



Multi-focus Image Fusion with Cooperative Image Multiscale Decomposition

Yueqi Tan and Bin Yang^(✉)

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

Abstract. Multi-focus image fusion plays an important role in the field of image processing for its ability in solving the depth-of-focus limitation problem in optical lens imaging by fusing a series of partially focused images of the same scene. The improvements on various fusion methods focus on the image decomposition methods and the fusion strategies. However, most decompositions are separately conducted on each image, which fails to sufficiently consider the nature of multiple images in fusion tasks, and insufficiently explores the consistent and inconsistent features of two source images simultaneously. This paper proposes a new cooperative image multiscale decomposition (CIMD) based on the mutually guided filter (MGF). With CIMD, two source multi-focus images are simultaneously decomposed into base layers and detailed layers through the iterative operation of MGF cooperatively. A saliency detection based on a mean-guide combination filter is adopted to guide the fusion of detailed layers and a spatial frequency-based fusion strategy is used to fuse the luminance and contour features in the base layers. The experiments are carried on 28 pairs of publicly available multi-focus images. The fusion results are compared with 7 state-of-the-art multi-focus image fusion methods. Experimental results show that the proposed method has the better visual quality and objective assessment.

Keywords: Multi-focus image fusion · Depth-of-focus · Mutually-guided filter · Cooperative image multiscale decomposition · Focus region detection

1 Introduction

Optical imaging has the limitation of the depth-of-focus, which makes it difficult to obtain an all-in-focus image with conventional digital cameras. The multi-focus image fusion technology which obtains an all-in-focus image by merging multiple optical images has an important practical significance. The fused images are widely used in the fields of human vision, computer vision, artificial intelligence, and so on.

According to the feature domains where the source images are fused, image fusion methods are roughly divided into two main categories, namely, spatial domain methods and transform domain methods. Spatial domain methods usually investigate the saliency information of pixels or regions to establish the fusion. Some of them are based on total variation, gradient optimization [4]. This type of methods is efficient. However, there may be artifacts and distortions in the fused image. Transform domain approaches are

consist of three steps, namely, decomposition, coefficient fusion, and reconstruction [1]. Pyramid transform [5][6], wavelet transform, and contourlet transform are commonly used to decompose the source images into different feature coefficients. These transform methods offer not only higher efficiencies but also offer greater flexibility to the fusion strategy design. The sparse representation [7] also attracts numerous researchers in image fusion applications due to its effective capability in underlying information extraction from an image. In recent years, deep learning has also been introduced to achieve multi-focus image fusion because of its potential in deep feature extraction [9].

Existing multiscale transforms have made various improvements in image fusion rules and image multiscale decomposition for resolving the problem of saliency detection and improving the fusion effect. Whereas the traditional multiscale decompositions do not sufficiently separate the consistent structure and inconsistent structure in the two images to be fused that have gradient amplitude differences. In this paper, we present a new fusion method based on the proposed cooperative image multiscale decomposition (CIMD). The CIMD is used to separate the consistent structure and the inconsistent structure in two images by the CIMD. With CIMD, two source images are decomposed into detailed layers and base layers cooperatively. The detailed layers which contain the inconsistent gradient structure of two source images are very helpful in adjusting the local clarity of two source images. Therefore, a saliency detection-based fusion rules are used to fusion the detailed layers that contain the inconsistent structure, and spatial frequency-based fusion rules are used to integrate the base layers. Finally, the fused base layers and detailed layers are merged together to obtain the final fused image. 28 pairs of publicly available multi-focus images are used to test the performances of the proposed method. The experimental results demonstrate that the proposed image fusion method demonstrates versatility across multiple types of multi-focus images.

The rest of this paper is organized as follows. we construct the CIMD-based MGF in Sect. 2. Sect. 3 describes the proposed image fusion algorithm in detail. The experiment results and discussions are presented in Sect. 4, and conclusions are given in Sect. 5.

2 Cooperative Image Multiscale Decomposition Based MGF

2.1 Mutually Guided Filter

Shen et al. [3] divided the mutual-structures of two input images into three types, including mutual structure, flat structure, and inconsistent structure. The mutual structure describes the part with both strong gradient structures, while the flat structure references the part of the consistent weak gradient structure. The two structures are collectively referred to as the consistent structure in this paper. The inconsistent structure represents the structure with divergence gradient amplitudes in two images at the same pixel position. The inconsistent structure would be vital for multi-focus image fusion applications. However traditional guide filter ignores the structural inconsistency when we need to filter two content correlative images simultaneously. To solve these problems, a mutually guided filter is proposed for the purpose to preserve the mutual structure, prevent misleading from inconsistent structure, and smooth flat regions between two input images [2]. Let T_0 and R_0 denote two input images that should be filtered simultaneously. The

filtered outputs T and R from the MGF are obtained by:

$$\arg \min_{T,R} \alpha \varphi[T, R, \varepsilon] + \|T - T_0\|_2^2 + \|R - R_0\|_2^2 \quad (1)$$

where

$$\varphi[T, R, \varepsilon] = \sum_{i \in \Omega} \sum_{d \in \{h,v\}} \left(\frac{|\nabla_d T_i|}{\max(|\nabla_d R_i|, \varepsilon)} + \frac{|\nabla_d R_i|}{\max(|\nabla_d T_i|, \varepsilon)} \right) \quad (2)$$

represents the mutual structure discrimination indicated by the gradient ratio between two images. The symbol Ω represents the pixel sets of the source images. T_i and R_i indicate the pixel values at i position. The symbol ε is a small positive constant to avoid division by zero. ∇_h and ∇_v are the horizontal and vertical derivative operators respectively. The first term in Eq. (1) is the regularization term which is the consistent structure restraint aiming to reduce the gradient discrepancy between two images to be filtered. The second and third terms are the fidelity term. The parameter α is regularization parameter given by the users.

From Eq. (1), we can easily deduce that the mutual structure and flat structure from two filter outputs toward identical. Figure 1 (a) shows two images obtained with different focus settings. The upper image shows rich texture in the left area while the bottom one focuses on the right area. Figure 1 (b), (c), and (d) illustrate filters employed by the mean filter, the guide filter [12], and MGF, respectively. All filter results are blurred in different various degrees. However, only the two outputs of MGF still maintain the consistent gradient structure. It demonstrates that MGF can preserve the contour information of multi-focus images.

2.2 Cooperative Image Multiscale Decomposition

It is known that a good decomposition method should separate the useful information, such as contour, texture, and structure information into the different sub-band images. MGF is a novel edge-preserving filter that offered an effective way of preserving the mutual structure and flat structure. The inconsistent structure that exhibits divergence gradient amplitudes between two source images is filtered out. For multi-focus image fusion, the inconsistent gradient structure is very helpful in adjusting the clarity of two source images. In this paper, the CIMD based on MGF is proposed for multi-focus image fusion. The CIMD can separate the inconsistent structure represented by texture information.

The CIMD is achieved by

$$[B_t, B_r] = \text{MuGuide}(I_t, I_r) \quad (3)$$

$$D_t = I_t - B_t \quad (4)$$

$$D_r = I_r - B_r \quad (5)$$

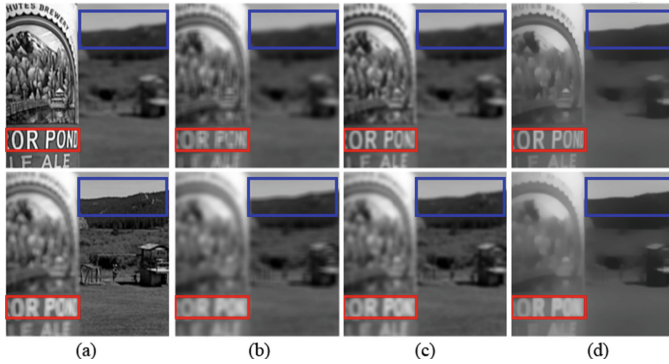


Fig. 1. Two multi-focus images and the output results with different filters. (a) The source multi-focus images; (b) Output from mean filter; (c) Output from guide filter; (d) Output from MGF.

where I_l and I_r represent two source multi-focus images; B_l and B_r are the base layers of the multi-focus images; D_l and D_r represent the detailed layers. The base layers are simultaneously obtained by the mutually guided filter with two source images being employed collaborative references. The base layers contain the consistent structure such as contour, brightness, and part of details of the source images. While the detailed layers obtained by subtracting base layers from the source images show the gradient discrepancy components of the source images.

Figure 2 shows an example of the CIMD on multi-focus images. In Fig. 2, the consistent structure such as the contour features of the source images is extracted separately into the base layers, while the inconsistent structure which shows local clarity of the source images is preserved in the detailed layers. Thus, the different amplitudes of inconsistent information can accurately discriminate the focus area of the source images. In addition, the separation property of consistent structure and inconsistent structure can effectively reduce the artifacts in the fused results.

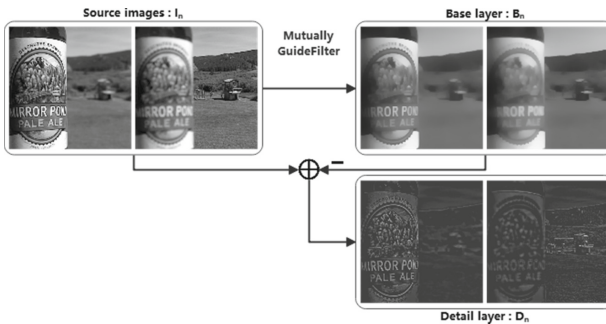


Fig. 2. Cooperative image multiscale decomposition on multi-focus images.

3 Image Fusion with CIMD

The schematic diagram of the proposed fusion strategy is illustrated in Fig. 3. Firstly, two multi-focus images are divided into the base layers and the detailed layers based on CIMD simultaneously. Then, the fused base layers and detailed layers are obtained by weight average fusion rules. Finally, the final fused image is a superimposition of the fused base and detailed layers.

3.1 Base Layers Fusion

Base layers images filtered by CIMD preserve the consistent gradient structure of the source images. However, the consistent structure does not guarantee identical local intensity or other image features in different source images. In particular, the difference between the two base layer images manifests as weak amplitude differences, and the weak amplitude differences can be measured by the local spatial frequency (LSF). Therefore, the difference in the LSF value can be implemented to discriminate the regions of interest. This paper adopts the LSF-based fusion rule to measure the blurred features of base layers images.

The LSF of a certain sub-band image B obtained by

$$LSF_B = \sqrt{Lrf_B^2 + Lcf_B^2} \quad (6)$$

where

$$Lrf_B(i, j) = \frac{1}{5} \sqrt{\sum_{a=-2}^2 \sum_{b=-2}^2 [B(i+a, j+b) - B(i+a, j+b-1)]^2} \quad (7)$$

$$Lcf_B(i, j) = \frac{1}{5} \sqrt{\sum_{a=-2}^2 \sum_{b=-2}^2 [B(i+a, j+b) - B(i+a-1, j+b)]^2} \quad (8)$$

where $B(i, j)$ is the pixel value at position (i, j) .

With Eq. (6), LSF_t and LSF_r of the base layers B_t and B_r are obtained respectively. Then the fusion weighted map \bar{m} for B_t and B_r is constructed as

$$\bar{m}(i, j) = \begin{cases} 1 & \text{if } LSF_t(i, j) > LSF_r(i, j) + TH \\ 0.5 & \text{elseif } \text{abs}(LSF_t(i, j) - LSF_r(i, j)) < TH \\ 0 & \text{Otherwise} \end{cases} \quad (9)$$

where TH is a predetermined threshold. To reduce the artifacts in the fused image, a threshold parameter is defined to restrict that the arithmetic mean fusion rule is used when the value of LSF difference between two images less than TH . In this paper, we set $TH = 0.001$ empirically. Furthermore, the guide filter defined in [12] is used to refined the weighted map as.

$$m = GuideFilter(\bar{m}, B_t) \quad (10)$$

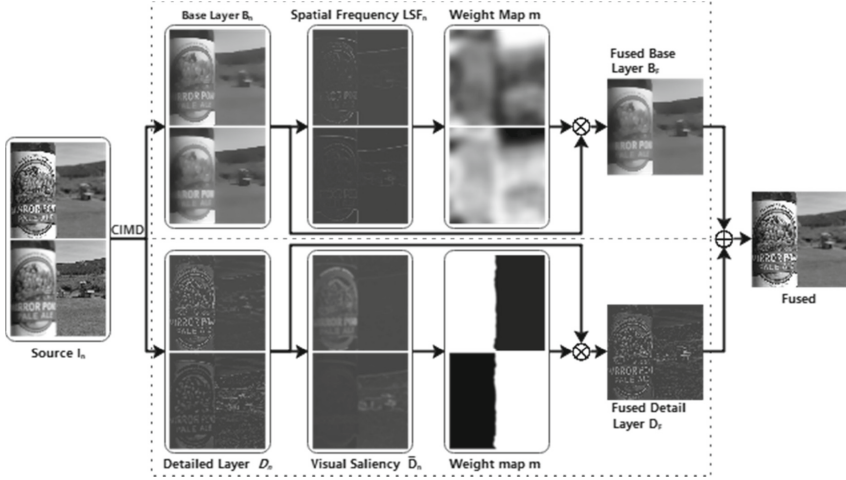


Fig. 3. Schematic diagram of the proposed algorithm.

The base image B_t is served as the guided image. Finally, the fused image B_F is obtained by

$$B_F = mB_t + (1 - m)B_r \tag{11}$$

3.2 Detailed Layers Fusion

The inconsistent structure such as texture information of multi-focused images determines the clarity of the fused image. A good fused result should contain all the inconsistent structure in images obtained by different focus sets. The detailed image filtered by CIMD preserves most of the inconsistent structure such as texture features in multi-focused images. Therefore, a weighted fusion rule based on saliency detection is proposed for distinguishing the saliency of inconsistent structure in detailed layers.

The saliency information map \bar{D} of a certain sub-band image D is obtained by the mean-guide filter as

$$\bar{S} = |D - D * f_m| \tag{12}$$

$$\bar{D} = GuideFilter(D, \bar{S}) \tag{13}$$

where f_m is a 7×7 mean filter operator and $||$ represents the absolute operator. The $*$ denotes the convolution operation. Furthermore, in order to enhance regional consistency of the map, the features saliency \bar{D} are obtained by the guide filter.

With Eq. (12) and Eq. (13), the features saliency information maps \bar{D}_t and \bar{D}_r of the sub-band images D_t and D_r are obtained. The initial weight map \bar{m} is determined by

$$\bar{m}(i, j) = \begin{cases} 1, & \text{if } \bar{D}_t(i, j) > \bar{D}_r(i, j) \\ 0, & \text{Otherwise} \end{cases} \tag{14}$$

Furthermore, to reduce artifacts or halos caused by unaligned source images or detection of edge misalignment problems, the optimized weight map m is obtained by

$$m = \text{GuideFilter}(\bar{m}D_t + (1 - \bar{m})D_r, \bar{m}) \quad (15)$$

The fused detailed layers D_F is obtained by

$$D_F = mD_t + (1 - m)D_r \quad (16)$$

3.3 Reconstruction

Finally, the fused image is constructed by

$$Fused = B_F + D_F \quad (17)$$

4 Experiment

4.1 Experiment Setup

28 pairs of publicly available multi-focus images are selected to test the proposed method. 22 pairs of images are collected from the Lytro multi-focus image dataset created by Nejati et al. in [13]. The remaining 6 pairs are collected from the website [14]. All of the source images are gray images. Figure 4 shows ten pairs of multi-focus images in the experiment dataset. All the experiments are implemented on a laptop equipped with an Intel(R) Core(TM) i5-10400 CPU (2.90GHz) and NVIDIA GeForce RTX 2070 GPU with 16GB RAM. And the software environment is MATLAB R2018b installed on Win 10 64-bit operating system.

To demonstrate the fusion effectiveness of the proposed algorithm, we quantitative compare the proposed method with 7 state-of-the-art image fusion algorithms, including guided filtering fusion (GFF) [5], sparse representation (SR) [7], gradient transfer fusion (GTF) [4], guided filter-based image detection fusion (GFDF) [6], multi-scale structural image decomposition (MSID) [10], Laplacian re-decomposition (LRD) [11], and residual neural network analysis image fusion (ResNet) [9]. For SR, we use the MST-SR-Fusion-Toolbox provided by Liu et al. [8]. The remaining algorithms are implemented based on publicly available codes provided by authors. For all methods, we use the optimal default parameters reported in the related publications.

To qualitatively analyze the effectiveness of different algorithms in terms of the perspective of visual effects, we select six evaluation metrics, including the gradient-based fusion performance $Q^{AB/F}$ [15], the structural similarity-based metric Q_S [19], Q_C [17], Q_Y [16], the human perception inspired metric Q_{CB} [20], the phase congruency-based metric Q_P [18].



Fig. 4. Ten pairs of multi-focus images were selected from the experiment dataset.

4.2 Comparison to Classical Fusion Method

To evaluate the sensitivities of the key parameters in the proposed method, we experimentally studied the influence of parameters α , ϵ in Eq. (1) and Eq. (2). Figure 5 shows the fusion quality measures with different parameter settings. The fusion performances are evaluated by the average value of $Q^{AB/F}$, Q_S , Q_C , Q_Y , Q_{CB} , and Q_P on 28 pairs source images. From Fig. 5(a)–(f), we can conclude that the performance of the proposed method is comprehensively optimal when the parameter α is in the range of [0.05 0.1] and the parameter ϵ is in the range of [0.01 0.02]. Moreover, we can see from Fig. 5(a) when the parameter ϵ is larger than 0.01, the fused results are all of the good quality. Figure 5(f) shows that the fused results are all of the lower quality when the parameter $\epsilon = 0.01$ and the parameter $\alpha = 0.1$. In the following experiments, the parameters are set as $\alpha = 0.05$ and $\epsilon = 0.01$.

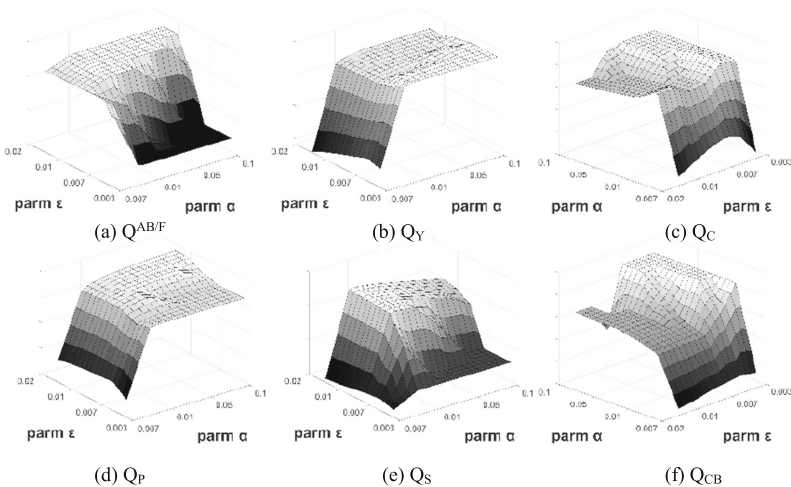


Fig. 5. Influence of parameters α and ϵ . (a)–(f) are the results of $Q^{AB/F}$, Q_Y , Q_C , Q_P , Q_S , and Q_{CB} .

Figure 6 shows an example of the fused images of different methods on the face image. The label regions by the blue rectangle are magnified and shown in the lower-left corner of each image. As shown in Fig. 6(a) and (h), the hair and skin texture of the

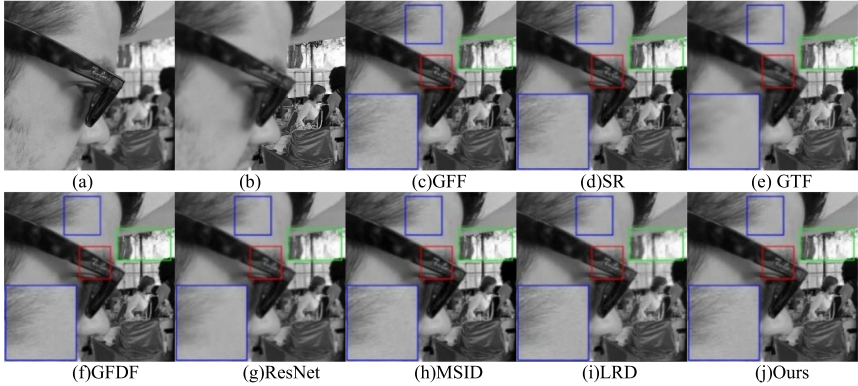


Fig. 6. Fusion results of the proposed and compared methods on the face image: (a) and (b) are the source images. (c)–(j) are the fusion results of GFF, SR, GTF, GFDF, ResNet, MSID, LRD, and the proposed method.

edge area are completely preserved. SR and ResNet in Fig. 6(b) and (e) preserve partial texture

(f)GFDF (g)ResNet (h)MSID (i)LRD (j)Ours.

information of the source image, but the texture of the skin disappear. Figure 6(c) generated by GTF suffers from edge clutter and texture loss. Moreover, there is a global blurring in Fig. 6(c). The edge information of hair is lost in Fig. 6(d), but the skin texture is preserved well. The person in Fig. 6(f) is blurred. Figure 6(g) from LRD fails to handle white balance around the edge of the eyeglass border. The characters on the temple are overexposed caused distortions and contrast decreased surrounding pixels. As seen in Fig. 6(h), the fused image of the proposed algorithm preserves all texture details of surface skin and no blur in the fused image.

Figure 7 shows another example of fusion results of the proposed and compared methods on the game machine image. There are apparent distortions and regional blurring in the gamepads areas of the GFF in Fig. 7(a). Artifacts are visible around the wires in the green rectangle area and the buttons in the blue rectangle area in the final result generated by SR in Fig. 7(b). The GTF in Fig. 7(c) has some blurring of the whole picture due to the lack of accurate focused regions detection, especially the gamepads where there is also brightness distortion. The corresponding shadow area of the gamepads in Fig. 7(d) and (f) is not completely preserved. The whole image generated by ResNet in Fig. 7(e) has over-exposure and blurring. The corresponding red rectangle area in Fig. 7(g) generated by the LRD is blurred, and the whole picture is noisy especially around the wires and the buttons. In contrast, our proposed method can obtain a better visibility effect.

The objective evaluation metrics of Fig. 6 and Fig. 7 are shown in Table 1. For each metric, the best performance result is labeled in bold. The underlined values indicate the second-largest values. Our method provides the highest $Q^{AB/F}$ and Q_Y for two experimental images. In addition, the Q_P and Q_S values of the proposed method are the best in Fig. 6. The Q_C and Q_{CB} values of the proposed method are the best in Fig. 7. The significant information in terms of gradient retention $Q^{AB/F}$, structural similarity Q_Y , and image phase consistency Q_P are achieved maximum preservation in our algorithm.

This is mainly attributable to the fact that the complete decomposition of the consistent structure and the inconsistent structure by CIMD and the appropriate fusion rules. In our experiment, the proposed method can effectively improve the evaluation of subjective effects and the advantage of objective metrics.

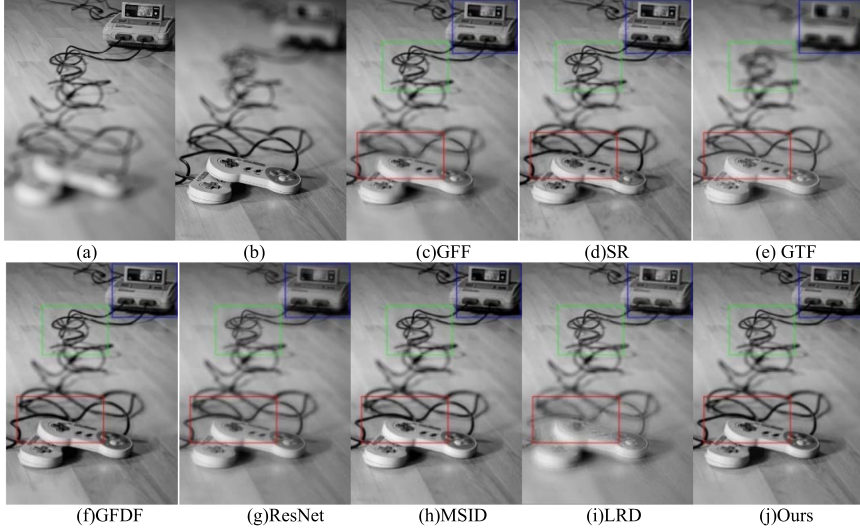


Fig. 7. Fusion results of the proposed and compared methods game machine image: (a) and (b) are the source images. (c)–(j) are the fusion results of GFF, SR, GTF, GFDF, ResNet, MSID, LRD, and the proposed method.

The average metric values and standard deviations on results of 28 pairs of source images with different methods are shown in Table 2. For each metric, the best performance result is labeled in bold, and the underlined values indicate the second-largest values. The smallest standard deviation is bolded and the second smallest is underlined. It can be seen that the proposed method is always the largest in the metrics of $Q^{AB/F}$, Q_Y , Q_C , Q_P , and Q_S for all images, which means that the proposed can well preserve the original information of different sources images. Q_Y and Q_{CB} have the smallest standard deviation. This means the proposed method is very stable in all experimental images.

Table 1. The objective metric values of fused images.

Figures	Metrics	GFF	SR	GTF	GFDF	ResNet	MSID	LRD	Ours
Figure 6	$Q^{AB/F}$	<u>0.7373</u>	0.6772	0.5760	0.7260	0.5663	0.7271	0.5649	0.7403
	Q_Y	0.9808	0.8775	0.9718	<u>0.9847</u>	0.8656	0.8979	0.8583	0.9853
	Q_C	0.8370	0.7368	0.7093	0.8293	0.8027	0.7880	0.7454	<u>0.8319</u>

(continued)

Table 1. (continued)

Figures	Metrics	GFF	SR	GTF	GDFD	ResNet	MSID	LRD	Ours
Figure 7	QP	<u>0.8650</u>	0.7714	0.7251	0.8621	0.7774	0.8423	0.5839	0.8652
	QS	<u>0.9440</u>	0.9341	0.8420	0.9428	0.9051	0.8925	0.9080	0.9444
	QCB	<u>0.7754</u>	0.6887	0.6409	0.8106	0.6212	0.6994	0.6231	0.7590
	Q ^{AB/F}	<u>0.7880</u>	0.6922	0.7058	0.7801	0.6423	0.7772	0.5627	0.7891
	QY	0.9824	0.8281	0.9798	<u>0.9834</u>	0.8879	0.9684	0.8936	0.9848
	QC	<u>0.8940</u>	0.7165	0.7467	0.8935	0.8361	0.8870	0.8144	0.8956
	QP	0.8706	0.6376	0.6348	0.8657	0.7513	0.8388	0.5463	<u>0.8701</u>
	QS	<u>0.9459</u>	0.9407	0.9101	0.9448	0.9304	0.9029	0.9494	0.9451
	QCB	0.7509	0.6470	0.6254	<u>0.7512</u>	0.6341	0.7072	0.5878	0.7611

Table 2. The average metric values and standard deviations on results of 28 pairs of source images with different methods.

Metrics	GFF	SR	GTF	GDFD	ResNet	MSID	LRD	Ours
Q ^{AB/F}	<u>0.7338</u> ± 0.0440	0.6705 ± 0.0558	0.6248 ± 0.0875	0.6913 ± 0.0978	0.5964 ± 0.0790	0.7329 ± 0.0314	0.5523 ± 0.0636	0.7454 ± 0.0316
QY	0.9699 ± 0.0183	0.8340 ± 0.1058	0.9614 ± 0.0291	<u>0.9782</u> ± 0.0105	0.8659 ± 0.0402	0.9392 ± 0.0245	0.8544 ± 0.0416	0.9795 ± 0.0094
QC	<u>0.8207</u> ± 0.0612	0.6962 ± 0.1058	0.7252 ± 0.0855	0.8186 ± 0.0597	0.8021 ± 0.0549	0.8188 ± 0.0523	0.7465 ± 0.0627	0.8231 ± 0.0582
QP	0.8231 ± 0.0880	0.7232 ± 0.1189	0.7183 ± 0.1536	<u>0.8432</u> ± 0.0613	0.7451 ± 0.0667	0.8114 ± 0.0546	0.5523 ± 0.1118	0.8463 ± 0.0560
QS	0.9196 ± 0.0388	0.9150 ± 0.0371	0.8576 ± 0.0632	0.9184 ± 0.0390	0.8941 ± 0.0418	<u>0.9199</u> ± 0.0453	0.8764 ± 0.0398	0.9203 ± 0.0372
QCB	0.7141 ± 0.0195	0.6635 ± 0.1249	0.6452 ± 0.0299	0.7703 ± 0.0111	0.6177 ± 0.0408	0.6512 ± 0.0255	0.6031 ± 0.0454	<u>0.7306</u> ± 0.0102

5 Conclusions

In this paper, we have presented a novel multi-focus fusion method based on MGF. To my knowledge, the proposed method is the first to introduce CIMD into the field of image fusion to obtain the two-scale representation. Exploiting the separation feature of CIMD on the inconsistent structure, and purposely proposing a detailed layers fusion method based on the mean-guide filter and a base layers' fusion method based on the spatial frequency. Finally, the fusion result was obtained by overlaying the fused detailed layers and the base layers. The experimental part compares 28 pairs of multifocal images of three different types and validates the experimental results in terms of both subjective visual evaluation and six sets of objective metrics. The proposed method can preserve more effective information and improving the effectiveness of multi-focus image fusion

compared with seven state-of-the-art methods. New advancement of introducing dual channels on the image fusion framework is achieved.

Acknowledgments. This paper is supported by the National Natural Science Foundation of China (No.61871210) and Chuanshan Talent Project of the University of South China.

References

1. Li, S., Kang, X., Fang, L., Hu, J.: Pixel-level image fusion: a survey of the state of the art. *Inf. Fusion* **33**, 100–112 (2017)
2. Guo, X., Li, Y., Ma, J., Ling, H.: Mutually guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**(3), 694–707 (2018)
3. Shen, X., Zhou, C., Xu, L., Jia, J.: Mutual-structure for joint filtering. In: *ICCV*, pp. 3406–3414. IEEE, Santiago (2015)
4. Ma, J., Chen, C., Li, C., Huang, J.: Infrared and visible image fusion via gradient transfer and total variation minimization. *Inf. Fusion* **31**, 100–109 (2016)
5. Li, S., Kang, X., Hu, J.: Image fusion with guided filtering. *IEEE Trans. Image Process.* **22**(7), 2864–2875 (2013)
6. Qiu, X., Li, M., Zhang, L., Yuan, X.: Guided filter-based multi-focus image fusion through focus region detection. *Signal Process. Image Commun.* **72**, 35–46 (2016)
7. Yang, B., Li, S.: Multifocus image fusion and restoration with sparse representation. *IEEE Trans. Instrument Measur.* **59**(4), 884–892 (2010)
8. Liu, Y., Liu, S., Wang, Z.: A general framework for image fusion based on multi-scale transform and sparse representation. *Inf. Fusion* **24**, 147–164 (2015)
9. Li, H., Wu, X.J., Durrani, T.S.: Infrared and visible image fusion with ResNet and zero-phase component analysis. *Infrared Phys. Technol.* **102**, 103039 (2019)
10. Li, H., Qi, X., Xie, W.: Fast infrared and visible image fusion with structural decomposition. *Knowl. Based Syst.* **204**, 106182 (2020).
11. Li, X., Guo, X., Han, P., Wang, X., Luo, T.: Laplacian re-decomposition for multimodal medical image fusion. *Trans. Instrument Measur.* **69**(9), 6880–6890 (2020)
12. He, K., Sun, J., Tang, X.: Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**(6), 1397–1409 (2013)
13. Nejati M, Lytro Multi-focus Dataset (2019). <https://mansournejati.ece.iut.ac.ir/content/lytro-multi-focus-dataset>. Accessed 6 Jan 2019
14. Pxleyes: Multi Focus Photography. <http://www.pxleyes.com/photography-contest/19726>. Accessed 20 Jan 2021
15. Xydeas, C.S., Petrovic, V.: Objective image fusion performance measure. *Electron. Lett.* **36**(4), 308–309 (2000)
16. Yang, C., Zhang, J., Wang, X., Liu, X.: A novel similarity based quality metric for image fusion. *Inf. Fusion* **9**, 156–160 (2008)
17. Cvejic, N., Loza, A., Bul, D., Canagarajah, N.: A similarity metric for assessment of image fusion algorithms. *Int. J. Signal Process.* **2**(3), 178–182 (2005)
18. Zhao, J., Laganiere, R., Liu, Z.: Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement. *Int. J. Innov. Comput. Inf. Control* **3**(6), 1433–1447 (2007)
19. Piella, G., Heijmans, G.: A new quality metric for image fusion. In: *Proceedings International Conference Image Processing*, pp. 3–173. IEEE, Barcelona (2010)
20. Chen, Y., Blum, R.S.: A new automated quality assessment algorithm for image fusion. *Image Vis. Comput.* **27**(10), 1421–1432 (2009)