Neural Networks 21 (2008) 1197–1204

Contents lists available at ScienceDirect

Neural Networks

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

2008 Special Issue

Neuromorphic VLSI vision system for real-time texture segregation

Kazuhiro Shimonomura [a], Tetsuya Yagi [b,*]

[a] The Center for Advanced Medical Engineering and Informatics, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

[b] Department of Electrical, Electronic and Information Engineering, Graduate School of Engineering, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Article history:
Received 2 October 2007
Received in revised form
14 July 2008
Accepted 28 July 2008

Keywords:
Analog VLSI
Neural image
Texture segregation
Primary visual cortex
Non-classical receptive field

ABSTRACT

The visual system of the brain can perceive an external scene in real-time with extremely low power dissipation, although the response speed of an individual neuron is considerably lower than that of semiconductor devices. The neurons in the visual pathway generate their receptive fields using a parallel and hierarchical architecture. This architecture of the visual cortex is interesting and important for designing a novel perception system from an engineering perspective. The aim of this study is to develop a vision system hardware, which is designed inspired by a hierarchical visual processing in V1, for real time texture segregation. The system consists of a silicon retina, orientation chip, and field programmable gate array (FPGA) circuit. The silicon retina emulates the neural circuits of the vertebrate retina and exhibits a Laplacian–Gaussian-like receptive field. The orientation chip selectively aggregates multiple pixels of the silicon retina in order to produce Gabor-like receptive fields that are tuned to various orientations by mimicking the feed-forward model proposed by Hubel and Wiesel. The FPGA circuit receives the output of the orientation chip and computes the responses of the complex cells. Using this system, the neural images of simple cells were computed in real-time for various orientations and spatial frequencies. Using the orientation-selective outputs obtained from the multichip system, a real-time texture segregation was conducted based on a computational model inspired by psychophysics and neurophysiology. The texture image was filtered by the two orthogonally oriented receptive fields of the multi-chip system and the filtered images were combined to segregate the area of different texture orientation with the aid of FPGA. The present system is also useful for the investigation of the functions of the higher-order cells that can be obtained by combining the simple and complex cells.

1. Introduction

The visual system of the brain can perceive an external scene in real-time, although the response speed of an individual neuron is considerably lower than that of a semiconductor device. In the visual pathway of mammals, an external image is first received by the retina, and then transmitted to the primary visual cortex (V1). Each neuron in this pathway produces a receptive field that filters the input image. The spatial receptive field structure of the bipolar and ganglion cells in the retina resembles a Laplacian–Gaussian function (Dowling, 1987, for review), and that of a V1 simple cell resembles a Gabor function (Jones & Palmer, 1987). These visual neurons produce the abovementioned receptive fields using a parallel and hierarchical architecture (Nicholls, Martin, Wallace, & Fuchs, 2001, for review). This architecture of the visual cortex is of great interest and is important for designing a novel perception system from an engineering perspective. We have developed an analog very large scale integrated (VLSI) microchip that emulates the structural and functional properties of the retinal and V1 simple cells (Kameda & Yagi, 2003; Shimonomura & Yagi, 2005a). These VLSI microchips can be used as a front-end for a vision system that is constructed hierarchically by mimicking the architecture of the visual cortex. Moreover, they aid in reducing the computational burden of subsequent processing devices.

The aim of this study is to develop a real-time texture segregation system that is based on a hierarchical computational model inspired by psychophysics and neurophysiology. For this purpose, we have constructed a vision system by combining neuromorphic analog VLSI microchips and digital circuits configured with field programmable gate arrays (FPGA). The front-end of the system is a previously developed multichip analog VLSI system comprising a silicon retina and orientation chip, which mimic the functions and structure of the retinal and V1 simple cells, respectively. The FPGA is used to implement a dedicated digital circuit for post-processing, following the analog multichip system.

In the field of neuromorphic engineering (Mead, 1990), several studies have been conducted on reconfigurable VLSI hardware systems for the simulation of neuron network models

* Corresponding author. Tel.: +81 6 6879 7785; fax: +81 6 6879 7784.
E-mail address: yagi@eei.eng.osaka-u.ac.jp (T. Yagi).

0893-6080/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2008.07.003

(Chicca et al., 2007; Choi, Merolla, Arthur, Boahen, & Shi, 2005; Liu et al., 2001; Merolla, Arthur, Shi, & Boahen, 2007; Serrano-Gotarredona et al., 2005; Shi, Tsang, Lam, & Meng, 2006; Vogelstein, Mallik, Vogelstein, & Cauwenberghs, 2007). These systems have designed as a hardware platform in order to verify and explore computational principle in neural networks, especially in spike domain. Therefore, most of these systems operate with an asynchronous spike representation that reflects the output of the neurons in the visual cortex by using the communication protocol, i.e., address event representation (AER), for interfacing the spike-based systems (Boahen, 2000). These systems are important for simulating the model in the spike domain, e.g., the dynamical behavior in synapses.

Our system operates with a synchronous frame-based architecture. The drawback of this method is that it does not reflect the spike representation of visual cortical neurons. Moreover, one cannot represent spike-level dynamics by using this method. However, some bottom-up models that were proposed in order to explain the functional roles of the V1 neurons represent their response with an analog value by considering only the firing rate of the spike, and do not necessarily require a spike representation for computation (e.g., Ohzawa, DeAngelis, and Freeman (1990) and Tohyama and Fukushima (2005)). The responses of such a model can be represented by using a frame-based architecture. In addition, it is useful to reconstruct the neural images because the output image of the system corresponds to the neural image at a particular instant. We have used an analog frame-based architecture in order to build a multichip VLSI vision system for a texture segregation system by modeling after the hierarchical structure of visual cortical computation rather than considering the detailed characteristics of individual neurons.

2. Model for V1 neurons

2.1. Simple and complex cells

We describe the model of the V1 simple and complex cells used in this study. Fig. 1 shows the schematic of the V1 neuron model. This is designed on the basis of the hierarchical feed-forward model by Hubel and Wiesel (Hubel & Wiesel, 1962). These V1 neurons are used as the fundamental unit for a higher-level computation.

First, an input image is filtered by center-surround spatial filters, similar to the receptive field of neurons of lateral geniculate nucleus (LGN) in mammalian visual systems. The V1 simple cells, which receive the signal from the LGN neurons, have an elongated receptive field and respond specifically to oriented bars or boundaries. This orientation-selective receptive field can be produced by aggregating the responses of the neurons that exhibit the center-surround receptive field (Hubel & Wiesel, 1962). The aggregation is carried out along a specific orientation. Such an orientation is termed the preferred orientation, and the response of the neurons to the input of this orientation is the maximum. Here, the aggregations along the 0° (horizontal) and 90° (vertical) orientations are shown. They provide a simple cell-like response, i.e., S_0 and S_{90}.

The orientation-selective outputs are full-wave rectified and transmitted to the spatial pooling stage. Here, we have employed a maximum operation as the mechanism of the spatial pooling (Riesenhuber & Poggio, 1999). The responses of the neighboring pixels aligned perpendicular to the preferred orientation are compared. The maximum response is selected for generating a complex cell-like response, C_{90} and C_0, as follows:

$$C_{90}(i,j) = \max\{|S_{90}(i-\delta,j)|, |S_{90}(i,j)|, |S_{90}(i+\delta,j)|\} \quad (1)$$

$$C_0(i,j) = \max\{|S_0(i,j-\delta)|, |S_0(i,j)|, |S_0(i,j+\delta)|\}. \quad (2)$$

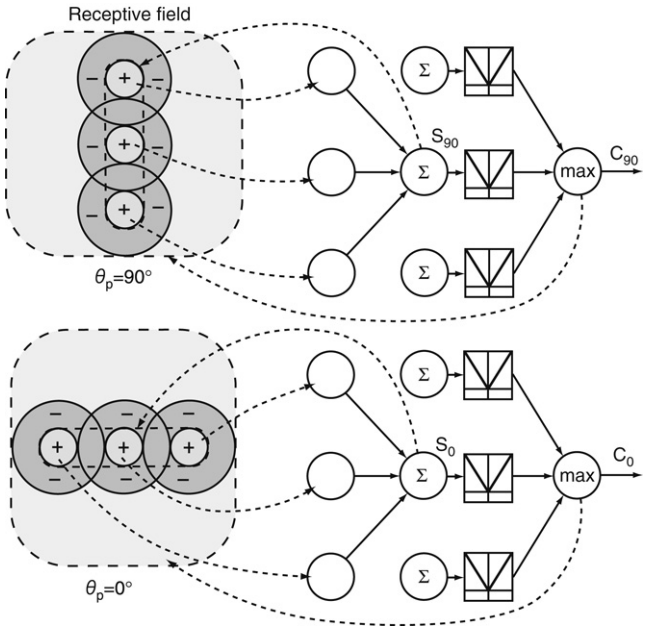

Fig. 1. Schematic of the model of V1 neurons. The input image is filtered by 90° orientation filters produced by the feed-forward aggregation of the center-surround receptive field. This output is full-wave rectified and transmitted to the spatial pooling stage to generate a complex cell-like response. The bottom figure shows the model of a 0° tuned complex cell.

2.2. Suppression outside CRF

It is known that some V1 neurons receive suppression outside the classical receptive field (CRF). CRF is defined as the region in which an optimal stimulus elicits a vigorous response from a neuron. In this context, the area that is indicated as the receptive field in Fig. 1 is a CRF. The areas outside the CRF can modulate the responses of the neurons in the V1, and these areas often exhibit suppression (Blakemore & Tobin, 1972). It is known that these surrounding areas are often spatially asymmetric and do not encircle the CRF (Walker, Ohzawa, & Freeman, 1999).

Here, we computed the neurons whose surrounding suppression area is asymmetric, using the complex cell responses. Fig. 2 shows the model utilized in this study. The response of the cell located at the center of the CRF is suppressed by the stimulus provided to the surrounding suppression area. The bar shown inside the CRF and the surrounding suppression area show the preferred orientation. Here, we prepare two types of suppression patterns shown in (A) and (B). In (A), the response is modulated by the suppression area on the right side of the CRF. The preferred orientation of the cell is 90°. In (B), the response is modulated by the suppression area on the upper side of the CRF. The preferred orientation of the cell is 0°. Therefore, the resultant responses C'_{90} and C'_{90} are obtained as follows:

$$C'_{90}(i,j) = C_{90}(i,j) - C_{90}(i+1,j) \quad (3)$$

$$C'_0(i,j) = C_0(i,j) - C_0(i,j-1). \quad (4)$$

Furthermore, we also produced a cell by the integration of two cells, shown as follows:

$$C_m(i,j) = C'_{90}(i,j) + C'_0(i,j). \quad (5)$$

3. Hardware implementation

3.1. System overview

We implement the model of the V1 neurons using analog and digital VLSI circuits. Fig. 3(A) shows the block diagram of