the first step, the pixel's sharpness of the input images is calculated by the proposed Gaussian filtering and point detection filter-based clarity measures. In the second step, initial focus segmentation map S is constructed by comparing the sharpness of the input images. The SCV is used to refine the initial focus segmentation map. Finally, the fused image I_F is constructed

by combining the source images I_A and I_B according to the final decision map D.

A. THE CLARITY MEASUREMENT WITH POINT DETECTION FILTER

In experimental practice, we found that if an image is digitally blurred with a blur kernel, the local quality of the focused areas will degrade more evidently than that of defocused areas.
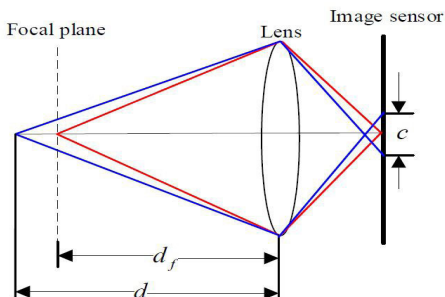

FIGURE 2. Geometric optics model.

Fig. 2 illustrates optical lens imaging model. As shown in Fig 2, when the distance between the object and the lens is equal to the focal length d_f, all the rays from the point of the object converge to a single sensor point and generate a sharp image in the sensor area. Otherwise, when $d \neq d_f$, these rays will generate a blurred image in the image sensor.

The defocus image can be modeled as the convolution of a sharp image with the point spread function (PSF) [33]. The PSF can be approximated by a Gaussian function $g(x, \sigma)$, where the standard deviation $\sigma = kc$ measures the defocus blur amount, the smaller the standard deviation, the better the image quality. As the amount of defocus blur is estimated at edge locations, we model two blurred edge with Gaussian function $g(x, \sigma_1)$ and $g(x, \sigma_2)$ as follows

$$f_1(x) = s(x) * g(x, \sigma_1) \tag{1}$$
$$f_2(x) = s(x) * g(x, \sigma_2) \tag{2}$$

where $s(x)$ is the step function. Note that the edge is located at $x = 0$. $f_1(x)$ and $f_2(x)$ are focus and defocus edges since $\sigma_1 < \sigma_2$. Fig 3 shows the overview of our blur estimation method. $f_1(x)$ and $f_2(x)$ are re-blurred using a Gaussian function $g(x, \sigma_3)$.

$$f_3(x) = f_1(x) * g(x, \sigma_3) \tag{3}$$
$$f_4(x) = f_2(x) * g(x, \sigma_3) \tag{4}$$

It can be clearly noted that the differences between $f_1(x)$ and $f_3(x)$ are larger than that between $f_2(x)$ and $f_4(x)$.

In order to verify the observation for 2D image, we selected a focused region Fig. 4(b) and a defocused region Fig. 4(c) from Fig. 4(a). Then both are blurred by Gaussian blur with standard deviation set from 0.5 to 4. The SF is calculated from each blurred result and illustrated in Fig. 5. The blue plot shows the SF in defocused region while the red plot shows
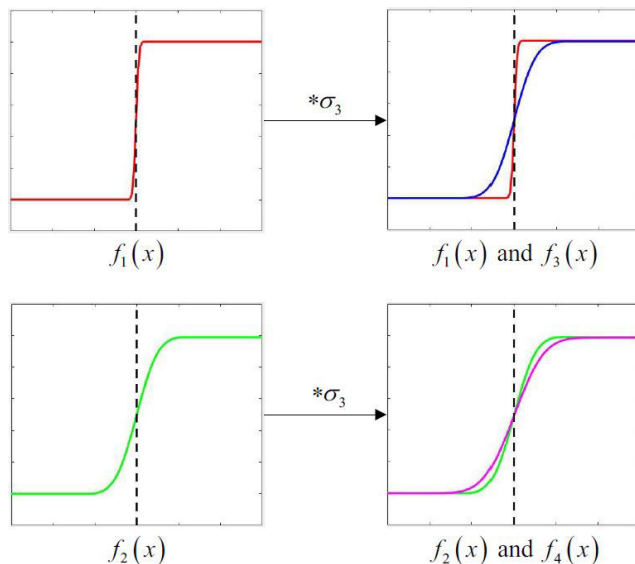

FIGURE 3. The overview of our blur estimation approach: here, is the convolution operators. The black dash line denotes the edge location.

the SF in defocused region. It is evident that the SF in focused region degrade more sharply than that in defocused region. This situation is the same as the 1D edge model.

Inspired by this observation, we proposed a new method to evaluate the sharpness of a pixel as following.

Step 1: Blurring the input source images by an average filter. For both source images I_A and I_B, their blurred versions M_q are obtained by Gaussian kernel as (5).

$$M_q = H * I_q, \quad q \in \{A, B\} \tag{5}$$

where I_q denotes the source image and H denotes the 5×5 Gaussian kernel with standard deviation σ.

Step 2: In multi-focus images, the texture of the focus region is rich relative to the defocus region, which means to a large extent that the difference between the pixels in the focus area is larger than that in the defocus area. Point detection filter can calculate the sum of the difference between the center pixels and their surrounding pixel. Thus when we convolute an image with a point detection filter, the larger the result, the sharper the pixel. Absolute difference of the original image and the blurred image yields the activity level map as (6).

$$T_q = \left| I_q * W \right| - \left| M_q * W \right|, \quad q \in \{A, B\} \tag{6}$$

where I and M are the original image and blurred image respectively. W is a point detection filter defined as (7).

$$W = \begin{pmatrix} -0.0417 & -0.0417 & -0.0417 & -0.0417 & -0.0417 \\ -0.0417 & -0.0417 & -0.0417 & -0.0417 & -0.0417 \\ -0.0417 & -0.0417 & 1 & -0.0417 & -0.0417 \\ -0.0417 & -0.0417 & -0.0417 & -0.0417 & -0.0417 \\ -0.0417 & -0.0417 & -0.0417 & -0.0417 & -0.0417 \end{pmatrix}$$
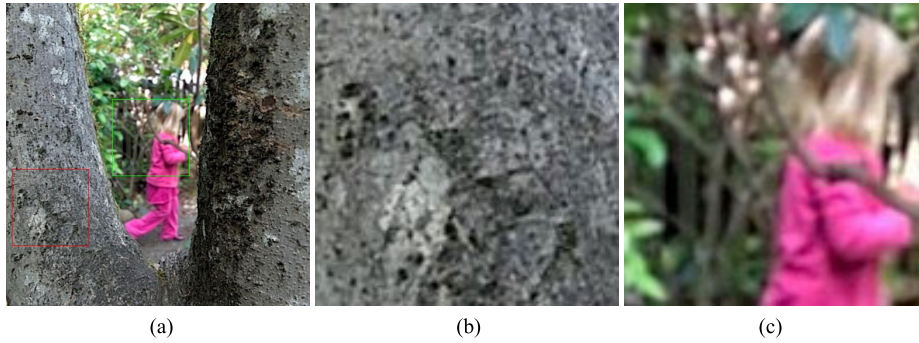$$\tag{7}$$

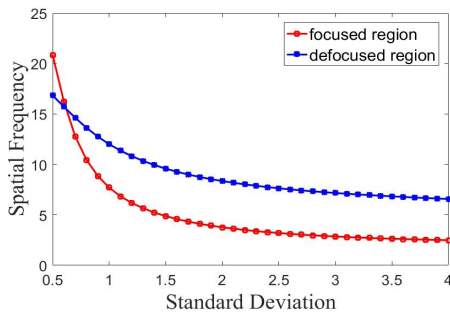**FIGURE 4.** An outdoor image with partial defocus. (a) original image; (b) focused region; (c) defocused region.

FIGURE 5. Spatial frequency in the focused region and defocused region different standard deviation.

FIGURE 6. Two gray test images 'clock' and 'disk'. The first column are ground-truth reference images (focus everywhere); The second and third columns are source images with focus on right and left respectively.

The evaluate strategy is based on the idea that image quality of focus areas is degraded significantly more than that of blurred areas when an image is digitally blurred.

We used the strategy in [22] to evaluate the validity of the novel sharpness measuring. Firstly, two multi-focus source images are decomposed into small blocks with equal size. Then, we use the proposed method as well as other widely used focus measuring methods to calculate the sharpness of each image block. Finally, the blocks with greater sharpness measures from each block pair are chose to construct the composite image. We use root mean square error (*RMSE*) between the composite image and ground truth reference image to evaluate the performance of the focus measures. Note that no consistency verification was used in this process, as it will disturb the results obtained by focus measures directly.

RMSE is defined as

$$RMSE = \sqrt{\frac{\sum_m \sum_n (R(m,n) - F(m,n))^2}{M \times N}} \quad (8)$$

where R and F denote ground-truth reference image and the composite image respectively. $F(m,n)$ is the pixel's intensity at position (m,n). M and N denote the rows and columns of the image.

Fig. 6 shows two test images 'clock' and 'disk' with sizes 512×512 and 640×480, respectively. In experiments, the block size is set to be 32×32 pixels and the standard deviation of Gaussian filter in (1) is set to 2. The main image clarity measures employed in spatial domain include SF, EOL, SML, and FSWM. Since Huang *et al* in [22] has proven that SML and EOL perform better than other three common focus measures, we only compared with SML, EOL and neighbor distance filter (ND) proposed in [34]. Fig. 7 shows the selection maps and fused images obtained by ND, EOL, SML, and our method, respectively. Apparently, the performance of SML and our method in clarity measure are better than those of ND and EOL. Fig. 8 shows the quantitative comparisons of different focus measures in term of *RMSE*. The proposed focus measure achieves the lowest *RMSE*, which implies the composite image obtained by our method is certainly closer to the reference image.

In this paper, we use the absolute value of sharpness measure in (6) to yield an initial focus segmentation map as.

$$S(m,n) = \begin{cases} 1 & |T_A(m,n)| > |T_B(m,n)| \\ 0 & elsewise \end{cases} \quad (9)$$

where $S(m,n) = 1$ denotes that pixel in I_A is focused, elsewise the pixel in I_B is focused.

Fig. 9(a) and (b) are two gray source images of the same scene. Fig. 9(a) is focused on the statue while Fig. 9(b) is focused on the background building. Fig. 9(c) and Fig. 9(d)
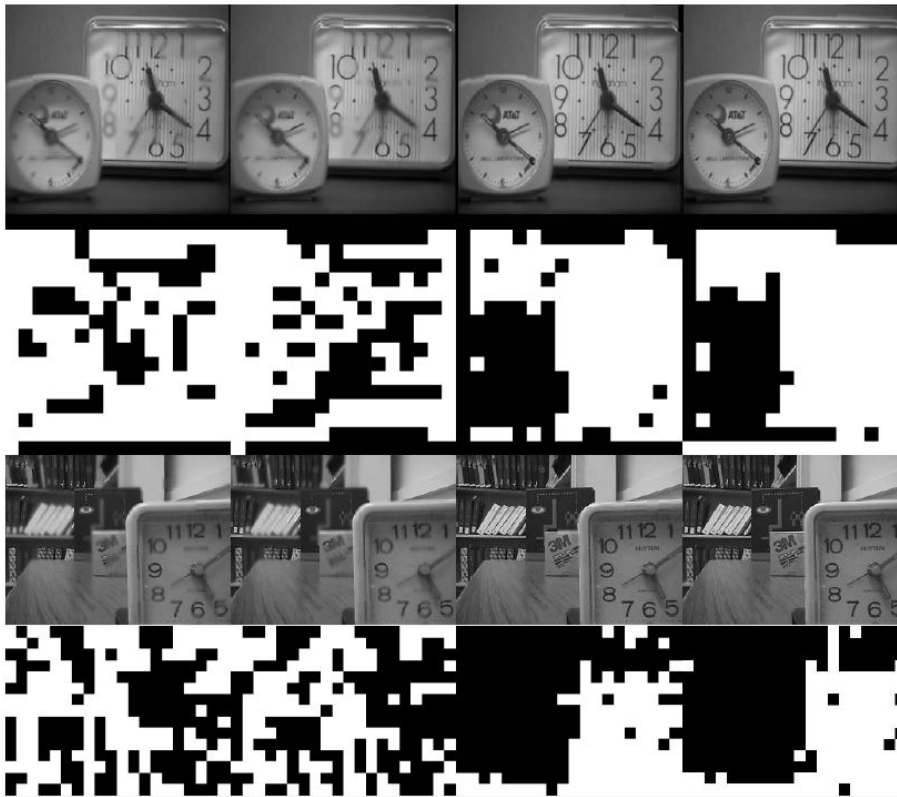
FIGURE 7. The fused 'clock' and 'disk' images and selection maps. From left to right are experiment results obtained by ND, EOL, SML and our method, respectively.
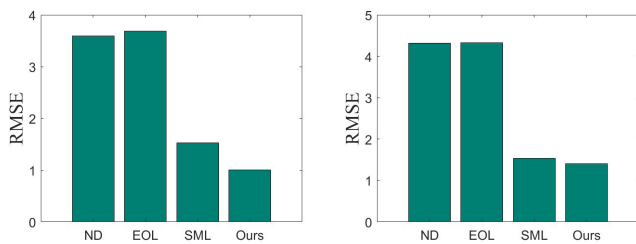

FIGURE 8. Evaluation of different focus measure by *RESM* in 'clock' and 'disk'.

FIGURE 9. An example of focus region detection. (a) and (b) are multi-focus images; (c) and (d) are Gaussian filtered images of two source images; (e)-(h) obtained based on W convoluting with (a)-(d) respectively; (i) initial focus segmentation map S; (j) the refined segmentation map filtered by guided filtering; (k) final decision map obtained by superpixel based consistency verification; (l) the final fused image.

are their Gaussian filtered versions. Fig. 9(e)-(h) are the processing results of Fig. 9(a)-(b) by point detection filter. As shown in Fig. 9(e) and (g), we can see the background of the two images has changed little, but the foreground has obvious changes. The same situation is illustrated in Fig. 9(f) and (h). Based on this observation, an initial focus segmentation map obtained by (9) illustrated in Fig 9(i), where the focus area and the non-focus area are roughly separated.

B. SUPERPIXEL BASED FUSION MAP CONSISTENCY VERIFICATION

The initial focus map is always not a perfect segmentation as shown in Fig. 9(i). The essential cause of situation is that the focus segmentation is usually a high-level vision task based on the understand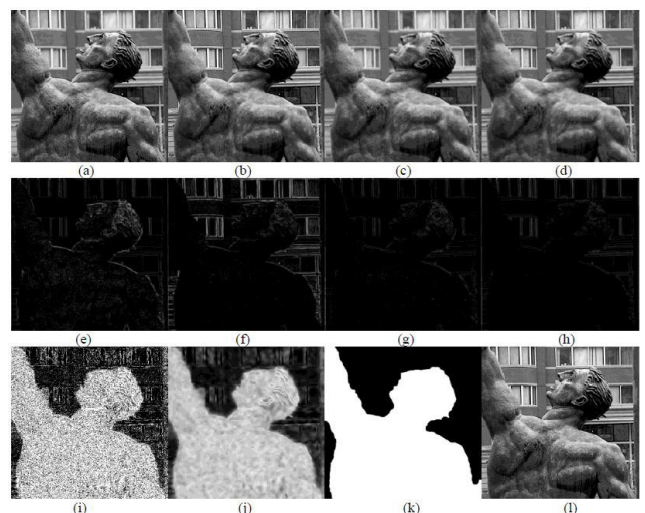ing the scene globally. However, most focus measures are based on local pixel values statistics features. Therefore, the initial focus map is usually post-processed with consistency verification to remove the misjudgment in the final fused results.
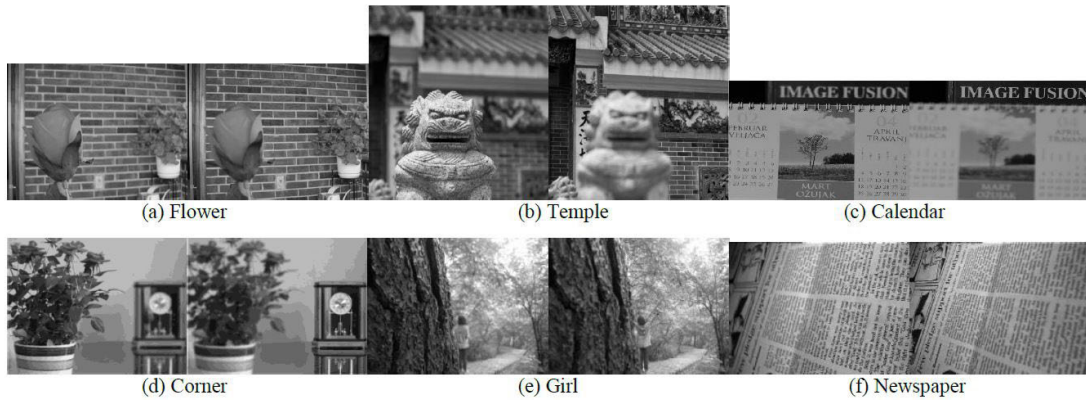
FIGURE 10. Multi-focus image pairs from Grayscale dataset.

In this paper, we comprehensively use superpixel, guided filters, and morphology operators to realize the robust consistency verification. More especially, the superpixel segmentation is used to cluster a set of image pixels that share similar visual characteristics. The superpixels represent perceptually meaningful image local features. The processing flow of our consistency verification scheme can be seen in Fig. 1. Firstly, initial focus segmentation map S is preprocessed with a guided image filtering and the source image is performed as guided images. This operation makes the focus map consistent with the source images content. Then, the filtered map is further segmented into small visual regions using SLIC which is an adaptation of k-means for superpixel generation. It operates faster and more memory efficient than existing methods. The parameter k adjusts desired number of approximately equally sized superpixels. The statistics number of '0' and '1' in ith superpixel region is denoted by NR_i. Finally, the final decision map D is constructed as

$$D_{j \in R_i} = \begin{cases} 1 & NR_i^1 \geq NR_i^0 \\ 0 & elsewise \end{cases} \quad (10)$$

where R_i denotes the ith superpixel region. NR_i^1 and NR_i^0 is the number of 1 and of 0 in R_i of the initial focus segmentation map S, respectively.

Furthermore, we use morphological operators to refine the decision map D. An example of decision map acquisition is demonstrated in Fig. 9. The initial focus segmentation map S is shown in Fig. 9(i) in which the misjudgments illustrated as noisy or cracks. Fig. 9(j) is the focus map processed by guided filtering. Fig. 9(k) is the final decision map obtained by superpixel-based consistency verification which yields a smooth and edge-aligned labeling.

C. FUSED RESULT

With the final decision map D, the fused image is obtained by (11)

$$I_F = I_A \dot{\times} D + I_B \dot{\times} (1 - D) \quad (11)$$

where $\dot{\times}$ denotes the point multiple operator.

Noticed that the proposed method can easily extend to achiever color image fusion. The input color images are translated into YUV color space. The intensity component Y is used as input gray images to get the clarity segmentation map. Then, we adopt the intuitive scheme that fuses three original RGB color channels separately.

III. EXPRIMENTS

In this section, we conduct numerical experiments to verify our main results. We first introduce the datasets used to test the performance evaluation of different methods. Then the influences of three free parameters on fusion performance are analyzed. Finally, the performance of the proposed fusion method is compared with other state-of-the-art alternatives. We also provide the experimental results discussion in details. The experiments are performed on an Intel(R) Core^{TM}i5-8250U @1.60GHz PC with the MATLAB R2016a simulation software.

A. IMAGE DATASETS

In experiments, we evaluate the proposed fusion algorithm on two image datasets. The first one is the grayscale multi-focus dataset which contains 6 pairs of commonly used grayscale images in many related papers. The second one is color image dataset composed of 6 pairs of color multi-focus images of size 520 × 520 which are selected from a multi-focus image dataset ''Lytro'' and is publicly available online [35]. Fig. 10 shows the grayscale multi-focus image dataset and Fig. 11 shows the color multi-focus image dataset.

B. OBJECTIVE EVALUATION METRICS

In our experiments, we exploit seven evaluation metrics to objectively evaluate the performance of fused images obtained by different algorithms, which are information theory based metric Q_{MI} [36], Q_{FMI} [37], image feature based metric $Q^{AB/F}$ [38], Q_E[39], structural similarity based metric Q_Y [40], Q_W[39], and human perception based metric Q_{CB}[41]. For all the seven evaluation metrics, a larger value indicates a better fused result.
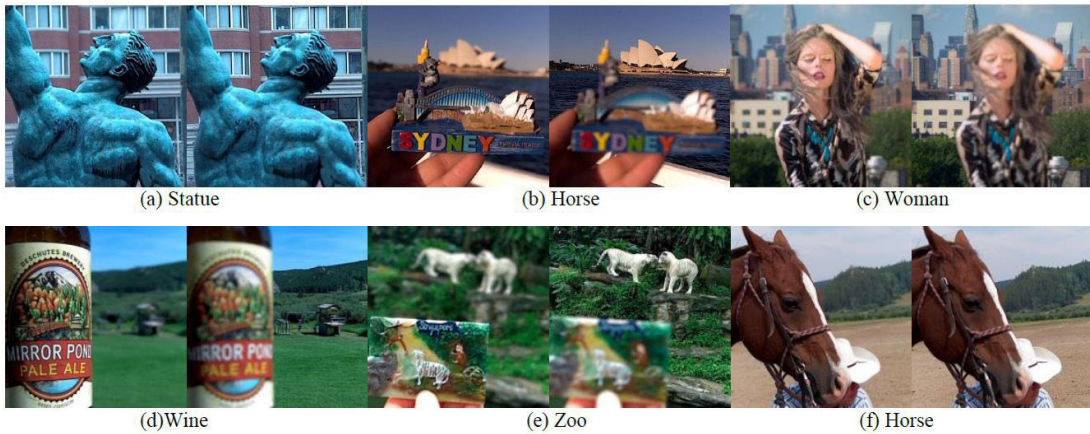
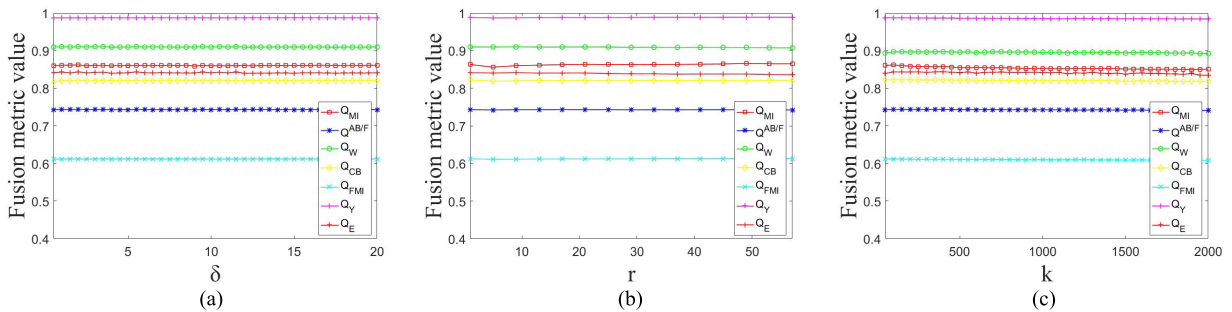**FIGURE 11.** Multi-focus image pairs from Lytro dataset.

FIGURE 12. Performance of the proposed method with different free parameters setting. (a) standard deviation δ; (b) block size r in guided filtering; (c) the number of superpixels k in SLIC.

C. ANALYSIS OF THE FREE PARAMETERS

In this subsection, the influences of different parameters to objective fusion performance are analyzed with an image dataset shown in Fig. 8 and Fig. 9. The fusion performances are evaluated by the average values of Q_{MI}, Q_{FMI}, $Q^{AB/F}$, Q_E, Q_Y, Q_W, Q_{CB}, and Q_{CB}. Obviously, the standard deviation σ in (1), the block size r in guided filtering and the number of superpixels k in SLIC are three free parameters in the proposed method. When analyzing the influence of σ, block size r is preset as 10 and the number of superpixels k is set as 100. In like manner, when we analyze the influence of block size r in guided filtering, we set the standard deviation σ to 2 and set the number of superpixels k to 100. Similarly, when analyzing the influence of the parameter k in SLIC, standard deviation σ is set to 2 and block size r is set to10. As shown in Fig. 12, the fusion performance show sign of levelling off with large rang. The evaluation metrics in Fig. 12(a)-(c) illustrated that our method does not rely heavily on the exact parameter choice. In this paper, the default parameters are set as $\sigma = 2$, $r = 10$, and $k = 100$.

D. EXPRIMENTAL RESULTS AND DISCUSSION

To confirm the effectiveness of the proposed method, we compare our algorithm with five state-of-the-art multi-focus image fusion approaches, including convolutional

neural network based image fusion algorithm [28], multi-scale focus measures and generalized random walk based image fusion algorithm [41], guided filter and mean filter based image fusion algorithm [27], guided filtering based image fusion algorithm [26], image matting based image fusion algorithm [43], and boundary finding based algorithm [44]. For convenience, convolutional neural network based image fusion algorithm is abbreviated as CNN; multi-scale focus measures and generalized random walk based image fusion algorithm is abbreviated as MFGR; guided filter and mean filter based image fusion algorithm is abbreviated as GFMF; guided filtering based image fusion algorithm is abbreviated as GF; image matting based image fusion algorithm is abbreviated as IM; boundary finding based algorithm abbreviated as BF. The source codes of CNN can be downloaded from website [45]. The source codes of MFGR can be downloaded from website [46]. The source code of GFMF can be downloaded from website [47]. The source codes of BF can be downloaded from website [48], and the source codes of IM and GF can be downloaded from website [49]. The optimal parameters reported in the related publications are used in our experiments.

The qualitative evaluation of multi-focus image fusion algorithm is completed by comparing the visual quality of fusion image. For Grayscale datasets, Flower and Temple
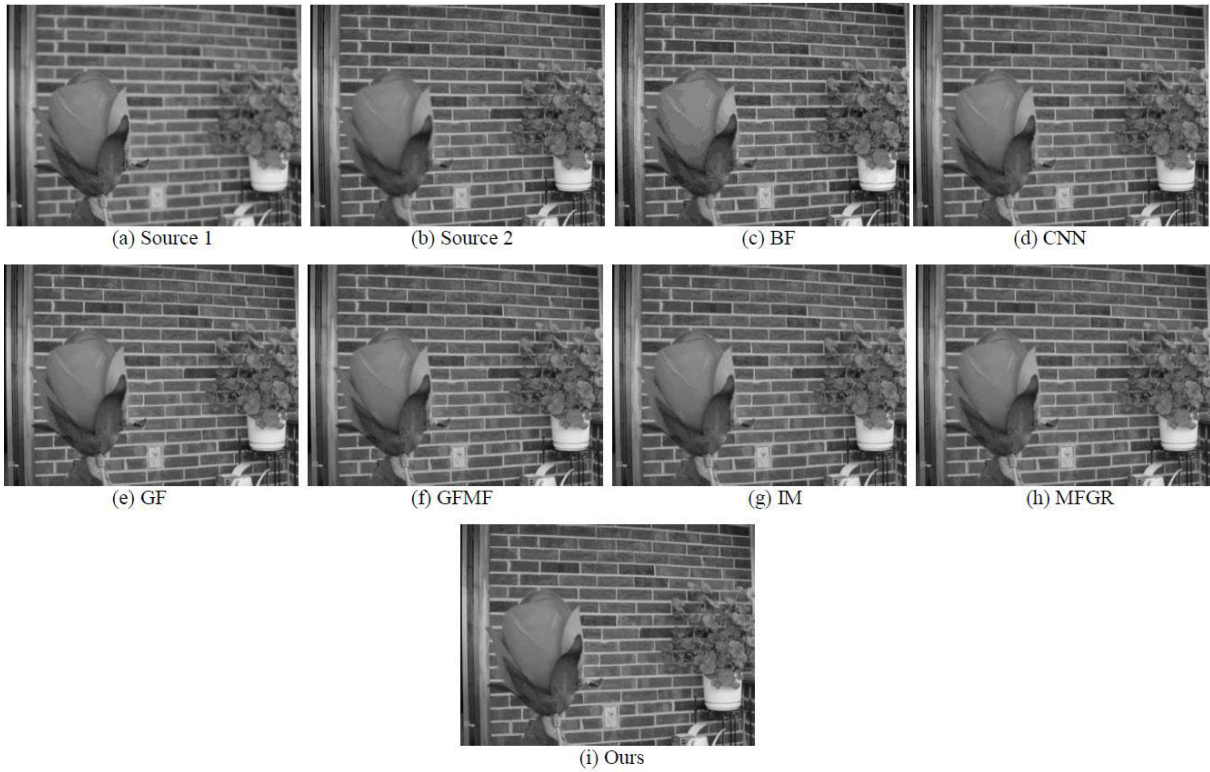
FIGURE 13. The Flower source image pair and their fused image obtained with different fusion methods.
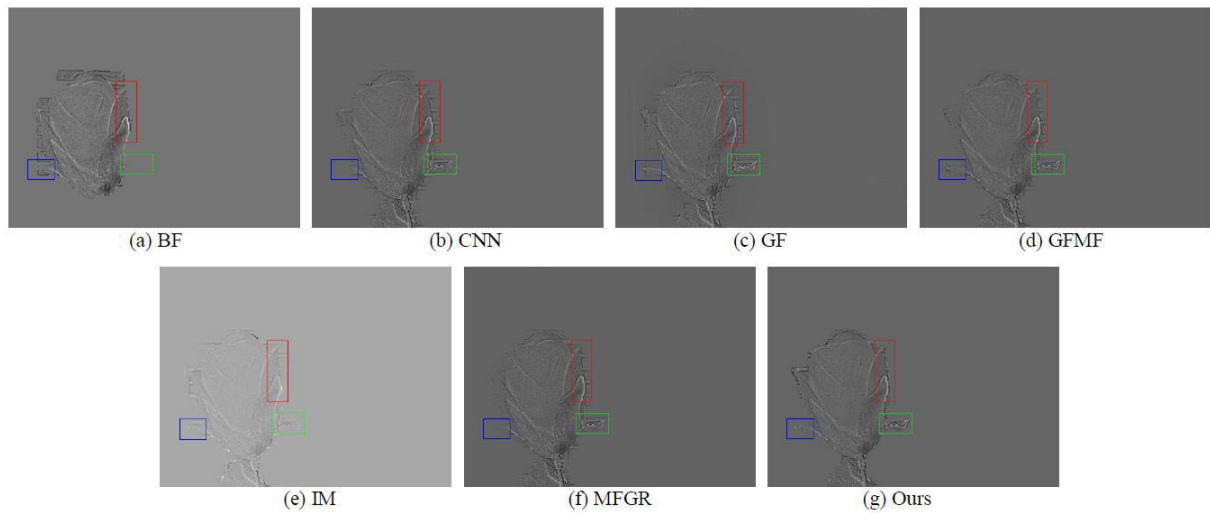
FIGURE 14. The normalized difference image between each fused image in Fig. 13 and the source image in Fig.13 (b).

source images in top row of Fig. 10 are chosen for visual comparison. For both 'flower' and 'temple' images, the left images are near focused, where the flower and the stone lion are in focus and clear in vision, whereas the wall and background temple are out of focus and blurred. The right images are far focused, and the situations for the flower and the background temple are contrary. The fusion results of Flower and Temple source images by different fusion methods are shown

in Fig. 13 and Fig.15, respectively. The fusion results in both Fig. 13 and Fig.15 show that all the algorithms could achieve the purpose of multi-focus image fusion. In order to show the fusion results more clearly, Fig. 14 and Fig. 16 show the difference image obtained by subtracting the second source image shown in Fig. 13(b) and Fig.15(b) from each fused image, and the values of each difference image is normalized to the range of 0 to 1. From Fig. 14(a)-(f), we can observe that
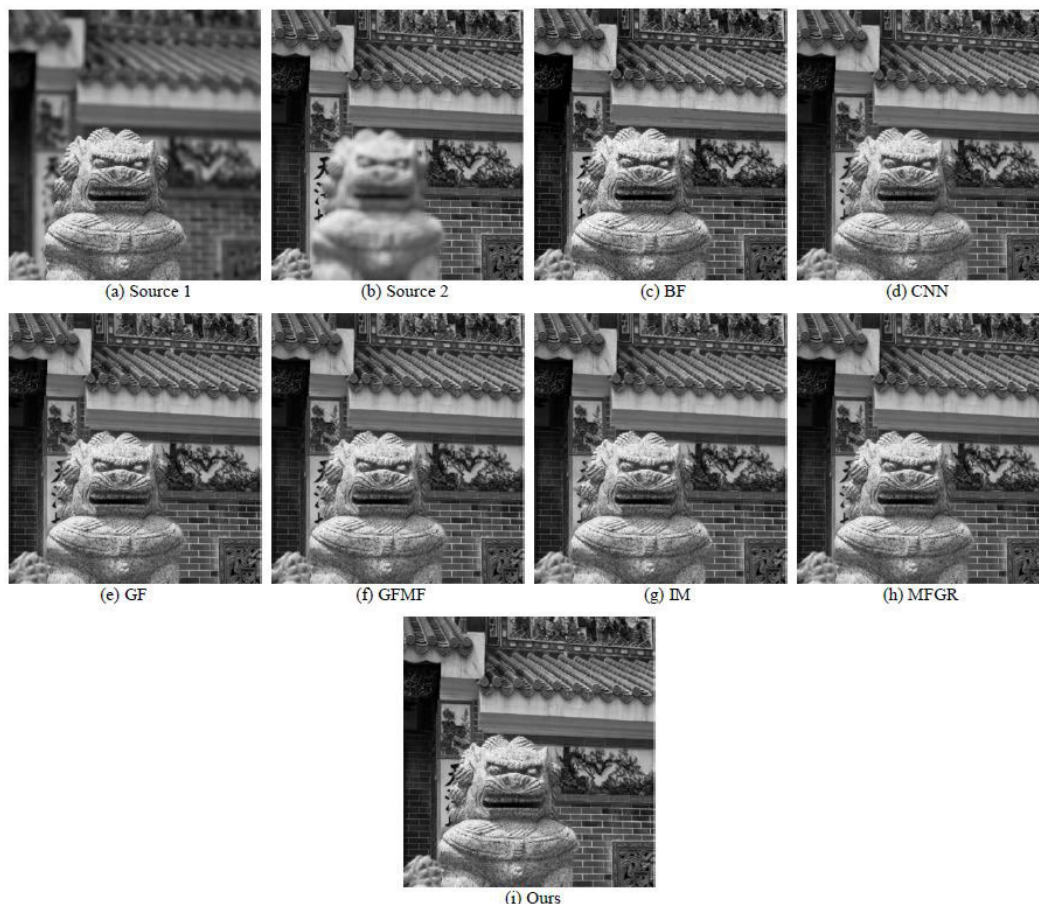
FIGURE 15. The Temple source image pair and their fused image obtained with different fusion methods.

the fused images of these compared methods produce significant undesirable block artifacts on the boundary between the flower and the background wall, and the fused image of BF loses some original image information in the lower part of the flower. As shown in Fig. 14(g), our method owns highest visual quality among other six methods. From Fig 16(a)-(g), it also can be seen that compared to other approaches, our algorithm can locate the boundary between the stone lion and the temple more accurate. Because of the lack of space, only the fused results of other gray images by our proposed method are presented in Fig. 17, which also provides high quality visual results. The object features from both source images are preserved clearly in the fused image and no reconstruction artifacts are produced.

We also conduct the proposed method on color image datasets shown in Fig. 11. For each pair images in Fig. 11, the left image near focused while the right image is far focus. Taking Fig. 11(a) as an example, the left image is near focused, where the statue is in focus and clear in vision, whereas the background wall is out of focus and blurred. The right image is far focused, and the situations for the statue and the background wall are contrary. Fig. 11(a) and Fig. 11(b) are also selected for visual comparison in detail. Fig. 18 and

Fig. 20 show the corresponding fusion results of proposed algorithm and other state-of-the-art approaches. As shown in Fig. 18 and Fig. 20, all fusion results of our method and other state-of-the-art approaches demonstrate high visual quality. To make better comparison, the normalized difference images obtained by subtracting the second source image shown in Fig. 18 (b) from each fused image is provided in Fig. 19. From Fig. 19(a)-(f), we can observe that the fused image of CNN, MFGR, GFMF, GF, and BF are overall high quality except that the regions around the statue are slightly blurred, but the fusion result of IM produce undesirable artifacts around the statue. The normalized difference image shown in Fig. 19(g) shows that our method owns highest visual quality in the boundary regions between the statue and the background wall among other six compared methods. Fig. 21 shows the normalized difference images obtained by subtracting the first source image shown in Fig. 20(a) from each fused image. From the area enclosed by a green rectangle in Fig. 21, we can note that a small focused region in Fig. 20(a) preserved by GF, IM, and MFGR, but their fusion quality around the boundary regions labeled by blue rectangle is poor, especially for IM. As shown in Fig. 21(a), the fusion results of BF misclassified on Koala, the CNN
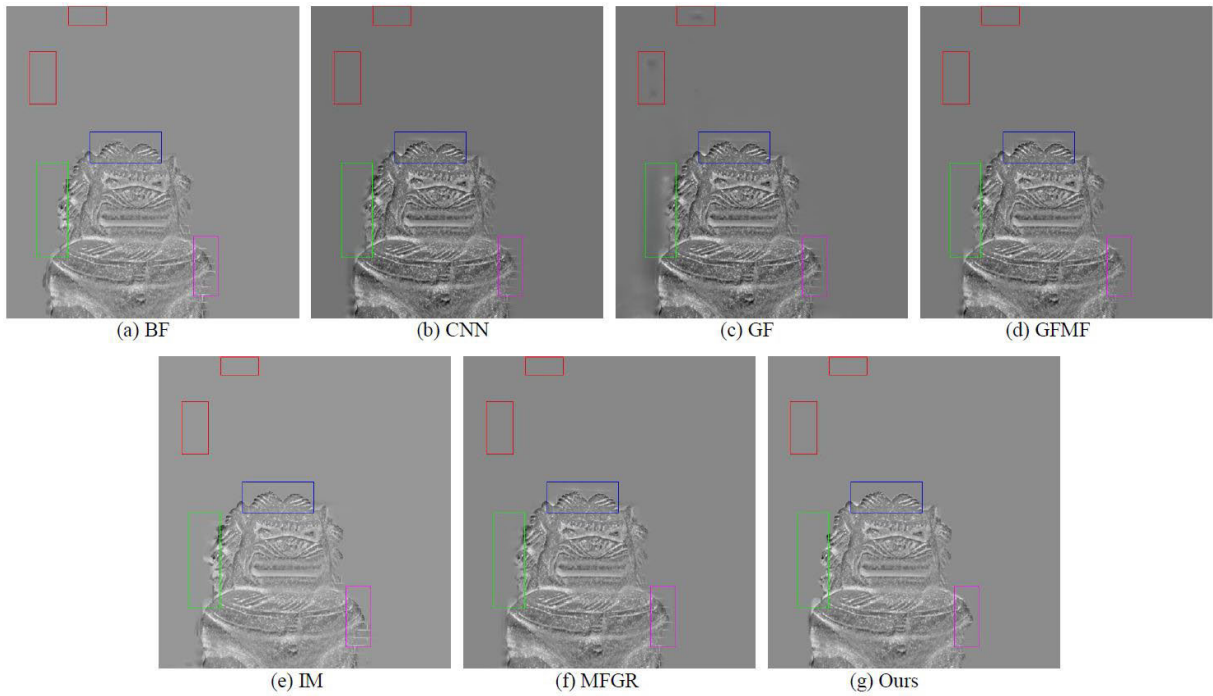
FIGURE 16. The normalized difference image between each fused image in Fig. 15 and the source image in Fig.15 (b).
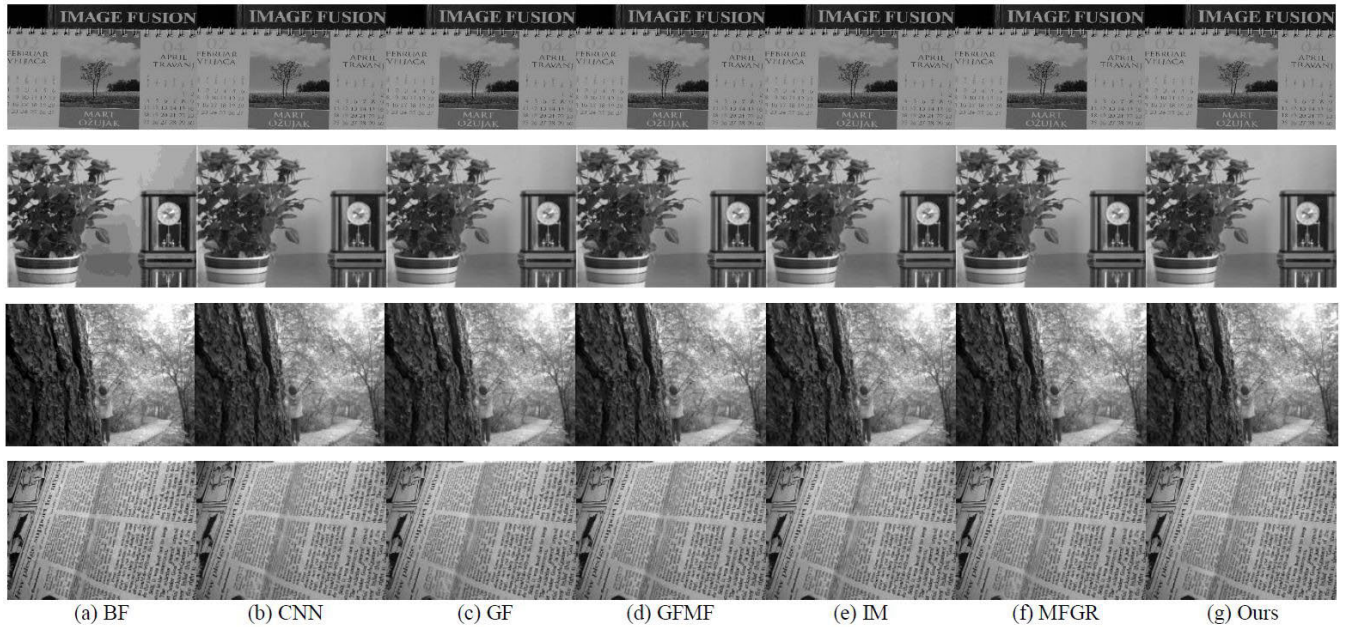
FIGURE 17. The fusion results of gray images from Grayscale dataset obtained with different fusion methods.

method also suffer from some artifacts in the blue region. The normalized difference image shown in Fig. 21(g) demonstrates that the fused image of our method owns highest visual quality among all these six methods. The fusion results of other test color images in Fig. 11 are shown in Fig. 22. All objects in scene are reserved clearly in the fused images.

In addition to the visual comparison, seven evaluation metrics, Q_{MI}, Q_{FMI}, $Q^{AB/F}$, Q_E, Q_Y, Q_W, and Q_{CB} are used to evaluate the fusion performance objectively. TABLE 2-8 show the Q_{MI}, Q_{FMI}, $Q^{AB/F}$, Q_E, Q_Y, Q_W, and Q_{CB} of the fused image obtained by different fusion algorithms on Gray dataset and Lyrto dataset, separately. The bold values in
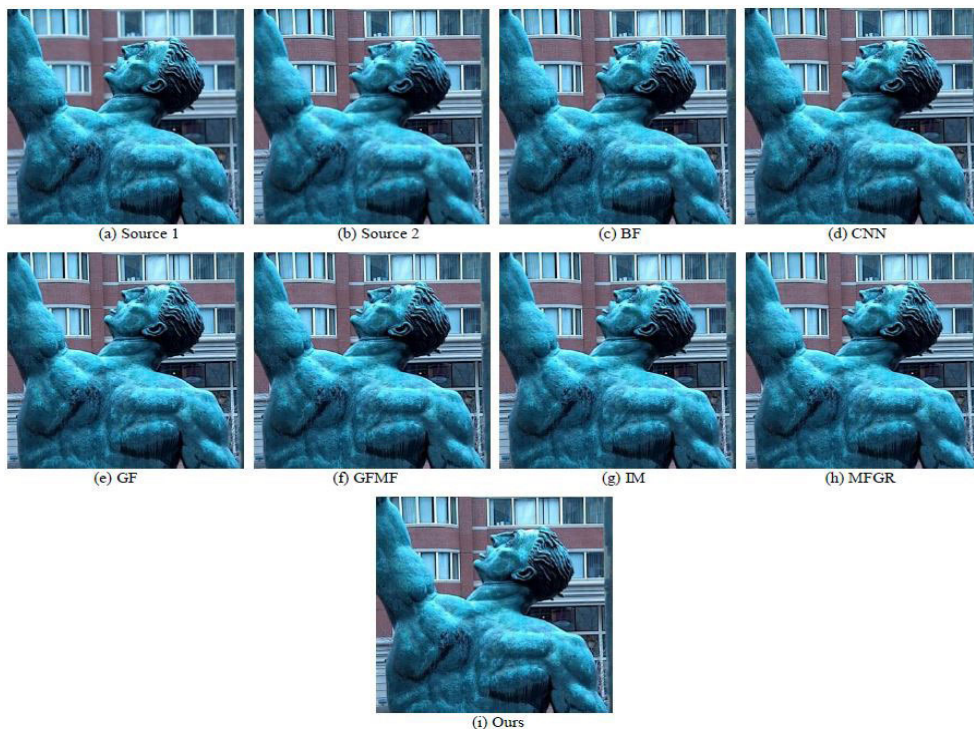
FIGURE 18. The statue source image pair and their fused image obtained with different fusion methods.
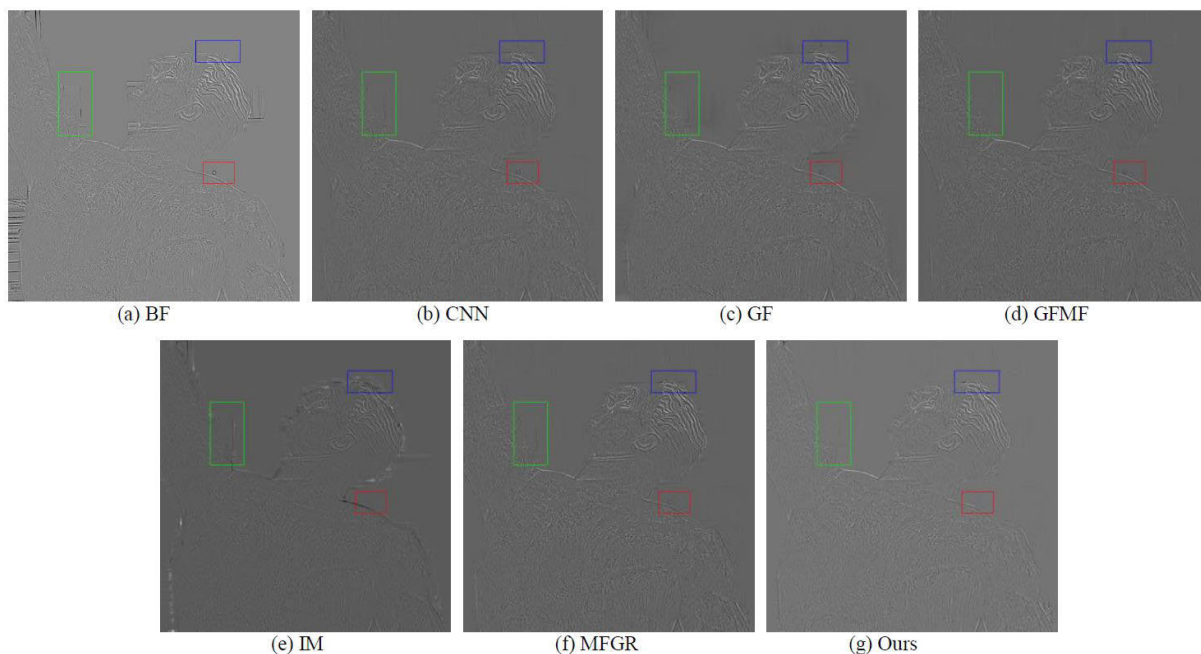
FIGURE 19. The normalized difference image between each fused image in Fig. 18 and the source image in Fig.18 (b).

these table indicate the maximum values in the corresponding rows, the underlined values in these table indicate the second largest values in the corresponding rows, and the greater value means the better performance of the corresponding algorithm. For the information based metrics Q_{MI} and Q_{FMI},

it can been seen that only the values of BF are better than our method in Girl, Flower, and Women. This means that our method can well preserve the original information of different sources image in most cases. In terms of the image feature-based metrics $Q^{AB/F}$, and Q_E, the proposed method

FIGURE 20. The Sydney source image pair and their fused image obtained with different fusion methods.
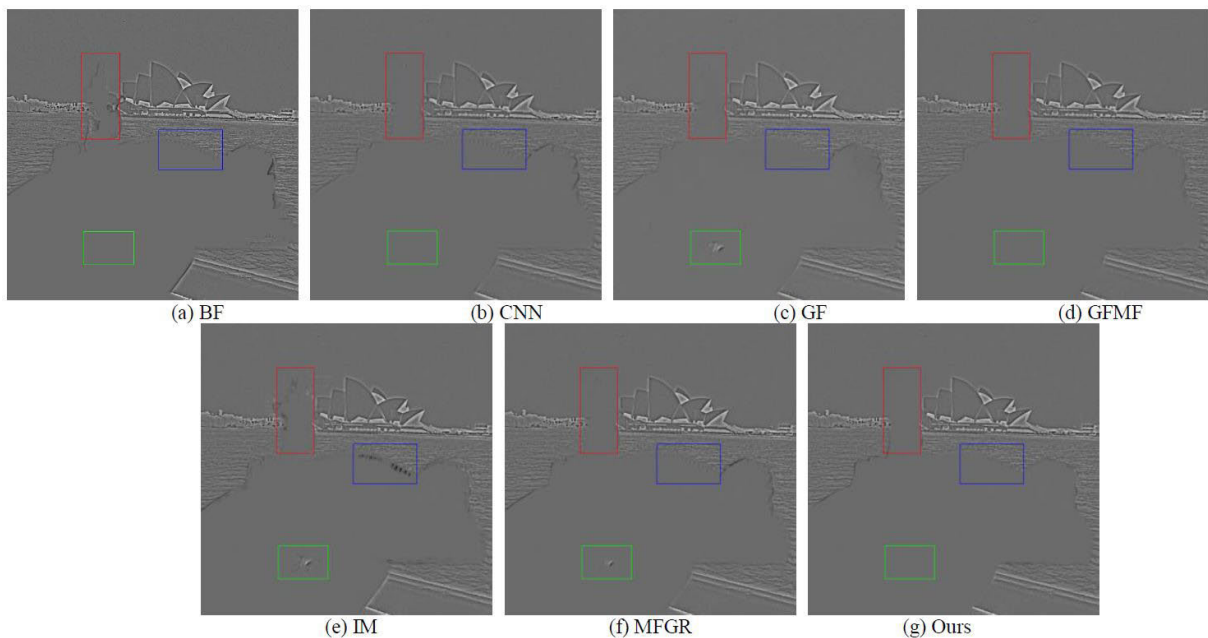
FIGURE 21. The normalized difference image between each fused image in Fig. 20 and the source image in Fig.20 (a).

also gives a largest value in most of the examples. It should be noticed that the proposed method, GFMF, and GF give a stable and better performance of Q_E, this is because the guided filtering based methods call inherit global salient edges effectively. As shown in TABLE 6 and TABLE 7, the proposed method outperform BF, CNN, IM and MFGR obviously,

FIGURE 22. The fusion results of color images from Lytro dataset obtained with different fusion methods.

TABLE 2. Quantitative assessments (Q_{MI}) of different multi-focus image fusion methods on different source images.

Source	Methods						
Images	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	**0.8863**	0.8415	0.7337	0.8525	0.8499	0.8502	0.8750
Temple	0.5847	0.5269	0.3355	0.5508	0.5527	0.5416	**0.5970**
Calendar	1.3091	1.2596	1.2371	1.2826	1.3006	1.2782	**1.3164**
Corner	1.1436	1.1390	1.0838	1.1391	1.1179	1.1318	**1.1493**
Girl1	**0.8200**	0.7915	0.6517	0.8046	0.7899	0.7860	0.8146
Newspaper	0.8301	0.7637	0.5694	0.7724	0.7640	0.7585	**0.8310**
Statue	0.8379	0.7955	0.7445	0.8159	0.7736	0.8124	**0.8497**
Sydney	0.8397	0.7991	0.7077	0.8234	0.7936	0.8110	**0.8598**
Woman	**0.8172**	0.7827	0.6982	0.7990	0.7589	0.7889	0.8150
Wine	0.7539	0.7301	0.6510	0.7432	0.7331	0.7385	**0.7564**
Zoo	0.5911	0.5513	0.4553	0.5660	0.5599	0.5598	**0.5947**
Horse	1.1035	1.0883	1.0822	1.0975	1.0670	1.0989	**1.1170**
Average	0.8764	0.8391	0.7458	0.8539	0.8384	0.8463	**0.8813**

and competitive with the GF and GFMF of the structural similarity based metrics Q_Y and Q_W, which means the fused image of our method, GF and GFMF can well preserve the complementary information of different source images. As shown in TABLE 8, our method also gives a stable and higher values for human perception-based metrics Q_{CB} in most cases, which indicates that the fused images of

proposed method have better effects in visual perception without introducing distortions. Although the proposed algorithm is not the most efficient multi-focus image fusion approach, the experiment results show that our method also is competitive with the state-of-the-art algorithms in terms of both subjective visual perception and objective evaluation metrics.

TABLE 3. Quantitative assessments (Q_{FMI}) of different multi-focus image fusion methods on different source images.

Source Images	Methods						
	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.6141	0.6135	0.6089	0.6147	0.6112	0.6147	**0.6170**
Temple	0.5754	0.5752	0.5675	0.5765	0.5742	0.5759	**0.5789**
Calendar	0.6221	0.6199	0.6058	0.6219	0.6222	0.6208	**0.6254**
Corner	0.6092	0.5959	0.5873	0.6060	0.5791	0.6035	**0.6100**
Girl1	0.6054	0.6053	0.5974	0.6057	0.6006	0.6036	**0.6059**
Newspaper	0.5823	0.5775	0.5704	0.5791	0.5744	0.5772	**0.5852**
Statue	0.6821	0.6803	0.6753	0.6818	0.6674	0.6810	**0.6835**
Sydney	0.5806	0.5797	0.5662	0.5814	0.5682	0.5790	**0.5849**
Woman	**0.6106**	0.6068	0.5795	0.6091	0.5839	0.6024	0.6098
Wine	0.5893	0.5879	0.5811	0.5886	0.5856	0.5885	**0.5897**
Zoo	0.5921	0.5909	0.5832	0.5918	0.5869	0.5901	**0.5924**
Horse	0.6655	0.6639	0.6598	0.6649	0.6529	0.6628	**0.6661**
Average	0.6107	0.6081	0.5985	0.6101	0.6006	0.6083	**0.6124**

TABLE 4. Quantitative assessments ($Q^{AB/F}$) of different multi-focus image fusion methods on different source images.

Source Images	Methods						
	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.6834	0.6880	0.6867	0.6883	0.6863	0.6886	**0.6889**
Temple	0.7179	0.7168	0.7144	0.7185	0.7193	0.7192	**0.7217**
Calendar	0.6879	0.6835	0.6769	0.6849	0.6882	0.6848	**0.6886**
Corner	**0.6567**	0.6461	0.6464	0.6546	0.6420	0.6544	0.6551
Girl1	0.6860	0.6864	0.6830	**0.6868**	0.6829	0.6865	0.6865
Newspaper	0.6464	0.6434	0.6408	0.6471	0.6425	0.6439	**0.6537**
Statue	0.7485	0.7568	0.7551	0.7567	0.7459	**0.7573**	0.7570
Sydney	0.7382	0.7397	0.7364	0.7415	0.7317	0.7423	**0.7456**
Woman	0.6816	0.6845	0.6740	0.6853	0.6686	0.6834	**0.6884**
Wine	0.7636	0.9642	0.9655	0.9656	0.9606	0.9651	**0.9657**
Zoo	0.7765	0.7754	0.7742	0.7760	0.7754	0.7770	**0.7779**
Horse	0.7534	0.7530	0.7536	0.7537	0.7446	0.7544	**0.7553**
Average	0.7117	0.7282	0.7256	0.7299	0.7240	0.7297	**0.7320**

TABLE 5. Quantitative assessments (Q_E) of different multi-focus image fusion methods on different source images.

Source Images	Methods						
	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.7515	0.7647	**0.7676**	0.7673	0.7574	0.7666	0.7661
Temple	0.8784	0.8861	0.8909	0.8967	0.8727	0.8916	**0.8990**
Calendar	0.8437	0.8404	0.8459	0.8475	0.8452	0.8480	**0.8489**
Corner	0.8260	0.8265	**0.8266**	0.8265	0.8129	0.8265	0.8263
Girl1	0.7557	0.7566	**0.7617**	0.7599	0.7413	0.7613	0.7594
Newspaper	0.6134	0.6147	0.6172	**0.6210**	0.6164	0.6170	0.6191
Statue	0.7906	0.8276	0.8279	**0.8288**	0.8057	0.8275	0.8261
Sydney	0.9009	0.9179	0.9263	0.9268	0.9155	0.9241	**0.9298**
Woman	0.8360	0.8432	0.8460	**0.8474**	0.8356	0.8461	0.8470
Wine	0.9440	0.9466	0.9493	0.9495	0.9385	0.9485	**0.9499**
Zoo	0.9403	0.9382	**0.9433**	0.9427	0.9386	0.9422	0.9424
Horse	0.9084	0.9107	**0.9119**	0.9092	0.8970	0.9114	0.9087
Average	0.8324	0.8394	0.8429	**0.8436**	0.8314	0.8426	**0.8436**

To evaluate the computational efficiency, the running time for processing 12 sources image sets by each algorithm is counted and listed in TABLE 9. In terms of the average time, it can be noticed that the CNN method suffered

TABLE 6. Quantitative assessments (Q_Y) of different multi-focus image fusion methods on different source images.

Source	Methods						
Images	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.9786	0.9802	0.9798	0.9807	0.9784	<u>0.9808</u>	**0.9811**
Temple	0.9926	0.9928	0.9883	<u>0.9944</u>	0.9923	0.9943	**0.9956**
Calendar	0.9900	0.9888	0.9830	0.9900	<u>0.9908</u>	0.9896	**0.9913**
Corner	0.9779	0.9728	0.9722	**0.9795**	0.9625	<u>0.9787</u>	0.9760
Girl1	0.9921	<u>0.9922</u>	0.9885	**0.9926**	0.9876	0.9919	0.9920
Newspaper	<u>0.9881</u>	0.9866	0.9824	0.9880	0.9834	0.9863	**0.9917**
Statue	0.9767	0.9780	0.9777	0.9780	0.9728	**0.9782**	<u>0.9781</u>
Sydney	0.9891	0.9896	0.9865	<u>0.9914</u>	0.9793	0.9910	**0.9936**
Woman	<u>0.9868</u>	0.9858	0.9730	**0.9870**	0.9693	0.9842	0.9867
Wine	0.9652	0.9642	0.9655	<u>0.9656</u>	0.9606	0.9651	**0.9657**
Zoo	**0.9961**	0.9953	0.9933	0.9959	0.9942	0.9955	<u>0.9960</u>
Horse	<u>0.9834</u>	0.9830	0.9814	0.9832	0.9724	0.9826	**0.9836**
Average	**0.9872**	0.9841	0.9810	0.9855	0.9786	0.9849	<u>0.9860</u>

TABLE 7. Quantitative assessments (Q_W) of different multi-focus image fusion methods on different source images.

Source	Methods						
Images	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.8484	0.8559	**0.8575**	<u>0.8570</u>	0.8519	0.8567	0.8564
Temple	0.9304	0.9340	0.9366	<u>0.9389</u>	0.9281	0.9371	**0.9397**
Calendar	0.9109	0.9097	0.9121	<u>0.9128</u>	0.9126	**0.9131**	0.9131
Corner	**0.8916**	0.8913	<u>0.8914</u>	0.8913	0.8886	<u>0.8914</u>	0.8912
Girl1	0.8619	0.8622	**0.8651**	0.8640	0.8520	<u>0.8650</u>	0.8641
Newspaper	0.8419	0.8431	<u>0.8444</u>	**0.8447**	0.8433	0.8440	0.8431
Statue	0.8867	0.9077	<u>0.9078</u>	**0.9080**	0.8963	0.9077	0.9064
Sydney	0.9401	0.9484	<u>0.9523</u>	0.9522	0.9470	0.9517	**0.9541**
Woman	0.8944	0.8985	<u>0.8996</u>	**0.8999**	0.8938	0.8993	0.8992
Wine	0.9633	0.9642	0.9655	<u>0.9656</u>	0.9606	0.9651	**0.9657**
Zoo	0.9633	0.9616	<u>0.9641</u>	0.9636	0.9619	0.9637	**0.9644**
Horse	0.9509	0.9518	**0.9524**	0.9512	0.9458	<u>0.9522</u>	0.9508
Average	0.9070	0.9107	**0.9124**	**0.9124**	0.9068	<u>0.9123</u>	**0.9124**

TABLE 8. Quantitative assessments (Q_{CB}) of different multi-focus image fusion methods on different source images.

Source	Methods						
Images	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	0.8345	0.8363	0.8314	**0.8367**	0.8319	<u>0.8366</u>	0.8359
Temple	0.8173	0.8159	0.7997	<u>0.8188</u>	0.8137	0.8177	**0.8200**
Calendar	0.8061	0.8057	0.7974	<u>0.8081</u>	0.8056	0.8065	**0.8092**
Corner	<u>0.8024</u>	0.7891	0.7853	**0.8029**	0.7794	0.8020	0.7942
Girl1	0.7957	0.7952	0.7891	<u>0.7959</u>	0.7898	0.7951	**0.7962**
Newspaper	**0.7475**	0.7422	0.7331	0.7445	0.7418	0.7412	<u>0.7446</u>
Statue	0.8603	0.8627	0.8565	**0.8637**	0.8442	0.8627	<u>0.8634</u>
Sydney	0.7853	0.7887	0.7745	<u>0.7926</u>	0.7756	0.7900	**0.7935**
Woman	0.8138	<u>0.8147</u>	0.7961	<u>0.8174</u>	0.7976	0.8129	0.8117
Wine	0.8536	0.8541	0.8478	<u>0.8558</u>	0.8504	0.8548	**0.8563**
Zoo	<u>0.8654</u>	0.8635	0.8561	<u>0.8654</u>	0.8628	0.8648	**0.8660**
Horse	0.7853	0.7887	0.7745	<u>0.7926</u>	0.7756	0.7900	**0.7935**
Average	0.8196	0.8184	0.8097	**0.8212**	0.8105	0.8197	<u>0.8203</u>

from the highest computation cost with 183.50 seconds. The GF method provides higher efficient than other methods with 0.40 seconds. The average time of our method is 3.02 seconds, it is also acceptable.

TABLE 9. Comparison of the computational costs (second).

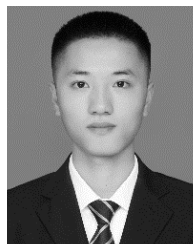| Source Images	Methods						
	BF	CNN	GF	GFMF	IM	MFGR	Ours
Flower	1.79	171.10	0.30	3.71	2.26	2.34	3.01
Temple	1.83	203.58	0.37	4.08	4.82	2.86	2.69
Calendar	1.18	57.57	0.10	3.84	1.94	1.06	2.61
Corner	1.19	83.74	0.07	1.75	1.14	1.38	2.41
Girl1	2.31	263.37	0.44	3.05	4.47	3.32	2.79
Newspaper	1.35	82.82	0.13	2.66	0.91	1.30	2.49
Statue	2.03	190.95	0.44	2.91	4.69	2.94	3.38
Sydney	1.97	244.21	0.65	2.94	5.93	3.01	3.44
Woman	2.16	233.98	0.92	4.11	6.10	3.07	3.32
Wine	1.90	229.78	0.46	3.44	5.64	3.03	3.34
Zoo	2.04	244.33	0.44	2.64	4.90	3.07	3.31
Horse	1.96	197.01	0.45	2.84	5.88	2.98	3.50
Average	1.81	183.50	0.40	3.16	4.06	2.53	3.02

IV. CONCLUSION

This paper proposes a new multi-focus image fusion method to detect the focus regions in multiple source images using point filter and Gaussian filter, and presents a superpixel-based consistency by introducing the SLIC algorithm into consistency verification to improve the fusion performance. Compared with several traditional multi-focus fusion methods, this proposed method achieves good results both subjectively and objectively.

REFERENCES

[1] L. Chen, J. Li, and C. L. Chen, "Regional multifocus image fusion using sparse representation," *Opt. Express*, vol. 21, no. 4, pp. 5182–5197, Feb. 2013.

[2] Y. Liu, S. Liu, and Z. Wang, "Multi-focus image fusion with dense SIFT," *Inf. Fusion*, vol. 23, pp. 139–155, May 2015.

[3] Z. Zhang and R. S. Blum, "A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application," *Proc. IEEE*, vol. 87, no. 8, pp. 1315–1326, Aug. 1999.

[4] Z. Liu, D. S. Forsyth, and M. S. Safizadeh, "A data-fusion scheme for quantitative image analysis by using locally weighted regression and Dempster–Shafer theory," *IEEE Trans. Instrum. Meas.*, vol. 57, no. 11, pp. 2554–2560, Nov. 2008.

[5] V. S. Petrovic and C. S. Xydeas, "Gradient-based multiresolution image fusion," *IEEE Trans. Image Process.*, vol. 13, no. 2, pp. 228–237, Feb. 2004.

[6] G. Pajares and J. Manuel de la Cruz, "A wavelet-based image fusion tutorial," *Pattern Recognit.*, vol. 37, no. 9, pp. 1855–1872, Sep. 2004.

[7] J. Duan, L. Chen, and C. L. P. Chen, "Multifocus image fusion with enhanced linear spectral clustering and fast depth map estimation," *Neurocomputing*, vol. 318, pp. 43–54, Nov. 2018.

[8] Z. Liu, K. Tsukada, K. Hanasaki, Y. K. Ho, and Y. P. Dai, "Image fusion by using steerable pyramid," *Pattern Recognit. Lett.*, vol. 22, no. 9, pp. 929–939, Jul. 2001.

[9] J. Tian and L. Chen, "Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure," *Signal Process.*, vol. 92, no. 9, pp. 2137–2146, Sep. 2012.

[10] Q. Zhang and B.-L. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Process.*, vol. 89, no. 7, pp. 1334–1346, Jul. 2009.

[11] S. Li and B. Yang, "Multifocus image fusion by combining curvelet and wavelet transform," *Pattern Recognit. Lett.*, vol. 29, no. 9, pp. 1295–1301, Jul. 2008.

[12] X. Jin, R. Nie, D. Zhou, Q. Wang, and K. He, "Multifocus color image fusion based on NSST and PCNN," *J. Sensors*, vol. 2016, pp. 1–12, 2016.

[13] M. Nejati, S. Samavi, and S. Shirani, "Multi-focus image fusion using dictionary-based sparse representation," *Inf. Fusion*, vol. 25, pp. 72–84, Sep. 2015.

[14] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 884–892, Apr. 2010.

[15] T. Wan, C. Zhu, and Z. Qin, "Multifocus image fusion based on robust principal component analysis," *Pattern Recognit. Lett.*, vol. 34, no. 9, pp. 1001–1008, Jul. 2013.

[16] S. Liu, J. Wang, Y. Lu, H. Li, J. Zhao, and Z. Zhu, "Multi-focus image fusion based on adaptive dual-channel spiking cortical model in non-subsampled shearlet domain," *IEEE Access*, vol. 7, pp. 56367–56388, 2019.

[17] S. Aymaz and C. Köse, "A novel image decomposition-based hybrid technique with super-resolution method for multi-focus image fusion," *Inf. Fusion*, vol. 45, pp. 113–127, Jan. 2019.

[18] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Commun.*, vol. 43, no. 12, pp. 2959–2965, Dec. 1995.

[19] S. Li, J. T. Kwok, and Y. Wang, "Combination of images with diverse focuses using the spatial frequency," *Inf. Fusion*, vol. 2, no. 3, pp. 169–176, Sep. 2001.

[20] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image Vis. Comput.*, vol. 26, no. 7, pp. 971–979, Jul. 2008.

[21] B. Yang and L. Guo, "Superpixel based fusion and demosaicing for multi-focus bayer images," *Optik*, vol. 126, no. 23, pp. 4460–4468, Dec. 2015.

[22] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 493–500, Mar. 2007.

[23] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 824–831, Aug. 1994.

[24] V. Aslantas and R. Kurban, "A comparison of criterion functions for fusion of multi-focus noisy images," *Opt. Commun.*, vol. 282, no. 16, pp. 3231–3242, Aug. 2009.

[25] Z. Zhou, S. Li, and B. Wang, "Multi-scale weighted gradient-based fusion for multi-focus images," *Inf. Fusion*, vol. 20, pp. 60–72, Nov. 2014.

[26] 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.

[27] X. Qiu, M. Li, L. Zhang, and X. Yuan, "Guided filter-based multi-focus image fusion through focus region detection," *Signal Process., Image Commun.*, vol. 72, pp. 35–46, Mar. 2019.

[28] Y. Liu, X. Chen, H. Peng, and Z. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36, pp. 191–207, Jul. 2017.

[29] C. Du, S. Gao, Y. Liu, and B. Gao, "Multi-focus image fusion using deep support value convolutional neural network," *Optik*, vol. 176, pp. 567–578, Jan. 2019.

[30] H. Li, B. S. Manjunath, and S. K. Mitra, "Multisensor image fusion using the wavelet transform," *Graph. Models Image Process.*, vol. 57, no. 3, pp. 235–245, May 1995.

[31] B. Yang and S. Li, "Multi-focus image fusion based on spatial frequency and morphological operators," *Chin. Opt. Lett.*, vol. 5, no. 8, pp. 452–453, Aug. 2007.

[32] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2281, Nov. 2012.

[33] P. Jiang, H. Ling, J. Yu, and J. Peng, "Salient region detection by UFO: Uniqueness, focusness and objectness," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 1976–1983.

[34] H. Zhao, Z. Shang, Y. Y. Tang, and B. Fang, "Multi-focus image fusion based on the neighbor distance," *Pattern Recognit.*, vol. 46, no. 3, pp. 1002–1011, Mar. 2013.

[35] *Lytro Image Dataset*. Accessed: May 8, 2019. [Online]. Available: https://mansournejati.ece.iut.ac.ir/content/lytro-multi-focus-dataset

[36] M. Hossny, S. Nahavandi, and D. Creighton, "Comments on 'Information measure for performance of image fusion,'" *Electron. Lett.*, vol. 44, no. 18, pp. 1066–1067, Feb. 2008.

[37] M. Haghighat and M. A. Razian, "Fast-FMI: Non-reference image fusion metric," in *Proc. IEEE 8th Int. Conf. Appl. Inf. Commun. Technol. (AICT)*, Oct. 2014, pp. 1–3.

[38] C. S. Xydeas and V. Petrović, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, 2000.

[39] G. Piella and H. Heijmans, "A new quality metric for image fusion," in *Proc. Int. Conf. Image Process.*, Barcelona, Spain, Sep. 2003, pp. 173–176.

[40] C. Yang, J.-Q. Zhang, X.-R. Wang, and X. Liu, "A novel similarity based quality metric for image fusion," *Inf. Fusion*, vol. 9, no. 2, pp. 156–160, Apr. 2008.

[41] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, Sep. 2009.

[42] J. Ma, Z. Zhou, B. Wang, and M. Dong, "Multi-focus image fusion based on multi-scale focus measures and generalized random walk," in *Proc. 36th Chin. Control Conf. (CCC)*, Dalian, China, Jul. 2017, pp. 5464–5468.

[43] S. Li, X. Kang, and J. Hu, "Image matting for fusion of multi-focus images in dynamic scenes," *Inf. Fusion*, vol. 14, no. 2, pp. 147–162, Apr. 2013.

[44] Y. Zhang, X. Bai, and T. Wang, "Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure," *Inf. Fusion*, vol. 35, pp. 81–101, May 2017.

[45] *The Source Code of CNN Method*. Accessed: Feb. 23, 2019. [Online]. Available: http://www.escience.cn/people/liuyu1/Codes.html

[46] *The Source Code of MFGR Method*. Accessed: Nov. 1, 2019. [Online]. Available: https://github.com/JinleiMa/Multi-focus-Image-Fusion-with-Multi-scale-Focus-Measures

[47] *The Source Code of GFMF Method*. Accessed: Nov. 1, 2019. [Online]. Available: https://github.com/bitname/Multi-focus-image-fusion-GFDF

[48] *The Source Code of BF Method*. Accessed: Nov. 1, 2019. [Online]. Available: https://github.com/uzeful/Boundary-Finding-based-Multi-focus-Image-Fusion

[49] *The Source Code of GF Method and IM Method*. Accessed: Jun. 29, 2019. [Online]. Available: http://xudongkang.weebly.com

QIANG CHEN received the B.S. degree in automation from the College of Information Engineering, Xiangtan University, Xiangtan, China, in 2018. He is currently pursuing the M.S. degree in control engineering with the University of South China, Hengyang, China. His current research interests include image fusion and machine learning.
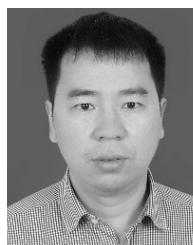
BIN YANG (Member, IEEE) was born in Henan, China, in 1980. He received the B.S. degree in automation from the Zhengzhou University of Light Industry, in 2005, and the Ph.D. degree in control science and engineering from Hunan University, in 2010. He is currently a Professor with the University of South China. He is the author of more than 30 journal articles. His current research interests include image processing, pattern recognition, and deep learning.

YUEHUA LI received the B.S. degree in electronic engineering from the East China University of Technology, in 1997, and the master's degree in signal and information processing from Nanchang University, in 2007. She is currently a University Lecturer with the Department of Electrical Engineering, University of South China. Her current research interests include signal and information processing, image processing, and artificial intelligent control systems.
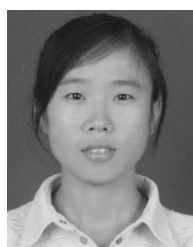
LIHUI PANG received the B.S. degree from the Henan University of Science and Technology, in 2008, and the Ph.D. degree from the University of Electronic Science and Technology of China, in 2015. She is currently a Lecturer with the University of South China. Her current research interests include blind signal processing, pattern recognition, and deep learning.

• • •