An Adaptive Infrared Image Segmentation Method Based on Fusion SPCNN

Zhengkun Guo\textsuperscript{1,2}, Yong Song\textsuperscript{1,2,*}, Yufei Zhao\textsuperscript{1,2}, Xin Yang\textsuperscript{1,2} and Fengning Wang\textsuperscript{1,2}

\textsuperscript{1}School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China
\textsuperscript{2}Beijing Key Laboratory for Precision Optoelectronic Measurement Instrument and Technology, Beijing 100081, China

*Corresponding author. Email: yongsong@bit.edu.cn

Abstract. Inspired by multiple information processing mechanisms of the human nervous system, a fusion simplified pulse coupled neural network (FSPCNN) model for infrared (IR) image segmentation is proposed in this paper. In the method based on FSPCNN, the time decay factor is set adaptively based on Stevens’ power law, and the synaptic weight is generated adaptively based on lateral inhibition (LI), without manual intervention. Meanwhile, according to fast linking mechanism, the similarity between adjacent iteration results is used to implement the automatic selection of optimal segmentation result and control iteration. Experimental results indicate that the proposed method has favorable robustness and segmentation performance.

Keywords: infrared image segmentation, pulse coupled neural network, adaptive parameter setting.

1. Introduction

Pulse coupled neural network (PCNN) is a novel neural network inspired by the phenomenon of neuron synchronization in the $\gamma$ band of the mammalian visual cortex systems. Eckhorn et al. were the first to discover the phenomena in the primary visual cortex of cats and established the Eckhorn neuron model [1]. Johnson then further improved the model and proposed the PCNN [2]. The core mechanism of PCNN is that the fired neurons capture the synchronous emission of neurons in a certain region through spatial proximity and brightness similarity [3]. In the last decade, PCNN has developed rapidly in many aspects of image processing, such as image shadow removal [4], feature extraction [5], image segmentation [6, 7], etc.

Currently, there are two main problems in PCNN applied to image segmentation [2, 3, 7]: (1) the quality of the segmentation result is mainly determined by the appropriate values of the parameters, but the parameters of the PCNN model are numerous and could only be set manually or estimated by extensive experimentation; (2) binary output is generated for each iteration, and it is usually necessary to manually select the optimal one among the plurality of segmentation results. These two aspects limit the further development of PCNN.
Furthermore, the quality of IR image may be poor due to some factors [8], for instance, complex background, low contrast, low signal-to-noise ratio (SNR), etc. Therefore, it is generally difficult to perfectly segment a target from an IR image.

In this paper, based on the neural system information processing mechanism and IR imaging characteristics, an FSPCNN model focusing on IR image segmentation is proposed. The model integrates three mechanisms into the network structure of simplified PCNN (SPCNN), and parameters in the model can be adaptively set while need not training or trials. Experimental results indicate that the proposed method can extract the target from a complex background satisfactorily, which has favorable effectiveness and robustness.

2. SPCNN model
Unlike general artificial neural networks, PCNN consists of a 2-D array of laterally linked pulse coupled neurons and does not require training [2, 3]. The neurons in the network correspond one-to-one with the image pixels. Because of the complex model and numerous parameters of the basic PCNN, many SPCNN models have emerged. Among them, Spiking Cortical Model (SCM) has lower computational complexity and higher accuracy and an SCM-based SPCNN model can be described as follows [3]:

\[
U_{ij}[n] = e^{-\alpha_f}U_{ij}[n-1] + S_{ij}(1 + \beta L_{ij}) + \sum_{k,l} W_{ijkl}Y_{ikl}[n-1] 
\]

(1)

\[
Y_{ij}[n] = \begin{cases} 
1, & \text{if } U_{ij}[n] > E_{ij}[n-1] \\
0, & \text{else}
\end{cases} 
\]

(2)

\[
E_{ij}[n] = e^{-\alpha_e}E_{ij}[n-1] + V_{ij}Y_{ij}[n] 
\]

(3)

The model retains six parameters (they are \(\alpha_f, W_{ijkl}, \beta, V_L, \alpha_e, V_E\)), which maintains the key features of PCNN such as high accuracy, while greatly improves efficiency. So it is chosen in this paper.

3. Fusion SPCNN
Based on the SPCNN model, an FSPCNN for IR image segmentation is proposed in this section. Fig. 1 shows the structure of FSPCNN model.

![Figure 1. The structure of FSPCNN model.](image-url)
### 3.1. Adaptive time decay factor

Stevens’ power law indicates that the perceived intensity is proportional to the involution of the magnitude of a physical stimulus [9]. The expression is as follows:

\[ S = K \times I^n \]  

where \( S \) is the psychological intensity, \( K \) is the constant, \( I \) is the physical intensity, and the index \( n \) varies depend on different objective conditions.

As shown in Fig. 2, it can also be understood as a nonlinear mapping process from spatial domain to visual domain — the grayscales of the high gray region in the spatial domain are compressed in the visual domain, while the grayscales of the low gray region are pulled up. Generally, the pixels of target in IR images are mostly high grayscale, since the grayscales are compressed, in order to avoid the misfire of the neurons corresponding to the pixels of non-target due to the excessive decay of \( E_{ij} \), as far as possible, a smaller time decay factor \( \alpha_e \) is needed to slow down the decay of \( E_{ij} \); On the contrary, the pixels of non-target are mostly in the low grayscale region, and the grayscales of the region are pulled up, so a larger \( \alpha_e \) can be used to improve the segmentation efficiency.

![Figure 2. The relationship between perceived brightness of subjective vision and actual brightness.](image)

After further analysis and a lot of experiments, it can be found that \( \alpha_e \) is negatively and positively correlated with \( T_{Otsu} \) and \( \sigma_b \) respectively. And the algorithm could balance the accuracy of the segmentation with the amount of computation when the three are in the relationship described below:

\[ \alpha_e = \left( \frac{\sigma_b}{T_{Otsu}} \right)^2 \]  

where \( \sigma_b \) denotes the standard deviation of background, and \( T_{Otsu} \) denotes the highest grayscale of background obtained by the Otsu.

### 3.2. Adaptive synaptic weight

LI is one of the basic principles of neural system information processing. When the excitation of a certain neuron accumulates to a certain intensity, the pulse of the other neurons in the same cortical module will be suppressed [10]. Based on LI, we make use of the grayscale and spatial information in IR images to develop a 1-tanh inhibition coefficient model, as shown in Fig. 3.
We set the neighborhood size to $3 \times 3$. The pixel spacing in the neighborhood is 1 or 1.4142, and the grayscale range of the IR image is $[0, 255]$. Considering that the inhibition coefficient should be in the range $[0, 1]$, the magnitude of the independent variable should be reduced, and then the 1-tanh function curve with the domain $[0, 3]$ is defined to simulate its distribution. The expression is as follows:

$$W_{ijk} = 1 - \tanh(I(i, j) \times d_{ijkl} \times 10^{-2})$$  \hspace{1cm} (6)

where $I(i, j)$ denotes the grayscale of the pixel corresponding to the central neuron, and $d_{ijkl}$ denotes the Euclidean distance between the two image neurons.

### 3.3. Adaptive output selection method

We introduce an adaptive output selection method according to fast linking [2], the core of which is to use the image similarity to achieve an automatic output of the optimal segmentation result and automatically stop the iteration. And the Hamming distance is used to characterize the image similarity.

As shown in Fig. 4, according to the fast linking, it can be found that the similarity change of the segmentation results in adjacent iterations is in the following three stages. In the iterative process, the algorithm will continuously search for the maximum value of similarity of the adjacent segmentation results, namely, the minimum value of the Hamming distance. And the segmentation result in the iterative round corresponding to the value will be output as the final result, and the iteration will be stopped.
4. Experiment

In this section, to verify the validity of the proposed method, a set of experiments is carried out on the original IR images mainly from the IEEE OTCBVS benchmark database. And the Otsu method, the FCM method, the FL-SCM model, the GA-PCNN model, and the ACO-PCNN model are chosen to evaluate the segmentation performance of the proposed method.

Some samples from the experimental images are selected to clarify our method and the comparison, the segmentation results are shown in Fig. 5. Column 8 shows the results of the proposed method, and it is evident that it provides a superior performance with more accurate and complete contour, compared to other segmentation methods in the experiment. It is believed that the proposed method is capable of effectively segmenting targets from the IR images with complex backgrounds, low contrasts and low SNRs.

![Figure 4. Schematic diagram of the similarity change of adjacent segmentation results.](image)

**Figure 4.** Schematic diagram of the similarity change of adjacent segmentation results.

![Figure 5. Original images, ground-truths and the segmentation results obtained by the proposed method and the compared methods.](image)

**Figure 5.** Original images, ground-truths and the segmentation results obtained by the proposed method and the compared methods.
5. Conclusion
In this paper, an adaptive IR image segmentation method based on FSPCNN is proposed. We primarily studied the mechanism of Stevens’ power law, LI and fast linking, and based on the characteristics of the IR image, integrated the three mechanisms into SPCNN to form FSPCNN. Briefly, FSPCNN is capable of satisfactorily segmenting the targets while processing the IR images with complex backgrounds, and effectively suppressing the background noise. Accordingly, it can be inferred that the method based on FSPCNN has potential and beneficial applications in flight navigation, autopilot and computer vision, etc.

Acknowledgement
We thank the Institute of Optics of the University of Rochester (Rochester, NY, USA), Jiamiao Yang of Department of Electrical Engineering and Medical Engineering at California Institute of Technology, and Qingyue Song of the High School Attached to Tsinghua University for their help.

References