IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2016, accepted August 20, 2016, date of publication August 30, 2016, date of current version September 16, 2016.

Digital Object Identifier 10.1109/ACCESS.2016.2604480

Multi-Focus Image Fusion Through Gradient-Based Decision Map Construction and Mathematical Morphology

XIANGZHI BAI[1,2], MIAOMING LIU[1], ZHIGUO CHEN[1], PENG WANG[1], AND YU ZHANG[1]

[1]Image Processing Center, Beijing University of Aeronautics and Astronautics, Beijing 100191, China
[2]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

Corresponding author: X. Bai (jackybxz@buaa.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61271023 in part by the Program for New Century Excellent Talents in Universities under Grant NCET-13-0020 in part by the Fundamental Research Funds for the Central Universities under Grant YWF-16-BJ-Y-28.

ABSTRACT Multi-focus image fusion is an important technique that extracts sharp regions from multiple images and composites them into a fully focused image. In this paper, a novel spatial domain-based fusion algorithm for multi-focus images through gradient-based decision map construction and mathematical morphology is proposed. The contributions of this paper are: 1) a weighted kernel based on image gradient is proposed to measure focus regions; 2) the boundaries between focus and defocus regions are adjusted by morphological operations and free boundary condition-based active contour model. Though the weighted kernel, the focus regions are roughly identified. Moreover, after the morphological operations and the adjustment of boundaries using active contour model, the true boundaries between the focused and defocused regions are extracted. The experimental results demonstrate that the proposed algorithm performs better than the other eight representative fusion algorithms in both the qualitative and quantitative evaluations.

INDEX TERMS Active contour model, decision map, mathematical morphology, multi-focus image fusion.

I. INTRODUCTION

IMAGE fusion is an important technique to extract as much as possible relevant information for image analysis applications, such as surveillance, target tracking, target detection and face recognition [1], [2]. Multi-focus image fusion is one branch of the image fusion, which combines all the focused scenes in the source images into a single fused image and enriches the image information. And the fused image is more suitable for human or machine perception. Multi-focus image fusion plays an important role in many applications, such as object recognition, feature extraction and segmentation [3], [4]. Recently, a variety of pixel-level fusion algorithms have been proposed for multi-focus image fusion, and these algorithms work in the transform domain or spatial domain [1], [5], [6].

The frequently-used transform domain based image fusion algorithms include principal component analysis (PCA) based fusion algorithms [7], pyramid based fusion algorithms [8]–[10] and wavelet based fusion algorithms [11], [12]. The transform domain based image fusion algorithms produce the fused image in a global way. Hence, there may be some mis-registrated regions in the results. Especially, when the contents in the same position of the source images are different, the fusion results would be unsatisfactory.

In recent years, spatial domain based image fusion algorithms have raised the attention of many researchers because they are intuitive and simple. The spatial domain based fusion algorithms mainly include the pixel based fusion algorithms [13], [14], the block based fusion algorithms [3], [15] and the region based fusion algorithms [16], [17]. The earliest and simplest pixel based fusion algorithm directly calculates the average of every corresponding pixel value of the source images. However, the direct average calculation cannot remove the effect of defocused pixels and this method usually comes to a bad fusion result. The idea of the block based or the region based fusion algorithms is to extract image blocks from the source images and composite them into a single image. However, the block based algorithms would not perform well if the extracted blocks contain both clear and blurred areas.

In this paper, we propose a novel spatial domain based fusion algorithm for multi-focus image fusion through gradient based decision map construction and mathematical morphology. The contributions of this paper are: (1) a weighted kernel based on image gradient is proposed to

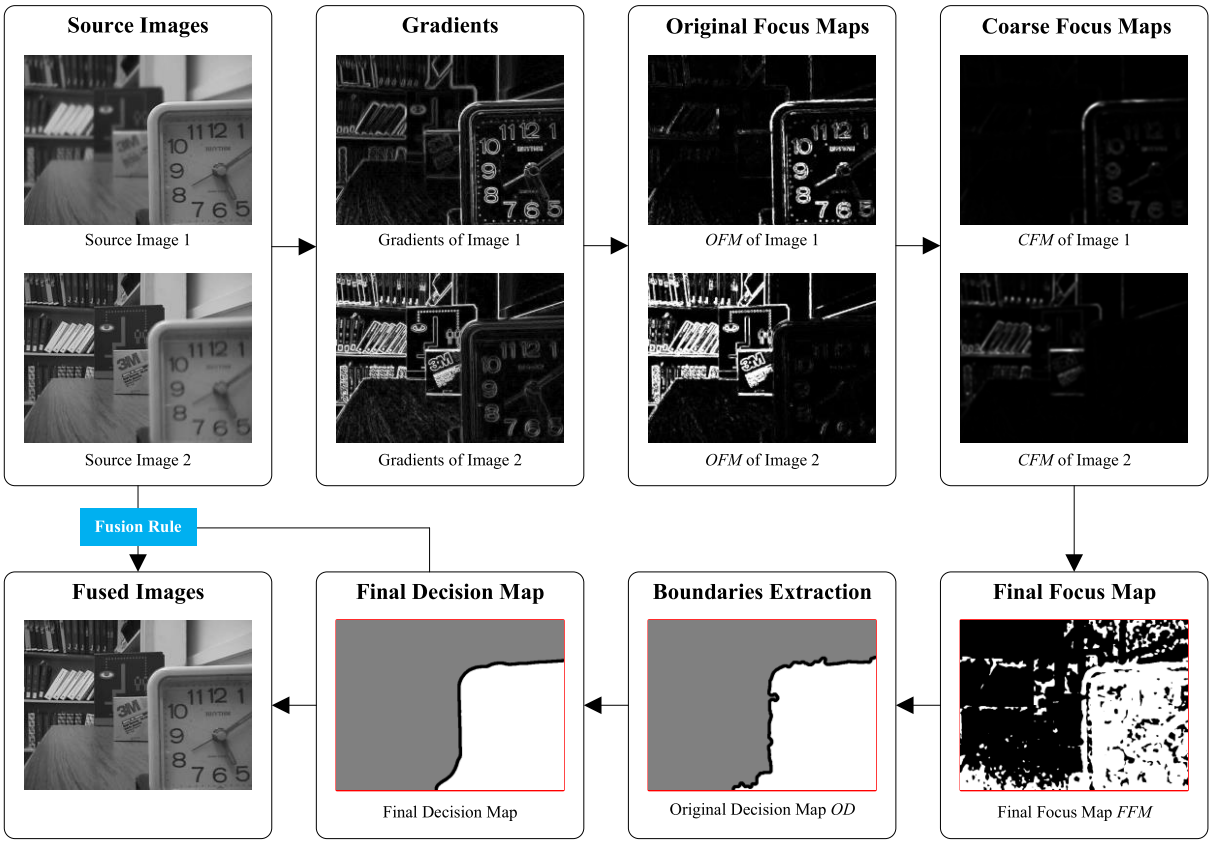


FIGURE 1. Scheme chart of the proposed algorithm. The boundaries in the final decision map and the original decision map are in bold for display.

measure focus regions; (2) the boundaries between focus and defocus regions are adjusted by morphological operations and free boundary condition based active contour model. Firstly, a weighted kernel is employed to filter the original focus maps and the coarse focus map is calculated, which produces the final focus map. Secondly, morphological operations and free boundary condition based active contour model are employed to obtain the exact boundaries. Finally, the fused image is obtained based on the final decision map and the fusion rule.

To demonstrate the performance of the proposed algorithm, experiments have been performed on the commonly used image sets. Moreover, the proposed algorithm is compared with eight representative fusion algorithms. And the results are evaluated in both the qualitative and quantitative ways, which indicates that the proposed algorithm performs better.

The rest of this paper is organized as follows. In Section II, the details of the proposed method are presented. Section III presents the experiment results and discussions. Finally, we conclude the paper in Section IV.

II. THE PROPOSED METHOD

In this paper, a novel multi-focus image fusion algorithm through gradient based decision map construction and mathematical morphology is proposed. The scheme chart of the proposed algorithm is shown in Fig. 1.

A. ORIGINAL FOCUS MAP

The gradient of an image indicates the sharpness information. For a certain region in the grayscale images, the change of gray value in focus region is usually greater than the change of gray value in defocus region. And the gradient represents the change of gray value. Thus, the gradient could be used to indicate the focus status. The gradient of every pixel in the focus region of the source image is usually greater than the gradient of corresponding pixel in the defocus region of other source images. Hence, the gradient maps are used to construct the original focus map in this subsection.

Firstly, the gradients $G_i(x, y)$ of the source images $f_i(x, y)$ $(i = 1, 2)$ are calculated as follows:

$$G_i(x, y) = \nabla f_i(x, y), \tag{1}$$

where (x, y) is the pixel coordinate of the image. The Sobel gradient operator [20] is used in our experiment.

Secondly, the differences between the gray value of every pixel (x, y) in the gradient map $G_i(x, y)$ and the gray value of the 3×3 neighborhood pixels are calculated. The original focus map $OFM_i(x, y)$ is defined as the quadratic sum of the differences, formulated as

$$OFM_i(x, y) = \sum_{(p,q)} (G_i(x, y) - G_i(x + p, y + q))^2, \quad p, q \in \{-1, 0, 1\}. \tag{2}$$

The gradient differences between a pixel and its neighbouring pixels represent the sharpness of the images. And the quadratic sum of the differences on the gradients could enlarge the differences between a pixel and its neighbouring pixels. This would be benefit for identifying focus and defocus regions. Thus, it would be effective to employ the quadratic sum of the gradient differences between one pixel and its eight neighbourhoods as the original focus map $OFM_i(x, y)$ to express the sharpness information.

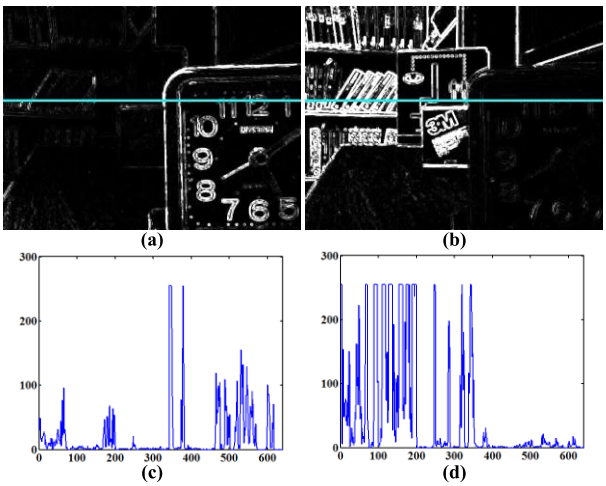


FIGURE 2. (a) The original focus map of the multi-focus image "Desk" focused on close shot. (b) The original focus map of multi-focus image "Desk" focused on distant view. (c) The gray value of pixels upon the cyan line in (a). (d) The gray value of pixels upon the cyan line in (b).

The original focus map $OFM_i(x, y)$ of multi-focus images "Desk" are presented in Fig. 2 (a) and (b), respectively. The gray values of pixels along the cyan line in Fig. 2 (a) and (b) are depicted in Fig. 2 (c) and (d). From Fig. 2 (a) and (c), we can find that most of the gray values of the focus region are much larger than that of the defocus region. Fig. 2 (b) and (d) shows that most of the pixels in the focus region have much larger gray values than the pixels in the defocus region, such as the books in the bookshelf. Compared with the gradients in Fig. 1, it could be observed that the gradient differences have been enlarged, which is in favour of identifying focus and defocus regions.

B. FINAL FOCUS MAP

1) COARSE FOCUS MAP

As the original focus map $OFM_i(x, y)$ is calculated based on the eight neighborhood pixels, $OFM_i(x, y)$ only represents the local information of the source images. Moreover, the quadratic sum of the differences between the gradient values of the pixels in the focus region and their eight-neighborhood pixels are very likely to be smaller than the quadratic sum of the gradient value differences between the corresponding pixels in the defocus region and their eight-neighborhood pixels. Then the $OFM_i(x, y)$ would fail to estimate the focus distribution. In order to have a more accurate estimation of the focus distribution and avoid the small isolate regions, a simple weighted kernel is constructed to compute the coarse focus map $CFM_i(x, y)$. The expression of coarse focus map $CFM_i(x, y)$ is formulated as

$$CFM_i(x, y) = \sum_{s=-k}^{k} \sum_{t=-k}^{k} w(x, y)\, OFM_i(x + s, y + t), \quad (3)$$

where $w(x, y)$ is the weight of pixel (x, y), which is defined as $w(x, y) = w^{(k)}(x, y)$, $k = 10$. The size of the weighted kernel is $(2k + 1) \times (2k + 1)$. In our experiment, k is set to 10. The value of the weighted kernel is generated by (4).

The weight kernel we constructed has the following characteristics:

(1) The weight of the center pixel is equal or greater than the other weights around the center. And if $k = 1$, the weights of the kernel are equal to 1. For $k > 1$, the weight kernel is represented as

$$w^{(k)}(x, y) = \sum_{s=-1}^{1} \sum_{t=-1}^{1} w^{(k-1)}(x + s, y + t), 1 < k \leq 10. \quad (4)$$

(2) The larger the distance between the pixel $w(x, y)$ and the center pixel is, the smaller weight of this pixel $w(x, y)$ is.

2) FINAL FOCUS MAP

The coarse focus maps could roughly represent the sharpness of the source images. And, in order to combine the focus status into one map, we construct the final focus map $FFM(x, y)$. For pixel (x, y) in two multi-focus images of the source image set, if $CFM_1(x, y) > CFM_2(x, y)$, then $FFM(x, y)$ equals to 1; otherwise, $FFM(x, y)$ equals to 0. Hence, the final focus map $FFM(x, y)$ is defined as

$$FFM(x, y) = \begin{cases} 1, & CFM_1(x, y) > CFM_2(x, y) \\ 0, & otherwise. \end{cases} \quad (5)$$

The final focus map $FFM(x, y)$ could represent the sharpness of every pixel in the source images. For a certain region in the image sets, if the value of the coarse focus map $CFM_i(x, y)$ is larger than the values of the other coarse focus maps, then this region in image $f_i(x, y)$ is focused; otherwise, this region in $f_i(x, y)$ is defocused.

C. MORPHOLOGICAL PROCESSING

The focus regions might contain some defocused patches, likewise, the defocus regions might contain some focused patches. To deal with these situations, the morphological small object removing operation is employed to remove the noise within the object region, which is based on opening and closing operations [18]. Moreover, the boundaries between focus and defocus regions may be winding and far away from the true boundaries. To make the boundaries become much closer to the true boundaries, morphological opening-and-closing operation [19] and small object removing operation are alternately employed to process the $FFM(x, y)$ after removing small patches.

1) SMALL PATCH REMOVING OPERATION

In this paper, if the pixel number of a certain region is smaller than NUM, then this region is regarded as small patches. And this region will be processed by morphological small object removing operation, expressed as

$$OD_1(x, y) = RSO(FFM(x, y), NUM), \tag{6}$$

where $OD_1(x, y)$ is the first stage of the original decision map; $RSO(\bullet)$ is the morphological small object removing operation [18]; NUM represents the maximum pixel number of the small patches removed. The parameter NUM is related to the pixel number of the source images. In our experiment, $NUM = [m \times n/40]$, where m, n is the height and width of the image $f_i(x, y)$, respectively; $[\cdot]$ is rounding operation.

2) MORPHOLOGICAL OPENING OPERATION AND SMALL PATCH REMOVING OPERATION

In this stage, morphological opening operation and small object removing operation are applied to the first stage of the original decision map $OD_1(x, y)$ with structuring element B, and this procedure is expressed as

$$OD_2(x, y) = RSO(OD_1(x, y) \circ B, NUM), \tag{7}$$

where $OD_2(x, y)$ is the second stage of the original decision map; B is a flat disk structuring element; $\circ$ is morphological opening operator [19].

After the morphological opening operation, there would be some small patches in the focus regions those are wrongly regarded as defocused. So the small object removing operation is necessary. After the removing operation, the small patches in focus regions would be eliminated.

3) MORPHOLOGICAL CLOSING OPERATION AND SMALL PATCH REMOVING OPERATION

Symmetry operations are performed in the defocus regions, which is expressed as

$$OD(x, y) = RSO(OD_2(x, y) \bullet B, NUM), \tag{8}$$

where $OD(x, y)$ is the original decision map; B is a flat disk structuring element; $\bullet$ is morphological closing operator [19].

In this stage, we could remove the small patches those are produced by morphological closing operation and wrongly regarded as defocused in the focus regions.

In order to better illustrate the procedure of morphological processing, an example is shown in Fig. 3. Fig. 3 (a) is the final focus map. Fig. 3 (b) is the first stage of the original decision map. Fig. 3 (c) is the second stage of the original focus map. Fig. 3 (d) is the original decision map. As depicted in Fig. 3 (a), we can find that there are some small patches in the final focus map. Though the small patch removing operation, almost all of them have been removed as shown in Fig. 3 (b). As shown in Fig. 3 (c), after processing by morphological opening operation and small object removing operation, the boundary is smoother. From Fig. 3 (d), the boundary becomes much closer to the true boundary though

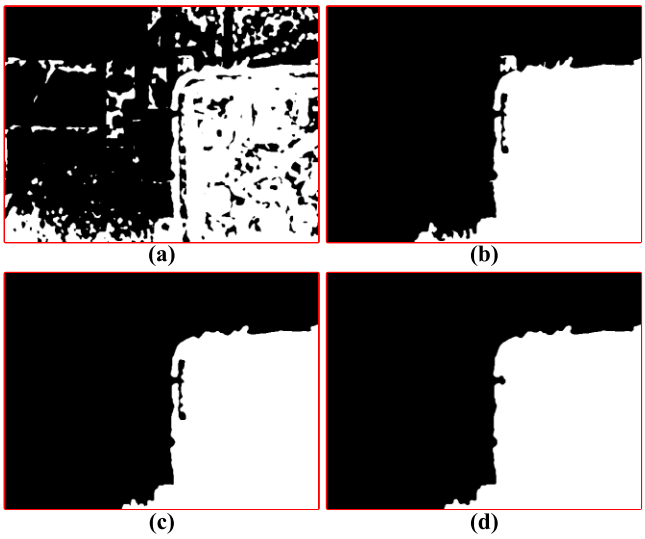


FIGURE 3. An example of morphological processing. (a) The final focus map. (b) The first stage of the original decision map. (c) The second stage of the original decision map. (d) The original decision map.

morphological closing operation and small object removing operation. This would benefit the multi-focus image fusion.

D. BOUNDARIES EXTRACTION

In this subsection, the boundaries between focus regions and defocus regions are extracted from the original decision map $OD(x, y)$.

Firstly, the boundaries are extracted from the original decision map $OD(x, y)$ though Sobel operator [20]. Secondly, the morphological bridging operation [21] is used to connect the gap interrupted by single pixel. Finally, the morphological thining operation [22] is applied to obtain the boundaries $L_j(j = 1, \ldots, l)$ whose width is single pixel.

E. FINAL DECISION MAP

In general, the extracted boundaries $L_j(j = 1, \ldots, l)$ cannot exactly represent the true boundaries between the focus and defocus regions. It is necessary to adjust the boundaries $L_j(j = 1, \ldots, l)$ for obtaining better results. Free boundary condition based active contour model [23] constructs the energy function depending on the image gradient, which is conformed to the characteristic of the focus and defocus region. And it would be effective to make the boundaries get closer to the true boundaries when using $L_j(j = 1, \ldots, l)$ as the initial contours instead of using the random initial contours.

The energy of active contour model is usually formulated as

$$E(C(s)) = \int [E_{int}(C(s)) + E_{ext}(C(s))]ds, \tag{9}$$

where $E_{int}(C(s))$ is internal energy term, $E_{ext}(C(s))$ is external energy term, $C(s)$ represents a differentiable curve in the image $I(x, y)$. The internal energy term controls the continuity and smoothness of the evolutionary curve while

the external energy term adjusts the curve to be consistent with image features. The energy function can be expressed more explicitly as [23], [24]

$$E\left(C\left(s\right)\right)=\int_{0}^{1}\left[\left|C'\left(s\right)\right|^{2}-\left|\nabla\left(G_{\sigma_0}*I\left(C\left(s\right)\right)\right)\right|\right]ds, \quad (10)$$

where G_{σ_0} signifies a Gaussian filter with standard deviation $\sigma_0=2$.

The procedure of searching the exact boundaries is described as follows.

Firstly, two end points P_0, P_1 of every active contour and two corresponding smooth boundary curves $B_0\left(q\right)$, $B_1\left(q\right)$ are defined. The two end points P_0 and P_1 of the active contour are defined by the two ends of the extracted boundary L_j, which are constrained lying on two smooth boundary curves $B_0\left(q\right)$ and $B_1\left(q\right)$, respectively. In our data, the end points always locate at the image boundaries. This makes the definition of the smoothed boundary curves much simpler. As a result, the smoothed boundary curve $B_0\left(q\right)$ is defined by a section of the image boundary with the end point as the center. We suppose that the boundaries we extracted are close to the true boundaries. The definition of another smoothed boundary curve $B_1\left(q\right)$ is similar to the definition of $B_0\left(q\right)$.

Secondly, the extracted boundaries $L_j(j=1,\ldots,l)$ are used as the initial contour of the active contour model. Then, the final boundary map L_{final} could be obtained by the active contour model.

Finally, the value of every region in the final decision map D_{final} is decided according to the original decision map $OD\left(x,y\right)$ and the final boundary map L_{final}. For a certain region R in the final boundary map L_{final}, if its value in the original decision map $OD\left(x,y\right)$ is 1, then its value in the final boundary map L_{final} equals to 1. That is, the value of the corresponding region in the final decision map D_{final} equals to 1 and the corresponding region in image $f_1\left(x,y\right)$ is copied to the region R. For a certain region R, if its value in the original decision map $OD\left(x,y\right)$ is 0, then its value in the final boundary map L_{final} equals to 0. This means the value of the corresponding region in the final decision map D_{final} equals to 0 and the corresponding region in image $f_2\left(x,y\right)$ is copied to the region R. Though this way, the final decision map D_{final} is obtained.

F. FUSION IMAGE

There might be some flaws in the boundary region if we directly copy the regions in the source images to the corresponding regions. These flaws are caused by the abrupt change of the gray values in the boundary region. So it is necessary to establish a fusion rule to fuse the source images. For simplicity, the final decision map D_{final} is smoothed by a Gaussian filter to make the fused image smooth in the boundary region. Then, the fused image $f\left(x,y\right)$ could be expressed as

$$f\left(x,y\right)=D_{finalG}\times f_1\left(x,y\right)+\left(1-D_{finalG}\right)\times f_2\left(x,y\right). \quad (11)$$

And (11) means the fused image is composed by the focus regions in the image $f_1\left(x,y\right)$ and $f_2\left(x,y\right)$. Though these steps, a fused image fully focused could be obtained.

III. EXPERIMENT RESULTS AND DISCUSSIONS

In this section, we conduct experiments in eighteen sets of multi-focus images to verify the performance of the proposed algorithm through the comparisons with the representative fusion algorithms in both the qualitative and quantitative ways. The experiment settings and experiment results are introduced as follows.

A. EXPERIMENT SETTINGS

The experiments are conducted in eighteen sets of multi-focus images, in which the ''Clock'', ''Lab'', ''Pepsi'' and ''Desk'' are obtained from [25], the ''OpenGL'', ''Flower'', ''Seascape'' and ''Temple'' are provided by [26], the ''Leopard'', ''Wine'', ''Balloon'', ''Calendar'', ''Corner'', ''Craft'', ''Leaf'', ''Newspaper'', ''Girl'' and ''Grass'' are acquired from [27]. And we appreciate them for providing the source images.

To verify the effectiveness of our algorithm, we perform the comparisons with eight representative image fusion algorithms in both the qualitative and quantitative aspects. These eight comparison algorithms include the discrete wavelet transform based fusion algorithm (DWT) [11], the shift invariant discrete wavelet transform based fusion algorithm (SIDWT) [27], the multi-scale transform and sparse representation based fusion algorithm (MST-SR) [29], the image matting based fusion algorithm (IMF) [30], the guided filter based fusion algorithm (GFF) [31], the multi-scale weighted gradient based fusion algorithm (MWGF) [32], the dense scale invariant feature transform based fusion algorithm (DSIFT) [33] and the quadtree based fusion algorithm (QUADTREE) [34].

The decomposition levels of DWT, SIDWT and MST-SR are 4. And the parameters of the other algorithms are set as the recommended values which are presented in the respective publications [11], [28]–[34]. The qualitative and quantitative evaluations are described in the following two subsections.

B. QUALITATIVE EVALUATION

Qualitative evaluation, i.e., visual comparison, is the simplest way to assess the performance of fusion algorithms. Some comparison examples are shown below. The comparison results of ''Desk'' among different fusion algorithms are shown in Fig. 4. And the enlarged images of the labeled regions in Fig. 4 are shown in Fig. 5. From Fig. 4 (a) and (b), we can see that DWT and SIDWT can achieve the fusion purpose. However, there are some blurs in their fusion results, such as the edges around the books in the bookshelf and the number on the clock. From Fig. 4 (d), we can find that the boundary between the focus and defocus regions of the source images is a little blur, which is caused by the decision map obtained by image matting. In Fig. 4 (c) and (e)-(h), the fused results obtained by MST-SR, GFF, MWGF, QUADTREE

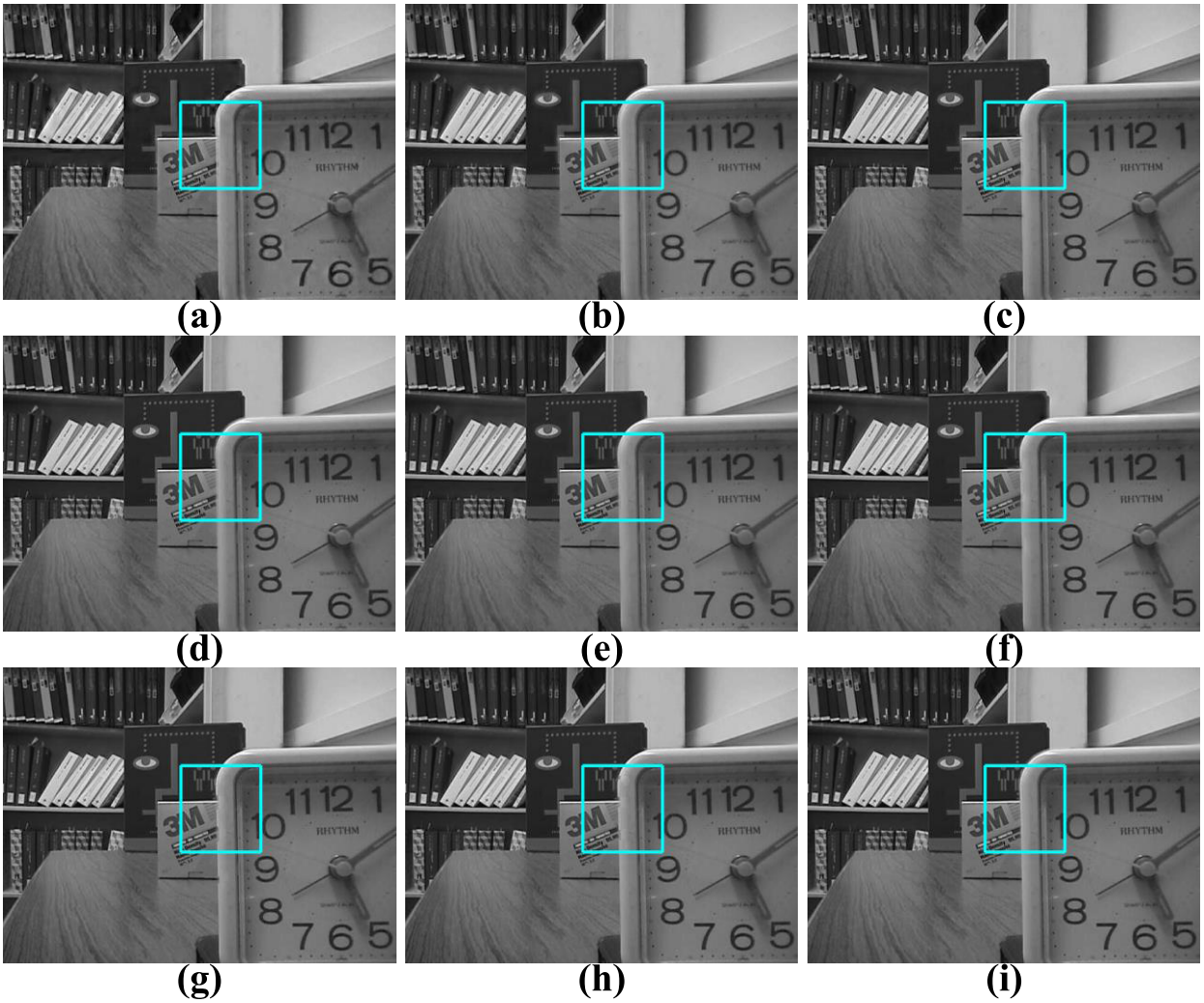


FIGURE 4. Fused images of the "Desk" image set. (a)-(i) are the fused images of DWT [11], SIDWT [28], MST-SR [29], IMF [30], GFF [31], MWGF [32], DSIFT [33], QUADTREE [34] and the proposed algorithm, respectively.

and DSIFT suffer from some artifacts around the boundary such as the labeled regions in Fig. 5. However, as shown in Fig. 4 (i), the fused result yielded by the proposed algorithm provides intact edges and obtains satisfying visual effect. Moreover, from Fig. 5, we can clearly see that the proposed algorithm obtains the best visual effect among these algorithms, which implies that the proposed algorithm obtains a satisfying result for image set "Desk".

The comparison results of "Temple" among different fusion algorithms are shown in Fig. 6. And the enlarged images of the labeled regions in Fig. 6 are shown in Fig. 7. From Fig. 6 (a)-(c), the fused results of DWT and SIDWT suffer from serious artifacts such as the text region of the left side of "stone lion" and the region above the text. This is because DWT and SIDWT use the global way to fuse images. As shown in Fig. 6 (c), the fused result of MST-SR generates less artifacts than the results of DWT and SIDWT, because MST-SR combines the merits of the MST and SR. For the results of IMF, GFF and MWGF in Fig. 6 (d)-(f), the artifacts are less. But there still exist obvious artifacts around the text. And according to Fig. 6 (g)-(i), QUADTREE, DSIFT and the proposed algorithm generate less artifacts than DWT, SIDWT, MST-SR, IMF, GFF and MWGF. This means, these three algorithms get better results. From Fig. 7 (a)-(c), we can clearly see that the artifacts in the enlarged patches getting from the results of DWT, SIDWT and MST-SR are serious. In Fig. 7 (d)-(f), the regions around the boundary in the fused images of IMF, GFF and MWGF are blurry. As shown in Fig. 7 (g)-(h), the regions around the boundary in the fused images of DISFT and QUADTREE introduce artifacts and are not smooth. However, the boundary in the fused image of the proposed algorithm is clear and smooth.

The comparison results of "Clock" among different fusion algorithms are shown in Fig. 8. In Fig. 8 (a), the fused result of DWT suffers from some artifacts. As for the results of SIDWT and MST-SR in Fig. 8 (b) and (c), we can find that there are some blurs on the numbers of clocks and shadows around the numbers of the clocks especially in the back clock.

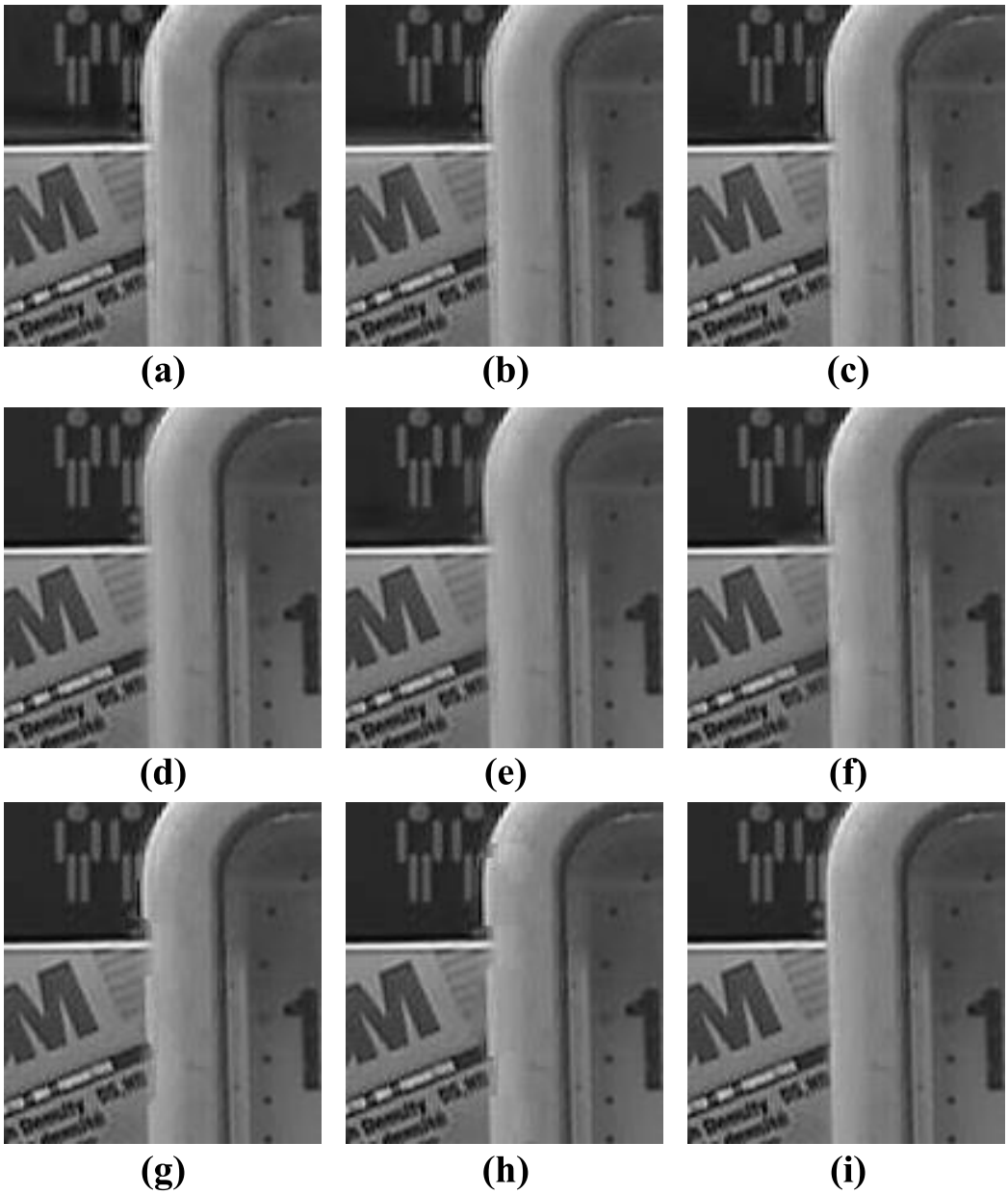


FIGURE 5. (a)-(i) the enlarged images of the labeled regions in Fig. 4, respectively.

The results of GFF, MWGF, DSIFT and QUADTREE introduce artifacts around the string ''8'' of the back clock as shown in Fig. 8(e)-(h). In Fig. 8 (d) and (i), the results of IMF and the proposed algorithm suffer from the least artifacts around the string ''8''.

Overall, the performance of the proposed algorithm is better than the other compared algorithms in visual aspect.

C. QUANTITATIVE EVALUATION

Since it is impossible to get the ground truth of the fused image in practice, metrics without reference image have been proposed to assess the performance of the fusion algorithms [35]–[43]. Five metrics are used to assess the fusion images: fast feature mutual information metric Q_{FFMI} [35], [36], Tsallis entropy based metric Q_{TE} [37], [38], nonlinear correlation information entropy metric Q_{NCIE} [39], [40], phase congruency metric Q_{PC} [41], and modified structural similarity metric Q_{MSSIM} [42].

1) FAST FEATURE MUTUAL INFORMATION METRIC Q_{FFMI}

Q_{FFMI} is defined as [35], [36]

$$Q_{FFMI} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{I_i(f,f_1)}{H_i(f)+H_i(f_1)} + \frac{I_i(f,f_2)}{H_i(f)+H_i(f_2)}\right), \tag{12}$$

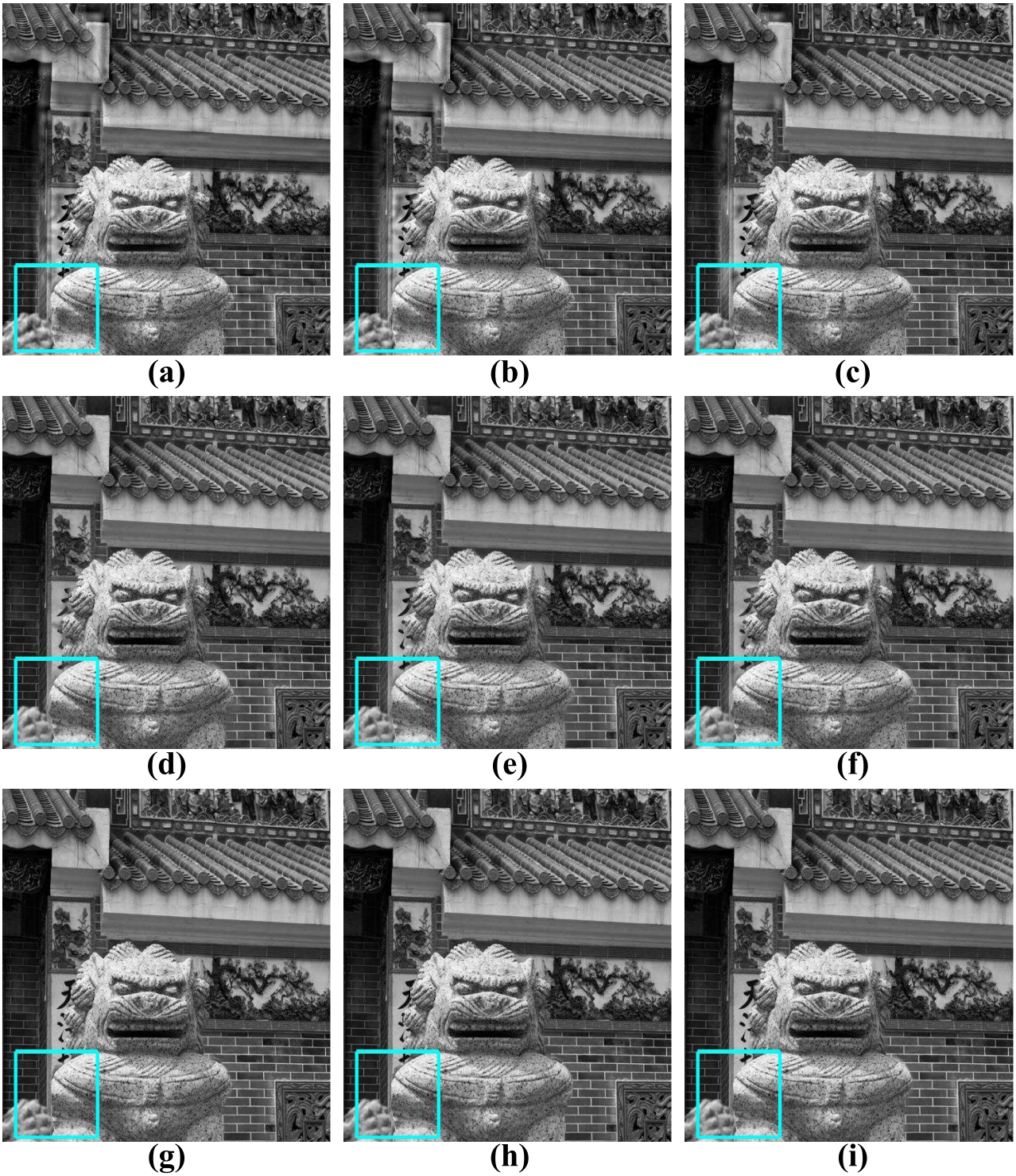


FIGURE 6. Fused images of the "Temple" image set by the compared algorithms. (a)-(i) are the fused image of DWT [11], SIDWT [28], MST-SR [29], IMF [30], GFF [31], MWGF [32], DSIFT [33], QUADTREE [34] and the proposed algorithm, respectively.

where n is the number of slide windows; $H_i(X)$ is the entropy of the ith windows from image X; $I_i(X, Y)$ is the regional mutual information between the ith windows of images X and Y. Q_{FFMI} evaluates how much of edge information the fused image obtained from the source images. The larger the value of Q_{FFMI} is, the better the quality of the fused image is.

2) Tsallis ENTROPY BASED METRIC Q_{TE}

Q_{TE} is defined as [37], [38]

$$Q_{TE} = \frac{I^q(f, f_1) + I^q(f, f_2)}{H^q(f_1) + H^q(f_2) - I^q(f_1, f_2)}. \tag{13}$$

where $H^q(X)$ is the marginal entropy of image X; $I^q(X, Y)$ is the Tsallis entropy between images X and Y;

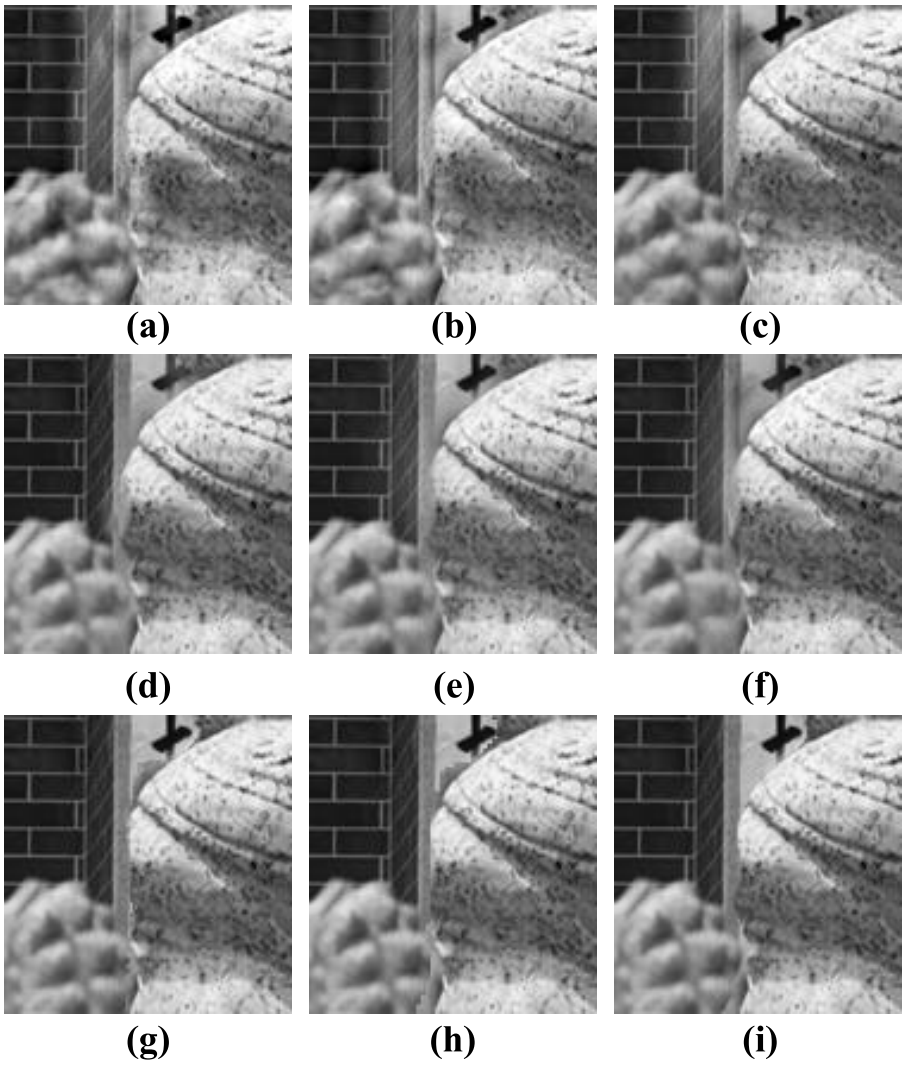


FIGURE 7. (a)-(i) the enlarged images of the labeled regions in Fig. 6, respectively.

order $q \neq 1$. Q_{TE} measures the amount of information transferred from the source images to the fused image. The larger value indicates the better performance of the fusion algorithm.

3) NONLINEAR CORRELATION INFORMATION ENTROPY METRIC Q_{NCIE}

Q_{NCIE} is defined as [39], [40]

$$Q_{NCIE} = 1 + \sum_{i=1}^{3} \frac{\lambda_i}{3} \log_b \frac{\lambda_i}{3}, \tag{14}$$

where b is the intensity level, i.e., $b = 256$; λ_i $(i = 1, 2, 3)$ are the eigenvalue of the nonlinear correlation matrix. Q_{NCIE} calculates the general correlation degrees between the source images and the fused image, respectively. The larger value of Q_{NCIE} indicates the better performance of the fusion algorithm.

4) PHASE CONGRUENCY METRIC Q_{PC}

Q_{PC} is formulated as [41]

$$Q_{PC} = (P_p)^{\alpha} (P_M)^{\beta} (P_m)^{\gamma}, \tag{15}$$

where p, M and m signify phase congruency, maximum, and minimum moments, respectively. Q_{PC} measures the extent that the salient features of the source images are preserved. The larger value means the better performance of the fusion algorithm.

5) MODIFIED STRUCTURAL SIMILARITY METRIC Q_{MSSIM}

Q_{MSSIM} is defined as [42]

$$Q_{MSSIM} = \begin{cases} \lambda(\omega) SSIM(f_1, f|\omega) + (1 - \lambda(\omega)) SSIM(f_2, f|\omega), \\ \quad if\ SSIM\ (f_1, f_2|\omega) \geq 0.75 \\ \max\{SSIM\ (f_1, f|\omega)\ , SSIM\ (f_2, f|\omega)\}, \\ \quad if\ SSIM\ (f_1, f_2|\omega) < 0.75, \end{cases} \tag{16}$$

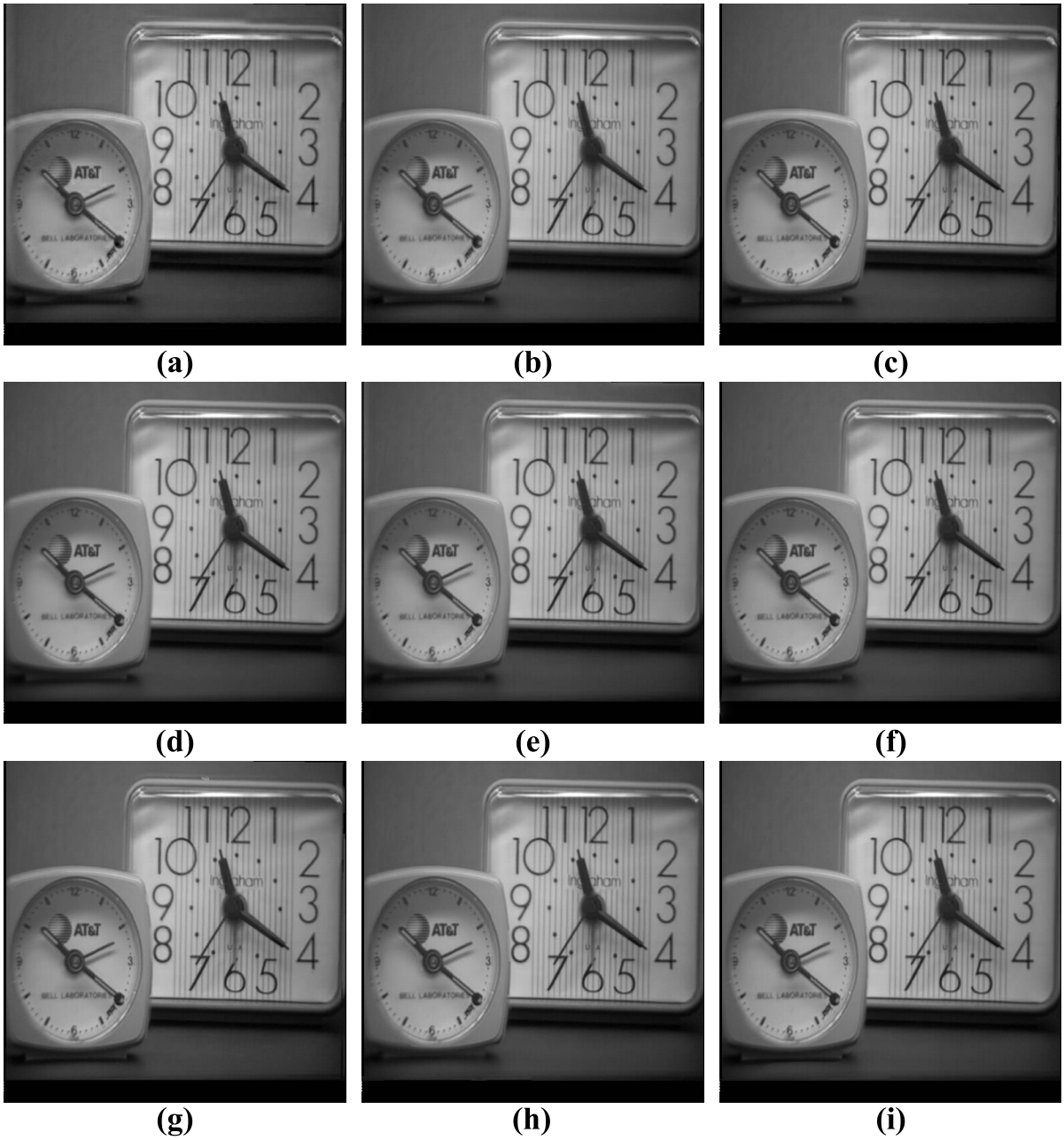


FIGURE 8. Fused images of the "Clock" image set by the compared algorithms. (a)-(i) are the fused image of DWT [11], SIDWT [28], MST-SR [29], IMF [30], GFF [31], MWGF [32], DSIFT [33], QUADTREE [34] and the proposed algorithm, respectively.

where *SSIM* is the structural similarity, $\lambda(\omega)$ is the local weight [42]. Q_{MSSIM} measures the amount of structural information preserved in the fused image. The larger value demonstrates the better performance of the fusion algorithm.

The averages of five evaluation metrics of each algorithm on eighteen multi-focus image sets are presented in Table 1. The largest number is shown in bold. From metric Q_{FFMI} in the first row of Table 1, we can find that the fused images of DWT, SIDWT, MST-SR, IMF, GFF, MWGF, DISIFT and QUADTREE obtain less edge information from the source images. In other words, the proposed algorithm gets the maximum value of metric Q_{FFMI}, and the amount of edge information in the fused images generated by the proposed algorithm is the most. As to metric Q_{TE} shown in the second row of Table 1, the fused images of the proposed algorithm inherits more information from the source images, which manifests the proposed algorithm performs better than the other eight algorithms. As to the metric Q_{NCIE} in the third row of Table 1, DWT, SIDWT, MST-SR, IMF, GFF, MWGF, DISIFT and QUADTREE obtain smaller values than the proposed

TABLE 1. Quantitative evaluation results.

Indices	DWT [11]	SIDWT [28]	MST-SR [29]	IMF [30]	GFF [31]	MWGF [32]	DSIFT [33]	QUADTREE [34]	Proposed algorithm
Q_{FFMI}	0.8576	0.8581	0.8670	0.8694	0.8691	0.8689	0.8696	0.8696	**0.8700**
Q_{TE}	0.3719	0.3803	0.3934	0.4178	0.4081	0.4108	0.4165	0.4198	**0.4201**
Q_{NCIE}	0.8220	0.8237	0.8287	0.8424	0.8356	0.8376	0.8438	0.8448	**0.8456**
Q_{PC}	0.6768	0.7169	0.7681	0.8103	0.8103	0.8101	0.8152	0.8119	**0.8185**
Q_{MSSIM}	0.8934	0.9221	0.9417	0.9829	0.9691	0.9770	0.9795	0.9850	**0.9880**

algorithm, which implies that the fused images of the proposed algorithm have stronger relationship with the corresponding source images. As for the metric Q_{PC} in the fourth row of Table 1, the amount of the preserved salient features of the fused images generated by the proposed algorithm is the most among these nine algorithms, which implies that the proposed algorithm outperforms the other eight algorithms. As shown in the last row of Table 1, the metric Q_{MSSIM} value of the proposed algorithm is higher than the other eight algorithms, which indicates that the proposed algorithm preserves more structural information than the other eight algorithms. This is mainly because the proposed algorithm could find the real boundaries and fuse the images more accurately.

The data set of multi-focus images includes sky, grass, seascape, temple, wall and other outdoor scenes. Also, the data set contains book, calendar, newspaper, clock, desk, lab and other indoor scenes. Moreover, the data set contains images with different contents in the same position. In a word, the fusion task of this data set is challenging. The post-processing morphological operations could deal with these various multi-focus images and the results are effective, which indicates the effectiveness of the morphological operations. All of these indicate that the proposed algorithm is effective for different types of multi-focus image fusion.

In conclusion, the proposed algorithm performs the best among these nine algorithms for multi-focus image fusion.

IV. CONCLUSIONS

This paper presents a novel algorithm for multi-focus image fusion through gradient based decision map construction and mathematical morphology. The contributions of this paper are: (1) a weighted kernel based on image gradient is proposed to measure focus regions; (2) the boundaries between focus and defocus regions are adjusted by morphological operations and free boundary condition based active contour model. From the qualitative and quantitative comparisons, it can be seen that the proposed algorithm is effective for multi-focus image fusion and it performs better than other eight representative fusion algorithms.

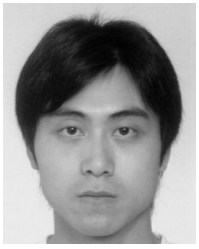
ACKNOWLEDGMENT

The part of this paper was presented at the International Conference [44].

REFERENCES

[1] A. A. Goshtasby and S. Nikolov, "Image fusion: Advances in the state of the art," *Inf. Fusion*, vol. 8, no. 2, pp. 114–118, 2007.
[2] Y. Xu, J. Wen, L. Fei, and Z. Zhang, "Review of video and image defogging algorithms and related studies on image restoration and enhancement," *IEEE Access*, vol. 4, pp. 165–188, 2016.
[3] 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, 2001.
[4] I. De, B. Chanda, and B. Chattopadhyay, "Enhancing effective depth-of-field by image fusion using mathematical morphology," *Image Vis. Comput.*, vol. 24, no. 12, pp. 1278–1287, 2006.
[5] 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.
[6] Y. Yang, S. Tong, S. Huang, and P. Lin, "Multifocus image fusion based on NSCT and focused area detection," *IEEE Sensors J.*, vol. 15, no. 5, pp. 2824–2838, May 2015.
[7] P. S. Chavez, Jr., and A. Y. Kwarteng, "Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis," *Photogramm. Eng. Remote Sens.*, vol. 55, no. 3, pp. 339–348, 1989.
[8] A. Toet, L. J. van Ruyven, and J. M. Valeton, "Merging thermal and visual images by a contrast pyramid," *Opt. Eng.*, vol. 28, no. 7, p. 287789, 1989.
[9] P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, Apr. 1983.
[10] 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.
[11] 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, 1995.
[12] J. J. Lewis, R. J. O'Callaghan, S. G. Nikolov, D. R. Bull, and N. Canagarajah, "Pixel- and region-based image fusion with complex wavelets," *Inf. Fusion*, vol. 8, no. 2, pp. 119–130, 2007.
[13] S. Li, J. T. Kwok, and Y. Wang, "Multifocus image fusion using artificial neural networks," *Pattern Recognit. Lett.*, vol. 23, no. 8, pp. 985–997, 2002.
[14] Z. Wang, Y. Ma, and J. Gu, "Multi-focus image fusion using PCNN," *Pattern Recognit.*, vol. 43, no. 6, pp. 2003–2016, 2010.
[15] V. Aslantas and R. Kurban, "Fusion of multi-focus images using differential evolution algorithm," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8861–8870, Dec. 2010.
[16] M. Li, W. Cai, and Z. Tan, "A region-based multi-sensor image fusion scheme using pulse-coupled neural network," *Pattern Recognit. Lett.*, vol. 27, no. 16, pp. 1948–1956, 2006.
[17] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image Vis. Comput.*, vol. 26, no. 7, pp. 971–979, 2008.
[18] S.-Y. Chien, S.-Y. Ma, and L.-G. Chen, "Efficient moving object segmentation algorithm using background registration technique," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 7, pp. 577–586, Jul. 2002.
[19] X. Bai and F. Zhou, "Analysis of new top-hat transformation and the application for infrared dim small target detection," *Pattern Recognit.*, vol. 43, no. 6, pp. 2145–2156, 2010.
[20] I. Sobel, "Camera models and machine perception," Ph.D. dissertation, Artif. Intell. Project, Stanford Univ., Stanford, CA, USA, 1970.

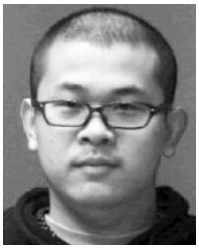
[21] P. Soille, *Morphological Image Analysis: Principles and Applications*. Berlin, Germany: Springer-Verlag, 2003.
[22] L. Lam, S.-W. Lee, and C. Y. Suen, "Thinning methodologies—A comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 9, pp. 869–885, Sep. 1992.
[23] M. Shemesh and O. Ben-Shahar, "Free boundary conditions active contours with applications for vision," in *Advances in Visual Computing*, vol. 6938. Berlin, Germany: Springer, 2011, pp. 180–191.
[24] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
[25] X. Qu. *Source Image for Image Fusion*, accessed on Sep. 5, 2016. [Online]. Available: http://www.quxiaobo.org/software/software_FusingImages.html
[26] Z. Zhou. *MWGF-Fusion*, accessed on Sep. 5, 2016. [Online]. Available: https://github.com/lsauto/MWGF-Fusion
[27] Y. Liu. *Liu Yu*, accessed on Sep. 5, 2016. [Online]. Available: http://home.ustc.edu.cn/~liuyu1/
[28] O. Rockinger, "Image sequence fusion using a shift-invariant wavelet transform," in *Proc. Int. Conf. Image Process.*, vol. 3. Oct. 1997, pp. 288–291.
[29] Y. Liu, S. Liu, and Z. Wang, "A general framework for image fusion based on multi-scale transform and sparse representation," *Inf. Fusion*, vol. 24, no. 1, pp. 147–164, Jul. 2015.
[30] S. Li, X. Kang, J. Hu, and B. Yang, "Image matting for fusion of multi-focus images in dynamic scenes," *Inf. Fusion*, vol. 14, no. 2, pp. 147–162, 2013.
[31] 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.
[32] 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.
[33] Y. Liu, S. Liu, and Z. Wang, "Multi-focus image fusion with dense SIFT," *Inf. Fusion*, vol. 23, no. 1, pp. 139–155, May 2015.
[34] X. Bai, Y. Zhang, F. Zhou, and B. Xue, "Quadtree-based multi-focus image fusion using a weighted focus-measure," *Inf. Fusion*, vol. 22, pp. 105–118, Mar. 2015.
[35] M. B. A. Haghighat, A. Aghagolzadeh, and H. Seyedarabi, "A non-reference image fusion metric based on mutual information of image features," *Comput. Elect. Eng.*, vol. 37, no. 5, pp. 744–756, 2011.
[36] M. Haghighat and M. A. Razian, "Fast-FMI: Non-reference image fusion metric," in *Proc. Int. Conf. Appl. Inf. Commun. Technol. (AICT)*, 2014, pp. 1–3.
[37] N. Cvejic, C. N. Canagarajah, and D. R. Bull, "Image fusion metric based on mutual information and Tsallis entropy," *Electron. Lett.*, vol. 42, no. 11, pp. 626–627, May 2006.
[38] R. Nava, G. Cristóbal, and B. Escalante-Ramírez, *Mutual Information Improves Image Fusion Quality Assessments*. Bellingham, WA, USA: SPIE, 2007.
[39] Q. Wang, Y. Shen, and J. Q. Zhangb, "A nonlinear correlation measure for multivariable data set," *Phys. D, Nonlinear Phenomena*, vol. 200, nos. 3–4, pp. 287–295, 2005.
[40] Q. Wang, Y. Shen, and J. Jin, "Performance evaluation of image fusion techniques," in *Image Fusion: Algorithms and Applications*, T. Stathaki, Ed. New York, NY, USA: Elsevier, 2008, ch. 19, pp. 469–492.
[41] J. Zhao, R. Laganiere, and Z. Liu, "Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement," *Int. J. Innov. Comput. Inf. Control*, vol. 3, no. 6, pp. 1433–1447, 2007.
[42] 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, 2008.
[43] Q. Wang, Y. Shen, Y. Zhang, and J. Q. Zhang, "Fast quantitative correlation analysis and information deviation analysis for evaluating the performances of image fusion techniques," *IEEE Trans. Instrum. Meas.*, vol. 53, no. 5, pp. 1441–1447, Oct. 2004.
[44] X. Bai, M. Liu, Z. Chen, P. Wang, and Y. Zhang, "Morphology and active contour model for multi-focus image fusion," in *Proc. IEEE Int. Conf. Digit. Image Comput. Techn. Appl. (DICTA)*, Nov. 2015, pp. 534–539.



XIANGZHI BAI received the B.S. and Ph.D. degrees from the Beijing University of Aeronautics and Astronautics (BUAA) in 2009 and 2003, respectively. He is currently a Full Professor with the Image Processing Center, BUAA. He holds eight national invention patents. He has authored over 100 international journal and conference papers in mathematical morphology, image analysis, pattern recognition, and bioinformatics. He also acts as an active Reviewer for around 50 international journals and conferences.

MIAOMING LIU received the B.S. degree from the Beijing University of Posts and Telecommunications in 2014. Her research interests include image fusion. She is currently pursuing her M.S. degree in Image Processing Center, BUAA.

ZHIGUO CHEN received the B.S. and M.S. degrees from the Beijing University of Aeronautics and Astronautics in 2013 and 2016, respectively. His research interests include image segmentation and analysis.



PENG WANG received the B.S. and M.S. degrees from the Beijing University of Aeronautics and Astronautics in 2012 and 2016, respectively. His research interests include image segmentation and analysis.

YU ZHANG received the B.S. degree from Dalian Maritime University in 2012, and the M.S. degree from the Beijing University of Aeronautics and Astronautics in 2015. His research interests include image segmentation and fusion.