Focus-pixel estimation and optimization for multi-focus image fusion

Kangjian He¹ · Jian Gong¹ · Dan Xu¹

Received: 26 November 2020 / Revised: 11 July 2021 / Accepted: 3 January 2022 /
Published online: 28 January 2022

© The Author(s) 2022

Abstract

To integrate the effective information and improve the quality of multi-source images, many spatial or transform domain-based image fusion methods have been proposed in the field of information fusion. The key purpose of multi-focus image fusion is to integrate the focused pixels and remove redundant information of each source image. Theoretically, if the focused pixels and complementary information of different images are detected completely, the fusion image with best quality can be obtained. For this goal, we propose a focus-pixel estimation and optimization based multi-focus image fusion framework in this paper. Because the focused pixels of an image are in the same depth of field (DOF), we propose a multi-scale focus-measure algorithm for the focused pixels matting to integrate the focused region firstly. Then, the boundaries of focused and defocused regions are obtained accurately by the proposed optimizing strategy. And the boundaries are also fused to reduce the influence of insufficient boundary precision. The experimental results demonstrate that the proposed method outperforms some previous typical methods in both objective evaluations and visual perception.

Keywords Multi-focus image fusion · Focus pixel estimation · Optimization · Focus-measure

1 Introduction

The development of sensor technology has promoted the expansion of image processing in many applications [49, 53]. Because of the limitation of depth of field (DOF) in sensor imaging, the optical elements can only capture the focused images of part of the DOF scene [2]. Since the defocused region of multi-focus image contains a lot of redundant information, which results in the failure of the direct application of the multi-focus image to other computer vision tasks. In order to solve above problems, multi-focus image fusion technology has emerged and attracted many researchers' attention. For example, in video surveillance, the

✉ Dan Xu
danxu@ynu.edu.cn

¹ School of Information Science and Engineering, Yunnan University, Kunming 650091, China

quality of the images captured by the camera can be improved by image fusion technology, which contributes to the realization of computer-vision based target detection, recognition and other tasks. In recent years, many multi-focus image fusion methods have been proposed [14, 35, 43, 51], among which the algorithms based on spatial and transform domain have been widely concerned and studied [1, 20, 22].

Most of the transform domain-based methods use different decomposition tools to decompose the source multi-focus image into different scales. After Laplacian pyramid (LP) based method was widely used, a lot of decomposition tools were proposed, such as wavelet transform (WT), discrete wavelet transform (DWT), nonsubsampling shearlet transform (NSST) and nonsubsampling contourlet transform (NSCT) [7, 8, 25, 27, 31, 39]. Most multi-scale transform (MST) domain-based methods decompose the source image into different detail layers and fuse the pixels by different fusion strategies. For these transform-domain-based methods, the details of pixels can be better analyzed in each layer of the transform domain, but the interference caused by noise pixels and image misregistration cannot be effectively eliminated.

The spatial domain-based methods usually include pixel based and block-based image fusion methods [11–13, 19, 28, 30, 36, 40, 42]. The latter uses different measurement strategies to analyze and fuse each block to produce fused image. The advantage of them is that all pixels in the fused block are preserved directly without loss of information. In spatial domain-based methods, spatial frequency (SF) [29], mutual information (MI) [45] and other indicators are often used to evaluate the quality of image. For block-based methods, the block size and the selected evaluation index will directly affect the final fusion quality. So, the focus-measure cannot always determine the focused region completely.

Recently, some deep learning-based approaches have been proposed and widely used in many applications, such as intrusion detection [24], 3D object recognition [5]. Liu *et al.* proposed a convolutional neural network (CNN) for multi-focus image fusion tasks [37], after which, many neural network-based models have been applied in image fusion tasks, such as U-Shaped networks-based method, and FuseGAN-based method [15, 16, 34]. The advantage of these network-based methods is that the decision map of focused region can be found effectively. However, still some problems need to be considered such as the requirement of huge training samples, a large amount of computing and memory resources.

To overcome the shortage of training data for deep learning-based methods, and the lack of focused region detection accuracy of pixel-level based fusion schemes, we propose a novel focus-pixel estimation and optimization based multi-focus image fusion framework. Firstly, we propose a multi-scale focus-measure based method to matting the focused pixels to obtain the focused map. Then, the focused regions are optimized by the proposed quality evaluation-based optimization method, and the boundary of focused and defocused regions are also fused to reduce the influence of insufficient boundary precision. The effectiveness of the proposed method is validated by comparing it with some existing methods. The experimental results show that the proposed method is superior to the compared methods in both objective evaluations and visual perception.

The rest of this paper is organized as follows: The imaging models and characteristics of multi-focus images and the evaluation methods are briefly reviewed in Section 2. The proposed fusion algorithm is presented in Section 3. Section 4 presents several groups of experiments and the corresponding analysis results. Finally, discussions and conclusions are summarized in Section 5.

2 Background

2.1 Multi-focus imaging model

Multi-focus image can be divided into focused region and defocused region and it can be modeled as follows:

$$I(x, y) = \delta(x, y)f(x, y) + (1 - \delta(x, y))f_{\xi}(x, y) \quad (1)$$

where $f(x, y)$ denotes the focused pixels and $f_{\xi}(x, y)$ denotes the defocused pixels; and $\delta(x, y)$ represents the properties of the pixels, and can be defined as:

$$\delta(x, y) = \begin{cases} 1, & \text{if } (x, y) \in S_F \\ 0, & \text{if } (x, y) \in S_D \end{cases} \quad (2)$$

where S_F denotes the focused regions and S_D denotes the union of defocused pixels. The defocused image $f_{\xi}(x, y)$ can be modeled approximately as follow:

$$f_{\xi} = h(x, y) \otimes (f(x, y) + \xi(x, y)) \quad (3)$$

where $\xi(x, y)$ denotes noise, $\otimes$ denotes convolution operating, and $h(x, y)$ denotes 2D Gaussian filter which is given by:

$$h(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (4)$$

where σ represents the parameter to control the blurriness [4, 18]. As shown in Fig. 1, the multi-focus image can be decomposed into focused region S_F , defocused region S_D and boundary region S_B .

The source image can be represented as:

$$S_I = S_F + S_B + S_D \quad (5)$$

In Fig. 1, if the segmentation of the S_F and S_D are accurate enough, that is, $S_B = 0$, the white region in the decision map is as small as possible, the ideal fused all-in-focus image I_F can be generated as follow:

Fig. 1 Segmentation result of different regions of multi-focus image

$$I_F = \sum_{i=1}^n S_{Fi} \quad (6)$$

where n denotes the number of multi-focus images. Since it is difficult to obtain the boundary accurately, in our proposed fusion scheme, the optimization-based boundary feathering method is used to make the obtained boundary region as accurate as possible. For better fusion result, we also fuse the boundary regions by guided filtering to reduce the influence of boundary segmentation imprecision.

2.2 Quality evaluation and analysis

There are many non-reference quality evaluation methods for single image, such as Sum-Modified-Laplacian *SML*, information entropy *IE*, mean gradient $\overline{G}$, and spatial frequency *SF* [6, 21, 23, 48, 52]. The mean gradient can reflect the detail differentiation of the image, and can be defined as:

$$\overline{G} = \left| \frac{f_{ij} - f_{i,j-1}}{2} + \frac{f_{ij} - f_{i-1,j}}{2} \right| \quad (7)$$

where $[i, j]$ denote the position of pixels. And the *SF* can be described as follow:

$$SF = \sqrt{RF^2 + CF^2} \quad (8)$$

where *RF* and *CF* are the spatial row and column frequency, and can be computed by:

$$RF = \sqrt{\frac{1}{m \times n} \sum_{i=1}^m \sum_{j=2}^n (f_{ij} - f_{i,j-1})^2} \quad (9)$$

$$CF = \sqrt{\frac{1}{m \times n} \sum_{i=2}^m \sum_{j=1}^n (f_{ij} - f_{i-1,j})^2} \quad (10)$$

Sum-Modified-Laplacian *SML* [3, 9, 46] can reflect the detailed texture contrast of the image, and can be described by:

$$\begin{aligned} \nabla_{ML}^2 f(x, y) = & \left| 2f(x, y) - f(x-n, y) - f(x+n, y) \right| \\ & + \left| 2f(x, y) - f(x, y-n) - f(x, y+n) \right| \end{aligned} \quad (11)$$

$$S_{ML}(i, j) = \sum_{x=n}^{x+n} \sum_{y=n}^{y+n} \nabla_{ML}^2 f(i, j) \quad (12)$$

where n represents the local window size, and $S_{ML}(i, j)$ is the obtained value within a local window of $n \times n$ centered on pixel (i, j) .

To verify the performance of the above evaluation method, we selected four images for testing. They are shown in Fig. 2. Figure 2a, c are multi-focus images, Fig. 2b, d

are fully focused images. By (3), we generated 50 images for each test image with the degree of blur deepening in turn. Then we calculated the normalized IE , $\overline{G}$, SF and SML values respectively, and the results are shown in Fig. 3. In Fig. 3, the horizontal axis represents the image indices, and the vertical axis represents the normalized evaluation values of each image. The normalized values of different evaluation metrics for each image can be calculated by:

$$\eta = \left(\frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n \eta(i, j) \right) / \eta_{\max} \quad (13)$$

where η denotes the obtained results by IE , $\overline{G}$, SF and SML .

As seen from Fig. 3a, b, the IE and $\overline{G}$ cannot effectively predict the clarities of multi-focus images. However, SF and SML values accurately reflect the blurring degrees. The red lines in Fig. 3c, d show that SML can discern variations more subtly in blurring degrees. Therefore, based on multi-scale SML , we proposed a focus-pixel matting algorithm to determine the focus regions, which will be detailed introduced in Section 3.

Different from non-reference image evaluation method, the reference quality evaluation-based method is usually used to measure the fusion quality in the fusion algorithm comparison. Mutual information (MI) [45], visual information fidelity (VIF) [47] and the gradient-based metric $Q^{AB/F}$ [50] are the commonly used reference quality evaluation metrics. In the optimization part, we use these three metrics to optimize our fusion results for higher evaluation scores.

Fig. 2 Test images. **a** Bee. **b** Dog. **c** Horse. **d** Tree. (a) and (c) are multi-focus images, (b) and (d) are fully focused images

Fig. 3 Performance comparison of different methods. **a** IE. **b** $\bar{G}c$ **c** SF **d** SML

3 Proposed fusion method

The schematic diagram of the proposed multi-focus image fusion method is shown in Fig. 4. Firstly, the source multi-focus images are processed by multi-scale SML to produce the focus-pixel matting maps. Next, the focused region is estimated based on the focus-pixel matting result. Then we refine the decision maps according to the evaluation function. Our goal is to obtain the fused image by (6).

Fig. 4 Schematic diagram of the proposed fusion framework. I_a and I_b are the source multi-focus images, I'_{ai} and I'_{bi} are the MSML results, F_{Mi} is the focus-pixel matting result, M is the focused region matting map, M' is the feathering map, D is the decision map, $f(Q)$ is the fusing quality evaluation function and I_F is the fused image

But given that the boundary region we got was larger than the real boundary, that is, $S_B \neq 0$, the obtained boundary regions of source images are also fused to reduce the influence of insufficient boundary precision. The ideal fused all-in-focus image I_F can be generated as follow:

$$I_F = \sum_{i=1}^n S_{Fi} + F_B \quad (14)$$

The final fusion result can be obtained by integrating the fused boundary region with the focused regions in source images.

3.1 Focused region estimation

It has been introduced in Section 2 that the clarity of the image can be measured by some quality evaluation metrics. Here we propose a focus-pixel matting method based on multi-scale SML and mathematical morphological processing. It has been shown in Fig. 3d that the SML values can measure the clarity of the image well, but the validity of the measurement is related to the size of source image. The normalized values of SML for test images with different sizes are given in Fig. 5 showing that the scores can effectively reflect the clarity when the size is large enough. Otherwise, when the size is small, for example, size of 2×3 or

Fig. 5 Normalized SML evaluation scores for different size of test images. **a** Bee. **b** Dog **c** Horse **d** Tree

even 8×12 , the scores cannot be used as the clarity evaluation. For the above reasons, we propose a multi-scale focus-pixel matting method.

We use the different size n in (11) to obtain the information in local windows with different scales. Then we integrate the results and process them further by mathematical morphology. Figure 6 shows the multi-scale SML results, and the third row of Fig. 6b–e is the focus-pixel matting result F_{Mi} which can be obtained by:

$$F_{Mi} = \begin{cases} 1, & \text{if } S_{ML}(I_a) > S_{ML}(I_b) \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

As shown in Fig. 6, more accurate boundary can be obtained with small local window size. But the accuracy of focused pixel matting is not enough. With the increase of local window size, the matting accuracy can be improved. However, the boundary drift will occur.

The roughly focused region matting map F_M can be produced as follow:

$$F_M = F_{M1} \oplus F_{M2} \cdots \oplus F_{Mi} \quad (16)$$

where $\oplus$ denotes exclusive OR, and the roughly focused region map F_M of Fig. 6 is given in Fig. 7a.

Due to the focused and defocused pixels are regional aggregation. We can improve the final matting accuracy by processing the pixel results in local region. Figure 7b are the partitioning results of Fig. 7a. Figure 7c shows the process of mathematical morphology operation mainly containing two operations, filtering out the independent focused pixels and filling the independent defocused pixels in the local region. The refinement result of the focused region is given in Fig. 7d.

Fig. 6 MSML results. The second to the fifth columns show the SML results with difference scale. **a** the first and second row are the source images and the third row is the ground truth. **b** Size 3×3 . **c** Size 7×7 . **d** Size 13×13 . **e** Size 31×31 . The third row from (b) to (e) are the matting results. The black represents the defocused pixels and the white denotes the focused pixels, and the red line represents the real boundary

Fig. 7 Focused region estimation result. **a** Focused region matting map obtained by (16). **b** The partitioning results of (a). **c** Focused region refining process. **d** The refinement result of the focused region

Here, we set the pixel q and p to the same property if $q \in N_8(p)$ and $\text{sum}(N_8(p)) \geq 5$, where p is the focused pixel. Then the different property partitioning results $\{\gamma_1, \gamma_2, \dots, \gamma_i\}$ can be obtained, where $\gamma_i \in F_M$. The discrete small regions in the red rectangle are removed firstly, which can be denoted as:

$$\gamma_\delta = \begin{cases} 1, & \gamma_\delta \in \{\gamma_1, \gamma_2, \dots, \gamma_i\} \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

where γ_δ is the focused region. Next, as shown in blue rectangle in Fig. 8b, those small regions within the biggest connected region are classified into this region, that is:

$$\chi_i = \begin{cases} 1, & \text{if } \chi_i \in \gamma_\delta, i \leq j \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

The final refinement result of focused region is shown in Fig. 7d.

Fig. 8 **a** Source image a. **b** Source image b. **c** Focused region result obtained by our method. **d** Boundary feathering result. **e** Focused region of image a. **f** Focused region of image b. **g** Fused boundary region. **h** Fusion result

3.2 Decision optimizing

In general, if we can obtain a more accurate estimation of focused region, that is, the boundary region S_B is very close to the real boundary, the fused result can be directly computed by (6). However, the boundary region obtained by our method is not exactly consistent with the real boundary sometimes, so we propose a fusing quality evaluation-based optimization method.

Figure 8 shows how to generate the final fusion image using the matting result. Figure 8c is the focused region obtained by our method. It is not the same as the real boundary. We feather Fig. 8c so that the boundary region can completely cover the real boundary. Figure 8d is the feathering result. Figure 8e, f are the focused regions. Figure 8g is the fused boundary region, and the fusion method will be introduced in the next Section. The fusion result Fig. 8h can be obtained by integrating Fig. 8e–g.

Since the final fusion effect depends on the accuracy of the boundary region, we propose a fusion quality evaluation-based method to optimize the feathering size. As shown in Fig. 9a, the red line represents the real boundary, and the yellow point represents the boundary pixel obtained by our matting algorithm. The boundary region obtained by feathering the yellow point with size 2 contains the real boundary. The pixel values within the window size $(2n + 1) \times (2n + 1)$ around the yellow point are computed by:

$$\bar{p} = \frac{\sum_{i=n}^{2n+1} \sum_{j=n}^{2n+1} p(i, j)}{(2n + 1) \times (2n + 1)} \quad (19)$$

here, the pixels with values outside 0 and 1 are regarded as the boundary pixels. Figure 9b shows the fusing scores obtained by different feathering size, and Fig. 9c is the fusion result with highest score. The feathering size is optimized by the fusing scores.

To demonstrate the effectiveness of decision optimizing strategy in the proposed fusion scheme, Table 1 gives the fusing scores of optimized fusion results with different feathering sizes. It can be seen that the fusion performance can be improved with a suitable feathering size. For Fig. 8, the best fusion result is obtained with $n = 21$.

Three most popular fusion quality metrics are used as the optimization functions in our proposed fusion scheme, MI, VIF and $Q^{AB/F}$. The MI metric of fused image F, source image A and B can be calculated as:

$$MI(F; A; B) = \frac{MI(F, A) + MI(F, B)}{2} \quad (20)$$

where MI of two images α and β can be computed by:

Fig. 9 Optimized fusion result. **a** Illustration of feathering. **b** Evaluation scores by the decision map with different feathering size, note that the feathering size goes never beyond image size. The real boundary can be covered with a smaller feathering size to obtain the highest evaluation score. **c** Fusion result with highest score

Table 1 Objective results of Fig. 8 with different feathering sizes

Images	Metric	n=2	n=5	n=10	n=15	n=20	n=25	n=30
Figure 8.	MI	9.0228	7.8076	7.6781	7.5929	7.5443	7.4921	7.4395
	$Q^{AB/F}$	0.6365	0.7473	0.7478	0.7488	0.7491	0.7491	0.7489
	VIF	0.6494	0.7266	0.7274	0.7275	0.7263	0.7265	0.7253

$$MI(\alpha, \beta) = H(\alpha) + H(\beta) - H(\alpha, \beta) \quad (21)$$

here $H(\cdot)$ denotes the entropy of the image, and is defined as:

$$H(\cdot) = - \sum_{i=1}^n p_i \log p_i \quad (22)$$

and VIF measure can be expressed as:

$$VIF = \frac{\sum_{j \in \text{subbands}} I(\vec{C}^{Nj}; \vec{F}^{Nj} | s^{Nj})}{\sum_{j \in \text{subbands}} I(\vec{C}^{Nj}; \vec{E}^{Nj} | s^{Nj})} \quad (23)$$

where $I(\vec{C}^{Nj}; \vec{F}^{Nj} | s^{Nj})$ and $I(\vec{C}^{Nj}; \vec{E}^{Nj} | s^{Nj})$ can be computed by:

$$I(\vec{C}^{Nj}; \vec{F}^{Nj} | s^{Nj}) = \frac{1}{2} \sum_{i=1}^N \log_2 \left(\frac{|g_i^2 s_i^2 C_U + (\sigma_v^2 + \sigma_n^2) I|}{|(\sigma_v^2 + \sigma_n^2) I|} \right) \quad (24)$$

Fig. 10 Illustration of fusion process. **a** and **b** are the source images. **c** is the decision map obtained by us. **d** and **e** are the focused regions of **(a)** and **(b)**. **f** to **h** are the focused region fusion results with different boundary sizes, where the white regions are the boundary. The corresponding white regions in **(a)** and **(b)** are fused by (28) to produce fused boundary. The final fusion result is given in **(i)**

Fig. 11 Some multi-focus images used in experiments

Fig. 12 Fusion results obtained by different methods on pair 1. **a** and **b** are source images. **c** NSCT. **d** MWGF. **e** IMF. **f** CBF. **g** GFF. **h** CNN. **i** BRW. **j** GFDF. **k** Proposed

Fig. 13 Fusion results obtained by different methods on pair 2. **a** and **b** are source images. **c** NSCT. **d** MWGF. **e** IMF. **f** CBF. **g** GFF. **h** CNN. **i** BRW. **j** GFDF. **k** Proposed

Fig. 14 Normalized difference images between the source image and each of the fused image of pair 2. **a** NSCT. **b** MWGF. **c** IMF. **d** CBF. **e** GFF. **f** CNN. **g** BRW. **h** GFDF. **i** Proposed

Fig. 15 Fusion results obtained by different methods on pair 3. **a** and **b** are source images. **c** NSCT. **d** MWGF. **e** IMF. **f** CBF. **g** GFF. **h** CNN. **i** BRW. **j** GFDF. **k** Proposed

Fig. 16 Normalized difference images between the source image and each of the fused image of pair 3. **a** NSCT. **b** MWGF. **c** IMF. **d** CBF. **e** GFF. **f** CNN. **g** BRW. **h** GFDF. **i** Proposed

Fig. 17 Fusion results with different boundary sizes. **a** Decision map obtained with different boundary sizes. **b** Evaluation metrics with different boundary sizes. The horizontal axis represents the boundary sizes and the vertical axis represents the evaluation values

$$I(\vec{C}^{Nj}; \vec{E}^{Nj} | s^{Nj}) = \frac{1}{2} \sum_{i=1}^N \log_2 \left(\frac{|s_i^2 C_U + \sigma_n^2 I|}{|\sigma_n^2 I|} \right) \quad (25)$$

The detailed explanations of these parameters can be found in Ref. [47]. The gradient-based metric $Q^{AB/F}$ is calculated as:

Fig. 18 Fusion results of other source images. **a** NSCT. **b** MWGF. **c** IMF. **d** CBF. **e** GFF. **f** CNN. **g** BRW. **h** GFDF. **i** Proposed

Table 2 Objective results of various fusion methods on pair1-pair3. The best results are shown in bold

Images	Metric	NSCT	MWGF	IMF	CBF	GFF	CNN	BRW	GFDF	Ours
Pair 1	<i>MI</i>	3.0753	6.1399	7.0506	5.3320	5.7845	6.8865	6.7198	6.7355	8.3487
	$Q^{AB/F}$	0.4286	0.7535	0.7575	0.7450	0.7571	0.7586	0.6517	0.6509	0.7589
	<i>VIF</i>	0.3231	0.6398	0.6586	0.5826	0.6546	0.6595	0.5757	0.5767	0.7375
Pair 2	<i>MI</i>	5.7619	7.7542	8.0861	6.9664	7.6265	8.1425	8.0708	8.0679	8.2500
	$Q^{AB/F}$	0.5497	0.7435	0.7432	0.7311	0.7437	0.8015	0.7449	0.7442	0.7466
	<i>VIF</i>	0.4804	0.7490	0.7518	0.6936	0.7546	0.7766	0.7590	0.7597	0.7878
Pair 3	<i>MI</i>	4.7477	7.2633	7.5627	6.2015	7.1716	8.9304	7.3990	7.3686	7.5305
	$Q^{AB/F}$	0.6300	0.7462	0.7416	0.7485	0.7489	0.7410	0.7465	0.7464	0.7492
	<i>VIF</i>	0.5269	0.7181	0.7110	0.6735	0.7271	0.7073	0.7238	0.7235	0.7264

$$Q^{AB/F} = \frac{\sum_{i=1}^W \sum_{j=1}^H Q^{AF}(x,y)w^A(x,y) + Q^{BF}(x,y)w^B(x,y)}{\sum_{x=1}^W \sum_{y=1}^H (w^A(x,y) + w^B(x,y))} \quad (26)$$

where $Q^{AF}(x,y) = Q_g^{AF}(x,y)Q_a^{AF}(x,y)$, $Q_g^{AF}(x,y)$ and $Q_a^{AF}(x,y)$ indicate the edge strength and orientation preservation values at pixel (x, y) . w^A and w^B are the weight coefficients of Q_g^{AF} and Q_a^{AF} .

Then the optimizing function $f(Q)$ is defined as:

$$f(Q) = \frac{1}{3} \left(\text{Nor}(MI) + VIF + Q^{AB/F} \right) \quad (27)$$

where $\text{Nor}(\cdot)$ denotes the normalized operations.

3.3 Boundary fusing

Because the boundary region we obtained is larger than the real boundary, so we fuse the obtained boundary region by guided filter (GF) [10, 17, 32] to improve the final fusion quality. The fused image can be obtained by:

$$O_i = a_k I_i + b_k, \forall i \in \omega_k \quad (28)$$

where ω_k is the local window entered pixel k , a_k and b_k are constant coefficients and can be computed by:

Table 3 Objective results of various fusion methods on pair4-pair 6. The best results are shown in bold

Images	Metric	NSCT	MWGF	IMF	CBF	GFF	CNN	BRW	GFDF	Ours
Pair 4	<i>MI</i>	8.1507	9.6563	9.7188	9.0457	8.4511	9.1273	9.7092	9.7185	9.7181
	$Q^{AB/F}$	0.6282	0.7357	0.7352	0.7285	0.7149	0.7231	0.7388	0.7389	0.7410
	<i>VIF</i>	0.6226	0.8020	0.7916	0.7441	0.6990	0.7074	0.8063	0.8066	0.8061
Pair 5	<i>MI</i>	5.0607	7.8477	8.1613	7.2747	7.7044	7.5264	7.8175	7.8129	8.2409
	$Q^{AB/F}$	0.4625	0.7982	0.7992	0.7919	0.8004	0.7493	0.7971	0.7968	0.8024
	<i>VIF</i>	0.4193	0.7726	0.7703	0.7337	0.7760	0.7276	0.7707	0.7708	0.7788
Pair 6	<i>MI</i>	4.8068	7.2235	7.2968	6.1211	6.8450	9.8028	7.0763	7.0762	9.4775
	$Q^{AB/F}$	0.5400	0.7657	0.7547	0.7529	0.7660	0.7414	0.7613	0.7613	0.7685
	<i>VIF</i>	0.4701	0.7193	0.6971	0.6496	0.7123	0.8086	0.7200	0.7209	0.7238

$$a_k = \frac{x(1/|\omega|) \sum_{i \in \omega_k} I_i P_i - \mu_k \overline{P_k}}{\sigma_k + \varepsilon} \quad (29)$$

$$b_k = \overline{P_k} - a_k \mu_k \quad (30)$$

where ε is the regularization parameter, μ_k and σ_k are the mean and variance of the pixels in local window ω_k , $\overline{P_k}$ is the mean value of input image P in local window ω_k .

Figure 10 shows how to obtain the final fusion result, where the white region is the boundary region that needs to be fused by (28). The best case is that the white region is 0, while the worst case is that the decision map we obtained is completely inaccurate with the white extending to the whole image. Hence, regardless of the accuracy of our focused region matting, our fusion result will not be worse than GF algorithm.

4 Experiments and analysis

To illustrate the performance of the proposed fusion scheme, some typical and efficient image fusion algorithms are compared in this section. All the experimental simulations are performed on a PC with 1.8GHZ Inter Core i7 CPU with 8GB RAM.

4.1 Experimental setup

All experiments are conducted on some widely used publicly available multi-focus image sets [41] (<http://mansournejati.ece.iut.ac.ir/content/lytro-multifocus-dataset>) (<https://github.com/lsauto/MWGF-Fusion>), some of which are shown in Fig. 11.

The proposed fusion method is compared with eight multi-focus image fusion algorithms namely Non-subsampled contourlet transform based method (NSCT) [51], multi-scale weighted gradient-based method (MWGF) [53], image matting-based method (IFM) [33], cross bilateral filter based fusion method (CBF) [26], guided filter based fusion method (GFF) [32], convolutional neural networks-based method (CNN) [37], boosted random walks-based method (BRW) [38], and guided filter-region detection based method (GFDF) [44]. MI, $Q^{AB/F}$ and VIF are used as the objective image fusion quality metrics in the experiments. MI measures the dependence between fused image and the source images. $Q^{AB/F}$ is the gradient-based evaluation metric and used to measure the details and gradient information. VIF measures the visual information fidelity between the fused image and the source images. The larger these metrics, the better the quality of the fused image.

4.2 Experimental results and analysis

To illustrate the visual effectiveness of the fused results obtained by our algorithm, some fusion results of different methods are presented in Fig. 12, 13, 14, 15, 16, 17 and 18.

Figure 12 shows the effect of the proposed fusion algorithm in dealing with unregistered images. It can be seen from Fig. 12a, b that the source image is not fully registered, which will lead to the artifacts in the fused results obtained by the pixel-based fusion methods. The magnified regions of different results are shown in red rectangles which shows that the visual

Table 4 The average objective assessments of each method. The best results are shown in bold

Metric	NSCT	MWGF	IMF	CBF	GFF	CNN	BRW	GFDF	Ours
MI	5.2671	7.6474	7.9793	6.8235	7.2638	8.4026	7.7987	7.7966	8.5839
$Q^{AB/F}$	0.5398	0.7571	0.7552	0.7496	0.7551	0.7524	0.7401	0.7397	0.7618
VIF	0.4737	0.7334	0.7307	0.6795	0.7205	0.7311	0.7259	0.7264	0.7425

effects of NSCT, CBF and GFF are worse than other algorithms. The evaluation metrics are given in Table 2. It shows that the quality of the fused image obtained by our algorithm is better than the other algorithms.

Figure 13 shows the fusion results by different methods on pair 2. To better illustrate the performance of these algorithms, the normalized difference images between the fused result and source images are given in Fig. 14. The white point represents that this region is not as same as the corresponding region in source image, indicating that the loss of the information is in this region. The MWGF algorithm cannot fuse the edge information effectively. Compared with GFF, IMF and CNN, the proposed method can integrate the information of source image more completely. Figure 15 presents the third set of experimental results. To better compare the fusion performance of the different algorithms, the normalized difference images between the fused result and source image of pair 3 are given in Fig. 16 which shows that the proposed method can fully detect the focused regions.

Figure 17 shows the optimization process of the boundary region by our proposed algorithm. According to the evaluation score computed by (27), the optimal boundary region is selected and the fusion result with higher score can be obtained. In general, when the value of the evaluation score decreases continuously, the optimizing iteration can be stopped and the computed highest value can be regarded as the best fusion result.

Figure 18 shows the fusion results of other image sets, and the evaluation metrics are given in Table 3. The average objective assessments of each method are given in Table 4, and the bar graph of the average objective assessments are given in Fig. 19. The maximum value in each row in Table 4 are marked in bold. It can be seen that our method can obtain the largest average values on MI , $Q^{AB/F}$ and VIF , which means that the fusion results by our method contain more information, more detail texture, and have better visual fidelity. The above subjective and objective evaluations show that the proposed fusion algorithm outperforms the compared methods.

Fig. 19 Bar graph of the average objective assessments

5 Conclusion

A novel focus-pixel matting and optimizing based multi-focus image fusion method is proposed in this paper. The proposed algorithm aims to integrate the focused pixels to generate all-in-focus image. We propose a multi-scale focus-measure based algorithm for the focused pixels matting and estimating the focused regions. Because the ideal fused image can be produced by integrating the focused regions of source images, a quality evaluation-based optimization strategy is employed to optimize the boundary regions to reduce the influence of insufficient focused boundary precision. Experimental results demonstrate that the proposed method outperforms some state-of-the-art fusion methods in both objective evaluation and visual perception.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant 62162068, Grant 61761046, Grant 61540062, and 62061049, in part by the Yunnan Province Ten Thousand Talents Program and Yunling Scholars Special Project under Grant YNWR-YLXZ-2018-022, in part by the Joint Fund of Yunnan Provincial Science and Technology Department and Yunnan University under Grant 2019FY003012.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ardeshtir Goshtasby A, Nikolov S (2007) Image fusion: advances in the state of the art. *Information Fusion* 8(2):114–118
2. Aslantas V, Kurban R (2010) Fusion of multi-focus images using differential evolution algorithm. *Expert Syst Appl* 37(12):8861–8870
3. Bai X, Zhang Y, Zhou F, Xue B (2015) Quadtree-based multi-focus image fusion using a weighted focus-measure. *Information Fusion* 22:105–118
4. Chen Y, Guan J, Cham W (2017) Robust multi-focus image fusion using edge model and multi-matting. *IEEE Trans Image Process* 27(3):1526–1541
5. Chowdhary CL (2019) 3D object recognition system based on local shape descriptors and depth data analysis. *Recent Patents on Computer Science* 12(1):18–24
6. Cvejic N, Canagarajah CN, Bull DR (2006) Image fusion metric based on mutual information and Tsallis entropy. *Electron Lett* 42(11):626–627
7. Da CA, Zhou J, Do MN (2006) The nonsubsampling Contourlet transform: theory, design, and applications. *IEEE Trans Image Process* 15:3089–3101
8. Dippel S, Stahl M, Wiemker R (2002) Multiscale contrast enhancement for radiographies: Laplacian pyramid versus fast wavelet transform. *IEEE Trans Med Imaging* 21(4):343–353
9. Du J, Li W, Xiao B, Nawaz Q (2016) Union Laplacian pyramid with multiple features for medical image fusion. *Neurocomputing* 194:326–339
10. Dundar T, Ince T (2018) Sparse representation-based hyperspectral image classification using multiscale superpixels and guided filter. *IEEE Geosci Remote Sens Lett* 16(2):246–250
11. Farid MS, Mahmood A, Al-Maadeed SA (2019) Multi-focus image fusion using content adaptive blurring. *Information fusion* 45:96–112

12. Favaro P, Soatto S (2005) A geometric approach to shape from defocus. *IEEE Trans Pattern Anal Mach Intell* 27(3):406–417
13. Gangapure VN, Banerjee S, Chowdhury AS (2015) Steerable local frequency based multispectral multifocus image fusion. *Information Fusion* 23:99–115
14. Gonzalo P, de La Cruz JM (2004) A wavelet-based image fusion tutorial. *Pattern Recogn* 37(9):1855–1872
15. Guo X, Nie R, Cao J, Zhou D, Qian W (2018) Fully convolutional network-based multifocus image fusion. *Neural Comput* 30(7):1775–1800
16. Guo X, Nie R, Cao J, Zhou D, Mei L, He K (2019) Fusegan: Learning to fuse multi-focus image via conditional generative adversarial network. *IEEE Transactions on Multimedia* 21(8):1982–1996
17. He K, Sun J, Tang X (2012) Guided image filtering. *IEEE Trans Pattern Anal Mach Intell* 35(6):1397–1409
18. He K, Zhou D, Zhang X, Nie R (2018) Multi-focus: focused region finding and multi-scale transform for image fusion. *Neurocomputing* 320:157–170
19. He K, Zhou D, Zhang X, Nie R, Jin X (2019) Multi-focus image fusion combining focus-region-level partition and pulse-coupled neural network. *Soft Comput* 23(13):4685–4699
20. Huang W, Jing Z (2007a) Multi-focus image fusion using pulse coupled neural network. *Pattern Recogn Lett* 28(9):1123–1132
21. Huang W, Jing Z (2007b) Evaluation of focus measures in multi-focus image fusion. *Pattern Recogn Lett* 28(4):493–500
22. Ji Z, Kang X, Zhang K, Duan P, Hao Q (2018) A two-stage multi-focus image fusion framework robust to image mis-registration. *IEEE Access* 7:123231–123243
23. Kellman P, McVeigh ER (2005) Image reconstruction in SNR units: a general method for SNR measurement. *Magn Reson Med* 54(6):1439–1447
24. Khare N, Devan P, Chowdhary CL, Bhattacharya S, Singh G, Singh S, Yoon B (2020) Smo-dnn: Spider monkey optimization and deep neural network hybrid classifier model for intrusion detection. *Electronics* 9(4):692
25. Kong W, Zhang L, Lei Y (2014) Novel fusion method for visible light and infrared images based on NSST-SF-PCNN. *Infrared Phys Technol* 65:103–112
26. Kumar BKS, Shreyamsha BK (2015) Image fusion based on pixel significance using cross bilateral filter. *SIViP* 9(5):1193–1204
27. Lewis JJ, O’Callaghan RJ, Nikolov SG, Bull DR, Canagarajah N (2007) Pixel-and region-based image fusion with complex wavelets. *Information Fusion* 8(2):119–130
28. Li S, Yang B (2008a) Multifocus image fusion by combining curvelet and wavelet transform. *Pattern Recogn Lett* 29(9):1295–1301
29. Li S, Yang B (2008b) Multifocus image fusion using region segmentation and spatial frequency. *Image Vis Comput* 26(7):971–979
30. Li M, Cai W, Tan Z (2006) A region-based multi-sensor image fusion scheme using pulse-coupled neural network. *Pattern Recogn Lett* 27(16):1948–1956
31. Li S, Yang B, Hu J (2011) Performance comparison of different multi-resolution transforms for image fusion. *Information Fusion* 12(2):74–84
32. Li S, Kang X, Hu J (2013a) Image fusion with guided filtering. *IEEE Trans Image Process* 22(7):2864–2875
33. Li S, Kang X, Hu J, Yang B (2013b) Image matting for fusion of multi-focus images in dynamic scenes. *Information Fusion* 14(2):147–162
34. Li HG, Nie RC, Cao JD, Guo XP, Zhou DM, He KJ (2019) Multi-focus image fusion using U-shaped networks with a hybrid objective. *IEEE Sensors J* 19(21):9755–9765
35. Liu W, Wang ZF (2020) A novel multi-focus image fusion method using multiscale shearing non-local guided averaging filter. *Signal Process* 166:107252
36. Liu Y, Jin J, Wang Q, Shen Y, Dong X (2014) Region level based multi-focus image fusion using quaternion wavelet and normalized cut. *Signal Process* 97:9–30
37. Liu Y, Chen X, Peng H, Wang Z (2017) Multi-focus image fusion with a deep convolutional neural network. *Information Fusion* 36:191–207
38. Ma J, Zhou Z, Wang B, Miao L, Zong H (2019) Multi-focus image fusion using boosted random walks-based algorithm with two-scale focus maps. *Neurocomputing* 335(28):9–20
39. Mallat SG (1989) A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans Pattern Anal Mach Intell* 11(7):674–693
40. Mitianoudis N, Stathaki T (2007) Pixel-based and region-based image fusion schemes using ICA bases. *Information Fusion* 8(2):131–142
41. Nejati M, Samavi S, Shirani S (2015) Multi-focus image fusion using dictionary-based sparse representation. *Information Fusion* 25:72–84

42. Piella G (2003) A general framework for multiresolution image fusion: from pixels to regions. *Information Fusion* 4(4):259–280
43. Pradhan PS, King RL, Younan NH (2006) Estimation of the number of decomposition levels for a wavelet-based multiresolution multisensor image fusion. *IEEE Transactions on Geoscience & Remote Sensing* 44(12):3674–3686
44. Qiu X, Li M, Zhang L, Yuan X (2019) Guided filter-based multi-focus image fusion through focus region detection. *Signal Process Image Commun* 72:35–46
45. Qu G, Zhang D, Yan P (2002) Information measure for performance of image fusion. *Electron Lett* 38(7): 313–315
46. Qu X, Yan J, Yang G (2009) Multifocus image fusion method of sharp frequency localized Contourlet transform domain based on sum-modified-Laplacian. *Opt Precis Eng* 17(5):1203–1212
47. Sheikh HR, Bovik AC (2006) Image information and visual quality. *IEEE Transactions on Image Processing* 15(2):430–444
48. Tsai DY, Lee Y, Matsuyama E (2008) Information entropy measure for evaluation of image quality. *Journal of Digital Imaging* 21(3):338–347
49. Xiao XH, Yao YS, Yin LJ, Wu SD, Li HC, Yang ZB (2018) Multi-focus image fusion based on probability filtering and region correction. *Signal Process* 153:71–82
50. Xydeas CS, Petrovic V (2000) Objective image fusion performance measure. *Electron Lett* 36(4):308–309
51. Zhang Q, Guo B (2009) Multifocus image fusion using the nonsubsampling Contourlet transform. *Signal Process* 89(7):1334–1346
52. Zheng Y, Essock EA, Hansen BC, Haun AM (2007) A new metric based on extended spatial frequency and its application to DWT based fusion algorithms. *Information Fusion* 8(2):177–192
53. Zhou Z, Li S, Wang B (2014) Multi-scale weighted gradient-based fusion for multi-focus images. *Information Fusion* 20:60–72

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.