Motion Detection Model Based on PCNN

Minoru YOSHIDA[1] Masaru TANAKA[1] Takio KURITA[2]
[1]Department of Information & Computer Sciences
Saitama university
Simo-ookubo255, Saitama-shi, Saitama, 338-8570 Japan
Tel. +81-48-858-3957, Fax.: +81-48-858-3716
E-mail: {yoshida, mtanaka}@mi.ics.saitama-u.ac.jp
[2]National Institute of Advanced Industrial Sciende and Technology
Umezono1-1-1,Tukuba-shi,Ibaraki, 305-8568 Japan
E-mail: takio-kurita@aist.go.jp

Abstract: Pulse-Coupled Neural Network (PCNN), which can explain the synchronous burst of neurons in a cat visual cortex, is a fundamental model for the biomimetic vision. The PCNN is a kind of pulse coded neural network models. In order to get deep understanding of the visual information processing, it is important to simulate the visual system through such biologically plausibie neural network model.

In this paper, we construct the motion detection model based on the PCNN with the receptive field models of neurons in the lateral geniculate nucleus and the primary visual cortex. Then it is shown that this motion detection model can detect the movements and the direction of motion effectively.

1. Introduction

Vision system, which analyzes and/or recognizes a image or a scene, is one of the most important fields for its usefulness and our curiosity. However, the present vision system dose not reach the level of human vision system yet at the following points such as robustness, and adaptability. Therefore constructing a vision system resembling the mammalian vision system, which demonstrates highly sophisticated capability based on a different framework from the present vision system, is one of the possible approaches in order to develop such an ideal vision system. The vision research field with such an approach is especially called a biomimetic vision.

Pulse-Coupled Neural Network (PCNN), which can explain the synchronous burst of neurons in a cat visual cortex, is a fundamental model for the biomimetic vision.[2] The PCNN is a kind of pulse coded neural network models. In order to get deep understanding of the visual information processing, it is important to simulate the visual system through such biologically plausible neural network model.

In this paper, we construct the motion detection model based on the PCNN, which resembles the visual information processing streams in the human visual cortex. This motion detection model consists of two extended PCNNs, which have the spatiotemporal structure in the receptive fields (RF) of the PCNN neurons. One is based on the RF model of neurons in the lateral geniculate nuleus (LGN), which is used to detect the movements of an object. The other is based on the RF model of neurons in the primary visual cortex (V1), which recieves the outputs of the movement detection PCNN as an input sginal and detects the direction of motion of the object. It is shown through some experiments that this motion detection model can detect the movements and the direction of motion effectively.

2. Pulse-Coupled Neural Network

The PCNN was originally presented to explain the synchronous burst of the neurons in the cat visual cortex by Eckhorn. Then Johnson and his colleagues have implemented the PCNN algorithms for the image processing changing the original PCNN into more suitable form for a computer processing and made pioneering works with a little bit modified PCNN. From now on, we call the PCNN for Johnson and his colleagues's modified PCNN. The PCNN receives the input stimuli such as an image through both feeding and linking connections. The PCNN has a lot of good properties for image processing, pattern recognition, etc.[1][2]

The PCNN neuron consists of three parts : the dendritic tree, the linking modulation, and the pulse generator. The dendritic tree part is further divided into two principal branches in order to make two distinct inputs together into the linking part of the neuron jk. One branch is for the feeding inputs (equation (1)) and the other for the linking inputs (equation (2)).

$$L_{jk}[i] = L_{jk}[i-1]\cdot e^{-\alpha_L} + V_L\cdot(K*Y[i-1])_{jk}, \tag{1}$$

$$F_{jk}[i] = S_{jk} + F_{jk}[i-1]\cdot e^{-\alpha_F} + V_F\cdot(M*Y[i-1])_{jk}, \tag{2}$$

where K and M are the connection weight matrices, $*$ means the convolution operation and the elements of Y contain the information whether a neuron jk has fired or not. α_L and α_F are the time decay constants. V_L and V_F are the linking and feeding constants, respectively. S_{jk} is the external stimulus for neuron jk.

For the linking modulation part, the feeding input F_{jk} and the linking input L_{jk} are got together in order to produce the internal activity U_{jk} for a neuron jk. The internal activity is given by

$$U_{jk}[i] = F_{jk}[i]\cdot(1+\beta L_{jk}[i]), \tag{3}$$

where β is the linking modulation constant.

The pulse generator of the neuron generates the binary output through a step function of the internal activity and changes the threshold value dynamically depending on the status of the neuron jk, fire or not. At time i, the output of the pulse generator Y_{jk} is set to 1 when the internal activity U_{jk} is greater than the threshold function Θ_{jk} and otherwise set to 0. If the neuron jk is fired, then the threshold function is charged by V_Θ (equations.(4), (5)).

$$Y_{jk}[i] = \begin{cases} 1 & \text{if } U_{jk}[i] > \Theta_{jk}[i] \\ 0 & \text{otherwise} \end{cases}, \quad (4)$$

$$\Theta_{jk}[i] = \Theta_{jk}[i-1] \cdot e^{-\alpha_\theta} + V_\Theta \cdot Y_{jk}[i-1], \quad (5)$$

where α_θ is the time decay constant and V_θ is the threshold constant.

3. Extension to PCNN

In the mammalian visual system, the visual information is mainly processed sequentially along the pathway form the retina through the LGN to the V1. Along this pathway, the visual information is decomposed into various forms such as the edges, fundamental shapes, colors and so on. Then, to construct the motion detection model based on the PCNN, we resemble the functions of the movement detection in the LGN and the detection of motion direction in the V1.

To resemble the functions of the LGN and the V1, the structures of the RFs of the neurons are considered. This is because the RF plays a central role in the conceptual framework to study the function of visually responsive neurons and it characterizes the transformation between the visual image and a neuronal activity. Although old traditional textbooks define the RF in spatialy static, the RF is inherently a function of both space and time.[3] Therefore, the RF for the external stimulus of a PCNN neuron is extended to the RF containing both spatial and temporal structures. More precisely, the external stimulus, S_{jk} (equation 2), is extended to the external stimulus at time i, $S'_{jk}[i]$, that is

$$S'_{jk}[i] = \sum_{t=0} (W[t] * I[i-t])_{jk}, \quad (6)$$

where I are intensities of the input sequential images around a neuron jk and W is a connection weight matrix which represents the arrangements of the excitatory or inhibitory region in the RF. A connection weight matrix W changes itself depending on the time, that is, the arrangements of the excitatory or inhibitory region in the RF changes depending on the time. In this paper, with adopting the profiles of the RFs of neurons in the LGN and the V1 for the connection weight matrices Ws, we construct the PCNN detecting the movements and direction of motion.

4. PCNN for the Movements Detection

Based on the spatiotemporal RF profile of the neuron in the LGN, the PCNN for the movements detection is constructed[3]. It is shown in Figure 1 that what profile the RF W for this PCNN has. In Figure 1, the excitatory region is expressed by the swelled shape and the inhibitory region corresponds to the depressed shape. At time t, the RF W is a positive Gaussian distribution, and at time $t-1$, it has the reversed shape of W at time t. This RF profile can respond to the change of intensity value caused by movements of stimuli in the RF. With this RF profiles for the connection weight matrix W (equation 6), the PCNN for the movements detection is constructed.

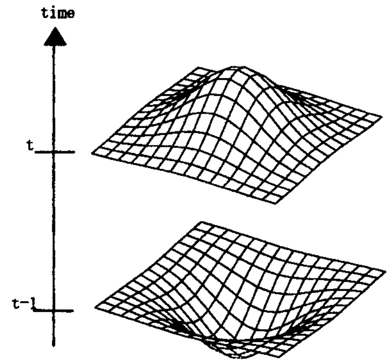

Figure 1. The spatiotemporal RF profile of the neuron of PCNN for the movements detection, based on the RF profile of the neuron in the LGN.

5. PCNN for the Detection of Motion Direction

In order to construct the PCNN for the detection of motion direction, the RF profile is adopted, which is based on the RF model of the neuron in the V1[3]. It is shown in Figure 2 that what profile the RF W for this PCNN has. In this case, if some leftward stimuli are given, that is, the stimuli are in the right hand side of the RF at time $t-1$ and move to the left hand side of the RF at time t, then the internal activity of the neuron has been very charged because the stimuli are in the excitatory region at each time.

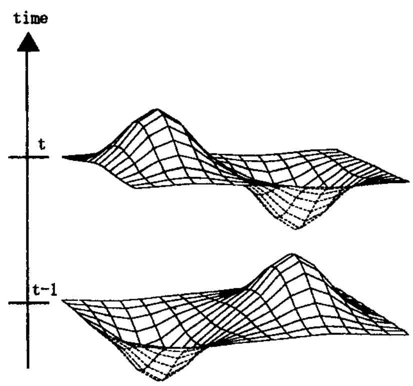

Figure 2. The spatiotemporal RF profile of the neuron of PCNN for the detection of motion detection, based on the RF profile of the neuron in the V1.

6. Motion Detection Model

Here we construct the motion detection model. This model consists of the PCNN for the movements detection and the PCNNs detecting four directions of motion, which are the PCNN detecting upward movement, the PCNN detecting downward movement, the PCNN detecting leftward movement, and the PCNN detecting rightward movement. First, the input sequential images are given to the PCNN for the movements detection, then the outputs from the PCNN for the movements detection are passed each PCNN detecting the direction of motion. Each PCNN detecting the direction of motion responds to the outputs from the PCNN for the movements detection, if the given inputs (movements) are preferable for them.

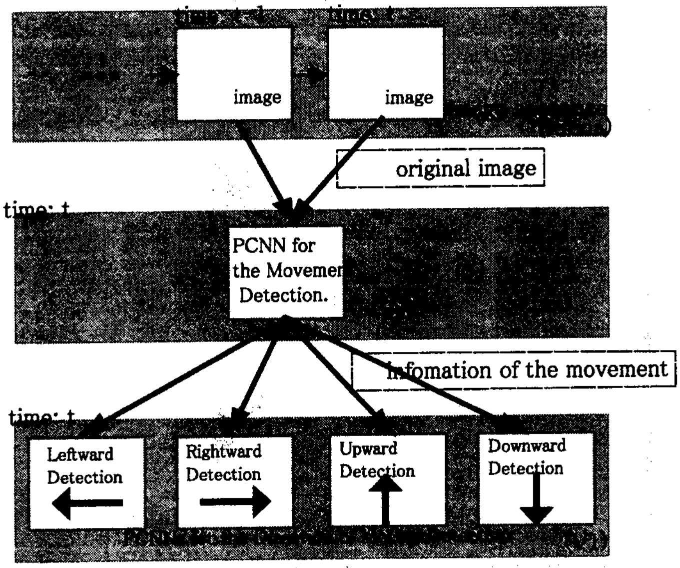

Figure 3. The Motion Detection Model

7. Experiments

Now we show some experiments. The input sequential images are obtained by the video camera. The results are shown in Figure 4–6. In each figure, the column (a) has the input sequential images from the top (time t) to the bottom (time $t+2$) in time order, the column (b) has the outputs of the PCNN for the movements detection in time order, the columns (c)–(f) show the output of the PCNN detecting leftward movement, the PCNN detecting rightward movement, the PCNN detecting upward movement, and the PCNN detecting downward movement, respectively.

In Figure 4, the original sequential images have two persons, one is moving rightward and the other is moving leftward. In Figure 4 (b), the outputs from the PCNN for the movements detection are shown. The thick edges of two persons show that both two persons are moving. In Figure 4(c), the PCNN detecting leftward movement responds to the person in the right hand side who is walking leftward and it does't respond to the person in the left hand side walking rightward. Similarly, Figure (d) shows responses to the rightward movements of the person in the left hand side. In this case, because there are almost no movements in the vertical direction, the conspicuous reaction can't be seen in Figure 4(e) and (f).

Figure 5 shows the result of detecting a motion in the depth direction. The PCNN for the movements detection responds to the movements in the depth direction. In Figure 5(b), the thick edges of the floppy disk show that the floppy disk is moving. From these responses, it is seen that each PCNN detecting the direction of motion responds to the only preferable direction of motion. Combining all outputs from the PCNNs detecting the direction of motion, this motion detection model can detect the motion in the depth direction.

In Figure 6, it is shown that the result of motion detection for the noisy images. Although the PCNN for the movements detection responds to both the actual movements and the noises, however, at the stage detecting the direction of motion, the random gaussian noises give little influence to the results. This is because the PCNN has the feeding and linking modules in itself and the RFs work like smooth filters.

8. Conclusion

We construct the motion detection model based on the PCNN with the RFs like ones of neurons in the LGN and the V1. This model can detect the movements and the direction of motion effectively. It requires no learning, no preprocess and has robustness for the noise.

To get the optical flows in more biologically plausible manner, the model we have presented here should be extended further. For instance, the extended model should detect more accurate direction not only four directions and the relative velocity of motion quantitatively.

This study was performed through Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

References

[1] Thomas Lindblad, Jason M.Kinser, "Image Processing using Pulse-Coupled Neural Networks", *Springer*, 1998.

[2] R.Eckhorn, H.J.Reitboech, M.Arndt, P.Dicke, "Feature Linking via Synchronization among Distributed Assemblies: Simulations of Results from Cat Visual Corlez", *Neural Comp*, vol.2, pp293-307, 1990.

[3] Gregory C. DeAngelis, Izumi Ohzawa, Ralph D. Freeman, "Receptive-field dynamics in the central visual pathways", ıTINS, vol.18,No.10, pp451-458, 1995.

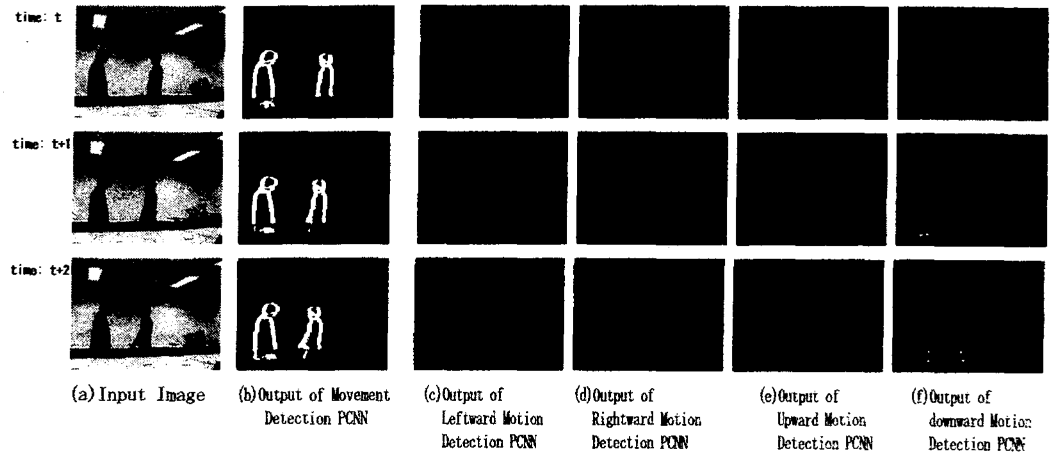

Figure 4. The result of motion detection for moving people.

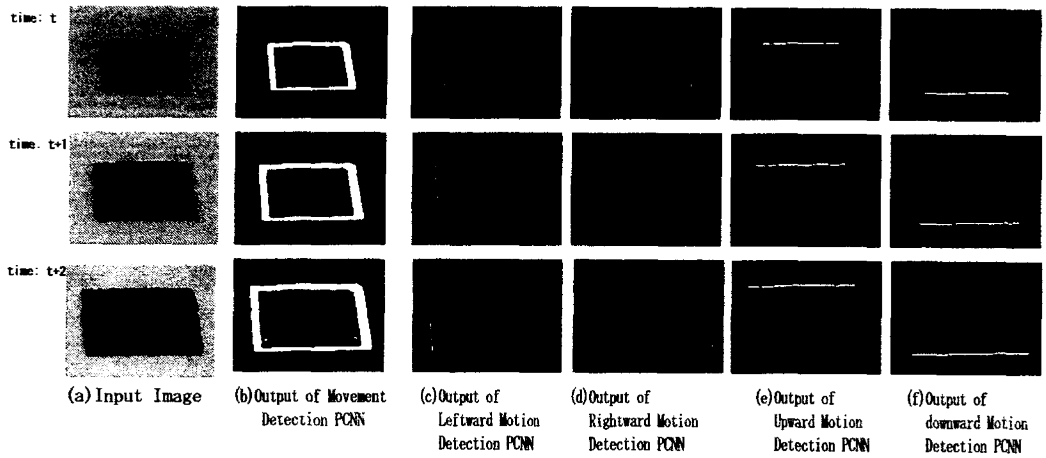

Figure 5. The result of detecting a motion in the depth direction.

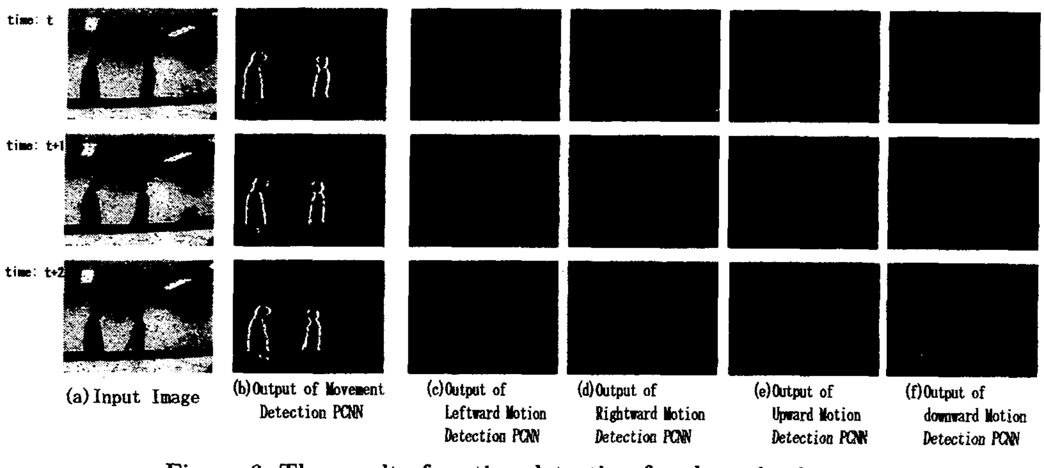

Figure 6. The result of motion detection for the noisy image.