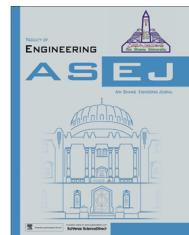




Ain Shams University

Ain Shams Engineering Journal

[www.elsevier.com/locate/asej](http://www.elsevier.com/locate/asej)  
[www.sciencedirect.com](http://www.sciencedirect.com)



## ELECTRICAL ENGINEERING

# Multi-focus image fusion using multi-scale image decomposition and saliency detection



Durga Prasad Bavirisetti <sup>\*</sup>, Ravindra Dhuli

School of Electronics Engineering, VIT University, Vellore 632014, India

Received 1 August 2015; revised 4 June 2016; accepted 21 June 2016  
Available online 9 July 2016

### KEYWORDS

Visual saliency;  
Weight map;  
Out-of-focus;  
Image fusion;  
Multi-scale

**Abstract** In this paper, we develop a new multi-focus image fusion method based on saliency detection and multi-scale image decomposition. Proposed method is very efficient, since the visual saliency explored in this algorithm is able to emphasize visually significant regions. Unlike most of the multi-scale fusion methods, an average filter is employed in our algorithm for multi-scale image decomposition. Hence it is computationally simple. A new weight map construction process based on visual saliency is developed. Weight maps of this algorithm are capable of detecting and identifying focused and defocused regions of the source images. We are able to integrate only focused and sharpened regions into the fused image. Performance of the proposed method is compared with that of the state-of-the-art multi-focus fusion methods. Proposed method outperforms them in terms of visual quality and fusion metrics. Our method requires considerably less computational time, thus making it preferable for real time implementation.

© 2016 Ain Shams University. Production and hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### 1. Introduction

Single sensor image cannot always provide entire information of a desired scene. Sometimes it is preferable to capture more than one image and all the necessary information of these multiple images has to be incorporated in a single image for better visual understanding of the scene. Image fusion [1] is the phenomenon of merging all the necessary information of source

images into a single image. Image fusion has numerous applications in diverse fields viz. digital photography [2], medical imaging [3,38], remote sensing [4], concealed weapon detection [5,6,37], military [3,36] and navigation [5,6].

Image fusion is classified in a generic way as in [7]: single-sensor image fusion (SSF) and multi-sensor image fusion (MSF). In SSF, multiple images of the targeted scene are captured using a single sensor. However, multiple sensors are used for the same purpose in MSF. These multiple images provide visually different or complementary information. For a human observer it is a tedious job to reliably combine and observe a composite image out of these multiple images. Therefore, for a better scene understanding, useful and complementary information of numerous source images should be integrated into a composite image. This combined image has to provide more details of the scene than any one of the individual source images. Digital photography applications namely multi-focus

\* Corresponding author.

E-mail addresses: [bdps1989@gmail.com](mailto:bdps1989@gmail.com) (D.P. Bavirisetti), [ravindradhuli@vit.ac.in](mailto:ravindradhuli@vit.ac.in) (R. Dhuli).

Peer review under responsibility of Ain Shams University.



Production and hosting by Elsevier

fusion [8] and multi-exposure fusion [9] fall under SSF; however, applications such as medical imaging, remote sensing, navigation, military and concealed weapon detection come under MSF. We concentrate only on multi-focus image fusion (MFF) in this paper.

Due to the inherent system limitations a single sensor is not capable of focusing more than one object of a scene at the same time. Thus multiple images with different focuses have to be captured and a multi-focus image has to be generated out of these several out-of-focus images in the process of MFF.

Fusion could be performed at three levels [1]: pixel, feature and decision. Fusion is done on each input image pixel by pixel at pixel level. However at feature level, fusion is executed on extracted features of the source images. At decision level, fusion is performed on probabilistic decision information of local decision makers. These decision makers are in turn derived from the extracted features. Pixel level fusion schemes are preferable for fusion compared to other level approaches because of their effectiveness and implementation ease. In this paper, our interest is only on pixel level fusion schemes.

The remaining paper is organized as follows: Section 2 reports the MFF literature. Section 3 describes the maximum symmetric surround saliency detection. Section 4 explains the proposed method. Section 5 details the fusion metrics. Experimental setup is presented in Section 6. Results are analyzed in Section 7. Section 8 concludes the paper.

## 2. Literature

The aim of any pixel level MFF scheme is to generate an all-in-one focus image with less computational time along with the following properties [1]:

1. It has to transfer entire complementary or useful information of input images into the composite image.
2. It should not lose source image information during fusion process.
3. It should not introduce artifacts into the fused image.

Over the past few decades numerous MFF algorithms have been developed to meet these requirements. MFF is broadly classified as spatial domain/single-scale fusion methods and multi-resolution/multi-scale fusion methods. In spatial domain, fusion is performed on the present scale source images without further decomposition. However in multi-scale, source images are decomposed into approximation and detail coefficients (images at several scales) at various levels. Fusion is performed on these decomposed coefficients by employing various fusion rules. All of these fused coefficients at different levels will be combined to obtain the fused image.

In a spatial domain fusion scheme principal component analysis (PCA) [10,11] tool is used for MFF. This method is computationally efficient. However, this tool is not able to give desirable results for most of the fusion datasets. Focus measure based approaches [12] are well known in this class. Here, source images are divided into blocks and various focus measures are employed to select the best among the image blocks. Variance, energy of image gradient, Tenenbaum's algorithm (tenengrad), energy of Laplacian (EOL), sum-modified-Laplacian (SML), spatial frequency (SF) are various focus measures successfully used for MFF [12]. From [12], it is

observed that SML gives superior performance compared to other focus measures. However, it is computationally expensive [13]. To address this problem, bilateral gradient-based sharpness criterion (BGS) [13] is used for MFF. But, it failed to produce a better focused image. In addition it is computationally demanding.

Next category in MFF is multi-scale fusion. Pyramid [14] and wavelet [11,15] based MFF methods are well known in this family. Discrete cosine transform with variance calculation (DCT + var) [16], discrete cosine transform with variance calculation and consistency verification (DCT + var + cv) [17], discrete cosine harmonic wavelet transform (DCHWT) [17], wavelet and adaptive block (DWT + AB) [19], cross bilateral filter (CBF) [5], guided image filter (GFF) [3] based methods are recently proposed MFF methods in multi-scale fusion category. DCHWT may produce blocking effects in the fused image. CBF may introduce gradient reversal artifacts in the composite image. DCT + var and DCT + var + cv methods are able to generate all-in-focus images. But they are not generating desirable focused image. DWT + DB method is unable to get better focused regions. Moreover, all of these methods are computationally expensive.

In addition to above fusion schemes, recently many researchers used visual saliency [20] for various applications of image fusion. Saliency preserving gradient is used for multi-focus fusion [21]. Visual saliency based on color is utilized for High time range imaging (HTRI) [22]. Frequency tuned saliency detection is employed for fusing the infrared and visible images in [23]. However, these approaches are different from our approach.

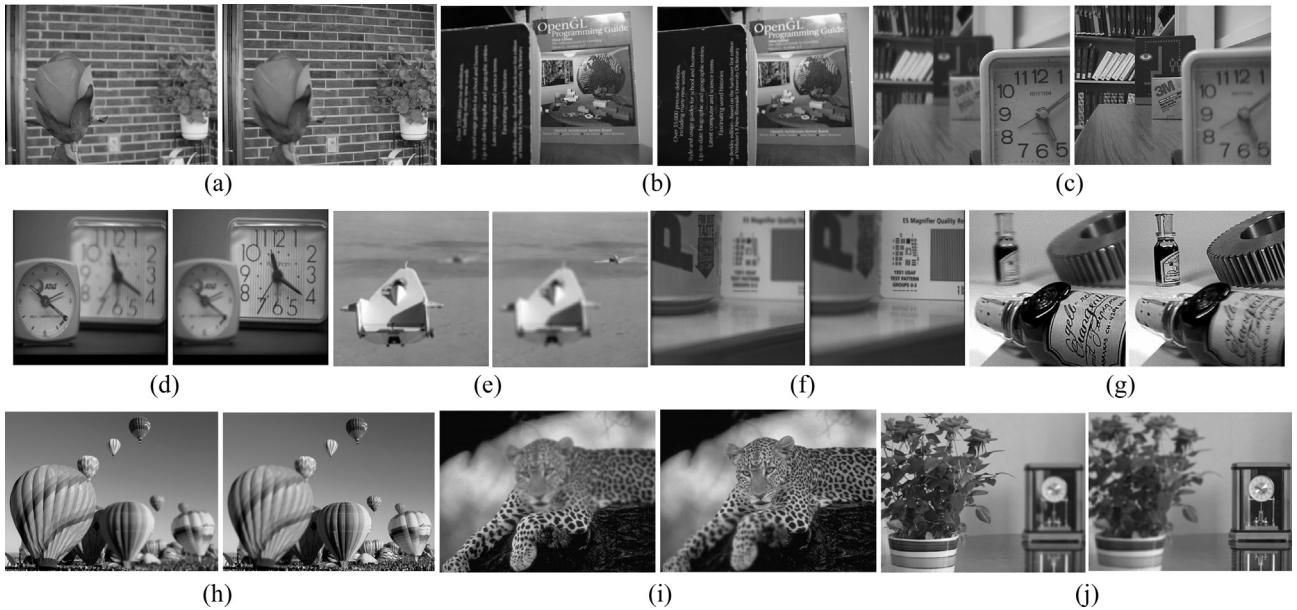
In this paper we propose a new MFF method based on maximum symmetric surround saliency detection (MSSS) and multi-scale image decomposition to address the problems of state-of-the-art MFF methods. This method is shortly referred as Saliency detection based MFF (SDMF). The highlights of the proposed method are as follows:

- SDMF can effectively integrate more focused (sharpened) regions of the source images compared to state-of-the-art MFF methods with less computational time.
- Unlike most of the multi-scale fusion methods, in our method a simple average filter is sufficient for multi-scale image decomposition.
- Saliency detection explored in our SDMF is able to emphasize visually significant regions.
- A simple and efficient weight map construction based on MSSS is proposed. This weigh map construction process is able to identify focus and defocus regions of source images very well.

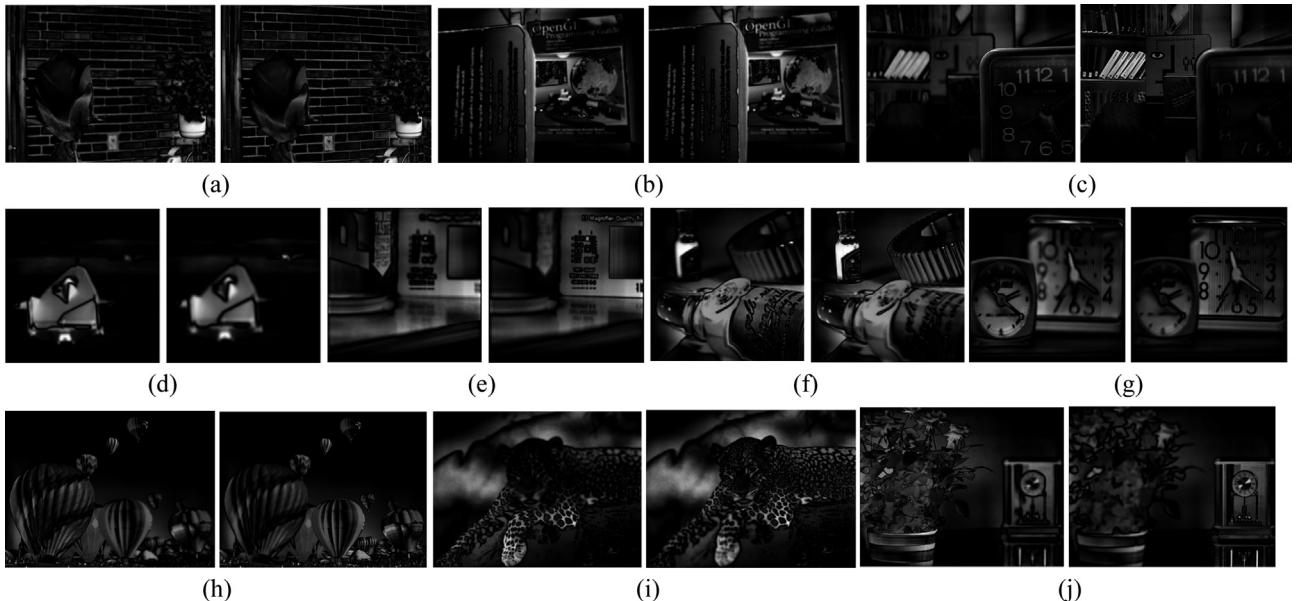
In the following section, we introduce the saliency detection used in our fusion algorithm.

## 3. Visual saliency detection

Saliency detection (SD) [20] is the way of detecting or identifying visually significant regions such as an object or a pixel, a person than their neighbors. These salient regions drag the human visual attention compared to other regions present in the image. SD is useful in many applications such as object segmentation, object reorganization and adaptive



**Figure 1** Multi-focus image datasets: (a) flower, (b) book, (c) book shelf, (d) clock, (e) aircraft, (f) pepsi, (g) bottle, (h) parachute, (i) leopard, (j) flower wage.



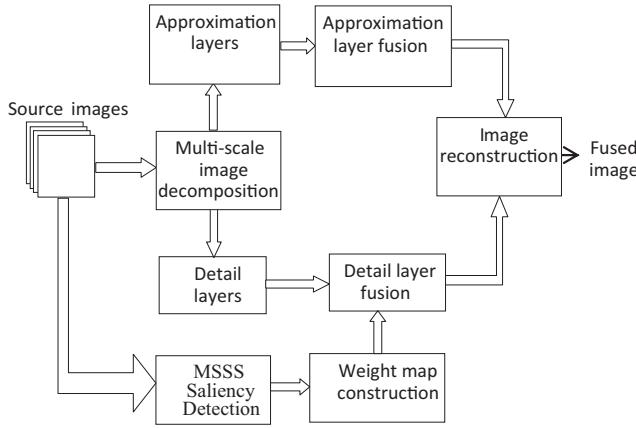
**Figure 2** Saliency maps of multi-focus image datasets: (a) flower, (b) book, (c) book shelf, (d) clock, (e) aircraft, (f) pepsi, (g) bottle, (h) parachute, (i) leopard, (j) flower wage.

compression. However, in this paper SD is employed for fusion. In [20], the properties of a good SD method are presented as follows:

- It should highlight large salient regions than small salient regions.
- It should uniformly highlight salient regions.
- Boundaries need to be well defined.
- It should ignore texture or noise artifacts.

In view of these properties we briefly review the existing SD methods. Some methods presented in [24–28] generate

saliency maps with low resolution. Some other SD algorithms in [25,26,28] produce ill-defined object boundaries. Because of these limitations, they are not useful to generate saliency maps for the purpose of fusion. To overcome the drawbacks of existing SD methods, Achanta et al. [20] introduced a frequency tuned saliency detection algorithm (FTSD). This algorithm is able to satisfy all the properties of a good SD method. But, this algorithm fails if image consists of complex background or large salient objects. To resolve these issues Achanta et al., suggested one more SD algorithm [29] called maximum symmetric surround SD algorithm (MSSS).



**Figure 3** General block diagram of the proposed algorithm.

In multi-focus images, focused regions provide visually more information than defocused regions. In other way, focused regions are more salient than defocused regions [30]. So we need to detect salient regions from these out-of-focus

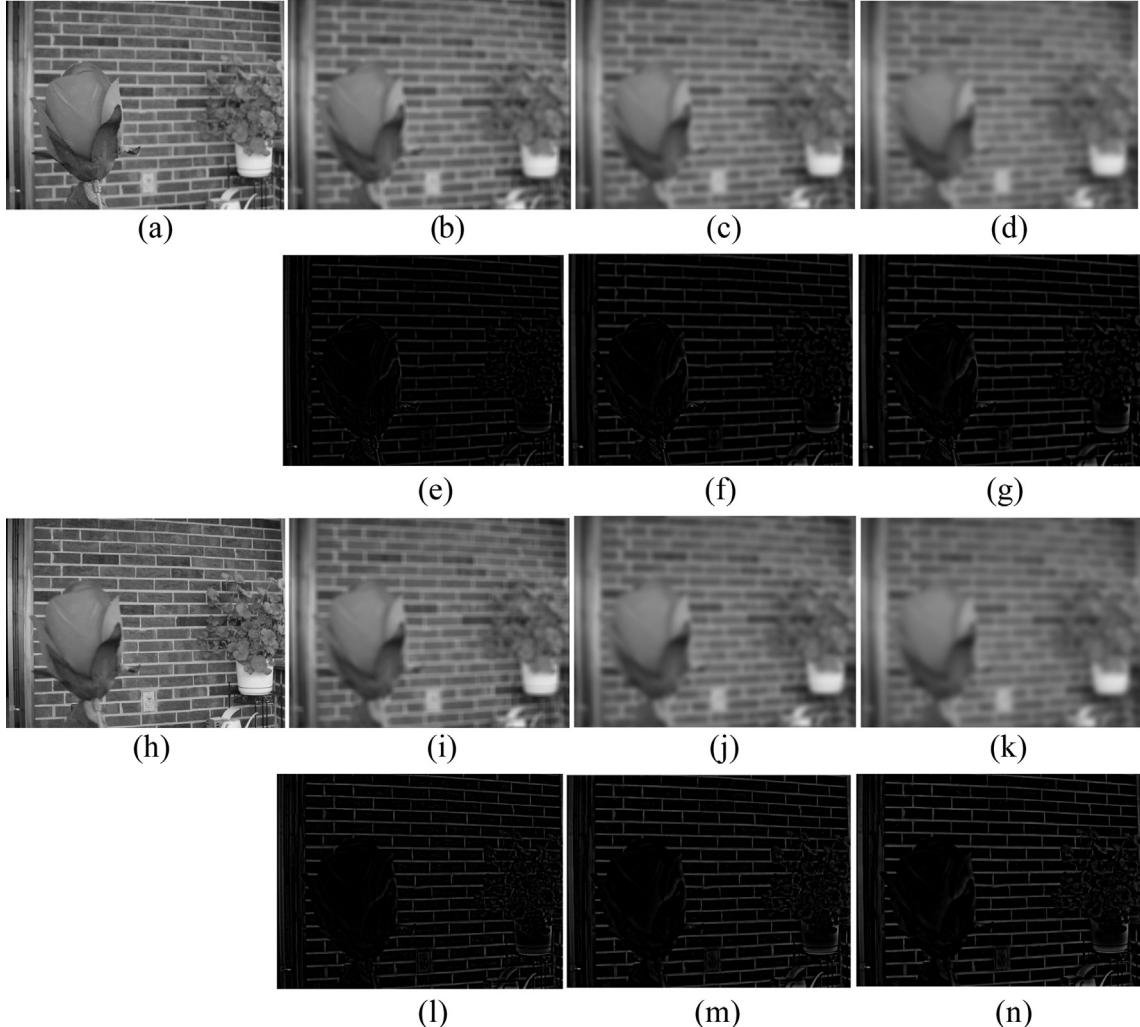
images using SD algorithms. We observe that the MSSS saliency detection algorithm is able to extract salient regions of the multi-focus images.

We prefer MSSS [29] compared to other SD methods due to the following facts:

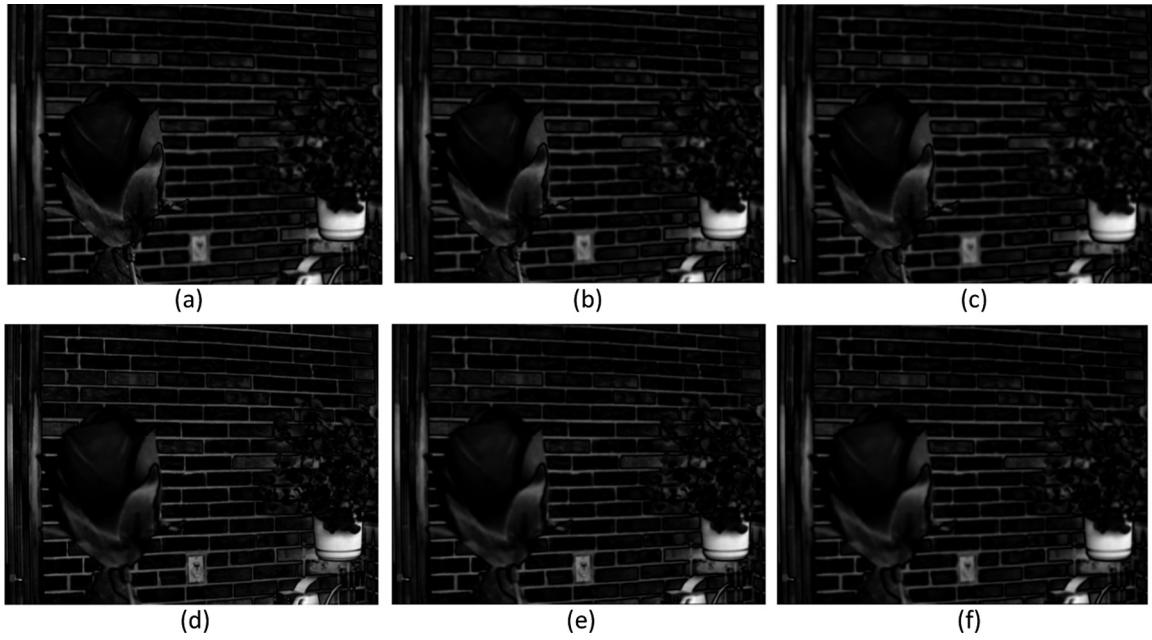
- (1) It generates full resolution saliency maps with well-defined boundaries.
- (2) It can effectively highlight the salient regions of images with complex background.

A brief review of MSSS detection algorithm theory is as follows.

Initially, the authors in [20] proposed a frequency-tuned saliency detection algorithm to use almost all low frequency information and a large portion of the high frequency information for acquiring perceptually qualitative saliency maps with full size. This saliency map is derived by taking the Euclidean separation between the mean of an image  $I_\mu$  and every pixel of the Gaussian blurred version  $I_f(u, v)$  of the same image.



**Figure 4** Approximation and detail layers of a flower dataset for three level decomposition ( $k = 3$ ): (a) left focused flower image, (b), (e) are approximation and detail layers of (a) at level-1, (c), (f) are approximation and detail layers of (a) at level-2, (d), (g) are approximation and detail layers of (a) at level-3, (h) right focused flower, (i), (l) are approximation and detail layers of (h) at level-1, (j), (m) are approximation and detail layers of (h) at level-2, (k), (n) are approximation and detail layers of (h) at level-3.



**Figure 5** Saliency maps of a flower dataset for three level decomposition: (a), (d) are saliency maps of left and right focused images at level-1, (b), (e) are saliency maps of left and right focused images at level-2, (c), (f) are saliency maps of left and right focused images at level-3.

$$S(u, v) = \|I_\mu - I_f(u, v)\|, \quad (1)$$

where  $S(u, v)$  is the saliency map at a pixel location  $(u, v)$ . Gaussian blur of size  $3 \times 3$  is chosen to get  $I_f(u, v)$ .

This algorithm fails if input image consists of a complex background. It assumes entire image as a common surround for all pixels in it [29] and detects the background as well as the saliency map. This is not preferable because for recognizing a pixel at the center of the salient object it needs to contain a little lower cutoff frequency. To identify a pixel close to the boundary, it is expected to contain a large high frequency cutoff. So as we approach image boundaries, we have to utilize nearby surround regions rather than common surround regions for distinguishing a given pixel [29]. It is possible by characterizing surround symmetry around the center pixel of its own sub image close to the boundary. It is achieved by the MSSS saliency map [29] defined for an image  $I$  of width  $w$  and height  $h$  as follows:

$$S_{ss}(u, v) = \|I_\mu(u, v) - I_f(u, v)\|, \quad (2)$$

where  $I_\mu(u, v)$  is the mean of the sub image centered at pixel  $(u, v)$  is denoted by the following:

$$I_\mu(u, v) = \frac{1}{A} \sum_{i=u-u_0}^{u+u_0} \sum_{j=v-v_0}^{v+v_0} I(i, j), \quad (3)$$

where  $u_0$ ,  $v_0$  indicate off-ssets and  $A$  denotes the area calculated as follows:

$$u_0 = \min(u, w - u), \quad (4)$$

$$v_0 = \min(v, h - v),$$

$$A = (2u_0 + 1)(2v_0 + 1).$$

The sub images obtained using Eqs. (3) and (4) are the maximum symmetric surround regions for a given central pixel. For a better visual understanding of the theory discussed so

far we consider multi-focus datasets shown in Fig. 1 and present their corresponding saliency maps in Fig. 2. We denote the MSSS visual saliency extraction process as follows:

$$S = \text{MSSS}(I), \quad (5)$$

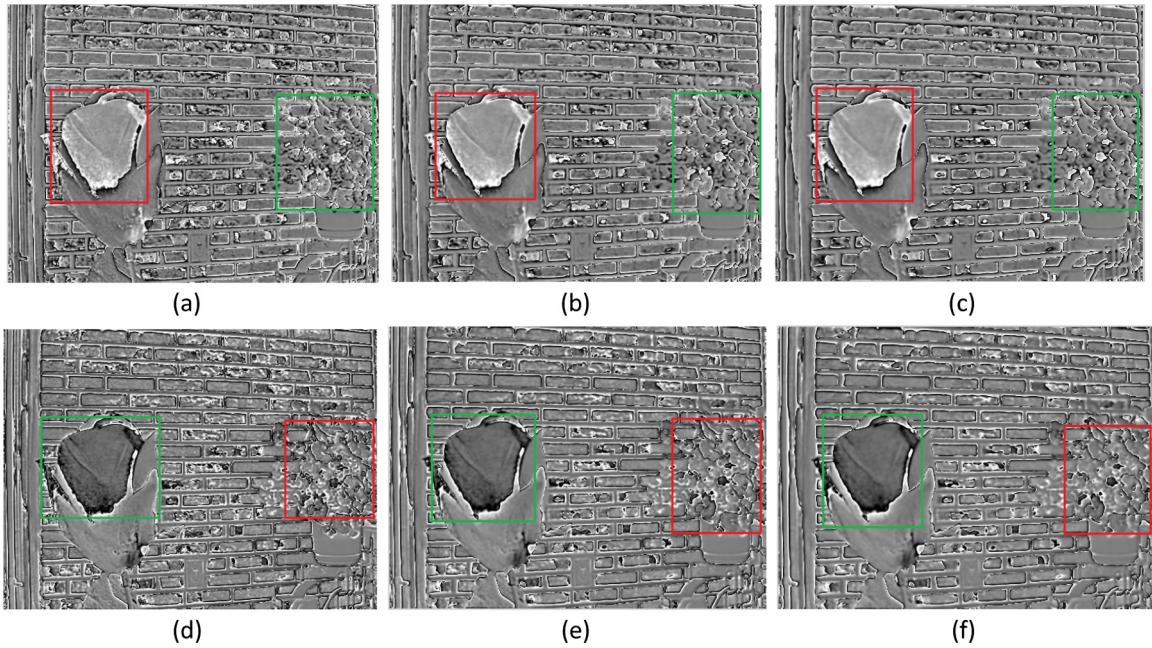
where  $I$  is the source image and  $S$  is the saliency map of it.

#### 4. Proposed methodology

The main objective of any MFF algorithm is to generate a more focused and sharpened fused image with less computational time. Here, we propose a simple and computationally efficient MFF algorithm based on saliency detection and multi-scale average filtering. In contrast to the existing multi-scale fusion methods (which use complex edge preserving filters [3,5,6,9], and transforms [8,15,17,18,31]) we use an average filter for this purpose. Hence, it is computationally simple. We exploit the computationally efficient MSSS detection algorithm [18] as reviewed in Section 3 for the purpose of fusion. We designed an optimal weight construction based on visual saliency with a simple normalization process, which is capable of identifying focused and defocused regions.

This algorithm is demonstrated in the schematic diagram illustrated in Fig. 3. It can be outlined in the following steps:

- Decompose the input images into approximation and detail layers of the required levels by employing an average filter.
- Compute visual saliences of each input image at different levels by using MSSS detection method.
- Determine weight maps from the extracted saliency maps by normalizing them.
- Multiply detail layers with the weight maps and merge weighted detail layers to obtain a final detail layer.



**Figure 6** Weight maps of a flower dataset for three level decomposition (red rectangle indicates the focused region, green rectangle indicates the defocused region): (a), (d) are weight maps of left and right focused images at level-1, (b), (e) are weight maps of left and right focused images at level-2, (c), (f) are weight maps of left and right focused images at level-3.

- E. Calculate the final approximation layer by averaging the approximation layers.
- F. Add final approximation and final detail layers to obtain the fused image.

This algorithm is detailed in the following subsections.

#### 4.1. Multi-scale image decomposition

Consider the input images  $\{I_n(x, y)\}_{n=1}^N$  of same size  $p \times q$  which are co-registered pixel by pixel. These  $N$ -images are decomposed into approximation layers containing large scale variations in intensity and detail layers containing small scale variations in intensity as follows:

$$B_n^{k+1} = B_n^k * A, \quad \text{where, } k = 0, 1, \dots, K \quad (6)$$

where  $B_n^{k+1}$  is the approximation layer of  $n$ -th source image at  $k+1$  level which depends on its previous level approximation layer  $B_n^k$ .  $B_n^0$  represents the  $n$ -th input image  $I_n$ . The convolution operation represented by  $*$ .  $A$  is an average filter.  $K$  is the number of levels. The detail layers  $D_n^{k+1}$  at present level  $k+1$  are obtained by subtracting approximation layers  $B_n^k$  at previous level  $k$  from approximation layers  $B_n^{k+1}$  at present level  $k+1$ .

$$D_n^{k+1} = B_n^{k+1} - B_n^k. \quad (7)$$

Multi-scale image decomposition of a flower for three level decomposition ( $k = 3$ ) is illustrated in Fig. 4.

#### 4.2. Visual saliency detection

Visual saliences of multi-focus images are obtained using MSSS detection algorithm [29]. This algorithm has been

reviewed in Section 3. The process of saliency extraction from the approximation layers  $B_n^k$  at  $k$  levels is represented as follows:

$$S_n^{k+1} = \text{MSSS}(B_n^k), \quad (8)$$

where  $S_n$  is the saliency map of  $n$ th source image.

To correlate this expression with actual images, visual saliences of a flower dataset for  $k = 3$  are displayed in Fig. 5.

#### 4.3. Weight map calculation

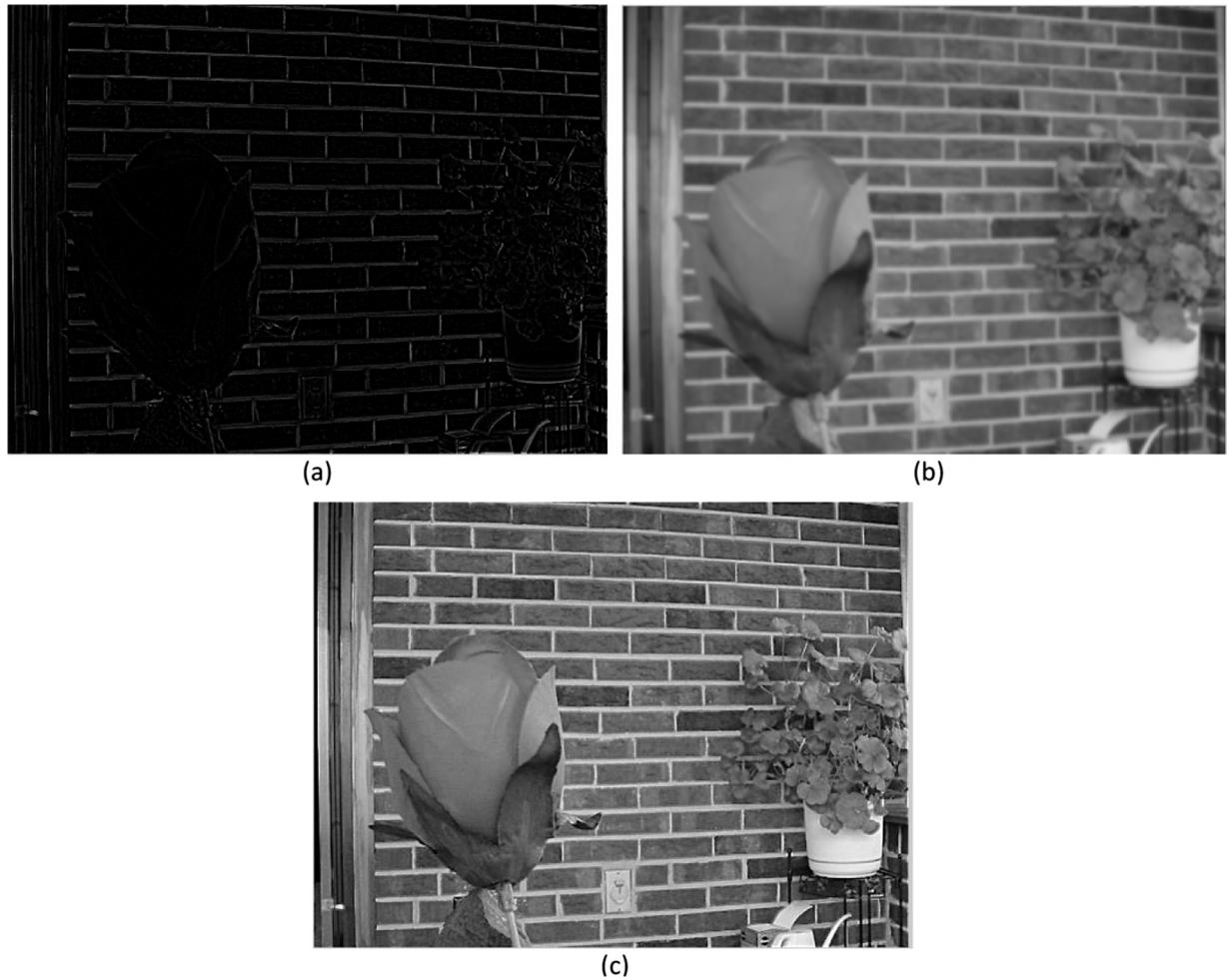
In MFF, each source image gives the information of a targeted focus region of a scene. We have to combine each focused region from these multiple images into a single image. This could be accomplished by assigning suitable weight maps to input images. These weight maps need to identify the focused and blurred (defocused) regions of the input images. Proposed weight maps are determined by normalizing the saliency maps as follows:

$$w_i^{k+1} = \frac{S_i^{k+1}}{\sum_{n=1}^N S_n^{k+1}}, \quad \forall i = 1, 2, \dots, N \quad (9)$$

As shown in Fig. 6, the proposed weight maps for a flower dataset (for three decomposition levels) are able to identify the focused and defocused regions. As an example, in Fig. 6 regions with red and green rectangles display focused and blurred regions of source images. These weight maps are complementary in nature.

#### 4.4. Detail layer fusion

For fusion the final detail layer  $\bar{D}$  (Fig. 7(a)) is derived as follows:



**Figure 7** Visual display of final approximation, detail layers and fused image. (a) Final detail layer, (b) final approximation layer, (c) proposed SDMF fused image.

$$\bar{D} = \sum_{k=1}^K \sum_{n=1}^N w_n^k D_n^k, \quad (10)$$

Clearly the detail layers  $D_n^k$  are given weight  $w_n^k$  to derive the final detail layer.

#### 4.5. Approximation layer fusion

Final approximation layer (Fig. 7(b)) is derived by averaging the approximation layers as follows:

$$\bar{B} = \frac{1}{NK} \sum_{k=1}^K \sum_{n=1}^N B_n^k. \quad (11)$$

#### 4.6. Fused image reconstruction

Fused image  $F$  (Fig. 7(c)) is generated by combining the final base ( $\bar{B}$ ) and final detail ( $\bar{D}$ ) layers as follows:

$$F = \bar{B} + \bar{D}. \quad (12)$$

## 5. Performance evaluation metrics

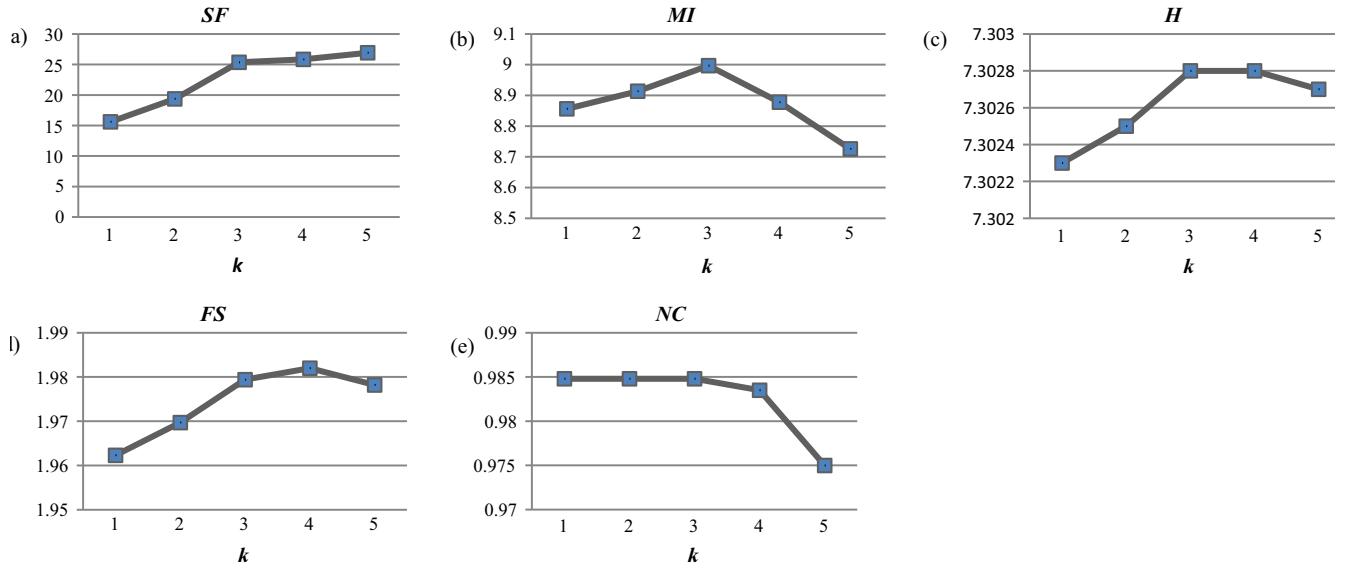
Assessment of a fusion method is really a challenging task when there is no ground truth available. Any image fusion algorithm can be assessed qualitatively by visual inspection and quantitatively by measuring the fusion metrics. We take fusion metrics such as spatial frequency [5,12,32], mutual information (MI) [33], entropy (H) [34], fusion symmetry (FS) [5,31,32] and normalized correlation (NC) [5,18,32].

For a fused image  $F(m, n)$  of size  $p \times q$ , we will be presenting various fusion metrics.

### 5.1. Spatial frequency (SF)

This metric [5,12,32] is used to find overall information level (activity level) present in the fused image and is given from the squares of row frequencies (RF) and column frequencies (CF).

$$SF = (RF^2 + CF^2)^{\frac{1}{2}},$$



**Figure 8** Affect of  $k$  on fusion metrics. (a)  $SF$ , (b)  $MI$ , (c)  $H$ , (d)  $FS$ , (e)  $NC$ .

$$\text{where, } RF = \sqrt{\frac{\sum_m \sum_n (F(m, n) - F(m, n-1))^2}{pq}}, \quad (13)$$

$$CF = \sqrt{\frac{\sum_m \sum_n (F(m, n) - F(m-1, n))^2}{pq}}.$$

### 5.2. Mutual information (MI)

It calculates the total information transferred from input images to the fused image [33] and it is denoted as follows:

$$MI = MI_{XF} + MI_{YF}, \quad (14)$$

where  $MI_{XF} = \sum_m \sum_n p_{X,F}(m, n) \log_2 \left( \frac{p_{X,F}(m, n)}{p_X(m)p_F(n)} \right)$  is the mutual information of input image  $X$  and fused image  $F$ . Here  $p_X(m)$  and  $p_F(n)$  denote the probability density functions of input images  $X$  and  $F$  respectively.  $p_{X,F}(m, n)$  is the joint probability density function of input image  $X$  and the fused image  $F$ . Similarly  $MI_{YF}$  is the mutual information between  $Y$  and  $F$ .

### 5.3. Entropy (H)

Entropy is the information quality index [22] given by

$$H = - \sum_{l=0}^{L-1} p_l \log_2(p_l), \quad (15)$$

where  $p_l$  is the probability of intensity value  $l$ .

### 5.4. Fusion symmetry (FS)

It measures the fused image closeness/symmetric nature with respect to the input images [5,32] or simply it represents the contribution of source image  $X$  (or  $Y$ ) to the fused image  $F$  [31].

$$FS = 2 - \left| \frac{MI_{XF}}{MI} - 0.5 \right|. \quad (16)$$

If both source images contribute equally to the fused image then the value of  $FS$  is closer to 2. So, the fused image will have better quality.

### 5.5. Normalized correlation (NC)

It indicates the amount of correlation of the fused image with the source images or it indicates the relevance of the resultant fused image to input images [5,18,32] and is measured as

$$NC = \frac{(r_{XF} + r_{YF})}{2}, \quad (17)$$

where  $r_{XF} = \frac{\sum_m \sum_n (x(m, n) - \bar{x})(f(m, n) - \bar{f})}{\sqrt{(\sum_m \sum_n (x(m, n) - \bar{x})^2)(\sum_m \sum_n (f(m, n) - \bar{f})^2)}}$ , is the normalized correlation between input image  $X$  and fused image  $F$ . Similarly,  $r_{YF}$  represents the correlation coefficient between input image  $Y$  and fused image  $F$ .

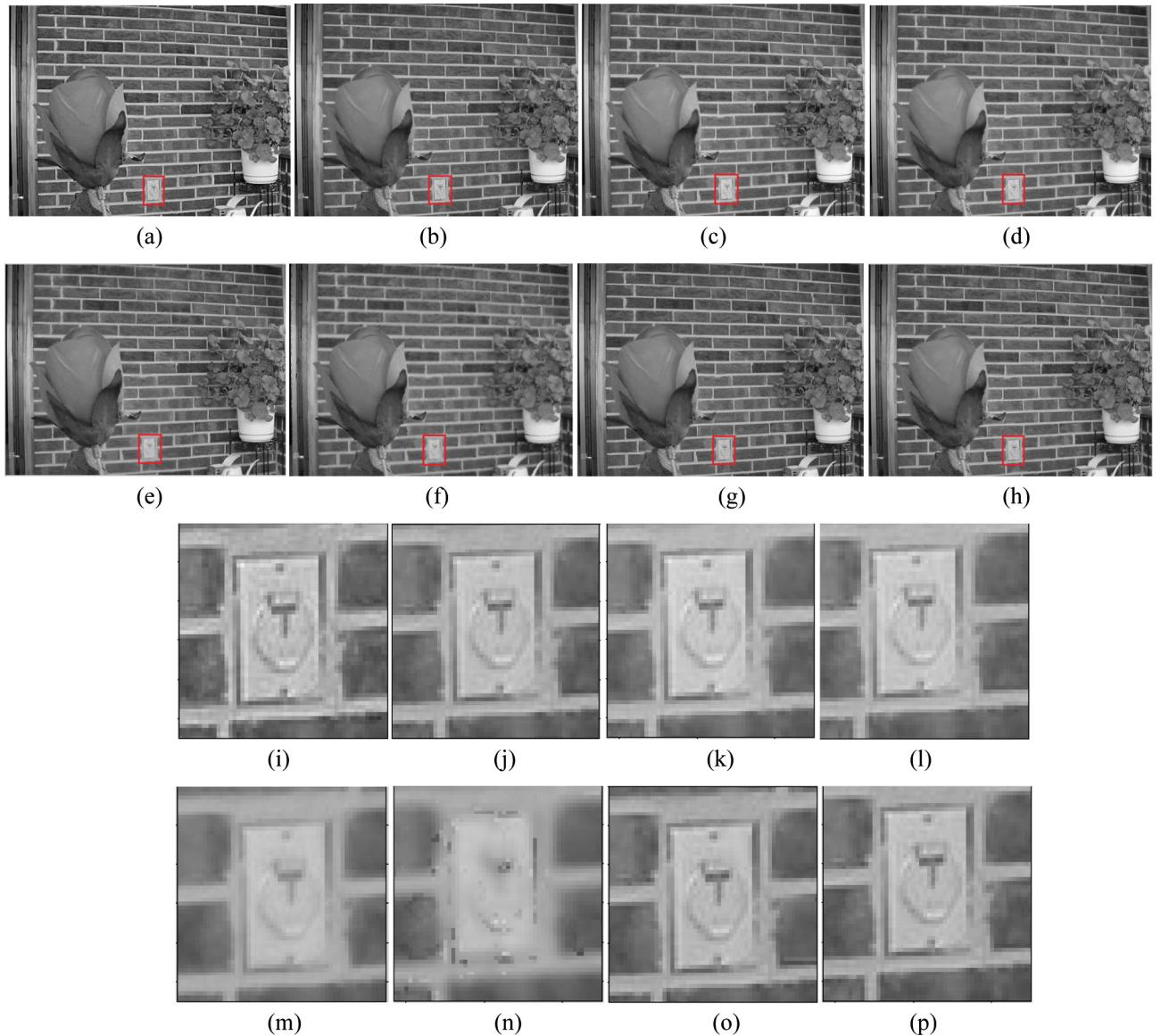
With these evaluation metrics, we assess the proposed method's performance. All these fusion metric values should be high for better visual quality.

## 6. Experimental setup

In this section, image database, existing MFF methods and free parameter analysis (effect on the performance of SDMF for change of  $k$ ) will be discussed.

### 6.1. Image database

Experiments are conducted on several multi-focus image datasets. Results and analysis for 10 image datasets viz. flower, leopard, book shelf, clock, air craft, pepsi, bottle, parachute, book, and flower wage are presented. These image datasets are shown in Fig. 1. They are available at [35].



**Figure 9** Comparison of visual quality of fused images of various methods for flower dataset (a) Proposed SDMF, (b) GFF, (c) CBF, (d) DCHWT, (e) DWT + AB, (f) BGS, (g) DCT + var + cv, (h) DCT + var. Subfigures (i)–(p) show the zoom version of switch portion of (a)–(h) respectively.

### 6.2. Other MFF methods for comparison

Our SDMF method is compared with spatial domain MFF method BGS [13] and multi-scale MFF methods DCT + var [16], DCT + var + cv [17], DCHWT [18], DWT + AB [19], BGS [13], CBF [5] and GFF [3]. Default parameter settings are adopted for all of these methods.

### 6.3. Free parameter analysis

In the proposed SDMF, fusion is done at multiple scales by going to various levels  $k$  with the help of an average filter. If the  $k$  increases then SDMF performance will also change. So, we have to analyze the performance of the SDMF with the change of  $k$ . This can be done by analyzing average fusion metrics  $SF$ ,  $MI$ ,  $H$ ,  $FS$ ,  $NC$  calculated over 10 multi-focus

image datasets (Fig. 1). As shown in Fig. 8, at  $k = 3$  our SDMF is giving better performance.

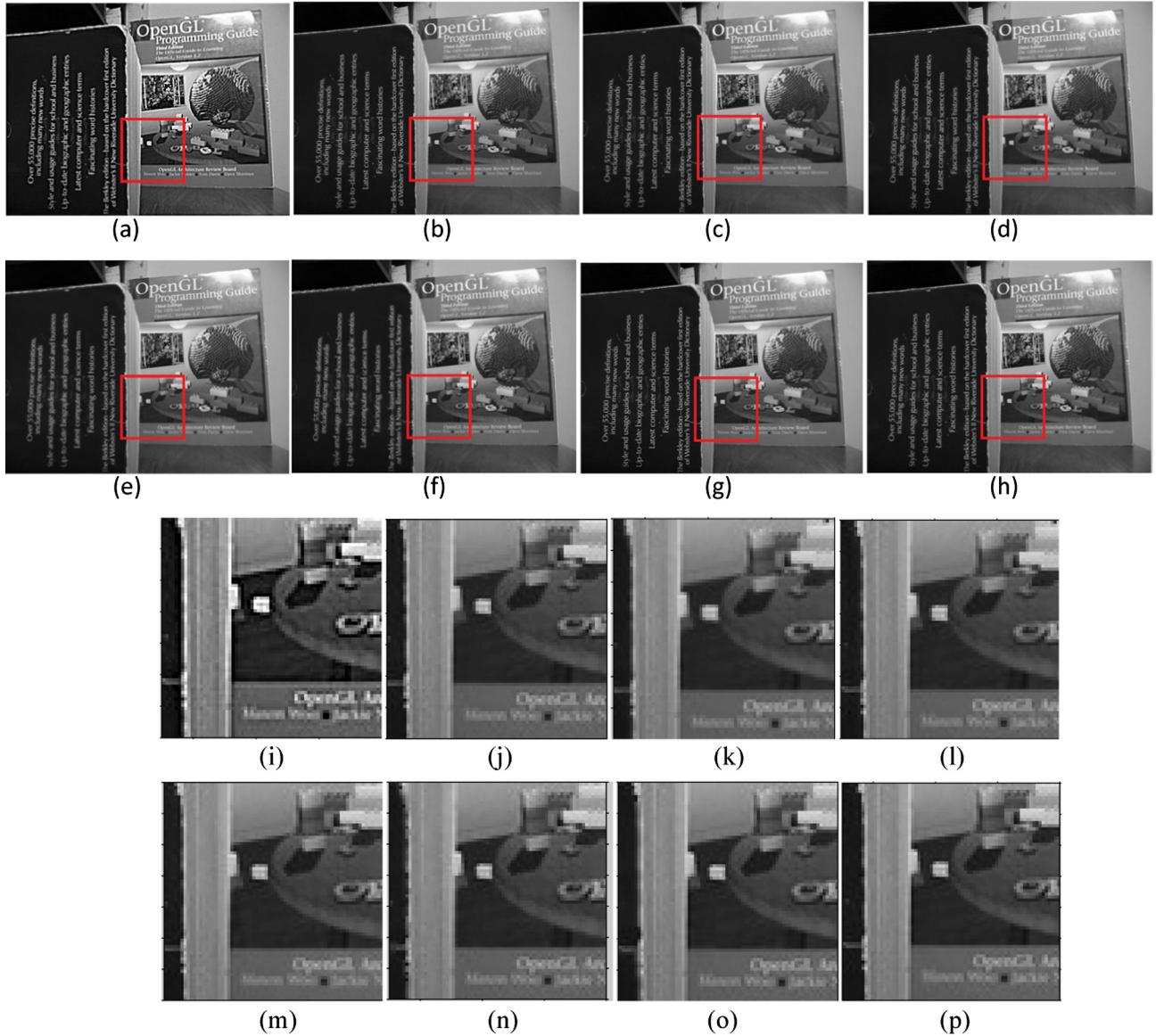
The size of the average filter is set to  $5 \times 5$  for better performance of the proposed SDMF algorithm.

## 7. Results and analysis

The aim of any MFF method is to obtain a properly focused image with less execution time. Performance of the MFF algorithm can be verified qualitatively by visual inspection, quantitatively using fusion metrics and by measuring the computational time.

### 7.1. Qualitative analysis

Qualitative analysis is presented for ten multi-focus image datasets namely flower, book, book shelf, clock, aircraft, pepsi,



**Figure 10** Comparison of visual quality of fused images of various methods for book dataset (a) Proposed SDMF, (b) GFF, (c) CBF, (d) DCHWT, (e) DWT + AB, (f) BGS, (g) DCT + var + cv, (h) DCT + var. Subfigures (i)–(p) show the zoom version of a particular focused region of (a)–(h) respectively.

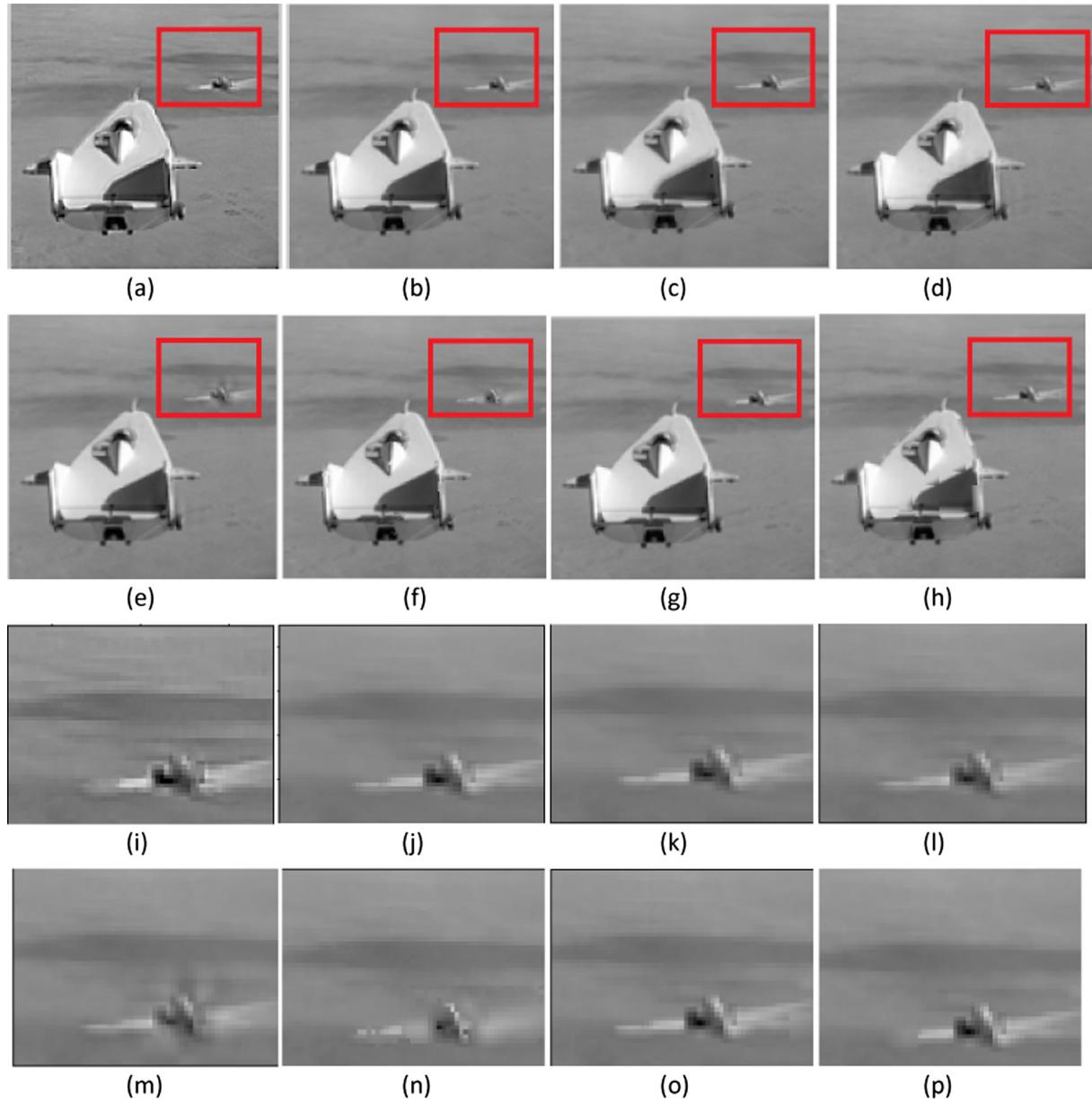
bottle, parachute, leopard, and flower wage datasets as shown in Fig. 1. Visual quality of the SDMF along with various MFF methods is presented in Figs. 9–11. For flower, book and aircraft datasets zoomed portion of a particular region of the fused image is also presented for in-depth qualitative analysis. In Figs. 9–11, subfigures (a)–(h) gives the visual display of our SDMF, GFF, CBF, DCHWT, DWT + AB, BGS, DCT + var + cv, and DCT + var respectively. The subfigures (i)–(p) illustrate the zoomed portions of (a)–(h) respectively.

In Fig. 9, the fused images for the flower dataset are displayed. As emphasized in Fig. 9(a) with red<sup>1</sup> rectangles, proposed SDMF is able to generate more focused regions (such as flower wage and switch). For an example, switch portions

(Fig. 9(m) and (n)) of DWT + AB and BGS methods are blurred. Remaining methods (Fig. 9(j)–(l) and Fig. 9(o) and (p)) are visually good but they are not able to provide more focused switch region. However, as shown in Fig. 9(i), SDMF method is able to get more sharpened switch region compared to the remaining methods.

Fig. 10 displays the visual quality of the results on book dataset. Zoomed portions of various MFF algorithms for book dataset highlighted in Fig. 10(a)–(h) (with red rectangles), are displayed in Fig. 10(i)–(p). Zoomed portions of GFF (Fig. 10(j)), CBF (Fig. 10(k)), DCHWT (Fig. 10(l)), DCT + var + cv (Fig. 10(o)) and DCT + var (Fig. 10(p)) are visually good. But, these methods are not reaching the desired quality. Zoomed portions of DWT + AB (Fig. 10(m)) and BGS (Fig. 10(n)) are visually distorted. However, as shown in Fig. 10(i), proposed SDMF is giving sharpened

<sup>1</sup> For interpretation of color in ‘Figs. 9 and 10’, the reader is referred to the web version of this article.



**Figure 11** Comparison of visual quality of fused images of various methods for aircraft dataset (a) Proposed SDMF, (b) GFF, (c) CBF, (d) DCHWT, (e) DWT + AB, (f) BGS, (g) DCT + var + cv, (h) DCT + var Subfigures (i)–(p) show the zoom version of a particular focused region of (a)–(h) respectively.

information about text, objects and book present in the fused image. Hence, SDMF integrates both foreground and background regions of source images in the fused image effectively compared to the state-of-the-art methods.

Similarly, for aircraft dataset as well, proposed SDMF combines more focused regions of source images with few artifacts compared to other methods as shown in Fig. 11. Zoomed portion of SDMF (Fig. 11(i)) provides more details of background and aircraft in the fused image when compared to the zoomed portions of the remaining methods (Fig. 11(j)–(p)).

Due to space constraint for remaining datasets (clock, book shelf, pepsi, bottle, parachute, flower wage, leopard) only fused images with corresponding source images are displayed in Fig. 12. As shown in Fig. 12, SDMF generates multi-focus

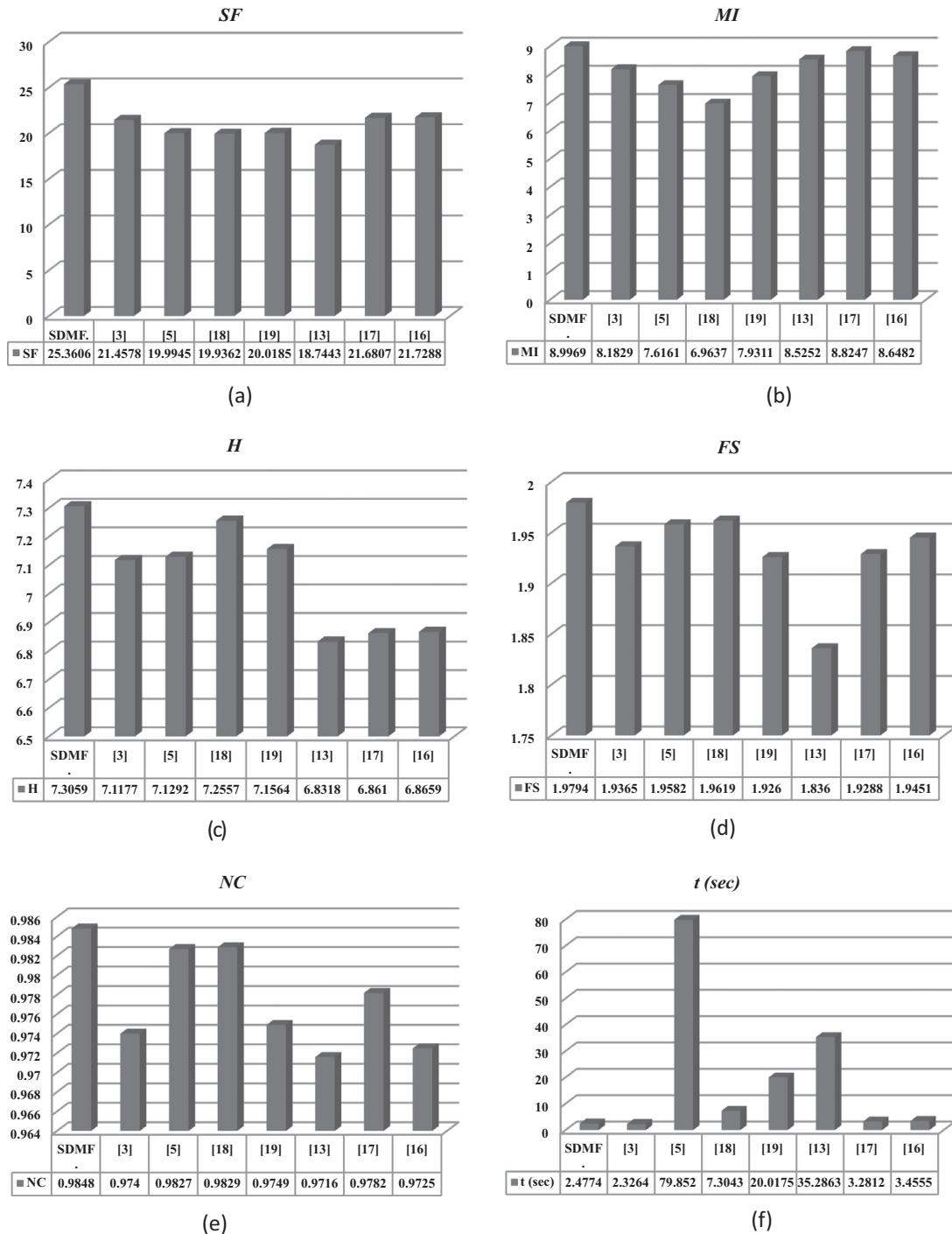
fused images for all the image datasets with more clarity and few artifacts.

## 7.2. Quantitative analysis

It is difficult to judge the performance of a fusion algorithm by visual inspection alone. Fusion algorithm has to be evaluated both qualitatively and quantitatively for better assessment. In the previous section qualitative analysis has been performed. Here, quantitative analysis is conducted by evaluating the fusion metrics (average  $SF$ ,  $MI$ ,  $H$ ,  $FS$ ,  $NC$  metrics calculated over 10 datasets) of other MFF algorithms (GFF [3], CBF [5], BGS [13], DCT + var [16], DCT + var + cv [17], DCHWT [18], DWT + AB [19]) along with our SDMF algo-



**Figure 12** Visual display of source and fused images of various datasets of the proposed SDMF method. (a)–(c) clock, (d)–(f) book shelf, (g)–(i) pepsi, (j)–(l) bottle, (m)–(o) parachute, (p)–(r) flower wage and (s)–(u) leopard.



**Figure 13** Quantitative analysis of various MFF methods along with the SDMF (average fusion metrics over 10 image datasets are considered). (a) SF, (b) MI, (c) H, (d) FS, (e) NC and (f)  $t$  in sec.

rithm. Bar chart comparison of the fusion metrics for various MFF algorithms is shown in Fig. 13(a)–(e). From this bar chart comparison, it is easy to observe that in all fusion metrics (Fig. 13(a)–(e)) our proposed SDMF got superior values.

### 7.3. Computational time $t$

Bar chart comparison of computational time  $t$  in seconds for various MFF methods is shown in Fig. 13(f). These experi-

ments are performed on a computer with 4 GB RAM and 2.2 GHz CPU. This  $t$  is averaged over 10 multi-focus image datasets. As shown in the bar graph, SDMF computational time is less than the algorithms proposed in [5,13,16–19] and more than that of [3].

From the above results and analysis, one can conclude that SDMF method is integrating more focused and sharpened regions of the source images. SDMF quantitative analysis against several state-of-the-art MFF methods [3,5,13,16–19],

using various fusion metrics (*SF*, *MI*, *H*, *FS*, *NC*) proves that it is outperforming the existing MFF methods. Its computational time is promising for a real time implementation.

## 8. Conclusion

In this paper, a new multi-scale multi-focus image fusion method using MSSS detection algorithm is proposed.

- Proposed SDMF generates fused images with more sharpened regions. It takes less computational time making it suitable for real time implementation.
- In contrast to most of the multi-scale fusion techniques, an average filter is employed for multi-scale decomposition purpose, which simplifies the implementation complexity.
- New MSSS detection is explored to extract visually significant regions from multi-focus images.
- A new weight map construction method based on visual saliency is developed which can detect and highlight focused and defocused regions of the source images at various scales.
- This method is both qualitatively and quantitatively compared and analyzed with recently proposed MFF methods. Results justify that our method yields better performance compared to existing methods.

## References

- [1] Goshtasby A Ardesir, Nikolov Stavri. Image fusion: advances in the state of the art. *Inform Fusion* 2007;8(2):114–8.
- [2] Zhang Zhong, Blum Rick S. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proc IEEE* 1999;87(8):1315–26.
- [3] Li Shutao, Kang Xudong, Hu Jianwen. Image fusion with guided filtering. *IEEE Trans Image Process* 2013;22(7):2864–75.
- [4] Choi Myungjin et al. Fusion of multispectral and panchromatic satellite images using the curvelet transform. *Geosci Rem Sens Lett IEEE* 2005;2(2):136–40.
- [5] Kumar BK Shreyamsha. Image fusion based on pixel significance using cross bilateral filter. *Signal Image Video Process* 2013;1–12.
- [6] Bavirisetti Durga Prasad, Dhuli Ravindra. Fusion of infrared and visible sensor images based on anisotropic diffusion and Karhunen-Loeve transform. *Sens J IEEE* 2016;16(1):203–9.
- [7] Stathaki Tania. *Image fusion: algorithms and applications*. Academic Press; 2011.
- [8] Zhang Qiang, Guo Bao-long. Multifocus image fusion using the nonsubsampled contourlet transform. *Signal Process* 2009;89(7):1334–46.
- [9] Singh Harbinder, Kumar Vinay, Bhooshan Sunil. Anisotropic diffusion for details enhancement in multiexposure image fusion. In: ISRN Signal Process.
- [10] Wan Tao, Zhu Chenchen, Qin Zengchang. Multifocus image fusion based on robust principal component analysis. *Pattern Recogn Lett* 2013;34(9):1001–8.
- [11] Naidu VPS, Rao JR. Pixel-level image fusion using wavelets and principal component analysis. *Def Sci J* 2008;58(3):338–52.
- [12] Huang Wei, Jing Zhongliang. Evaluation of focus measures in multi-focus image fusion. *Pattern Recogn Lett* 2007;28(4):493–500.
- [13] Tian J, Chen L, Ma L, Yu W. Multi-focus image fusion using a bilateral gradient-based sharpness criterion. *Opt Commun* 2011;284(1):80–7.
- [14] Rockinger O. *Multiresolution-Verfahren zur Fusion dynamischer Bildfolgen*. Ph.D. Thesis. Berlin: Technische Universität; 1999, ISBN 3-933342-77-5.
- [15] Rockinger Oliver. Image sequence fusion using a shift-invariant wavelet transform. In: International conference proceedings on image processing, vol. 3. IEEE; 1997. p. 288–91.
- [16] Haghigat Mohammad Bagher Akbari, Aghagolzadeh Ali, Seyedarabi Hadi. Real-time fusion of multi-focus images for visual sensor networks. In: Machine vision and image processing (MVIP), 2010 6th Iranian. IEEE; 2010.
- [17] Haghigat Mohammad Bagher Akbari, Aghagolzadeh Ali, Seyedarabi Hadi. Multi-focus image fusion for visual sensor networks in DCT domain. *Comput Electr Eng* 2011;37(5):789–97.
- [18] Kumar BK Shreyamsha. Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform. *Signal Image Video Process* 2013;7(6):1125–43.
- [19] Liu Yu, Wang Zengfu. Multi-focus image fusion based on wavelet transform and adaptive block. *J Image Graph* 2013;18(11):1435–44.
- [20] Achanta Radhakrishna et al. Frequency-tuned salient region detection. In: IEEE conference on computer vision and pattern recognition, 2009. CVPR 2009. IEEE; 2009.
- [21] Hong Richang et al. Salience preserving multi-focus image fusion. *IEEE international conference on multimedia and expo*, 2007. IEEE; 2007.
- [22] Gouiffès Michèle, Planes Bertrand, Jacquemin Christian. HTRI: high time range imaging. *J Vis Commun Image Represent* 2013;24(3):361–72.
- [23] Cui Guangmang et al. Detail preserved fusion of visible and infrared images using regional saliency extraction and multi-scale image decomposition. *Opt Commun* 2015;341:199–209.
- [24] Ma Yu-Fei, Zhang Hong-Jiang. Contrast-based image attention analysis by using fuzzy growing. Proceedings of the eleventh ACM international conference on multimedia. ACM; 2003.
- [25] Itti Laurent, Koch Christof, Niebur Ernst. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans Pattern Anal Mach Intell* 1998;20(11):1254–9.
- [26] Frintrop Simone, Klodt Maria, Rome Erich. A real-time visual attention system using integral images. In: International conference on computer vision systems.
- [27] Hou Xiaodi, Zhang Liqing. Saliency detection: a spectral residual approach. In: IEEE conference on computer vision and pattern recognition, 2007. CVPR'07. IEEE; 2007.
- [28] Harel Jonathan, Koch Christof, Perona Pietro. Graph-based visual saliency. Advances in neural information processing systems 2006.
- [29] Achanta Radhakrishna, Süstrunk Sabine. Saliency detection using maximum symmetric surround. 2010 17th IEEE international conference on image processing (ICIP). IEEE; 2010.
- [30] Khan Rizwan Ahmed, Konik Hubert, Dinet E. Enhanced image saliency model based on blur identification. 2010 25th International conference of image and vision computing New Zealand (IWCNZ). IEEE; 2010.
- [31] Arivazhagan S, Ganesan L, Kumar TG Subash. A modified statistical approach for image fusion using wavelet transform. *Signal Image Video Process* 2009;3(2):137–44.
- [32] Shah Parul, Merchant Shabbir N, Desai Uday B. Multifocus and multispectral image fusion based on pixel significance using multiresolution decomposition. *Signal Image Video Process* 2013;7(1):95–109.
- [33] Cover Thomas M, Thomas Joy A. *Elements of information theory*. John Wiley & Sons; 2012.
- [34] Leung Lau Wai, King Bruce, Vohora Vijay. Comparison of image data fusion techniques using entropy and INI. Paper presented at the 22nd Asian conference on remote sensing, vol. 5; 2001.
- [35] <http://home.ustc.edu.cn/~liuyu1/>.

- [36] Bavirisetti Durga Prasad, Dhuli Ravindra. Multi sensor image fusion using saliency map detection. *International review on computers and software (IRECOS)* 2015;10(7):757–63.
- [37] Bavirisetti Durga Prasad, Dhuli Ravindra. Two-scale image fusion of visible and infrared images using saliency detection. *Infrared Phys Technol* 2016;76:52–64.
- [38] Bavirisetti Durga Prasad, Dhuli Ravindra. Multi sensor image fusion based on visual saliency detection. *Int J Tomogr Simulat* 2016;29(2):102–11.



**Durga Prasad Bavirisetti** received B.Tech degree in electronics and communication engineering from Jawaharlal Nehru Technological University, Kakinada, India, in 2010, and M.Tech degree in communication engineering from VIT University, India in 2012. At present, he is pursuing Ph.D in school of electronics engineering, VIT University, India. His research interests are in signal and image processing.



**Ravindra Dhuli** was born in Tuni, Andhra Pradesh, India. He received his B.Tech in electronics and communication engineering from Bapatla engineering college, Andhra Pradesh, India in 2005, and Ph.D in signal processing from the department of electrical engineering, Indian institute of technology Delhi in 2010. He is currently an associate professor with the school of electronics engineering, VIT University, India. His research interests include multirate signal processing, statistical signal processing, image processing and mathematical modeling.