

Multi-Chip Implementation of a Biomimetic VLSI Vision Sensor Based on the Adelson-Bergen Algorithm

Erhan Ozalevli and Charles M. Higgins

Department of Electrical and Computer Engineering
The University of Arizona, Tucson, AZ 85721
{erhan,higgins}@ece.arizona.edu

Abstract. Biological motion sensors found in the retinas of species ranging from flies to primates are tuned to specific spatio-temporal frequencies to determine the local motion vectors in their visual field and perform complex motion computations. In this study, we present a novel implementation of a silicon retina based on the Adelson-Bergen spatio-temporal energy model of primate cortical cells. By employing a multi-chip strategy, we successfully implemented the model without much sacrifice of the fill factor of the photoreceptors in the front-end chip. In addition, the characterization results proved that this spatio-temporal frequency tuned silicon retina can detect the direction of motion of a sinusoidal input grating down to 10 percent contrast, and over more than a magnitude in velocity. This multi-chip biomimetic vision sensor will allow complex visual motion computations to be performed in real-time.

1 Introduction

Every organism in nature must struggle to survive within its perceived truth by the help of its senses. What organisms perceive is only a noisy flow of sensation and this does not necessarily represent the perfect truth of their environment, but in fact provides information on their indispensable sensed truth. The biological strategies they have adopted have been proven to be effective and reliable by evolution. In this respect, when building artificial vision systems these biological models can be taken as a reference to deal with real life tasks such as target tracking and object avoidance. These systems employ massively parallel processing in dealing with complex and dynamic visual tasks and by incorporating the same strategy with parallel VLSI design principles in artificial systems, reliable low-power real-time neuromorphic systems can be built.

In this study we present a multi-chip system implementation of an analog VLSI hardware visual motion sensor based on the Adelson-Bergen motion energy model [1]. The Adelson-Bergen spatiotemporal energy model is a biological model and is often used as a model to describe primate cortical complex cells [2],[3]. In addition, it is classified under the correlation-based algorithms which are utilized to explain the optomotor response in flies [4] and direction selectivity in a rabbit's retina [5].

Multi-chip implementations are very suitable especially to achieve 2D optical flow motion computation [8]. These kind of systems integrate analog circuitry with an asynchronous digital interchip communication which is based on the Address-Event Representation (AER) protocol proposed by Mahowald [9] and revised by Boahen [13]. We utilized the same protocol in our implementation of the Adelson-Bergen algorithm to achieve the communication between sender and receiver chips.

A software version of the Adelson-Bergen algorithm was implemented on general-purpose analog neural computer by Etienne-Cummings [6]. Later, Higgins and Korrapati [7] implemented a monolithic analog VLSI sensor based on this algorithm. The main reason why we employed the Adelson-Bergen algorithm in motion computation is that it responds to real-world stimuli better than other algorithms and is therefore more amenable for robotic applications. Here we show that by incorporating the advantages of the model with a modular strategy the computational overload in the front-end chip can be reduced noticeably and the fill factor of the photoreceptors can be kept high. In addition the analog hardware implementation of the model exploits the subthreshold behavior of MOS transistors and therefore consumes little power. Furthermore, it works in real time, adapts to its environment and utilizes a massively parallel processing biological strategy by aligning in parallel arrays in the silicon retina.

2 Algorithm

The Adelson-Bergen algorithm is a spatiotemporal frequency tuned algorithm which obtains its direction selectivity by integrating quadrature filters with a nonlinearity. This algorithm is used in motion computation to extract the Fourier energy in a band of spatiotemporal frequencies regardless of phase of the stimulus. In this study, the Adelson-Bergen algorithm has been implemented (in a modified form; see Figure-1) by making use of Neuromorphic principles [10] and simplified without modifying the basic idea of spatiotemporal energy model.

Firstly, the spatial filtering in the model is trivialized by simply taking photoreceptor outputs separated by a $\Delta\Phi$ spatial distance between adjacent photoreceptors. Secondly, temporal filters in the model are implemented by employing an integrating circuit. Here we demonstrate that the integrating receiver circuitry can be used to attain a phase difference close to 90 degrees. This novel technique of using an integrator instead of a low-pass filter as a temporal filter enables us to exploit the advantages of multi-chip strategy in motion computation and to decrease the computational overload. Finally, the nonlinearity required to realize the algorithm is attained in the implementation by making use of the mathematical properties of rectification.

3 Hardware Architecture

In this section, the hardware architecture of the Adelson-Bergen algorithm is explained in detail. The overall organization of the multi-chip system is illus-

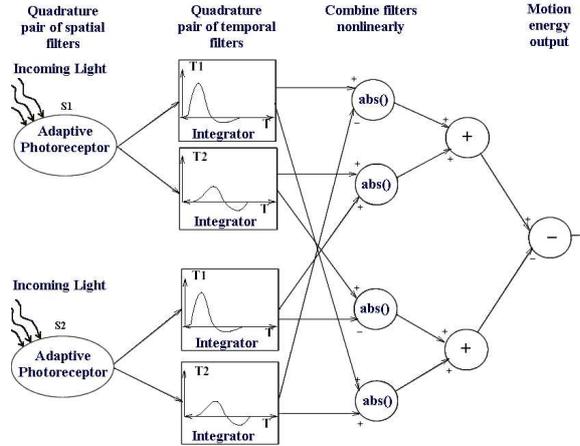


Fig. 1. Modified Adelson-Bergen spatiotemporal energy model

trated in Figure-2. The system consists of two chips, namely sender and receiver chips. These chips communicate by using the AER protocol [13] (Figure-4a) and perform all computations in current mode to minimize the space and maximize the fill-factor of the front-end chip.

In this implementation, the sender chip is employed to sense the intensity changes in the environment with respect to its adapted background and send this information to the receiver chip. It includes an array of pixels that discretize space and incorporates a photoreceptor, transconductance amplifier, rectifier and AER interface circuitry. In order to sense the intensity changes in the environment we used adaptive photoreceptors (Figure-3a) by Liu [11]. This photoreceptor provides a continuous-time signal that has a low gain for static signals and a high gain for transient signals. In addition to the photoreceptor output, an internal feedback voltage in this circuit is utilized in the motion computation to obtain information on the change of the adapted background level of its environment. This feedback voltage represents the running average of the illumination level of the background and the time interval of the averaging operation can be changed by altering the adaptation time. In the next stage, the response of the photoreceptor is compared with its feedback voltage and converted to current by making use of a transconductance amplifier shown in Figure-3b. In this way, we obtain a bandpass characteristic from the photoreceptor and transconductance amplifier pair. This characteristic ensures that very high frequencies are attenuated and offset or in this case the background level, is removed. After that, the output current of the transconductance amplifier is rectified by utilizing a full-wave rectifier (Figure-3c) in order to acquire separate currents for negative and positive intensity changes. Lastly, the communication interface circuitry that sends these intensity changes to the corresponding pixels

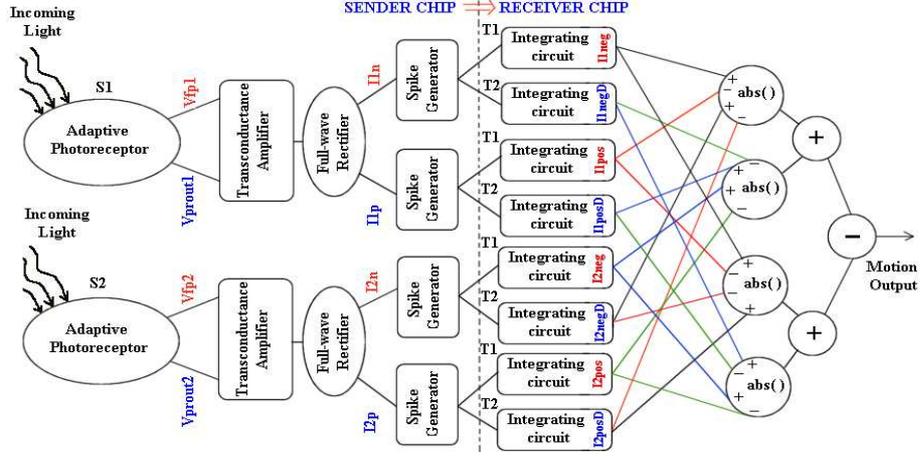


Fig. 2. The multi-chip implementation of the Adelson-Bergen algorithm. V_{prouit} and V_{fb} represent the photoreceptor's output and feedback response, respectively. In addition, I_p and I_n refer to positive and negative parts of the rectified signal, and in the receiver part, I_{pos} and I_{neg} represent the integrated versions of these signals. Lastly, I_{posD} and I_{negD} are the signals that are delayed relative to I_{pos} and I_{neg} .

in the receiver chip is implemented. This interface circuitry generates spikes with a frequency proportional to the amplitude level of the intensity change.

The receiver chip is utilized to achieve small field motion computation by making use of the information obtained from the sender chip. It is composed of the AER communication circuitry and corresponding pixels of the sender chip. In this implementation, the integrating circuit is employed not only to integrate the incoming spikes but also to attain necessary delays. The integrating receiver circuit has a particular spike frequency f_0 for which the charging and discharging current are the same on average keeping output voltage at a steady state value. In Figure-4c, the relationship between the voltage output and incoming spike frequency is illustrated. Similar to low-pass characteristics, this circuit can be incorporated to obtain the necessary delay and is therefore amenable to be used as a temporal filter. Accordingly, the positive or negative part of the signal is integrated by two integrating circuits tuned to different temporal frequencies in order to obtain a similar configuration as it was achieved with monolithic implementation [7]. As a result, the integrated positive and negative parts of the signal are summed, subtracted and absolute valued to perform the motion computation illustrated in Figure-2. The formulation of the final output is shown below.

$$I_{out} = |I_{1neg} + I_{2negD} - I_{1pos} - I_{2posD}| + |I_{2neg} + I_{1posD} - I_{1negD} - I_{2pos}| \\ - |I_{1neg} + I_{2posD} - I_{1pos} - I_{2negD}| - |I_{2neg} + I_{1negD} - I_{2pos} - I_{1posD}| \quad (1)$$

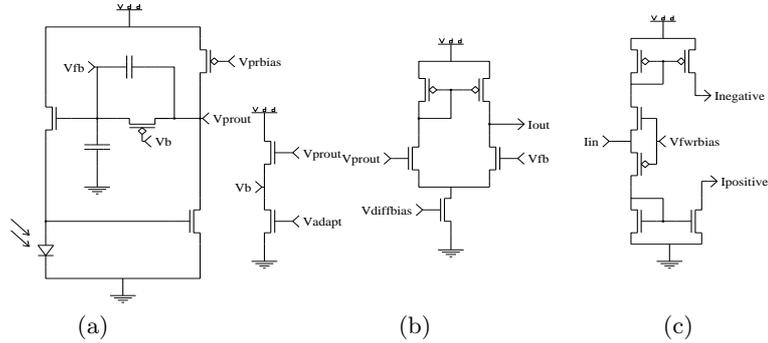


Fig. 3. (a) Adaptive photoreceptor circuit by Shih-Chii Liu. (b) Transconductance amplifier circuit. (c) Full-wave rectifier circuit.

4 Characterization Results

In this section, characterization results of the multi-chip implementation are presented. The experiments are performed by using computer-generated sinusoidal grating stimuli on an LCD screen. In the experiments, each time one parameter is changed the others are held constant. In order to remove the phase dependence of the sensor and prevent artifacts, output voltages are averaged over 10 temporal periods of stimuli. The output of the sensor is obtained in current mode and converted to voltage by utilizing current sense amplifier with a 3.9 megohm feedback resistor.

The first experiment is performed by using sinusoidal grating stimuli to test the direction selectivity of the sensor for preferred, null, orthogonal motion, and no-motion cases. As can be seen in Figure-5a, the mean response of the sensor quite clearly indicates the direction of the stimulus. In the second experiment, it is proved that the sensor shows sinusoidal dependence to orientation sweep of sinusoidal grating (Figure-5b) as expected from the theoretical results of Adelson-Bergen algorithm. At 90 degrees the motion output gives its positive peak response and at 270 degrees the sensor output reaches its negative maximum. In the last experiment, the sensor is tested to acquire the spatial and temporal frequency characteristics of the sensor. The response of the sensor to a temporal frequency sweep is shown in Figure-6a. The output of the sensor peaks at around 1 Hz and, as is obvious from the temporal frequency response, the sensor responds to a velocity range of more than one order of magnitude. These responses justify the use of integrating circuit as a temporal filter in the motion computation. The response of the sensor to a spatial frequency sweep is illustrated in Figure-6b. The plot of the multi-chip sensor peaks around 0.4 cycles/pixel and the sensor shows a strong spatial aliasing around 0.9 cycles/pixel. Lastly, in Figure-6c, the spatiotemporal response of the sensor is illustrated. This plot shows the mean output of the model in response to sinusoidal gratings varying in both spatial and temporal frequency. The mean output is plotted for spatial frequencies on the X-axis versus temporal frequencies on the Y-axis. It

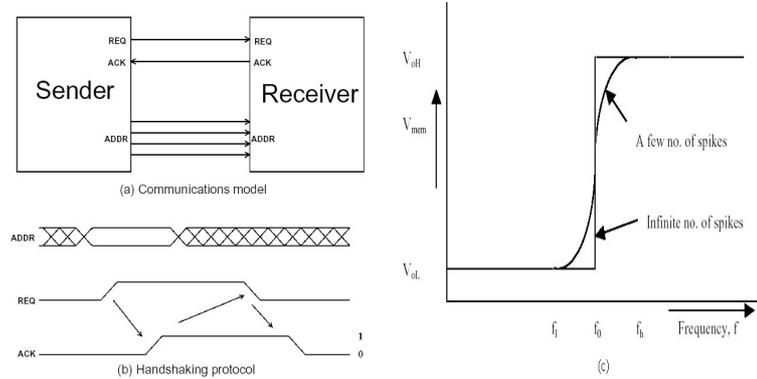


Fig. 4. AER protocol summary. (a) The model for AER transmission: a sender chip communicates with a receiver chip via request, acknowledge and address lines. (b) The handshaking protocol for transmission using the control and address lines: a request with a valid address leads to an acknowledgment, which in turn leads to falling request and falling acknowledge. (c) V_{out} vs. frequency sketch.

is obvious from the graph that the model responds best to a particular spatiotemporal frequency for which it is tuned and the response decreases at other frequencies.

5 Discussion

We have described and characterized a novel spatiotemporal frequency tuned multi-chip analog VLSI motion sensor and presented a new technique to realize the temporal filters needed for motion computation. The characterization results clearly elucidate the fact that by using this technique, we can obtain a reliable, low-power and real time multi-chip Neuromorphic motion processing system while retaining many of the advantages of monolithic implementation.

The multi-chip sensor responds to optimal spatial frequencies over a velocity range of more than an order of magnitude. Since the motion sensor is tuned to spatiotemporal frequencies, the main concern is to increase the range of the spatial and temporal frequencies and the contrast level to which it can respond. As seen from the spatiotemporal plot (Figure-6c), the response range of the multi-chip motion sensor is increased in comparison to monolithic implementation by Higgins and Korrapati [7]. The area of this region increased two times compared to the monolithic implementation. Besides, the immunity of the sensor to low contrast levels is improved. These results clearly indicate the improvement in the performance while the motion computation capability of the system is vastly increased compared to any space limited monolithic system. Furthermore, 2D motion computation can be easily achieved with this implementation by using a single sender chip with two two receivers and manipulating the x and y connections of the address lines of the second receiver.

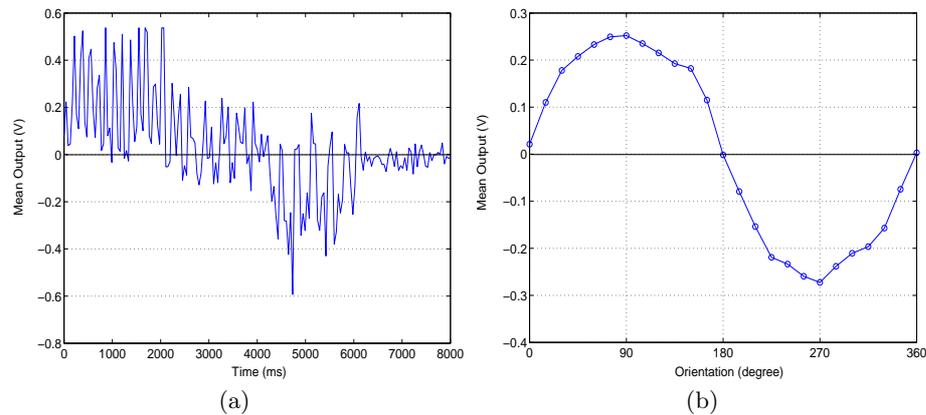


Fig. 5. (a) Output of the motion sensor. In the interval 0-2 seconds, a preferred direction sinusoidal stimulus is presented. Between 2-4 seconds, a sinusoidal stimulus moving orthogonal to the sensor orientation is presented. After that, sensor is exposed to a null direction sinusoidal stimulus. Lastly, no stimulus is presented between 6-8 seconds. (b) Orientation sweep. A sinusoidal stimulus is presented at varying directions relative to the motion sensor, which is optimally oriented for a stimulus at 90 degrees.

By using a multi-chip system, we obtained 36 pixels in a standard $1.5\mu\text{m}$ CMOS process and 2mm-by-2mm die size. In a new designed sender chip, positive and negative parts of the rectified signal will be sent through with a novel technique implemented in the AER system. With this implementation, we are expecting to exceed the number of pixels that can currently be realized on a single chip.

References

1. E.H. Adelson and J.R. Bergen: "Spatiotemporal energy models for the perception of motion," *J. Opt. Society of America*, vol. 2, no. 2,(1985), pp. 284-299
2. S.J. Nowlan and T.J. Sejnowski: "Filter selection model for motion segmentation and velocity integration," *J. Opt. Soc. America*, vol. 11, no. 12,(1994), pp. 3177-3200
3. DJ Heeger, EP Simoncelli, and JA Movshon: "Computational models of cortical visual processing," *Proc. Natl. Acad. Sci.*, vol. 93,(1996), pp. 623-627
4. Hassenstein, B. and W. Reichardt: "Systemtheoretische analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*. Z." *Naturforsch.*, (1956), 11b: 513-524
5. Barlow, H.B. and W.R. Levick: "The mechanism of directionally selective units in the rabbit's retina." *J. Physiology*, (1965), 178: 447-504
6. R Etienne-Cummings, J. Van der Spiegel, and P. Mueller: "Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters," *IEEE Transactions on Circuits and Systems-II*, vol. 46, no. 9, (1999), pp. 1121-1136
7. Higgins, C.M. and S. Korrapati: "An analog VLSI motion energy sensor based on the Adelson-Bergen algorithm." In *Proceedings of the International Symposium on Biologically Inspired Systems*, (2000)

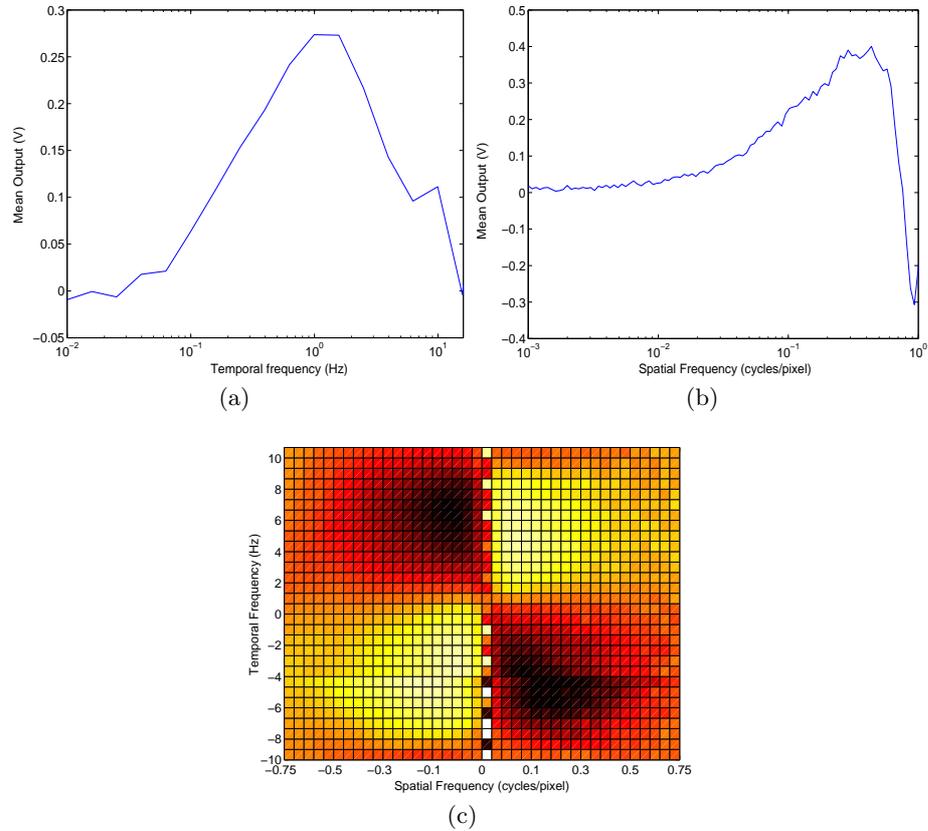


Fig. 6. Spatio-temporal frequency tuning of the sensor. (a) Temporal frequency sweep. (b) Spatial frequency sweep: note the onset of spatial aliasing. (c) Spatio-temporal frequency plot: light colors indicate positive and dark colors indicate negative average response.

8. C. M. Higgins and C. Koch: "A Modular Multi-Chip Neuromorphic Architecture for Real-Time Visual Motion Processing," *Analog Integrated Circuits and Signal Processing* 24(3), (September, 2000), pp 195-211
9. M.A. Mahowald: "VLSI analogs of neuronal visual processing: A synthesis of form and function." PhD thesis, Department of Computation and Neural Systems, California Institute of Technology, Pasadena, CA., (1992)
10. C. A. Mead: "Neuromorphic electronic systems," *Proceedings of the IEEE*, vol. 78, (1990), pp. 1629-1636
11. S-C. Liu: "Silicon retina with adaptive filtering properties," *Analog Integrated Circuits and Signal Processing*, 18 (2/3), (1999), pgs 1-12
12. Mead, C.A.: "Analog VLSI and Neural Systems." Addison-Wesley, Reading, (1989)
13. K. Boahen: "A throughput-on-demand 2-D address-event transmitter for Neuromorphic chips," in *Proc. Of the 20th Conference on Advanced Research in VLSI*, Atlanta, GA, (1999)