

# Technique for multi-focus image fusion based on fuzzy-adaptive pulse-coupled neural network

Yong Yang<sup>1</sup> · Yue Que<sup>1</sup> · Shu-Ying Huang<sup>2</sup> · Pan Lin<sup>3</sup>

Received: 18 April 2016 / Revised: 29 July 2016 / Accepted: 31 August 2016 / Published online: 16 September 2016  
© Springer-Verlag London 2016

**Abstract** Multi-focus image fusion technique can solve the problem that not all the targets in an image are clear in case of imaging in the same scene. In this paper, a novel multi-focus image fusion technique is presented, which is developed by using the nonsubsampling contourlet transform (NSCT) and a proposed fuzzy logic based adaptive pulse-coupled neural network (PCNN) model. In our method, sum-modified Laplacian (SML) is calculated as the motivation for PCNN neurons in NSCT domain. Since the linking strength plays an important role in PCNN, we propose an adaptively fuzzy way to determine it by computing each coefficient's importance relative to the surrounding coefficients. Combined with human visual perception characteristics, the fuzzy membership value is employed to automatically achieve the degree of importance of each coefficient, which is utilized as the linking strength in PCNN model. Experimental results on simulated and real multi-focus images show that the proposed technique has a superior performance to series of exist fusion methods.

**Keywords** Image fusion · Fuzzy logic · Nonsubsampling contourlet transform · Sum modified-laplacian · Adaptive pulse-coupled neural network

## 1 Introduction

Due to the limited depth of the field of optical lenses, the sensors cannot generate images of all objects at various distances with equal clarity, which means a scene cannot simultaneously focus on all targets well, i.e., only the objects at focus plane would appear sharp, while others would appear blur [1]. An effective solution to this problem is through multi-focus image fusion technology, which can create a single 'all-in-focus' image with all the objects are in focus [2]. The fused image is more suitable for human or machine perception as well as the computer processing tasks such as segmentation, feature extraction, and target recognition.

The multi-focus image fusion algorithm is divided into two categories: spatial domain-based algorithm and transform domain-based algorithm [3]. Currently, the transform domain-based multiscale image fusion algorithm has become a mainstream class fusion method [4]. The traditional discrete wavelet transform (DWT) has been successfully used in image fusion [5]. However, DWT lacks shift-invariance due to its underlying down-sampling process [6], which means the fusion is sensitive to registration. Unser [7] proposed a shift-invariant DWT (SIDWT) to overcome the shortcomings of DWT. However, the two-dimensional wavelet is isotropic and cannot effectively describe the abrupt transitions such as line and curve singularities. In addition, the wavelet can only be limited to capture three directions of the information [8].

---

✉ Shu-Ying Huang  
shuyinghuang2010@126.com

Yong Yang  
greatyang@126.com

Yue Que  
qki120@163.com

Pan Lin  
tiger.lin9906@gmail.com

<sup>1</sup> School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China

<sup>2</sup> School of Software and Communication Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China

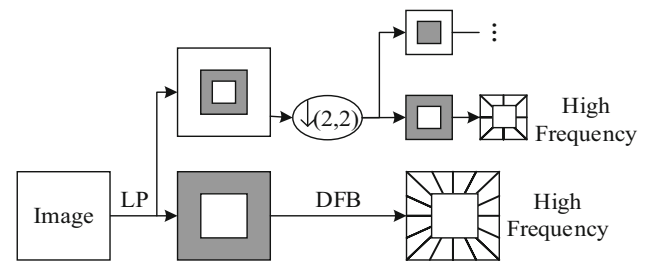
<sup>3</sup> Institute of Biomedical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

To overcome the shortcomings of the wavelet, Do and Vetterli developed a true two-dimensional image representation method, namely the contourlet transform (CT) [9]. Compared with DWT, the CT not only has multiscale parameters and localization but also multi-direction and anisotropy. As a result, the CT can represent edges and other singularities along curves more efficiently [10]. However, the up- and down-sampling process of CT results in lacking shift-invariance and having pseudo-Gibbs phenomena in the fusion. In 2006, Da Cunha et al. [11] proposed an over complete transform, namely, the nonsubsampling CT (NSCT). NSCT inherits the advantages of CT, while also possessing shift-invariance and effectively suppressing Pseudo-Gibbs phenomena [12]. Thus, the NSCT is more suitable for image fusion.

Currently, artificial neural networks have been successfully used in image processing. The pulse-coupled neural network (PCNN) was proposed in 1990 by Eckhorn to simulate the processing mechanism of cat's visual cortex, which results in a new type of neural network model [13]. Based on the PCNN bionic mechanism and pulse synchronization of neurons characteristics, PCNN plays an important role in image fusion [14]. However, the optimal parameters that set PCNN for different images have poor versatility when applied in image fusion. For most image fusion methods based on PCNN, the parameters are set to fixed values. However, visual processing for obvious characteristic areas should react more strongly than the generic areas. Hence the parameters for the PCNN neurons should reflect the importance of each pixel's characteristic of the image [15,16]. Therefore, we need a way to adaptively set the parameters of the PCNN-based fusion method.

In addition, the selection of the fusion rules is also the key problem in the transform domain-based fusion algorithm. The conventional fusion rule that used Max-Average technique for selecting high–low subbands has some disadvantages such as reduce the contrast of the fused image [4]. In this paper, an adaptive PCNN is proposed to act as the fusion rule after the source images are decomposed by NSCT. In order to better measure the pixel importance of the source images to the fused image, we employ fuzzy logic combined with the human visual perception characteristics to build an adaptive PCNN model and propose sum-modified Laplacian (SML) to motivate the PCNN neurons. The experimental results indicate that the proposed method can obtain better subjective and objective results than series of existing fusion methods.

The rest of this paper is organized as follows. The related theories are briefly introduced in Sect. 2. The proposed image fusion method is described in Sect. 3. The experimental results and analysis are depicted in Sect. 4, and the final conclusions are given in Sect. 5.



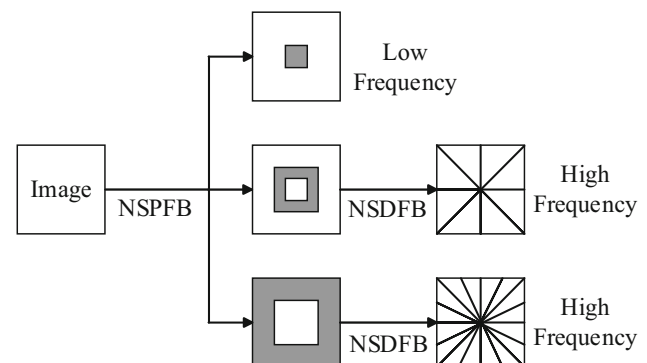
**Fig. 1** Contourlet decomposed schematic diagram

## 2 Preliminaries

### 2.1 Nonsubsampling contourlet transform

The CT process is divided into two stages. First, the Laplacian Pyramid (LP) is utilized to capture the point singularities and to decompose the original images into low-frequency and high-frequency sub-images. Then, the Directional Filter Bank (DFB) offers an efficient directional multiresolution image representation and divides the high-frequency subbands into directional subbands [9]. A contourlet decomposed schematic diagram is shown in Fig. 1. According to the sampling theorem, the pseudo-Gibbs phenomena would appear in low- and high-frequency sub-images in LP domain. Directional subbands which come from the high-frequency sub-images by DFB filtering would also appear the pseudo-Gibbs phenomena. These phenomena would weaken the directional selectivity of the CT-based method.

The NSCT is proposed based on the CT which inherits the advantages of the CT. The whole implementation of NSCT process consists of two phases, the multiscale decomposition and the directional decomposition [11]. The most difference from CT is that NSCT is composed of two shift invariant parts, i.e., nonsubsampling Pyramid filter banks (NSPFB) and nonsubsampling directional filter banks (NSDFB). An NSCT decomposed schematic diagram is shown in Fig. 2. The size of different sub-images decomposed by NSCT is identical, so it is easy to find the connection among sub-images of



**Fig. 2** Nonsubsampling contourlet decomposed schematic diagram

different images, which is beneficial to design fusion rules. Additionally, NSCT-based fusion can effectively reduce the impacts of misregistration on the results.

## 2.2 Pulse-coupled neural network

PCNN was proposed in the 1990s and was gradually replacing the traditional neural networks. PCNN does not need to learn or train and can extract the useful information from the complex background. It has become the most widely studied neural network model, but the theory should be further improved. In practical applications, a number of parameters in PCNN need to be set, which increase the difficulty and complexity of its usage. Consequently, the simplified model of PCNN has often been used. In this paper, we use an improved PCNN [17], with neurons after  $q$  iterations, whose mathematical formula is shown below.

$$\begin{cases} F_{ij}(q) = I_{ij} \\ L_{ij}(q) = e^{-\alpha L} L_{ij}(q-1) + \sum_{m,n} W_{ij,mn} Y_{mn}(q-1) \\ U_{ij}^k(q) = F_{ij}^k(q)(1 + \beta L_{ij}^k(q)) \\ \theta_{ij}^k(q) = e^{-\alpha \theta} \theta_{ij}^k(q-1) + V_{\theta} Y_{ij}^k(q-1) \\ Y_{ij}^k(q) = \begin{cases} 1, & U_{ij}^k(q) > \theta_{ij}^k(q) \\ 0, & \text{otherwise} \end{cases} \end{cases} \quad (1)$$

where  $q$  denotes the iteration times,  $ij$  represents a pixel in the image matrix position,  $I_{ij}$  is an external input stimulus signal,  $Y_{ij}$  and  $U_{ij}$  are the output neurons of the external and internal state information,  $F_{ij}$  and  $L_{ij}$  are the feed input and the input connectors,  $W_{ij,mn}$  is the synaptic gain strength and subscripts  $m$  and  $n$  are the size of the linking range in PCNN.  $\beta$  is the linking strength.  $\theta_{ij}$  is the threshold.  $\alpha_{ij}$  is the decay constants.  $V_L$  and  $V_{\theta}$  are the amplitude gain.

## 3 Proposed method

In this section, we will provide a more detailed description of the proposed fusion method based on NSCT and fuzzy-adaptive PCNN. Suppose A and B are two registered source images. After  $k$ -level NSCT decomposition, the corresponding subbands of different scales and directions can be obtained, i.e., A:  $\{L^A, H_{k,l}^A\}$  and B:  $\{L^B, H_{k,l}^B\}$ , where  $L^A$ ,  $L^B$  are the low-frequency sub-images and  $H_{k,l}^A$ ,  $H_{k,l}^B$  represent the high-frequency sub-images at level  $k$  and direction  $l$ .

### 3.1 SML motivated PCNN neurons

In most multiscale fusion algorithms based on PCNN, a single pixel is often directly used to motivate a neuron in multiscale decomposition domain. Actually, the human visual

nervous system is often highly sensitive to the edge information and the directional features [17]. Thus, pure single pixel motivation for PCNN is not sufficient. The SML can reflect the marginalized mutation and clarity information of an image [18]. Therefore, compared to spatial frequency (SF), variance, gradient energy and Laplace energy, SML is more suitable to motivate PCNN neurons. The definition of SML is defined as follows:

$$\text{SML}(i, j) = \sum_{m=-M_1}^{M_1} \sum_{n=-N_1}^{N_1} [\text{ML}(i+m, j+n)] \quad (2)$$

for a window with size  $(2M_1+1)(2N_1+1)$ , where  $\text{ML}(i, j)$  is the Modified Laplacian (ML), which is defined as:

$$\text{ML}(i, j) = |2C(i, j) - C(i - \text{step}, j) - C(i + \text{step}, j)| + |2C(i, j) - C(i, j - \text{step}) - C(i, j + \text{step})| \quad (3)$$

where step is a variable spacing between coefficients. In this paper, step always equals 1.  $C(i, j)$  denotes the pixel value of one coefficient located at  $(i, j)$ .

Hence, in the proposed fusion method, the SML is used to motivate the PCNN neurons, while human visual perception characteristics are used to determine the fuzzy-adaptive linking strength.

### 3.2 Fuzzy-adaptive linking strength determination

As can be seen from references [19,20], the linking strength  $\beta$  of PCNN has a certain relationship with the pixel characteristic of an image, and  $\beta$  should reflect the importance of a pixel relative to its surrounding pixels. The larger of  $\beta$  means the PCNN neuron is captured and fired more easily and quickly. Meanwhile, the human visual nonlinearity characteristics can determine the pixel relative to its surrounding pixels whether they are visually important or not. However, the uncertainty and subjectivity exists in the process of visual response.

Fuzzy logic can effectively address this problem and the uncertainty contribution of each source image pixel. By fuzzy if-then rules and membership functions applied for the image data set, the fuzzy logic approach can model and combine the images to enhance the contrast of the fused image [21]. Hence, we propose a novel fuzzy logic based method to determine the linking strength of PCNN, which can be adaptively determined by measuring each coefficient's importance to its surrounding coefficients.

For the low-frequency subbands, there is a nonlinear relationship between the contrast sensitivity threshold and the brightness of the background by considering human visual perception characteristics. A concept of visibility that can

effectively measure the image uniformity is introduced in [22]. Based on the visibility concept, a local visibility is defined, which is used to measure the clarity of the low-frequency subbands. The local visibility is defined as:

$$L_{V1}^x(i, j) = \begin{cases} \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N \frac{|L^x(i, j) - \tau^x(i, j)|}{\tau^x(i, j)^{1+\alpha}} & \tau^x(i, j) \neq 0 \\ L^x(i, j) & \text{otherwise} \end{cases} \quad (4)$$

where  $(M \times N)$  is the window size,  $\tau^x(i, j)$  denotes the mean low coefficient located at  $(i, j)$  of the image  $x$ , and  $\alpha$  is a visual constant ranging from 0.6 to 0.7. If the local visibility of a coefficient is larger, then the coefficient is more important in the image. The fuzzy membership values  $\mu_L$  corresponding to the low-frequency subband coefficient's local visibility is thus can be calculated as follows:

$$\mu_L(L_{V1}^x(i, j)) = \frac{1}{1 + e^{-(L_{V1}^x(i, j) - a_1)}} \quad (5)$$

where  $a_1 = \text{average}(L_{V1}^x(i, j))$ .

For the high-frequency subbands, because human visualization is sensitive to the changes in local contrast of the image. Meanwhile, considering the advantages of SML in distinguishing clearness and blurriness of the image block and the human visual perception characteristics, the local visual features contrast is defined as:

$$L_{VC}^{x,k,l}(i, j) = \begin{cases} \left(\frac{1}{\vartheta_k^x(i, j)}\right)^\alpha \frac{\text{SML}_{k,l}^x(i, j)}{\vartheta_k^x(i, j)} & \vartheta_k^x(i, j) \neq 0 \\ \text{SML}_{k,l}^x(i, j) & \text{otherwise} \end{cases} \quad (6)$$

where,  $\vartheta_k^x(i, j)$  is the average low coefficient located at  $(i, j)$  of the image  $x$  at the scale  $k$ , and  $\text{SML}_{k,l}^x(i, j)$  denotes the

SML located at  $(i, j)$  of the image  $x$  at the scale  $k$  and direction  $l$ ,  $x \in \{A, B\}$ . If the local visual feature contrast of a coefficient is large, then the coefficient has more importance in the image. Hence, the fuzzy membership values  $\mu_H$  corresponding to the high-frequency subband coefficient's local visual features contrast can be defined as:

$$\mu_H(L_{VC}^{x,k,l}(i, j)) = \frac{1}{1 + e^{-(L_{VC}^{x,k,l}(i, j) - a_2)}} \quad (7)$$

where  $a_2 = \text{average}(L_{VC}^{x,k,l}(i, j))$ .

Therefore, the linking strength  $\beta$  of low-frequency subbands and high-frequency subbands are presented as  $\beta_L^x(i, j) = \mu_L(L_{V1}^x(i, j))$  and  $\beta_H^{x,k,l}(i, j) = \mu_H(L_{VC}^{x,k,l}(i, j))$ , respectively, representing their importance in the corresponding source image.

### 3.3 Algorithm

The block diagram of the proposed fusion scheme is shown in Fig. 3. This method can be easily extended to more than two images. The fusion process consists of the following steps:

*Step 1* Decompose the  $n$ th source images  $I_n$  by using NSCT to obtain one low-frequency subband and a series of high-frequency subbands at each level  $k$  and direction  $l$ .

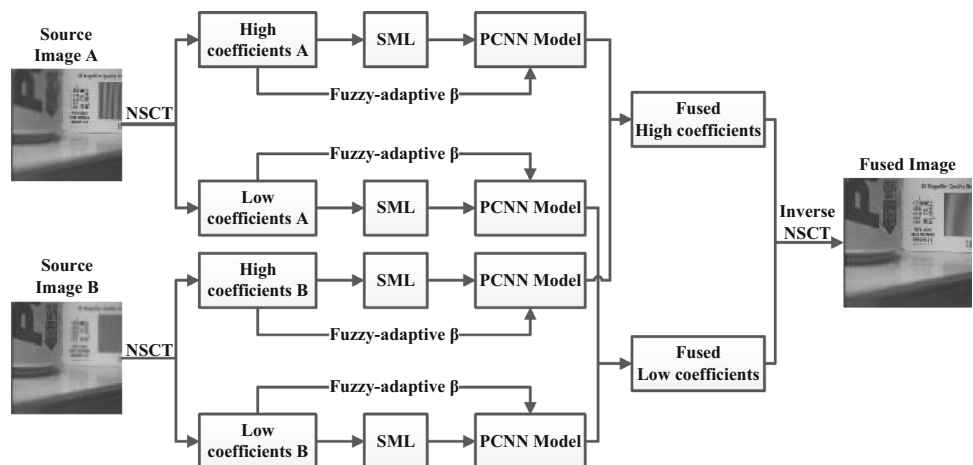
*Step 2* Compute the SML of the subbands as the input of PCNN using (2).

*Step 3* Compute the linking strength  $\beta_L^n$  and  $\beta_H^{n,k,l}$  of the low-frequency subbands and the high-frequency subbands respectively, as described in Sect. 3.2.

*Step 4* Input the SML to motivate the PCNN with the linking strength of step 3 to generate pulses of neurons using (1), and compute the firing times as follows:

$$T_{i,j}^{n,k,l} = T_{i,j}^{n,k,l}(q - 1) + Y_{i,j}^{n,k,l}(q) \quad (8)$$

**Fig. 3** The framework of the proposed fusion algorithm



**Step 5** When the number of iterations reaches  $q$ , the maximum firing time rules for fusion subband coefficients is defined as follows:

$$D_n^{k,l}(i, j) = \begin{cases} 1, & \text{if } T_{i,j}^{n,k,l} = \max(T_{i,j}^{1,k,l}, T_{i,j}^{2,k,l}, \dots, T_{i,j}^{n,k,l}) \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

$$P_F^{k,l}(i, j) = \begin{cases} P_n^{k,l}(i, j), & \text{if } D_n^{k,l}(i, j) = 1 \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

where,  $P_F^{k,l}$  and  $P_n^{k,l}$  denote the coefficients of the fused image and  $n$ th source image, respectively.

**Step 6** Use the inverse NSCT on the fused coefficients to obtain the final fused image.

## 4 The experimental results and analysis

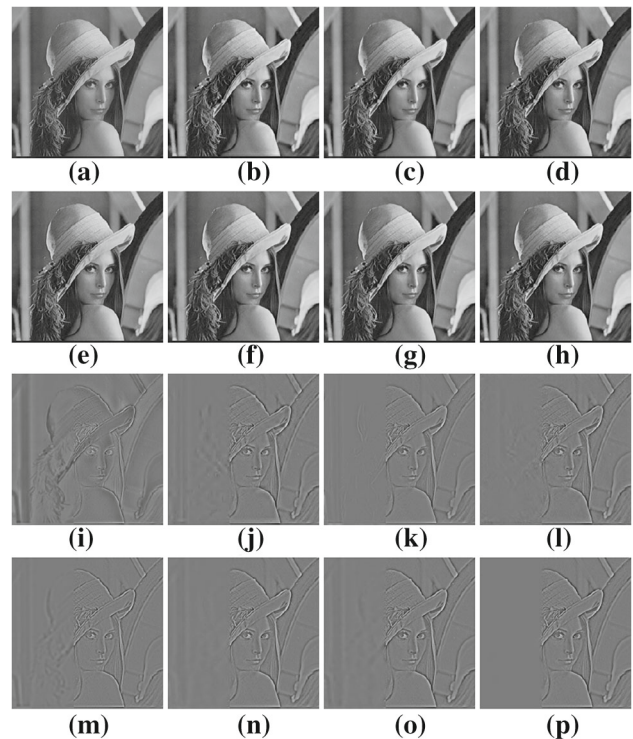
In this section, we compare the performance of our technique with the existing image fusion methods based on Gradient Pyramid (GP), DWT, SIDWT, CT, NSCT and NSCT-PCNN. In addition, this method was also compared with the method of literature [17], which uses SF to motivate the PCNN neurons (SF\_PCNN). All the methods are the three layers decomposition. All the NSCT-based methods have the same multi-direction decomposition levels are 1, 2 and 8. The parameters of PCNN are set as  $W = [0.707 \ 1 \ 0.707; \ 1 \ 0 \ 1; \ 0.707 \ 1 \ 0.707]$ ,  $\alpha_L = 0.01$ ,  $\alpha_\theta = 20$ ,  $m \times n = 3 \times 3$ , and  $q=200$ . The size of the window for computing SML,  $L_{VI}$  and  $L_{VC}$  was set as  $3 \times 3$ .

### 4.1 Experiments on simulated multi-focus image fusion

The first experiment is conducted using an extensive set of artificially generated multi-focus images. Figure 4a shows the referenced Lena image and two blurred artificial images obtained by Gaussian filtering as can be seen from Fig. 4b, c, respectively. The fusion experiments are then carried out by using the eight fusion methods mentioned above. The fusion



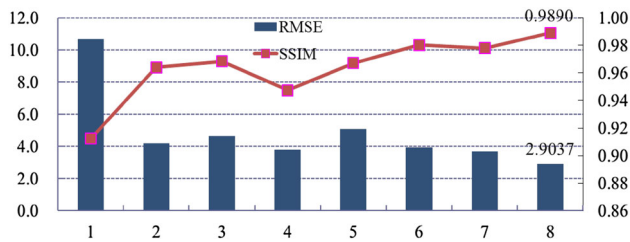
**Fig. 4** Referenced and blurred Lena images. **a** Referenced image; **b** Blurred on right; **c** Blurred on left



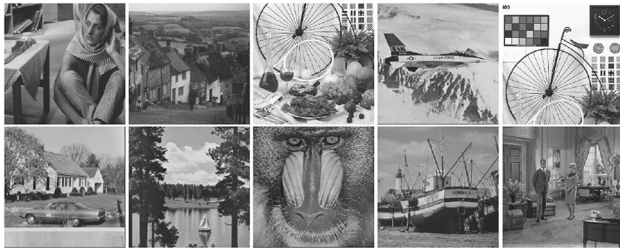
**Fig. 5** The experimental results of fusing simulated multi-focus image: **a–h** the fusion results of GP, DWT, SIDWT, CT, NSCT, NSCT+PCNN, NSCT+SF\_PCNN and the proposed method, respectively; **i–p** the residual images between the eight fusion results and the source image Fig. 4b

results and their corresponding residual images of different methods are depicted in Fig. 5. In residual images, the lower residue of the features means the more detail information in the focus region has been transferred to the fused image, i.e., the corresponding fusion method is better. As seen from Fig. 5, we can observe that the residual image between the result of the proposed method and Fig. 4b is almost zero on the left region. In other words, the proposed method has extracted the most information from the clear/focus region (the left side of the image) of Fig. 4b. From the fusion and residual results, we can also easily find that the GP method has the worst performance in the eight methods.

In order to objectively evaluate the performance of those algorithms, two evaluation criteria including the root mean square error (RMSE) [4] and the structural similarity measure (SSIM) [23] are introduced and employed in the paper. RMSE and SSIM are the metrics that measure the error and similarity between a referenced image and a fused image, respectively. The smaller the RMSE, the better the performance of the fusion algorithm. However, the larger the SSIM, the better the performance of the fusion algorithm. Figure 6 shows the values of RMSE and SSIM obtained from the fusion results of Fig. 5a–h and the referenced image Fig. 4a. As seen from Fig. 6, the RMSE value of the proposed method



**Fig. 6** Comparison on RMSE and SSIM of different fusion methods on simulated multi-focus images (methods 1–8 are GP, DWT, SIDWT, CT, NSCT, NSCT+PCNN, NSCT+SF\_PCNN and the proposed method, respectively)



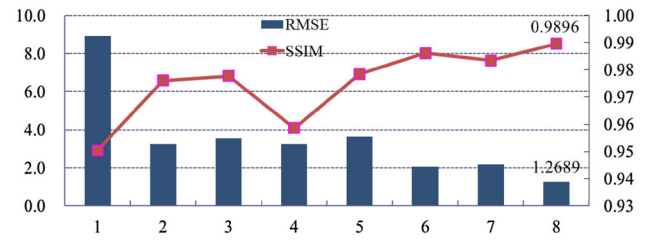
**Fig. 7** Standard test images

**Table 1** Comparison on objective criteria of different methods in simulated multi-focus images

Method	MI	$Q^{AB/F}$
GP	5.5978	0.6990
DWT	6.6028	0.7126
SIDWT	7.0733	0.7397
CT	6.6583	0.7041
NSCT	6.7839	0.7302
NSCT-PCNN	7.4501	0.7412
NSCT-SF_PCNN	8.0737	0.7422
Proposed	8.6791	0.7531

is the minimum, and the SSIM value of the proposed method is the maximum, which means our method can obtain the fused result that it is closer to the standard image compared to other seven fusion methods, i.e., the proposed method can achieve the best fusion result.

In addition to above evaluation, Table 1 gives two other frequently used metrics, mutual information (MI) [24] and edge-based similarity measure ( $Q^{AB/F}$ ) [25] to assess the results obtained by those fusion methods. MI indicates how much of the information the fused image conveys about the source images. The greater the value of MI, the better the fusion effect.  $Q^{AB/F}$  shows the similarity between the edges that are transferred from the input images to the fused image. The value of  $Q^{AB/F}$  is more close to one, which indicates a better fused image. From Table 1, we can observe that the values of MI and  $Q^{AB/F}$  of the proposed method are larger than other seven fusion methods.



**Fig. 8** Comparison on RMSE and SSIM of different fusion methods on standard test images (methods 1–8 are GP, DWT, SIDWT, CT, NSCT, NSCT+PCNN, NSCT+SF\_PCNN and the proposed method)

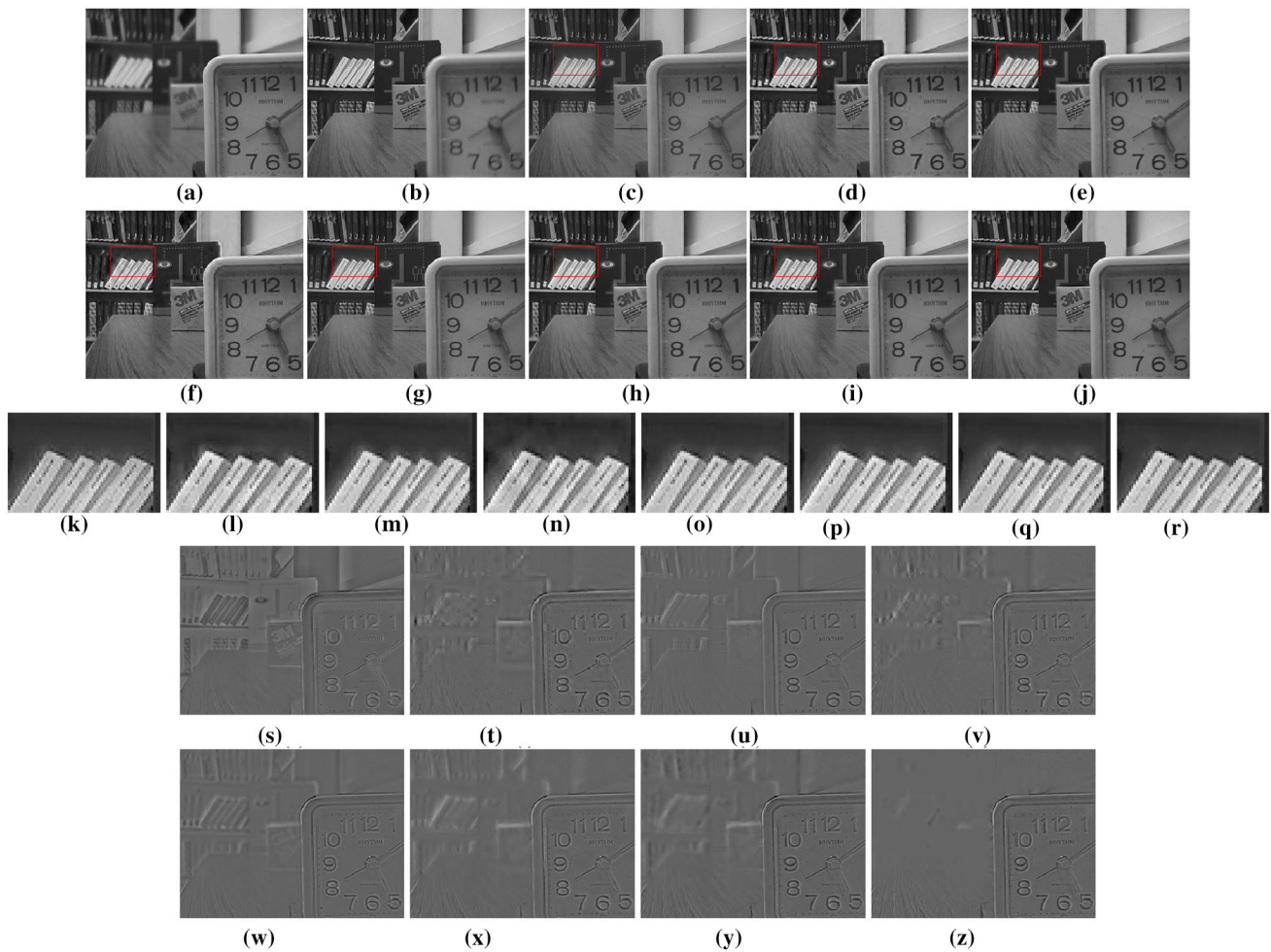
**Table 2** Comparison on objective criteria of different methods in standard test images

Method	MI	$Q^{AB/F}$
GP	4.6321	0.6970
DWT	5.3482	0.7191
SIDWT	5.7913	0.7322
CT	5.3665	0.7049
NSCT	5.4124	0.7216
NSCT-PCNN	6.0773	0.7399
NSCT-SF_PCNN	6.4303	0.7403
Proposed	7.0787	0.7486

To further verify the effectiveness of the proposed method, tests were then realized on ten groups of synthetic multi-focus images, which are created by convolving Gaussian blurring on ten standard images as shown in Fig. 7. Then, the ten groups of simulated images are fused by the previously listed eight methods. Figure 8 shows the average values of RMSE and SSIM of the different fusion methods on standard test images. It is easy to find that in ten sets of fusion results, the SSIM and RMSE values of the proposed method results outperform other seven methods. In addition, Table 2 shows the average values of MI and  $Q^{AB/F}$  of different fusion methods. Similarly, we can also observe that the values of MI and  $Q^{AB/F}$  of the proposed method are all the largest and optimal in the eight fusion methods. Therefore, the four optimal criteria values presented here can indicate that the fused image obtained by the proposed method can retain much more focused information and achieve higher similarity and correlation to the source images.

## 4.2 Experiments on real multi-focus image fusion

The second part of the experiment is conducted on three sets of generally used real multi-focus images, including *pepsi*, *lab*, and *disk*, which are used to further verify the validity of the proposed method. Every set includes two source images that focus on the right and focus on the left, respectively. An example of experimental data, fusion results and their corresponding residual images between the fused result and



**Fig. 9** Source images “disk” and the fusion results. **a** The first source image with focus on the right. **b** The second source image with focus on the left. **c** GP result. **d** DWT result. **e** SIDWT result. **f** CT result. **g** NSCT result. **h** NSCT-PCNN result. **i** NSCT-SF\_PCNN result. **j** Result

of the proposed algorithm. **k–r** Zoom in regions of the eight fused results marked by *red rectangles*. **s–z** The residual images between the eight fusion results and source image **b**

**Table 3** Comparison on objective criteria of different methods in real multi-focus images

Images	Evaluation	GP	DWT	SIDWT	CT	NSCT	NSCT-PCNN	NSCT-SF_PCNN	Proposed method
pepsi	MI	6.5207	6.6335	7.0450	6.5573	7.1904	7.2165	7.8127	8.2353
	$Q^{AB/F}$	0.6896	0.6829	0.7166	0.6681	0.7082	0.7153	0.7330	0.7371
lab	MI	6.3615	6.7254	7.1326	6.4033	7.0403	7.1793	7.7654	7.8414
	$Q^{AB/F}$	0.7144	0.6972	0.7276	0.6816	0.7103	0.7241	0.7302	0.7458
disk	MI	5.579	5.7166	6.0652	5.4463	6.0822	5.9719	6.3985	7.0477
	$Q^{AB/F}$	0.6562	0.6464	0.6879	0.6257	0.6717	0.6743	0.6784	0.6972

the source image focused on left (foreground) are show in Fig. 9, respectively. From the residual images, it is easy to find that the proposed method has extracted much more information than other seven methods in the focused left areas, and almost all of the useful information in the focus area of the source images has been transferred to the fused image. Considering the example of the set of ‘disk’ images in Fig.

9, we can see that GP-based fusion result has lower contrast, DWT- and CT-based fusion results have a significant Pseudo-Gibbs phenomenon as marked in the red rectangle. However, it is difficult to visually discriminate the difference among the proposed method and the SIDWT and other NSCT-based fusion results because all these methods have shift-invariance. But from the residual images in Fig. 9, we

can easily see that residue between the result of the proposed method and the source image is smallest as can be seen from the foreground of Fig. 9z.

In order to better evaluate these fusion methods on three groups of real multi-focus images, quantitative assessments of the performance of the eight methods are needed. However, since in real applications no reference images can be provided, only MI and  $Q^{AB/F}$  are then used to evaluate the performance. The quantitative results of the fusion results are given in Table 3. From Table 3, it is easy to see, just as Table 1 and Table 2, that the performance of the proposed method is the best in the eight methods because it has the biggest values of both MI and  $Q^{AB/F}$ , which means the fused image generated by proposed method contains more information from the source image, and the details of the two source images are reflected more accurately.

## 5 Conclusions

In this paper, we propose a novel multi-focus image fusion technique based on the NSCT and a fuzzy-adaptive PCNN. After the sources images to be fused are decomposed by NSCT, SML is utilized as the motivation for PCNN neurons, and the linking strength of PCNN is automatically determined by calculating the fuzzy membership value of each pixel. The fuzzy logic is developed by considering the human visual perception characteristics to different frequency subbands; the proposed fuzzy-adaptive PCNN model is thus formed. Experimental results indicate that the proposed method outperforms several popular widely used fusion methods, which means our method is effective and promising. Although the NSCT and PCNN algorithms are time-consuming, we believe that with a more efficient implementation approach such as C++, the running time of the method can be significantly reduced in the future work.

**Acknowledgements** This work is supported by the National Natural Science Foundation of China (Nos. 61262034, 61462031, 61473221 and 61662026); Natural Science Foundation of Jiangxi Province (Nos. 20151BAB207033 and 20161ACB21015); Project of the Education Department of Jiangxi Province (Nos. KJLD14031, GJJ150461 and GJJ150438).

## References

1. Nejati, M., Samavi, S., Shirani, S.: Multi-focus image fusion using dictionary-based sparse representation. *Inf. Fusion* **25**, 72–84 (2014)
2. Kumar, B.K.S.: Image fusion based on pixel significance using cross bilateral filter. *Signal, Image Video Process.* **9**(5), 1193–1204 (2015)
3. Liang, J., Yang, H., Ding, L., Zeng, X.: Image fusion using higher order singular value decomposition. *IEEE Trans. Image Process.* **21**(5), 2898–2909 (2012)
4. Yang, Y., Tong, S., Huang, S., Lin, P.: Multifocus image fusion based on NSCT and focused area detection. *IEEE Sens. J.* **15**(5), 2824–2838 (2015)
5. Pajares, G., Cruz, J.M.D.L.: A wavelet-based image fusion tutorial. *Pattern Recognit.* **37**(9), 1855–1872 (2004)
6. Yin, M., Liu, W., Zhao, X., Yin, Y., Guo, Y.: A novel image fusion algorithm based on nonsubsampling shearlet transform. *Optik-Int. J. Light Electron Opt.* **125**(10), 2274–2282 (2014)
7. Unser, M.: Texture classification and segmentation using wavelet frames. *IEEE Trans. Image Process.* **4**(11), 1549–1560 (1995)
8. Bhatnagar, G., Wu, Q.M.J., Liu, Z.: Directive contrast based multimodal medical image fusion in NSCT domain. *IEEE Trans. Multimed.* **15**(5), 1014–1024 (2013)
9. Do, M.N., Martin, V.: The contourlet transform: an efficient directional multiresolution image representation. *IEEE Trans. Image Process.* **14**(12), 2091–2106 (2005)
10. Yang, L., Guo, B.L., Ni, W.: Multimodality medical image fusion based on multiscale geometric analysis of contourlet transform. *Neurocomputing* **72**(1–3), 203–211 (2008)
11. Da Cunha, A.L., Zhou, J., Do, M.N.: The nonsubsampling contourlet transform: theory, design, and applications. *IEEE Trans. Image Process.* **15**(10), 3089–3101 (2006)
12. Kong, W., Lei, Y., Ni, X.: Fusion technique for grey-scale visible light and infrared images based on non-subsampling contourlet transform and intensity-hue-saturation transform. *IET Signal Process.* **5**(1), 75–80 (2011)
13. Eckhorn, R., Reitboeck, H., Arndt, M., Dicke, P.: Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex. *Neural Comput.* **2**(3), 293–307 (1990)
14. Xiang, T., Yan, L., Gao, R.: 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 (2015)
15. Kinsler, J.M.: Foveation by a pulse-coupled neural network. *IEEE Trans. Neural Netw.* **10**(3), 621–625 (1999)
16. Buchsbaum, G.: An analytical derivation of visual nonlinearity. *IEEE Trans. Biomed. Eng.* **27**(5), 237–242 (1980)
17. Qu, X.B., Yan, J.W., Xiao, H.Z., Zhu, Z.Q.: Image fusion algorithm based on spatial frequency-motivated pulse coupled neural networks in nonsubsampling contourlet transform domain. *Acta Autom. Sin.* **34**(12), 1508–1514 (2008)
18. Huang, W., Jing, Z.: Evaluation of focus measures in multi-focus image fusion. *Pattern Recognit. Lett.* **28**(4), 493–500 (2007)
19. Miao, Q.G.: A novel image fusion algorithm based on local contrast and adaptive PCNN. *Chin. J. Comput.* **31**(5), 875–880 (2008)
20. Wang, Z.H., Zhao, B.J., Shen, T.Z.: Image fusion based on MMPN and adjustable linking strength. *Acta Electron. Sin.* **38**(5), 1162–1166 (2010)
21. Sharma, K.K., Sharma, M.: Image fusion based on image decomposition using self-fractional Fourier functions. *Signal, Image Video Process.* **8**(7), 1335–1344 (2014)
22. Bhatnagar, G., Wu, Q.M.J., Liu, Z.: Human visual system inspired multi-modal medical image fusion framework. *Expert Syst. Appl.* **40**(5), 1708–1720 (2013)
23. Zhou, W., Alan Conrad, B., Hamid Rahim, S., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
24. Qu, G., Zhang, D., Yan, P.: Information measure for performance of image fusion. *Electron. Lett.* **38**(7), 313–315 (2002)
25. Xydeas, C.S., Petrovic, V.: Objective image fusion performance measure. *Electron. Lett.* **36**(4), 308–309 (2000)