Simulation of Visual Attention Using Hierarchical Spiking Neural Networks

QingXiang Wu[1,2], T. Martin McGinnity[1], Liam Maguire[1],
Rongtai Cai[2], and Meigui Chen[2]

[1] Intelligent Systems Research Center, School of Computing and Intelligent Systems
University of Ulster at Magee, Londonderry, BT48 7JL, Northern Ireland, UK
{q.wu,tm.mcginnity,lp.maguire}@ulster.ac.uk
[2] School of Physics and OptoElectronics Technology, Fujian Normal University
Fuzhou, 350007, China
{qxwu,rtcai,mgchen}@fjnu.edu.cn

Abstract. Based on the information processing functionalities of spiking neurons, a hierarchical spiking neural network model is proposed to simulate visual attention. The network is constructed with a conductance-based integrate-and-fire neuron model and a set of specific receptive fields in different levels. The simulation algorithm and properties of the network are detailed in this paper. Simulation results show that the network is able to perform visual attention to extract objects based on specific image features. Using extraction of horizontal and vertical lines, a demonstration shows how the network can detect a house in a visual image. Using this visual attention principle, many other objects can be extracted by analogy.

Keywords: Visual attention, spiking neural network, receptive field, visual system.

1 Introduction

The biological brain, with its huge number of neurons, displays powerful functionality in information processing, vision, reasoning, and other intelligent behaviours. Spiking neurons are regarded as essential components in the neural networks in the brain. Neurobiologists have found that various receptive fields exist in the visual cortex and play different roles [1, 2]. Visual Attention enables the visual system to process potentially important objects by selectively increasing the activity of sensory neurons that represent the relevant locations and features of the environment [3]. Given the complexity of the visual environment, the ability to selectively attend to certain locations, while ignoring others, is crucial for reducing the amount of visual information to manageable levels and for optimizing behavioral performance and response times. However, there are relatively few explanations of visual attention using spiking neural networks in the literature. Based on a spiking neuron model, a simulation of visual attention is demonstrated in this paper. The principles and simulation algorithms are presented in detail.

D.-S. Huang et al. (Eds.): ICIC 2011, LNBI 6840, pp. 26–31, 2012.

2 Spiking Neural Network for Simulation of Visual Attention

Biological findings show that the visual system can use feedback signals to highlight the relevant locations[4][5]. In this paper, a spiking neural network, as shown in Fig.1, is proposed to simulate such visual attention. Suppose that a visual image is presented to the retina and an edge firing rate map [6] is obtained as the input for the network. The line detection layer contains two pathways. The horizontal pathway contains a neuron array N_h with the same size as the input neuron array. Each neuron has a receptive field corresponding to a horizontal synapse strength matrix W^h. The vertical pathway contains a neuron array N_v with the same size as input neuron array. Each neuron has a receptive field corresponding to a vertical synapse strength matrix W^v. Therefore, the spike rate map of the neuron arrays N_h and N_v represent horizontal and vertical lines respectively.

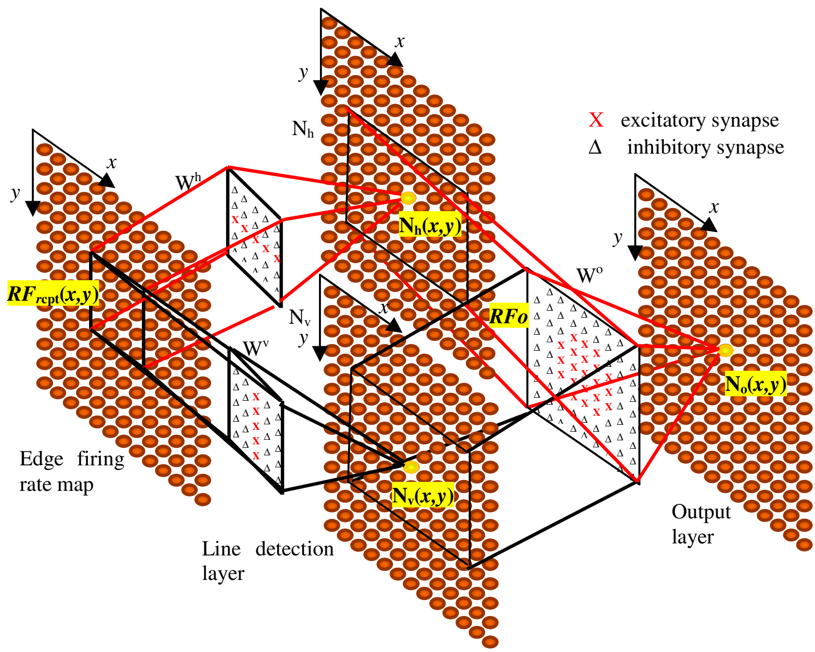

Fig. 1. Spiking Neural Network Model for Attention on Vertical and Horizontal Lines

The output layer is a neuron array with the same size as the edge firing rate map. Each neuron has two receptive fields on the horizontal and vertical neuron arrays respectively. The two receptive fields have the same synapse strength matrix W^o. If there is any horizontal line or vertical line around Neuron $N_o(x,y)$, the neuron will be activated, and a horizontal or vertical lines area can be obtained from the output layer. If these signals are regarded as feedback to an earlier layer of the visual system, objects in the area can be extracted and objects in other areas can be ignored. The simulation algorithms are shown in following sections.

3 Simulation Algorithms of the Spiking Neural Network

In this work a conductance-based integrate-and-fire model has been used to simulate the spiking neural networks since its behaviour is very close to the Hodgkin and Huxley neuron model but yet its computational complexity is much less than that neuron model [7]. Let $S_{x,y}(t)$ represent a spike train that produced by neuron (x,y). If neuron (x,y) fires at time t, $S_{x,y}(t)=1$, otherwise $S_{x,y}(t)=0$. In order to detect edges, the spike train array $S_{x,y}(t)$ can be obtained using the spiking neural network in [6]. Let q^{ex} and q^{ih} represent the peak conductance for excitatory synapse and inhibitory synapse respectively, $g_{x,y}^{hex}(t)$ and $g_{x,y}^{hih}(t)$ represent conductance for excitatory and inhibitory synapses respectively for neuron $N_h(x,y)$. $g_{x,y}^{vex}(t)$ and $g_{x,y}^{vih}(t)$ represent conductance for excitatory and inhibitory synapses respectively for neuron $N_v(x,y)$. In the conductance-based integrate-and-fire neuron model, we have

$$\frac{g_{x,y}^{hex}(t)}{dt} = -\frac{1}{\tau_{ex}} g_{x,y}^{hex}(t) + S_{x,y}(t)q^{ex}, \qquad \frac{g_{x,y}^{hih}(t)}{dt} = -\frac{1}{\tau_{ih}} g_{x,y}^{hih}(t) + S_{x,y}(t)q^{ih}, \tag{1}$$

$$\frac{g_{x,y}^{vex}(t)}{dt} = -\frac{1}{\tau_{ex}} g_{x,y}^{vex}(t) + S_{x,y}(t)q^{ex}, \qquad \frac{g_{x,y}^{vih}(t)}{dt} = -\frac{1}{\tau_{ih}} g_{x,y}^{vih}(t) + S_{x,y}(t)q^{ih}, \tag{2}$$

where τ_{ex} and τ_{ih} are time constants for excitatory and inhibitory synapses, *ex* is short for excitatory, and *ih* for inhibitory. If the neuron generates a spike, the conductance of excitatory and inhibitory synapses increase an amount of q^{ex} and q^{ih} respectively. If the neuron does not generate spike at time t, $g_{x,y}^{hex}(t)$, $g_{x,y}^{hih}(t)$, $g_{x,y}^{vex}(t)$ and $g_{x,y}^{vih}(t)$ decay with time constants τ_{ex} and τ_{ih} respectively. The conductance changes lead to different currents that are injected to neurons $N_h(x, y)$ and $N_v(x,y)$. The membrane potential $v_{h(x,y)}(t)$ and $v_{v(x,y)}(t)$ of neurons $N_h(x, y)$ and $N_v(x,y)$ are governed by the following equations.

$$\begin{aligned} c_m \frac{dv_{h(x,y)}(t)}{dt} &= g_l(E_l - v_{h(x,y)}(t)) + \sum_{(x',y')\in RF_{rcpt}(x,y)} \frac{w_{x',y'}^{hex} g_{x',y'}^{hex}(t)}{A_{ex}}(E_{ex} - v_{h(x,y)}(t)) \\ &+ \sum_{(x',y')\in RF_{rcpt}(x,y)} \frac{w_{x',y'}^{hih} g_{x',y'}^{hih}(t)}{A_{ih}}(E_{ih} - v_{h(x,y)}(t)), \end{aligned} \tag{3}$$

$$\begin{aligned} c_m \frac{dv_{v(x,y)}(t)}{dt} &= g_l(E_l - v_{v(x,y)}(t)) + \sum_{(x',y')\in RF_{rcpt}(x,y)} \frac{w_{x',y'}^{vex} g_{x',y'}^{vex}(t)}{A_{ex}}(E_{ex} - v_{v(x,y)}(t)) \\ &+ \sum_{(x',y')\in RF_{rcpt}(x,y)} \frac{w_{x',y'}^{vih} g_{x',y'}^{vih}(t)}{A_{ih}}(E_{ih} - v_{v(x,y)}(t)), \end{aligned} \tag{4}$$

where E_{ex} and E_{ih} are the reverse potential for excitatory and inhibitory synapses respectively, c_m represents the capacitance of the membrane, g_l represents the conductance of the membrane, A_{ex} is the membrane surface area connected to a excitatory synapse, A_{ih} is the membrane surface area connected to an inhibitory synapse, $w^{hex}_{x',y'}$ represents excitatory synapse strength in matrix W^h, and $w^{hih}_{x',y'}$ represents inhibitory synapse strength in matrix W^h. If the edge in the receptive field matches the pattern in W^h, the neuron $N_h(x, y)$ receives a strong input current through the excitatory synapses and the membrane potential $v_{h(x,y)}(t)$ increases. When the potential reaches a threshold v_{th}, the neuron $N_h(x,y)$ generates a spike and moves into the refractory period for a time τ_{ref}. After the refractory period, the neuron receives inputs and prepares to generate another spike. By analogy, $w^{vex}_{x',y'}$ represents excitatory synapse strength in matrix W^v, and $w^{vih}_{x',y'}$ represents synapse inhibitory strength in matrix W^v. If the edge in the receptive field matches the pattern in W^v, the neuron $N_v(x, y)$ receives a strong input current through the excitatory synapses and generates spikes. The matrix pair of synapse strength distribution for W^h and W^v is determined by the following equations.

$$w^{hex}_{x',y'} = \begin{cases} e^{-(\frac{(x'-x)^2}{\delta_{hx}{}^2}+\frac{(y-y')^2}{\delta_{hy}^2})} & if \sqrt{(x'-x)^2+(y-y')^2} \le R_{RF} , \\ 0 & if \sqrt{(x'x)^2+(y'y)^2} > R_{RF} \end{cases} \tag{5}$$

$$w^{hih}_{x',y'} = \begin{cases} (e^{-(\frac{(x'-x)^2}{\delta_{hx}^2}+\frac{(y'-y-\Delta)^2}{\delta_{hy}^2})} + e^{-(\frac{(x'-x)^2}{\delta_{hx}^2}+\frac{(y'-y+\Delta)^2}{\delta_{hy}^2})}) & if \sqrt{(x'-x)^2+(y'-y)^2} \le R_{RF} \\ 0 & if \sqrt{(x'-x)^2+(y'-y)^2} > R_{RF} \end{cases} \tag{6}$$

$w^{vex}_{x',y'}$ and $w^{vih}_{x,y}$ have the same distribution as $w^{hex}_{x',y'}$ and $w^{hih}_{x,y}$ replacing δ_{hx} and δ_{hy} with δ_{vx} and δ_{vy}, where (x ,y) is the centre of receptive field, (x',y') is a neuron position within the receptive field, R_{RF} is a radius of the receptive field, for horizontal line detection $\delta_{hx}>>\delta_{hy}$ and for vertical line detection $\delta_{vx}<<\delta_{vy}$. The output spike train from the horizontal neuron array is represented by $S_{h(x,y)}(t)$ and $S_{v(x,y)}(t)$. The neuron $N_o(x,y)$ in output layer receives spike trains from both the horizontal and vertical pathways, and the behaviours are governed by the following equations.

$$\frac{g^{oex}_{x,y}(t)}{dt} = -\frac{1}{\tau_{ex}} g^{oex}_{x,y}(t) + \sum_{(x',y')\in RF_o} w^{oex}_{x',y'} S_{h(x',y')}(t) q^{oex} + \sum_{(x',y')\in RF_o} w^{oex}_{x',y'} S_{v(x',y')}(t) q^{vex},$$

$$\frac{g^{oih}_{x,y}(t)}{dt} = -\frac{1}{\tau_{ih}} g^{oih}_{x,y}(t) + \sum_{(x',y')\in RF_o} w^{oih}_{x',y'} S_{h(x',y')}(t) q^{hih} + \sum_{(x',y')\in RF_o} w^{oih}_{x',y'} S_{v(x',y')}(t) q^{vih},$$

$$c_m \frac{dv_{o(x,y)}(t)}{dt} = g_l(E_l - v_{h(x,y)}(t)) + \frac{g^{oex}_{x,y}(t)}{A_{ex}}(E_{ex} - v_{o(x,y)}(t)) + \frac{g^{oih}_{x,y}(t)}{A_{ih}}(E_{ih} - v_{o(x,y)}(t)), \tag{7}$$

where the synapse strength matrix is determined by the following equations. If $(x'-x)^2+(y-y')^2 \le R^2_{RFo}$, $w^{oex}_{x',y'} = \exp(-((x'-x)^2+(y-y')^2)/\delta^2)$ and $w^{oih}_{x',y'} = \exp(-(root((x'-x)^2+(y-y')^2)-\lambda)^2/\delta^2)$, otherwise $w^{oex}_{x',y'} =$ 0 and

$w^{oih}_{x',y'} = 0$. Let $S_{o(x,y)}(t)$ represent a spike train generated by Neuron $N_o(x,y)$ in the output layer. The firing rate for neuron $N_o(x,y)$ is calculated by the following expression.

$$r_{o(x,y)} = \frac{1}{T}\sum_{t-T}^{t} S_{o(x,y)}(t) \tag{8}$$

By plotting this firing rate as an image with a colour bar, areas of horizontal and vertical lines are obtained. Suppose that pixel values of the input image are represented by $I(x,y)$. Suppose that the output signals are regarded as feedback to filter out the attention area as follows.

$T(x,y)=I(x,y)r_{o(x,y)}/r_{max}$ $\quad$ (r_{max} =max($r_{o(x,y)}$ |$(x,y)\in$ all pixels in the image)) (9)
The attention area image is obtained by a plot of image $T(x,y)$.

4 Simulation Results

The network model was implemented in Matlab using a set of parameters for the network: v_{th} = -60 mv. v_{reset} = -70 mv. E_{ex}= 0 mv. E_{ih}= -75 mv. E_l= -70 mv. g_l =1.0 μs/mm^2. c_m=10 nF/mm^2. τ_{ex}=4 ms. τ_{ih}=10 ms. A_{ih}=0.028953 mm^2. A_{ex}=0.014103 mm^2. τ_{ref} = 6 ms for neurons in line detection layer. τ_{ref} = 25 ms for output layer. A_{ih}, A_{ex}, q^{ex} and q^{ih} can be adjudged from a large range for example [0.0001,1]. These parameters are consistent with biological neurons [8]. Parameters for the receptive fields are set as follows. δ_{hx} =20. δ_{hy} = 2. δ_{vx} = 2. δ_{vy} = 20. Δ=2. R_{RF} =15. δ = 30. R_{RFo} =40. λ=24.

The proposed spiking neural network is combined with the edge detection network model in [6]. If a visual image as shown in Fig.2 (A) presents to the edge detection network, edges of the image are reflected in the output neuron array in [6]. The spike trains from this neuron array are regarded as inputs for the proposed network in this paper. In our simulation, the firing rate map is recorded as shown in Fig.2 (B).

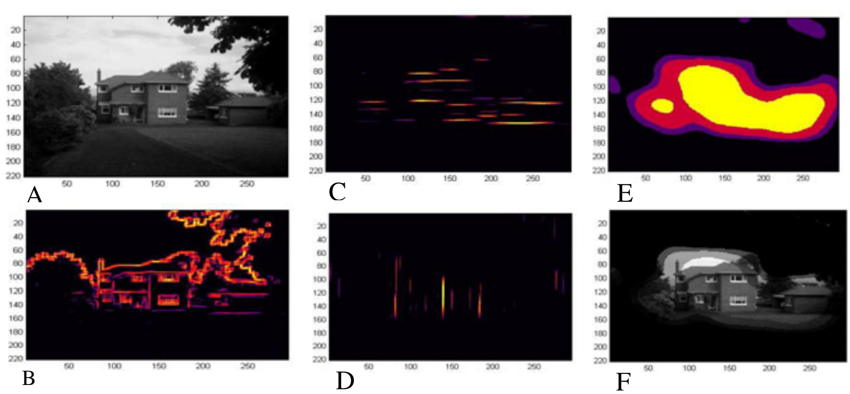

Fig. 2. Simulation results for extraction of attention area

The spike trains are transferred to horizontal and vertical line pathways according to the network architecture in Fig.1, the firing rates of the horizontal and vertical line neuron arrays are recorded as in Fig.2 (C) and (D). The spike trains from horizontal and vertical line neuron arrays are transferred to the output layer. The firing rates of the output neuron array are shown in Fig.2 (E). The output firing rates can be regarded as feedback signals to obtain the attention areas that are based on the key features of horizontal and vertical lines. The image corresponding to the attention area is obtained in Fig. 2 (F). It can be seen that the region around the house has been strengthened and other areas are ignored.

5 Discussion

Various receptive fields and hierarchical structures of spiking neurons enable a spiking neural network to perform very complicated computation tasks, learning tasks and intelligent behaviours in the human brain. This paper proposed a spiking neural network model that can perform visual attention based on image features of horizontal and vertical lines. Simulations show that the proposed network is able to obtain attention area based on the inherent horizontal and vertical lines, for example a house. Based on this principle, other more complicated image features can also be used in this attention mechanism to focus on more complicated objects. We will address this topic in further study.

References

1. Hosoya, T., Baccus, S.A., Meister, M.: Dynamic predictive coding by the retina. Nature 436, 71–77 (2005)
2. Kandel, E.R., Shwartz, J.H.: Principles of neural science. Edward Amold (Publishers) Ltd., London (1981)
3. Anderson, J.R.: Cognitive psychology and its implications, 6th edn. Worth Publishers, Belmont (2004)
4. Saalmann, Y.B., Pigarev, I.N., Vidyasagar, T.R.: Neural Mechanisms of Visual Attention: How Top-Down Feedback Highlights Relevant Locations. Science 316(5831), 1612–1615 (2007)
5. Lauritzen, T.Z., D'Esposito, M., Heeger, D.J., Silver, M.A.: Top-down flow of visual spatial attention signals from parietal to occipital cortex. Journal of Vision 9(13), 18, 1–14 (2009)
6. Wu, Q., McGinnity, M., Maguire, L.P., Belatreche, A., Glackin, B.: Edge Detection Based on Spiking Neural Network Model. In: Huang, D.-S., Heutte, L., Loog, M. (eds.) ICIC 2007. LNCS (LNAI), vol. 4682, pp. 26–34. Springer, Heidelberg (2007)
7. Wu, Q.X., McGinnity, T.M., Maguire, L.P., Belatreche, A., Glackin, B.: Adaptive co-ordinate transformation based on a spike timing-dependent plasticity learning paradigm. In: Wang, L., Chen, K., S. Ong, Y. (eds.) ICNC 2005. LNCS, vol. 3610, pp. 420–428. Springer, Heidelberg (2005)
8. Gerstner, W., Kistler, W.: Spiking Neuron Models: Single Neurons, pulations, Plasticity. Cambridge University Press, Cambridge (2002)