

An Analog VLSI Motion Energy Sensor Based on the Adelson-Bergen Algorithm

Charles M. Higgins and Sudhir Korrapati
Neuromorphic Vision and Robotic Systems Laboratory
Department of Electrical and Computer Engineering
The University of Arizona
Tucson, AZ 85721
`[higgins,sudhir]@ece.arizona.edu`

**International ICSC Symposium
on Biologically-Inspired Systems**
December 12-15, 2000

An Analog VLSI Motion Energy Sensor Based on the Adelson-Bergen Algorithm

Charles M. Higgins and Sudhir Korrapati
Neuromorphic Vision and Robotic Systems Laboratory
Department of Electrical and Computer Engineering
The University of Arizona
Tucson, AZ 85721
[higgins,sudhir]@ece.arizona.edu

Abstract

The low-level representation of visual motion utilized by biological organisms from insects to primates is fundamentally different from that used in conventional computer vision systems. Rather than an optical flow vector field, banks of nonlinear spatio-temporal frequency tuned filters are used, giving rise to a representation which naturally supports transparent motion and shear at occlusion boundaries, and which facilitates the solution of the aperture problem. Because of the amount of parallel computation required to detect motion in this fashion, a biological representation of motion requires efficient implementation of elementary motion detectors. In this paper we describe the low-power continuous-time analog VLSI implementation of a biologically-inspired visual motion sensor. This sensor, based on the motion energy algorithm of Adelson and Bergen, is a simple hardware model of the motion response of a primate cortical complex cell but can also be shown to be equivalent to the Reichardt model of insect motion detection. Characterization results show that this spatio-temporal frequency tuned sensor can discriminate the direction of motion of a sinusoidal grating down to less than 5% contrast, and over more than an order of magnitude in velocity. In addition to providing a real-time hardware biological model for investigation of spatial motion integration algorithms, this sensor will be a fundamental building block for experimentation into biologically-inspired visual navigation architectures for underwater, airborne, and land-based autonomous robots.

1 INTRODUCTION

The architecture of biological visual systems has something to teach conventional engineering about system-level design. Representations of visual modalities including depth, motion, color, and form are quite unlike those employed by conventional computer vision systems. And yet biological organisms trivially accomplish visual tasks such as object recognition, obstacle avoidance, and target tracking which continue to challenge artificial systems. By building artificial vision systems true to the representations used by neurobiology, we may be able to suggest testable hypotheses as to how

biological systems accomplish these tasks, while at the same time producing novel highly capable artificial systems. However, the primary strategy used by neurobiology is massively parallel processing. In order to make effective use of such a strategy in artificial systems, an efficient parallel implementation must be conceived.

This paper describes an analog VLSI hardware visual motion sensor based on the Adelson-Bergen motion energy model [1], which is often cited as a basic model of primate cortical complex cells [2, 3, 4]. This model can be shown to be equivalent to the Reichardt [5] model of insect motion detection, and thus is applicable over a wide range of taxa. The implementation is continuous in time, like the biological system that it models. Because of the use of analog circuitry with subthreshold MOS transistors, the power consumption of this implementation is very low. The sensor is designed to be fabricated in parallel arrays, and thus is ideal for experimentation with biological representations of motion.

2 RELATED WORK

Because low-level visual motion processing is very well matched to continuous-time fully parallel focal plane arrays, a large number of integrated sensors of this type have been fabricated [6, 7, 8, 9, 10, 11, 12, 13] based on a variety of algorithms. Etienne-Cummings et al. [14] have implemented a large-scale version of the Adelson-Bergen algorithm on a general-purpose analog neural computer.

The most closely related sensors to the present work are those of Deutschmann et al. [12] and Harrison et al. [11]. While the sensor of Deutschmann et al. is based on a multiplication of temporal and spatial derivatives, and the sensor of Harrison et al. is based on an elaborated version of the Reichardt model of fly motion detection, both sensors are spatio-temporal frequency tuned, based almost purely on analog circuitry, and have the same basic response properties shown in this

paper. However, both sensors are quite different from the sensor presented here both in algorithm and implementation, and thus may have different advantages and disadvantages.

3 ALGORITHM

3.1 Canonical Form

In a seminal 1985 paper [1], Adelson and Bergen published a model of the activity of a primate complex cell in response to a motion stimulus. They showed how it was possible to build a direction-selective unit by combining linear quadrature filters with a nonlinearity. Such models are responsive to the Fourier energy in a band of spatio-temporal frequency, independent of the phase of the stimulus, and are thus referred to as motion energy models.

The one-dimensional model, shown in Figure 1, is constructed by combining the responses of four simple-cell models. The spatial impulse responses of each quadrature pair of simple cells are modeled as Gabor functions with identical spatial receptive fields but quadrature relative phase. Because visual motion involves changes in both space and time, two temporal impulse responses must be considered, as well. Adelson and Bergen modeled these as second and third derivatives of Gaussians, but the primary criterion is that the temporal responses have quadrature phase at the temporal frequency of interest.

In order to compute motion energy, it is necessary to combine each spatial impulse response with both temporal impulse responses as shown in Figure 1. By making sums and differences of linear separable spatiotemporal filters and squaring their outputs, it is possible to synthesize a nonseparable motion energy filter.

3.2 Modified Form

In order to implement this algorithm in analog VLSI hardware, it was necessary to modify the canonical algorithm somewhat.

In the hardware implementation, the spatial quadrature pair of filters is approximated by using neighboring photoreceptors, each of which has an antagonistic center-surround spatial impulse response very similar to a Gabor function. The width of these spatial impulse responses can be adjusted to approximate the quadra-

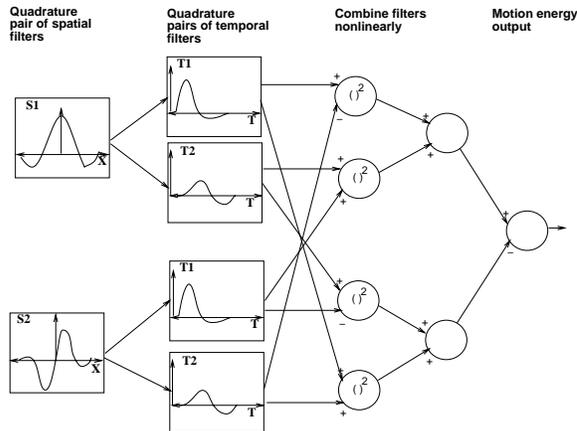


Figure 1: Adelson-Bergen motion energy model. Each of a quadrature pair of spatial filters (at the same spatial location) is filtered with a quadrature pair of temporal filters. By combining these linear spatiotemporal filters in the correct order and squaring their outputs, it is possible to synthesize a nonseparable motion energy filter.

ture Gabor functions used by Adelson and Bergen. In fact, in this implementation the Gabor function shape is not necessary for direction-selectivity as long as the DC value of the intensity signal is removed. The spatial frequency tuning obtained is actually a result of the phase relationship caused by the photoreceptor spacing.

The simplest possible pair of temporal filters required to obtain direction-selectivity would create a phase difference of as close to 90 degrees as possible at the temporal frequency of interest. In the current implementation, a single-pole lowpass filter is used to create the required phase delay. That is, one temporal pathway is unfiltered, and the other passes through a temporal lowpass filter.

Computing the mathematical square of a signal may be considered as rectification followed by an expansive nonlinearity. In this implementation, we have employed the rectification but neglected the expansive nonlinearity, replacing the square in Figure 1 with an absolute value.

3.3 Expected Response

Taking into account the modified Adelson-Bergen algorithm, it is possible to derive the output of the sensor in response to a moving sinusoidal grating. If the input is taken to be a two-dimensional sinusoidal grating with maximum amplitude A , contrast C , spatial fre-

quency f_s , orientation θ relative to the sensor preferred orientation, and temporal frequency f_t ,

$$I(x, y, t) = A(1 + C \sin(2\pi f_t \cdot t + 2\pi f_s (\cos \theta \cdot x + \sin \theta \cdot y)))$$

then the output of an X-oriented sensor with photoreceptor separation δ can be computed as

$$A^2 P(f_t)^2 C^2 \cdot \sin(2\pi f_s \delta \cos \theta) \cdot H(f_t) \cdot \sin \phi_t(f_t)$$

where $P(f_t)$ is the magnitude of the photoreceptor temporal frequency response, and $H(f_t)$ and $\phi_t(f_t)$ are respectively the magnitude and phase of the lowpass filter's response. The squaring operation has been used instead of the absolute value in order to obtain a closed-form solution.

From this formula we expect a square-law response to contrast variation and a sinusoidal response to the variation of stimulus orientation, which gives a large positive response for the preferred orientation, a large negative response for the null orientation, and zero responses for orthogonal orientations.

We expect the spatial frequency tuning to peak at the frequency at which $2\pi f_s \delta \cos \theta = \pi/2$. Expressing spatial frequency in cycles/pixel this can be written $f_s \cos \theta = 0.25$ cycles/pixel. Because of the Nyquist sampling theorem, we expect the sensor to spatially alias at $f_s = 0.5$ cycles/pixel.

The temporal frequency tuning can be written from the expression above as $P(f_t)^2 \cdot H(f_t) \cdot \sin(\phi_t(f_t))$. For a single-pole lowpass filter, the product of magnitude and sine of phase can be shown to be symmetric in log frequency, and peak at the cutoff (3 dB) frequency of the lowpass filter. Because the photoreceptor frequency response has much wider bandwidth than the rest of the expression, it will form an envelope for temporal frequency response as we vary the temporal tuning. Thus we expect the temporal frequency tuning to peak at the cutoff frequency of the lowpass filter and show an envelope corresponding to the photoreceptor temporal frequency tuning.

4 HARDWARE ARCHITECTURE

The hardware implementation of the Adelson-Bergen algorithm is diagrammed in Figure 2. All of the circuitry shown operates the MOSFETs below threshold to make use of the exponential current-voltage relationship. The spatial filters are implemented by using Delbrück adaptive photoreceptors [15] coupled with a Liu-Boahen diffuser network [16] (shown in Figure 3(a)),

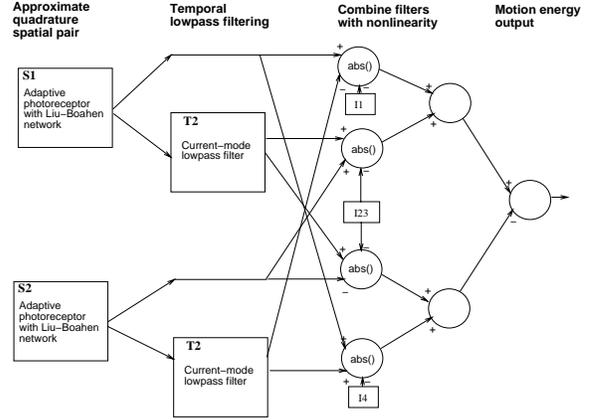


Figure 2: Analog circuit implementation of Adelson-Bergen motion energy model. Spatial filtering of illumination signals is accomplished with adaptive photoreceptors and a diffuser network. Temporal filtering is done with a current-mode lowpass filter. Sums and differences are easily accomplished in current mode.

which results in an antagonistic center-surround spatial impulse response very similar to a Gabor function. The width of this spatial impulse response can be adjusted by the use of biases. The output of the photoreceptor is taken as a current, in order to use current-mode computation for the rest of the circuit. Current-mode computation was chosen due to the large number of sums and differences required to implement the desired algorithm. One disadvantage of current-mode computation is the necessity to copy (mirror) a current in order to use it more than once in the computation. Note that the photoreceptor itself has a bandpass temporal characteristic which limits the temporal frequency response of the motion circuit as a whole.

The temporal current-mode lowpass filter is implemented as shown in Figure 3(b). Neglecting the Early effect, the current signal coming into the circuit is reduced by a multiplicative factor of e^{-V_{iau}/V_T} where V_T is the thermal voltage. The signal is then passed through a current mirror with capacitor C , which in the small-signal approximation is a single-pole lowpass filter with time constant

$$\tau = \frac{C r_o}{1 + g_m r_o}$$

in seconds/radian where r_o is the small-signal output resistance of the (nominally identical) current mirror transistors, and g_m is their transconductance. Clearly this time constant depends on the DC current level seen by the current mirror, and thus V_{iau} controls the time constant. Again neglecting the Early effect, the final output mirror multiplies the filtered signal by e^{V_{iau}/V_T} ,

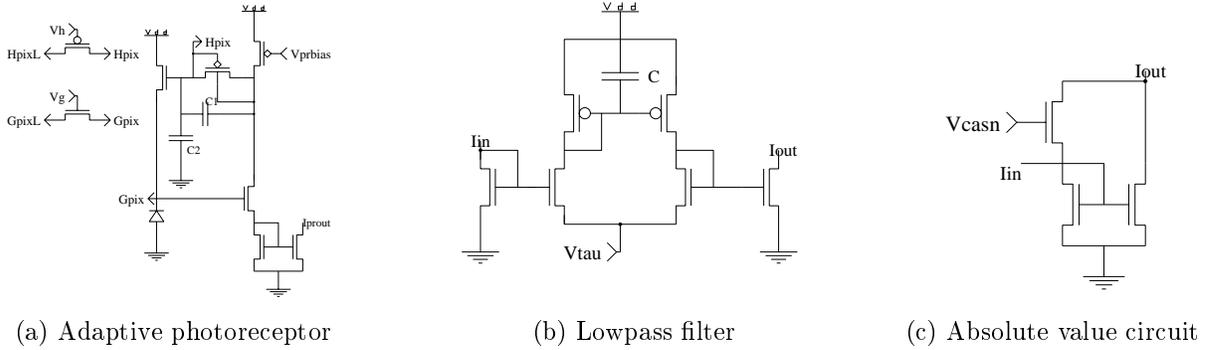


Figure 3: Circuits used in the implementation. (a) Adaptive photoreceptor: this circuit transduces light into a current, and uses diffusers to create an antagonistic center-surround receptive field. (b) Current-mode lowpass filter: the input current is reduced in magnitude, processed through a ‘slow’ current mirror, and return to its original magnitude. (c) Absolute value circuit: a bidirectional input results in a rectified current at the output.

returning it to its original amplitude range. This circuit is clearly not a linear filter for large signals, but is quite compact and suffices for our application.

Due to the current-mode nature of the computation, the sums shown in Figure 2 are accomplished by wiring together the appropriate signals. Differences are accomplished by the use of two-transistor current mirrors. The current-mode absolute value circuit is shown in Figure 3(c), and was inspired by the above-threshold circuits of Bult and Wallinga [17]. This circuit takes a bidirectional current input. Currents going into the circuit are passed through the current mirror (reversing their direction) and coupled to the output current. Currents leaving the circuit are pulled directly from the output current through the V_{casn} transistor.

Directly before the computation of the nonlinearity, it is essential that the signals have no DC component for maximum direction selectivity. In this circuit, it was necessary to subtract off manually-tuned offset currents to accomplish this task, shown in Figure 2 as I_1, I_{23} , and I_4 . These compensation currents required a single transistor each with an off-chip bias.

In total, the motion sensor implementation required 41 transistors and 3 capacitors, not including scanout circuitry. At the bias settings used for data collection, the pixel circuitry draws less than $32 \mu A$ at 5 volts.

5 CHARACTERIZATION RESULTS

This section details the characterization of a single motion sensor. Bias settings are held constant for all experiments unless otherwise stated. Although the sensor

directly outputs a current, output quantities are quoted in millivolts from a current sense amplifier with a 4.7 megohm feedback resistor. Output voltages are averaged over 10 temporal periods of the stimulus to remove the phase dependence of the sensor. When a parameter is swept, all other parameters are held constant at a standard value. All experiments were performed by presenting computer-generated sinusoidal grating stimuli on an LCD screen.

Figure 4 shows the output of the sensor for four stimulus conditions: no input, orthogonal sinusoid grating, preferred and null direction sinusoid gratings. The raw output of the sensor is plotted lightly in the background, and shows large phase-dependent oscillations. The filtered output of the sensor is plotted boldly in the foreground and shows that the output for no stimulus and an orthogonal stimulus have approximately the same average value, and that preferred and null direction stimuli can be clearly resolved.

Figure 6 shows the spatio-temporal frequency tuning of the sensor. In (a), we see as expected that the strongest response is at a particular combination of spatial and temporal frequency. The sensor is direction selective because of the difference in its response to positive and negative temporal frequencies. Parts (b) and (c) show sections through the plot of part (a), but with points logarithmically spaced. The spatial frequency sweep of (b) shows a peak at 0.15 cycles/pixel and the onset of spatial aliasing (reporting the wrong direction) at approximately 0.4 cycles/pixel. While the ideal tuning would be at 0.25 cycles/pixel with aliasing at 0.5 cycles/pixel, the relative delay between the two analog photoreceptor circuits combined by the sensor shifts these frequencies in the real circuit. This relative delay

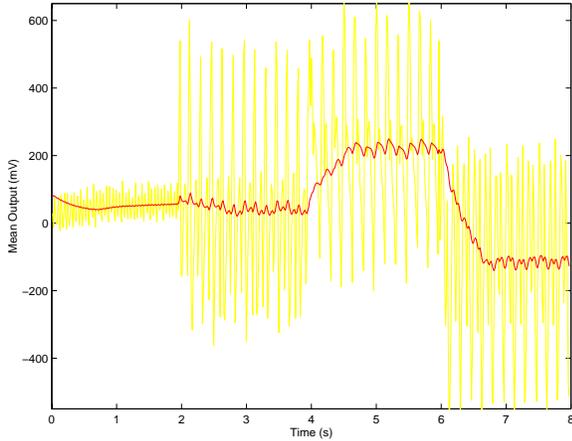


Figure 4: Output of the motion sensor. Raw output is shown lightly in the background, filtered output boldly in the foreground. For the interval 0-2 seconds, no stimulus is presented. Flicker response to fluorescent lighting is seen. Between 2-4 seconds, a sinusoidal stimulus moving orthogonal to the sensor orientation is presented. Between 4-6 and 6-8 seconds respectively, a preferred and null direction sinusoidal stimulus is presented.

elongates the spatio-temporal frequency tuning shown in part (a), stretching it along the constant-velocity direction. The temporal frequency plot of (c) shows a peak at 6 Hz. With a fixed spatial frequency, the sensor is responsive to a velocity range of more than one order of magnitude.

Figure 5 shows the response of the sensor as the orientation of the stimulus is varied 360 degrees. The expected sinusoid shape is seen, with an asymmetry due to mismatch in the circuitry. Figure 7 shows the range of temporal frequency tuning that can be achieved by varying the bias V_{tau} . As predicted, the envelope of this set of curves is the photoreceptor frequency response. The tuning frequency can be varied over the entire range allowed by the photoreceptor. The response to varying contrast for both preferred and null directions of motion is shown in Figure 8. We see the expected square-law response to contrast except at very high contrasts, where the curve saturates. It is possible to distinguish the direction of motion even at contrasts less than 5%.

6 DISCUSSION

We have described and characterized a novel spatio-temporal frequency tuned analog VLSI motion sensor. The sensor as fabricated responds to optimal spatial frequencies over a velocity range of more than an or-

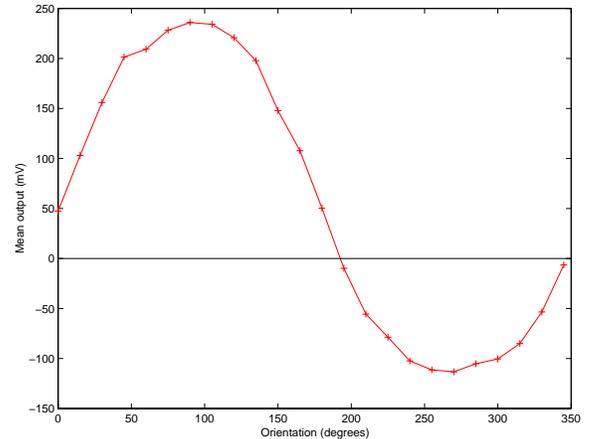


Figure 5: Orientation sweep. A sinusoidal stimulus is presented at varying directions relative to the motion sensor, which is optimally oriented for a stimulus at 90 degrees.

der of magnitude, and can discriminate direction down to less than 5% contrast. The spatial frequency tuning is determined by the optics used to focus an image onto the chip, and can be varied over a factor of at least a decade. The temporal frequency tuning can be controlled by bias settings over a range of more than a decade, limited by the photoreceptor frequency response.

While the success of the first implementation of this sensor proves the applicability of the algorithm used, there are a number of drawbacks to the current circuit. The worst of these is the necessity to compensate for the DC value of the sums before they enter the absolute value stage. These compensating biases are very sensitive, and in the current implementation the optimal bias settings for neighboring pixels are quite different. These biases also make the circuit sensitive to temperature variations and variations in the average illumination seen by the photoreceptors. A second problem is in the distribution of currents between pixels, which requires multiple current mirrors to copy current-mode signals. This circuit has already been redesigned and submitted for fabrication with two major changes. The first is the use of a voltage-mode lowpass filter, which allows distribution of voltages rather than currents. The second and most crucial is the adoption of a complementary current representation of each signal, in which two currents represent each signal and can be subtracted to obtain the full signed quantity represented. In this new version, no manual compensation will be required.

Because this motion sensor is an oriented tuned spatio-

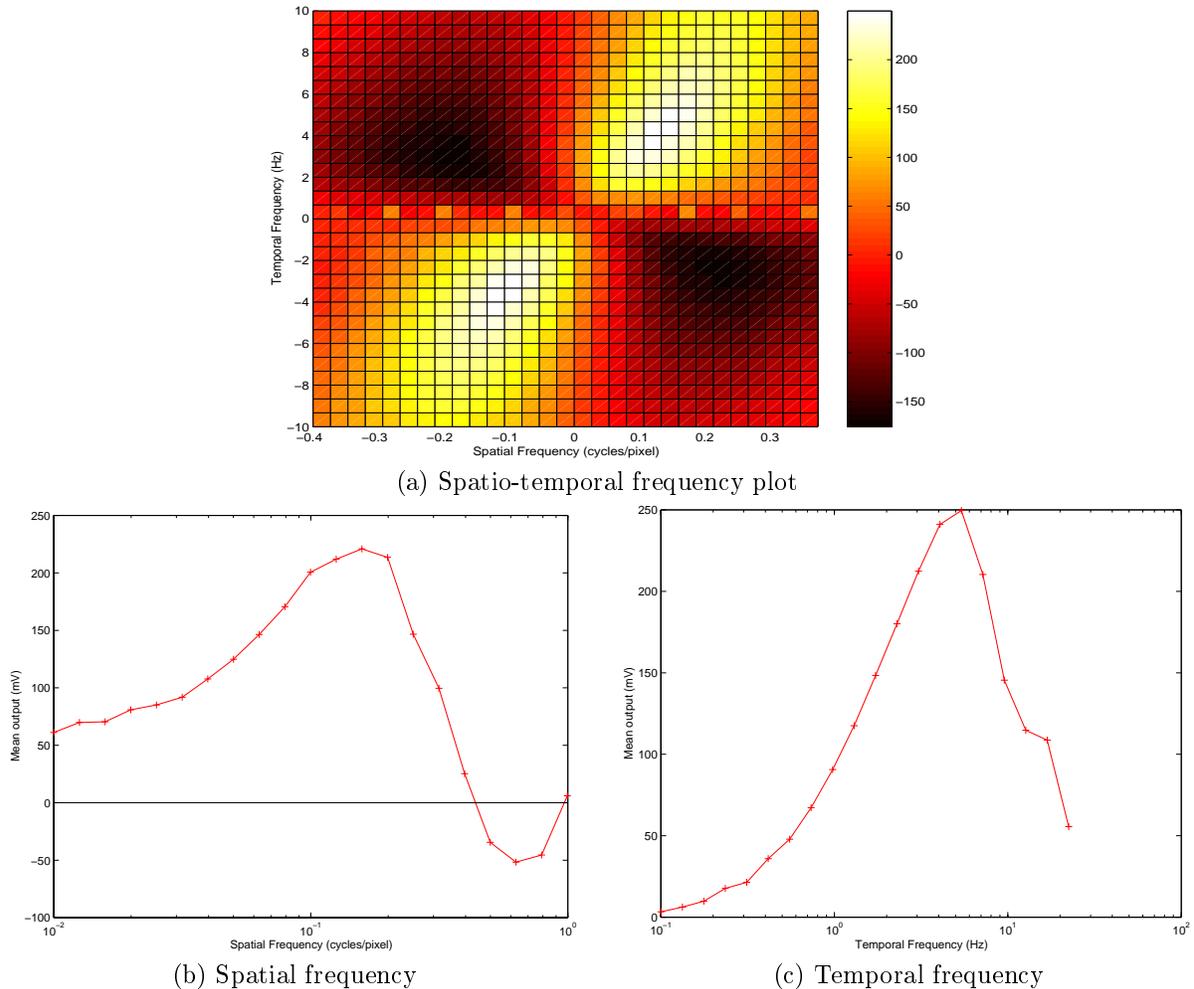


Figure 6: Spatio-temporal frequency tuning of the sensor. (a) Spatio-temporal frequency plot: light colors indicate positive average response and dark colors indicate negative. (b) Spatial frequency sweep: note the onset of spatial aliasing. (c) Temporal frequency sweep.

temporal filter, it cannot respond to a wide range of spatial frequencies, orientations or velocities, and thus must be considered as part of a filter bank approach to system-level visual motion processing. Design is underway to incorporate this sensor into a biologically-inspired multi-chip implementation with a single retina front end and multiple parallel motion-sensitive filters. This multi-chip motion processing system will be highly sensitive to motion over a wide range of stimulus parameters.

Acknowledgments

This research was begun by CMH while a postdoctoral research fellow in the laboratory of Christof Koch at the

California Institute of Technology, where he was supported by the Caltech Center for Neuromorphic Systems Engineering as a part of the National Science Foundation’s Engineering Research Center program, as well as by the Office of Naval Research. At the University of Arizona, this work is supported by the Office of Naval Research under grant number N00014-97-1-0970.

References

- [1] E.H. Adelson and J.R. Bergen, “Spatiotemporal energy models for the perception of motion,” *J. Optical Society of America A*, vol. 2, no. 2, pp. 284–299, 1985.
- [2] N. Qian, RA Andersen, and EH Adelson, “Transparent motion perception as detection of unbalanced motion signals. III. modeling,” *J. Neuroscience*, vol. 14, no. 12, pp. 7381–7392, 1994.

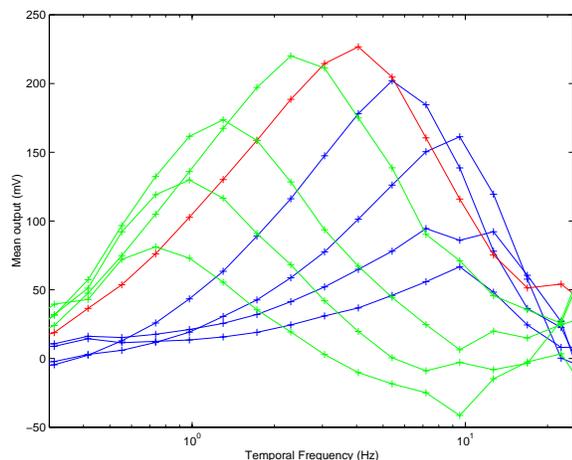


Figure 7: Variation of temporal frequency tuning. By changing the bias V_{tau} , it is possible to change the temporal frequency tuning of the sensor. Each curve in this plot is at a different bias setting. The overall envelope of all plots is the photoreceptor temporal frequency response, the low frequency cutoff of which is due to adaptation, and the high frequency cutoff of which is due to the capacitance driven at its output.

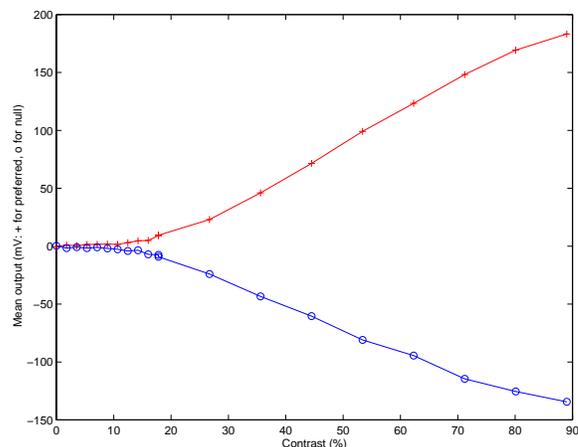


Figure 8: Contrast sweep. As contrast decreases, the mean outputs seen for the preferred and null directions of motion get closer to each other. However, it is still possible to discern the direction of motion even at contrasts less than 5%.

- [3] S.J. Nowlan and T.J. Sejnowski, "Filter selection model for motion segmentation and velocity integration," *J. Optical Soc. America*, vol. 11, no. 12, pp. 3177–3200, 1994.
- [4] DJ Heeger, EP Simoncelli, and JA Movshon, "Computational models of cortical visual processing," *Proc. Natl. Acad. Sci.*, vol. 93, pp. 623–627, 1996.
- [5] Jan P. H. van Santen and George Sperling, "Elaborated Reichardt detectors," *Journal of the Optical Society of America A*, vol. 2, pp. 300–320, 1985.
- [6] J. Tanner and C. Mead, "An integrated analog optical motion sensor," in *VLSI Signal Processing, II*, New York, Ed., pp. 59–76. IEEE Press, 1986.
- [7] T. Horiuchi, J. Lazzaro, A. Moore, and C. Koch, "A delay-line based motion detection chip," in *Advances in Neural Information Processing Systems*, R.P. Lippman, J.E. Moody, and D.S. Touretzky, Eds., pp. 406–412. Morgan Kaufmann, San Mateo, CA, 1991.
- [8] T. Delbrück, "Silicon retina with correlation-based, velocity-tuned pixels," *IEEE Trans. Neural Networks*, vol. 4, pp. 529–541, May 1993.
- [9] R. Etienne-Cummings, J. Van der Spiegel, and P. Mueller, "A focal plane visual motion measurement sensor," *IEEE Trans. on Circuit and Systems I*, vol. 44, no. 1, pp. 55–66, 1997.
- [10] C. M. Higgins, R. A. Deutschmann, and C. Koch, "Pulse-based 2D motion sensors," *IEEE Transactions on Circuits and Systems II*, vol. 46, no. 6, pp. 677–687, June 1999.
- [11] RR Harrison and C Koch, "An analog VLSI model of the fly elementary motion detector," in *Advances in Neural Information Processing Systems 10*, M. I. Jordan, M. J. Kearns, and S. A. Solla, Eds., pp. 880–886. MIT Press, 1998.
- [12] R. A. Deutschmann and C. Koch, "Compact real-time 2-D gradient based analog VLSI motion sensor," in *Proceedings of the Int. Conf. on Advanced Focal Plane Arrays and Electronic Cameras*, Zurich/Switzerland, 1998.
- [13] A. Stocker and R. Douglas, "Computation of smooth optical flow in a feedback connected analog network," in *Advances in Neural Information Processing Systems*, M. S. Kearns, S. A. Solla, and D. A. Cohn, Eds., Cambridge, MA, 1999, vol. 11, MIT Press.
- [14] 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, pp. 1121–1136, 1999.
- [15] T. Delbrück and C. Mead, "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," Tech. Rep. 30, Department of Computation and Neural Systems, California Institute of Technology, 1993.
- [16] S. Liu and K. Boahen, "Adaptive retina with center-surround receptive field," in *Advances in neural information processing systems*, D.S. Touretzky, M.C. Mozer, and M.E. Hasselmo, Eds. 1996, vol. 8, MIT Press.
- [17] K Bult and H Wallinga, "A class of analog CMOS circuits based on the square-law characteristic of a MOS transistor in saturation," *IEEE Journal of Solid-State Circuits*, vol. SC-22, no. 3, pp. 357–365, 1987.