

Region growing for image segmentation using an extended PCNN model

Dongguo Zhou¹ , Yanhua Shao²

¹School of Power and Mechanical Engineering, Wuhan University, Wuhan 430072, People's Republic of China

²School of Information Engineering, Southwest University of Science and Technology, Sichuan 621010, People's Republic of China

 E-mail: dongguozhou@gmail.com

Abstract: Pulse-coupled neural network (PCNN) is a biologically-inspired algorithm suited for image processing. However, determining a set of parameters involved in the alteration of the neural behaviour remains a prevalent research for further application. To apply the model into image segmentation, this study proposes an extended PCNN model by using a strategy of the decision tree, and establishes links between the parameters and image characteristics. Particularly, the adjustable threshold term, interacted with the estimation of the global neural threshold, enables the proposed model to obtain the better results with the use of the fuzzy set theory. Through iterative computation, the proposed model can be considered as a region growing approach for multilevel image segmentation, thus named as an extended PCNN model. Finally, experiments on synthetic and natural images demonstrate the efficiency of the proposed model. Moreover, comparisons with some existing PCNN-based models, and recently graph-based methods, normalised cuts, show that the proposed model can extract regions with more similarity.

1 Introduction

Pulse-coupled neural network (PCNN), motivated by the observation of synchronous pulse burst in the visual cortex of mammals such as cats [1], has been identified as having interesting application prospects in the field of image segmentation [2–4]. Up to now, PCNN has been applied into the segmentation of an extensive range of real-world images, including infrared images [5–9], embryonic plant cells [10], magnetic resonance images [11], and fabric defect [12], and it is likely to be continuously developed, like level set, watershed, and so on [13–15].

Generally speaking, PCNN model for image segmentation is highly dependent on the values of network parameters. Among these parameters are the amplitudes and decay constants in each leaky integrator, two weight matrices, and a linking coefficient, making the user flexible to alter the neural behaviour. However, the inner mechanism of parameter settings still remains uncertain because of its biological background [2].

To alleviate the complex task of determining the values of parameters, many studies have focused on the simplified or modified version of PCNN model, and various automatically setting approaches have also been proposed. The classic modified model developed by Kuntimad and Ranganath [16] was considered as a perfect segmentation method for an image after the intensity ranges of adjacent regions are obtained. Following this model, Raya *et al.* [17] considered the linking coefficient and primary neural threshold as global and local values, respectively, and then obtained them directly from image statistics. However, such available image statistics rely on the criterion of maximising the variance within the class, rather than the dynamic properties of neurons. Recently, to determine the relationship between the dynamic properties of neurons and image characteristics, Chen *et al.* [18] derived several criteria to deal with the parameter settings of a simplified model [19]. In [5, 6], efficient methods were proposed to establish an automatic and simple way to segment the infrared image. Nevertheless, the derived conditions have relation with the image characteristic like the distribution of classes. The methods [9, 20] we proposed recently utilise the local information to preserve the characteristic of the synchronous pulse inherent in PCNN.

Alternatively, there are some simplified or modified models, which take advantage of a basic and the specific mechanism for producing the outputs via setting of the neural threshold as well as the inner parameter. Wei *et al.* [21] established the threshold decay time constant of the overall features of images based on the subjective feeling of the human eye logarithmically related to the actual light intensity. The works in literatures [22–25] then used a strategy to linearly decrease the threshold value. However, the neural threshold, which periodically decays, seems to be unsuitable for segmenting other types of images. Kong *et al.* [8] thereby entirely abandoned this mechanism and instead used the average grey value of a peak in the histogram for the corresponding threshold. Karvonen [26] used a fixed classwise threshold value which is derived from the histogram of a given image through the expectation-maximisation method. Li *et al.* [27] developed a neural threshold setting method in which the threshold value is determined by using the water region in the area histogram. In the literatures [28–30], the behaviour of PCNN model was considered as a region growing method. Our previous works [5, 6, 9] were trying to use the image characteristics, allowing the modified neural threshold to be updated adaptively. However, an important issue that should be further considered is that, during separating the whole image into three or more meaningful regions using PCNN model, the mechanism of the modified neural threshold is still somewhat sensitive to other inner parameter, such as the linking coefficient.

In this paper, we present a modified version of PCNN with proper parameters setting, and then extend it into a region growing model for separating the whole into several meaningful regions. In contrast to some existing PCNN-based models, the main contributions of our model are as follows. The first contribution is related to the threshold term for adjusting the output, which is more useful in removing the negative effect of the dynamic threshold and providing the remedy for blurring characteristic of regional boundary. The second contribution involves the critical problem of synchronous pulse inherent in PCNN for capturing neurons with soft decision. Lastly, the model is extended by using a strategy of the decision tree, making it possible to divide a given image into reasonable clusters for further analysis, such as image representation, object recognition, and so on.

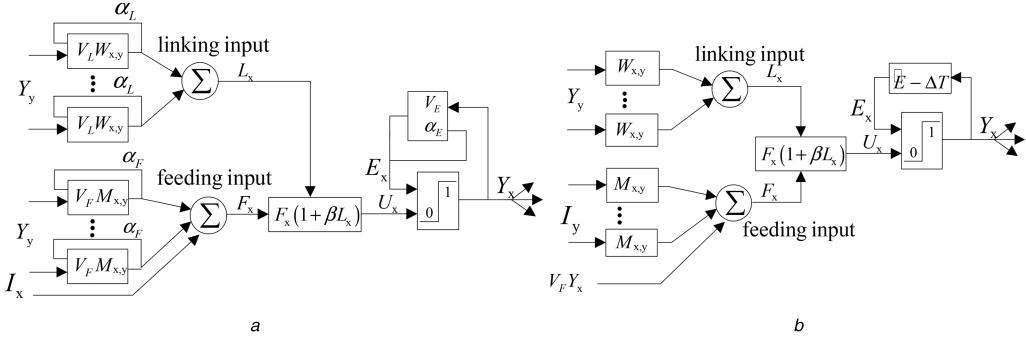


Fig. 1 Schematic representation of PCNN

(a) General pulse-coupled neuron, (b)Modified pulse-coupled neuron

The organisation of the paper is as follows. In Section 2, we will present a brief review of the original PCNN. Section 3 introduces our modified version of PCNN and its behaviour to image segmentation. The parameters are then set by coordinating with image statistics, and its relationship between them is also built in Section 4. In Section 5, experimental results of the proposed model on synthetic and real-world images are given, and the results are then compared with some existing PCNN-based methods, and widely-used method, including graph-based methods, normalised cuts, to demonstrate the efficiency of our model in terms of segmentation performance. Finally, conclusions are drawn in Section 6.

2 Principle of PCNN for image segmentation

PCNN is a two-dimensional neural network. Each neuron in the processing layer is directly related to a corresponding pixel in the image and takes into account the pulse from its neighbouring neurons. Generally, a model of a pulse-coupled neuron (Fig. 1a) consists of three parts: (i) acceptance of the external stimulus and the pulse output from neighbouring neurons, (ii) modulation of inputs to increase the internal activity, and (iii) generation of a pulse via comparison.

In contrast to other types of neural networks for image segmentation, PCNN can be considered as a new-generation neural network without any training [31], and its fundamental idea of segmentation is to operate by iterating the following equations:

$$F_x(n) = e^{-\alpha_F} F_x(n-1) + I_x + V_F \sum_{y \in N_x} M_{x,y} Y_y(n-1), \quad (1)$$

$$L_x(n) = e^{-\alpha_L} L_x(n-1) + V_L \sum_{y \in N_x} W_{x,y} Y_y(n-1), \quad (2)$$

$$U_x(n) = F_x(n) \cdot [1 + \beta L_x(n)], \quad (3)$$

$$Y_x(n) = \begin{cases} 1, & U_x(n) > E_x(n-1) \\ 0, & \text{otherwise} \end{cases}, \quad (4)$$

$$E_x(n) = e^{-\alpha_E} E_x(n-1) + V_E Y_x(n), \quad (5)$$

where subscript x denotes a neuron, corresponding a pixel in the image. Feeding input denoted as F mainly consists of three parts: (i) the input from external stimulus I (e.g. intensity); (ii) the input from the pulse output of eight-neighbouring neurons N_y through the weight matrix W , adjusted by the amplitude V_F ; and (iii) the input from the previous state, where the parameter α_F denoted as the decaying factor is used to enable the feeding input to keep the part information of the previous state. Similarly, linking input L has the part (ii) and part (iii) with another group of parameters, amplitude V_L , decaying factor α_L , and weight matrix M . Here, the weight matrix M and W are often set as follows:

$$M_{x,y}/W_{x,y} = \begin{cases} 0, & x = y \\ 1/\|x - y\|^2, & x \neq y \end{cases}. \quad (6)$$

In the pulse generator part, the threshold E also has the part (ii) and part (iii) with the decaying factor α_E and amplitude V_E , regardless of the weight M/W . Since two inputs are coupled together by linking coefficient β , the internal activity U_x is then generated. Basically, the neuron fires or generates a pulse ($Y_x = 1$) as soon as the accumulated internal activity U_x of a corresponding neuron surpasses its inner dynamic threshold E_x . The result will then alter the threshold, the linking and feeding inputs, and the behaviour of neighbouring neurons at the next iteration. Through iterative computation, PCNN will produce a sequence of binary images which contains a desired result of segmentation.

3 Modified PCNN and its application to image segmentation

3.1 Modified PCNN

Fig. 1b shows a modified version of PCNN model. Generally, the major changes of this model lie in the feeding and linking inputs and the neural threshold, as seen in the following iterative circle:

$$F_x(n) = \sum_{y \in N_x} M_{x,y} I_y + V_F Y_x(n-1), \quad (7)$$

$$L_x(n) = \sum_{y \in N_x} W_{x,y} Y_y(n-1), \quad (8)$$

$$U_x(n) = F_x(n) \cdot [1 + \beta L_x(n)], \quad (9)$$

$$Y_x(n) = \begin{cases} 1, & U_x(n) > E_x(n-1) \\ 0, & \text{otherwise} \end{cases}, \quad (10)$$

$$E_x(n) = \hat{E}_x(n) - \Delta T_x, \quad (11)$$

where all the notations have the same meaning as those of (1)–(5), except for the notations ΔT_x and \hat{E}_x in (11). \hat{E}_x denotes the estimation of the global neural threshold. ΔT_x is associated with two factors and is defined as follows:

$$\Delta T_x = C \cdot \exp(-D^2/\gamma). \quad (12)$$

One factor is the distance D , which is closely related to the unfired region Ω_1 and the fired region Ω_2 , and is defined as

$$D = \arg \min_{x \in \Omega_2} \text{dist}(y, x), \quad y \in \Omega_1. \quad (13)$$

The other factor is related to the transition between the regions Ω_1 and Ω_2 and is controlled by the positive value of C . For the parameter γ , whose aim is to scale the distance, it is often set as a fixed value. ΔT_x decreases significantly from a positive number to zero as D increases, so that, ΔT_x can provide a good remedy for minor magnitude variation if the blur boundary occurs and the fired neurons can continue firing once it begins.

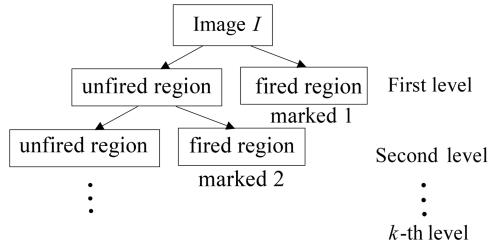


Fig. 2 Extending our model using a decision tree for multilevel image segmentation

3.2 Behaviour of the modified model

The synchronous pulse derived from the linking input L is a special behaviour inherent in PCNN. Once the neighbouring neurons receive a similar stimulus, synchronous pulse emanates from that perimeter of the fired group, capable of forcing neighbouring neurons with similarity to generate a pulse together, i.e. seeded region growing.

Let $Y(n)$ be a pulse output at the n th iteration, where $Y_z(n) = 1(Y_z(n) = 0)$ denotes the neuron z in fired region Ω_2 (unfired region Ω_1). The output of pulsing information will then be delivered to neighbouring neurons at the $n+1$ th iteration denoted as a set X in the following equation:

$$X = \{z | L_z(n+1) > 0\} \cap \{z | Y_z(n) = 0\}. \quad (14)$$

In this group X , each unfired neuron receiving the information from fired neurons will tend to emit the pulse once its internal activity U surpasses the neural threshold designed in our model. In this sense, this behaviour of our model can be viewed as a region growing approach.

3.3 Its application to multilevel image segmentation

As discussed previously, our modified PCNN is capable of grouping the pixels with similarity through iterative computation. Consequently, a given image can be adaptively decomposed into a series of regions by extending our model with the strategy of the decision tree shown in Fig. 2. In this strategy, each level contains a pulse output $Y(n)$ yielded by our model. The unfired region of the previous level is further divided into two regions by our modified model. Although this process is somewhat similar to Chen's strategy [18], the difference is that, in our model, the iteration will continue until the distance between the cluster centres of the unfired and fired region is less than threshold δ at the k th level, which is motivated by the subjective feeling of the human eye less sensitive to the low intensity [4]. Here, we set $\delta = 10$, and then merge the result of the unfired and fired regions at the k th level. Meanwhile, we set the mean value of fired regions for the whole region at the k th level.

In summary, our model can be considered as a region growing method by adjusting fewer parameters which facilitate the adjustment of neural behaviour for image segmentation. The next section will discuss the specification of the parameters in detail.

4 Parameter setting for modified PCNN

In this section, we will present a method of parameter setting for adjusting the behaviour of modified pulse-coupled neurons. Actually, we need to consider a total of five parameters, although some of the parameters are straightforward. Accordingly, the way to set each parameter is presented as follows.

4.1 Weight M and W

Weights M and W are used to connect the neighbouring neurons in the feeding and linking channels, which are responsible for transmitting some information from neighbours. Weight M can be characterised as the following normalised Gaussian function of the centre of the square matrix [9]:

$$M_{x,y} = C_s \exp(-\|x - y\|^2/\sigma_s^2), \quad (15)$$

where y is the spatial position surrounding the centre x of the local window (e.g. 3×3 window), σ_s denotes the scale factor as it determines the change in distance between x and y , and C_s is the normalised constant. In practice, this scheme significantly reduces noise and enables our model to cope with the variation between the gaps to some extent.

Weight W is usually set to

$$W_{x,y} = \begin{cases} C_w \cdot 1/\|x - y\|_2, & x \neq y \\ 0, & x = y \end{cases}, \quad (16)$$

where $\|\cdot\|$ denotes the Euclidean distance and C_w is a normalised positive constant that can be combined with linking coefficient β to derive βC_w , which always acts as a whole factor. As such, we set $C_w = 1$ for simplification.

4.2 Amplitude V_F

V_F can be considered as an additional parameter when compared with some simplified or modified PCNN models [2, 3, 21, 32]. In our study, this value could be deduced from the expression of neural threshold E and external stimulus. As shown in (7), V_F scales the input when the corresponding neuron generates the pulse. In particular, the neighbouring neurons, which are captured using the synchronous pulse, will continue firing in the successive iteration. Therefore, the value of V_F can be expressed as follows:

$$[V_F(n)]_x = \begin{cases} 0, & x \in \Omega_1 \\ E(n-1) + C - F_x, & x \in \Omega_2 \end{cases}. \quad (17)$$

4.3 Neural threshold E

Neural threshold E is essential to control the pulse behaviour of neurons. In our model, this term consists of the adjustable term ΔT_x and \hat{E}_x , which have been described previously and are usually used to fine-tune the neural threshold based on the image properties. \hat{E}_x , in practice, plays a key role in preventing the neurons with dissimilarity from firing. In previous studies, two kinds of methods can be used to determine \hat{E}_x . One method is to fix the value of \hat{E}_x derived from the high intensity in the unfired region. Pixels with similarity are then grouped based on the synchronous pulse via increasing the value of β at each PCNN iteration. This method is similar to the method proposed by Stewart [28]. The other method is to update the \hat{E}_x value as follows:

$$\hat{E}_x(n) = \sum_{x \in \Omega_2} \sum_{y \in N_x} M_{x,y} I_y / \sum_{x \in \Omega_2} 1_x, \quad \Omega_2 = \{x | Y_x(n-1) = 1\}. \quad (18)$$

Pixels with similarity are then clustered together.

In comparison, the adoption of (18) for setting \hat{E}_x has several advantages for region growing as follows: (i) it acts as a cluster centre of the region Ω_2 to force the neurons with similarity intensities to pulse together; (ii) it can significantly reduce the number of iterations.

4.4 Linking coefficient β

Linking coefficient β provides a simple and effective method for controlling the influence of the linking input on the internal activity of the neuron, thus forcing the neighbouring neurons with brightness similarity to pulse together. In other words, linking coefficient β encourages the set X obtained using (14) at the n th iteration to be separated into two groups. One group starts pulsing, which satisfies $U_x(n) > E(n)$, whereas the other does not. In previous studies, the neighbouring neurons are often separated via the hard decision [5] or directly using fuzzy clustering method [32], so it may be affected by the uncertainty and robustness

against noise. Inspired by the fuzzy set theory that classifies the pixels into different groups via a membership graded between 0 and 1, the choice of designing the method proposed in this study is defined as follows:

$$\begin{aligned} \min_{\beta} \quad & f(\beta) = \sum_{x \in X} \mu_B(S_x)(F_x - m_1(n))^2 + \mu_O(S_x)(F_x - m_2(n))^2 \\ \text{s.t.} \quad & 0 < \beta \leq \xi, \quad \mu_B(S_x) + \mu_O(S_x) = 1 \end{aligned}, \quad (19)$$

In the equation, ξ is the upper bound set for β , which should be large enough to encourage the neighbouring neurons with similarity to fire; $\mu_O(\cdot)$ and $\mu_B(\cdot)$ denote the membership of the first and second groups, respectively; and $m_i(n)$ is set as follows:

$$m_i(n) = \sum_{x \in \Omega_i} \sum_{y \in N_x} M_{x,y} I_y / \sum_{x \in \Omega_i} 1_x, \quad i = 1, 2. \quad (20)$$

Here, the value of β is not apparent in the cost function, but is implied in the value of S_x as follows:

$$S_x = \begin{cases} F_x(1 + \beta L_x(n)), & U_x(n) > E_x(n) \\ F_x, & \text{otherwise} \end{cases}, \quad x \in X. \quad (21)$$

For better understanding, here we elaborate on the meaning of cost function f in (19). First, $f(\beta)$ is a sum of fuzzy within-class variance of pixels surrounding the different cluster centres, with $\mu_O(S_x)$ and $\mu_B(S_x)$ as the membership assigned to corresponding neurons in separate groups. Second, considering the membership $\mu_O(S_x)$ and $\mu_B(S_x)$, the contribution of β inherent in S_x may guide the membership bias towards the right side, thus reducing the misclassification error.

In fact, the membership function of fuzzy sets can be defined in any number of ways as long as they follow the rules of the definition of a fuzzy set. In most cases, the decision on which type of membership function to use is dependent on the purpose or application [33–36]. In this study, to partition pixels into the background and object, we use the membership function associated with the distance to the obtained centre of the object (background) as follows:

$$\begin{aligned} \mu_O(g) &= \begin{cases} 0, & g \leq a \\ \frac{1}{(1 + \|g - b\|^2/\|g - a\|^2)}, & a < g \leq b, \\ 1, & g > b \end{cases} \\ \mu_B(g) &= \begin{cases} 1, & g \leq a \\ \frac{1}{(1 + \|g - a\|^2/\|g - b\|^2)}, & a < g \leq b, \\ 0, & g > b \end{cases} \end{aligned} \quad (22)$$

where g is a variable in the intensity domain, a is set as $m_1(n)$, and b is varied based on the value of $m_2(n)$ at the n th iteration. The intensity range $[a \ b]$ is usually considered a fuzzy interval. We then transform a set of pixels in X from the intensity domain into the fuzzy domain, assigning large membership value to the pixel whose intensity is close to the centre.

To fast obtain the value of β , the choice is very rich in optimisation methods such as population-based search methods (genetic algorithm, particle swarm optimisation etc.) and iterative search methods (Newton iterative method, golden section method etc.). Here, we adopt the golden section method frequently used as a rapid method for determining the optimal solution in one variable, and a sufficiently large value $\xi=2$ is set initially at each PCNN iteration.

In summary, the parameter setting of our model is potentially associated with the characteristic of synchronous pulse inherent in the PCNN. The weight M is set by the value of neighbouring neurons. With the use of W , the neighbouring neurons surrounding with fired neurons are linked together, in order to generate the

synchronous pulse by setting linking coefficient β according to the fuzzy clustering method. Besides, amplitude V_F and neural threshold E interact with the characteristic of the synchronous pulse. Through PCNN iteration, our method can group the similarity neurons, and get the multilevel segmentation result. The following is the complete algorithm description, together with the way to set the parameter and the iterative procedures of PCNN (see Fig. 3).

5 Experiments and discussions

In this section, experiments on synthetic and natural images were carried out to assess the effectiveness of the proposed method. The results obtained by the proposed method were then compared with the results yielded by two types of segmentation methods, namely, the PCNN-based methods, such as Stewart's [28], Chen's [18], and the model [9] (called SPCNN¹) and the model [32] (called SPCNN²) which were extended by using the strategy of the decision tree, and the approaches, including graph-based method [37], and normalised cuts [38]. For comparison with the graph-based method, we used the Matlab code downloaded from <http://cn.mathworks.com/matlabcentral/fileexchange/25866-efficient-graph-based-image-segmentation>, which is built on an adjacent neighbourhood model or k -nearest neighbourhood model. In the following experiments, we use the adjacent neighbourhood model in the graph-based method. For the normalised cuts, the code downloaded from http://www.timotheecour.com/software/nCut_multiscale/nCut_multiscale.html is used for image segmentation. We notice that the segmentation results of normalised cuts are somewhat dependent on the choice of the number of segments. Here, we use the number of segments derived from our method for comparison, and set the class mean value for each class.

Additionally, in the experiments, unless specified otherwise, we set $d=1$, $\delta\beta=0.01$, $\delta\mu_{\max}=0.15$, and $\beta_{\max}=1$ for Stewart's method. However, no post-processing step was conducted to merge the regions smaller than 50 pixels in Stewart's method. In our model, we set $\sigma_s=1$, $\delta=10$, $C=20$, and $\gamma=18$ and initialised the pulse output $Y(0)$ in which the neurons receive the highest intensity at each level. All the algorithms were implemented in Matlab 7.10 and performed on a computer with Intel(R) Core(TM) 2 Duo 2.4 GHz i5 CPU 2G RAM, and windows 64bit operation system.

In order to provide the numerical assessment for segmentation performance, we use the modified Dice coefficient for coarse evaluation as follows:

$$D = \frac{1}{k} \sum_{i=1}^k \frac{2|X_i \cap Y_i|}{|X_i| + |Y_i|} \quad (23)$$

where k is the number of classes; X_i denotes the i th class generated by the segmentation method; Y_i denotes the i th true class; $|.|$ is the cardinality of a set. Usually, the higher value of D implies the better quality of the result, while the lower value indicates that the method obtains the worse segmentation result.

5.1 Synthetic image segmentation

In the first experiment on the synthetic image, as shown in Fig. 4a, all labelled regions were characterised by Gaussian distribution with the same standard deviation constant (i.e. $\sigma_i=15$, $i=1, \dots, 4$), but different class means (i.e. $\mu_1=150$, $\mu_2=110$, $\mu_3=70$, $\mu_4=45$), to generate the degree of overlap, as shown in Figs. 4b and c. Thus, this image is expected to be divided into four classes by the use of those segmentation methods. Figs. 4d-j illustrate the results of Stewart's, Chen's, SPCNN¹, SPCNN², graph-based method, normalised cuts, and the proposed model, respectively. We can observe that, Stewart's method separates the whole image into several regions, and misclassifies some parts of the region 3 to the region 2. Chen's method does not obtain accurate results visually because of its coarse strategy for parameters setting. The results of SPCNN¹ and SPCNN² to automatically segment the regions are as

```

1 Input          -- test images;
                -- PCNN parameter  $M/W$  in Eq.(15) and Eq. (16)
                -- the parameter  $C, \gamma$  in Eq. (12) for setting the transition parameter of gray value from the
                  pulse region to un-pulsed region;
                -- the value  $\xi$  for obtaining the  $\beta$  via an optimization procedure in Eq. (19)
                --the parameter  $\delta$  in Eq. for the stop criteria that the iteration is stopped until the distance
                  between the cluster centers of the unfired and fired region is less than this value;
                --the initial PCNN neural threshold  $E$  whose value is the highest intensity of the image



---


2 PCNN iteration Repeat
    Repeat
        Calculate  $F, L, U, Y$  through iteration, as seen in Eqs. (7)-(10)
        Compute the parameter  $\beta$  in Eq. (21), and update neural threshold  $E$  in Eq. (11)
        until the pulse region in  $Y$  does not change any more
        Calculate  $m_1, m_2$  by using  $Y$ 
        Remove the pulsed region and marked in the  $R$ 
        Set the initial neural threshold  $E$ 
        until  $|m_1-m_2|<\delta$ 
        Group the remaining region together, and mark in the  $R$ 


---


3 output:      --the marked region  $R$  as the final segmentation result

```

Fig. 3 Algorithm: our PCNN model for image segmentation

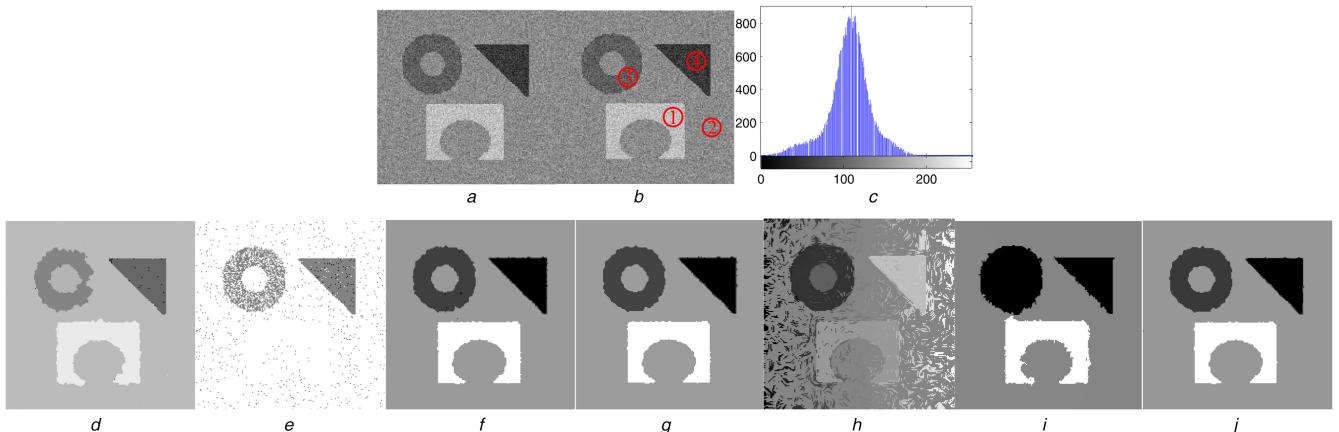


Fig. 4 Segmentation result of the synthetic image

(a) Original image, size of 200×200 , (b) Labelled region, (c) Histogram, (d)–(j) Stewart's, Chen's, SPCNN¹, SPCNN², graph-based method, normalised cuts, and the proposed method, respectively

good as that of the proposed method, as shown in Figs. 4f and g. Graph-based method, however, generates a lot of small regions because of noise in the synthetic image, and normalised cuts method does not segment the part three accurately.

The results in terms of the number of clusters, cluster centres, Dice coefficient, and running time obtained by applying the segmentation methods to segment the synthetic image are shown in Table 1. We observed that all the PCNN-based methods can segment the image into four clusters, except for Stewart's method because it generates a lot of small parts during region growing. In addition, the cluster centres obtained by the proposed model are the closest to the real cluster centres. However, our model requires a bit more time than that of Chen's, but less than that of all other methods. Nevertheless, our model has higher value of Dice coefficient, which demonstrates the higher performance of

multilevel image segmentation than that of those PCNN-based methods, graph-based method, and normalised cuts.

5.2 Berkeley image segmentation

The following experiments were performed on natural images from the Berkeley Segmentation Dataset [39], proving the ability of our method to segment natural images. All the images have a size of 321×481 , as shown in Fig. 5a. In some of these images, the number of clusters is uncertain, as shown in the corresponding histogram in Fig. 5b. Again, in this experiment, the results imply that the proposed model can automatically obtain the acceptable multi-classes by using the strategy of the decision tree described above, as shown in Fig. 5c. The number of classes, respectively, is four, five, four, six, four, five from image 1 to image 6.

Table 1 Performance of each segmentation method for synthetic image

	Cluster number	Cluster means	Dice coefficient	Time, s
Stewart's	—	—	—	18.6179
Chen's	4	(1.0000; 13.6515; 58.7589; 113.7576)	0.0552	0.0252
SPCNN ¹	4	(50.3504; 74.8259; 109.6545; 148.3359)	0.9739	19.0791
SPCNN ²	4	(48.4365; 74.6843; 109.6439; 148.2633)	0.9709	5.4050
graph-based	—	—	—	0.6106
normalised cuts	4	(70.2453; 109.5530; 109.6996; 145.5338)	0.5036	0.7308
proposed	4	(45.3886; 74.7606; 109.6433; 148.2247)	0.9743	0.4312

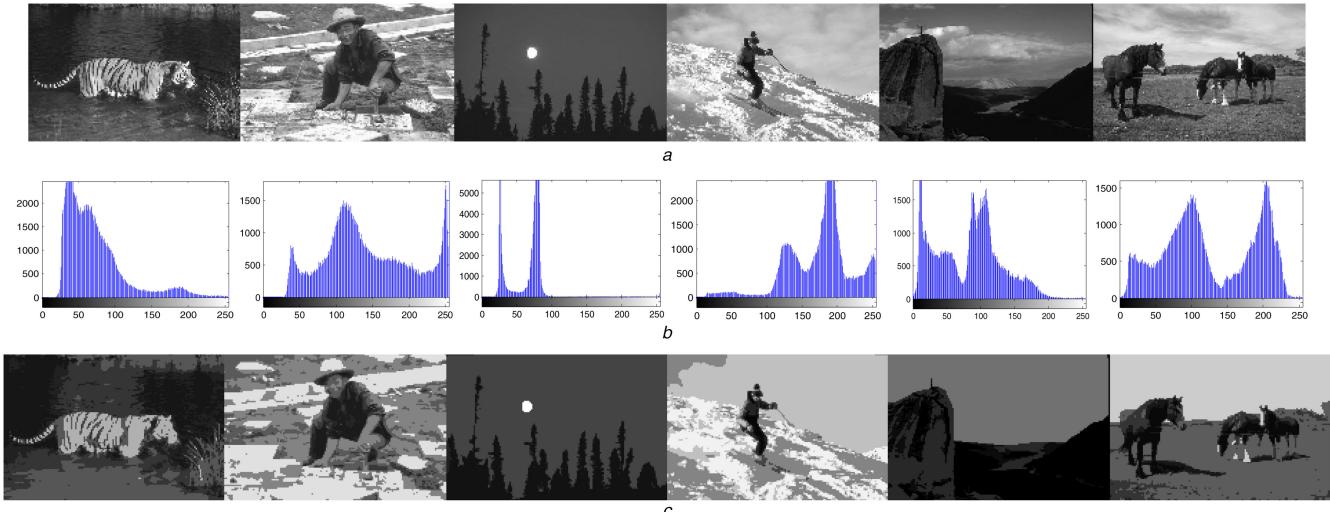


Fig. 5 Experiments on natural images

(a) Natural images, (b) Its histogram, (c) Segmentation results of the proposed model

For comparison, Fig. 6 shows the results yielded by each aforementioned method. Among the PCNN-based methods, results of SPCNN¹ method and SPCNN² are similar to each other, and are more competitive compared with those yielded by the proposed method. However, our method uses neither such fine strategy in SPCNN¹ nor a fuzzy clustering method for keeping the characteristic of synchronous pulse, so that it runs faster than that of SPCNN¹ and SPCNN², as shown in Table 2. Stewart's method seems to separate the image into several more similarity regions under three conditions for region growing. In Chen's method, the results for images 2, 4, and 5 are reasonable, showing a specific ability to divide the image into a set of clusters. However, a coarse strategy for segmentation may result in undersegmentation, as observed in the result of image 1. Additionally, a small portion of the moon is missed in the result of image 3. From the result of graph-based method, we found that the errors of this method are even more obvious because of the adjacent neighbourhood model used in the model. Besides, we notice that normalised cuts seems to coarsely separate the whole image into similarity regions, and shows worse performance under the same number of segments.

5.3 Further discussions

5.3.1 Discussion on the linking coefficient: In our model, the way to set parameter β is more robust against a fixed setting way. Fig. 7 illustrates the segmentation results of synthetic image for example, corresponding to several fixed values of β , i.e. $\beta = 0.01, \beta = 0.05, \beta = 0.10, \beta = 0.3$, and the value designed in Chen's method. Roughly speaking, setting a relatively large value of β makes the neighbouring neuron with low brightness produce the synchronous pulse, resulting in poor performance, as shown in Figs. 7c–e, whereas setting a smaller value of β keeps the characteristic of synchronous pulse, as shown in Figs. 7a and b. This implies that, a fixed value of β in entire procedure of PCNN image segmentation may have more sensitive to intensity distribution of neighbouring neurons. This further demonstrates the efficiency of the way to set the value of β automatically during PCNN iteration.

5.3.2 Discussion on the neural threshold: In the PCNN model, neural threshold often plays a significant role in image segmentation. In our model, the neural threshold consists of two parts, named $\hat{E}_x(n)$ and ΔT_x , respectively. In order to demonstrate the better performance of the setting way than that of SPCNN¹, we set the threshold E as that of SPCNN¹ (named as proposed²) for comparison. Table 3 illustrates the results of Dice coefficient, cluster means, and times, obtained by the proposed method (named as proposed¹) and proposed² method. The result of E_x through iterative computation is shown in Fig. 8a, where $\hat{E}_x(n)$ is in red colour and E_x is in blue colour. Correspondingly, the linking coefficient changes varied with the neural threshold, as shown in Fig. 8b. It can be found that the proposed¹ method takes less number of iterations than that of proposed² method, although the result of proposed² seems to be the same as that of proposed¹, as shown in Fig. 8c. So, we remark that, the part of ΔT_x interacted with the estimation of global neural threshold enables the model to be faster iteration, and to be getting the better results with the use of our setting way for linking coefficient.

5.3.3 Discussion on comparison with SPCNN¹ and SPCNN² for binary image segmentation: Frankly speaking, our model extended by a decision tree can divide the whole image into several regions with similarity. In order to further demonstrate the advantage in terms of parameter setting and the performance of binary image segmentation, we perform the experiments on infrared images for the test, and the results obtained by our method are then compared with some existing PCNN-based methods, such as SPCNN¹ and SPCNN². Fig. 9a illustrates five representative infrared images, size of 240×360 , 240×320 , 246×328 , 240×320 , 240×320 from top to down. The results of SPCNN¹, SPCNN², and our method are shown in Figs. 9c–e, respectively. Our method almost obtains the higher performance of segmentation, with comparison to the foreground manually delineated from the images shown in Fig. 9b. From the comparison

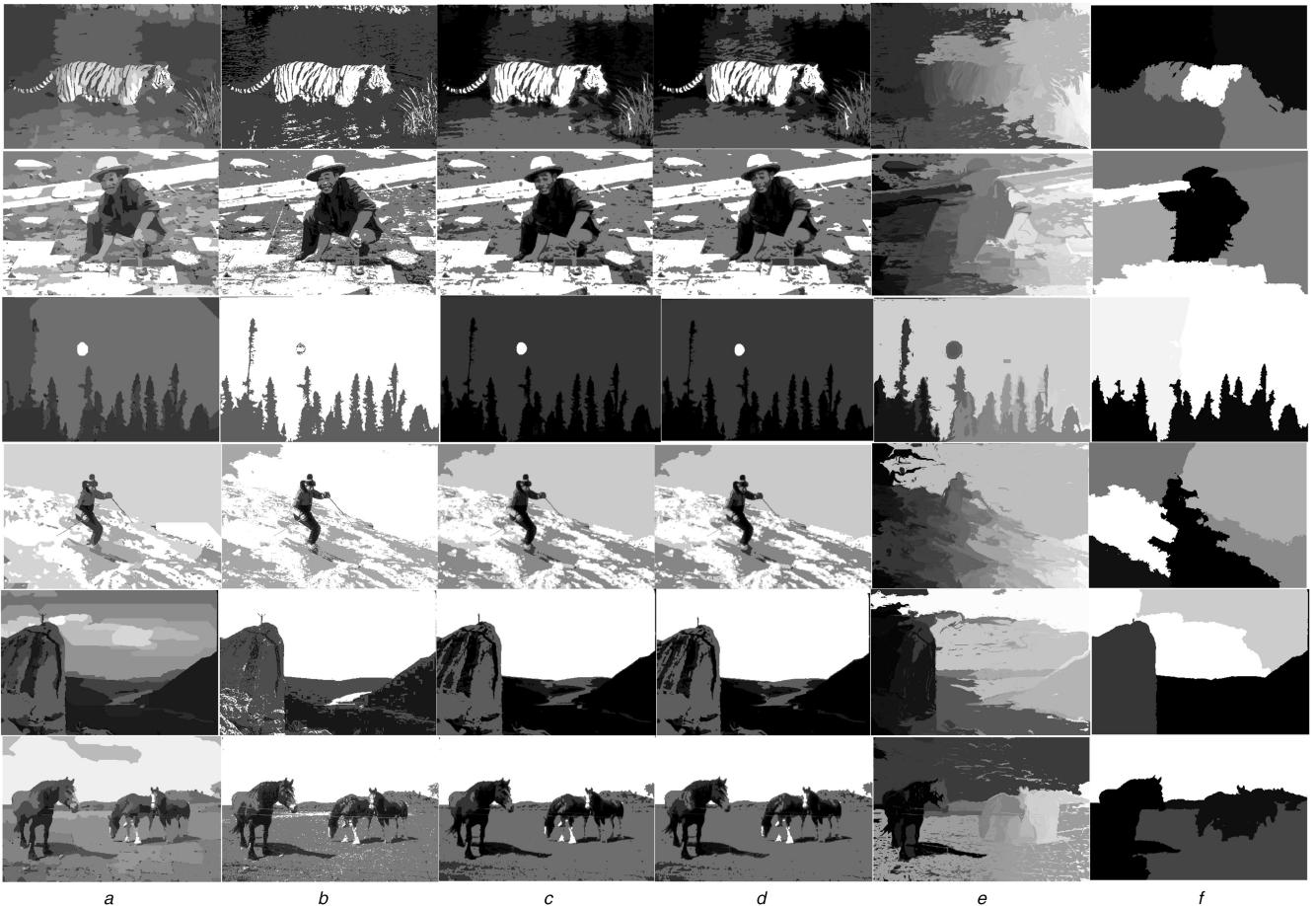


Fig. 6 Segmentation results yielded by each segmentation method for comparison

(a)–(f) Results of Stewart's, Chen's, SPCNN¹, SPCNN², graph-based method, and normalised cuts, respectively

Table 2 CPU time (in second) for the proposed model, SPCNN¹ model, and SPCNN² model

	Image 1	Image 2	Image 3	Image 4	Image 5	Image 6
SPCNN ¹	243.7262	241.3745	101.6403	264.6733	122.2837	171.3781
SPCNN ²	113.8863	56.6031	44.2811	80.8311	29.6803	40.7120
proposed	11.6482	17.3412	5.3624	13.2697	10.7319	15.8346

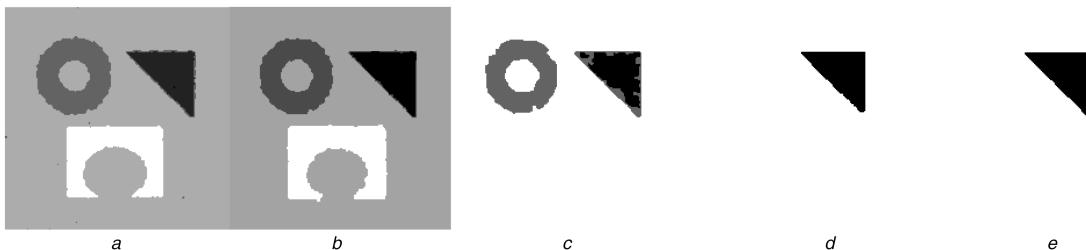


Fig. 7 Segmentation results of our method with different value of β

(a) $\beta = 0.01$, (b) $\beta = 0.05$, (c) $\beta = 0.10$, (d) $\beta = 0.30$, (e) The value designed in the Chen's method

Table 3 Results in terms of Dice coefficient and CPU time (in second)

Method	D	Cluster means	Time, s
proposed ¹	0.9743	(45.3886, 74.7606, 109.6433, 148.2247)	0.4312
proposed ²	0.9688	(50.7840, 75.1856, 109.6532, 148.1412)	0.6968

in terms of Dice coefficient and time consumption, our method shows the higher Dice coefficient and less time consumption than that of SPCNN¹ and SPCNN², as shown in Table 4. This further demonstrates the efficiency of our method in terms of parameters setting.

5.3.4 Discussion on characteristics of our model: PCNN for image segmentation is heavily dependent on its inner characteristic, such as synchronous pulse. The objective of our work is to build a useful segmentation tool in automatic parameter setting way, to ensure the similarity neurons keep synchronous pulse. In contrast to other method, the novelty aspects of our model are summarised as follows:

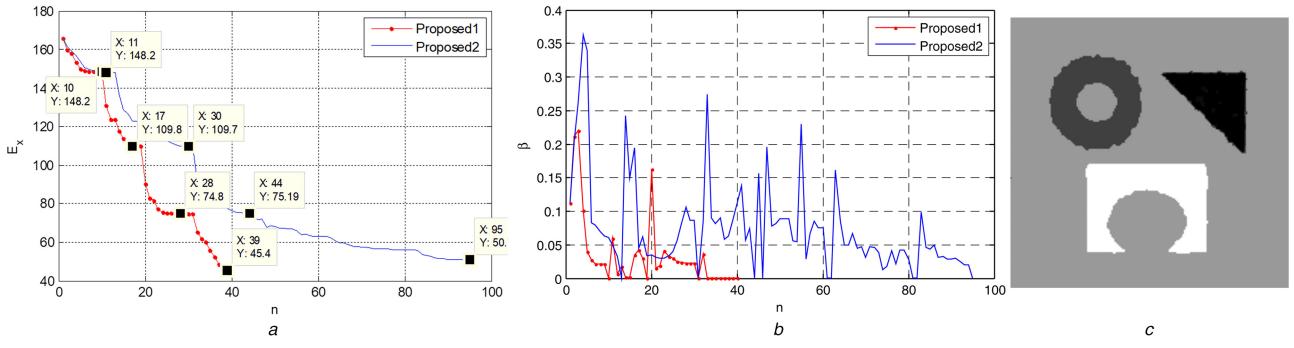


Fig. 8 Results of synthetic image segmentation

(a) Curve of neural threshold, (b) Curve of linking coefficient, (c) Results of proposed² method

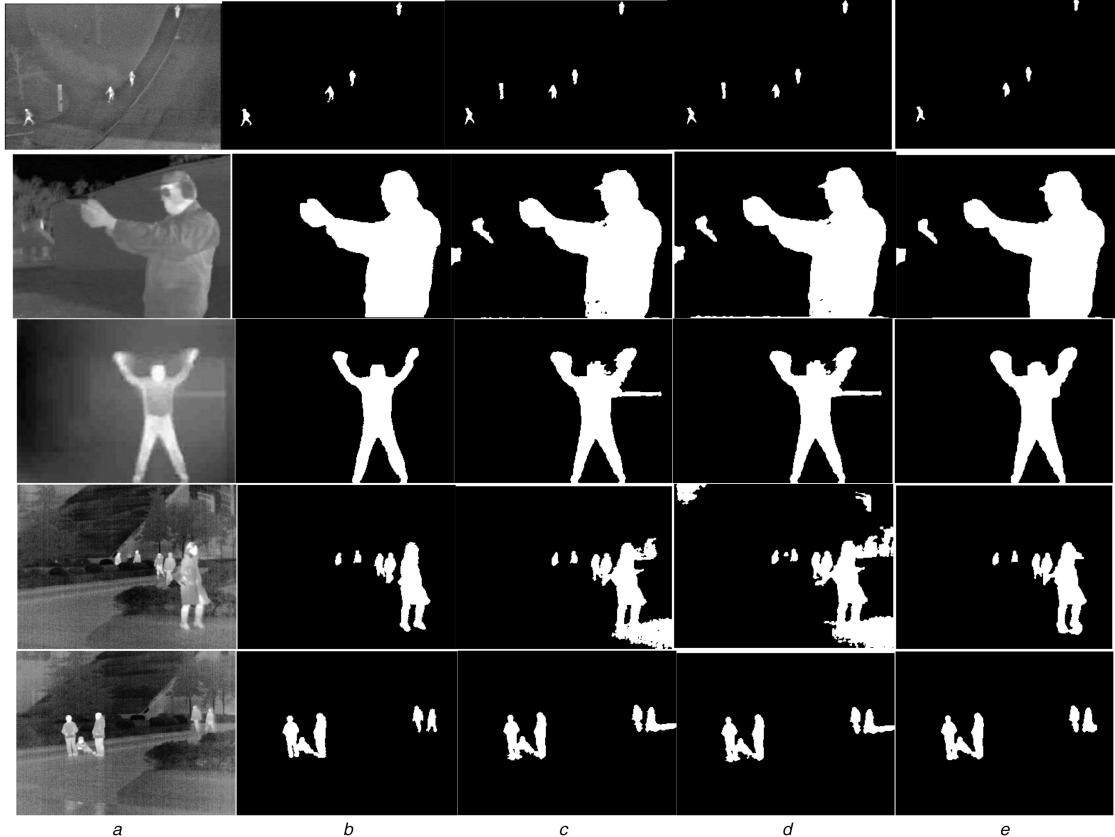


Fig. 9 Experiments on the infrared images

(a) Original infrared image, (b) Foregrounds delineated from the image, (c) Segmentation results of SPCNN¹, (d) Segmentation results of SPCNN², (e) Segmentation results of our proposed model

Table 4 Results in terms of Dice coefficient and CPU time (in second)

Method	Results	Image 1	Image 2	Image 3	Image 4	Image 5
SPCNN ¹	D	0.9104	0.9701	0.9376	0.8466	0.9435
	t	1.1894	10.7637	8.9465	6.0524	1.9625
SPCNN ²	D	0.9077	0.9687	0.9364	0.7762	0.9413
	t	0.6104	19.7149	8.2613	10.4933	2.5398
proposed	D	0.9409	0.9721	0.9424	0.9666	0.9609
	t	0.1152	2.2343	0.4627	0.3214	0.1979

(i) The mechanism of inner parameter inherent in PCNN is described in more detail, and gives the extended model by incorporating a strategy of decision tree for image segmentation.

(ii) The parameter setting of our model is potentially associated with the characteristic of synchronous pulse inherent in the PCNN, and the cost of time in determining the parameters in our work seems less than that of SPCNN¹ and SPCNN².

(iii) We introduce the fuzzy clustering method for obtaining the proper value of β , enabling the model have the characteristic of fuzzy clustering.

6 Conclusions

In this study, we introduced a method for parameter setting of PCNN model based on the image characteristics and behaviour associated with synchronous pulse, and then incorporated a

strategy using decision tree into our model for multilevel image segmentation. In contrast to some existing PCNN-based methods, in our model, the feeding input, neural threshold, and linking coefficient are set by interacting with each other via analysis of the image characteristics and neural behaviour, allowing the model to adaptive region growing. Especially, the adjustable term in our method interacted with the estimation of global neural threshold enables the model to be faster iterated, and to be getting the better results with the use of linking coefficient setting way inspired by the fuzzy set theory. Experiments were finally performed on synthetic and natural images, especially the Berkeley images. The results yielded by the proposed model were better than those yielded by some existing PCNN-based methods, graph-based method, and normalised cuts method. In the near future, we will attempt to the post-processing step using PCNN model, such as image presentation and object detection.

7 Acknowledgments

The authors thank Y. L. Chen and S. Wei for their helpful discussion and suggestion. This work was supported by the Fundamental Research Funds for the Central Universities under grant no. 2042015kf0006, the Fund of Robot Technolgy Used for Special Environment Key Laboratory of Sichuan Province under grant no. 15kftk05, and the Fund of National Natural Science Foundation of China under grant no. 61601382.

8 References

- [1] Echhorn, R., Reitboeck, H.J., Arndt, M., et al.: ‘Feature linking via synchronization among distributed assemblies: simulations of results from cat visual cortex’, *Neural Comput.*, 1990, **2**, (3), pp. 293–307
- [2] Wang, Z., Ma, Y., Cheng, F., et al.: ‘Review of pulse coupled neural networks’, *Image Vis. Comput.*, 2010, **28**, pp. 5–13
- [3] Subashini, M.M., Sahoo, S.K.: ‘Pulse coupled neural networks and its applications’, *Expert Syst. Appl.*, 2014, **41**, pp. 3965–3974
- [4] Lindblad, T., Kinser, J.M.: ‘Image processing using pulse-coupled neural networks’ (Springer, Berlin Heidelberg, 2005, 2nd edn.)
- [5] Gao, C., Zhou, D.G., Guo, Y.C.: ‘Automatic iterative algorithm for image segmentation using a modified pulse-coupled neural network’, *Neurocomputing*, 2013, **119**, pp. 332–338
- [6] Zhou, D.G., Gao, C., Guo, Y.C.: ‘Simplified pulse coupled neural network with adaptive multilevel threshold for infrared human image segmentation’, *J. Comput.-Aided Des. Comput. Graph.*, 2013, **25**, (2), pp. 208–214
- [7] Zhou, D.G., Gao, C., Guo, Y.C.: ‘Automatic segmentation of infrared image using simplified pulse coupled neural network’, *ICIC Express Lett.*, 2012, **6**, (12), pp. 3111–3116
- [8] Kong, X.W., Huang, J., Shi, H.: ‘Infrared image multi-threshold segmentation algorithm based on improved pulse coupled neural networks’, *J. Infrared Millim. Waves*, 2011, **20**, (5), pp. 365–379
- [9] Zhou, D.G., Gao, C., Guo, Y.C.: ‘A coarse-to-fine strategy of iterative segmentation using simplified pulse-coupled neural network’, *Soft Comput.*, 2014, **18**, (3), pp. 557–570
- [10] Ma, Y., Dai, R., Li, L., et al.: ‘Image segmentation of embryonic plant cell using pulse-coupled neural networks’, *Chin. Sci. Bull.*, 2002, **47**, (2), pp. 167–172
- [11] Murugavel, M.: ‘Magnetic resonance image segmentation using pulse coupled neural networks’. Ph. D. dissertation, Mechanical Engineering, Worcester Polytechnic Institute, USA, 2009
- [12] Shi, M.H., Jiang, S.S., Wang, H.R., et al.: ‘A simplified pulse-coupled neural network for adaptive segmentation of fabric defects’, *Mach. Vis. Appl.*, 2009, **20**, (2), pp. 131–138
- [13] Abdelsames, M.M., Gnecco, G., Gaber, M.M.: ‘An efficient self-organizing active contour model for image segmentation’, *Neurocomput.*, 2015, **149**, (Part B), pp. 820–835
- [14] Zhang, S.Q., Wang, W.L., Lu, J.F., et al.: ‘Image segmentation using a new scalar texture descriptor based on extended structure tensor’, *J. Inf. Hiding Multimedia Signal Process.*, 2016, **7**, (5), pp. 1020–1030
- [15] Wu, Y.P., Peng, X.Q., Jiang, Q., et al.: ‘An improved method of flotation froth image segmentation based on watershed transformation’, *J. Inf. Hiding Multimedia Signal Process.*, 2017, **8**, (1), pp. 238–247
- [16] Kuntimad, G., Ranganath, H.S.: ‘Perfect image segmentation using pulse coupled neural networks’, *IEEE Trans. Neural Netw.*, 1999, **10**, (3), pp. 591–598
- [17] Raya, T.H., Bettaiah, V., Ranganath, H.S.: ‘Adaptive pulse coupled neural network parameters for image segmentation’, *World Acad. Sci., Eng. Technol.*, 2011, **73**, pp. 1046–1056
- [18] Chen, Y.L., Park, S.K., Ma, Y.D., et al.: ‘A new automatic parameter setting method of a simplified PCNN for image segmentation’, *IEEE Trans. Neural Netw.*, 2011, **22**, (6), pp. 880–892
- [19] Kun, Z., Hongjuan, Z., Ma, Y.: ‘New spiking cortical model for invariant texture retrieval and image processing’, *IEEE Trans. Neural Netw.*, 2009, **20**, pp. 1980–1986
- [20] Gao, C., Zhou, D.G., Guo, Y.C.: ‘An iterative thresholding segmentation model using a modified pulse coupled neural network’, *Neural Process. Lett.*, 2014, **39**, (1), pp. 81–95
- [21] Wei, S., Hong, Q., Hou, M.S.: ‘Automatic image segmentation based on PCNN with adaptive threshold time constant’, *Neurocomputing*, 2011, **74**, pp. 1485–1491
- [22] Bi, Y.W., Qiu, T.S., Li, X.B., et al.: ‘Automatic image segmentation based on a simplified pulse coupled neural network’, *Lect. Notes Comput. Sci.*, 2004, **3174**, pp. 405–410
- [23] Zhu, S.W., Hao, C.Y.: ‘An approach for fabric defect image segmentation based on the improved conventional PCNN model’, *Acta Electron. Sin.*, 2012, **40**, (3), pp. 611–617
- [24] Peng, Z.M., Jiang, B., Xiao, J., et al.: ‘A novel method of image segmentation based on parallelized firing PCNN’, *Acta Autom. Sin.*, 2008, **34**, (9), pp. 1169–1173
- [25] Gu, X.D., Guo, S.D., Yu, D.H.: ‘A new approach for automated image segmentation based on unit-linking PCNN’. Proc. of Int. Conf. on Machine Learning and Cybernetics, Beijing China, 2002, pp. 175–178
- [26] Karvonen, J.A.: ‘Baltic sea ice SAR segmentation and classification using modified pulse-coupled neural networks’, *IEEE Trans. Geosci. Remote Sens.*, 2004, **42**, (7), pp. 1566–1574
- [27] Li, M., Cai, W., Li, X.Y.: ‘An adaptive image segmentation method based on a modified pulse coupled neural network’, *Lect. Notes Comput. Sci.*, 2006, **4221**, pp. 471–474
- [28] Stewart, R.D., Fermin, I., Opper, M.: ‘Region growing with pulse-coupled neural networks: an alternative to seeded region growing’, *IEEE Trans. Neural Netw.*, 2002, **13**, (6), pp. 1557–1562
- [29] Lu, Y.F., Miao, J., Duan, L.J., et al.: ‘A new approach to image segmentation based on simplified region growing PCNN’, *Appl. Math. Comput.*, 2008, **205**, (2), pp. 807–814
- [30] Yan, C.M., Guo, B.L., Ma, Y.D., et al.: ‘New adaptive algorithm for image segmentation using the dual-level PCNN model’, *J. Optoelectron. Laser*, 2011, **22**, (7), pp. 1102–1106
- [31] Kinser, J.M., Lindblad, T.: ‘Implementation of pulse-coupled neural networks in a CNAPS environment’, *IEEE Trans. Neural Netw.*, 1999, **10**, (3), pp. 584–590
- [32] Zhou, D.G., Zhou, H., Gao, C., et al.: ‘Simplified parameters model of PCNN and its application to image segmentation’, *Pattern Anal. Appl.*, 2016, **19**, (4), pp. 939–951
- [33] Zadeh, L.A.: ‘Calculus of fuzzy restrictions’, in Zadeh, L.A., Fu, K.S., Tanaka, K., Shimura, M. (Eds.): ‘Fuzzy sets and their applications to cognitive and decision processes’ (Academic Press, New York, 1975), pp. 1–39
- [34] Tao, W., Tian, J., Liu, J.: ‘Image segmentation by three-level thresholding based on maximum fuzzy entropy and genetic algorithm’, *Pattern Recognit. Lett.*, 2003, **24**, pp. 3069–3078
- [35] Tizhoosh, H.R.: ‘Fuzzy image processing’ (Springer-Verlag, 1997)
- [36] Tizhoosh, H.R.: ‘Image thresholding using type II fuzzy sets’, *Pattern Recognit.*, 2005, **38**, pp. 2363–2372
- [37] Felzenszwalb Pedro, F., Huttenlocher Daniel, P.: ‘Efficient graph-based image segmentation’, *Int. J. Comput. Vis.*, 2004, **59**, (2), pp. 167–181
- [38] Cour, T., Beneit, F., Shi, J.: ‘Spectral segmentation with multiscale graph decomposition’. IEEE Int. Conf. on Computer Vision and Pattern Recognition, San Diego, California, June 2005, pp. 1124–1131
- [39] Martin, D., Fowlkes, C., Tal, D., et al.: ‘A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics’. Proc. of 8th Int. Conf. Computer Vision, Vancouver, BC, Canada, July 2001, pp. 416–423