May 2003

An orientation selective 2D AER transceiver

Thomas Yu Wing 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

Copyright 2003 IEEE. Reprinted from *Proceedings of the 2003 International Symposium on Circuits and Systems (ISCAS 2003)*, Volume 4, pages IV-800 - IV-803.
Publisher URL: http://ieeexplore.ieee.org/xpl/tocresult.jsp?isNumber=27140&page=13

An orientation selective 2D AER transceiver

Abstract
This paper describes an address event representation (AER) transceiver chip that accepts 2D images and produces 2D output images equal to the input filtered by even and odd symmetric orientation selective spatial filters. Both input and output are encoded as spike trains using a differential ON/OFF representation, conserving energy and AER bandwidth. The spatial filtering is performed by symmetric analog circuits that operate on input currents obtained by integrating the input spike trains, and which preserve the ON/OFF representation. This chip is a key component of a multi-chip system we are constructing that is inspired by the visual cortex. We present measured results from a 32 x 64 pixel prototype, which was fabricated in the TSMC0.25 μm process on a 3.84mm by 2.54mm die. Quiescent power dissipation was 3mW.

An Orientation Selective 2D AER Transceiver

Thomas Yu Wing CHOI[1], Bertram E. SHI[1], and Kwabena BOAHEN[2]

[1]Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, {eethomas, eebert}@ee.ust.hk

[2]Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104-6392, kwabena@neuroengineering.upenn.edu

ABSTRACT

This paper describes an address event representation (AER) transceiver chip that accepts 2D images and produces 2D output images equal to the input filtered by even and odd symmetric orientation selective spatial filters. Both input and output are encoded as spike trains using a differential ON/OFF representation, conserving energy and AER bandwidth. The spatial filtering is performed by symmetric analog circuits that operate on input currents obtained by integrating the input spike trains, and which preserve the ON/OFF representation. This chip is a key component of a multi-chip system we are constructing that is inspired by the visual cortex. We present measured results from a 32 x 64 pixel prototype, which was fabricated in the TSMC0.25μm process on a 3.84mm by 2.54mm die. Quiescent power dissipation was 3mW.

1. INTRODUCTION

Moving from the retina to higher levels of visual processing in the cortex, neurons become progressively more selective to more complex stimuli. Cells in the retina are sensitive along stimulus dimensions of position, spatial frequency (size), temporal frequency and color. In the primary visual cortex, additional selectivity along the dimensions of orientation, direction of motion and binocular disparity emerges. Subsequent areas are selective to higher order dimensions such as curvature and illusory contours. Concurrently, there is a progressive increase in the size of the receptive field along stimulus dimensions established earlier, e.g. spatial position. Thus, neurons in V2 respond to visual stimuli in a much larger spatial area than ganglion cells in the retina.

A functional model that seems to account for the responses of a large proportion of cells in the primary visual cortex consists of a linear spatio-temporal filtering stage and three nonlinear mechanisms: contrast normalization, half-wave rectification and expansive exponentiation[1][2][3]. Linear spatio-temporal filtering determines the neural selectivity along different stimulus dimensions. Contrast normalization accounts for the observed saturation of the neural response with increasing contrast. The saturation occurs at a fixed contrast, independent of the response level, enabling neurons to retain selectivity over a wide input contrast range. Half-wave rectification conserves metabolic energy by mapping mean levels to a low quiescent spike rate. Signals above and below the mean are carried by complementary channels. The expansive exponent enhances stimulus selectivity.

In this work, we describe a silicon chip that implements two components of this model: linear orientation selective spatial filtering and half-wave rectification. The impulse response of the spatial filters approximate even and odd symmetric Gabor functions, which are commonly used to model the spatial receptive field profiles of visual cortical neurons[4]. Complementary ON/OFF channels carry positive and negative parts of all input, internal and output signals, which are processed using analog continuous time circuits. As in biological systems, this representation improves energy efficiency.

This chip is intended to serve as one component of a multi-chip system that takes input from a silicon retina and implements more complex visual information processing inspired by that found in the visual cortex. To facilitate construction of this system, input and output signals are encoded as spike trains, which are communicated on and off chip using the asynchronous Address-Event Representation (AER) communication protocol[7]. The AER protocol is more efficient than scanning when the spike activity within the array is sparse, as we expect here since only a few image locations will contain edges near the orientation selected by each chip and the quiescent spike activity in the array is low due to the ON/OFF signal representation. This combination of continuous time analog processing and digital communication circuits, which directly allocates power to salient areas in the image, results in better power efficiency than a conventional DSP approach.

We focus upon orientation as a first step in constructing this a system, as this dimension seems to be a fundamental primitive from which selectivity along other stimulus dimensions can be constructed. For example, linear direction-selective spatio-temporal filters can be obtained by cascading these orientation selective filters with bandpass temporal filters and combining their outputs[5]. Filters tuned to binocular disparity can be obtained by combining orientation selective filtered images from the left and right eye[6].

0-7803-7761-3/03/$17.00 ©2003 IEEE

This paper describes the architecture of this chip, as well as measurement results from a prototype.

2. CHIP ARCHITECTURE

2.1 AER Interface

The chip is a transceiver, containing both an AER transmitter and an AER receiver (see Chips A and B of Fig. 1a). The AER protocol was developed to communicate continuous time spike activity from an array of silicon neurons in one chip to another chip over a digital bus. The transmitter signals a spike occurrence in the array by placing the location (address) of the spiking neuron onto the bus. The receiver takes the address that appears on the bus and feeds a spike to the corresponding neurons in its array. The protocol is asynchronous, with the time that the address appears on the bus encoding the spike time directly. Collisions between simultaneous spikes from two neurons on the array are handled by arbitration.

The AER interface includes routing circuitry to facilitate the construction of a multi-chip network. Fig. 1a illustrates a three chip network. The split circuit enables fan out by splitting the incoming AER address stream into two: one sent to the pixel array and the other sent to an output, which can be fed into the input of another chip. A merge circuit enables fan in by combining the output of the array with an AER stream provided through a second input into a single serial output stream. In Fig. 1a, the merge output of Chip B encodes spikes from both Chips A and B.

Addresses are placed onto the bus in "bursts," where each burst encodes all of the simultaneous spikes from neurons within a given row and a given chip. We use a word serial format, where each burst is a sequence of addresses. As shown in Fig. 1b, the transmitter signals the start of a burst by placing an address identifying the source chip onto the address lines (Addr) and taking the request signal ReqY high. Subsequent addresses are signalled by taking _ReqX low. The second address identifies the row. Each of the remaining addresses identifies one of the columns containing a neuron that spiked. The transmitter signals the end of the burst by taking ReqY low. The receiver acknowledges receipt of each address by a transition on the Ack line.

We use absolute addressing to identify rows and columns within a chip, but relative addressing to identify each chip. Each chip signals its own activity with bursts whose chip addresses are set to zero. Every time a chip relays a burst from its split or a merge input, it increments the chip address. For example, a chip address of 1 at the merge output of Chip B in Fig. 1a indicates the spikes in the burst come from Chip A.

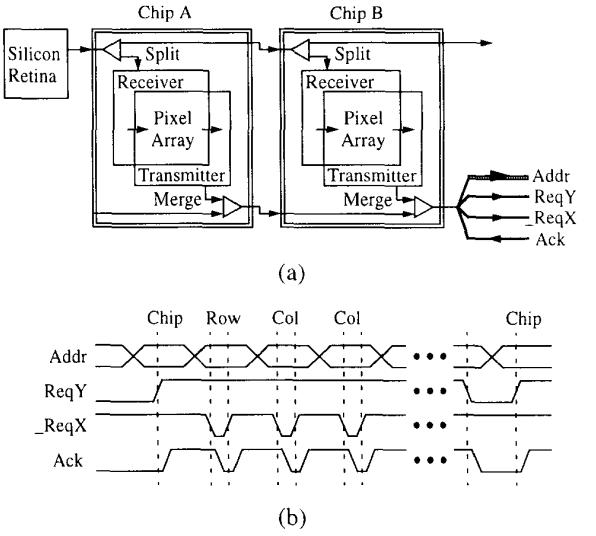

Fig. 1: (a) A three chip system where the output of a silicon retina[8] is fanned out to two orientation selective chips (Chip A and Chip B) tuned to different orientations. (b) A signal diagram of the merge output of chip B showing the addressing scheme.

2.2 Pixel level processing

The pixel processing array filters the incoming image with the transfer function:

$$H(e^{j\omega_x}, e^{j\omega_y}) = \frac{H_\Omega}{1 + \frac{(2 - 2\cos(\omega_x - \Omega_x))}{(\Delta\Omega_x)^2} + \frac{(2 - 2\cos(\omega_y - \Omega_y))}{(\Delta\Omega_y)^2}}$$

where H_Ω is the gain at resonance, (Ω_x, Ω_y) is the center spatial frequency, and $\Delta\Omega_x$ and $\Delta\Omega_y$ are the 6dB half bandwidth in the x and y directions. Pixel values at input and output are in general complex valued. For a real valued input image, the real and imaginary parts of the output equal the input image convolved with even and odd symmetric filters. Similar to Gabor functions, the impulse responses are cosine or sine waves modulated by an envelope that decays with distance from the origin. However, the envelope of these filters decays more sharply at the origin and slower at the tails than the Gaussian envelope of a Gabor function.

Each pixel within the array receives four spike trains, corresponding to the ON and OFF components of the real and imaginary parts of the input. Current mode integrators[9] convert the spike trains into currents that are approximately proportional to the incoming spike rates. In the addressing scheme, the real and imaginary parts are encoded by the least significant bit of the row address and the ON and OFF components by the least significant bit of the column address.

The four input currents are then processed by an analog neural network that produces four output currents, corresponding to the ON and OFF components of the real and imaginary parts of the filter output. The ON and OFF components of the output are interconnected in opponency, so that they mutually inhibit each other. At any time, only one component of the output is positive, the other being close to zero.

This network is a two dimensional extension of that described in [10], except that the "ON/OFF block" is placed at the output of the diffuser networks, rather than at the input. This improves the ON/OFF representation at the output by reducing the common-mode activity in the complementary channels. This change required that the diffuser or pseudo-resistor networks be implemented with NMOS rather than PMOS transistors. By adjusting analog bias voltages controlling pseudo-conductance ratios and current gains within the array, we can tune the array to arbitrary spatial frequencies and orientations between 0 and 90 degrees. Other orientations can be obtained by remapping input and output addresses to flip the array horizontally and/or vertically.

Each of the four output currents is passed to a spiking neuron circuit similar to that described in [11], which encodes each current by a spike train whose rate is proportional to the current amplitude. As activity is sparse, most pixels' input or output is zero, which can be encoded by the lack of activity in either channel, conserving power dissipated by spiking and preserving bandwidth on the AER bus.

3. EXPERIMENTAL RESULTS

We designed and fabricated an array of 32 x 64 pixels in the TSMC0.25um mixed signal/RF process available through MOSIS. This process contains 5 metal layers and 1 poly layer, uses non-epitaxial wafers, and is intended for 2.5V applications.

The array layout was generated by tiling metapixels whose layout is shown in Fig. 2. Each metapixel contains the circuits required for two pixels stacked vertically. Each metapixel is 103μm by 49μm (860λ by 390λ for $\lambda = 0.12\mu m$). The die size was 3.84mm by 2.54mm.

We laid out the metapixels to minimize interference between the analog spatial filtering circuits and the digital communication circuits (the spiking neurons, current mode integrators and the AER interface). The analog filtering circuits lie in the middle of the metapixel, sandwiched between digital circuits on the top and bottom. Within the digital parts, the integrators lie next to the analog circuits. The spiking neurons, which generate the most switching noise, lie at the top and bottom, farthest from the analog processing circuits. Guard rings are also inserted in between the Gabor cells, the integrators and the spiking neurons. The digital and analog circuits use separate power and ground lines. Bias lines connected to source voltages controlling current mirror gains run wide on the top metal layer to reduce impedance.

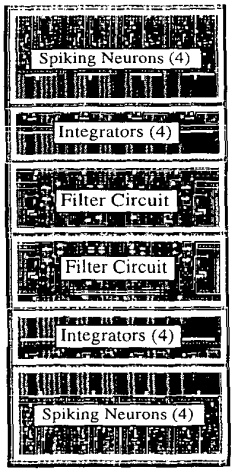

Fig. 2: Layout of one metapixel, containing the circuits necessary to process two vertically stacked pixels.

To test the response of the array, we excited pixel (16,32) with a 20kHz spike train from a pattern generator. All other inputs were silent. A logic analyzer collected the spike train at the merge output, which is digitally processed for analysis. Fig. 3(a,b,c,d) shows the average spike rates of the ON and OFF components of the real and imaginary parts of the output when the array is tuned to vertical orientations. The difference between the ON and OFF spike rates are shown in Fig. 3(e,f). Fig. 3(g,h) shows similar data when the array is tuned to horizontal orientations.

The power consumption increases with the total spike activity at the input and output. We measured the power dissipation of the chip while stimulating pixel (16,32) with spike trains ranging in frequency from 0Hz to 100kHz and plot the results in Fig. 4 as a function of average output activity per neuron, which is much higher than the input activity. The power increases linearly with the output activity. The quiescent power consumption with no input, but an average output activity of 14Hz, is about 3mW. The pads account for around 75% of the total power consumption. The digital communication circuits account for around 24%. The analog circuits consume less than 1%.

4. CONCLUSION

We have successful designed, fabricated and tested a 2D AER transceiver chip that performs orientation selective image filtering. Our initial characterizations of the chip indi-

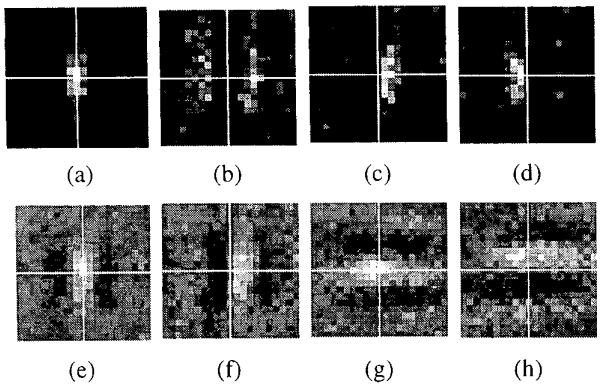

Fig. 3: Measured responses from a 21 by 21 pixel window to a spike train applied at pixel (16,32). (a) ON component of the real part of the output for the array tuned to vertical orientations. (b) OFF component of the real part. (c,d) ON and OFF components of the imaginary part. (e,f) The difference between the ON and OFF components of the real part and imaginary parts. (g,h) Similar data for the array tuned to horizontal orientations. White/black corresponds to a spike rates of (a) 267/0, (b) 98/0, (c) 142/0, (d) 149/0, (e) 267/-267, (f) 149/-149, (g) 173/-173 and (h) 134/134 Hertz. Crosshairs indicate pixel (16,32).

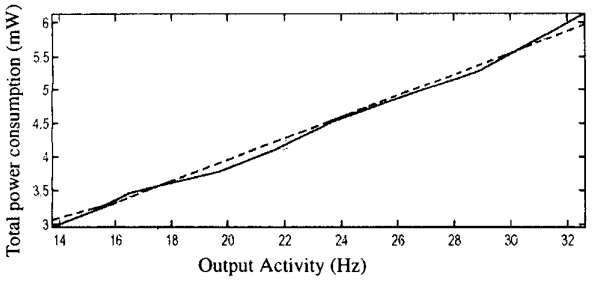

Fig. 4: The solid line plots total power consumption versus the average output activity per neuron. The dotted line is a linear least squares fit to the data, which has slope 0.16mW/ Hz and vertical offset 0.77mW.

cate that it functions as expected. Our ongoing work seeks to integrate this chip into a multi-chip architecture for visual information processing.

ACKNOWLEDGEMENTS

This work was supported by the Hong Kong Research Grants Council.

REFERENCES

[1] 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.

[2] D. J. Heeger, "Normalization of cell responses in cat striate cortex," *Visual Neuroscience*, vol. 9, pp. 181-197, 1992.

[3] D. J. Heeger, "Half-squaring in responses of cat striate cells," *Visual Neuroscience*, vol. 9, pp. 427-443, 1992.

[4] 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.

[5] A. B. Watson and J. A. J. Ahumada, "Model of human visual-motion sensing," *J. Optical Society of America A*, vol. 2, pp. 322-342, Feb. 1985.

[6] I. Ohzawa, G. C. DeAngelis and R. D. Freeman, "Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors," *Science*, vol. 249, pp. 1037-1041, 31 Aug. 1990.

[7] K. A. Boahen, "Point-to-point connectivity between neuromorphic chips using address events," *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 47, no. 5, pp. 416-434, May 2000.

[8] K. A. Zaghloul, *A silicon implementation of a novel model for retinal processing*, Ph.D. thesis,. University of Pennsylvania, 2001.

[9] K. A. Boahen, "The retinomorphic approach: Pixel-parallel adaptive amplification, filtering and quantization," *Analog Integrated Circuits and Signal Processing*, vol. 13, pp. 53-68, 1997.

[10]B. E. Shi, T. Y. W. Choi and K. Boahen, "On-off differential current mode circuits for Gabor-type spatial filtering," *Proc. IEEE Intl. Symp. On Circuits and Systems*, Phoenix, AZ, vol. II, pp. 724-727, May 2002.

[11]E Culurciello, R Etienne-Cummings, and K. Boahen, "Arbitrated Address Event Representation Digital Image Sensor", *IEEE International Solid-State Circuits Conference*, pp 92-93, San Francisco CA, February 2001.