



University of Pennsylvania
ScholarlyCommons

Departmental Papers (BE)

Department of Bioengineering

May 2004

A multi-chip implementation of cortical orientation hypercolumns

Thomas Y. W. Choi

Hong Kong University of Science and Technology

Bertram E. Shi

Hong Kong University of Science and Technology

Kwabena A. Boahen

University of Pennsylvania, boahen@seas.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/be_papers

Recommended Citation

Choi, T. Y., Shi, B. E., & Boahen, K. A. (2004). A multi-chip implementation of cortical orientation hypercolumns. Retrieved from http://repository.upenn.edu/be_papers/13

Copyright 2004 IEEE. Reprinted from *Proceedings of the 2004 International Symposium on Circuits and Systems (ISCAS 2004)*, Volume 3, pages 13-16.
Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=29374>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

This paper is posted at ScholarlyCommons. http://repository.upenn.edu/be_papers/13
For more information, please contact libraryrepository@pobox.upenn.edu.

A multi-chip implementation of cortical orientation hypercolumns

Abstract

This paper describes a neuromorphic implementation of the orientation hypercolumns found in the mammalian primary visual cortex. A hypercolumn contains a group of neurons that respond to the same retinal location, but with different orientation preferences. The system consists of a single silicon retina feeding multiple orientation selective chips, each of which contains neurons tuned to the same orientation, but with different receptive field centers and spatial phases. All chips operate in continuous time, and communicate with each other using spikes transmitted by the asynchronous digital Address Event Representation communication protocol. This enables us to implement recurrent interactions between neurons within one hypercolumn, even though they are located on different chips. We demonstrate this by measuring shifts in orientation selectivity due to changes in the feedback.

Keywords

neurophysiology, recurrent neural nets, visual perception, asynchronous digital address event representation, communication protocol, cortical orientation hypercolumns, mammalian primary visual cortex, multichip implementation, multiple orientation selective chips, neuromorphic implementation, orientation selectivity, receptive field centers, recurrent neuron interaction, retinal location, silicon retina, spatial phases

Comments

Copyright 2004 IEEE. Reprinted from *Proceedings of the 2004 International Symposium on Circuits and Systems (ISCAS 2004)*, Volume 3, pages 13-16.

Publisher URL: <http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=29374>

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the University of Pennsylvania's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org. By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

A MULTI-CHIP IMPLEMENTATION OF CORTICAL ORIENTATION HYPERCOLUMNS

Thomas Y. W. Choi¹, Bertram E. Shi¹ and Kwabena A. Boahen²

¹Dept. of EEE, HKUST, {eethomas,eebert}@ee.ust.hk and ²Dept. of Bioengineering, UPenn, boahen@seas.upenn.edu

ABSTRACT

This paper describes a neuromorphic implementation of the orientation hypercolumns found in the mammalian primary visual cortex. A hypercolumn contains a group of neurons that respond to the same retinal location, but with different orientation preferences. The system consists of a single silicon retina feeding multiple orientation selective chips, each of which contains neurons tuned to the same orientation, but with different receptive field centers and spatial phases. All chips operate in continuous time, and communicate with each other using spikes transmitted by the asynchronous digital Address Event Representation communication protocol. This enables us to implement recurrent interactions between neurons within one hypercolumn, even though they are located on different chips. We demonstrate this by measuring shifts in orientation selectivity due to changes in the feedback.

1. INTRODUCTION

Neurons in the mammalian primary visual cortex (V1) are predominantly orientation selective [1]. In other words, they respond strongest when the retina is stimulated by a bar at a preferred location and orientation. Although V1 neurons are also selective along other stimulus dimensions (e.g. direction/speed of motion and binocular disparity), orientation selectivity seems to be a fundamental primitive from which selectivity along other stimulus dimensions can be constructed.

A functional model that accounts for the responses of a large proportion of simple cells in the primary visual cortex consists of a linear spatio-temporal filtering stage followed by several nonlinear mechanisms [2]. Linear spatio-temporal filtering determines the neural selectivity. The linear filter for an orientation selective neuron is commonly modelled using a spatial Gabor filter [3], whose impulse response is a 2-D sinusoid modulated by a Gaussian. Physiological measurements indicate that neighboring cortical neurons often share similar filter tunings, except for a phase difference in the sinusoid of $\pi/2$ [4]. The first nonlinear mechanism, half-wave rectification, conserves metabolic energy by mapping mean levels to a low quiescent spike rate. The positive and negative parts of a signal are encoded on two separate neurons, denoted by ON and OFF.

Hubel and Wiesel suggested that neurons serving a particular retinal location are grouped into a hypercolumn [1]. In their “ice-cube” model, illustrated in Figure 1(a), the visual cortex is thought of as a thin 2-D sheet. This sheet is cut into large blocks, corresponding to different hypercolumns, which are arranged retinotopically: hypercolumns serving neighboring retinal regions

occupy neighboring cortical regions. Each hypercolumn is subdivided into smaller cubes, each containing neurons sensitive to a particular orientation.

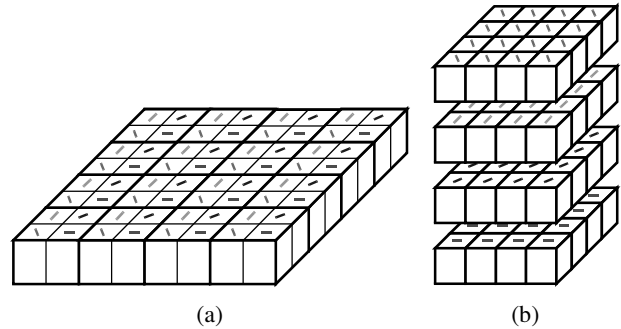


Fig. 1: (a) The “ice-cube” model of the visual cortex. Thick lines divide hypercolumns. Thin lines divide columns with preferred orientations indicated by the oriented bars. (b) The multi-chip architecture used here. Each layer represents one chip containing neurons tuned to the same orientation but different retinal positions. Vertically aligned neurons receive inputs from the same retinal locations.

This paper describes a model of orientation selective hypercolumns in the visual cortex that captures the following properties. First, it contains neurons with different preferred orientations and receptive field centers. Second, the model neurons have receptive field profiles similar to Gabor functions. Third, each orientation is served by two sets of neurons whose spatial phases differ by $\pi/2$. Fourth, the neural responses include half wave rectification, with ON and OFF neurons carrying the positive and negative components of the linear filter output. Fifth, the neurons operate in parallel and in continuous time, which enables the incorporation of both feedforward and feedback interconnections between neurons. Although feedforward models of orientation selectivity can account for much of the responses of cortical neurons, their shortcomings, as well as the ubiquity of feedback interconnections in cortex have prompted the development of feedback models of orientation selectivity [5].

Our approach is similar to those suggested by Serrano-Gotarredona et. al. [6] and Venier et. al. [7]. It differs primarily in the orientation selective receptive fields. The approach proposed in [6] can implement only spatial filters with x - y separable convolution kernels, leading to only horizontal or vertical orientation selectivity. The system reported in [7] implements neurons with purely excitatory and even symmetric RF profiles, and cannot

implement neurons with the phase quadrature Gabor-like receptive fields observed in cortex.

It differs from that of Cauwenberghs and Waskiewicz [8] because we split orientation selective neurons among different chips, rather than combining them on the same chip. It also differs from the multi-chip architecture described by Liu et. al. [9], where neurons tuned to the same retinal position but different orientations are on the same chip. It differs from the Cellular Neural Network Universal Machine [10] and the computation on readout architecture [11], which multiplex the same circuits in time for multiple calculations, because it computes all maps in parallel enabling recurrent interactions between them.

2. MODEL IMPLEMENTATION

Figure 1(b) illustrates the general architecture, which consists of multiple chips, each containing layers/arrays of neurons with the same orientation tuning but with different receptive field centers and spatial phases. These orientation selective chips have been described in detail elsewhere [12], so we only recapitulate their salient characteristics here. Each chip contains four 32×64 arrays of neurons. Each position in the array corresponds to a different receptive field center. For each receptive field center location, all four neurons have Gabor-like receptive field profiles with the same orientation/spatial frequency tuning and bandwidth. They can be grouped into two pairs of ON and OFF neurons whose receptive field profiles have spatial phase offsets of 0 radians (referred to as EVEN neurons) and $-\pi/2$ radians (referred to as ODD neurons). The neurons operate in continuous time. Their orientation and spatial frequency tuning is continuously adjustable via externally supplied analog bias voltages. The neurons receive spike rate encoded inputs on ON and OFF channels, and produce spike rate encoded outputs, where each spike has fixed width and fixed amplitude. Spike activity is communicated between chips using the Address Event Representation Protocol (AER) [14]. The chips were implemented in the TSMC0.25um process available through MOSIS, and have a quiescent power dissipation of 3mW.

Representing the two retinal dimensions on the 2-D surface of the silicon substrate enables us to build high resolution arrays. Since each neuron has the same orientation selectivity, its structure is identical to the other neurons on the chip, which is constructed by tiling identical circuit blocks. The system is also both modular and expandable. Adding more chips increases the resolution at which we cover the orientation dimension.

We implemented both a feedforward and a feedback model of orientation selective hypercolumns. Figure 2(a) shows the block diagram of the feedforward implementation. The output of the silicon retina described in [13] is fanned out to a chain of four Gabor chips that operate independently. This system implements a 32×64 retinotopic array of hypercolumns, with each hypercolumn containing neurons tuned to orientations 0° , 45° , 90° and 135° .

Each Gabor chip contains routing circuits that facilitate the construction of this system. The split circuits enable fan out by creating two copies of their input: one that is sent to the neuron array

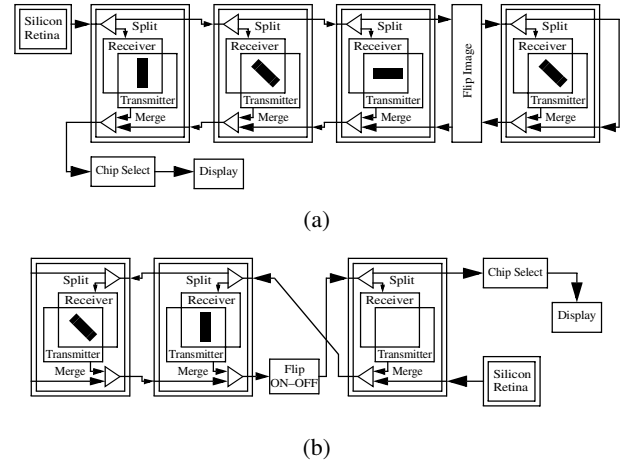


Fig. 2: (a) The feedforward implementation of orientation selective hypercolumns. Each box with the double line border represents a chip containing a retinotopic array of neurons. Gabor filtering chips are represented by the larger boxes with the dark bar indicating the tuned orientation of the neurons on the chip. Boxes with single line borders indicate circuits that manipulate AER encoded spike trains. Because the chips can only be tuned to orientations between 0° and 90° , the “Flip Image” block flips the image horizontally to obtain neurons tuned to the orientation 135° . The “Chip Select” block passes only spikes originating from a desired chip for display. (b) The feedback implementation. The “Flip ON-OFF” block inverts the polarity of the images. The missing bar in the third chip indicates that its spatial filtering is disabled.

for processing and the other that is sent on to the next Gabor chip in the chain. Routing complexity expands automatically to accommodate the number of chips in the system. The merge circuits enable us to combine spike trains from multiple chips onto the same digital bus. Spike activity is encoded on the bus in such a way that we can distinguish spikes originating from different chips.

Figure 2(b) shows the block diagram of the feedback implementation of orientation hypercolumns. As described below, the interactions between neurons tuned to different orientations is inhibitory. With a large number of orientations, these inhibitory interactions will sharpen orientation selectivity. Here, we study the interactions between only two orientations for the sake of simplicity. In this case, the dominant effect is a shift in the orientation selectivity of the neurons. The system can be expanded easily to include more orientations by adding more orientation selective chips.

In the feedback system, each neuron in the orientation selective arrays is driven by the residual signal, which is defined at each retinal location as the difference between the silicon retina input and the sum of the outputs of the orientation selective neurons centered at that retinal location. Intuitively, the summed output over all the neurons tries to fit the input. The residual measures the quality of the fit. If the residual has a large component near the tuned orientation of a chip, that chip increases its response.

The feedback system uses the merge circuits to implement summation of spike trains. The merge circuit of the Gabor chip tuned to vertical orientation combines the output spike trains from itself and the chip tuned to diagonal orientations. Spikes from the two chips are summed by the third chip, since all spikes appearing at its split input are sent to the neuron array irrespective of which chip they originated from. This third chip has its spatial filtering circuits disabled so that it combines signals at each retinal location independently. Similarly, the merge circuit of the third chip combines the sum of the outputs with the silicon retina input to represent the residual signal.

3. EXPERIMENTAL RESULTS

To test the feedforward system, we presented the retina with circular Gabor patches of varying orientation using an LCD monitor placed 20cm in front of a 4mm lens attached to the retina. To avoid edge effects, we chose the radius of the modulating Gaussian so that the Gabor patch was entirely confined within the field of view of the retina. The radius of the patch was around 2.5cm, corresponding to a radius of 14 pixels on the retina. The period of the sine wave grating on the retina was approximately the same period as the sinusoidal variation of the Gabor filters, about six pixels. The input orientations ranged from 0° to 157.5° in steps of 22.5° . The other orientations (180° to 337.5°) are equivalent to the orientations from 0° to 157.5° .

To characterize the response of the neurons, we measured the average spike rate in the population of ODD neurons. By measuring the population response irrespective of receptive field center and neural polarity, we eliminate any dependency on the spatial phase of the stimulus. For each input orientation, we collected approximately 200 spikes from each chip, counted the total number of ODD spikes, and divided it by the total time it took to generate these spikes. We plot our results using polar plots in Figure 3. The maximum response occurs at the tuned orientation, and the minimum response at the orthogonal orientation.

We also characterized neurons in the feedback configuration. The dominant effect of the inhibitory feedback interaction when two orientations are coupled is a repulsion of the tuning curves. Consider two neurons, labelled 1 and 2, whose receptive field centers are the same, but that are tuned to different orientations in the absence of coupling. When these two neurons are coupled together, the peak of the tuning curve of Neuron 1 will shift away from the preferred orientation of Neuron 2 because the strong response of Neuron 2 to orientations near its preferred orientation reduces the response of Neuron 1.

We measured this peak shift in our system by tuning one array to respond maximally to orientations of 45° when uncoupled and varying the uncoupled orientation tuning of the other array between 0° , 22.5° , 67.5° and 90° . We then coupled the two arrays together as shown in Figure 2(b). We used a pattern generator to generate square-wave gratings with variable orientations as an external signal source to the system, and measured the population response of the ODD neurons on each chip using the same technique as in the feedforward system.

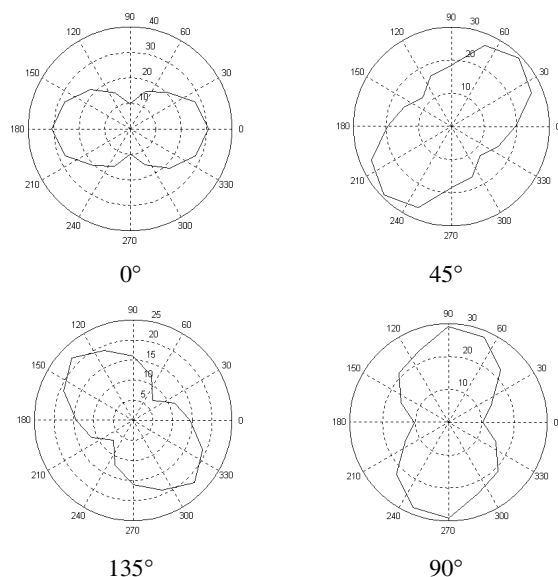


Fig. 3: Polar plots showing the average spike rate of the ODD neurons from chips tuned to different orientations. For each point, the angle represents the input orientation and the radial magnitude represents the total output spike rate of the population (~4000 neurons) minus the background spike rate measured in response to a blank screen. Spike rate is measured in kHz.

Figure 4 shows the polar plots of the responses from the 45° array. We quantify the tuning of the array using the resultant

$$R = \left(\sum_k r_k e^{j2\theta_k} \right) \left(\sum_k r_k \right)^{-1}$$

where r_k is the response to the input orientation θ_k . We double the angle θ_k in the calculation of the resultant to take into account the fact that orientations that differ by 180° are equivalent, and undo this by defining the tuned orientation of the array to be $0.5\angle R$. The feedback from the second array shifts the tuning of the neurons in the 45° array in the direction away from the tuning of the second array. The magnitude of the shift increases with the difference between the two orientations, which is consistent with our theoretical predictions (not given here due to space constraints).

4. CONCLUSION

We have constructed a neuromorphic system that models the columnar organization of orientation selective neurons in the mammalian visual cortex. Although orientation selectivity is only one function of V1, because the receptive field properties of the silicon neurons closely match those of orientation selective visual cortical neurons, we can use these neurons as building blocks to implement selectivity along other stimulus dimensions associated with V1. For example, we have shown that these orientation selective neurons can be used to construct neurons tuned to binocular disparity [15]. Another paper in these proceedings describes our approach to constructing velocity tuned neurons [16]. The long-term goal in this work is to construct retinotopic arrays of

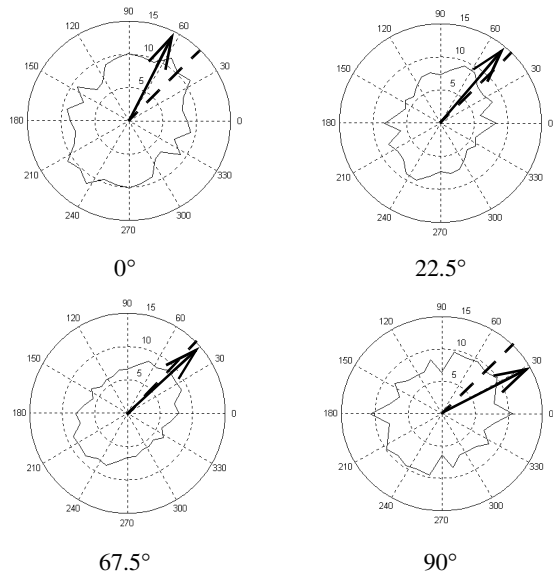


Fig. 4: Measured orientation tuning curve of the neuron array with a uncoupled peak tuning at 45° when coupled with a second array containing neurons with the uncoupled tunings listed. The arrow indicates the tuned orientation as determined by the resultant. The dotted line shows tuned orientation without feedback. The radial magnitude is the total population spike rate measured in kHz.

neurons selective along all stimulus dimensions associated with V1.

Despite the fact that neurons within one hypercolumn physically reside on different chips, we have demonstrated the system is still capable of implementing feedback interactions between them. This capability is essential in modelling V1, where feedback connections are pervasive. For example, physiologically observed shifts in orientation tuning physiologically in response to learning and adaptation, can account for performance improvements in orientation discrimination, have been attributed to changes in feedback strength [17]. Thus, although the current system does not include learning or adaptation, our present results indicate that we will be able to include these effects in the future.

5. ACKNOWLEDGEMENTS

We would like to thank Paul Merolla and John Arthur for their work in designing the transmitter and receiver circuits for the AER interface, Kai Hynna and Brian Taba for their assistance in generating the chip layout and Kareem Zaghoul for providing the silicon retina. This work was supported in part by the Hong Kong Research Grants Council under grant HKUST6218/01E and in part by an NSF CAREER grant (ECS00-93851).

REFERENCES

[1] D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," *J. Physiology*, vol. 160, pp. 106-154, 1962.

[2] D. G. Albrecht and W. S. Geisler, "Motion selectivity and the contrast response function of simple cells in the visual cortex," *Visual Neuroscience*, vol. 7, pp. 531-546, 1991.

[3] J. P. Jones and L.A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *Journal of Neuroscience*, vol. 58, no.6, pp. 1233-1258, Dec. 1987.

[4] D. A. Pollen and S. Ronner, "Visual cortical neurons as localized spatial-frequency filters," *IEEE Trans. on Systems, Man and Cybernetics*, vol. 13, pp. 907-916, 1973.

[5] D. Ferster and K. D. Miller, "Neural mechanisms of orientation selectivity in the visual cortex," *Annual Review of Neuroscience.*, vol. 23, pp. 441-71, 2000.

[6] T. Serrano-Gotarredona, A. G. Andreou, A.G. and B. Linares-Barranco, "AER image filtering architecture for vision-processing systems," *IEEE Transactions on Circuits and Systems I*, vol. 46, no. 9, pp. 1064 -1071, Sept. 1999.

[7] P. Venier, A. Mortara, X. Arreguit, and E. A. Vittoz, "An integrated cortical layer for orientation enhancement," *IEEE J. of Solid-State Circuits*, vol. 32, no. 2. pp. 177-186. Feb. 1997.

[8] G. Cauwenberghs and J. Waskiewicz, "Focal-plane analog VLSI cellular implementation of the boundary contour system," *IEEE Trans. Circuits and Systems-I*, vol. 46, no. 2, pp. 327-334, Feb. 1999.

[9] S. C. Liu, J. Kramer, G. Indiveri, T. Delbruck, T. Burg, and R. Douglas, "Orientation selective aVLSI spiking neurons," *Neural Networks*, vol. 14, pp. 629-643, 2001.

[10] T. Roska and L. O. Chua, "The CNN Universal Machine: An Analogic Array Computer," *IEEE Transactions on Circuits and Systems-II*, vol. 40, pp. 163-173, 1993.

[11] V. Gruev and R. Etienne-Cummings, "Implementation of steerable spatiotemporal image filters on the focal plane," *IEEE Transactions on Circuits and Systems-II: Analog & Digital Signal Processing*, vol. 49, pp. 233-44, 2002.

[12] T. Y. W. Choi, B. E. Shi, and K. Boahen, "An ON-OFF Orientation Selective Address Event Representation Image Transceiver Chip," *IEEE Transactions on Circuits and Systems-I*, Feb. 2004.

[13] K. A. Zaghoul, *A silicon implementation of a novel model for retinal processing*, Ph.D. thesis, Dept. of Neuroscience, University of Pennsylvania, 2001.

[14] M. Sivilotti, "Wiring considerations in Analog VLSI Systems with Application to Field Programmable Networks", *Ph.D. thesis*, California Institute of Technology, 1991.

[15] E. K. C. Tsang and B. E. Shi, "A Neuromorphic Multi-chip Model of a Disparity Selective Complex Cell," in *Advances in Neural Information Processing Systems*, vol. 16, S. Thrun, L. Saul, and B. Schoelkopf, Eds. Cambridge, MA: MIT Press, in press.

[16] B. E. Shi, E. K. C. Tsang, and P. S. P. Au, "An ON/OFF temporal filter circuit for visual motion analysis," *Proc. IEEE Intl. Symp. on Circuits and Systems*, Vancouver, BC, May 2004.

[17] A. Schoups, R. Vogels, N. Qian, and G. Orban, "Practising orientation identification improves orientation coding in V1 neurons," *Nature*, vol. 412, pp. 549-53, 2001.