



Multi-focus image fusion combining focus-region-level partition and pulse-coupled neural network

Kangjian He¹ · Dongming Zhou¹ · Xuejie Zhang¹ · Rencan Nie¹ · Xin Jin¹

Published online: 8 March 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Multi-scale transforms (MST)-based methods are popular for multi-focus image fusion recently because of the superior performances, such as the fused image containing more details of edges and textures. However, most of MST-based methods are based on pixel operations, which require a large amount of data processing. Moreover, different fusion strategies cannot completely preserve the clear pixels within the focused area of the source image to obtain the fusion image. To solve these problems, this paper proposes a novel image fusion method based on focus-region-level partition and pulse-coupled neural network (PCNN) in nonsampled contourlet transform (NSCT) domain. A clarity evaluation function is constructed to measure which regions in the source image are focused. By removing the focused regions from the source images, the non-focus regions which contain the edge pixels of the focused regions are obtained. Next, the non-focus regions are decomposed into a series of subimages using NSCT, and subimages are fused using different strategies to obtain the fused non-focus regions. Eventually, the fused result is obtained by fusing the focused regions and the fused non-focus regions. Experimental results show that the proposed fusion scheme can retain more clear pixels of two source images and preserve more details of the non-focus regions, which is superior to conventional methods in visual inspection and objective evaluations.

Keywords Multi-focus image fusion · Focus region partition · Gaussian blurred · Pulse-coupled neural network · Nonsampled contourlet transform

1 Introduction

As an important branch of image fusion, multi-focus image fusion has been widely used in the field of machine vision, object recognition and artificial intelligence. Due to the limitation of the focusing ability of imaging devices, many natural scenes are not homogeneously focused, which means that there is clarity and detail information only in the focus region, and other regions are not easily observed and understood for human visual perception (Liu et al. 2015; Wang et al. 2010). In order to get a clear image which contains more targets and detail information of true scene, many different fusion techniques are proposed by researchers (Yang

2011; Jin et al. 2016; Zheng et al. 2007; Jin et al. 2017; Tian and Chen 2012).

The key point of image fusion is to extract useful information from the images which are collected from multi-source channels in the same scene so that the fused image has higher accuracy, quality and reliability compared with single image (Zhang and Guo 2009). In recent years, multi-focus image fusion methods based on multi-scale transforms (MST) have become popular, such as Laplacian pyramid (LP), wavelet transform (WT), discrete wavelet transform (DWT) and non-sampled contourlet transform (NSCT) (Amolins et al. 2007; Zhong et al. 2014; Kavitha and Thyagarajan 2016) (He et al. 2017; Li et al. 2017). These MST-based methods fuse pixels directly according to different fusion strategies, in which the visual effects of the fused image are better compared with other traditional methods. However, these pixel-based fusion methods which only operated the pixels of the source images by different fusion strategies were easily disturbed by noise and misregistration; moreover, most of pixel-based methods cannot completely retain the clear pixels and more detail information of the focused region to the fused

Communicated by V. Loia.

✉ Dongming Zhou
zhoudm@ynu.edu.cn

✉ Xuejie Zhang
xjzhang@ynu.edu.cn

¹ Information College, Yunnan University, Kunming 650504, China

image (Aslantas and Toprak 2014; Li et al. 2006). Another popular multi-focus fusion method is block-based, the main idea of this method is to extract the focused region from the source images, and then these regions were processed and integrated to produce new image. The advantage of this method is that the sharpness and accuracy of the corresponding region in the fused image can be maximally achieved. However, the key problem of this method is how to distinguish the sharpness of the block image and whether it is the focused region. If the block which contains the focused region and non-focus regions is fused into the final image directly, the detail information of non-focus region will be lost (Li and Yang 2008; Li et al. 2004). According to these problems, we proposed a new fusion scheme named multi-focus image fusion combining focus-region-level partition and pulse-coupled neural network (PCNN). The proposed fusion scheme has the advantages of multi-scale transforms (MST)-based method and block-based method.

First step, we differentiate the different regions of the source image based on the sharpness. In order to achieve this objective, we construct an evaluation function to measure these regions. In the two fully registered images of the same scene, the clearer image contains more edges and detail information. And we can measure the edges and detail information of the image by using edge intensity function and gray difference, so that by analyzing the difference between the two indexes of the images, we can determine that which one is the focused image and which one is the non-focus image. In addition, using principal component analysis (PCA) of the image, the obtained eigenvalues and eigenvectors can also intuitively reflect the basic characteristics of the image data matrix. We use the maximum eigenvalue of the image as the third index to measure the sharpness degree. In order to prove the effectiveness of the three indexes, we choose a completely focused image and then process Gaussian blurred to get a new image, which can be regarded as non-focus image. Finally, three indexes of the two images are calculated, respectively, and the experimental results prove the feasibility of the indexes. This is equally valid for measuring the focused regions and the corresponding non-focus regions of two images, which the method is described in Sect. 2.

Back-propagation (BP) neural network has been widely used in the artificial neural networks area because of its excellent characteristics since it was proposed (Wang and Jeong 2017). Next, we use BP neural network to learn and train the three indexes, so as to use the training model to measure the sharpness of the regions in the source image. After dividing the regions of the source image, we obtain the focused regions and the non-focus regions, and then different fusion strategies are selected to fuse the non-focus regions. Since the focused region contains most of the information and clear pixels compared with the non-focus region, we fuse it directly into the final fused image. For the non-focus regions, since

the regions contain clear pixels and fuzzy pixels which cannot fuse them to the final image directly, we use MST-based method to process them.

For these MST-based methods, NSCT is used in processing the non-focus regions in our scheme. Since NSCT was proposed by da Cunha et al. (2006), it has been widely used in image fusion due to its many advantages; NSCT is a multi-scale and multi-direction decomposition tool, which has many good properties, such as time–frequency localization, shift invariance, anisotropy; more importantly, it can effectively avoid frequency aliasing phenomenon (Li et al. 2013; Zhao et al. 2015). The non-focus regions of two source images are decomposed by NSCT to get one lowpass subimage and a series of bandpass subimages, respectively. In dealing with lowpass subimage, due to the characteristics of low-frequency image, Gaussian blurred is used in the lowpass subimage in the proposed fusion algorithm. In the bandpass subimage, the fusion method based on PCNN is used to process it. PCNN was proposed as a neural network model which has been widely used in image processing (Johnson and Padgett 1999); due to its excellent characteristics, it has been widely applied in image processing areas, such as image segmentation, image enhancement, image edge detection, image fusion (Qu et al. 2008; Jin et al. 2016a, b) and the PCNN-based methods in image fusion are effective (Xiang et al. 2015). We use the spatial frequency (SF) metric of high-frequency subimages as external incentive information of the PCNN model, which makes it better to deal with overexposed or weak exposure images.

The remaining sections are organized as follows: clarity evaluation functions and focus-region-level partition are presented in Sect. 2. The nonsampled contourlet transform and pulse-coupled neural network model are briefly reviewed, and the introduction of proposed fusion algorithm in detail is given in Sect. 3. Experimental results and analysis are given in Sect. 4. Finally, conclusion is shown in Sect. 6.

2 Focus-region-level partition

The first step of the proposed algorithm is to divide the source multi-focus image into the focused regions and the non-focus regions according to different focusing levels. In order to evaluate the definition of the image, we selected three evaluation functions, which will be described in detail as follows.

2.1 Principal component analysis

The principal component analysis (PCA) is a commonly used method of data analysis which is mainly used to extract the main characteristic components of data (Kaya et al. 2017). By finding a projecting relation, PCA projects high-dimensional data onto low-dimensional data subspace, which purpose is

to find the orthogonal directions of strong variability in the source data. Assume a set of data X has n -dimensional independent vectors $\{x_i\}$, where $i = 1, \dots, n$, the dimension reduction data y can be obtained by

$$y_i = A^T(x_i - \mu), \tag{1}$$

where $\mu = \frac{1}{n} \sum_{i=1}^n x_i$ is the sample mean of $\{x_i\}$, A is the orthogonal transformation matrix, which is consisted of the orthonormal eigenvectors of the sample covariance matrix, and the covariance matrix can be defined as

$$S = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T. \tag{2}$$

And the eigenvalues and their corresponding eigenvectors of S can be calculated. Then the eigenvalues are arranged in order from large to small; we use the largest eigenvalue to measure the different resolution images of the same scene; the following experiment shows the feasibility and effectiveness of that.

Figure 2 shows the images, the first image is the source image, which is a focused image, and the second to fourth is processed by Gaussian blurred with different σ , which can be regarded as non-focus image. To show the Gaussian blurred image which can be used to simulate non-focus image, the experimental simulation results are given in Fig. 1. From Fig. 1, it can be seen that there is a small difference between the non-focus image and the Gaussian blurred image, and it is feasible to use Gaussian blurred image to simulate the non-focus image (Kumar et al. 2018). It can be seen that with the increasing of σ , the image becomes increasingly blurred. And the largest eigenvalues of these images are given in Table 1. It can be seen from Table 1, the more blurred image is, and

Table 1 The largest eigenvalues of the images

Source images	Focused images	$\sigma = 1$	$\sigma = 2$	$\sigma = 5$
First	94.9817	108.5568	113.5932	114.7155
Second	59.7529	64.2693	66.4521	67.0978

the greater the corresponding largest eigenvalue value is. The source image in Fig. 2 can be seen as the focused region in the multi-focus image, and the blurred image can be seen as the non-focus region; then we can calculate the eigenvalues of the corresponding regions of two multi-focus images to measure which regions are the focused regions.

2.2 Edge intensity function

Sobel operator is a discrete differential operator, which is mainly used for edge detection (Al-Nima et al. 2017; Gupta et al. 2016), and can be computed by

$$G = \sqrt{G_x^2 + G_y^2}, \tag{3}$$

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I, \tag{4}$$

where I is the source image. And (4) can be written as

$$G_x = [f(x - 1, y + 1) + 2f(x, y + 1) + f(x + 1, y + 1)] - [f(x - 1, y - 1) + 2f(x, y - 1) + f(x + 1, y - 1)], \tag{5}$$

$$G_y = [f(x - 1, y - 1) + 2f(x - 1, y) + f(x - 1, y + 1)] - [f(x + 1, y - 1) + 2f(x + 1, y) + f(x + 1, y + 1)], \tag{6}$$

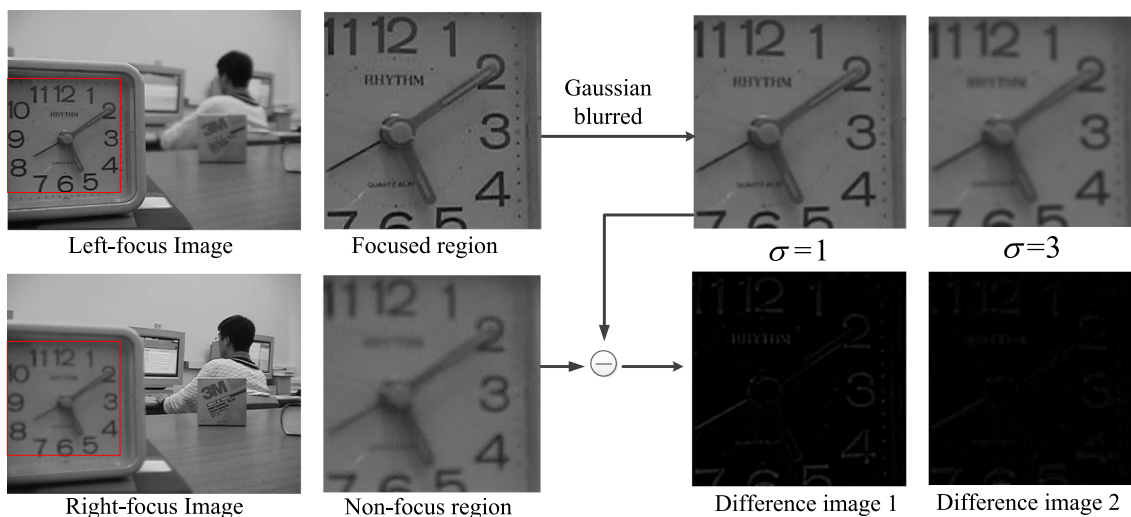


Fig. 1 Comparison between the non-focus image and the simulated non-focus images

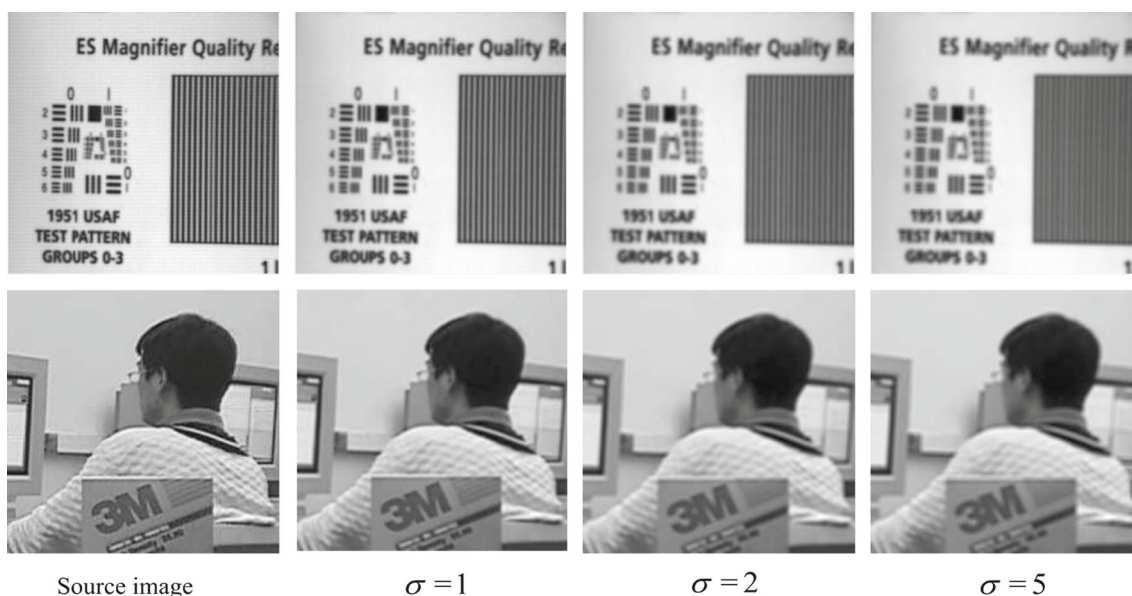


Fig. 2 The images with different Gaussian blurred

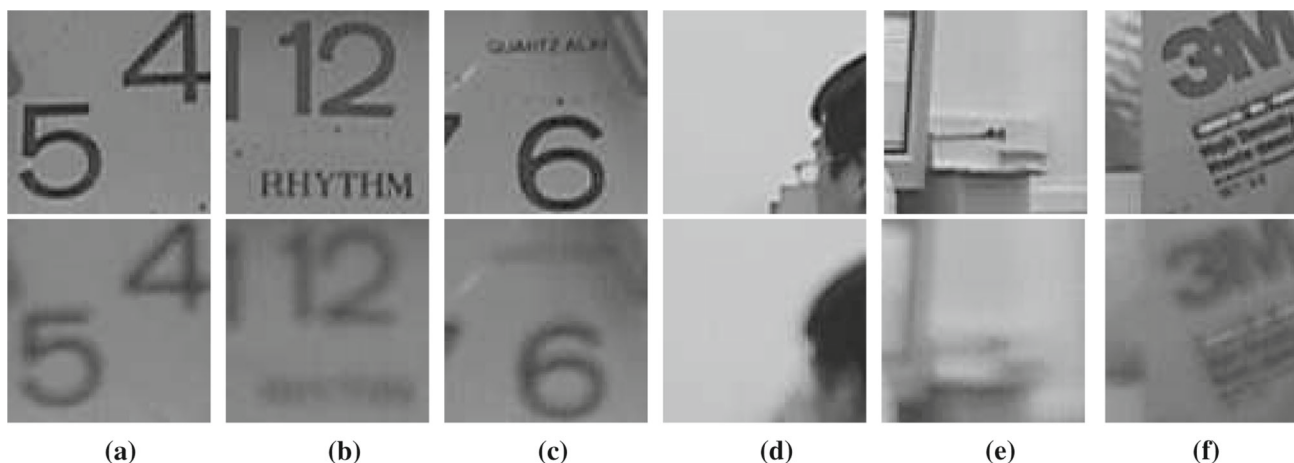


Fig. 3 Focused images and corresponding non-focus images

where $f(x, y)$ is the gray value at pixel (x, y) in the source image I . For each pixel (x, y) , we can obtain the corresponding value of G . Assume that the source image $I=(a_{ij})_{m \times n}$, as shown in Fig. 3, the edge intensity image $G=(G_{ij})_{m \times n}$ can be obtained by (3), which is shown in Fig. 4. We define $g = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n G_{ij}$, and use g to represent the intensity of the image edge. The edge intensity values of g for Fig. 4 are given in Table 2.

The Sobel edge operator can provide more edges information, and it contains more accurate edges direction information of the image pixels where the gradient is the highest. In general, the focused image will contain more edges' detail information compared with the non-focus image, so using edge intensity to measure the sharpness is feasible. The experiments are shown in Figs. 3 and 4, the focused region images and the corresponding non-focus region images of six

multi-focus images are given in Fig. 3, and the edge intensity images are shown in Fig. 4.

It can be seen from Fig. 3 that the focused image contains more detail and clearer edge information. Table 2 shows the edge intensity values of Fig. 3, in which a larger edge intensity value means that the image contains more edge details.

2.3 Gray-level difference

The gray image is usually obtained by measuring the brightness of each pixel in the visible light spectrum; therefore, the gray value of the image can reflect the local feature; and the gray-level difference is commonly used in image enhancement and image segmentation (Ji et al. 2016). In this paper, we improve the image gray-level difference as the

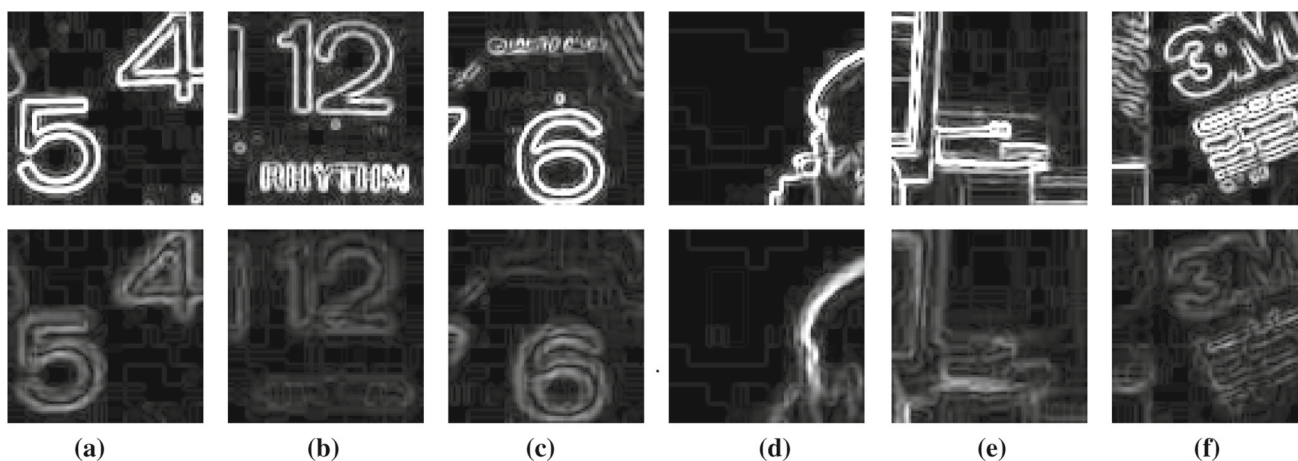


Fig. 4 Edge intensity images of Fig. 3

Table 2 Edge intensity values of Fig. 4

Images in Fig. 4	(a)	(b)	(c)	(d)	(e)	(f)
Focused	52.5592	60.0059	56.6855	81.3162	29.4241	70.5593
Non-focus	22.8651	26.0865	19.2199	25.1976	19.4316	28.0627

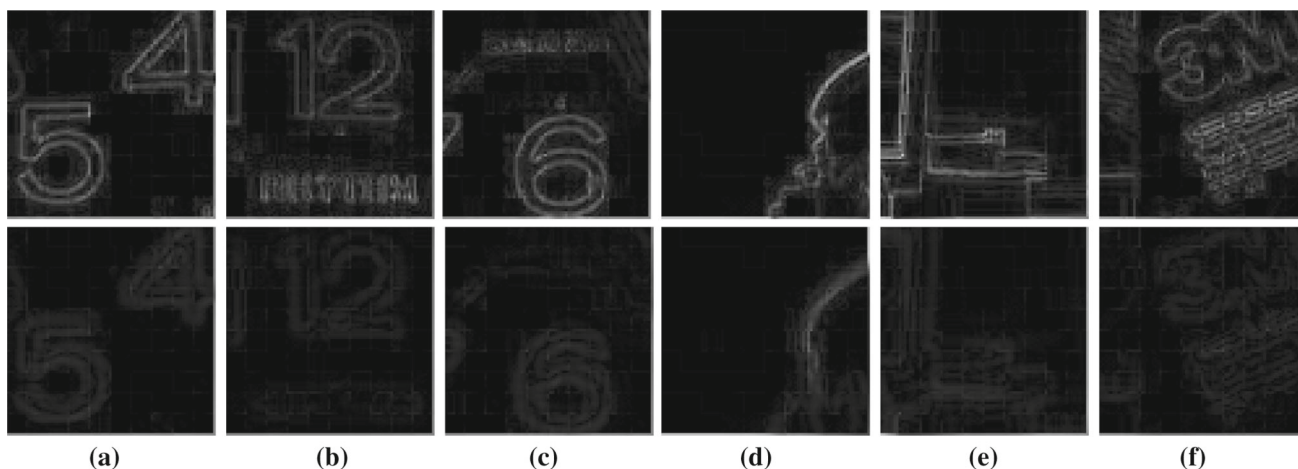


Fig. 5 The gray-level difference images of Fig. 3

Table 3 Absolute values of gray-level difference of Fig. 3

Images in Fig. 3	(a)	(b)	(c)	(d)	(e)	(f)
Focused	10.3697	11.1413	11.2891	9.6659	12.2416	16.5095
Non-focus	3.8742	4.3217	3.1120	3.3003	4.4238	4.5688

third evaluation index. Suppose that an image is $m \times n$, the absolute value of gray-level difference d can be defined as follows:

$$I_x = I(x + 1, y) - I(x, y), \tag{7}$$

$$I_y = I(x, y + 1) - I(x, y), \tag{8}$$

$$d = \frac{\sum_{x=1}^m |I_x| + \sum_{y=1}^n |I_y|}{m \times n}. \tag{9}$$

The gray-level difference images of Fig. 3 are shown in Fig. 5, and the d of them is given in Table 3.

2.4 Focus-region-level partition by BP neural network

The key idea of BP neural network is to revise the network weighted by training sample data to minimize the errors, so as to make the results approximate the desired output. It has

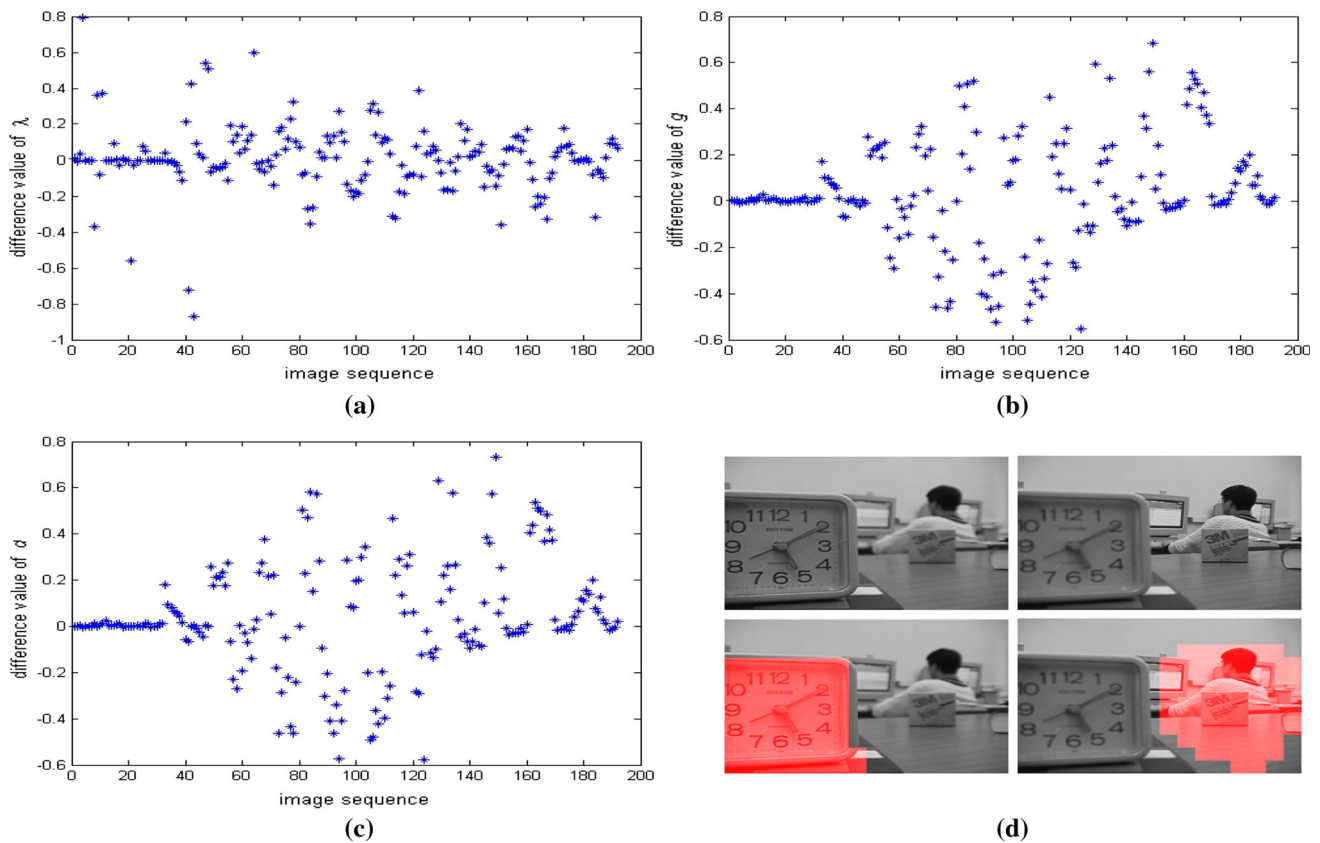


Fig. 6 Three-parameter diagram of two multi-focus images. **a** λ of the source image. **b** g of the source image. **c** d of the source image. **d** Source image and focus-region-level partition

been widely used in function approximation, pattern recognition, classification and other fields. BP neural network can be simply described by (10–13):

$$X_i = \sum \omega_{ij}u_j, \tag{10}$$

$$Y_i = f(X_j), \tag{11}$$

where X is the input, Y is the output of hidden layer, and the transformation function is:

$$f(x) = \frac{1}{1 + e^{-x}}. \tag{12}$$

And the output O of the output layer can be defined as follows:

$$S_k = \sum \omega_{jk}Y_j, \tag{13}$$

$$O_k = f(S_k), \tag{14}$$

where Y is the input of the output layer, and assume the desired output is P , then the output error is:

$$E = \frac{\sum_k (O_k - P_k)^2}{2}. \tag{15}$$

From Sects. 2.1 to 2.3, we can learn the three parameters (λ , g , d) and measure the definition of the image. In order to train the BP neural network, the differences of the three parameters between the two images are used as BP training parameters. An experimental result of the multi-focus image is given in Fig. 6, and the size of the source image is 640×480 . We divide it into the 40×40 size image and get 192 images. And then we calculate the differences of normalized λ , g and d of two source images, which are described as:

$$\lambda = \lambda_1 / \max(\lambda_1) - \lambda_2 / \max(\lambda_2), \tag{16}$$

$$g = g_1 / \max(g_1) - g_2 / \max(g_2), \tag{17}$$

$$d = d_1 / \max(d_1) - d_2 / \max(d_2). \tag{18}$$

The three parameters of two multi-focus images are given in Fig. 6. The x -axis shows the sequence number of images, and the y -axis represents the difference of parameter between the two corresponding multi-focus images. The values of λ , g and d can be used to measure which region is the focused region. In order to show the differences of these parameters in detail, the detail values of g in Fig. 6c are shown in Fig. 7.

In Fig. 6 the relationship between the three indexes and their corresponding subimage’s clarity is shown. We use the

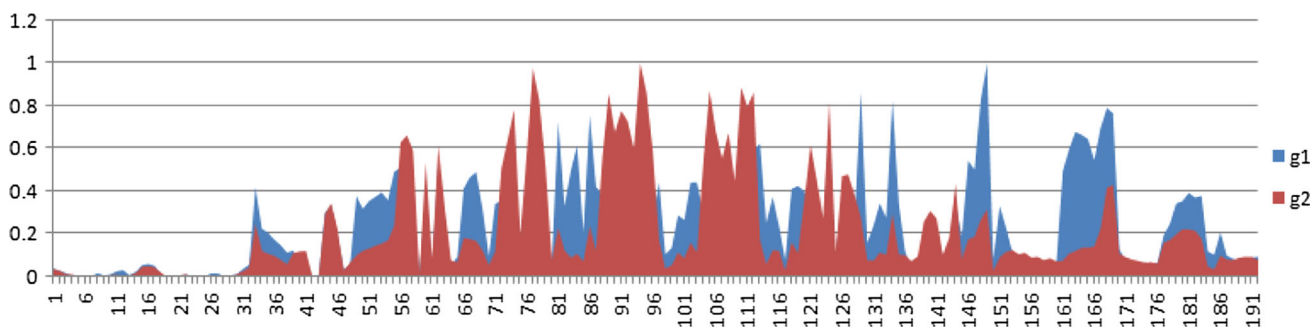


Fig. 7 Edge intensity g of two multi-focus images

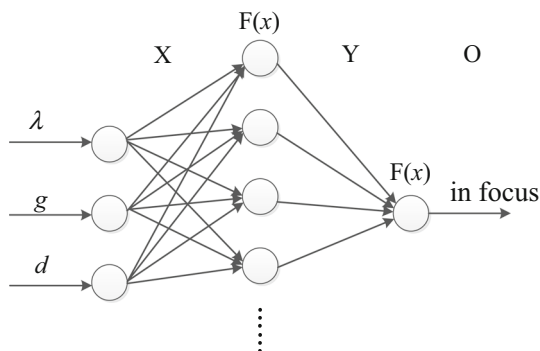


Fig. 8 BP neural network model

three indexes as the training data of BP, and BP neural network model is shown in Fig. 8.

The BP neural network training parameters are set to the expected error is 10^{-6} , and the maximum number of training iteration of network is 5000. We selected ten fully focused images, processed Gaussian blurred, then divided each image into the 40×40 size image and obtained 192 image sequences, and then 1920 pairs of focused image sequences and blurred image sequences are obtained, which are used to

train BP neural network. We create the BP neural network: $net=newff(minmax(P),[6,6,1],\{'tansig', 'tansig', 'purelin'\}, 'trainlm')$. BP consists of three parts: one input layer which has three neurons, two hidden layers which each layer has six neurons and one output layer which is to determine whether the input image is focused image or not. And we set *tansig* as the transfer function of the hidden layers, which is described as:

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \tag{19}$$

Then we can divide the focused region of the multi-focus image by the trained BP neural network. And the experimental results of focus region partition are given in Fig. 9.

3 Proposed fusion scheme

The proposed fusion scheme will be discussed in detail in this section, which is shown in Fig. 10. As discussed above, the focused region of the source multi-focus image which we have separated should be fused into the final image as much as possible. Next we discuss the non-focus region fusion rules.

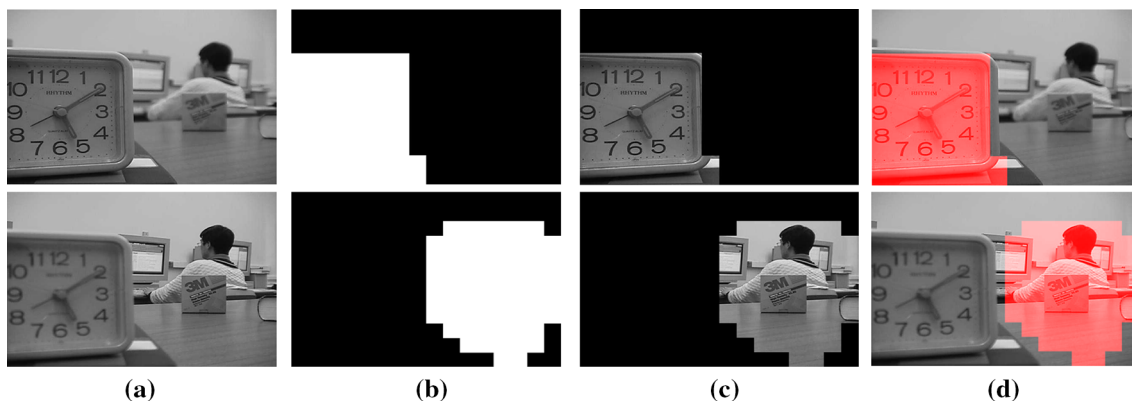


Fig. 9 Experimental results of focus region partition. **a** Multi-focus images. **b** Location of the focused region. **c** Focused region of the multi-focus image. **d** Focus-region-level partition

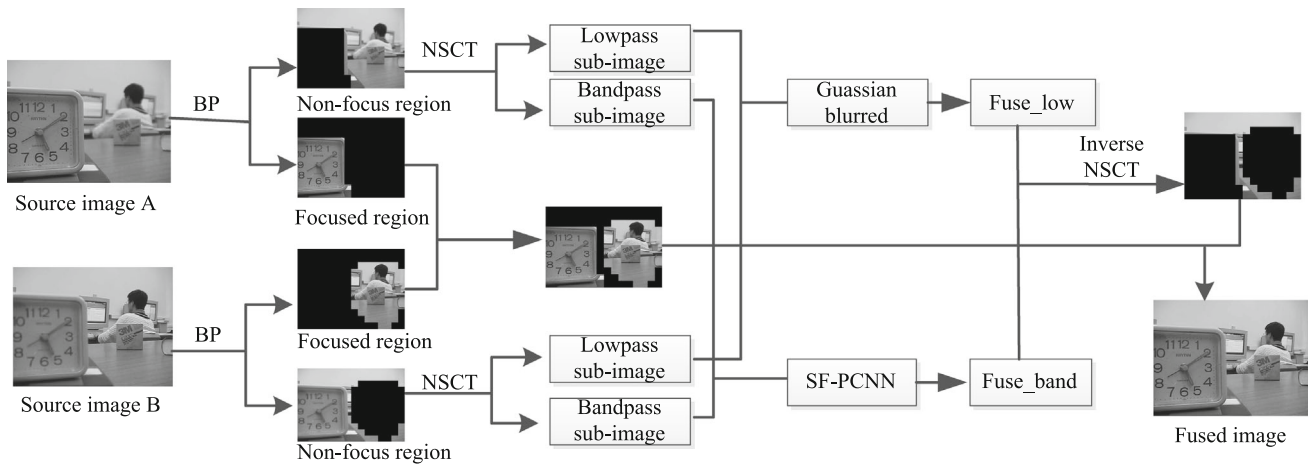


Fig. 10 Proposed fusion method flow chart

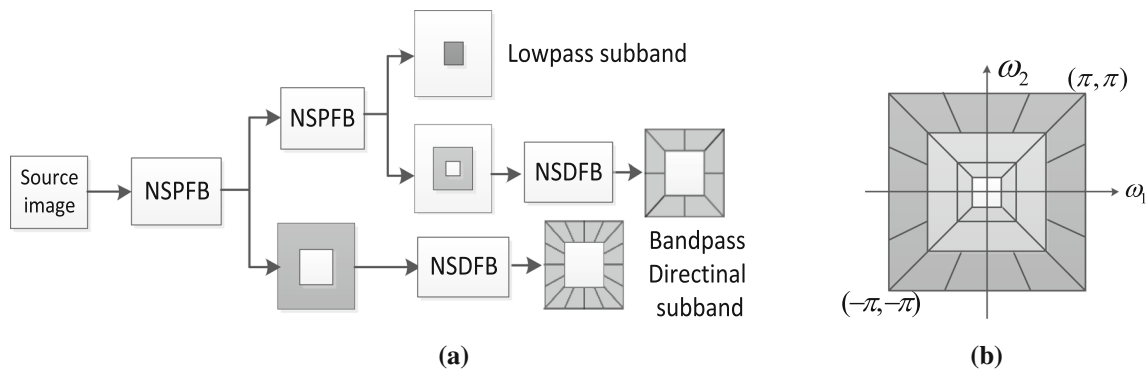


Fig. 11 NSCT framework. a The decomposition framework of NSCT. b Ideal frequency partitioning

For the non-focus regions, we will fuse them by Gaussian blurred and PCNN-based method in the NSCT domain. We will make a brief introduction for NSCT and PCNN as follows.

3.1 Nonsubsampled contourlet transform

Nonsubsampled contourlet transform (NSCT) is an effective tool for image decomposition, which is derived from contourlet transform (CT), and the NSCT frame is shown in Fig. 11.

As shown in Fig. 11, NSCT contains two parts: the nonsubsampled pyramids filter banks (NSPFB) and nonsubsampled directional filter banks. Therefore, it has many advantages, such as avoiding frequency aliasing phenomenon. The two filter banks endow the ability of multi-scale and multi-direction decomposition in NSCT. One image is decomposed by NSCT to obtain a lowpass image and a series of bandpass images, and the subimages are same size as the source image. Figure 12 gives an example of NSCT decomposition.

3.2 Pulse-coupled neural network (PCNN)

The basic neuron of PCNN can be divided into three parts: the receptive field, the modulation field and the pulse generator, which are shown in Fig. 13.

The receptive filed, modulation field and pulse generator can be described in detail by (20)-(25)

$$F_{ij}(n) = V_F \sum_{k,l} W_{kj} Y_{kl}(n-1) + F_{ij}(n) e^{-\alpha_F} + S_{ij}, \tag{20}$$

$$L_{ij}(n) = e^{-\alpha_L} L_{ij}(n) + V_L \sum_{kl} M_{kj} Y_{ijkl}(n-1), \tag{21}$$

$$U_{ij}(n) = F_{ij}(n) [1 + \beta L_{ij}(n)], \tag{22}$$

$$\theta_{ij}(n) = e^{-\alpha_\theta} \theta_{ij}(n-1) + V_T Y_{ij}(n-1), \tag{23}$$

$$Y_{ij}(n) = \begin{cases} 1, & U_{ij}(n) > \theta_{ij}(n) \\ 0, & \text{otherwise} \end{cases}, \tag{24}$$

$$T_{ij}(n) = \begin{cases} n, & \text{if } Y_{ij}(n) = 1, \text{ for the first time} \\ T_{ij}(n-1), & \text{otherwise} \end{cases}. \tag{25}$$

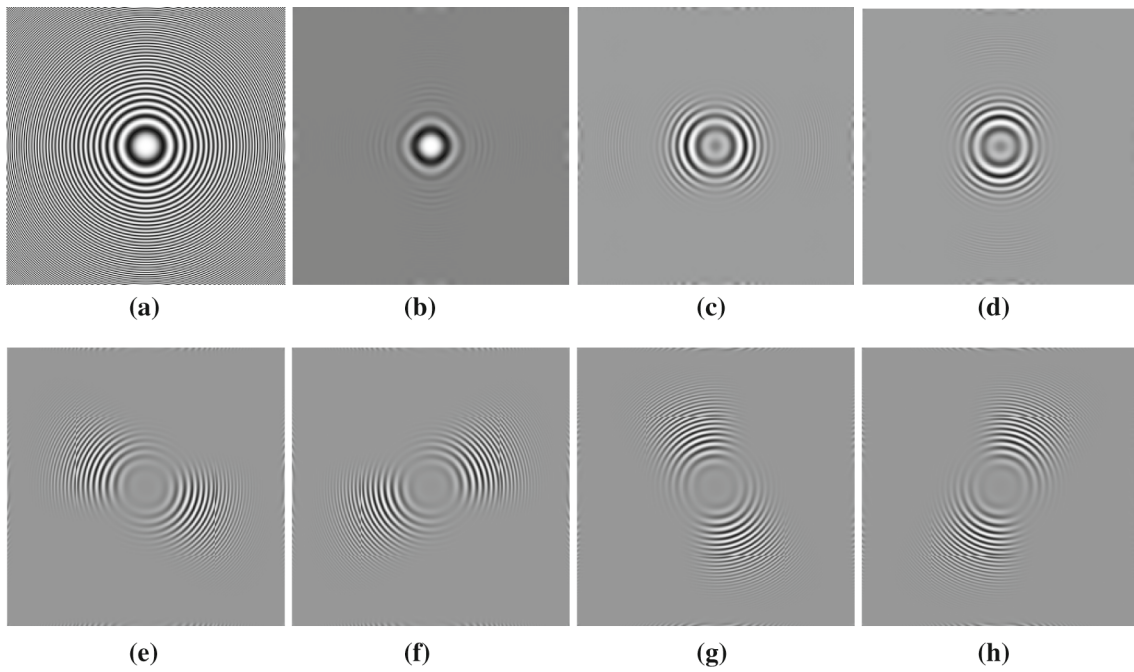


Fig. 12 NSCT decomposition example. **a** Source image. **b** Lowpass image. **c,d** Images of the detailed coefficients at level 1. **e,f,g,h** Images of the detailed coefficients at level 2

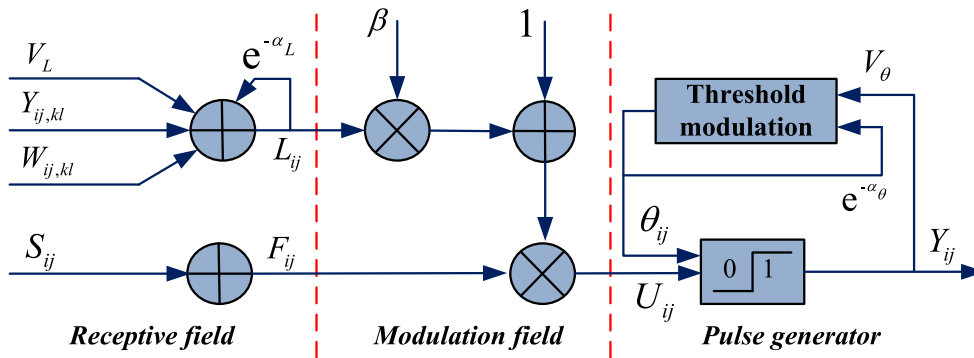


Fig. 13 The typical structure of PCNN

In (20) and (21), S_{ij} is the input stimulus at pixel (i, j) in the source image. F_{ij} is the feeding input of the pixel. Matrices M and W are the constant synaptic weighted. β is the linking strength of the neuron. α_F and α_L are the time constants. And Y_{ij} is the output of the neuron at (i, j) . The linking coefficient β is a key parameter in the PCNN model, which can vary the weighting of the linking channel. And in our proposed fusion method, the spatial frequency (SF) of the bandpass subimages will be used as the linking coefficient β .

3.3 Non-focus region fusion rules

For the non-focus regions of the source multi-focus images, one lowpass image and a series of bandpass images are

obtained by NSCT decomposition firstly. Then we fuse them with different rules.

3.3.1 Lowpass subband fusion rules

The common processing methods in low-frequency domain are weighted average-based methods; these lowpass subimage fusion rules usually cannot retain the low-frequency information to the final results perfectly, which will loss the partial details of the source image. To solve this problem, we use Gaussian blurred method (Jamal and Karim 2012) to fuse the lowpass subimage in our proposed method, which can be described as follows:

$$C_F^L(i, j) = w_A(i, j) \times C_A^L(i, j) + w_B(i, j) \times C_B^L(i, j), \tag{26}$$

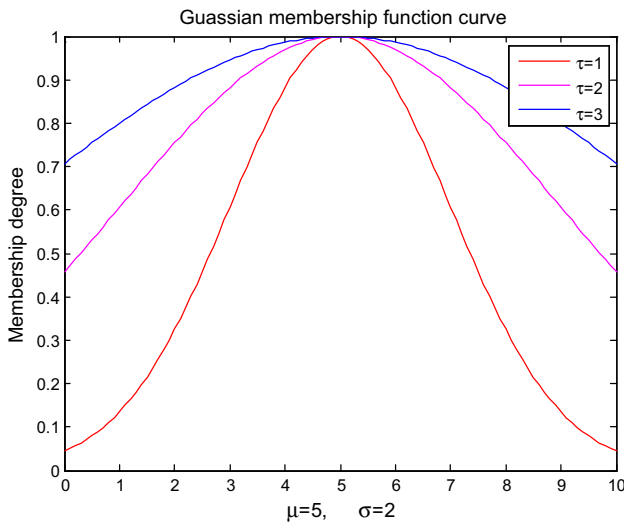


Fig. 14 Gaussian blurred curve with different τ

$$w_A(i, j) = \exp \left[-\frac{(C_B^L(i, j) - \mu)^2}{2(\tau\sigma)^2} \right], \tag{27}$$

$$w_B(i, j) = 1 - w_A(i, j). \tag{28}$$

In (26), $C_F^L(i, j)$ is the low-frequency coefficient of the final result, $C_A^L(i, j)$ is the low-frequency coefficient of the non-focus regions in the source image A, $C_B^L(i, j)$ is the low-frequency coefficient of the non-focus regions in the source image B. Equations (27) and (28) are the fusing weighted coefficients. In (27), μ and σ are the mean and variance of the non-focus regions in the source image B, and τ is the adjustment factor of Gaussian blurred function. And the Gaussian blurred function curve is shown in Fig. 14. In the proposed method, we set $\tau = 1$.

3.3.2 Bandpass subband fusion rules

We can see from Fig. 12 that the high-frequency details of information, texture and edge are included in the bandpass images. In the proposed method, PCNN is used to process these bandpass images. In the PCNN model, due to the spatial frequency (SF) can reflect the overall definition level of the source image, we use the SF of the input image to determine the linking strength β . SF is described as follows:

$$SF = \sqrt{RF^2 + CF^2}, \tag{29}$$

$$RF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=2}^N [F(i, j) - F(i, j - 1)]^2}, \tag{30}$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=2}^N [F(i, j) - F(i, j - 1)]^2}. \tag{31}$$

In (29), RF is the spatial low frequency, CF is the spatial column frequency, F is the image, the size of F is $M \times N$, and the fused bandpass coefficients $C_{F,ij}$ can be determined

as follows:

$$C_{F,ij} = \begin{cases} C_{A,ij}, T_{A,ij}(n) \geq T_{B,ij}(n) \\ C_{B,ij}, T_{A,ij}(n) < T_{B,ij}(n) \end{cases}, \tag{32}$$

$C_{A,ij}$ and $C_{B,ij}$ are the bandpass coefficients of the non-focus regions in source image A and B. $T_{A,ij}(n)$ and $T_{B,ij}(n)$ denote time matrix of each neuron, which is obtained by (25).

3.4 Fusion steps

The fusion framework presented in this paper is shown in Fig. 10, which can be described concretely as follows:

Input: Source multi-focus images A and B.

- Step 1* Perform the focus-region-level partition by trained BP neural network and get the focused regions and the non-focus regions of the source images.
- Step 2* Perform NSCT in the non-focus regions and then obtain one lowpass subband image and a series of bandpass subband images for each source image.
- Step 3* For the lowpass subband image, Gaussian blurred fusion algorithm is used to produce the fused low-pass coefficient, which is described by (26–28).
- Step 4* For the bandpass subband images, SF-PCNN is used to produce the fused bandpass coefficients, which is described by (29–31).
- Step 5* Fused non-focus regions are produced by NSCT reconstruction.
- Step 6* Fuse the focused regions and the fuse non-focus regions to produce the final fusion image.

4 Experimental results and analysis

In this section, we use three groups of experiments to illustrate the feasibility and practicability of the proposed fusion algorithm. And all simulations are conducted in MATLAB 2014a. We first introduce the experimental parameters setting and then discuss the fusion results compared with other methods.

5 Experiments introduction

In all experiments, the parameters of PCNN are set as $\alpha_\theta = 0.2$, $\alpha_L = 0.05$, $V_L = 0.02$, $V_\theta = 40$, $N = 200$, $M = W = [0.70710.707; 101; 0.70710.707]$, and NSCT-based method, we set “pkva” and “9-7” as the pyramid and the direction filter. Moreover, we choose the other six current fusion methods as the comparison algorithms: PCA-based method, discrete wavelet transform (DWT)-based method, PCNN-based method, NSCT-based method, Laplacian pyramid transform

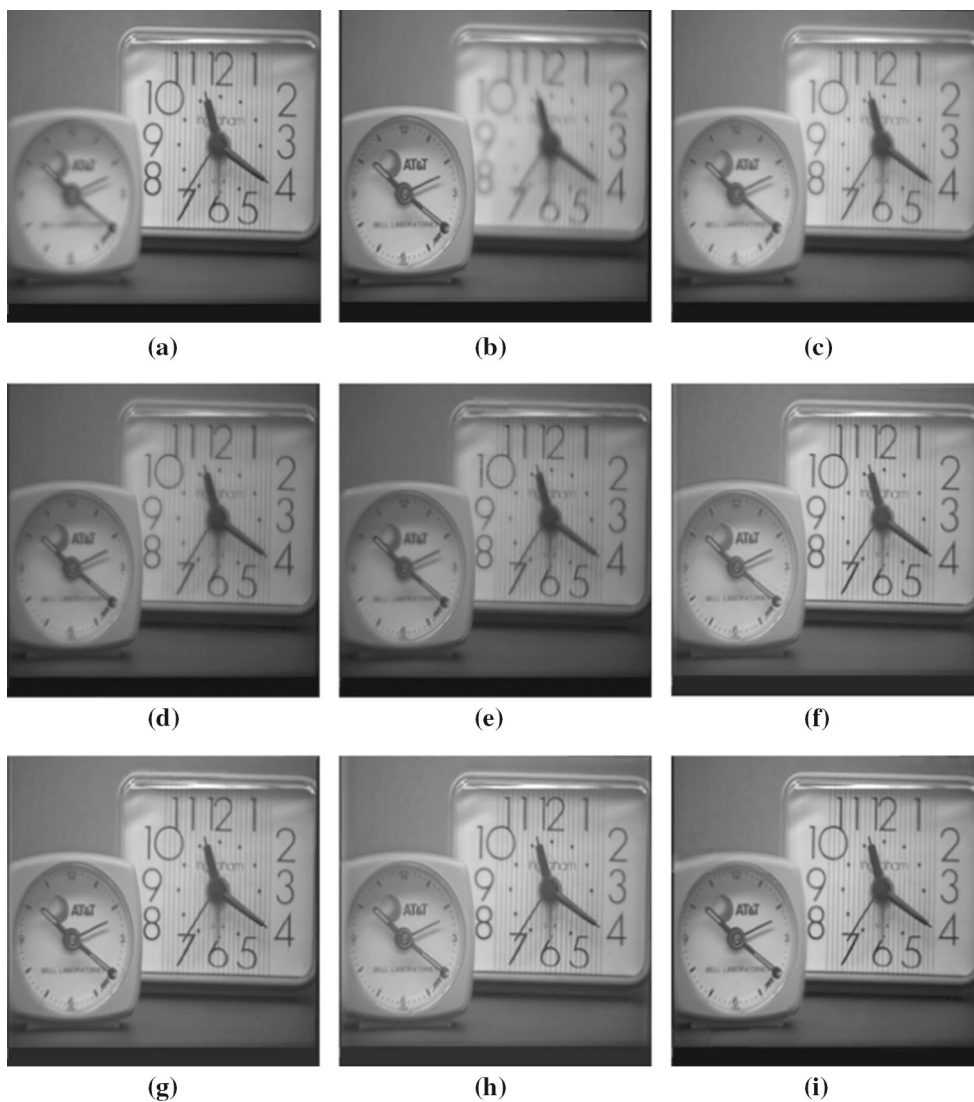


Fig. 15 First experimental results using different methods. **a, b** Multi-focus images. **c** PCA. **d** DWT. **e** PCNN. **f** NSCT. **g** LP-PCNN. **h** NSCT-PCNN. **i** Proposed method

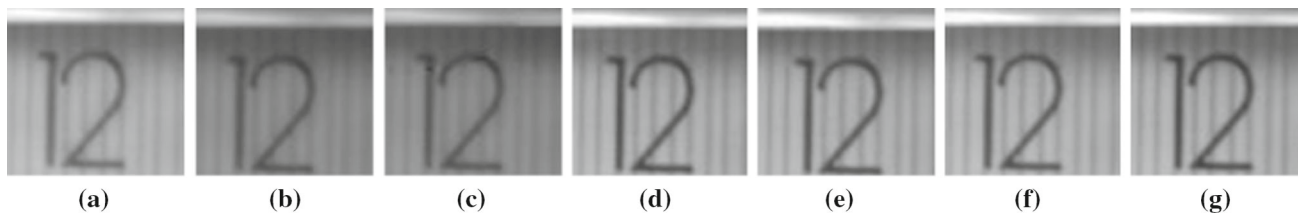


Fig. 16 Details of enlarged scale. **a** PCA. **b** DWT. **c** PCNN. **d** NSCT. **e** LP-PCNN. **f** NSCT-PCNN. **g** Proposed method

(LP)-PCNN-based method and NSCT-PCNN-based method. And for multi-scale decomposition methods, the decomposition level is set to 3, and we use “averaging” to fuse the lowpass subimages, and the bandpass subimages are fused by “absolute maximum choosing.”

In order to objectively illustrate the fusion results, we adopt three objective indicators as the evaluation indexes: mutual information (MI), pixel of visual information (VIF) and edge gradient operator ($Q^{AB/F}$). MI is used to measure the amount of the information of the source images retained in the fused image. VIF is an evaluation index for human

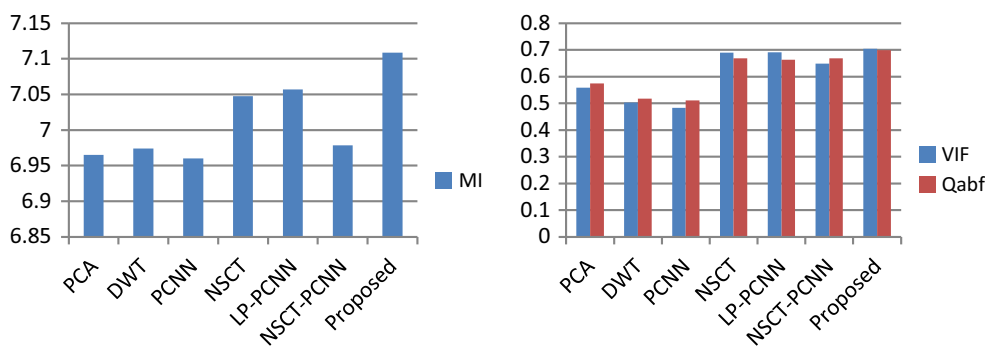


Fig. 17 Bar chart comparison of MI, VIF and $Q^{AB/F}$ values for the first example

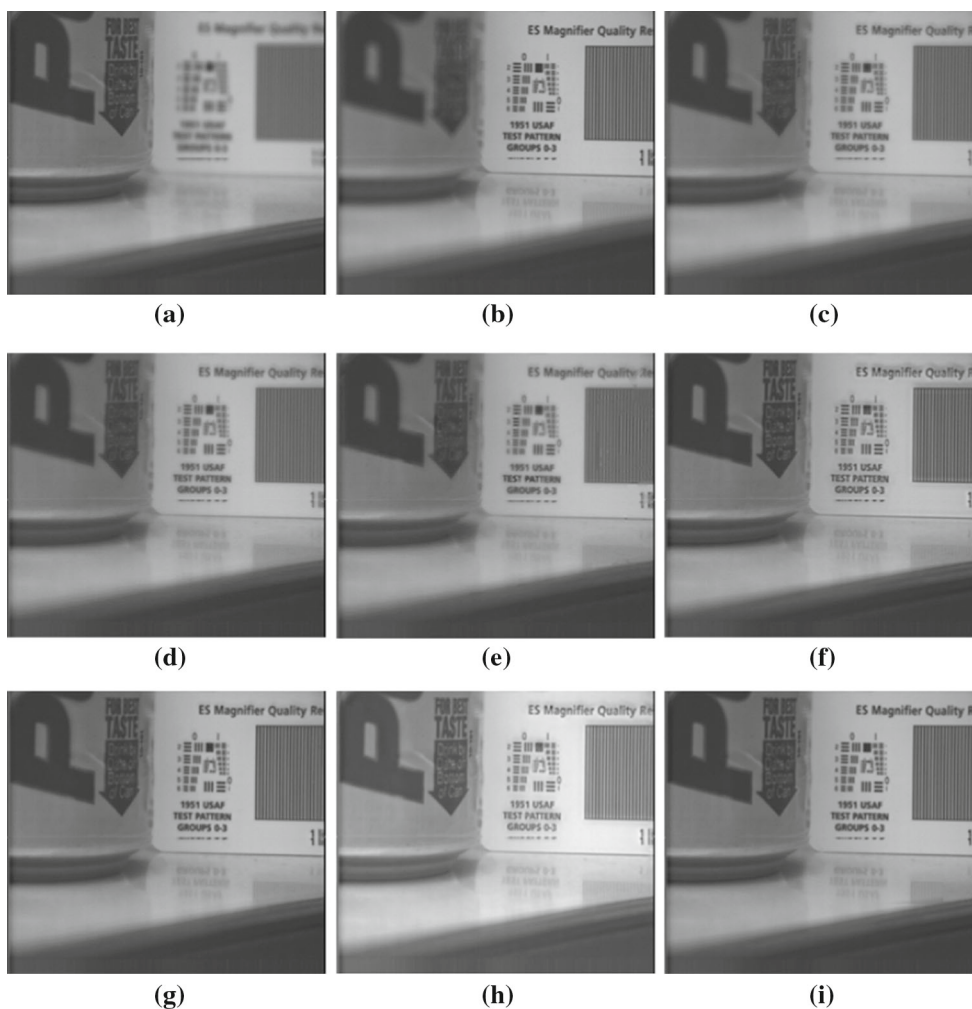


Fig. 18 Second experimental results using different methods. a, b Multi-focus images. c PCA. d DWT. e PCNN. f NSCT. g LP-PCNN. h NSCT-PCNN. i Proposed method

visual system. $Q^{AB/F}$ is used to measure the fused image how much edges information is obtained from the source images (Sheikh and Bovik 2006). Generally, the larger the values of these three indexes are, the better the quality of the fused image is.

5.1 Fusion results and discussion

The first set of results are given in Fig. 15, where Fig. 15a, b shows the multi-focus images, and (c–i) shows the fusion results using PCA, DWT, PCNN, NSCT, LP-PCNN, NSCT-

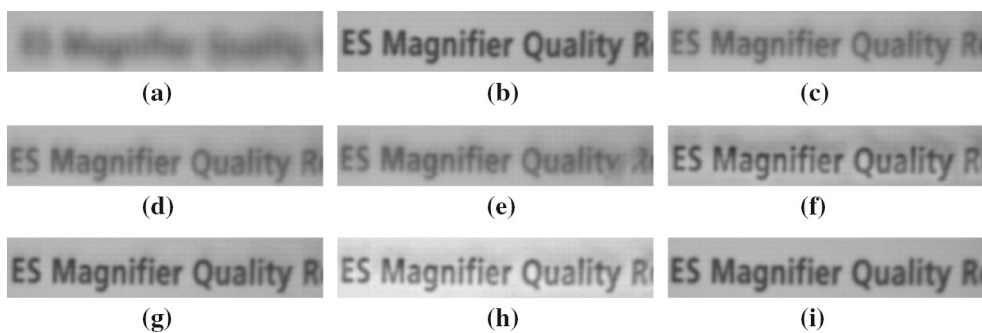


Fig. 19 Details of enlarged scale. **a** Source image A. **b** Source image B. **c** PCA. **d** DWT. **e** PCNN. **f** NSCT. **g** LP-PCNN. **h** NSCT-PCNN. **i** Proposed method

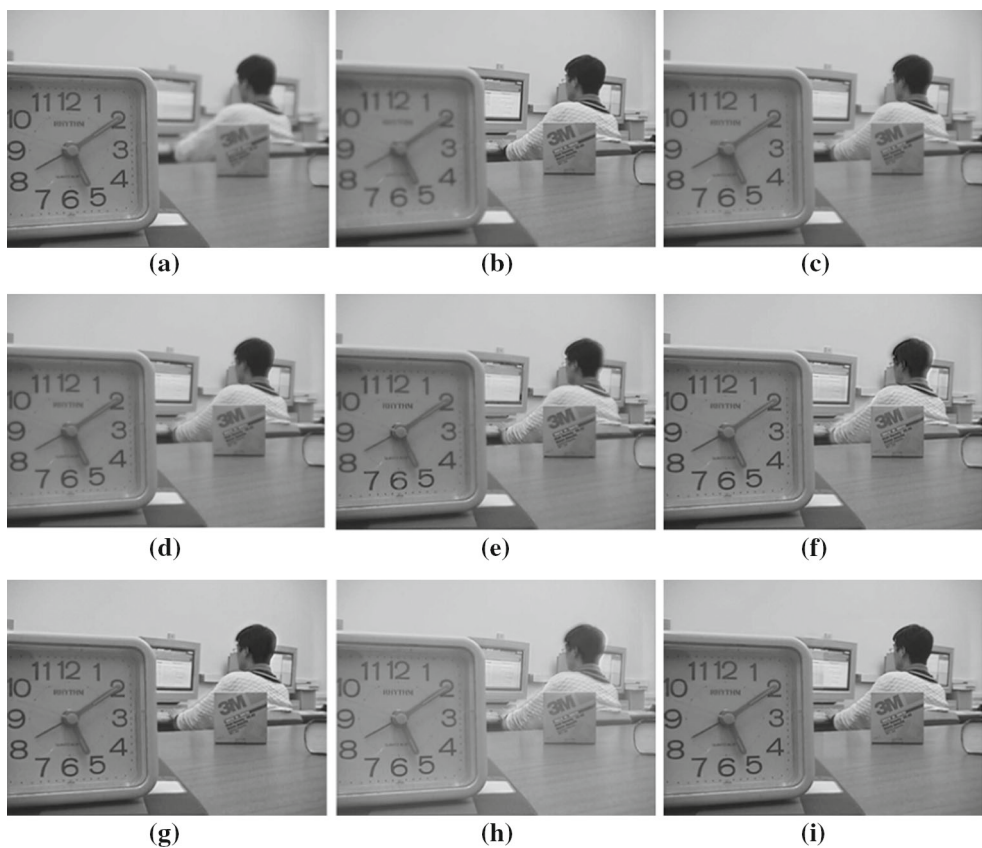


Fig. 20 Third experimental results using different methods. **a, b** Multi-focus images. **c** PCA. **d** DWT. **e** PCNN. **f** NSCT. **g** LP-PCNN. **h** NSCT-PCNN. **i** Proposed method

PCNN and proposed method. We can see that from Fig. 15a, b, only the focused region is clear, and the other regions are fuzzy, which causes the lack of the information of the multi-focus image and the poor quality of single source multi-focus image. The fused images have more information compared with the source images and contain two source images details, and the fused image’s quality and reliability were improved compared with the single source image, as shown in Fig. 15c–i. However, the quality of the fused image obtained by different methods is not exactly the same.

In order to better reflect their differences, Fig. 16 shows the detail with enlarged scale of Fig. 15.

Figure 17 shows the comparison of MI, VIF, $Q^{AB/F}$ values between different methods for first example. The x -axis denotes the different fusion methods, and the y -axis shows the values of MI, VIF, $Q^{AB/F}$. From Figs. 16 and 17, we can see that the proposed fusion method can outperform other traditional fusion methods, and the fused image of our proposed method has better visual effect; it contains more details and more information of the focus region in the source image are preserved.

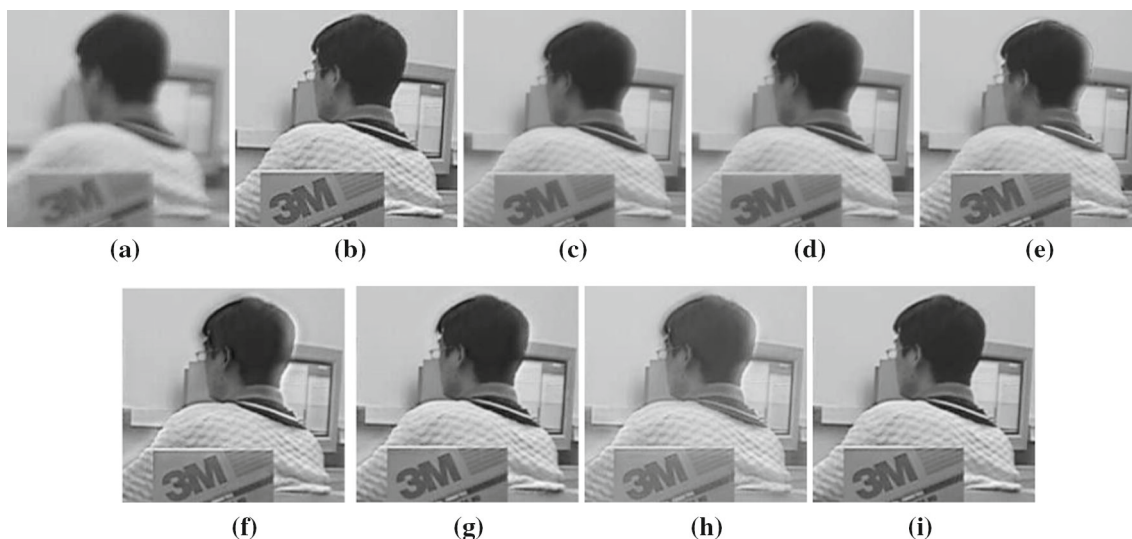


Fig. 21 Details of enlarged scale. **a** Source image A. **b** Source image B. **c** PCA. **d** DWT. **e** PCNN. **f** NSCT. **g** LP-PCNN. **h** NSCT-PCNN. **i** Proposed method

Table 4 Objective evaluation indexes for various fusion results

Evaluation indexes	Experiments	Fusion methods						
		PCA	DWT	PCNN	NSCT	LP-PCNN	NSCT-PCNN	Proposed
MI	First	6.9648	6.9741	6.9603	7.0473	7.0570	6.9784	7.1088
	Second	7.0895	7.0818	7.0763	7.1256	7.1190	7.0880	7.1253
	Third	6.9658	6.9522	6.9661	7.0260	7.0597	6.8756	7.0275
VIF	First	0.5582	0.5027	0.4833	0.6897	0.6908	0.6489	0.7043
	Second	0.6825	0.5767	0.5371	0.8041	0.8177	0.7120	0.8182
	Third	0.5184	0.4963	0.4924	0.6370	0.6302	0.5973	0.7125
$Q^{A/B/F}$	First	0.5750	0.5170	0.5106	0.6680	0.6627	0.6684	0.6994
	Second	0.6357	0.4961	0.4922	0.7419	0.7529	0.5443	0.7841
	Third	0.5652	0.5625	0.5908	0.6781	0.6891	0.5796	0.7348

Second and third experiments are tested by different multi source images, and the results are given in Figs. 18, 19, 20, 21. Figures 19 and 21 are the entails of enlarged scale images of Figs. 18 and 20, respectively.

From Figs. 19i and 21i, we can see that the fused image obtained by the proposed method contains more edges information and details, which is easily observed by the human visual system. The objective evaluation indexes are given in Table 4. Through these experiments, from Table 4, we can summarize that the proposed fusion algorithm is feasible, and it outperforms than other traditional fusion algorithms.

The best result for each evaluation is highlighted in bold

6 Conclusion

This paper presents a novel fusion scheme for the multi-focus image based on focus-region-level partition. The first step is

to put forward three kinds of evaluation functions to measure the definition of the image, then the focused region of the source multi-focus image is divided by BP neural network, and the focused regions and the non-focus regions are obtained, respectively. Next the non-focus regions will be fused in NSCT domain. The Gaussian blurred is used to produce lowpass coefficients, SF-PCNN is used to produce bandpass coefficients, and then the fused non-focus regions are produced by NSCT reconstruction. Finally, the fused image is produced by fusing the focus regions and the fused non-focus regions. Experimental results show that the proposed fusion scheme not only retains more clear pixels of two source images, but also preserves more details and edges information of the non-focus regions. By comparison between other traditional fusion methods, the proposed method can achieve superior results in visual inspection and objective evaluations.

Acknowledgements The authors thank the editors and the anonymous reviewers for their careful works and valuable suggestions for this study. This study was supported by the National Natural Science Foundation of China (No. 61463052 and No. 61365001).

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interests regarding the publication of this paper.

Human and animal rights This article does not contain any studies with human participants performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Al-Nima RRO, Abdullah MAM, Al-Kaltakchi MTS et al (2017) Finger texture biometric verification exploiting multi-scale sobel angles local binary pattern features and score-based fusion. *Digit Signal Proc* 70:178–189
- Amolins K, Zhang Y, Dare P (2007) Wavelet based image fusion techniques—an introduction, review and comparison. *ISPRS J Photogramm Remote Sens* 62(4):249–263
- Aslantas V, Toprak AN (2014) A pixel based multi-focus image fusion method. *Opt Commun* 332(4):350–358
- da Cunha AL, Zhou J, Do MN (2006) The nonsubsampling contourlet transform: theory, design, and applications. *IEEE Trans Image Process* 15(10):3089–3101
- Gupta S, Gore A, Kumar S et al (2016) Objective color image quality assessment based on Sobel magnitude. *Signal Image Video Process* 11(1):1–6
- He K, Zhou D, Zhang X, Nie R et al (2017) Infrared and visible image fusion based on target extraction in the nonsubsampling contourlet transform domain. *J Appl Remote Sens* 11(1):015011. <https://doi.org/10.1117/1.JRS.11.015011>
- Jamal S, Karim F (2012) Infrared and visible image fusion using fuzzy logic and population-based optimization. *Appl Soft Comput* 12(3):1041–1054
- Ji W, Qian Z, Xu B et al (2016) Apple tree branch segmentation from images with small gray-level difference for agricultural harvesting robot. *Optik* 127(23):11173–11182
- Jin X, Nie R, Zhou D et al (2016) A novel DNA sequence similarity calculation based on simplified pulse-coupled neural network and Huffman coding. *Phys A Stat Mech Appl* 461:325–338
- Jin X, Zhou D, Yao S et al (2016c) Remote sensing image fusion method in CIE Lab color space using nonsubsampling shearlet transform and pulse coupled neural networks. *J Appl Remote Sens* 10(2):025023
- Jin X, Nie R, Zhou D, Wang Q, He K (2016) Multifocus color image fusion based on NSST and PCNN. *J Sens* 2016:8359602. <https://doi.org/10.1155/2016/8359602>
- Jin X, Zhou D, Yao S, Nie R et al (2017) Multi-focus image fusion method using S-PCNN optimized by particle swarm optimization. *Soft Comput* 1298:1–13. <https://doi.org/10.1007/s00500-017-2694-4>
- Johnson JL, Padgett ML (1999) PCNN models and applications. *IEEE Trans Neural Netw* 17(3):480–498
- Kavitha S, Thyagarajan KK (2016) Efficient DWT-based fusion techniques using genetic algorithm for optimal parameter estimation. *Soft Comput* 2016:1–10. <https://doi.org/10.1007/s00500-015-2009-6>
- Kaya IE, Pehlivanl AA, Sekizkarde EG et al (2017) PCA based clustering for brain tumor segmentation of T1w MRI images. *Comput Methods Programs Biomed* 140:19–28
- Kumar A, Hassan MF, Raveendran P (2018) Learning based restoration of Gaussian blurred images using weighted geometric moments and cascaded digital filters. *Appl Soft Comput* 63:124–138
- Li S, Yang B (2008) Multifocus image fusion using region segmentation and spatial frequency. *Image Vis Comput* 26(7):971–979
- Li ST, Kwok JTY, Tsang IWH et al (2004) Fusing images with different focuses using support vector machines. *IEEE Trans Neural Netw* 15(6):1555–1561
- Li M, Cai W, Tan Z (2006) A region-based multi-sensor image fusion scheme using pulse-coupled neural network. *Pattern Recogn Lett* 27(16):1948–1956
- Li H, Chai Y, Li Z (2013) Multi-focus image fusion based on non-sampled contourlet transform and focused regions detection. *Optik* 124(1):40–51
- Li S, Kang X, Fang L et al (2017) Pixel-level image fusion: a survey of the state of the art. *Inf Fusion* 33:100–112
- Liu Y, Liu S, Wang Z (2015) A general framework for image fusion based on multi-scale transform and sparse representation. *Inf Fusion* 24:147–164
- Qu X, Yan J, Xiao H et al (2008) Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain. *Acta Autom Sin* 34(12):1508–1514
- Sheikh HR, Bovik AC (2006) Image information and visual quality. *IEEE Trans Image Process* 15(2):430–444
- Tian J, Chen L (2012) Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure. *Sig Process* 92(9):2137–2146
- Wang J, Jeong J (2017) Wavelet-content-adaptive BP neural network-based deinterlacing algorithm. *Soft Comput*. <https://doi.org/10.1007/s00500-017-2968-x>
- Wang Z, Ma Y, Gu J (2010) Multi-focus image fusion using PCNN. *Pattern Recogn* 43(6):2003–2016
- Xiang T, Yan L, Gao R (2015) A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain. *Infrared Phys Technol* 69:53–61
- Yang Y (2011) A novel DWT based multi-focus image fusion method. *Proc Eng* 24(1):177–181
- Zhang Q, Guo BL (2009) Multifocus image fusion using the non-sampled contourlet transform. *Sig Process* 89(7):1334–1346
- Zhao C, Guo Y, Wang Y (2015) A fast fusion scheme for infrared and visible light images in NSCT domain. *Infrared Phys Technol* 72:266–275
- Zheng S, Shi WZ, Liu J et al (2007) Multisource image fusion method using support value transform. *IEEE Trans Image Process* 16(7):1831–1839
- Zhong F, Ma Y, Li H (2014) Multifocus image fusion using focus measure of fractional differential and NSCT. *Pattern Recogn Image Anal* 24(2):234–242