IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 20, 2019, accepted March 30, 2019, date of publication April 9, 2019, date of current version April 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2909591

Multifocus Image Fusion Based on Fast Guided Filter and Focus Pixels Detection

FUQIANG ZHOU[1], XIAOSONG LI[1], JUAN LI[1], RUI WANG[1], AND HAISHU TAN[2]
[1]School of Instrumentation and Optoelectronic Engineering, Beihang University, Beijing 100083, China
[2]School of Physics and Optoelectronic Engineering, Foshan University, Foshan 528000, China

Corresponding author: Fuqiang Zhou (zfq@buaa.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 61471123 and Grant 61673039.

ABSTRACT As an effective technique of information fusion, multifocus image fusion has attracted increasing attention for image processing and computer vision. The goal of multifocus image fusion method is integrating all focus information from source images into fused result. In this paper, we propose a novel simple and effective multifocus image fusion technique based on fast guided filter and focus pixels detection. In order to detect the focused pixels correctly from source images, we develop a new multi-scale sum modified Laplacian technique. For the decision maps learning stage, based on the block consistency verification and faster guided filter techniques, we generate a series of binary focus decision maps to express the focus property of each pixel. In the fusion stage, we employ neighbor distance filter to extract detail pixels of source images, then the informative highpass images and the energetic lowpass images can be generated. The fused results are developed by constructing the corresponding decision maps and the neighbor distance filtered images. Compared with some state-of-the-art fusion methods, experimental results clearly demonstrate the superiority of the proposed method in terms of both comprehensive subjective assessment and some well-known quantitative evaluations.

INDEX TERMS Multifocus image fusion, multi-scale sum modified Laplacian, block consistency verification, fast guided filter.

I. INTRODUCTION

The constraint of the limited depth of field in optical lens motivates the development of multifocus image fusion technique which can integrate all focus information from all input images into one all focused image [1]. Up to now, multifocus image fusion has witnessed various significant applications in real products, such as medical, military, monitoring and robot guidance et al. In recent decades, a plethora of multifocus image fusion approaches have been proposed [2]. As a whole, multi-scale transform (MST) based methods and spatial domain (SPD) based methods are the dominating trend [3].

For the MST-based methods, the source images are first conducted by a multi-resolution decomposition tool, and then the different spectral information is selected by some rules. Lastly, reconstructing the fused image by employing an inverse multi-resolution transform. The way of MST-based methods processes images information is similar to that of the human vision system, which has a recognition process from coarse to fine at different scales and directions.

At the same time, the way of multi-scale analysis can effectively transfer the useful information from the source image to the fusion image, even in the case of spatial overlap of image features [4]. The superiority motivates these methods can achieve good fusion performance. Therefore, over the years, many effective MST-based methods have been developed. In this class of methods, the way of MST -is an important factor that affects the quality of fusion results. Popular MSTs include Laplace pyramid (LP) [5], discrete wavelet transform (DWT) [6], stationary wavelet transform (SWT) [7], [8], dual tree complex wavelet transform (DTCWT) [9], [10], contour transform (CT) [11], Shearlet [12], Curvelet [13] and non-subsample contour transform (NSCT) [14], [15] et al. These MSTs tools have been widely used in the field of image fusion. Among them, NSCT can achieve competitive characteristics of localization, anisotropy, multi-scale, multi-direction, and shift invariance in multifocus image fusion, so it will conduce to good performance in multifocus image fusion.

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

Apart from the way of MST, another important factor of MST-based algorithms is the fusion rules. In recent years, many fusion rules have been proposed [16]. Commonly used fusion rules are designed on pixels, local window neighborhood and local region algorithm. For the low frequency subband coefficient, the simplest way is the weighted averaging method. However, this manner will reduce the contrast of the fused image and lose some details of the image, thus reducing the quality of the fused image. For bandpass subband (high frequency) coefficients, generally, the way of the absolute value of pixels is adopted, because pixels with larger pixel values are more likely to correspond to sharp areas of the image. However, this method only considers a single pixel value and thus ignores the correlation between pixels.

Recently, many fusion rules have been put forward, Zhang and Guo [17] proposed an improved average scheme in the NSCT domain to deal with low-frequency coefficients and designed a method based on contrast to deal with pass-through sub-band coefficients. Nikolaos and Stathaki [18] designed a new region-based fusion rule by a transform domain in independent component analysis method, and the fusion results was greatly improved. Redondo *et al.* [19] proposed the selection rule based on multi-size window in the Log Gabor domain, and used the window with a small size to select which coefficient the local region originated from. This algorithm can comprehensively utilize the respective advantages of different Windows and retain the useful information on different sub-band coefficients in the fusion results. These fusion rules, that is, rules based on pixel, neighborhood window and region, can produce better fusion results. Generally, pixel-based algorithms are weaker than those methods that based on neighborhood Windows and region segmentation, because the latter two algorithms take into account the correlation between surrounding pixels.

However, the pixel-based rules do not take into account the correlation between the central pixel and its surrounding neighborhood pixels. For region segmentation based algorithms, the fusion performance is highly relying on the segmentation algorithm. At the same time, due to the difference and complexity of the image content, the segmentation algorithms are hard to be applied to all the source images adaptively. The integrity of the object in segmented based algorithm may not be considering sufficiently, high-quality fusion results would not be generated due to the incomplete object segmentation. For window-based fusion rules, window size directly affects the fusion results. Large Windows may include both focused and unfocused pixels and areas, whereas small Windows may judge smooth areas in focus as fuzzy, so it is important to select the appropriate window size [20].

Although some reasonable fusion rules have been proposed to for MST based methods, there still much room for improvement in fusion performance. The reason is that due to the complexity and fuzziness of the multifocus image content, some coefficients corresponding to the important features in the source image cannot be recognized correctly. Unfortunately, MAT based methods usually subject to the misfortune of losing useful information from source images for the sake of decomposing and reconstructing procedure.

In contrast to MST based methods, SPD based techniques fuse the input images directly in the spatial domain [21], [22]. For SPD based multifocus image fusion algorithm, the corresponding pixels (individual pixels or pixels in the neighborhood window), image blocks or image areas of source image are firstly segmented or extracted out through the correlation algorithm. Then, some focus measures (FM) are adopted to determine the focused pixels or areas and identify which pixel in the source image is focused. Finally, all the clear pixels, blocks or areas in each source image are copied directly to the corresponding position of the fusion image to obtain the final fusion image. Generally, these methods can be categorized into pixel based methods, block based methods and region based methods. For the pixel based techniques, the sharp pixels are processed independently to produce the fused image. That is to say, these operations just utilize individual pixels while not consider the connection among the surround pixels and the local and neighbor features on image. In the block based approaches, however, the suitable size of block is an extreme challenge, because a large block is more easily including both sharp and blur information. While small blocks may not be distinguished by the focus measures when they are focused but located in the smooth regions. Even though the adaptive methods could optimize the block-size choosing problem, but the block based methods still cannot get rid of blocking effect due to the complex content and variable information of images. Meanwhile, for the region based methods, most of the segmentation algorithms are complex and the performance usually depends on the segmentation algorithm [23], which may produce a suboptimal fusion result on account of the case of mistaken identity of pixel focusing properties.

To address the challenges mentioned above and accurately detect the focused information of source images, recent years, some carefully designed methods were proposed, which could generate pleasing fusion performance [20], [24]–[29]. Yang *et al.* [24] proposed a hybrid multifocus image fusion method based on NSCT and focused area detection, the final fused image can be obtained by the initial decision maps. Liu *et al.* [25] took the first step for utilizing deep convolutional neural network into the field of multifocus image fusion, they trained a fixed fusion model to fuse the source images. Li *et al.* [20] proposed a fixed window technique of multiscale image analysis and a new weighted fusion strategy by employing non-local means filtering. The artifacts and erroneous information are significantly decreased in the fused results in these method. In the method of [26], a multiple visual features measurement with gradient domain guided filtering was proposed, some key visual factors include contrast saliency, sharpness, and structural saliency are considered in this fusion scheme. In the work of [27], they employed

filtering techniques to detect focus regions and refine them by morphological techniques and consistency verification, but the experiments were not very sufficient. In the technique mentioned by Li *et al.* [28], they designed a multifocus image fusion method based on mixed-order structure tensors and multi-scale neighborhood. This method takes full advantage of fractional order and integer order structure tensors to detect focus information by a multi-scale window technique. Yu *et al.* [29] proposed a new multifocus image fusion method by combining the support vector machine and dual-tree complex wavelet transform. However, some of them either too complex (heavily rely on parameters setting) or cannot obtain a better fusion result. For the leaning based method [25], due to the complexity and diversity of nature source images content, a fixed model may be could not generate good fusion performance in some cases. In addition, many of these methods [24], [26]–[29] only performed on grayscale source images, which hinders the popularity and applications of them.

In this work, we propose a novel multifocus image fusion method based on fast guided filter (*FGF*) and focus pixels detection. As far as we know, this is the first time that *FGF* is applied to weight construction for image fusion. We consider the multifocus image fusion can be divided into two stages include focused decision maps learning and image fusion. In the learning phase, we develop a new multi-scale sum-modified-Laplacian (*NSML*) to detect focused pixels from source images, then generate the initial decision maps. By performing block consistency verification (*BCV*) technique, the wrong choice of small pixels and regions in initial decision maps will be eliminated effectively, then generates the middle decision maps. Lastly, we can obtain the final decision maps by taking the source images as guide into middle decision maps. In the fusion phase, source images are decomposed into highpass imges and lowpass images by neighbor distance filter (*ND* filter) directly, then the detail information and outline information will be presented in highpass and lowpass image, respectively. The fused result can be obtained by combing the corresponding final decision maps and the highpass and lowpass images. The main contribution of the proposed method are as following:

(1) We develop a new multiscale focus pixels detection scheme to generate the focus decision maps.
(2) We introduce block consistency verification and fast guided filter technique to remove wrong focus pixels and generate refined maps.
(3) Experimental results demonstrate the superiority of our method in both subjective visual effects and objective evaluation criteria.

The remainder of this paper is organized as follows. Section II illustrates the related work about our techniques. The proposed model is described in detail in section III. Simulations on multifocus grayscale and color source images are presented and the analysis of experimental results are given in Section IV. Section V is about conclusion.

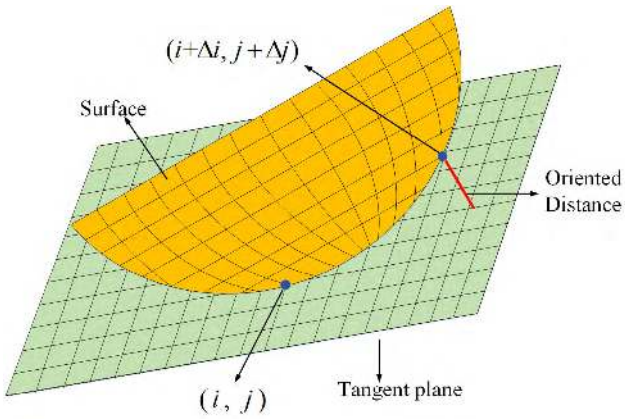


FIGURE 1. The illustration of oriented distance from point.

II. RELATED WORKS

A. ND FILTER

The primary task of multifocus image fusion is evaluating the focused information. Generally, the surface of focused object is sharper than those in out of focus region. Zhao *et al.* [30] developed the oriented distance (*OD*) to evaluate the blending degree of a surface in a given direction, and deduced a low computation complexity called *ND* filter shown in Fig.1 which can strong perception in measuring sharpness. For image surface, curve surface would lead the large directional distance between points, and the big change among image pixels. On the contrary, the smaller the grayscale change in pixel points, the flatter the image surface around the point, so the directional distance between pixel point and each pixel around it is smaller.

For a digital gray image $S = z(i, j)$, the oriented distance (*OD*) from point (i, j) to $(i + \Delta i, j + \Delta j)$ is formulated by

$$OD((i,j),(i+\Delta i,j+\Delta i)) = \frac{1}{2}l_{ii}(i,j)\Delta i\Delta i + 2l_{ij}(i,j)\Delta i\Delta j + l_{jj}(i,j)\Delta j\Delta j \quad (1)$$

where $l_{\mathbf{ij}}$, $(\mathbf{i}, \mathbf{j} = i, j)$ is given by

$$l_{\mathbf{ii}}(i,j) = z_{\mathbf{ij}}(i,j)/\sqrt{1+z_i^2(i,j)+z_j^2(i,j)} \quad (2)$$

Let Θ is the neighbor distance (*ND*) centered at point (i, j), then the neighbor distance of (i, j) is defined as

$$ND(i,j) = \sum\nolimits_{k\in\Theta} OD((i,j)(i+i_k, j+j_k)) \quad (3)$$

$\{(i_k, j_k)\}_{k=1}^{8} = \{(i+m, j+n) : m, n = 0, \pm 1\}$ denotes the neighborhood with eight pixels around (i, j), the neighbor distance of point (i, j) on image surface can be expressed as

$$\begin{aligned} ND(i,j) &= \sum\nolimits_{k=1}^{8} OD((i,j)(i+i_k, j+j_k)) \\ &= \sum\nolimits_{m=-1}^{1}\sum\nolimits_{n=-1}^{1} OD((i,j)(i+m, j+n)) \\ &= 2(l_{ij}(i,j) + l_{jj}(i,j)) \end{aligned} \quad (4)$$

More detail illustrations about *ND* filter can be found in [30].

B. FGF FILTER

Guided filter (*GF*) is a powerful technique for edge-preserving smoothing [31]. Due to the low complexity and the border refinement ability between focus and defocus, *GF* has been successful introduced into some image processing application.

GF is given by a linear model in a local neighbor around the pixel k.

$$q_i = a_k I_i + b_k, \quad \forall i \in \Theta_k \tag{5}$$

where Θ_k denotes a square window of a radius r, q and I are the output image and guidance image, respectively. Giving the input image p, the constant a_k and b_k in (5) can be calculated by minimizing the reconstruction error between p and q.

$$a_k = \frac{\frac{1}{|w|}\sum_{i\in\Theta_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \varepsilon} \tag{6}$$

$$b_k = \bar{p}_k - a_k u_k \tag{7}$$

where μ_k and σ_k are the mean and variance of I in the window Θ_k, $|w|$ is the number of pixels in Θ_k, ε is a regularization parameter set by the user to control the degree of smoothness. $\bar{p}_k$ is the mean of p in Θ_k.The filtering output can be estimated by

$$q_i = \bar{a}_i I_i + \bar{b}_i \tag{8}$$

To reduce the running time while maintaining the effectiveness of the *GF*, *FGF* was introduced as a speedup *GF* version [32]. By subsampling the input image and guidance image as a subsampling ratio s, *FGF* can decrease the time complexity form $O(N)$ to $O(N/s^2)$ and with almost no visible degradation at the same time. Algorithm 1 illustrates the detailed implementation of *FGF*. More detail information about *FGF* can be found in [31]–[33].

Algorithm 1 Fast Guided Filter(FGF)

1. $I' = f_{\text{subsample}}(I, s)$
   $p' = f_{\text{subsample}}(p, s)$
   $r' = r/s$
2. $\text{mean}_I = f_{\text{mean}}(I', r')$
   $\text{mean}_p = f_{\text{mean}}(p', r')$
   $\text{corr}_I = f_{\text{mean}}(I'.*I', r')$
   $\text{corr}_{I_p} = f_{\text{mean}}(I'.*p', r')$
3. $var_I = \text{corr}_I - \text{mean}_I.*\text{mean}_I$
   $cov_{I_p} = corr_{I_p} - \text{mean}_I.*\text{mean}_p$
4. $a = \text{cov}_{I_p}./(var_I + \varepsilon)$
   $b = mean_p - a.*\text{mean}_I$
5. $\text{mean}_a = f_{\text{upsample}}(f_{\text{mean}}(a, r'), s)$
   $\text{mean}_b = f_{\text{upsample}}(f_{\text{mean}}(b, r'), s)$
6. $q = \text{mean}_a.*I + \text{mean}_b$

III. PROPOSED FUSION ALGORITHM

In this paper, only two or three input multifocus images are tested in the proposed model. Our algorithm can easily extend to the situation of more than three source images. Some symbols will be used in the following discussion, So we would like to provide a summary for them. A, B, F, and X are the image names; (x, y) denotes the pixel coordinate; X_H and X_L are the highpass and lowpass components of image X, respectively. *IDM* is the initial decision map, *MDM* is the middle decision map, *FDM* is the final decision map, *NSML* denote the proposed novel *SML* to detect focus pixels.

A. DECISION MAPS LEARNINGS

Huang and Jing [34] compared some image focus measures and demonstrated *SML* can lead the first rank performance than other measures include variance, spatial frequency (*SF*), energy of image gradient (*EOG*), Tenenbaum's algorithm (Tenengrad), Energy of Laplacian (*EOL*). Therefore, *SML* is employed in source images and then the high-pass image and low-pass image are obtained. The *SML* image can be achieved as

$$SML(x, y) = \sum_{g=-P}^{P}\sum_{h=-Q}^{Q} [\nabla_{SML} f(x+g, y+h)]^2 \tag{9}$$

$$\begin{aligned}\nabla_{SML} f(x, y) = {} & |2f(x, y) - f(x-\tau, y) - f(x+\tau, y)| \\ & + |2f(x, y) - f(x, y-\tau) - f(x, x+\tau)|\end{aligned} \tag{10}$$

where τ is the step value which usually set as 1, P and Q controlling the window of size $(2P+1)\times(2Q+1)$. The center pixel in (10) is only related to neighboring pixels in both horizontal and vertical directions. To address this problem, a refined version of *SML* is generated which consider the connection of two diagonal direction, then (10) is rewritten as:

$$\begin{aligned}\nabla_{SML} f(x, y) = {} & |2f(x, y) - f(x-\tau, y) - f(x+\tau, y)| \\ & + |2f(x, y) - f(x, y-\tau) - f(x, x+\tau)| \\ & + |2f(x, y) - f(x-\tau, y-\tau) - f(x+\tau, y+\tau)| \\ & + |2f(x, y) - f(x-\tau, y+\tau) - f(x+\tau, y-\tau)|\end{aligned} \tag{11}$$

To reduce the computing complexity, generally, both of the parameters P and Q are set to the same value. So (9) can be modified as

$$SML_{k_1}(x, y) = \sum_{g=-k_1}^{k_1}\sum_{h=-k_1}^{k_1} [\nabla_{SML} f(x+g, y+h)]^2 \tag{12}$$

where k_1 denotes the size of neighbor surrounded by pixel (x, y), called scale factor. However, only one size of the neiborhood is considered in (12), in order to take full advantage of the different neighborhoods, we propose a new multi-scale *SML* (*NSML*) as follow:

$$\begin{aligned}NSML(x, y) = {} & \left|SML_{k_2}(x, y) - SML_{k_1}(x, y)\right| \\ & + \left|SML_{k3}(x, y) - SML_{k_1}(x, y)\right|\end{aligned} \tag{13}$$

where k_1, k_2, k_3 are the scale parameters. Then the initial decision maps (*SM*) are generated by comparing the *NSML* value of each pixel in A, B.

$$IDM_A(x,y) = \begin{cases} 1, & NSML_A(x,y) \geq NSML_B(x,y) \\ 0, & NSML_A(x,y) < NSML_B(x,y) \end{cases} \quad (14)$$

$$IDM_B(x,y) = \begin{cases} 1, & NSML_A(x,y) < NSML_B(x,y) \\ 0, & NSML_A(x,y) \geq NSML_B(x,y) \end{cases} \quad (15)$$

In order to refine the initial decision maps (*IDM*s), IDM_A and IDM_B are first segmented into non-overlapping block with size $n \times n$,. Then update the pixels in each block based on the following constraint in (16) to obtain a suite of binary images $B_i(x_k, y_k)$.

$$B_i(x_k,y_k) = \begin{cases} 1, & if \sum IDM_i(x_k,y_k) > \frac{n \times n}{2} \\ 0, & otherwise \end{cases} \quad (16)$$

where $B_i(x_k, y_k)$ denotes the k-th block at pixel (x_k, y_k), $k = 1, 2, \cdots N$, $i = A$ *or* B, and N is the sum of blocks. $B_i(x_k, y_k) = 1$ indicates all values of pixels in the k-th block are 1, on the contrary, $B_i(x_k, y_k) = 0$ means all pixels in k-th (x_k, y_k) are 0.

Update the pixels in each block based on the following rule

$$\hat{B}_i(x_k,y_k) = \begin{cases} 1, & |\Omega_i(\tilde{x}_k,\tilde{y}_k)| > \frac{r \times s}{2} \\ 0, & otherwise \end{cases} \quad (17)$$

where

$$\Omega_i(\tilde{x}_k,\tilde{y}_k) = \{(\tilde{x}_k,\tilde{y}_k)|B_i(\tilde{x}_k,\tilde{y}_k) = 1\}$$

$\hat{B}_i(\hat{x}_k, \hat{y}_k)$ denotes a neighborhood of $B_i(x_k, y_k)$ centered in (x_k, y_k), $|\cdot|$ is cardinality of $\Omega_i(\tilde{x}_k, \tilde{y}_k)$, $r \times s$ is the size of $B_{i*}(\tilde{x}_k, \tilde{y}_k)$, generally, $r \times s$ is 3×3 or 5×5.

At last, enlarge $\hat{B}_i(x_k, y_k)$ to make their size is the same with source images

$$\hat{B}_i(x_k,y_k) = \begin{cases} \text{ones}(8,8), & \hat{B}_i(x_k,y_k) = 1 \\ \text{zeros}(8,8), & \hat{B}_i(x_k,y_k) = 0 \end{cases} \quad (18)$$

To eliminate some small holes, thin protrusions, thin gulfs, narrow breaks in $\hat{B}_i(x_k, y_k)$, the area opening operation '*bwareaopen*' is performed, and the mistake holes will be removed significantly, and the *MDMs* can be generated.

$$MDM_i(x,y) = bwareaopen(\hat{B}_i(x_k,y_k), th_i) \quad (19)$$

where th_i is a pre-define threshold, $i = I_A, I_B$.

The above formulas (16)-(19) are the illustration of *BCV* [20].Traditional block based techniques are inevitable suffer from the uncertain size of block problems, and result in block effect in fused image. *BCV* technique can refine the binary decision maps and decrease the block effect in fused result. Despite the above *BCV* operation could generate some pleasing fusion decision maps, however, this block-relative technique usually subject to block effect in the border between focus and defocus more or less.

Algorithm 2 The Overall of the Proposed Method

Inputs: Multifocus source images A and B, initialize
Decision maps learning
- a) Compute the *NSML* of source images by (13).
- b) Obtain the *IDM*s by comparing *NSML* in (14) and (15).
- c) Optimize*IDM*s by *BCV* in (16)-(18) to generate *MDM*s.
- d) Generate the *MDM*s by *bwareaopen* operation in (19)
- e) Perform *FGF* to obtain the *FDM*s in (20) and (21).

Image fusion
- a) Extract the highpass components via *ND* filter in (22)
- b) Calculate the lowpass by (23)
- c) Obtain the fused highpass and lowpass image by combing *DFM*s in (24) and (25).
- d) Generate the fused image by (26)

Outputs: Fused image F

Let source images as guide to optimize MDM_i, four corresponding fast guider decision maps are developed

$$FDM_{i,H} = G_{r_1,\varepsilon_1,s}(MDM_i, i) \quad (20)$$

$$FDM_{i,L} = G_{r_2,\varepsilon_2,s}(MDM_i, i) \quad (21)$$

where $i = A, B$, r_1, r_2, ε_1, ε_2 and s are parameters of *FGF*, $FDM_{i,H}$ and $FDM_{i,L}$ are the resulting decision maps of highpass and lowpass images of source image. Actually, similar strategies of guided filtering has been used in many image fusion tasks [20], [26], [35].

B. IMAGE FUSION

The highpass and lowpass images of source images can be obtained by *ND* filter

$$X_H = ND * X \quad (22)$$

$$X_L = X - X_H \quad (23)$$

where X_H and X_L are the highpass and lowpass images of image X, $X = A, B$.

Combining *FDM*s with the base layer and detail layer in (23) and (22), the highpass and lowpass images of source images can be generated by

$$F_{\mathrm{H}}(x,y) = FDM_{A,H}(x,y) \times A_H(x,y) + FDM_{B,H}(x,y) \times B_H(x,y) \quad (24)$$

$$F_L(x,y) = FDM_{A,L}(x,y) \times A_L(x,y) + FDM_{B,L}(x,y) \times B_L(x,y) \quad (25)$$

Finally, the fused image can be obtained by the following

$$F(x,y) = F_H(x,y) + F_L(x,y) \quad (26)$$

The mechanisms of our fusion technique is described in Algorithms 2. The main procedure of the proposed fusion scheme is summarized in Fig. 2. In the decision maps

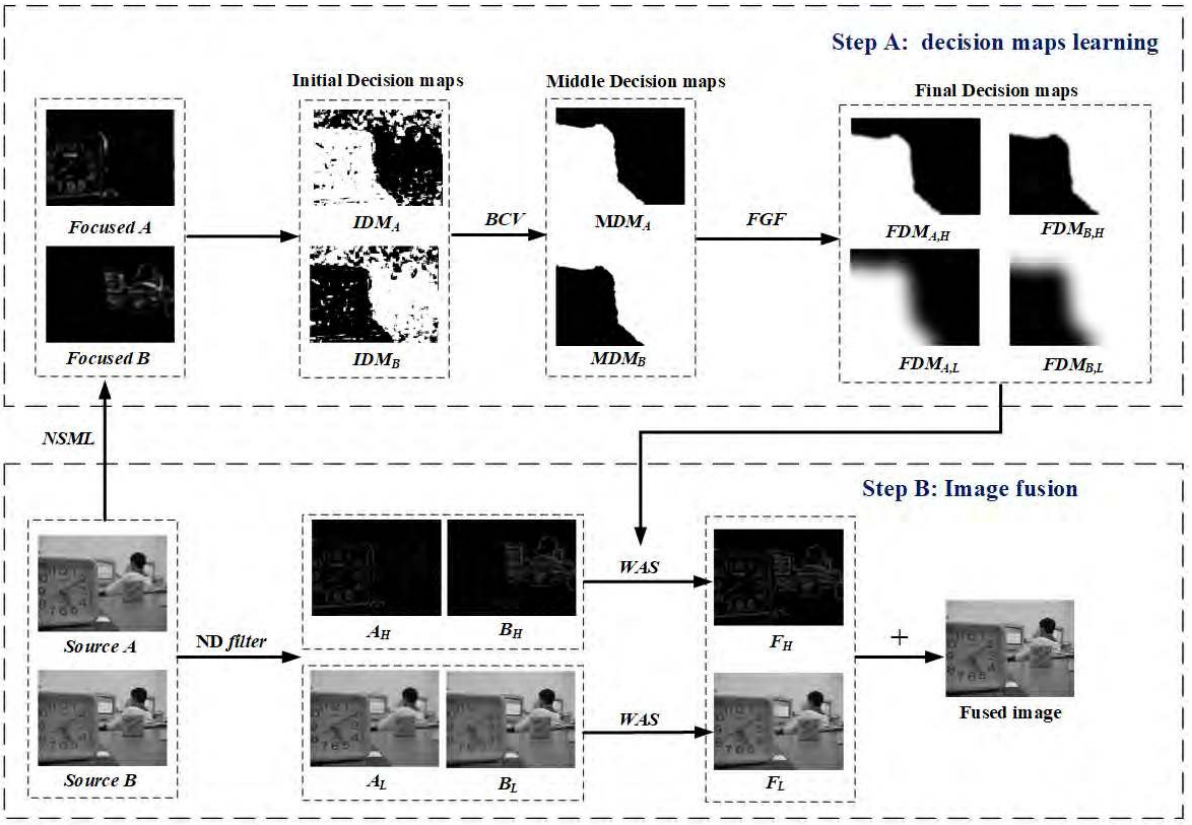


FIGURE 2. Schematic diagram of the proposed multifocus image fusion algorithm.

learning, the salient maps are generated first by *NSML*, and the *BCV* technique is performed to construct the initial decision maps. The final binary decision maps can be developed by guided filtering. In the fusion progress, highpass image and lowpass image of source image are obtained by *ND* filter, then we can reconstruct the final fused image by the learned maps and *ND* filtered images using weighting averaging scheme (*WAS*).

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In the experiments, all of these tests are implemented under Matlab 2012b environment and on PC with an Intel(R) Core(TM) i7-3770 CPU@ 3.4 GHz (8 CPU) and 12GB RAM. The proposed methods will compare with other seven multifocus image fusion methods, nine sets of source images include gray, color and triple color multifocus source images will be utilized. Subjective and objective evaluation include computational efficiency on fused results will be analyzed in detail.

A. DISCUSSIONS OF PARAMETERS SELECTION

In this subsection, the parameters in the proposed method are provided in detail. Among them, According to the suggestions of literatures [32], [35], and combined with experimental verification and analysis, the parameters include r_1, ε_1, r_2, ε_2, s in fast guided filter are set 25, 0.1, 10, 0.000001, 5 respectively. The step parameter τ in *NSML* is equal to 1, the scale factors k_1, k_2 and k_3 are 3, 5, 7 respectively. The parameters of *ND* filter are same with the publication [30]. The threshold th_i in (19) is 12000. The size of $B_{i*}(\tilde{x}_k, \tilde{y}_k)$ in *BCV* is set to 8×8. These parameters setting can obtain good results for all images used in this paper.

B. EXPERIMENTAL SETTINGS

1) TEST SOURCE IMAGES

To verify the fusion performance of the proposed method, in this section, nine pairs of testing multifocus source images shown in Fig.3 are utilized in experiments. In these source images, four sets of grayscale images are first conducted. To verify the proposed method is applicable to color image, four pairs of color images are employed. Furthermore, a pair of multifocus color images which has three source images is provided to demonstrate our method can handle more than two source images. All of them are perfect registration and displayed and these source images are public available at http://imagefusion.org/ and http://mansournejati.ece.iut.ac.ir/content/lytro-multi-focus-dataset from.

2) THE SUPERIORITY OF NSML THAN SML

In this section, to verify the advantage of *NSML* than *SML*, we compare the proposed *NSML* scheme described in Section III with *SML* scheme, meanwhile, other parts of these

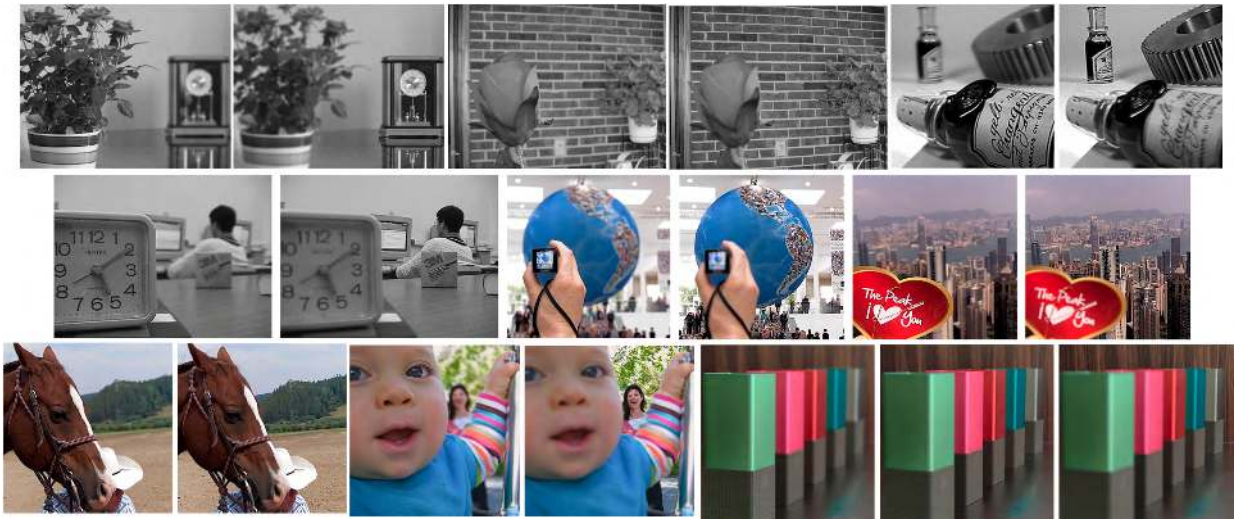


FIGURE 3. Nine pairs of multifocus source images used in our experiments.

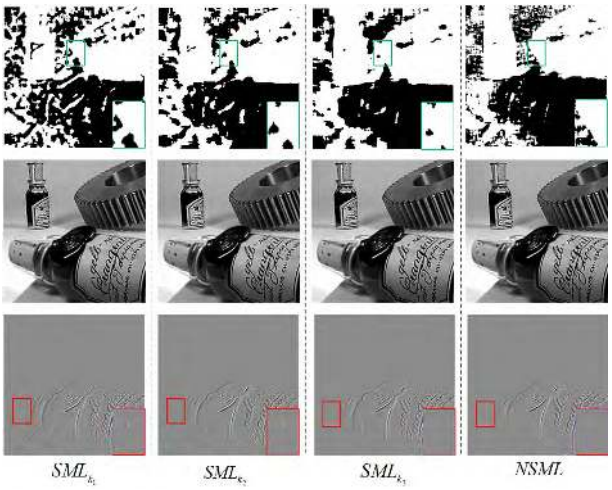


FIGURE 4. Fusion results by SML and NSML.

fusion methods are same with the proposed fusion methods. For simplicity, we utilized the multifocus ''Bottle'' images shown in Fig. 4 in this experiment. The initial decision maps, fused results and difference images are listed in each row in Figure 4. Furthermore, in Figure 4, the first three column are the results produced by SML_{k_1}, SML_{k_2} and SML_{k_3} schemes, respectively, and the results in the last column are obtained by the proposed *NSML* scheme. For the parameters in *SML* and *NSML*, k_1, k_2 and k_3 are set 3, 5, 7 respectively.

As can be seen from these initial decision maps, we can see the proposed *NSML* can accurately detect the focus information than *SML* scheme, the reason lies in *NSML* integrates the multiscale property of different window size instead of single size of *SML*. Besides, with carefully observe the difference images in the last row in Figure 4, we can find the proposed *NSML* scheme has the least information in focus region than other three *SML* schemes, which means the *NSML* scheme can transfer more focused information from source images than *SML* scheme.

3) COMPARING METHODS

To demonstrate the superiority of the proposed technique, Seven state-of-the-art image fusion methods including Convolutional Sparse Representation based method (CSR) [36], Guided filter based method (GF) [35], NSCT PCNN based method (NSCT_PCNN) [23], NSCT-contrast based method (NSCT_Con) [17], SR method (Yang_SR) [37], NSCT-SR method (NSCT_SR) [38], Multiscale curvature method (Curvature) [39] are compared. For these seven algorithms, all fusion rules and main parameters setting are determined by the corresponding authors' suggestion in the respective publications.

4) OBJECTIVE EVALUATION METRICS

The way to evaluate the results of multifocus image fusion is usually first through the subjective visual evaluation by human vision system (HVS). However, due to the complexity and diversity of image content, as well as the subjective bias of HVS evaluation. In general, it is necessary to quantitatively evaluate the objective indicators of the images.

Up to now, there has not been a unified and objective evaluation index can apply to any image quality evaluation. Related researches has been going on. In order to reduce the incompleteness of a single metric as much as possible, many related papers use multiple indexes to evaluate the image quality. This paper employs six popular fusion image quality metrics to evaluate the objective fusion performance. These metrics include mutual information metric (MI) [40], edge information transferred degree metric $Q^{AB/F}$ [41], Nonlinear Correlation Information Entropy metric Q_{NCIE} [42], Multiscale Scheme metric Q_M [43], Phase Congruency metric

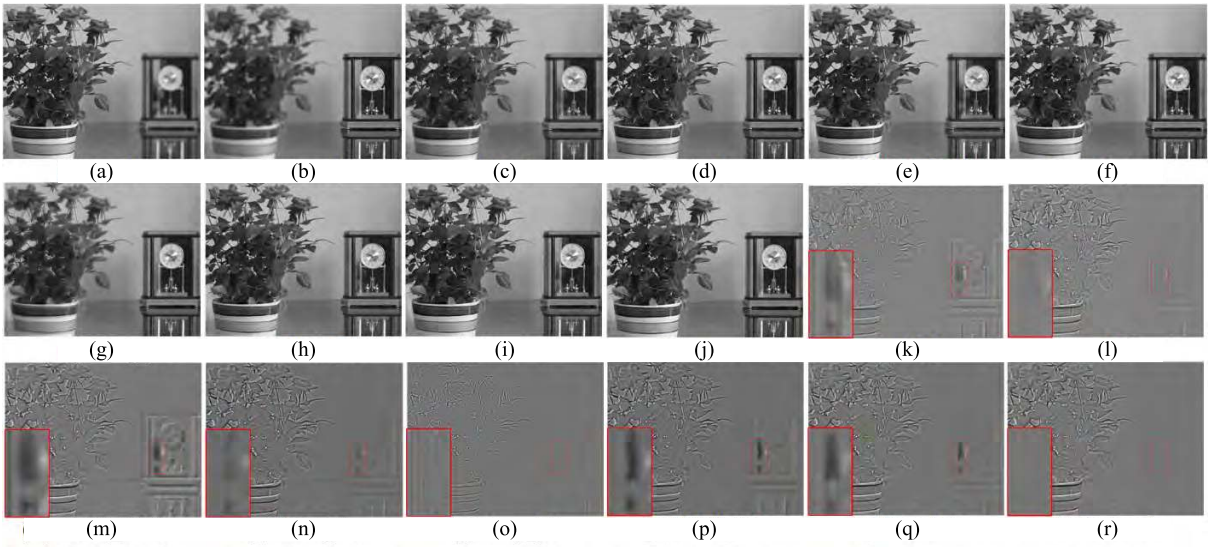


FIGURE 5. Fused results of different methods on processing the "Leaf" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (b).

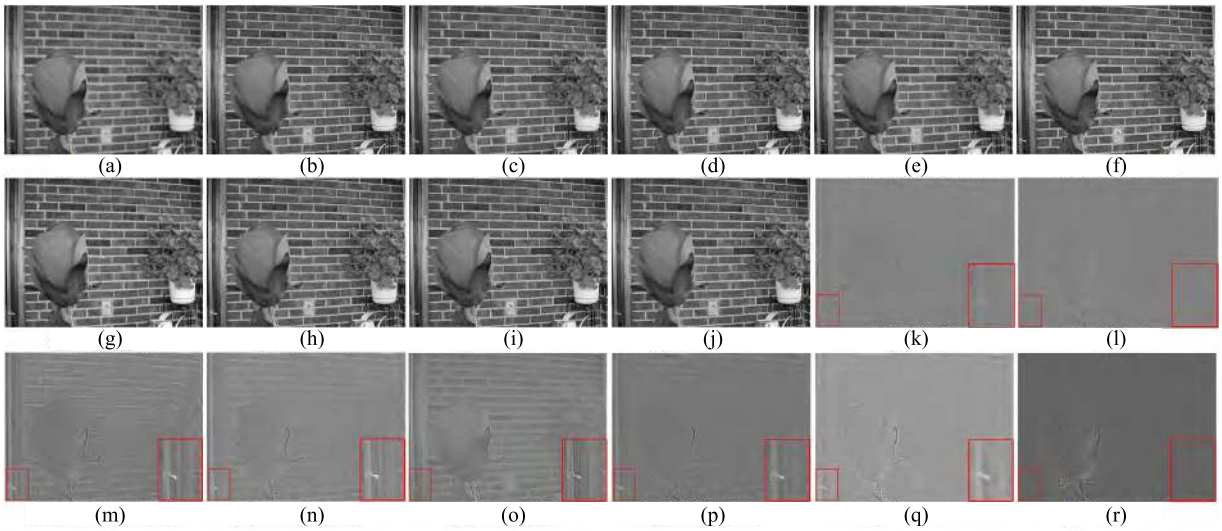


FIGURE 6. Fused results of different methods on processing the "Flower" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (b).

Q_P [44], [45], Chen-Blum Metric Q_{CB} [46]. Among these metrics, MI and Q_{NCIE} are information theory-based metrics, Q_M and Q_P are image feature-based metrics, $Q^{AB/F}$ is image detail-based metric, Q_{CB} is based on human perception inspired metric. Many of them are analyzed and summarized in [47], for all of them, the larger values indicate the better fusion performance.

C. ANALYSIS OF EXPERIMENTAL RESULTS

1) FUSION OF MULTIFOCUS GRAYSCALE IMAGES

The first experiments was performed on multifocus grayscale source images. Four sets of source images including the "Leaf", "Flower", "Lab", and "Bottle" images are tested, the size of them are 944×736, 512×384, 640×480 and 256×256, the fusion results of different methods are displayed in Fig.4, Fig.5, Fig.6, and Fig.7, respectively. Generally, the aim of multifocus image fusion is to preserve all focused information from different source images into the fused image. From the perspective of human eyes, it is usually difficult to make an impartial visual comparison among these fused all-focused images. Fortunately, difference image between fused image and source image can address this question. In the ideal fusion case, the area focused in source image ought to be no information in the difference image due to the successful copy of useful information. That is to say, the amount of residual information in difference image can reveal the fusion performance of algorithm. As can be seen from these fusion results shown in (k)-(r) of Fig. 5-8,

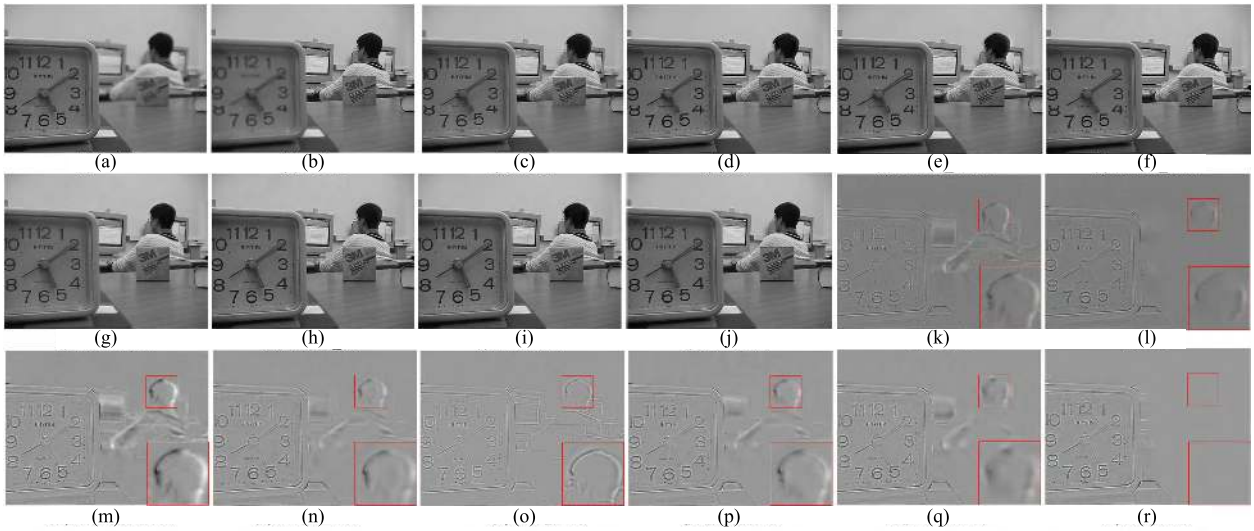


FIGURE 7. Fused results of different methods on processing the "Lab" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (b).

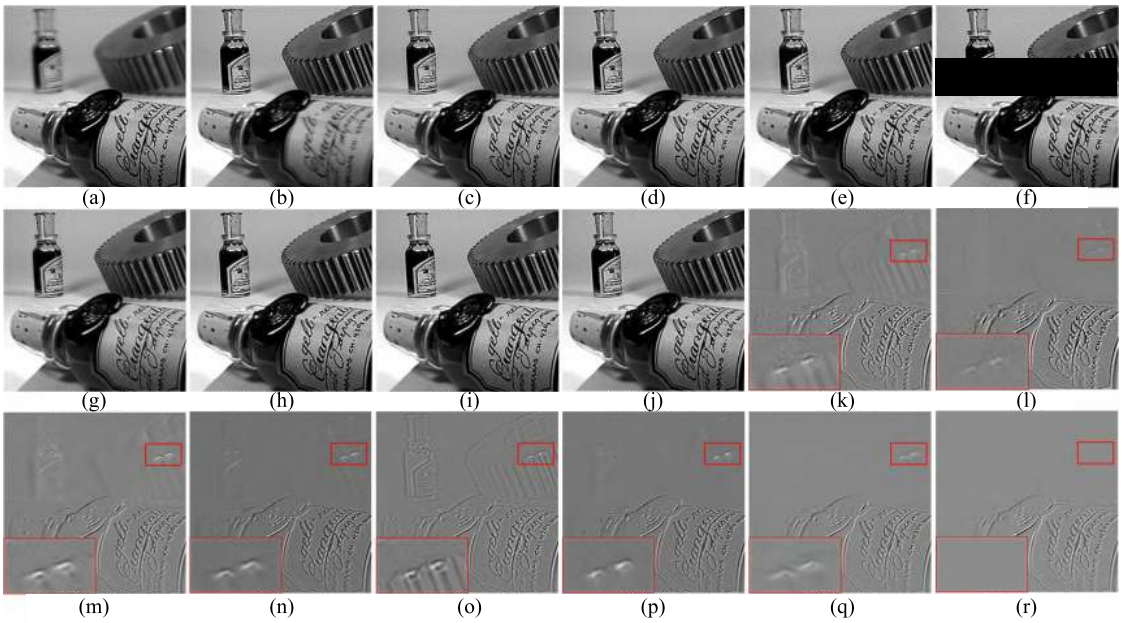


FIGURE 8. Fused results of different methods on processing the "Bottle" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (b).

the differences produced by the fused image and corresponding source image. For the convenience of comparison, a same region in the differences is selected by the red line rectangle and enlarged and presented in the bottom of left (or right) corner of the respective difference.

As can be seen from these difference images, CSR, NSCT_PCNN, NSCT_Con, Yang_SR and NSCT_SR methods introduced some un-focused information from source images into the fused results, which indicate some detail information such as texture, edge et al. of source images cannot be successful copied into fused images. In contrast. For GF and Curvature methods, the majority of useful information from source images are retained in fused results, and the boundaries between the focused and non-focused regions relatively clear. However, it can be clearly seen that our method can generate the better visual effect. As shown in Fig. 5(r), Fig. 6(r), Fig. 7(r) and Fig. 8(r), there is almost no residual information in focus area of the difference images, that is, the focus information in our fused results from source images is successfully copied into the fusion images. At the same time, our algorithm can also be effectively applied to the fusion problem of lab images with ''unregistered

TABLE 1. The quantitative assessment of different fusion methods for "Leaf" and "Flower" source images .

Image	Methods	MI	$Q^{AB/F}$	Q_{NCIE}	Q_M	Q_P	Q_{CB}
Leaf	CSR	6.1182	0.65576	0.8226	0.9941	0.7113	0.5355
	GF	7.3002	0.7179	0.8313	2.0895	0.7849	0.6821
	NSCT_PCNN	7.6867	0.6880	**0.8389**	1.9667	0.7245	0.6977
	NSCT_Con	6.5483	0.6950	0.8257	1.5911	0.7722	0.6750
	Yang_SR	6.0746	0.6324	0.8226	1.0328	0.6675	0.5273
	NSCT_SR	6.2551	0.6921	0.8238	1.6377	0.7567	0.6445
	Cuvature	6.6794	0.7020	0.8267	1.9456	0.7729	0.6849
	Proposed	**8.1914**	**0.7335**	0.8386	**2.4046**	**0.8181**	**0.7655**
Flower	CSR	5.7173	0.7036	0.8204	0.9176	0.7555	0.7721
	GF	7.2740	**0.7270**	0.8357	**2.2191**	0.7648	0.8231
	NSCT_PCNN	5.0839	0.6737	0.8162	0.7687	0.7311	0.7586
	NSCT_Con	5.4894	0.7015	0.8191	1.2393	0.7718	0.7880
	Yang_SR	5.4359	0.7164	0.8183	0.8230	**0.7823**	0.7853
	NSCT_SR	5.7783	0.7062	0.8212	1.6673	0.7487	0.7966
	Cuvature	6.3758	0.7180	0.8262	1.9383	0.7572	0.8095
	Proposed	**7.7383**	0.7249	**0.8419**	2.2122	0.7611	**0.8236**

TABLE 2. The quantitative assessment of different fusion methods for "Lab" and "Bottles" source images.

Image	Methods	MI	$Q^{AB/F}$	Q_{NCIE}	Q_M	Q_P	Q_{CB}
Lab	CSR	7.1652	0.7081	0.8308	0.8711	0.6982	0.5902
	GF	7.9114	0.7506	0.8360	2.3666	0.7822	0.6921
	NSCT_PCNN	7.3668	0.7067	0.8330	1.2923	0.6769	0.6630
	NSCT_Con	7.1621	0.7250	0.8308	1.5664	0.7369	0.6635
	Yang_SR	6.9940	0.7085	0.8299	0.7620	0.7227	0.6065
	NSCT_SR	7.1809	0.7330	0.8309	1.9174	0.7375	0.6648
	Cuvature	7.3058	0.7366	0.8317	2.0995	0.7455	0.6627
	Proposed	**8.4738**	**0.7538**	**0.8405**	**2.5067**	**0.7999**	**0.7408**
Bottles	CSR	5.8690	0.6801	0.8212	0.4558	0.8588	0.6970
	GF	6.7716	0.7081	0.8264	1.5115	**0.8853**	0.7538
	NSCT_PCNN	5.7756	0.6681	0.8210	0.5917	0.8381	0.7173
	NSCT_Con	5.7024	0.6731	0.8204	0.6222	0.8492	0.7338
	Yang_SR	5.7772	0.6588	0.8208	0.2098	0.8390	0.6961
	NSCT_SR	6.0526	0.6905	0.8223	1.2810	0.8635	0.7470
	Cuvature	6.0135	0.67874	0.8220	1.4475	0.8545	0.7468
	Proposed	**7.7364**	**0.72015**	**0.8337**	**1.6915**	0.8849	**0.7926**

information'' (the head positions of the man in the two source images of ''Lab'' are different), which indicate that the proposed algorithm has the best subjective visual effect.

In addition to the subjective visual analysis, we also conduct a detailed objective index evaluation on these fusion results of Fig.5-8. Six image fusion evaluation metrics including MI, $Q^{AB/F}$, Q_{NCIE}, Q_M, Q_P, Q_{CB} are used to quantitatively evaluate the image quality of the fused results obtained by different algorithms. The results are given in Table 1 and Table 2 and the largest value of each group is shown in bold for facilitate comparison. It can be clearly seen that our method has the highest values in most of objective evaluation, and outperforms the other methods in terms of all metrics when the proposed technique is performed on the mis-registration 'Lab' source images. The evaluation results of different metrics show that the proposed method has an excellent performance in objective evaluation criteria. By the comprehensive visual comparison and objective evaluation, we can conclude that the proposed algorithm has the best subjective and objective evaluation on gray image fusion.

2) FUSION OF MULTIFOCUS COLOR IMAGES

In order to demonstrate the superiorities that the proposed method can fuse the color source images, the second set of experiments is performed on multifocus color images, i.e., 'Tellurion', 'Heart shape', 'Horse', 'Infant', images, the size of them are 256×256, the fusion results of different methods and their corresponding differences are displayed in Fig. 9, Fig. 10 Fig. 11, and Fig. 12, respectively. To facilitate comparisons, local regions enclosed by red and yellow

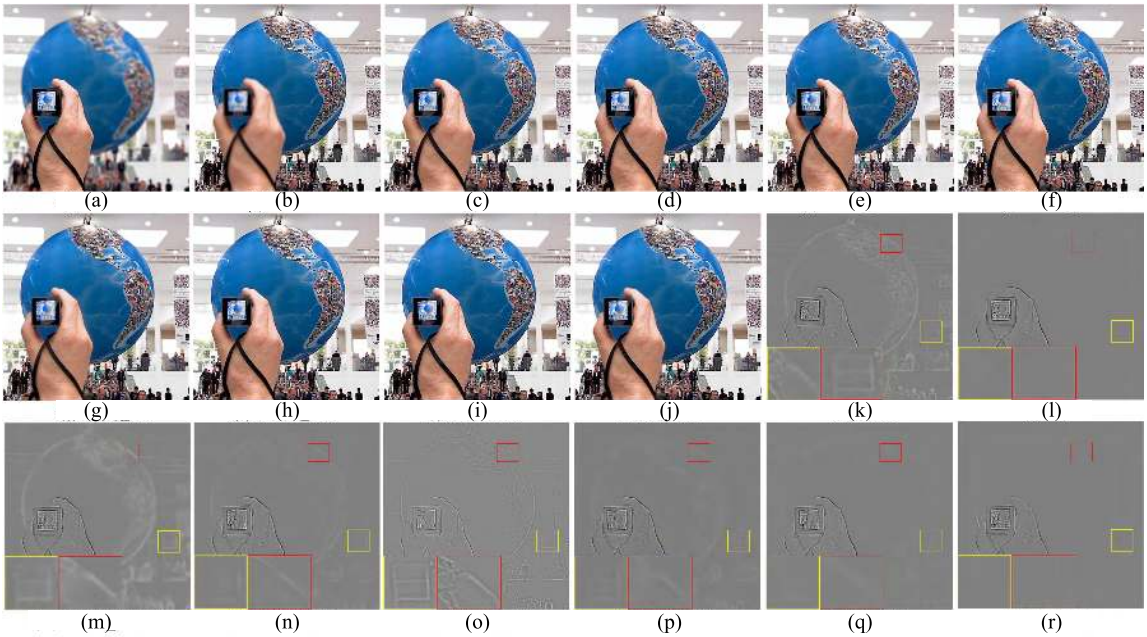


FIGURE 9. Fused results of different methods on processing the "Tellurion" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (b).

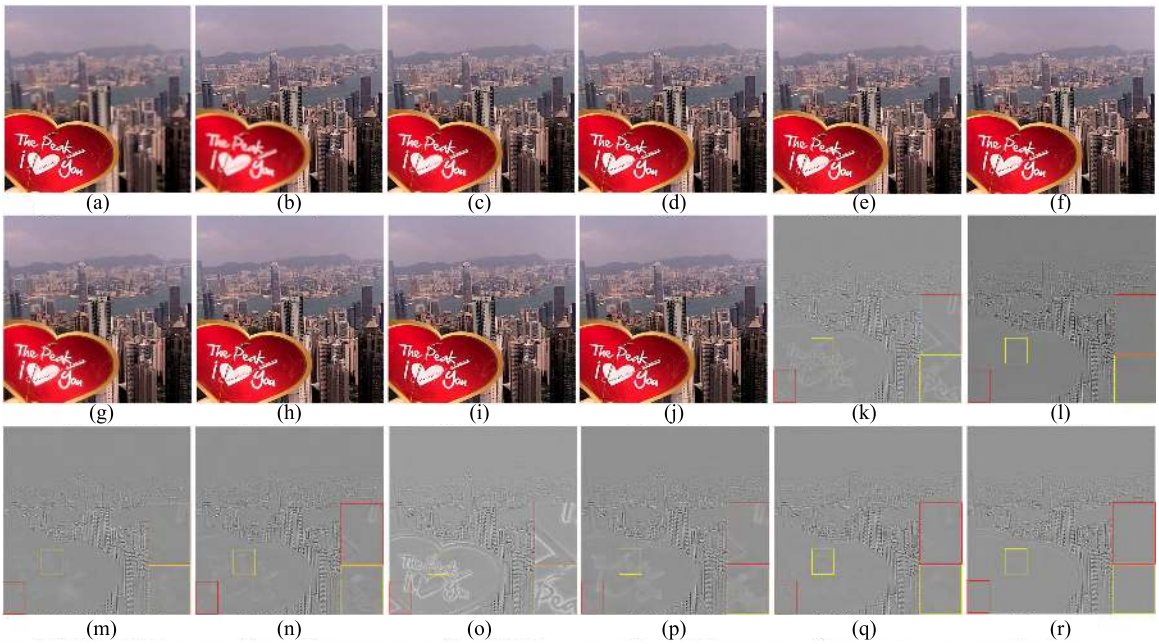


FIGURE 10. Fused results of different methods on processing the "Heart shape" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (a).

colored rectangle in Fig.9-12 are enlarged and presented in the bottom left corners in same color. From these magnified views of the colored rectangle regions, we can notice that the fusion results produced by CSR, NSCT_PCNN, NSCT_Con, Yang_SR, NSCT_SR and Curvature methods suffer from residual information in the focused regions of their differences. These illustrate that some useful information, i.e. contour, texture, boundary, from source images cannot be comprehensive transferred into the fused images by these algorithms due to the misjudgment of focused pixel. By contrast, the results produced be GF method can well detect the focused region. But with carefully observe the difference images Fig. 10(l) and Fig. 12(l), we can notice little wrong information have introduced in these results.

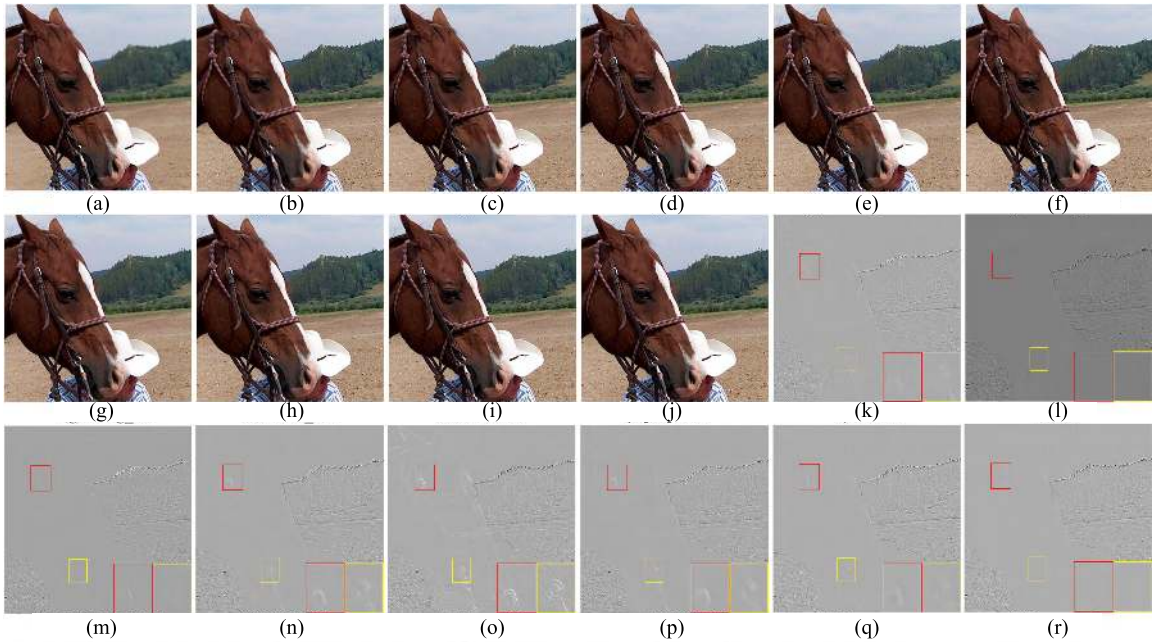


FIGURE 11. Fused results of different methods on processing the "Horse" images and their respective differences. (a)-(b) are sourceimages, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (a).

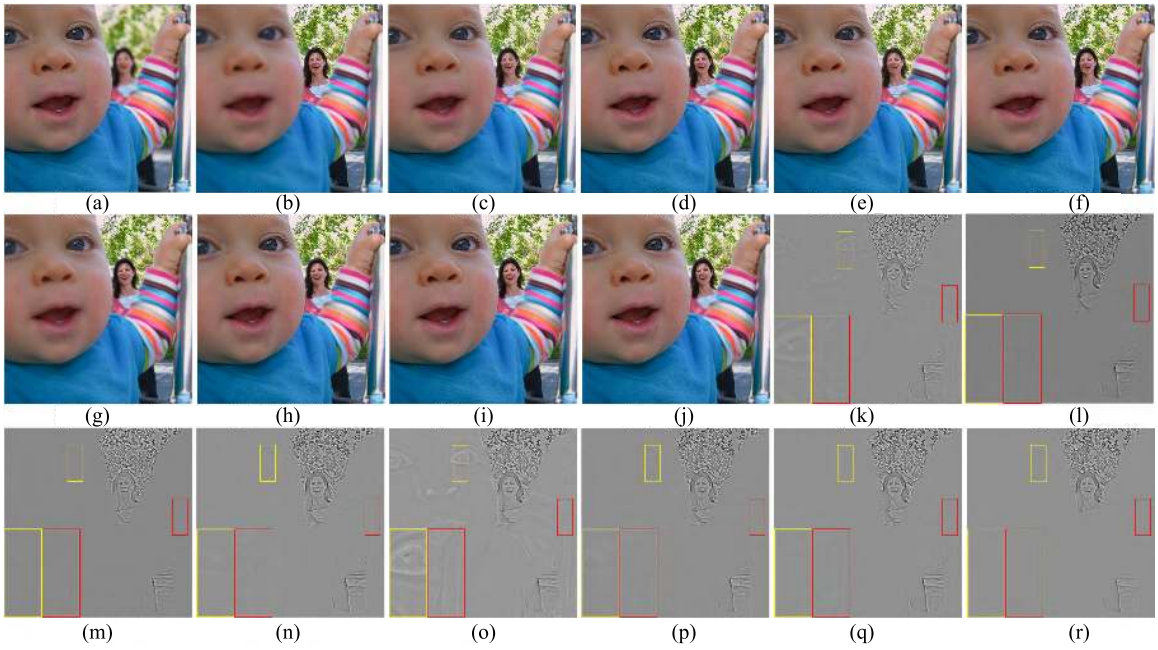


FIGURE 12. Fused results of different methods on processing the "Infant" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (a).

Meanwhile, from the Fig. 9(r), Fig. 10(r), Fig. 11(r) and Fig. 12(r), we can see that little information in these differences due to the successful detect the focus pixels in our scheme, which indicate the proposed method can obtain the best visual effect.

To evaluate the fusion performance more objectively, the quantitative objective evaluation metrics, i.e. MI, $Q^{AB/F}$, Q_{NCIE}, Q_M, Q_P, Q_{CB} are used in this experiments. For facilitate and clear comparisons, the evaluation values obtained by different fusion methods are displayed as eight bar charts with different color in Fig.13. According to these charts, we note that the proposed method performed slightly worse than the GF method when processing the ''Infant'' source images in terms of Q_M, as well as slightly lower than Curvature method

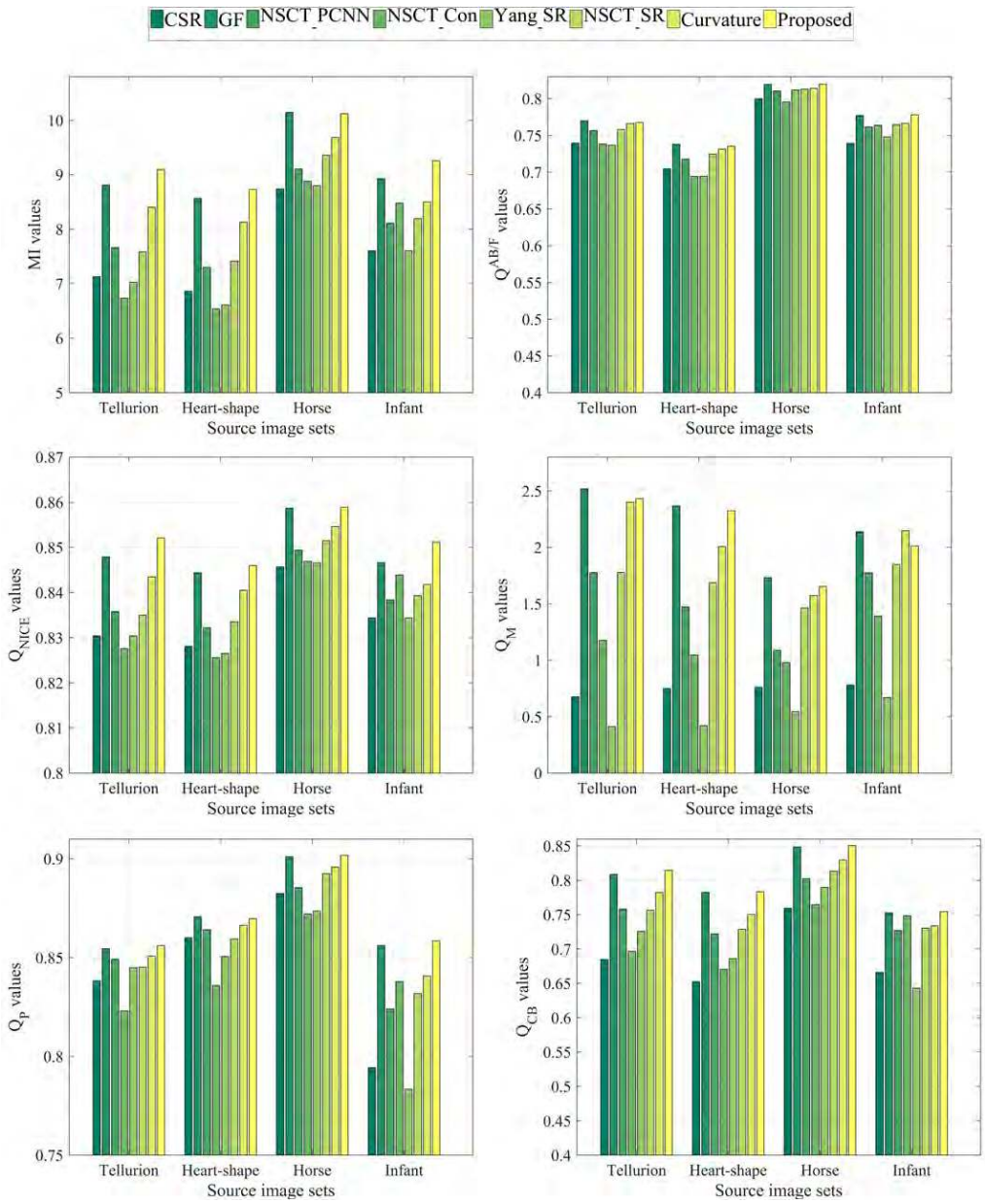


FIGURE 13. Performance comparison based on the color images fused results using MI, $Q^{AB/F}$, Q_{NCIE}, Q_M, Q_P, Q_{CB}.

for the "Infant" source images in terms of Q_M. Whereas it also performed similar with the GF in terms of $Q^{AB/F}$ and Q_{CB}. However, using only one metric with the highest evaluation score cannot reflect the fusion performance in an objective manner, the proposed method had the best performance in terms of more than half of the metrics. Thus we can conclude that the proposed method yields the best results for the color images in terms of visual quality and objective evaluation.

3) FUSION OF MULTIPLE COLOR IMAGES

To demonstrate our method can handle an image set containing more than two source images, the third experiment using three "Boxes" color source images are performed. The fused results produced by the different fusion methods and their corresponding differences are presented in Fig. 14. For the fusion process of these methods, it should be noted that the first and second source images were merged first, then the temporary fused image was fused with the third source image to obtain the final fusion results. The local regions enclosed by the red and yellow rectangle in differences are extracted and enlarged shown in the bottom right corners. With careful observation, we can find that there are more or less artifacts appearing in the transition area between the focused and defocused image in other methods, whereas the fusion results obtained by our proposed method contained fewer artifacts

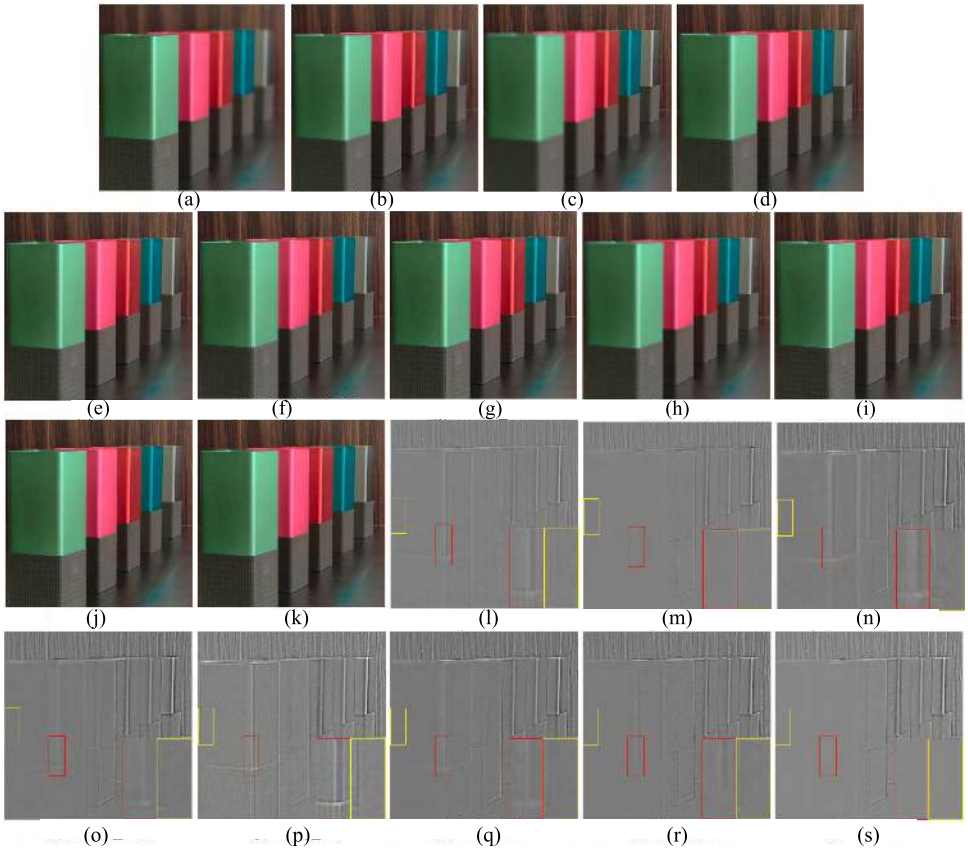


FIGURE 14. Fused results of different methods on processing the "Boxes" images and their respective differences. (a)-(b) are source images, (c)-(j) are fused results by different methods, (k)-(r) are differences between (c)-(j) and (a).

TABLE 3. The quantitative assessment of different fusion methods for "Boxes" source images.

Image	Methods	MI	$Q^{AB/F}$	Q_{NCIE}	Q_M	Q_P	Q_{CB}
Boxes	CSR	7.8036	0.6251	0.8364	1.3504	0.6346	0.5958
	GF	8.4685	0.7113	0.8423	**2.2492**	0.7956	0.7037
	NSCT_PCNN	7.5190	0.6338	0.8340	1.5934	0.6516	0.6233
	NSCT_Con	7.7489	0.6723	0.8360	1.8086	0.7536	0.6897
	Yang_SR	7.8488	0.6646	0.8370	1.3899	0.7068	0.6487
	NSCT_SR	7.8419	0.6782	0.8365	1.9232	0.7606	0.6903
	Cuvature	8.0000	0.6850	0.8381	1.9494	0.7605	0.6900
	Proposed	**8.6741**	**0.7178**	**0.8446**	2.2211	**0.8002**	**0.7046**

and exhibited the best visual effects Similarity, we used all of the evaluation criteria discussed above in this experiment. The quantitative assessments results of different methods using "Boxes" source images are presented in Table 3. We can easily find that our method outperform other methods in terms of MI, $Q^{AB/F}$, Q_{NCIE}, Q_M, Q_P, Q_{CB}. Based on the visual effect and objective analysis, we concluded that our methods are with the best performance compared with others in fusion of more than two source images.

According to the above comprehensive analysis and comparing, we can conclude that our method can detect the focused pixels accurately, and transfer the focused information from source images into the fused results successfully. The proposed method can generate pleasing fusion results in grayscale, color, triple multifocus images. We demonstrate the effectiveness of our method outperform some state-of-the-art fusion methods include the CSR, GF, NSCT_PCNN, NSCT_Con, Yang_SR, NSCT_SR and Curvature methods in terms of subjective visual effect as well as quantitative metrics.

D. COMPUTATIONAL EFFICIENCY ANALYSIS

The computational efficiency of different fusion algorithms are shown in Table 4. In these methods, NSCT-Con

TABLE 4. The computational complexity of different fusion methods (TIME/s).

	CSR	GF	NSCT_PCNN	NSCT_Con	Yang_SR	NSCT_SR	Curvature	Proposed
Leaf	479.14	1.69	691.18	335.86	9741.00	116.31	1265.32	14.83
Flower	122.03	0.66	177.97	93.73	2941.78	39.42	329.11	4.26
Bottles	49.17	0.45	48.51	32.03	962.99	13.58	124.31	2.11
Lab	194.29	0.87	294.86	147.65	3215.73	45.17	582.53	6.62
Tellurion	318.24	1.52	759.28	388.48	11102.68	158.95	1428.90	14.79
Heart shape	310.61	1.58	793.94	392.64	11508.25	132.96	1235.25	15.15
Horse	327.03	1.56	774.15	386.88	11788.43	134.70	1445.40	15.24
Infant	315.64	1.59	756.45	376.26	11323.35	146.89	1324.40	15.20
Boxes	546.03	3.55	1509.38	761.65	23156.41	240.64	2516.46	28.10

method is implemented by ourselves, while for CSR, GF, NSCT_PCNN, Yang_SR, NSCT_SR, Curvature, we use the original codes from respective publications provided by their authors. For the ''Boxes'' source images, we first fuse the first and the second images, and then the fused result is merged with the last one. From these time consumption of each fusion methods in Table 4, we can notice that the propose method is significantly efficient than CSR, GF, NSCT_PCNN, Yang_SR, NSCT_SR and Curvature methods. Meanwhile, GF method perform more efficient than the proposed method, however, the computational efficiency of our method is still acceptable, if we fully optimize our code and implement it in a more efficient manner such as, the running time of the proposed algorithms can be significantly decreased.

V. CONCLUSION

Various multifocus image fusion methods have been developed in recent years, nevertheless, this technique is remain a big challenge. In this paper, we propose a novel fusion scheme which includes two parts: decision maps learning and image fusion. For the decision maps learning stage, we develop a focus pixel detect metric *NSML* to extract the focus information of source images and generate a series of salient maps, then the initial decision maps are obtained by comparing the absolute value of pixels. In addition, we introduce the *BCV* technique to refine the initial decision maps to generate the middle decision maps. By taking sources as guide in *FGF* filter, the final decision maps can be obtained. In the image fusion step, in order to take full advantage of the focus information of source images in spatial domain, we employ the *ND* filter to generate the corresponding highpass and lowpass images. Thus the fused results can be obtained by the *ND* filtered images and corresponding final decision maps. Experimental results on nine sets of different multifocus images to demonstrate that the proposed approach is superior to some state-of-the-art methods in terms of human visual perceptions and objective criteria.

ACKNOWLEDGMENT

The part of this paper was presented at the International Conference [27]. The authors sincerely want to express heartfelt thanks to the editors and anonymous reviewers for the valuable comments and suggestions. We would like to thank Dr. Huafeng Li, Hengjun Zhao, Xiaobo Qu, Yu Liu, Shutao Li, Bin Yang, and Zheng Liu for providing the experimental codes for their methods.

REFERENCES

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

[2] H. Li, X. He, D. Tao, Y. Tang, and R. Wang, ''Joint medical image fusion, denoising and enhancement via discriminative low-rank sparse dictionaries learning,'' *Pattern Recognit.*, vol. 79, pp. 130–146, Jul. 2018.

[3] Y. Yang, M. Yang, S. Huang, M. Ding, and J. Sun, ''Robust sparse representation combined with adaptive PCNN for multifocus image fusion,'' *IEEE Access*, vol. 6, pp. 20138–20151, Apr. 2018.

[4] H. Li, X. Liu, Z. Yu, and Y. Zhang, ''Performance improvement scheme of multifocus image fusion derived by difference images,'' *Signal Process.*, vol. 128, pp. 474–493, Nov. 2016.

[5] J. Ma, Y. Ma, and C. Li, ''Infrared and visible image fusion methods and applications: A survey,'' *Inf. Fusion*, vol. 45, pp. 153–178, Jan. 2019.

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

[7] S. Li, ''Fusion using stationary wavelet transform: Effects of basis and decomposition level,'' *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 6, no. 1, pp. 37–50, Jan. 2008.

[8] Q. Jiang, X. Jin, S.-J. Lee, and S. Yao, ''A novel multi-focus image fusion method based on stationary wavelet transform and local features of fuzzy sets,'' *IEEE Access*, vol. 5, pp. 20286–20302, 2017.

[9] I. W. Selesnick, R. G. Baraniuk, and N. C. Kingsbury, ''The dual-tree complex wavelet transform,'' *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[10] R. Singh and A. Khare, ''Fusion of multimodal medical images using Daubechies complex wavelet transform—A multiresolution approach,'' *Inf. Fusion*, vol. 19, no. 3, pp. 49–60, Sep. 2014.

[11] M. N. Do and M. Vetterli, ''The contourlet transform: An efficient directional multiresolution image representation,'' *IEEE Trans. Image Process.*, vol. 14, no. 12, pp. 2091–2106, Dec. 2005.

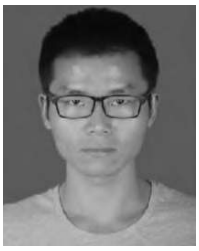
[12] L. Wang, B. Li, and L.-F. Tian, ''Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients,'' *Inf. Fusion*, vol. 19, no. 9, pp. 20–28, 2014.

[13] F. Nencini, A. Garzelli, S. Baronti, and L. Alparone, ''Remote sensing image fusion using the curvelet transform,'' *Inf. Fusion*, vol. 8, no. 2, pp. 143–156, Apr. 2007.

[14] A. L. D. Cunha, J. Zhou, and M. N. Do, ''The nonsubsampled contourlet transform: Theory, design, and applications,'' *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, Oct. 2006.

[15] H. Li, H. Qiu, Z. Yu, and Y. Zhang, "Infrared and visible image fusion scheme based on NSCT and low-level visual features," *Infr. Phys. Technol.*, vol. 76, pp. 174–184, May 2016.
[16] S. Li, B. Yang, and J. Hu, "Performance comparison of different multi-resolution transforms for image fusion," *Inf. Fusion*, vol. 12, no. 2, pp. 74–84, Apr. 2011.
[17] Q. Zhang and B.-L. Guo, "Multifocus image fusion using the nonsubsampled contourlet transform," *Signal Process.*, vol. 89, no. 7, pp. 1334–1346, 2009.
[18] N. Mitianoudis and T. Stathaki, "Pixel-based and region-based image fusion schemes using ICA bases," *Inf. Fusion*, vol. 8, no. 2, pp. 131–142, 2007.
[19] R. Redondo, F. Šroubek, S. Fischer, and G. Cristóbal, "Multifocus image fusion using the log-Gabor transform and a multisize windows technique," *Inf. Fusion*, vol. 10, pp. 163–171, Apr. 2009.
[20] H. Li, H. Qiu, Z. Yu, and B. Li, "Multifocus image fusion via fixed window technique of multiscale images and non-local means filtering," *Signal Process.*, vol. 138, pp. 71–85, Sep. 2017.
[21] 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.
[22] Y. Liu, S. Liu, and Z. Wang, "Multi-focus image fusion with dense SIFT," *Inf. Fusion*, vol. 23, pp. 139–155, May 2015.
[23] X.-B. Qu, J.-W. Yan, H.-Z. Xiao, and Z.-Q. Zhu, "Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampled contourlet transform domain," *Acta Autom. Sinca*, vol. 34, no. 12, pp. 1508–1514, 2008.
[24] 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.
[25] Y. Liu, X. Chen, H. Peng, and Z. F. Wang, "Multi-focus image fusion with a deep convolutional neural network," *Inf. Fusion*, vol. 36, pp. 191–207, Jul. 2017.
[26] Y. Yang, Y. Que, S. Huang, and P. Lin, "Multiple visual features measurement with gradient domain guided filtering for multisensor image fusion," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 4, pp. 691–703, Apr. 2017.
[27] X. Li, F. Zhou, and J. Li, "Multi-focus image fusion based on the filtering techniques and block consistency verification," in *Proc. IEEE 3rd Int. Conf. Image Vis. Comput. (ICIVC)*, Chongqing, China, Jun. 2018, pp. 453–457.
[28] H. Li, X. Li, Z. Yu, and C. Mao, "Multifocus image fusion by combining with mixed-order structure tensors and multiscale neighborhood," *Inf. Sci.*, vols. 349–350, pp. 25–49, Jul. 2016.
[29] B. Yu *et al.*, "Hybrid dual-tree complex wavelet transform and support vector machine for digital multi-focus image fusion," *Neurocomputing*, vol. 182, pp. 1–9, Mar. 2016.
[30] H. Zhao, Z. Shang, Y. Y. Tang, and B. Fan, "Multi-focus image fusion based on the neighbor distance," *Pattern Recognit.*, vol. 46, no. 3, pp. 1002–1011, Mar. 2013.
[31] K. He, J. Sun, and X. Tang, "Guided image filtering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 6, pp. 1397–1409, Jun. 2013.
[32] K. He and J. Sun, "Fast guided filter," pp. 1–2, May 2015. [Online]. Available: https://arxiv.org/abs/1505.00996
[33] K. He, J. Sun, and X. Tang, "Guided image filtering," in *Proc. Eur. Conf. Comput. Vis.*, Heraklion, Greece, Sep. 2010, pp. 1–14.
[34] W. Huang and Z. Jing, "Evaluation of focus measures in multi-focus image fusion," *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 493–500, 2007.
[35] 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.
[36] Y. Liu, X. Chen, R. K. Ward, and Z. J. Wang, "Image fusion with convolutional sparse representation," *IEEE Signal Process. Lett.*, vol. 23, no. 12, pp. 1882–1886, Dec. 2016.
[37] 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.
[38] 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, pp. 147–164, Jul. 2015.
[39] X. Li, H. Li, Z. Yu, and Y. Kong, "Multifocus image fusion scheme based on the multiscale curvature in nonsubsampled contourlet transform domain," *Opt. Eng.*, vol. 54, no. 7, Jul. 2015, Art. no. 073115.
[40] G. Qu, D. Zhang, and P. Yan, "Information measure for performance of image fusion," *Electron. Lett.*, vol. 38, no. 7, pp. 313–315, Mar. 2002.
[41] C. S. Xydeas and V. Petrovic, "Objective image fusion performance measure," *Electron. Lett.*, vol. 36, no. 4, pp. 308–309, Feb. 2000.
[42] Q. Wang, Y. Shen, and J. Jin, "Performance evaluation of image fusion techniques," in *Image Fusion: Algorithms Applications*, vol. 19. Amsterdam, The Netherlands: Elsevier, T. Stathaki, Jun. 2008, pp. 469–492.
[43] P.-W. Wang and B. Liu, "A novel image fusion metric based on multi-scale analysis," in *Proc. 9th Int. Conf. Signal Process.*, Beijing, China, Oct. 2008, pp. 965–968.
[44] 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.
[45] Z. Liu, D. S. Forsyth, and R. Laganière, "A feature-based metric for the quantitative evaluation of pixel-level image fusion," *Comput. Vis. Image Understand.*, vol. 109, no. 1, pp. 56–68, Jan. 2008.
[46] Y. Chen and R. S. Blum, "A new automated quality assessment algorithm for image fusion," *Image Vis. Comput.*, vol. 27, no. 10, pp. 1421–1432, Sep. 2009.
[47] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganiere, and W. Wu, "Objective assessment of multiresolution image fusion algorithms for context enhancement in night vision: A comparative study," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 1, pp. 94–109, Jan. 2012.

FUQIANG ZHOU received the B.S., M.S., and Ph.D. degrees in instrument, measurement and test technology from Tianjin University, China, in 1994, 1997, and 2000, respectively. He joined the School of Automation Science and Electrical Engineering, Beihang University, China, as a Postdoctoral Research Fellow in 2000, where he is currently a Professor with the School of Instrumentation and Optelectronic Engineering. His research interests include computer vision, image processing, and optical metrology.



XIAOSONG LI received the B.S. degree from Changchun University, in 2013, and the M.S. degree from the Kunming University of Science and Technology, in 2016. He is currently pursuing the Ph.D. degree in measurement technology and instruments with the Key Laboratory of Precision Opto-Mechatronics Technology, Ministry of Education, Beihang University, Beijing, China. His research interests include image processing, pattern recognition, and information fusion.

JUAN LI received the B.S. degree in measurement Control technology and instrumentation from Tianjin University, Tianjin, China, in 2016. She is currently pursuing the M.S. degree in measurement technology and instruments with the Key Laboratory of Precision Opto-Mechatronics Technology, Ministry of Education, Beihang University, Beijing, China. Her current research interests include object detection and deep learning.

RUI WANG received the B.S. and M.S. degrees in optical instruments from the Department of Precision Instruments and Mechanology, Tsinghua University, China, in 1988 and 1990, respectively, and the Ph.D. degree in machine vision of precision instrument and mechanics from the School of Instrumentation and Opto-Electronics Engineering, Beihang University, in 2006, where she is currently an Associate Professor. Her current research interests include machine vision, pattern recognition and tracking, optical sensing, and image processing.

HAISHU TAN received the B.S. and Ph.D. degrees in optical engineering from Tianjin University, China, in 1994 and 1998, respectively. He is currently a Professor with the School of Physics and Optoelectronic Engineering, Foshan University, China. His research interests include computer vision, photoelectric measurement, and optical metrology.