Infrared Physics & Technology 69 (2015) 53–61

Contents lists available at ScienceDirect

Infrared Physics & Technology

journal homepage: www.elsevier.com/locate/infrared

A fusion algorithm for infrared and visible images based on adaptive dual-channel unit-linking PCNN in NSCT domain

Tianzhu Xiang [a], Li Yan [a,*], Rongrong Gao [b]

[a] School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China
[b] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

HIGHLIGHTS

- The nonsubsampled contourlet transform is used in this image fusion method.
- The adaptive dual-channel unit-linking PCNN is devised and given in detail.
- The different features of image are employed to motivate the PCNN neuron.
- The proposed algorithm is more effective compared with other current ones.

ARTICLE INFO

Article history:
Received 16 October 2014
Available online 14 January 2015

Keywords:
Image fusion
Nonsubsampled contourlet transform
Dual-channel PCNN
Unit-linking PCNN
Infrared image

ABSTRACT

In this paper, a novel fusion algorithm based on the adaptive dual-channel unit-linking pulse coupled neural network (PCNN) for infrared and visible images fusion in nonsubsampled contourlet transform (NSCT) domain is proposed. The flexible multi-resolution and directional expansion for images of NSCT are associated with global coupling and pulse synchronization characteristic of dual-PCNN. Compared with other dual-PCNN models, the proposed model possesses fewer parameters and is not difficult to implement adaptive, which is more suitable for image fusion. Firstly, the source images were multi-scale and multi-directional decomposed by NSCT. Then, to make dual-channel PCNN adaptive, the average gradient of each pixel was presented as the linking strength, and the time matrix was presented to determine the iteration number adaptively. In this fusion scheme, a novel sum modified-Laplacian of low-frequency subband and a modified spatial frequency of high-frequency subband were input to motivate the adaptive dual-channel unit-linking PCNN, respectively. Experimental results demonstrate that the proposed algorithm can significantly improve image fusion performance, accomplish notable target information and high contrast, simultaneously preserve rich details information, and excel other typical current methods in both objective evaluation criteria and visual effect.

1. Introduction

Image fusion is to synthesize information from the same scenes' multiple images which come from different kinds of image sensors or the same one that functions in diverse modes, to obtain a composite image which is more suitable for further image processing tasks. Compared with the source images, fused images which integrate complementary and redundant information from multiple images contain a "better" description of the scene than any of the single source images [1,2]. Up to now, image fusion plays an important part in many fields such as computer vision, medical image, military, remote sensing and so on. The fusion of visible and infrared images is one of the most useful cases.

Visible images obtained by spectral reflection have high resolution, good image quality and rich details of background, while they cannot detect the objects in hidden or in low light and nighttime conditions; infrared images acquired through infrared thermal radiation are capable of showing the concealed objects of interest in some poor environment, while they are insensitive to the changes of the brightness in the scene which results in poor image quality and lacking details about the scene. The fused image, constructed by the combination of features, can assimilate the advantages of the visible and infrared images, improve detection and recognizable localization of a target in the infrared image with respect to its background in the visible image, and make it suitable for subsequent processing tasks.

* Corresponding author. Tel.: +86 13907152721.
E-mail address: lyan@sgg.whu.edu.cn (L. Yan).

http://dx.doi.org/10.1016/j.infrared.2015.01.002
1350-4495/

During the last decade, the multi-scale transform based methods and pulse couple neuron network (PCNN) based methods have attracted more and more researchers' attention and are becoming the focus of study. Among the tools of multi-scale geometrical analysis, nonsubsampled contourlet transform [3] (NSCT), proposed by Da Cunha et al., is referred as the most "optimal" representation of two dimension (2D) image, and can take full advantage of the geometric regularity of image intrinsic structures and obtain the asymptotic optimal representation. Compared with the contourlet transform [4] (CT), NSCT inherits the perfect properties of CT and wavelet, such as the characteristics of time–frequency localization, multidirection, and anisotropy, and meanwhile possesses the shift-invariance which contourlet transform lacks of, leading to better frequency selectivity and regularity than contourlet transform and get rid of pseudo-Gibbs phenomena along the edges to some extent. NSCT is a fully shift-invariant, multi-scale and multi-direction expansion that has a fast implementation and has been used in image denosing and enhancement. When the NSCT is introduced to the image fusion field, more information for image fusion can be obtained and the impacts of mis-registration on the fused results can also be reduced effectively [5]. So, the NSCT is more suitable for image fusion.

Pulse coupled neural network (PCNN) is a visual cortex-inspired neural network developed by Eckhorn et al. in 1990 and founded on the experimental observations of synchronous pulse bursts in cat and monkey visual cortex [6,7]. Because of its characteristics of global coupling and pulse synchronization of neurons, PCNN has attracted sufficient attention and been extensively used in the field of image fusion. Although these fusion methods achieve excellent results, they are complex and inefficient because they cannot deal with different kinds of images. Most PCNN models have only one stimulus and one PCNN cannot deal with the whole process of image fusion, so image fusion generally requires multiple PCNNs. As a result, the standard PCNN structurally limits its application in image fusion to some extent.

To make PCNN more appropriate for image fusion, many researchers improved the original PCNN model to the dual-channel PCNN model, which is greatly suitable for image fusion and can solve the problem of complication and inefficiency of PCNN-based methods very well. Wang et al. [8–10] pioneered a novel dual-channel PCNN, which overcomes the defects mentioned above, and successfully applied in medical and multi-focus image fusion. In this dual-PCNN model, all source images are input into the same PCNN at the same time, and these data is weighted and mixed in the information fusion pool according to the weighting coefficients, and the output of dual-PCNN is the fused image, which makes the dual-PCNN fusion model simple and fast. Literature [11–14] improved and developed the dual-channel PCNN model.

These methods accomplish excellent performance, but Wang's model does not consider the linking channel for the input image which may result in false fired pulses. This model cannot put out the fusion image directly, and it must employ the time matrix and linearly transformation, which makes it inconvenient enough. Therefore, Chai et al. [15] proposed another dual-channel PCNN model. Unlike Wang's model, two source images are input into the same PCNN at the same time, and select the big one in the information fusion pool for the subsequent processing. What is more, the algorithm introduces the lifting stationary wavelet transform (LSWT) and the image features to fuse images. But the limitations of LSWT lead to that the 2D and higher-dimensional information of images could not be effectively depicted. So El-taweel G. S. [16] applied nonsubsampled contourlet transform to replace LSWT to fuse multi-focus images, medical images and infrared and visible images, and achieve good performance. However, by analyzing the Chai's dual-channel PCNN model, it is found that the Chai's model also has too many uncertain parameters, which are difficult to set and lack automation. Li et al. [17] utilized particle swarm optimization (PSO) to set the parameters adaptive, but it requires large number of iterations. And there hardly any work based on dual-channel PCNN algorithm for the fusion of visible and infrared images.

To make up for these defects and obtain better fusion performance, an improved simplified adaptive dual-channel PCNN in NSCT domain for infrared and visible images fusion is proposed in this paper based on Chai's model. The proposed model introduces the unit-linking PCNN, which decreases the number of parameters a lot and achieves satisfactory performance of image processing. For the sake of implementing the dual-channel PCNN model adaptive, the model utilizes the average gradient of each pixel in images as the linking strength, and the time matrix to determine the iteration times adaptively. After decomposing the source images by NSCT, we use the new simplified dual-channel unit-linking PCNN, which not only inherits the properties of the global coupling and pulse synchronization of the PCNN neurons but also overcomes the above-mentioned problems and the limits of the original model in the image fusion, to select the coefficients of the fused image. The concrete selection principles of the low-frequency sub-band images and the high-frequency directional sub-band images are discussed in detail in the paper, respectively. Experimental results show the proposed method does well in the fusion of infrared and visible images and can preserve not only the spectral information of the visible image, but also the thermal target information of the infrared image, and the fused result has high contrast, remarkable target and rich background information.

The remaining sections of this paper are organized as follows. Section 2 reviews the theory of the NSCT in brief. Section 3 presents the unit-linking PCNN briefly, and then the novel simplified dual-channel unit-linking PCNN is introduced in detail. Section 4 describes the image fusion algorithm using NSCT and the new dual-channel PCNN, and the exhaustive fusion steps are described. Experimental results and discussion are given in Section 5. Some conclusions are summarized in Section 6.

2. Nonsubsampled contourlet transform

Nonsubsampled contourlet transform is a multi-resolution, multi-directional expansion for images derived from contourlet transform. As a new two-dimension image analysis tool, NSCT cannot only deal with the problem that higher-dimensional singularities cannot be effectively represented by the wavelets, but can also overcome the drawbacks of contourlet transform that does not own the property of shift-invariance.

Differently from contourlet transform, NSCT employs nonsubsampled pyramid (NSP) decomposition and nonsubsampled directional filter banks (NSDFB). The nonsubsampled pyramid structure is realized by using nonsubsampled 2D filter banks, which can reach the subband decomposition structure similar to Laplacian pyramid in CT. The NSDFB is achieved by switching off the down-samplers/up-samplers in the DFB tree structure and up-sampling the filters accordingly instead. This algorithm is conceptually similar to the nonsubsampled wavelet transform computed with the àtrous algorithm [5]. As a result, NSCT gains shift-invariant except for the properties of multi-scale, multi-direction and localization, which can effectively represent the image information of edge and contour and get rid of the frequency aliasing of the CT so as to overcome the pseudo-Gibbs phenomena around singularities in image fusion, which leads to better frequency selectivity and regularity than contourlet transform. Because of the shift-invariance, NSCT can greatly reduce the effects of mis-registration on the fused results, and the size of different subbands decomposed by NSCT is the same to the original images,

which make it easy to find the correspondence between different subbands and is beneficial for designing fusion rules.

The NSCT construction can be divided into two parts like CT: (1) a nonsubsampled pyramid structure which ensures the multi-scale property and (2) a nonsubsampled directional filter bank structure that provides directionality. The image is first decomposed by NSP, and one low-frequency image and one high-frequency image can be produced at each NSP decomposition level. Subsequent NSP decomposition is carried out to decompose the low-frequency component iteratively to capture the singularities in the image. If the level of decomposition is k, so NSP can generate $k+1$ subbands, which consist of one low-pass subband and k high-pass subbands, whose sizes are as same as the source images. And then the high-frequency subbands from NSP at each scale are decomposed by NSDFB with l stages, and this can produce 2^l directional subbands with the same size as the source image. Through this step, NSCT can extract more precise directional details information which can benefit image fusion. The NSCT construction is illustrated in Fig. 1. Fig. 1(a) is a schematic diagram of two stages decomposition frameworks of NSCT, and Fig. 1(b) denotes the structure consists of a bank of filters which splits the 2D frequency plane into the subbands.

3. Adaptive dual-channel unit-linking pulse couple neuron network

In this section, a brief review of the original unit-linking PCNN is given, and then the novel simplified dual-PCNN model based on unit-linking PCNN which can put out the fusion result directly is proposed. Compared with the original dual-channel PCNN model, it improves the linking input and decreases the numbers of parameters, and can adaptively choose the value of the parameters, such as linking strength and iteration number.

3.1. Unit-linking PCNN model

Known as third-generation artificial neural network, the pulse-coupled neural network, which has important biological backgrounds, has incomparable superiority over other current methods when applied in image processing because of the characteristics of global coupling and pulse synchronization of neurons. In the last decade, a wide variety of fusion methods based on PCNN has been proposed [7,18]. But the performance of image fusion methods using PCNN is limited due to its complexities, and some inherent drawbacks still cannot be ignored. The original PCNN model has so many parameters which are sometimes difficult to assign. At present, the parameters are commonly determined through the numerous experiments and experience. And usually, these parameters are only suitable for some certain situation. So a series of modified and simplified PCNN models have been proposed including the intersecting cortical model, the spiking cortical model [19] and the unit-linking PCNN model [2].

Among them, the unit-linking PCNN (UL-PCNN) proposed by Gu et al. [20,21], a simplified model of PCNN, has attracted many researchers' attention by virtue of its simple assignment of parameters and better performance for image processing. UL-PCNN improves the linking input, and decreases the number of parameters a lot, and also retains main characteristics of PCNN.

Like original PCNN model, UL-PCNN is a feedback network and each neuron in UL-PCNN model consists of three parts: the receptive field, modulation field and pulse generator. The UL-PCNN model is shown in Fig. 2.

As shown in Fig. 2, the neuron receives the input signals from other neurons and external stimulus S through the receptive field. In general, the signals from other neurons are pulses, and the signals from external stimulus are often the normalized gray level of image pixels. Then, the inputs are divided into two channels: feeding input F and linking input L. The difference between these two channels is that the feeding connection has a slower response time constant than that of the linking connection. In the modulation field, internal activation element U combines the feeding input with the linking input. It firstly adds a constant positive bias to the L firstly and then multiplies L by F, and the bias is taken to be unity. β is the linking strength. The internal activity U is the result of modulation and is inputted to the pulse generator. The value of internal activity element is compared with a dynamic threshold θ that gradually decreases at iteration. The internal activity element accumulates the signals until it surpasses the dynamic threshold and then fires the output element, and the dynamic threshold increases simultaneously strongly. The output of the neuron Y is then iteratively fed back to the linking input with a delay of one iteration. The output of each neuron contains two states, namely firing and non-firing. The mathematical model of the UL-PCNN is described as follows:

$$F_{ij}(n) = S_{ij}(n) \tag{1}$$

$$L_{ij}(n) = \begin{cases} 1 & \text{if } \sum_{(K,l)\in N(i,j)} Y_{kl}(n-1) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

$$U_{ij}(n) = F_{ij}(n)(1 + \beta L_{ij}(n)) \tag{3}$$

$$Y_{ij}(n) = \begin{cases} 1 & \text{if } U_{i,j}(n) > \theta_{ij}(n-1) \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

$$\theta_{ij}(n) = \exp(-\alpha_\theta)\theta_{ij}(n-1) + V_\theta Y_{ij}(n) \tag{5}$$

where the indexes (i, j) refer to the pixel location in the image or the neuron location in the UL-PCNN, n denotes the current iteration. $N(i, j)$ refers to the neighbor field of neuron (i, j). And α_θ indicates

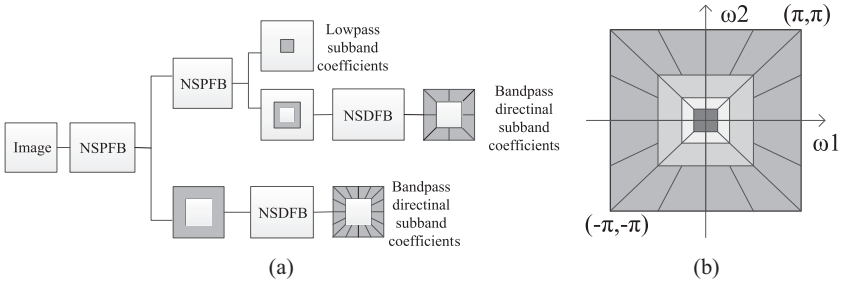

Fig. 1. Decomposition frameworks of nonsampled contourlet transform. (a) Two stages decomposition frameworks of NSCT and (b) idealized frequency partitioning obtained with the proposed structure.

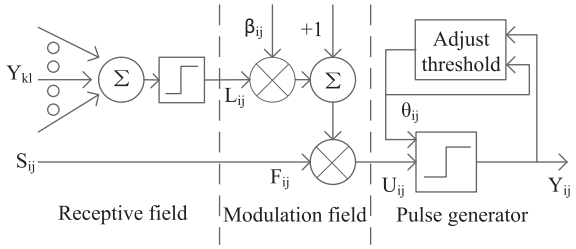

Fig. 2. Unit-linking PCNN model.

the rate of decay of the threshold in the iterative process, V_θ decides the threshold value of fired neuron.

Differently from common simplified PCNN model, the UL-PCNN improves the linking channel that to each neuron, the value of channel L_{ij} will be 1 as long as there is a fired neuron in its neighborhood (not including the center neuron itself), otherwise the value of L_{ij} will be 0. In other words, a fired neuron will fire any unfired neurons with similar input in its neighborhood. So it makes the linking inputs uniform so that the impulse expanding behaviors of networks consisting of unit-linking neurons are clear and easy to analyze and control.

In traditional UL-PCNN, the threshold function is exponentially decayed and periodically changed. Although it accords with the non-linear characteristics of the response of the human visual system to the brightness intensity, it is not a good mode for controlling the impulse and thus easily leads to computation complexity. What is more, exponential decay causes partial treatment between high-intensity and low-intensity pixels that the threshold decays fast at the high intensity and slowly at the low intensity, which makes the varied gray range process differently [22].

Therefore, in order to get over the above drawbacks and make it more suitable for image fusion, the traditional exponential decay mechanism of the threshold is replaced by a kind of simple and rational mode that threshold function θ_{ij} is dropped off linearly with time n. The expressions of θ is described as follows:

$$\theta_{ij}(n) = \theta_{ij}(n-1) - \Delta + V_\theta Y_{ij}(n) \tag{6}$$

where Δ is a positive constant for controlling the decline extent of the dynamic threshold. And Δ should be enough small to make the threshold attenuate slowly in order to partition those pixels with the similar intensity at the different fire time. The larger Δ will results in losing some image information. And V_θ is usually set as a relatively huge value to guarantee that the firing times of each neuron is not more than once at most.

When unit-linking PCNN is used for image processing, it is a single-layer two-dimensional array of laterally linked neurons. The number of neurons in PCNN is equal to that of pixels in the input image. The relation between image pixels and the network neurons is a one-to-one correspondence, and the gray value of each pixel is often input to the F channel as external stimulus of each neuron. Meanwhile, each neuron is connected with neurons in its neighboring field by the L channel.

3.2. Model of dual-channel unit-linking PCNN

Undoubtedly, there requires more than one PCNN model for image fusion using unit-linking PCNN, which makes the method complex and time-consuming. Obviously, only one stimulus for each neuron is an obstacle for multiple-image fusion using PCNN.

Due to these defects of single-channel PCNN for image fusion, inspired by Chai's model, there proposes a novel simplified dual-channel PCNN model based on unit-linking PCNN, which lessens the parameters a lot and is easy to make it adaptive, as shown in Fig 3.

Like the original UL-PCNN, each simplified dual-channel PCNN neuron also consists of three parts: the receptive field, information fusion part and pulse generator. Two kinds of inputs including external stimulus and surrounding neuron stimulus are received in the receptive field. Information fusion part is the place where all images are fused, and can be utilized to restore the fused image. The pulse generator is to generate output pulses. In simplified dual-channel PCNN model, both stimuli can be input into the model at the same time, and the output of dual-channel PCNN is the fused image. The mathematical equations of dual-channel unit-linking PCNN can be described as follows:

$$F_{ij}^1(n) = S_{ij}^1 \tag{7}$$

$$F_{ij}^2(n) = S_{ij}^2 \tag{8}$$

$$L_{ij}(n) = \begin{cases} 1 & \text{if } \sum_{(K,I)\in N(i,j)} Y_{kl}(n-1) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

$$U_{ij}(n) = \max\left\{F_{ij}^1(n)(1+\beta_{ij}^1 L_{ij}(n)), F_{ij}^2(n)(1+\beta_{ij}^2 L_{ij}(n))\right\} \tag{10}$$

$$Y_{ij}(n) = \begin{cases} 1 & \text{if } U_{i,j}(n) > \theta_{ij}(n-1) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$\theta_{ij}(n) = \theta_{ij}(n-1) - \Delta + V_\theta Y_{ij}(n) \tag{12}$$

where S_{ij}^1 and S_{ij}^2 denote the corresponding external stimulus of two images such as the normalized gray level of image pixels at (i, j) position, F_{ij}^1 and F_{ij}^2 stand for two symmetrical feed inputs, is the linking input. And β_{ij}^1 and β_{ij}^2 are the linking strength. U_{ij} is the internal state of the neuron. According to U_{ij}, we can obtain the decision map which can decide the fused image. Other parameters of the dual-channel PCNN model are the same as parameters in the original UL-PCNN model.

A remarkable characteristic of dual-channel PCNN is that two images can be input into the model at the same time, and the output of dual-channel PCNN is the fused image, which is simple and fast for image fusion. And the new dual-channel PCNN model inherits the properties of original PCNN model. The dual-channel PCNN also possesses global coupling and pulse synchronization characteristics, which take full advantage of local image information and benefit image fusion.

3.3. Adaptive parameters setting in dual-channel PCNN

In traditional PCNN based image fusion, parameters mainly depend on the large number of experiments or experience. And a

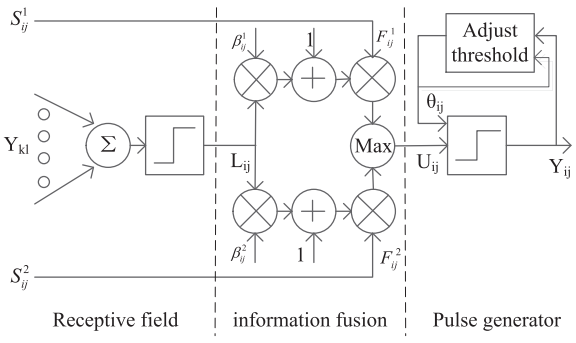

Fig. 3. The model of a dual-channel unit-linking PCNN.

set of parameters which achieve good performance in certain situation may be unsuitable for other applications. To implement the automation of PCNN parameters, Li et al. employed PSO to set the parameters adaptive, but it requires large number of iterations and may be complex. In literature [15], Y. Chai utilized orientation information as the linking strength, but the other parameters were set by manually. So Chai's model just implements part of the automation. In this paper, we apply the unit-linking PCNN model and improve it to dual-channel PCNN model, which reduces parameters a lot. In dual-channel PCNN, the number of parameters requiring setting is just four in all, namely β, Δ, V_θ and n. Among them, Δ and V_θ are easy to determine as described previously. With regard to the step Δ , we can set it as 0.01 to ensure that the decayed speed of the dynamic threshold is moderate and acceptable. In order to limit the firing times of each neuron to no more than one, the parameter V_θ can be simply assigned by a comparatively larger value such as 20. And taking account of the parameters β and n, which are significant for image fusion, the average gradient of each pixel in images is employed to determine the linking strength β adaptively, and the time matrix of the images is presented to determine the iteration number n adaptively.

The parameter β, namely linking strength, reflects the pixel characteristics and can adjust the weighting of the linking channel in the internal activity, so it plays an important role in the proposed image fusion method. Usually, most literatures assign the same constant to the linking strength of each neuron through experiments or experiences, such as 0.2 [23]. This setting manner can sharply simplify the problem in basic PCNN model. However, according to the analysis of literature [24], it does not accord with the actuality of images, because the linking strength of each neuron in PCNN model should be relevant to the features of the corresponding pixels of the images. And this shortcoming is also a big limit to the automatic process. So many researchers choose the clarity of each pixel, which denotes the notable features of images, to adjust the linking strength, such as spatial frequency [25], orientation information [15], standard deviation [26] etc. In this paper, considering that the gradient of each pixel is a notable feature of the local neighborhood region and reflects the edges of images, so we utilize the average gradient to determine the linking strength. The average gradient of image represents the change of the gray value, and is one of the key criterions to reflect the clarity of image, and what is more, it can denote the details of edge, texture. The mathematical expression of average gradient is as follows:

$$\bar{g}(i,j) = \frac{1}{9}\sum_{m=-1}^{1}\sum_{n=-1}^{1}\sqrt{[g_1(i+m,j+n)+g_2(i+m,j+n)]/2} \tag{13}$$

$$g_1(i,j) = [I(i,j) - I(i+1,j)]^2 \tag{14}$$

$$g_2(i,j) = [I(i,j) - I(i,j+1)]^2 \tag{15}$$

where $\bar{g}(i,j)$ is the average gradient at position (i,j) in the 3×3 local region, $g_1(i,j)$ and $g_2(i,j)$ illustrate the change of the gray value in the horizontal and vertical direction respectively.

In general, the value of β is between 0 and 1. Therefore, we exploit the sigmoid function to normalize the average gradient. Suppose that $\beta(i,j)$ denotes the linking strength at (i, j) neuron, so the linking strength is defined as follows:

$$\beta_{ij} = \frac{1}{1+\exp(-\bar{g}(i,j))} \tag{16}$$

From formula (16), we can see that β reflects the change of gray value in the image window. The larger the average gradient is, the better the clarity of the local region is, and the larger the value of β is, so the earlier the neuron fires. The linking strength can adaptive adjust according to the average gradient, which makes the PCNN model preserve the image details effectively and improve the performance of image fusion.

Generally speaking, it is hard to set the iteration number n in PCNN image processing. In most literatures, the iteration number n is decided by experiments or experience. In Das et al.'s [23] paper, the iteration number n of PCNN is set as 200 times. It is known to all, if n is too large, it will be time-consuming, and if n is set to be too small, the synchronous impulse characteristic of PCNN will not be taken full advantage of to get optimal image processing effect. And those methods whose iteration number n is set as a fixed constant are only applicable to certain specific occasions.

In order to set the iteration number n properly, enlightened by literature [27], we adopt the time matrix T whose size is equivalent to the external input S to determine the iteration number n adaptively. T is defined as follows:

$$T_{ij}(n) = \begin{cases} n, & \text{if } Y_{ij}(n) = 1 \text{ for the first time} \\ T_{ij}(n-1), & \text{otherwise} \end{cases} \tag{17}$$

where T_{ij} denotes each neuron's first firing time. The time matrix is a mapping from the spatial image information to the time information, providing a genuine storage of the information about the firing time of each neuron. It is necessary to explain and note three points about time matrix: (1) T_{ij} is set to be zero if the neuron has never fired; (2) T_{ij} is set to be n if the neuron fires for the first time at the time of n; (3) T_{ij} remains invariable if the neuron has fired. The iteration goes on until all neurons have been fired, that is to say, each element value of T is nonzero. Moreover, in the time matrix T, those pixels whose image intensity values are similar in the source image often share the same or similar firing time. Thus by using time matrix, the dual-channel unit-linking PCNN model can not only describe time information of each neuron, but also preserve space information including image gray distributing, which will benefit further processing much more.

4. Fusion method

In this section, the proposed fusion scheme will be discussed in detail. Firstly, NSCT has been introduced into the fusion algorithm in this paper. After the multi-scale and multi-direction decompositions of the source image, the low-frequency sub-image denotes the approximate component, which represents the main information of the source image, whereas a series of high-frequency sub-images that reflect the details components, contains the edge details information of the source image from different directions and scales. So, the fusion rules for the subband images decomposed by NSCT are very important for the quality of fusion.

At present, fusion rules that average or weighted average are commonly used in low-frequency domain, and max absolute is often used in high-frequency domain. For the fusion of infrared and visible images, these methods may lose part of spectrum information and details of source images, and reduce the contrast. Considering the global coupling and pulse synchronization characteristics of PCNN, we take the dual-channel unit-linking PCNN, which make full use of local image information and can effectively extraction the details of image, to make an intelligent decision on the selection of high- and low-frequency NSCT coefficients.

In most of the multi-scale analysis and PCNN based algorithms, the value of single pixel in multi-scale decomposition domain is directly used to motivate one neuron. In fact, human's visual system in most time is highly sensitive to edges, directional features, etc., yet insensitive to real luminance at independent positions. So a pure use of single pixel is not enough. It will be more reasonable to exploit features, rather than the value of pixels, to motivate PCNN neurons. In this paper, we utilize a novel

sum-modified-Laplacian [28] (NSML), which can represent the edge features of the low-frequency sub-images in NSCT domain, and a modified spatial frequency [28], which stands for gradient energy of the high-frequency sub-images in NSCT domain, to motivate the dual-channel PCNN neurons. The detailed fusion rules can be described below.

4.1. Lowpass subband fusion rule

In this paper, we use a novel sum-modified-Laplacian (NSML), which can reflect the edge details of the lowpass subband images, to motivate the PCNN neurons. Supposed that $C(i, j)$ denotes the coefficients of the low-frequency subband image at (i, j) position, the modified Laplacian energy (ML) and NSML are described as follows:

$$\begin{aligned} \mathrm{ML}(i,j) &= |2C(i,j) - C(i-\mathrm{step},j) - C(i+\mathrm{step},j)| \\ &\quad + |2C(i,j) - C(i,j-\mathrm{step}) - C(i,j+\mathrm{step})| \end{aligned} \tag{18}$$

$$\mathrm{NSML}k(i,j) = \sum_a \sum_b W_1(a,b)[ML(i+a,j+b)]^2 \tag{19}$$

where 'step' is the variable spacing between the coefficients and is equal to 1 in this paper, a and b denotes the size of the widow for computing the NSML, usually 3×3, 5×5 or 7×7, and 3×3 in this paper. And W_1 is a weighted template, which is given as:

$$W_1(a,b) = \frac{1}{15}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \tag{20}$$

4.2. Highpass subband fusion rule

For the high-frequency sub-images, we introduce a modified spatial frequency (MSF) which used as the gradient features of images to motivate dual-channel PCNN and generate pulse of neurons. Spatial frequency (SF) is measured by using slipping window of coefficients in subbands. It measures the entire activity in the window-based coefficients via the gradient energy in rows and columns. In this paper, the spatial frequency is extended, in addition to the calculation of changes in horizontal and vertical directions, the frequency changes of the two diagonal directions are increased. And so the image information is reflected more comprehensively.

We take $C_{l,k}(i,j)$ denotes the coefficient located at (i, j) in the k-th directional subband at the l-th decomposition level of high-frequency sub-images, the modified spatial frequency (MSF) is defined as:

$$\mathrm{MSF} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(\mathrm{RF} + \mathrm{CF} + \mathrm{MDF} + \mathrm{SDF}) \tag{21}$$

with

$$\mathrm{RF} = [C_{l,k}(i,j) - C_{l,k}(i,j-1)]^2 \tag{22}$$

$$\mathrm{CF} = [C_{l,k}(i,j) - C_{l,k}(i-1,j)]^2 \tag{23}$$

$$\mathrm{MDF} = [C_{l,k}(i,j) - C_{l,k}(i-1,j-1)]^2/\sqrt{2} \tag{24}$$

$$\mathrm{SDF} = [C_{l,k}(i,j) - C_{l,k}(i-1,j+1)]^2/\sqrt{2} \tag{25}$$

where RF, CF, MDF and SDF, express the row frequency, the column frequency, the main angular frequency and the auxiliary diagonal frequency respectively. And M and N denote the size of the window for computing MSF. In this paper, we take 3×3 neighbor region.

4.3. Fusion step

The schematic diagram of the proposed fusion scheme is shown in Fig. 4. Before fusion, all source images must be spatially registered. The detailed fusion process consists of the following four steps:

(1) Perform the NSCT on the registered source infrared and visible images, respectively, and obtain one lowpass subband image and a series of highpass directional subband images.
(2) The coefficients of the sub-images from infrared and visible images are all normalized between 0 and 1. And calculate NSML as described in formula (18)–(20) on lowpass subband and MSF as described in formula (21)–(25) on highpass subband. Take them respectively as the feeding input to stimulate PCNN.
(3) Select fusion NSCT coefficients for the low-frequency subband coefficients and the high-frequency subband coefficients via dual-channel unit-linking PCNN. The fusion process can be described below:
   a. Initialize $U_{ij}(0) = Y_{ij}(0) = T_{ij}(0) = 0, \theta_{ij}(0) = 1$ meanwhile, each neuron does not fire. The mode of setting $\theta_{ij}(0)$ above is to make neurons be activated as soon as possible, which can prevent unnecessary "void" iterations.
   b. Take the value of average gradient of each pixel as the linking strength, according to the formula (16).
   c. Take the NSML of low-frequency sub-images and MSF of high-frequency sub-images, as the external stimulus to motivate the dual-channel PCNN neurons, respectively. And then compute $U_{ij}(n), \theta_{ij}(n), Y_{ij}(n), T_{ij}(n)$ according to formula (7)–(12) and (17).
   d. Implement step (c) iteratively until all neurons have been activated, namely each element in T is nonzero. The fused coefficients can be selected as follows:

$$C_F(i,j) = \begin{cases} C_I(i,j) & \text{if } U_{ij}(N) = U_{ij}^I(N) \\ C_V(i,j) & \text{if } U_{ij}(N) = U_{ij}^V(N) \end{cases} \tag{26}$$

with

$$U_{ij}^I(N) = F_{ij}^I(N)(1 + \beta_{ij}^I L_{ij}(N)) \tag{27}$$

$$U_{ij}^V(N) = F_{ij}^V(N)(1 + \beta_{ij}^V L_{ij}(N)) \tag{28}$$

where $C_F(i, j)$, $C_I(i, j)$ and $C_V(i, j)$ denote the coefficients of the fused image, the infrared image and the visible image, respectively. And $U_{ij}^I(N), U_{ij}^V(N)$ stand for the internal state of the neuron of infrared and visible image, respectively. N denotes the total fire times.

(4) Reconstruct the fused image by using an inverse NSCT.

5. Experimental results and analyses

To demonstrate effectiveness of the proposed fusion method, some experiments have been performed. In this section, we test the proposed algorithm through two typical groups of infrared and visible images.

5.1. Experimental introduction

For comparison purposes, the proposed algorithm is assessed and compared with other three current fusion methods: the DWT method, the NSCT method and the NSCT–PCNN method [15]. The former two methods take the absolute maximum choosing rule for high-frequency coefficients and averaging combination rule for low-frequency coefficients. As for NSCT–PCNN method, the averaging scheme and the typical PCNN-based scheme fusion rule

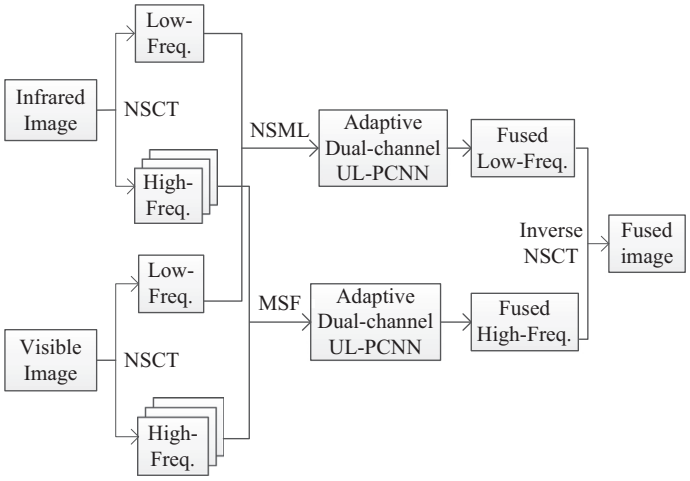

Fig. 4. The schematic diagram of the proposed fusion scheme.

are used to fuse the lowpass subband coefficients and the highpass subband coefficients in NSCT domain respectively. And all multi-scale transform take 3 layers decomposition except NSCT, other parameters are as follows: DWT takes 'db2'; NSCT takes "9/7" and "pkva", together with a decomposition level of 4, and the number of direction from coarser to finer scale is set to be [0,2,3,4], respectively. Parameters of the PCNN used in NSCT–PCNN method are set as: $\alpha_L = 0.06931$, $\alpha_\theta = 0.2$, $V_L = 1$, $V_\theta = 20$, $\theta = 0.2$, $N = 200$, and linking weight $W = [0.707,1,0.707;1,0,1;0.707,1,0.707]$.

For the purpose of better evaluating the effect of the fused images, besides visual observation, this paper has adopted the objective evaluation to evaluate results of all kinds of fusion methods objectively. We adopt five objective metrics: information entropy (IE), mutual information (MI), standard deviation (SD), Q_E and $Q^{AB/F}$ [29]. IE quantifies the richness of information in the fused image. The larger the IE value is, the more abundant the information amount of the fusion image is. MI essentially computes how much information from source images is transferred to the fused image. With the increase of the value of mutual information, fused image can get richer information from the source image. The SD shows the statistical distribution of fused image and describes the contrast of the fused image. The bigger the SD is, the more dispersed the distribution of gray level in image is, and the greater the contrast is, the better the visualization of the fused image is. Q_E uses correlation, luminance distortion and contrast distortion to measure the fused images. The higher the Q_E is, the more the saliency information of source images is contained in the fused image. $Q^{AB/F}$ computes and measures the amount of edge information transferred from the source images to the fused images using a Sobel edge detector to calculate strength and orientation information at each pixel in both source and the fused images. The larger the $Q^{AB/F}$ is, the more edge information the fused image preserves, and the better the performance of fusion is.

5.2. Performance evaluation

The first experiment is performed on the 'UN Camp' infrared and visible images which have been registered perfectly. Fig. 5 illustrates the source images and the fusion results obtained by the above different methods. Fig. 5(a) and (b) are visible light image and infrared image which obtained in the same scene. And fusion results using DWT, NSCT, NSCT–PCNN and our proposed method are displayed in Fig. 5(c)–(f).

From Fig. 5(a) and (b), we can see that the pedestrian in infrared image is in clearly sight, and infrared image is favorable for the recognition of the interesting targets, whereas the background information in infrared image is very poor and the scene about pedestrian is a little heavily blurred, e.g., it is difficult to recognize the shrubs and fences. And in the visible image, it is nearly impossible to distinguish the pedestrian owing to the dim light and the shelter of trees, but it has rich background information and details of edges texture, and is clearly to identify the roads, bushes, fences etc.

As can be seen from Fig. 5(c)–(f), four methods successfully fuse the infrared and visible images, and all the fused images contain the target information and background information. However, we can find that the fused result using DWT has low contrast and loses many details. The target of pedestrian is not prominent, and details of roads, bushes etc. are blurred. Evidently, many artifacts are introduced in the fused image because of the lack of shift-invariance which causes pseudo-Gibbs phenomena. Yet the NSCT based and NSCT–PCNN based methods achieve a better performance than DWT-based method. Due to the shift-invariant of NSCT, the pseudo-Gibbs phenomenon has been well eliminated and the fused images are clearer and more natural than the DWT fused results. It is proven that shift-invariant methods can overcome the pseudo-Gibbs phenomena successfully and improve the quality of the fused image around the edges. Furthermore, target information is more prominent and background information such as road, shrubs is much abundant compared with the DWT-based method. But we can see that these two methods still have some defects. The fusion results lose much spectral information, such as the trees of bottom left and bottom right in the fused images. Obviously, the proposed method provides best visual effects. Almost all the useful information of the source images has been transferred to the fused image, and meantime, fewer artifacts are introduced during the fusion process. The proposed method not only obtains the higher contrast and highlights the target of pedestrian, but also possesses rich spectral information and preserves the edges and detailed information well, especially the trees in the bottom left of fused image which are injected much more texture information from visible image.

Table 1 shows evaluation results of four methods in Fig. 5. The IE of the fused image obtained by the proposed method is maximum, which means that the fused image contains the largest amount of information and has relatively better fusion result than others. The MI achieves the best result of all, which illustrates that the fused image based on the proposed method extracts more information from the original images. And the SD of the fused image by our method is maximum, which represents that the fused image has the best contrast and the notable target. What is more, the Q_E of the proposed method gains the maximum, and the $Q^{AB/F}$

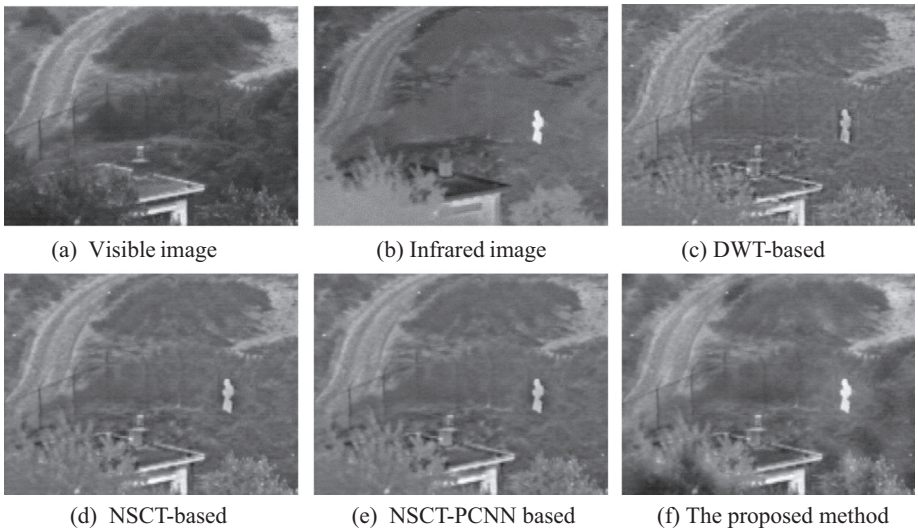

Fig. 5. Fusion results of "UN Camp" images; (a) source visible image; (b) source infrared image; (c)–(f) fusion results of DWT-based, NSCT-based, NSCT–PCNN based and the proposed method.

Table 1
Evaluation results of four methods for "UN Camp".

Methods	IE	MI	SD	Q_E	$Q^{AB/F}$
DWT	6.4646	1.4691	25.4912	0.1492	0.3966
NSCT	6.5536	1.5246	27.0326	0.2000	0.4591
NSCT–PCNN	6.5269	1.5338	26.5558	0.2052	0.4464
Proposed	6.9434	1.7881	33.5729	0.2127	0.4464

of the proposed method acquires the second maximum, which denotes that the method in our paper extracts more edge information from source images and preserves the detailed information effectively. The objective evaluation agrees with the visual observation.

The second experiment is performed on the 'tank' infrared and visible images. The source images and fused results are displayed in Fig. 6. In Fig. 6, (a) and (b) are visible image and infrared image. And Fig. 6(c)–(f) are the results of four fusion methods. From the Fig. 6(a) and (b), the tank in Fig. 6(a) is almost hard to recognize, while the background such as grass, trees can be observed clearly. And the infrared image concerning the same scene in Fig. 6(b) is easy to identify the tank, but the background is very poor.

From Fig. 6(c)–(f), four methods successfully synthesize the target information of the tank in the infrared image and the background information in the visible image, and reach good fusion performance. But the contrast of the fused image based on DWT is relatively low, and the DWT-based result loses much infrared information such as the tail and wheels of the tank and spectral information in the visible image such as trees. What is worse, the fused image is troubled by the pseudo-Gibbs phenomena, especially at the wheels of tank. While the methods of NSCT-based and NSCT–PCNN based improve the fusion effect greatly. They remove the pseudo-Gibbs phenomena, and achieve high contrast. Regrettably these two methods cannot fully extract the details from the source images. For instance, the fusion of the tail of tank is not sufficient, and the texture of trees is not abundant. The proposed method is evidently better than other methods mentioned above. It fully synthesizes the infrared and visible images, highlights the target information of tank. The proposed method

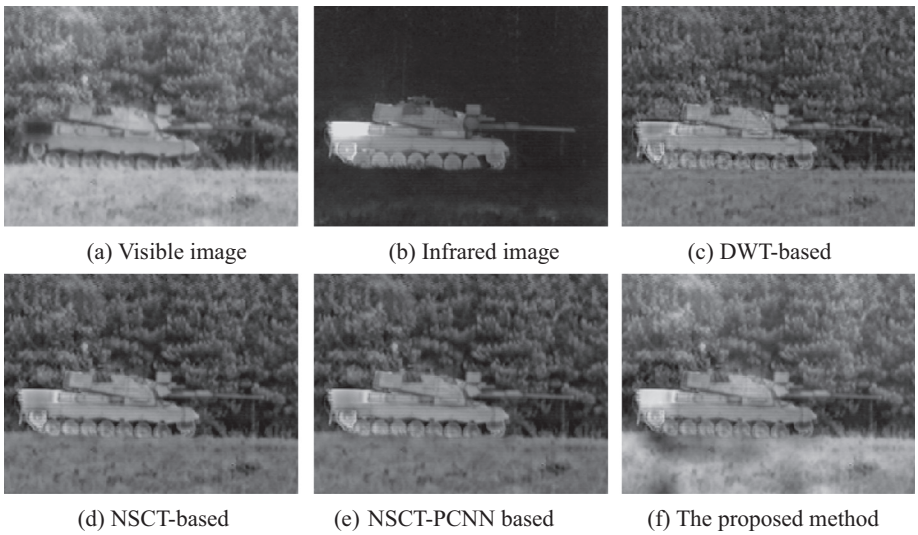

Fig. 6. Fusion results of "Tank" images; (a) source visible image; (b) source infrared image; (c)–(f) fusion results of DWT-based, NSCT-based, NSCT–PCNN based and the proposed method.

Table 2
Evaluation results of four methods for "Tank".

Methods	IE	MI	SD	Q_E	$Q^{AB/F}$
DWT	6.9422	2.4997	31.9431	0.3212	0.5421
NSCT	7.0257	2.5335	33.7976	0.4193	0.6380
NSCT–PCNN	7.0076	2.5341	33.3793	0.4166	0.5985
Proposed	7.3777	3.6892	41.6931	0.5090	0.6528

extracts the trees, tank etc. better and preserves the useful information better than other methods. Meantime, fewer artifacts are introduced during the fusion process.

In comparison, objective criteria used to compare the fusion results are listed in Table 2. These five objective criteria receive the maximum compared with the other three methods, which proves that the fused image of the proposed method contains abundant image information, has a high contrast, preserves more image features and is strongly correlated with the source images.

As can be seen from these two experiments, the objective evaluation results coincide with the visual effect very well. From what has been discussed above, it is can be concluded that the proposed method does well in the fusion of infrared and visible images and outperforms the DWT, NSCT, and NSCT–PCNN fusion algorithms, whether in visual observation or objective evaluation criterion.

6. Conclusions

In this paper, according to original PCNN, a novel adaptive dual-channel unit-linking PCNN for the fusion of infrared and visible images is proposed in the NSCT domain. Compared with the dual-channel PCNN models proposed in [8] and [15], our proposed dual-channel PCNN model is devised to be more simple and adaptive, which possesses much fewer parameters and is suitable for image fusion, and it can put out the fused image directly. To make dual-channel unit-linking PCNN adaptive, we take the average gradient of each pixel in images as the linking strength, and the time matrix to determine the iteration number adaptively. In the proposed algorithm, a novel sum-modified-Laplacian, which stands for edge features in low frequency sub-images in NSCT domain, is used to motivate the adaptive dual-channel unit-liking PCNN neurons. For the high frequency sub-images, a modified spatial frequency which represents the gradient energy of sun-images is presented to motivate the adaptive dual-channel unit-liking PCNN neurons. The flexible multi-resolution and directional expansion for images of NSCT are associated with global coupling and pulse synchronization characteristics of PCNN. Two groups of experiments on evaluating the fusion performance have been conducted and the results show that the proposed algorithm can effectively fuse infrared and visible images, has high contrast and remarkable target information, and preserve rich information of texture and details well. The proposed method is superior to the other current fusion algorithms in terms of both visual quality and quantitative evaluation.

Conflict of interest

No conflict of interest.

Acknowledgements

The authors would like to thank the editor, associate editor and the anonymous reviewers for their careful work and invaluable suggestions for an earlier version of this paper. This work was jointly supported by the National Natural Science Foundation of China (Grant No. 41271456). We are also grateful to TNO Human Factors Research Institute in the Netherlands and Song Le Ph.D. for providing the experiment images.

References

[1] X.Q. Zhang, Q.L. Chen, T. Men, Comparison of fusion methods for the infrared and color visible images , in: Second IEEE International Conference on Computer Science and Information Technology, Beijing, 2009, pp. 421–424.
[2] W.W. Kong, Y.J. Lei, Y. Lei, et al., Image fusion technique based on nonsubsampled contourlet transform and adaptive unit-fast-linking pulse-coupled neural network, IET Image Process. 5 (2) (2011) 113–121.
[3] A.L. Da Cunha, Zhou Jianping, M.N. Do, The nonsubsampled contourlet transform: theory, design, and applications, IEEE Trans. Image Process. 15 (10) (2006) 3089–3101.
[4] M.N. Do, M. Vetterli, The contourlet transform: an efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.
[5] Li Huafeng, Chai Yi, Li Zhaofei, Multi-focus image fusion based on nonsubsampled contourlet transform and focused regions detection, Int. J. Light Electron Optic. 124 (1) (2013) 40–51.
[6] R. Eckhorn, H.J. Reitboeck, Arndt Mt, Dicke P. Feature linking via synchronization among distributed assemblies: Simulations of results from cat visual cortex, Neural Comput. 2 (3) (1990) 293–307.
[7] M. Monica Subashini, Sahoo Sarat Kumar, Pulse coupled neural networks and its applications, Expert Syst. Appl., 41(8) (2014) 3965–3974.
[8] Wang Zhanbin, Ma Yide. Dual-channel PCNN and its application in the field of image fusion, in: Third IEEE International Conference on Natur. Comput., Jinan, 2007, pp. 755–759.
[9] Wang Zhaobin, Ma Yide, Medical image fusion using m-PCNN, Inform. Fusion. 9 (2) (2008) 176–185.
[10] Wang Zhaobin, Ma Yide, Gu Jason, Multi-focus image fusion using PCNN, Pattern Recognit. 43 (6) (2010) 2003–2016.
[11] Chang Weiwei, Guo Lei, Fu Zhaoyang, et al., Hyperspectral multi-band image fusion algorithm by using pulse coupled meiral networks, J. Infrared Milli. waves. 29(03) (2010) 205–209.
[12] Fu Zhaoyang, Guo Lei, Chang Weiwei, Fusion algorithm of hyperspectral images based on wavelet transform and multi-channel PCNN, J. Jilin Univ. (Eng. Technol. Ed.). 41 (03) (2011) 838–843.
[13] Guo Xinpin, Duan Xianhua, Xia Jiaxing, Application and research based on two-channel PCNN, Sci. Technol. Eng. 12 (34) (2012) 9225–9233.
[14] Zhao Yaqian, Zhao. Qinping, Hao. Aimin, Multimodal medical image fusion using improved multi-channel PCNN, Bio-Med. Mater. Eng. 24 (1) (2014) 221–228.
[15] Y Chai, H.F. Li, J.F. Qu, Image fusion scheme using a novel dual-channel PCNN in lifting stationary wavelet domain, Opt. Commun. 283(19) (2010) 3591–3602.
[16] El-taweel Ghada Sami, Helmy Ashraf Khaled, Image fusion scheme based on modified dual pulse coupled neural network, IET Image Process. 7 (5) (2013) 407–414.
[17] Li Yi, Wu Xiaojun. A novel image fusion method using self-adaptive dual-channel pulse coupled neural networks based on pso evolutionary learning, Acta Electronica Sinica. 42(2) (2014) 217–222.
[18] Wang Zhaobin, Ma Yide, Cheng Feiyan, etc. Review of pulse-coupled neural networks, Image Vis. Comput. 28(1) (2010) 5–13.
[19] Wang Nianyi, Ma Yide, Zhan Kun, Spiking cortical model for multifocus image fusion, Neurocomputing 130 (2014) 44–51.
[20] Gu Xiaodong, Zhang Liming, Yu Daoheng. General design approach to unit-linking PCNN for image processing, in: IEEE International Joint Conference on Neural Networks, Montreal, Canada, 2005, pp. 1836–1841.
[21] Gu Xiaodong, Feature extraction using unit-linking pulse coupled neural network and its applications, Neural Process. Lett. 27 (1) (2008) 25–41.
[22] Cui Kebin, Li Baoshu, Yuan Jinsha, Wang Ping, An improved unit-linking PCNN for segmentation of infrared insulator image, Appl. Math. Inform. Sci. 8 (6) (2014) 2997–3004.
[23] Das Sudeb, Kundu Malay Kumar, NSCT-based multimodal medical image fusion using pulse-coupled neural network and modified spatial frequency, Med. Biol. Eng. Comput. 50 (10) (2012) 1105–1114.
[24] Y. Chai, H.F. Li, M.Y. Guo, Multifocus image fusion scheme based on features of multiscale products and PCNN in lifting stationary wavelet domain, Opt. Commun. 284 (5) (2011) 1146–1158.
[25] Kong Weiwei, Zhang Longjun, Lei Yang, Novel fusion method for visible light and infrared images based on NSST–SF–PCNN, Infrared Phys. Technol. 65 (2014) 103–112.
[26] Xi Cai, Wei Zhao, Fei Gao, Image fusion algorithm based on adaptive pulse coupled neural networks in curvelet domain, in: IEEE 10th International Conference on Signal Processing, Beijing, 2010, pp. 845–848.
[27] Liu Qiong, Ma Yide, A new algorithm for noise reducing of image based on PCNN time matrix, J. Electron. Inform. Technol. 8 (2008) 1869–1873.
[28] Huang Wei, Jing Zhongliang, Evaluation of focus measures in multi-focus image fusion, Pattern Recognit. Lett. 28 (4) (2007) 493–500.
[29] M.B.A. Haghighat, A. Aghagolzadeh, H. Seyedarabi, A non-reference image fusion metric based on mutual information of image features, Comput. Electr. Eng. 37 (5) (2011) 744–756.