Multi-focus Image Fusion based on Multi-scale Focus Measures and Generalized Random Walk

Jinlei Ma, Zhiqiang Zhou, Bo Wang, Mingjie Dong

School of Automation, Beijing Institute of Technology, Beijing 100081, China
E-mail: zhzhzhou@bit.edu.cn

Abstract: Multi-focus image fusion aims to produce an all-in-focus image by integrating a series of partially focused images of the same scene. A small defocused (focused) region is usually encompassed by a large focused (defocused) region in the partially focused image, however, many state-of-the-art fusion methods cannot correctly distinguish this small region. To solve this problem, we propose a novel multi-focus image fusion algorithm based on multi-scale focus measures and generalized random walk (GRW) in this paper. Firstly, the multi-scale decision maps are obtained with multi-scale focus measures. Then, multi-scale guided filters are used to make the decision maps accurately align the boundaries between focused and defocused regions. Next, the GRW is introduced to effectively combine the advantages of the decision maps in different scales. As a result, our method can effectively distinguish the small defocused (focused) regions encompassed by large focused (defocused) regions, and the boundaries can also be aligned accurately. Experimental results demonstrate that our method can achieve a superior performance compared with other fusion methods in both subjective and objective assessments.

Key Words: Multi-focus Image, Image Fusion, Multi-scale Focus Measures, Generalized Random Walk

1 Introduction

The depth of field in optical lens of conventional cameras is usually limited. Hence, only the objects falling in the depth of field are in focus and appear to be sharp, while other objects out of the depth of field are blurred. However, for accurately interpreting and analyzing images, it is important to obtain images with all objects in focus. Multi-focus image fusion is an effective technique to solve this problem by merging the focused objects from a series of images with different focus settings to construct a single all-in-focus image. The multi-focus image fusion has been widely applied in various applications, such as remote sensing [1], biochemical analysis [2] and so on.

In recent years, many multi-focus image fusion methods have been proposed. According to the fusion domain, these methods can be classified into two categories: spatial domain based methods and transform domain based methods. The spatial domain based methods usually fuse images via pixels or image blocks, such as multi-scale weighted gradient-based fusion algorithm (MWGF) [3], dense SIFT based algorithm [4], deep convolutional neural network based algorithm (CNN) [5]. The transform domain based methods produce transform coefficients first, and then these coefficients are fused according to some rules. Lastly, the fused image is reconstructed from these composite coefficients. The transform domain based methods include Laplacian pyramid [6], nonsubsampled contourlet transform [7], guided filtering based fusion algorithm [8], multi-scale transform and sparse representation based algorithm [9].

A large focused (defocused) region usually encompasses a small defocused (focused) region or hole in the partially focused image. However, many state-of-the-art fusion methods cannot correctly distinguish this small region [3, 5, 10] (for example, Fig. 1 (c)). In order to solve this problem, we propose a novel multi-focus image fusion method in spatial domain based on multi-scale focus measures and generalized

random walk (GRW). Our fusion method can effectively distinguish this small region, and the decision map is able to accurately align the boundaries between focused and defocused regions (for example, Fig. 1 (d)).

This work mainly includes two contributions. 1) By combining the advantages of multi-scale focus measures, our method can effectively distinguish the small defocused (focused) regions encompassed by the large focused (defocused) regions. 2) With multi-scale guided filters and GRW, our decision map can accurately align the boundaries between focused and defocused regions.

The rest of this paper is organized as follows. In Section 2, SML and GRW are introduced. Section 3 describes our multi-focus image fusion method based on multi-scale SML and GRW. Experimental results and comparisons are given in Section 4. In Section 5, we conclude this paper.

2 Sum of the Modified Laplacian (SML) and Generalized Random Walks (GRW)

2.1 Sum of the Modified Laplacian (SML)

The focus measures are used to measure the image clarity in the field of multi-focus image fusion. The focus measures mainly include variance, energy of image gradient (EOG), energy of Laplacian of the image (EOL), sum of the modified Laplacian (SML), spatial frequency (SF) [11]. In [11], SML has been experimentally proved to outperform the other focus measures. Thus, SML is chosen as the used focus measure in this paper.

In [12], Nayar noted that in the case of the Laplacian, the second derivatives in the x- and y-directions can have opposite signs and tend to cancel each other. Therefore, he proposed the modified Laplacian (ML).

The expression for the discrete approximation of ML is:

$$\nabla^2_{ML}f(x,y) = |2f(x,y) - f(x-step,y) - f(x+step,y)| \\ + |2f(x,y) - f(x,y-step) - f(x,y+step)| \tag{1}$$

In order to accommodate for possible variations in the size of texture elements, Nayar used a variable spacing ($step$) between the pixels to compute ML. $step$ is set as 1

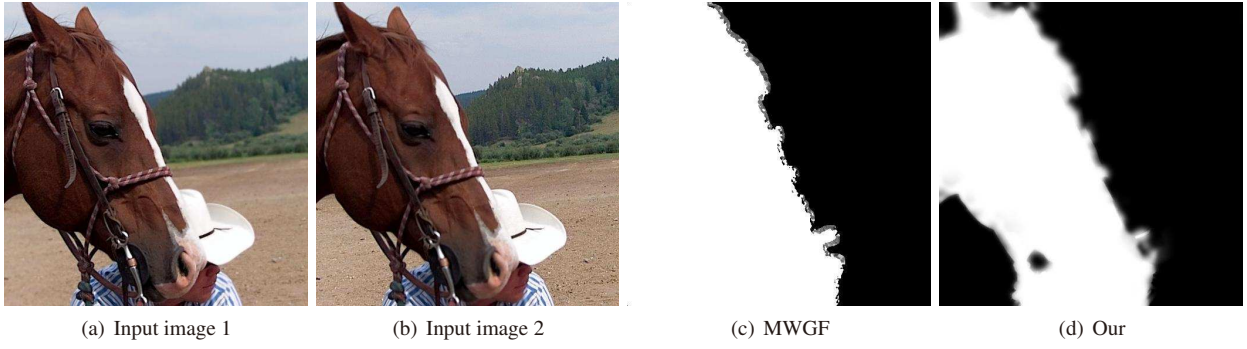(a) Input image 1   (b) Input image 2   (c) MWGF   (d) Our

Fig. 1: The decision maps from MWGF and our method. (a) Input image 1 focuses on the foreground. (b) Input image 2 focuses on the background. (c) The decision map obtained with MWGF. (d) The decision map obtained with our method.

in this paper. SML is defined as

$$
\begin{aligned}
& SML\left(f\left(x,y\right),N\right)= \\
& \sum_{i=x-N}^{i=x+N}\sum_{j=y-N}^{j=y+N}\nabla_{ML}^2 f\left(i,j\right) \quad for \quad \nabla_{ML}^2 f\left(x,y\right)\ge T
\end{aligned}
\tag{2}
$$

where N is the radius of the window and T is a threshold value set as 0 in this paper.
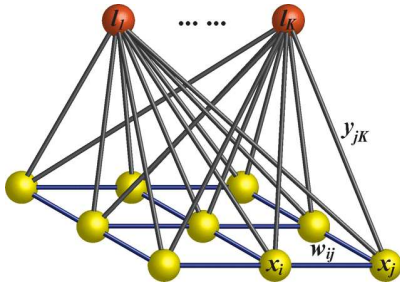

Fig. 2: The graph used in GRW. The yellow nodes are scene nodes and the orange nodes are label nodes.

2.2 Generalized Random Walks (GRW)

The random walk (RW) algorithm has been used extensively for interactive image segmentation in the image processing and computer vision [14]. The RW algorithm computes the first arrival probability that a random walk starts at one pixel first reaches one of the seeds with each label, and then that pixel is denoted as the same label with maximum probability of the corresponding seed. According to RW, the generalized random walk (GRW) is proposed to fuse multi-expose images [13].

The graph used in GRW is shown in Fig. 2. The variable set $\mathcal{X}$ and the label set $\mathcal{L}$ are represented in a weighted undirected graph. Each variable is associated with a pixel location, and each label is associated with an input image in our case. The graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$ is constructed as $\mathcal{V}=\mathcal{L}\bigcup\mathcal{X}$ and $\mathcal{E}=\mathcal{V}\times\mathcal{V}$ including edges both within $\mathcal{X}$ and between $\mathcal{X}$ and $\mathcal{L}$. The yellow nodes are scene nodes and the orange nodes are label nodes. For a scene node $x\in\mathcal{X}$, edges $\mathcal{E}_\mathcal{X}$ are drawn between it and each of its immediate neighbors in $\mathcal{X}$. In addition, edges $\mathcal{E}_\mathcal{L}$ are drawn between a scene node and each label node. $\omega_{ij}=\omega(x_i,x_j)$ is a function defined on $\mathcal{E}_\mathcal{X}$ that models the similarity between nodes x_i and x_j, and $y_{ik}=y(x_i,l_k)$ is a function defined on $\mathcal{E}_\mathcal{L}$ that models the similarity between x_i and l_k.

Let $\mathcal{V}$ be arranged in a way that the first K nodes are label nodes, i.e., $\{v_1,\ldots,v_K\}=\mathcal{L}$, and the rest N nodes are scene nodes, i.e., $\{v_{K+1},\ldots,v_{K+N}\}=\mathcal{X}$. We can define a node similarity function $c(\cdot,\cdot)$ on $\mathcal{E}$ with the following form:

$$
c_{ij}\triangleq c\left(v_i,v_j\right)=\left\{
\begin{array}{ll}
y_{i-K,j}, & \left(v_i,v_j\right)\in\mathcal{E}_\mathcal{X}\wedge v_j\in\mathcal{L} \\
\omega_{i-K,j-K}, & \left(v_i,v_j\right)\in\mathcal{E}_\mathcal{X}
\end{array}
\right.
\tag{3}
$$

Since the graph is undirected, we have $c_{ij}=c_{ji}$. Let $u(v_i)$ denote the potential associated with v_i. Based on the relationship between RW and electrical networks [15], the total energy of the system given in Fig. 2 is

$$
E=\frac{1}{2}\sum_{(v_i,v_j)\in\mathcal{E}}c_{ij}\left(u\left(v_i\right)-u\left(v_j\right)\right)^2
\tag{4}
$$

The goal of the GRW is to find a function $u(\cdot)$ defined on $\mathcal{X}$ that minimizes this quadratic energy with boundary values $u(\cdot)$ defined on $\mathcal{L}$. $u(\cdot)$ can minimize such quadratic energy E when $\nabla^2 u=0$ [14].

The function $u(\cdot)$ can be computed efficiently using matrix operations. Similar to [14], a Laplacian matrix L with size $(K+N)\times(K+N)$ can be constructed as follows:

$$
L_{ij}=\left\{
\begin{array}{ll}
d_i, & i=j \\
-c_{ij}, & \left(v_i,v_j\right)\in\mathcal{E} \\
0, & \text{otherwise}
\end{array}
\right.
\tag{5}
$$

where $d_i=\sum_{v_j\in\mathcal{N}_i}c_{ij}$ is the degree of the node v_i defined on its neighborhood $\mathcal{N}_i$. Then, formula (4) can be rewritten in matrix form as

$$
\begin{aligned}
E &=\left(\begin{array}{c}u_\mathcal{L}\\u_\mathcal{X}\end{array}\right)^T L\left(\begin{array}{c}u_\mathcal{L}\\u_\mathcal{X}\end{array}\right) \\
&=\left(\begin{array}{c}u_\mathcal{L}\\u_\mathcal{X}\end{array}\right)^T\left(\begin{array}{cc}L_\mathcal{L}&B\\B^T&L_\mathcal{X}\end{array}\right)\left(\begin{array}{c}u_\mathcal{L}\\u_\mathcal{X}\end{array}\right)
\end{aligned}
\tag{6}
$$

where $u_\mathcal{L}=\left(u\left(l_1\right),\ldots,u\left(l_K\right)\right)^T$ and $u_\mathcal{X}=\left(u\left(x_1\right),\ldots,u\left(x_N\right)\right)^T$; $L_\mathcal{L}$ is the upper left $K\times K$ submatrix of L that encodes the interactions within $\mathcal{L}$; $L_\mathcal{X}$ is the lower right $N\times N$ submatrix that encodes the interactions within $\mathcal{X}$; and B is the upper right $K\times N$ submatrix that encodes the interactions between $\mathcal{L}$ and $\mathcal{X}$. Hence, the minimum energy solution can be obtained by setting $\nabla E=0$ with respect to $u_\mathcal{X}$, i.e., solving the following equation:

$$
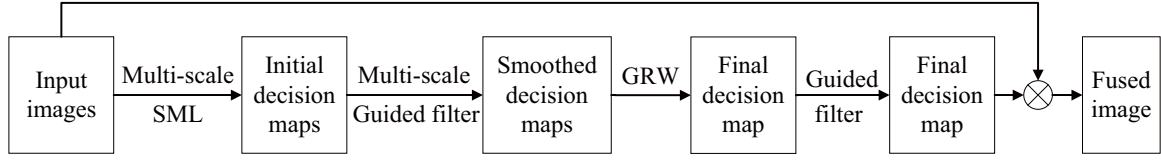L_\mathcal{X}u_\mathcal{X}=-B^T u_\mathcal{L}
\tag{7}
$$

Fig. 3: The schematic diagram of the proposed multi-focus image fusion algorithm.
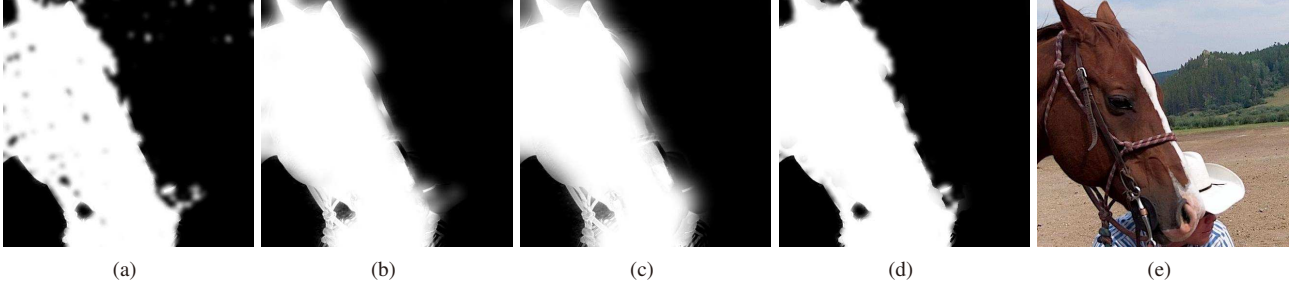

Fig. 4: (a)-(c) The decision maps obtained by three-scale guided filtering. (d) The decision map obtained with GRW. (e) Fused image.

3 Proposed Fusion Method

Fig. 3 shows the schematic diagram of our multi-focus image fusion algorithm. I_1 and I_2 are two input partially focused images. The multi-scale SMLs can be obtained with different window sizes:

$$M_{ij}(p, q) = SML(I_i(p, q), N_j), \quad i = 1, 2 \quad (8)$$

where (p, q) is pixel location, M_{ij} denotes the SML of image I_i in scale j. In this paper, we use the three-scale SMLs (i.e., $j = 3$), which is good enough to produce desired fusion result. We set $N_j = \{5, 11, 17\}$ for $j = \{1, 2, 3\}$ to get SMLs in different scales.

Then, we can get initial decision maps by comparing SMLs of input images:

$$Map_j = \begin{cases} 1, & M_{1j} > M_{2j} \\ 0, & otherwise \end{cases} \quad (9)$$

In order to make the maps accurately align the boundaries between focused and defocused regions, we introduce multi-scale guided filters to smooth these maps:

$$Gmap_j = GuidedFilter(I_1, Map_j, r_j, \varepsilon_j) \quad (10)$$

where $GuidedFilter(I_1, Map_j, r_j, \varepsilon_j)$ denotes the guided filtering [16] in which I_1 serves as the guidance image, Map_j is the input image, r_j and ε_j denote the spatial and range weights, respectively. To align the boundaries effectively, we set $r_j = \{5, 13, 21\}$ and $\varepsilon_j = \{0.01, 0.001, 0.0001\}$ for $j = \{1, 2, 3\}$ [8].

Figs. 4 (a)-(c) show the decision maps obtained by three-scale guided filtering. We can see that the small-scale map (Fig. 4 (a)) can accurately align the boundaries, and distinguish the small defocused (focused) regions encompassed by the large focused (defocused) regions, however, some small focused (defocused) regions in focused (defocused) regions are incorrectly identified as defocused (focused) regions. On the contrary, the small focused (defocused) regions in focused (defocused) regions can be correctly identified in the

large-scale decision maps (Figs. 4 (b)-(c)), while the boundaries are blurred, and the small defocused (focused) regions encompassed by the large focused (defocused) regions cannot be effectively distinguished. Based on this observation, we can conclude that the brightest and darkest regions in Figs. 4 (a)-(c) are usually the desired focused and defocused regions, respectively. Hence, the brightest and darkest regions need to be selected into the final decision map, and other regions can be abandoned. Generalized random walk (GRW) is particularly applicable to this selection problem [13]. As described in Section 2.2, to get $u_{\mathcal{X}}$, we need to define the similarity function ω_{ij} and y_{ik} in Fig. 2.

Similar to [14], we set ω_{ij} as

$$\omega_{ij} = \exp(-\frac{(g_i - g_j)^2}{\sigma^2}), \quad g = I_1 \quad (11)$$

where σ is set as 0.01 in this paper. Since there are three input images (Figs. 4 (a)-(c)) in GRW, $K = 3$ and $k = \{1, 2, 3\}$. y_{ik} is defined as

$$y_{ik} = \begin{cases} Gmap_i^k, & Gmap_i^k > 0.8 \\ 1 - Gmap_i^k, & Gmap_i^k < 0.2 \\ 0, & otherwise \end{cases} \quad (12)$$

where $Gmap_i^k$ denotes the intensity value of the pixel i in the image $Gmap^k$. According to the above analysis, the pixels with the largest or smallest intensity values would be more likely to be selected into final decision map. Thus, when $Gmap_i^k > 0.8$ ($Gmap^k$ is normalized to $[0, 1]$), we set $y_{ik} = Gmap_i^k$. When $Gmap_i^k < 0.2$, we set $y_{ik} = 1 - Gmap_i^k$. When $Gmap_i^k$ is between 0.2 and 0.8, we set $y_{ik} = 0$. The experiments have shown that thresholds 0.2 and 0.8 are suitable to produce good fusion results. Then, by solving formula (7), we can get $u_{\mathcal{X}}$, which is the $N \times K$ matrix (N is the pixel number of $Gmap^k$ and $K = 3$). $u_{\mathcal{X}}(i, k)$ can be seen the probability that the pixel i in input image $Gmap^k$ is selected into final decision map. Thus, we can get final decision map $Fmap$ (Fig. 4 (d)) by

$$Fmap_i = Gmap_i^k, \quad u_{\mathcal{X}}(i, k) > \max_{m, m \neq k} \{u_{\mathcal{X}}(i, m)\} \quad (13)$$

(a) Input image 1   (b) Input image 2   (c) MWGF   (d) BF   (e) Our
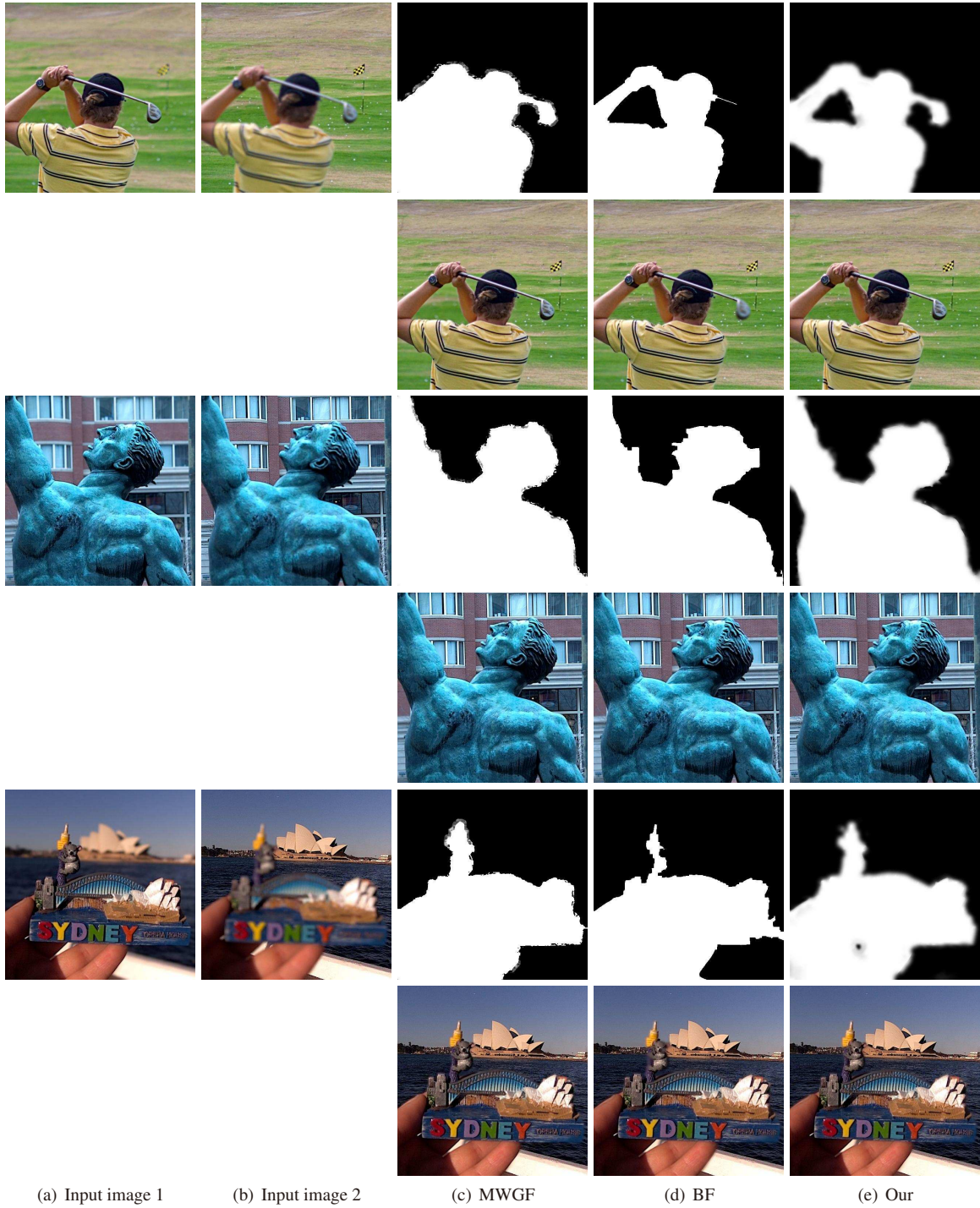
Fig. 5: (a)-(b) The input multi-focus images. (c) The decision maps and fused images obtained with MWGF. (d) The decision maps and fused images obtained with BF. (e) The decision maps and fused images obtained with our method.

In order to make the boundaries between focused and defocused regions in the fused image look more natural, we use the guided filter with small spatial weight to smooth $Fmap$:

$$Fmap = GuidedFilter(I_1, Fmap, r, \varepsilon) \qquad (14)$$

where $r = 3$ and $\varepsilon = 0.01$. Finally, the fused image F (Fig. 4 (e)) can be obtained by

$$F = FmapI_1 + (1 - Fmap) I_2 \qquad (15)$$

4   Experimental Results and Comparisons

The proposed method is compared with two recently proposed methods MWGF [3] and boundary finding based image fusion method (BF) [10] in both subjective and objective assessments. Ten pairs of multi-focus images are used as test images. We objectively evaluate the performance of different fusion methods using four quality metric, including entropy (EN), Piella's metric Q_S, structural similarity based metric Q_C, Chen-Blum metric Q_{CB} [17].

Three pairs of test images are shown in Fig. 5 to compare the performance of these three fusion algorithms. Figs. 5

(a) and (b) are input multi-focus images. Figs. 5 (c)-(e) are the decision maps and their corresponding fused images, obtained with MWGF, BF and our method, respectively. As can be seen from Fig. 5, MWGF and BF cannot effectively distinguish the defocused regions encompassed by focused regions, as a result, their fused images are blurred in these defocused regions. On the contrary, our method is able to accurately identify all the focused and defocused regions. In addition, owing to the multi-scale guided filters and GR-W, the decision maps from our method can accurately align the boundaries between focused and defocused regions. The boundaries in the decision maps from BF are damaged heavily.

Table 1 shows the objective assessments of different fusion methods. The average values of fusion metrics over ten pairs of test images are computed, and the best results are highlighted in bold. It can be seen that our method outperforms the other two fusion methods in terms of these metrics.

Table 1: Objective assessments of different fusion methods

Metric	MWGF	BF	Our
EN	7.5550	7.5519	**7.5555**
Q_S	0.9376	0.9418	**0.9455**
Q_C	0.8183	0.8191	**0.8230**
Q_{CB}	0.7842	0.7953	**0.7957**

5 Conclusion

In this paper, a novel multi-focus image fusion algorithm based on multi-scale focus measures and GRW is proposed. By combining the advantages of multi-scale focus measures, our method is able to effectively distinguish the small defocused (focused) regions encompassed by the large focused (defocused) regions, while it is very difficult for many state-of-the-art multi-focus image fusion methods. In addition, with multi-scale guided filters and GRW, our decision map is able to align the boundaries between focused and defocused regions accurately. Experimental results of both objective and subjective evaluations prove that our method can achieve better performance compared with other fusion methods.

References

[1] G. Pajares, A wavelet-based image fusion tutorial, *Pattern recognition*, 2004, 37(9): 1855-1872.

[2] Van de Plas R, Yang J, Spraggins J, et al. Image fusion of mass spectrometry and microscopy: a multimodality paradigm for molecular tissue mapping. *Nature methods*, 2015, 12(4): 366-372.

[3] Zhou Z, Li S, Wang B. Multi-scale weighted gradient-based fusion for multi-focus images. *Information Fusion*, 2014, 20: 60-72.

[4] Liu Y, Liu S, Wang Z. Multi-focus image fusion with dense SIFT. *Information Fusion*, 2015, 23: 139-155.

[5] Liu Y, Chen X, Peng H, et al. Multi-focus image fusion with a deep convolutional neural network. *Information Fusion*, 2017, 36: 191-207.

[6] P. Burt, E. Adelson, The laplacian pyramid as a compact image code, *IEEE Transactions on communications* 31 (4) (1983) 532-540.

[7] Zhang Q, Guo B. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Processing*, 2009, 89(7): 1334-1346.

[8] Li S, Kang X, Hu J. Image fusion with guided filtering. *IEEE Transactions on Image Processing*, 2013, 22(7): 2864-2875.

[9] Liu Y, Liu S, Wang Z. A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion*, 2015, 24: 147-164.

[10] Zhang Y, Bai X, Wang T. Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure. *Information Fusion*, 2017, 35: 81-101.

[11] Huang W, Jing Z. Evaluation of focus measures in multi-focus image fusion. *Pattern recognition letters*, 2007, 28(4): 493-500.

[12] Nayar S K, Nakagawa Y. Shape from focus. *IEEE Transactions on Pattern analysis and machine intelligence*, 1994, 16(8): 824-831.

[13] Shen R, Cheng I, Shi J, et al. Generalized random walks for fusion of multi-exposure images. *IEEE Transactions on Image Processing*, 2011, 20(12): 3634-3646.

[14] Grady L. Random walks for image segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 2006, 28(11): 1768-1783.

[15] Doyle P G, Snell J L. Random walks and electric networks. *Mathematical Association of America*, 1984.

[16] He K, Sun J, Tang X. Guided image filtering. *IEEE transactions on pattern analysis and machine intelligence*, 2013, 35(6): 1397-1409.

[17] Liu Z, Blasch E, Xue Z, et al. Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: a comparative study. *IEEE transactions on pattern analysis and machine intelligence*, 2012, 34(1): 94-109.