

Analog VLSI Implementation of Spatio-Temporal Frequency Tuned Visual Motion Algorithms

Charles M. Higgins, *Senior Member, IEEE*, Vivek Pant, and Rainer Deutschmann

Abstract—The computation of local visual motion can be accomplished very efficiently in the focal plane with custom very large-scale integration (VLSI) hardware. Algorithms based on measurement of the spatial and temporal frequency content of the visual motion signal, since they incorporate no thresholding operation, allow highly sensitive responses to low contrast and low-speed visual motion stimuli. We describe analog VLSI implementations of the three most prominent spatio-temporal frequency-based visual motion algorithms, present characterizations of their performance, and compare the advantages of each on an equal basis. This comparison highlights important issues in the design of analog VLSI sensors, including the effects of circuit design on power consumption, the tradeoffs of subthreshold versus above-threshold MOSFET biasing, and methods of layout for focal plane vision processing arrays. The presented sensors are capable of distinguishing the direction of motion of visual stimuli to less than 5% contrast, while consuming as little as 1 μ W of electrical power. These visual motion sensors are useful in embedded applications where minimum power consumption, size, and weight are crucial.

Index Terms—Analog very large-scale integration (VLSI), biomimetic, spatio-temporal frequency, vision chip.

I. INTRODUCTION

WHEN the visual image of a dynamic three-dimensional world is projected onto a two-dimensional (2-D) retina, the true motion of objects in the world relative to the viewer becomes complex to interpret. Estimating a measure of this motion is one of the many tasks of a visual system. Biological visual systems, as well as most artificial vision systems, begin this process by detecting the image motion in an array of small local regions of the visual image. The representation of this local motion information is key to the visual system's ability to reconstruct the motion of objects in the world.

A common representation in artificial vision systems is *optical flow*, in which visual motion in the image is represented by an array of 2-D vectors, each representing the speed and direction of motion in a local region of the image. However, many situations occur

in nature that the optical flow representation is insufficient to represent. If a local image region contains portions of two objects moving in different manners, which vector is correct? If two objects move across one another and one is transparent (for example, when dust is blown across the road while in a moving vehicle), which motion should the optical flow vector represent? The flaw in the optical flow representation is the necessity of only a single motion represented in each local image region.

Perhaps for this reason, biological visual systems universally begin motion processing by using an altogether different representation. Biological visual systems from flies to primates compute the spatio-temporal frequency content of local image patches [1], [2], representing the motion of objects in the activation of an array of oriented spatio-temporal frequency filters [3], [4]. This use of a *place code* (relative activation of a bank of filters) rather than a *value code* (numerical velocity) has powerful implications. The activation of one filter representing leftward motion does not preclude the activation of a filter representing a different direction at a different spatio-temporal frequency tuning, and thus multiple motions in a local image patch are explicitly allowed, so long as they can be distinguished in frequency content. This greatly facilitates the solution of the aperture problem [5] (determination of the true motion of an extended object from local image measurements), since local motion measurements that are all consistent with one object need not conflict with the measurements of another moving object or the background. The advantages of this representation over optical flow are obvious, but if estimation of local image velocity is required these advantages come at the cost of the necessity of a bank of filters for each local image region. However, it is not clear that it is necessary to estimate image velocity to generate useful visually-guided behavior in real-world scenes. While it has been suggested that human visual systems may combine banks of spatio-temporal filters to estimate velocity [6], this seems unlikely to be true in insects due to their relatively small number of neurons [7]. Rather, it seems more likely that insect behavior is generated from spatio-temporal frequency tuned motion detectors [8] or from a prior intermediate stage of motion processing [9].

Regardless of the representation used, real-time estimation of visual motion requires significant computation. Performing this computation in special-purpose hardware is highly advantageous to reduce the load on later processing stages, potentially reducing power, weight, and size requirements over a conventional CCD/DSP combination. Very large-scale integration (VLSI) implementation of visual motion algorithms has been accomplished by a large number of authors over the last two decades [10]–[35] varying in implementation from digital block

Manuscript received January 12, 2004; revised June 18, 2004. The work of C. M. Higgins and V. Pant was supported by the Office of Naval Research under Grant N68936-00-2-0002. The work on the Barlow–Levick sensor was supported by the Office of Naval Research. This paper was recommended by Associate Editor G. Cauwenberghs.

C. M. Higgins is with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ 85721 USA and also with the Arizona Research Laboratories, Division of Neurobiology, The University of Arizona, Tucson, AZ 85721 USA (e-mail: higgins@ece.arizona.edu).

V. Pant is with the Department of Electrical and Computer Engineering, The University of Arizona, Tucson, AZ 85721 USA.

R. Deutschmann is with the McKinsey & Company, Inc., Munich 80331, Germany.

Digital Object Identifier 10.1109/TCSI.2004.841601

matching chips with serial inputs to analog VLSI motion sensors. One significant advantage of the analog VLSI approach for visual motion processing is continuous-time operation: the lack of temporal sampling removes the possibility of temporal aliasing. Secondly, the true parallel approach of fabricating a dedicated processor to compute motion in each local region of the image makes higher resolution processing only a matter of fabricating a larger VLSI chip. It is very difficult to realistically compare the performance of these many implementations characterized in many different ways with different criteria of optimality, even among monolithic analog VLSI sensors. For this reason, in this work, we describe three sensors and compare them on an equal basis to highlight issues in their design.

One significant advantage of spatio-temporal frequency based visual motion algorithms for analog VLSI implementation is the complete lack of any thresholding step. Many motion algorithms which compute a form of optical flow require a “feature detection” stage [23] to compute image speed independently of spatial frequency and contrast. The threshold required to detect a feature is the weak point of these algorithms, reducing their ability to respond to low-contrast and low-speed moving objects. Such algorithms must raise the threshold above the level of noise to avoid detecting spurious features, making it *impossible* for them to respond to stimuli which do not reach the threshold. However, implementations of spatio-temporal frequency motion algorithms (along with other algorithms which lack a thresholding step, many of which are mathematically equivalent to spatio-temporal frequency based methods) are limited in their low speed and contrast response only by signal-to-noise ratio. Even if a low-contrast object creates a response below the noise level, if the motion is sustained it is possible in principle to average out the random noise and still detect it.

With the goal of endowing small autonomous robots with powerful visual systems, in this paper we present and compare monolithic analog VLSI implementations of the three most prominent spatio-temporal frequency based visual motion algorithms, namely the Adelson–Bergen (AB) motion energy model [4], the Barlow–Levick (BL) model [36], and the Hassenstein–Reichardt (HR) model [37]. Most of the MOSFET circuits in these implementations are operated in the subthreshold regime for low power consumption. Since the human visual system cannot process frequencies higher than about 60 Hz (as evidenced by computer monitor refresh rates of 70–85 Hz), and even the faster insect visual system has cut off by 200–300 Hz [38], clearly the high frequency bandwidths of above-threshold MOSFET circuits are not required to generate useful visually-guided behavior. We begin by describing the algorithms used by each sensor, then show their analog circuit implementations, and finally present experimental characterization results from each sensor using the same apparatus. All three sensors are suitable for use in embedded applications where size, weight, and power consumption are important design factors. A method for creating a bank of such spatio-temporal filters is addressed in a related paper [39].

II. VISUAL MOTION ALGORITHMS

The three visual motion sensors described in this paper are based on algorithms which are equivalent at an appropriate level of abstraction [40], in that they all measure the energy in a

band of spatial and temporal frequency. The theoretical basis for this similarity is presented in this section, while details of each sensor’s response are derived in the Appendix. Each algorithm incorporates an essential nonlinearity required to make the mean of the output signal proportional to the direction of stimulus motion. Despite the high-level similarities between the algorithms, the steps required to implement each algorithm differ significantly. In addition, two of the algorithms have been modified from their canonical form to facilitate hardware implementation.

A. Adelson–Bergen Algorithm

The AB algorithm [4] is commonly used as a mathematical model of early visual motion processing in primates [41]. Other authors have published sensors based on this algorithm [31], but not in a monolithic VLSI implementation. The algorithm, shown in Fig. 1(a), is altered from the canonical form in four ways [33]. Firstly, the steady-state photoreceptor response is removed from the signal (using a high-pass filter) to enhance direction selectivity. Secondly, the explicit even and odd Gabor spatial filtering used in the canonical AB model is replaced by the spatial filtering implicit in a local visual sampling point. The distance between two neighboring visual sampling points is utilized to achieve the necessary spatial phase difference. Thirdly, rather than using two explicit temporal filters, one temporal pathway is unfiltered, and the second passed through a first-order low-pass filter. Finally, a full-wave rectification (absolute value) is used to approximate the squaring operation of the original algorithm. These alterations significantly simplify implementation of the algorithm without altering the fact that the sensor mean output is spatio-temporal frequency tuned. The spatial frequency tuning is dependent upon the spacing of neighboring visual sampling points, and the temporal frequency tuning upon the time constant of the low-pass filter.

In terms of the signals shown in Fig. 1, the output of the AB algorithm is

$$O_{AB} = |S_{11} - S_{22}| + |S_{12} + S_{21}| - |S_{11} + S_{22}| - |S_{12} - S_{21}|. \quad (1)$$

A derivation of the mean response of this algorithm to sinusoidal gratings is given in the Appendix. However, if we replace the absolute value with the squaring operation that it approximates, we can obtain

$$\hat{O}_{AB} = 4 \cdot (S_{21} \cdot S_{12} - S_{11} \cdot S_{22}) = 4 \cdot O_{STF} \quad (2)$$

where O_{STF} is a spatio-temporal frequency-tuned motion output term shared by the canonical versions of all three motion sensors. In the special case of a one-dimensional moving sinusoidal grating stimulus with contrast C , temporal frequency ω_t , and spatial frequency ω_x , O_{STF} can be shown to be (see Appendix)

$$O_{STF} = -\frac{C^2}{4} \cdot \sin(\phi_s) \cdot h_1^2 \cdot h_2 \cdot \sin(\phi_2). \quad (3)$$

$h_1(\omega_t)$ is the magnitude response of the high-pass filter, $h_2(\omega_t)$ and $\phi_2(\omega_t)$ are respectively the magnitude and phase response

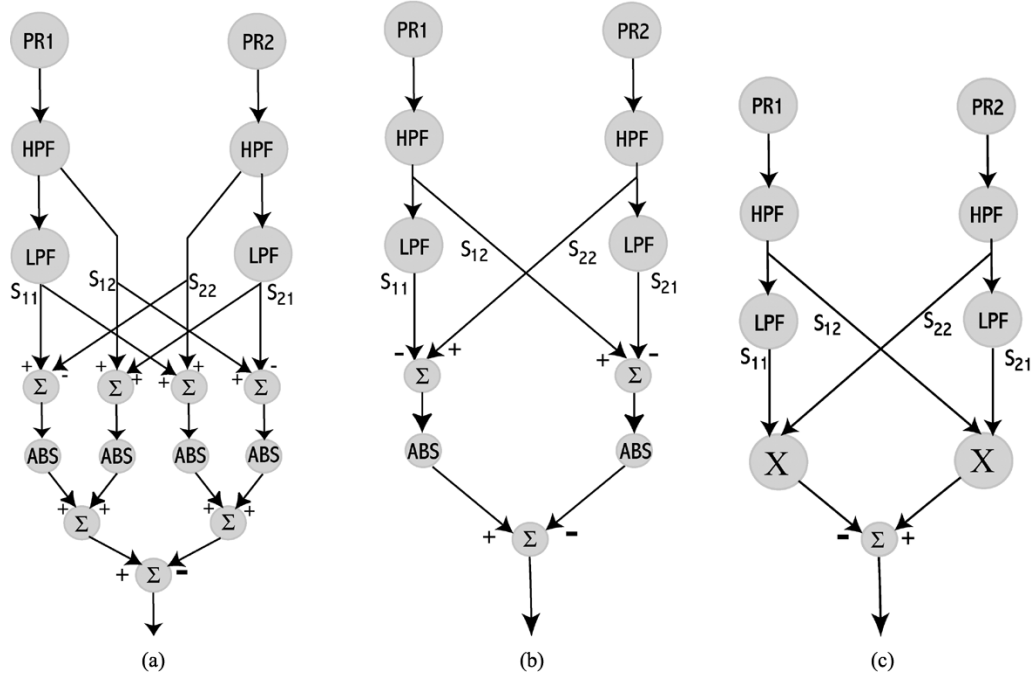


Fig. 1. Visual motion algorithms. (a) AB. (b) BL. (c) HR. PR1 and PR2 are inputs from two neighboring visual sampling points. HPF and LPF denote respectively high-pass and low-pass temporal filters. Σ indicates a summation. ABS denotes the absolute value operation. The signals shown, common to all three motion algorithms, are S_{11} (delayed response of photoreceptor one), S_{12} (undelayed response of photoreceptor one), S_{21} (delayed response of photoreceptor two), and S_{22} (undelayed response of photoreceptor two). For each algorithm, the preferred direction (which generates a positive motion response) is leftward.

of the low-pass filter, and the spatial phase factor ϕ_s is computed as

$$\phi_s = \omega_x \cdot \Delta \quad (4)$$

where Δ is the photoreceptor separation. Note that this output is constant over time. This equation can clearly be seen to have bandpass tuning in spatial frequency with its strongest response at $\omega_x = \pi/(2\Delta)$, that is, one quarter spatial cycle per photoreceptor spacing. Since the temporal response magnitude is the product of a high-pass and low-pass filter it also has a bandpass tuning, and in the case where both filters are first order with time constant τ , it can be shown that the peak temporal frequency of O_{STF} occurs at $\omega_t = 1/\tau$.

B. Barlow–Levick Algorithm

The BL algorithm [36] was originally proposed as the neuronal circuitry for direction selectivity in the rabbit retina. This algorithm has previously been the inspiration for an analog VLSI motion sensor [16], [18] which required a stimulus to cross multiple pixels to be strongly detected. Shown in Fig. 1(b), the algorithm incorporates the same high-pass filtering stage as in the AB algorithm, but then combines direct “excitation” from one photoreceptor with delayed “inhibition” from the neighbor. This creates a motion subunit sensitive to motion from the first photoreceptor toward the second. A mirror symmetrical unit is created, and the two units are combined after taking the absolute value of each to preserve the directional properties. Expressed in terms of the signals in Fig. 1, the output of this algorithm is

$$O_{BL} = |S_{22} - S_{11}| - |S_{12} - S_{21}|. \quad (5)$$

A derivation of the mean response of this algorithm to sinusoidal gratings is given in the Appendix. However, if the absolute value is replaced with a square, the output can be shown to be

$$\hat{O}_{BL} = 2 \cdot O_{STF} + S_{22}^2 - S_{12}^2 + S_{11}^2 - S_{21}^2. \quad (6)$$

In the case of a drifting sinusoidal grating stimulus, the last four terms sum to a zero-mean sinusoid and thus do not alter the fact that the mean output of the sensor is spatio-temporal frequency tuned.

C. Hassenstein–Reichardt Algorithm

The HR algorithm [3], [37] was proposed in the 1950’s to model the optomotor response in the beetle *Chlorophanus*, and is currently the leading mathematical model of insect elementary motion detection [42]. Several authors have fabricated monolithic VLSI versions of this algorithm [12], [17], [32], the closest to the present implementation being that of Harrison and Koch [32], whose algorithm is similar to the present one but whose implementation is compared in the next section. Like the previous algorithm, the HR algorithm [shown in Fig. 1(c)] synthesizes two subunits sensitive to motion in opposite directions by combining the delayed response from one photoreceptor with the undelayed response from a second. However, the HR algorithm combines these channels using a multiplication, followed by synthesis of the final output from the two subunits. Expressed in terms of the signals in Fig. 1, the output of this algorithm is

$$O_{HR} = O_{STF} \quad (7)$$

and is spatio-temporal frequency tuned.

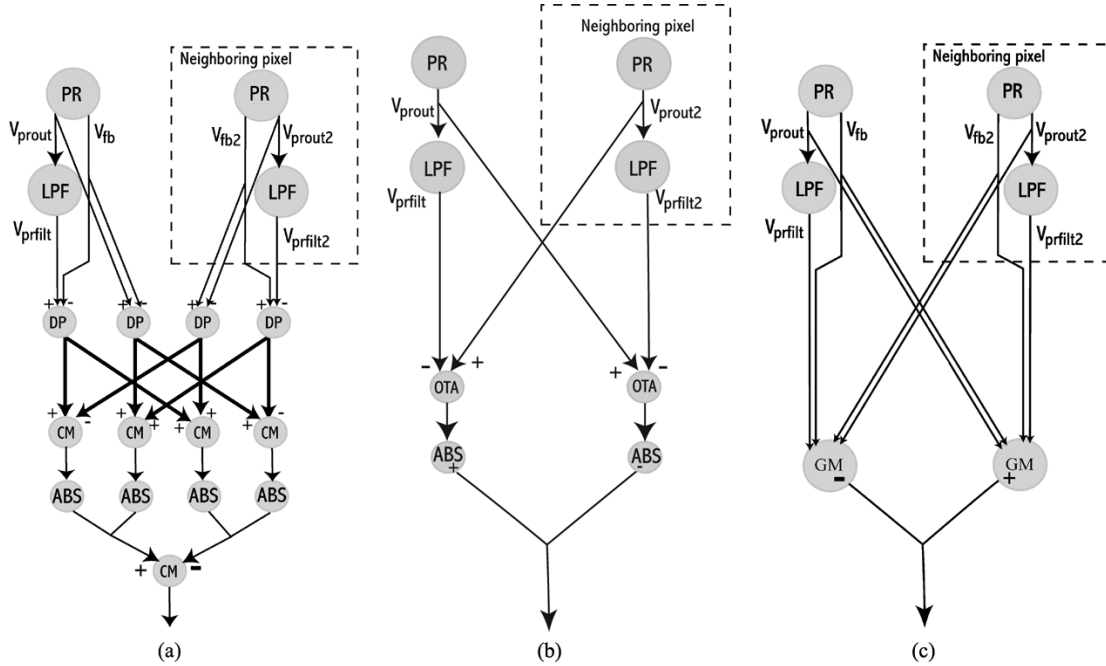


Fig. 2. Block diagram of analog circuit implementations of visual motion algorithms. (a) AB. (b) BL. (c) HR. PR indicates a photoreceptor circuit which transduces local light intensity into a voltage (see Fig. 3). LPF indicates a first-order g_m -C low-pass filter using a 5-transistor transconductance amplifier. DP indicates a differential pair used to convert a differential voltage into a differential current. Bold arrows in panel a indicate a differential current signal. CM indicates a current mirror, used to change the sign of a current. ABS indicates a current-mode absolute value circuit (Fig. 3). OTA indicates a 7-transistor wide-range operational transconductance amplifier (Fig. 3). Finally, GM represents a Gilbert multiplier, used to multiply two differential voltages.

D. Nonlinearities in the Algorithms

The absolute value operation is used in the modified AB and BL algorithms to approximate a (harder to implement) squaring function. The squaring operation on a sinusoid results in an output with two frequency components: one at dc proportional to the square of the amplitude (the crucial component for the sensor output) and one at double the frequency of the input signal (which will be removed by time-averaging). A rectification produces these same two frequency components (although the dc component is directly proportional to amplitude), but also produces harmonics at higher frequencies, thus providing less magnitude in the dc signal. As the Appendix derivations show, the fact that the absolute value operation is used in the implementation of the AB and BL algorithms rather than the square has two primary effects. First, the response of each sensor is directly proportional to contrast, rather than the square of contrast. Second, the shape of the spatio-temporal frequency response of these sensors is slightly distorted relative to O_{STF} .

While each algorithm is necessarily nonlinear, the canonical version of each algorithm can be characterized as a linear system by looking only at the mean of the sensor output. The mean output of each algorithm above to a sinusoidal grating input is a scaled version of O_{STF} . Consider the mean output of the algorithms to two distinct sinusoidal grating stimuli

$$\bar{O}_{STF1} = -\frac{C_1^2}{4} \cdot \sin(\phi_{s1}) \cdot h_1(\omega_{t1})^2 \cdot h_2(\omega_{t1}) \cdot \sin(\phi_2(\omega_{t1})) \quad (8)$$

$$\bar{O}_{STF2} = -\frac{C_2^2}{4} \cdot \sin(\phi_{s2}) \cdot h_2(\omega_{t2})^2 \cdot h_2(\omega_{t2}) \cdot \sin(\phi_2(\omega_{t2})). \quad (9)$$

The mean output of the canonical algorithms to the sum of the two sinusoidal grating stimuli can be shown to be

$$\bar{O}_{SUM} = \bar{O}_{STF1} + \bar{O}_{STF2}. \quad (10)$$

Thus, the mean output of the sensor in response to the sum of two sinusoidal grating stimuli is the sum of the mean outputs in response to the individual stimuli, and the algorithms can be termed “linear in the mean.” For this reason, we characterize each sensor by its mean output.

III. HARDWARE ARCHITECTURE

Each of the three algorithms has been implemented in analog VLSI hardware, with a design emphasis on demonstrating the low speed and low contrast response possible with this type of algorithm, while retaining small pixel size and low power consumption. A separate VLSI chip has been fabricated to evaluate each algorithm, with the algorithm implemented in each pixel of a focal plane image processing array. Each chip incorporates scanners [43] to allow individual pixel readout. A full block diagram of each of the three pixel circuit implementations is shown in Fig. 2; circuit diagrams of each sensor are given in Fig. 3.

A. Common Circuit Design Elements

Used for phototransduction in all three motion pixels is the adaptive photoreceptor by Delbrück and Mead [44]. This circuit is used in its original form in the BL pixel. The elaborated photoreceptor by Liu [45], which allows control of the adaptation time constant through the biases V_{well} and V_{adapt} , is used in the AB and HR pixels. The ratio of the vertical to the lateral capacitor value (to which the transient gain is proportional) is

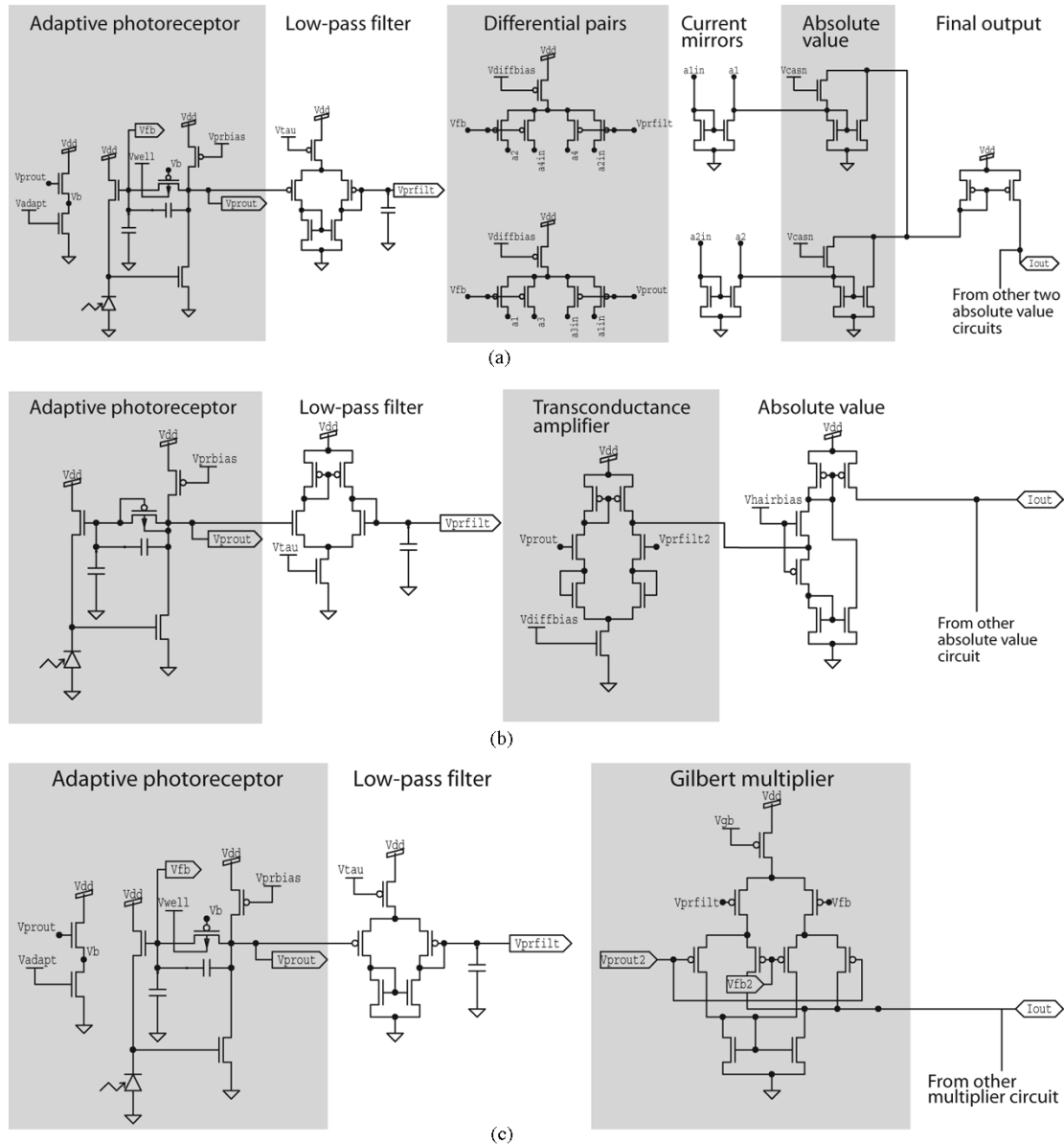


Fig. 3. Pixel circuit diagrams of visual motion sensor implementations. (a) AB sensor. Not shown (refer to Fig. 2) are two additional differential pair, current mirror, and absolute value circuits used to complete the computation with signals from the neighboring pixel. (b) BL sensor. Not shown are an additional transconductance amplifier and absolute value circuit. (c) HR sensor. Not shown is a second Gilbert multiplier. $V_{\text{prout}2}$ and $V_{\text{fb}2}$ are the photoreceptor signals from a neighboring pixel.

approximately 15 for the AB and HR pixels, and 10 in the BL pixel. All three sensors use this circuit to provide high-pass filtering; however, since the upper frequency cutoff is determined by the bias setting V_{prbias} , this circuit has an overall bandpass characteristic which limits the high-frequency response of each sensor. This circuit provides two voltage outputs: V_{prout} , with a high sensitivity to transient changes in contrast and a low sensitivity to ambient illumination levels; and V_{fb} , which is primarily sensitive to the long-term illumination level. The BL pixel uses V_{prout} directly as a high-pass filtered signal, since its response to sustained illumination is reduced relative to transient illumination. To improve the high-pass filtering operation, the AB and HR pixels use further circuitry to take the difference of V_{prout} and V_{fb} . This difference has near zero response to mean luminance and thus has effectively been high-pass filtered with the

time constant controlled by the adaptation time constant of the photoreceptor.

In all three pixel circuits, V_{prout} is low-pass filtered using a first order g_m -C filter [46] to produce V_{prfilt} . The AB and HR pixels use a P-type transconductance amplifier, while the BL pixel uses an N-type amplifier. In each case, the capacitor value is approximately 1.7 pF. The bias V_{tau} controls the low-pass filter time constant.

B. Adelson–Bergen Sensor

In the AB sensor [Fig. 3(a)], P-type differential pairs are used to subtract local signals V_{prout} and V_{prfilt} from V_{fb} and also to do the same with these three signals from the neighboring pixel (differential pairs not shown). Since each current must go to

TABLE I
MOTION SENSOR CHIP SPECIFICATIONS*

Sensor	Array Size (pixels)	Pixel transistor count	Pixel area (μm^2)	Photodiode area (μm^2)	Fill factor (%)	Process resolution (μm)	Die size (mm^2)	Pixel power (μW)	Total power (μW)
AB	12×40	53	29950	1002	3.3	1.6	20.2	1.04	500
BL	24×1	35	20233	576	2.8	2.0	4.41	27	650
HR	5×13	29	17620	1335	7.6	1.6	4.41	3.3	200

*In addition to transistor count shown, each pixel also contains a photodiode (area given above) and three capacitors (see text). power consumption shown is as biased for best performance. Power-supply voltage was 5 V. Larger power consumption of BL sensor is due to above-threshold photoreceptor bias current (see text).

two places in the computation, doubled differential pairs are required, with the doubled transistors on each side sharing equally in the current if the Early effect is neglected. These differential pairs convert the voltage differences into differential currents, which are then properly routed to perform the AB computation in current mode. The bias $V_{diffbias}$ controls the magnitude of sensor current output. To implement the rest of the algorithm, current mirrors are used to change the signs of currents where a difference is necessary, and Kirchoff's current law (KCL) to perform additions. Four three-transistor current-mode absolute value circuits [47] are used to implement the necessary nonlinearities. A summary of the specifications of the AB sensor is given in Table I.

C. Barlow–Levick Sensor

The BL sensor [Fig. 3(b)] uses an n-type transconductance amplifier, with its linear differential input range extended by source degeneration of the differential pair, to subtract the local voltage signal V_{pout} from the filtered signal from the neighboring photoreceptor V_{pfil2} . A second transconductance amplifier (not shown) does the same with the two signals V_{pout2} and V_{pfil} . The bias $V_{diffbias}$ controls the magnitude of sensor output current. Current-mode rectifier circuits [48] (only one shown) are used for computing the absolute value of each current. The bias $V_{hairbias}$ is set at a dc level such that both output transistors of the transconductance amplifier can stay in saturation. This rectifier circuit can be used with either the P-type or N-type mirror providing the output current to obtain either current output sign. By taking the output of the two rectifier circuits with different current signs, KCL is used to perform the final subtraction. A summary of the specifications of the BL sensor is given in Table I. Despite the fact that the BL pixel was laid out in a one-dimensional array, scaling to two dimensions poses no technical challenges since no additional pixel wiring would be required to support two-dimensional scanner circuitry.

D. Hassenstein–Reichardt Sensor

The HR sensor [Fig. 3(c)] is the simplest of all sensors shown. It uses a Gilbert multiplier to produce a current proportional to the product of $V_{pfil} - V_{fb}$ with $V_{pout2} - V_{fb2}$. The bias V_{gb} controls the magnitude of the sensor current output. A second Gilbert multiplier (not shown) produces a current proportional to the product of $V_{pfil2} - V_{fb2}$ with $V_{pout} - V_{fb}$. The final

current mirror on the second Gilbert multiplier is switched input for output to provide the opposite current sign from the first multiplier, so that the final subtraction can be performed with KCL. A summary of the specifications of the HR sensor is given in Table I.

The subtraction of the long-term mean voltage V_{fb} from the photoreceptor signal is the primary difference between our implementation and that of Harrison [32], who uses a lateral dc-clamped capacitor to perform the high-pass filtering operation.

IV. EXPERIMENTAL CHARACTERIZATION OF MOTION SENSORS

A single pixel of each VLSI sensor was visually characterized by presenting computer-generated moving visual stimuli on an LCD screen (used for its low flicker). An image of this stimulus was focused onto each chip using a camera lens. Despite the fact that some of the sensors were laid out in two-dimensional arrays, each individual sensor on all chips responds maximally to visual stimuli moving along the axis of the two photoreceptors involved in its motion computation. We refer to stimuli which move in that direction which elicits the largest positive mean current output from the sensor as *preferred direction stimuli*, and stimuli which move in that direction which elicits the largest negative mean current output from the sensor as *null direction stimuli*. Each chip was biased to maximize the difference between mean outputs in response to preferred and null direction stimuli (which we use as a definition of good performance) at a differential pair bias current in the nanoamperes. For the BL chip, this required biasing the photoreceptor well above threshold to satisfy the saturation condition for the bias transistor of the N-type transconductance amplifier in the next stage. All other sensor circuits were biased in the subthreshold regime to save power.

In this section, we detail the visual stimuli used and present results of our characterizations.

A. Visual Stimuli

Unless otherwise stated, the stimuli were all sinusoidal gratings [refer to (A.1)]. At a given bias setting, we define the “optimal” stimulus (or stimulus parameter value) as the one that maximizes the difference between mean outputs to preferred and null direction stimuli. Whenever one or more parameters

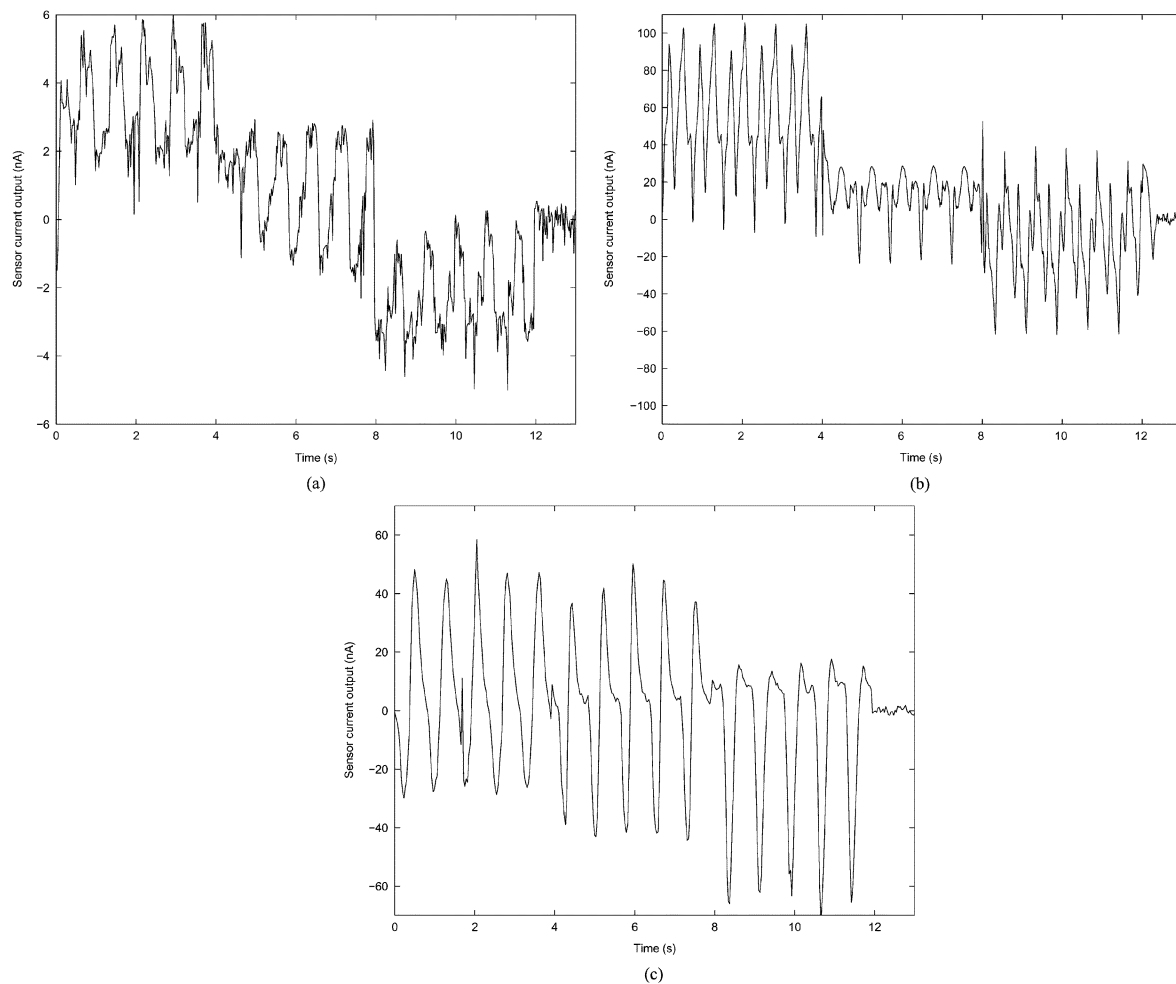


Fig. 4. Raw current output (in nanoamperes) of individual analog VLSI motion sensors to a moving sinusoidal grating stimulus. (a) AB. (b) BL. (c) HR. The stimulus moves leftward for the first four seconds (yielding a positive mean output), orthogonal to the sensor orientation between 4 and 8 s (zero mean output), rightward (negative mean output) between 8 and 12 s, and then stops (zero output). The magnitude of these outputs can be controlled by pixel biases.

were being varied, all other parameters were held at their optimal value.

For each chip, a fixed mean illumination level m was held for all sinusoidal visual stimuli. The AB and BL sensors were characterized with a mean illumination of 50% of the maximal LCD brightness, which corresponded to $0.739 \mu\text{W}/\text{m}^2$. Due to a light-induced leakage problem with the HR sensor (see Discussion), it was characterized with a lower fixed mean illumination level of 20% of the maximum possible, which corresponded to $0.323 \mu\text{W}/\text{m}^2$.

Due to the brightness of the “black” regions of the LCD screen, the maximum possible contrast was approximately 80%. Optimal stimulus contrast for the AB and BL chips was maximum. Due to light leakage effects on the HR sensor, it was characterized for an optimal contrast of 20%.

For an equal comparison between chips, spatial frequency was calibrated in units of *cycles per chip pixel* (that is, the number of spatial periods of a sinusoid that fit in the distance from the center of one photoreceptor to its neighbor) for each chip. The optimal spatial frequency for the AB and BL sensors was 0.25 cycles per chip pixel, at which setting their mean outputs are theoretically predicted to be maximal. Due to the light-induced leakage problem with the HR sensor, and the

corresponding shift of spatial frequency characteristic [refer to Fig. 5(c), shown later], the HR sensor was characterized at an optimal spatial frequency of 0.15 cycles per chip pixel.

Temporal frequencies were varied from zero to about 20 Hz, with each sensor electronically tuned to have maximum mean output at a temporal frequency of 1.4 Hz.

Outputs of the chips are in the form of currents in the nanoampere range, and were converted to voltages using off-chip current sense amplifiers for data collection. Data were averaged for ten temporal periods of the stimulus, which was long enough that further averaging did not significantly change the data.

B. Results

The raw current output of each sensor to an optimal sinusoidal grating stimulus is shown in Fig. 4. Visual stimuli move first in the preferred direction, then move orthogonal to the sensor orientation (which stimulates both photoreceptors identically and should elicit zero mean response), and finally in the null direction. The AB and BL sensors achieve a large difference in mean value between the three stimulus directions relative to the amplitude of the overall waveform, while the responses of the HR sensor are much more overlapping. Theoretically, due to the symmetry in each motion algorithm, each sensor should have no

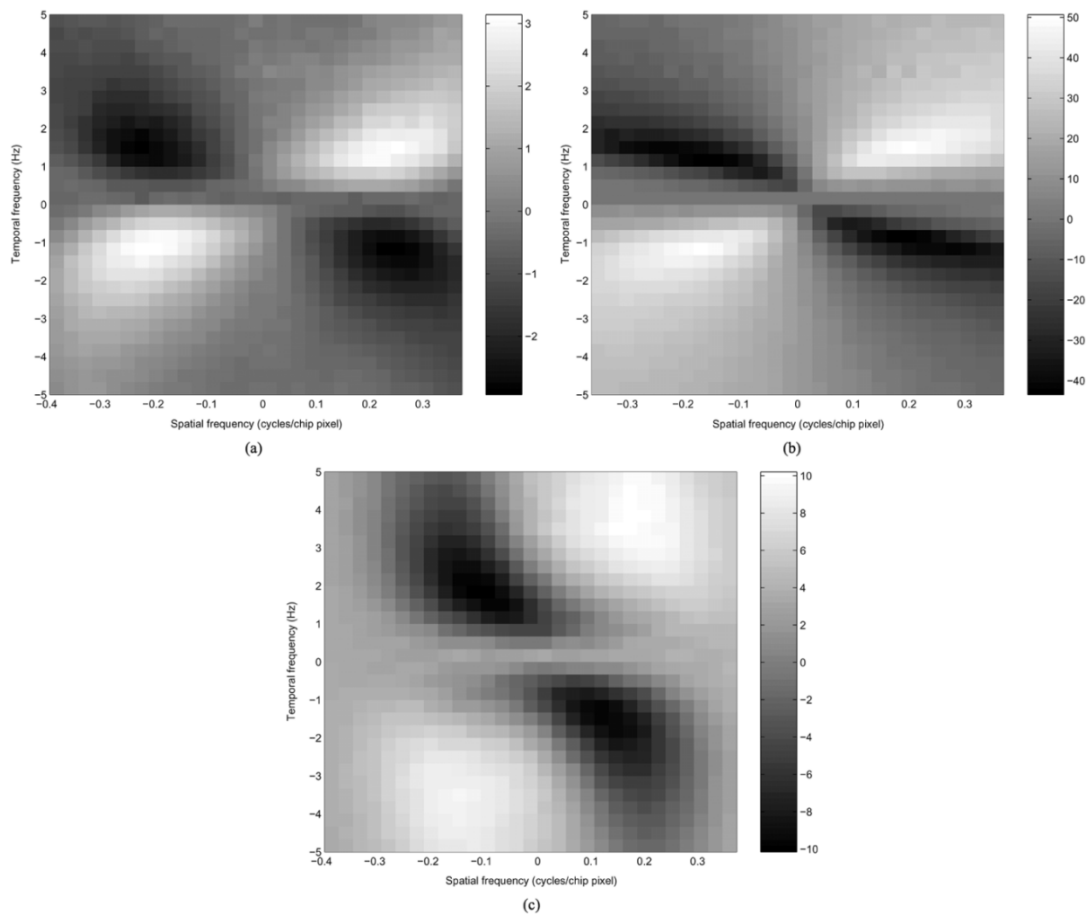


Fig. 5. Mean output of analog VLSI motion sensor implementations to variation of spatial and temporal frequency, shown in nanoamperes. (a) AB. (b) BL. (c) HR. Brighter gray levels indicate stronger responses. Each point in the plot corresponds to the mean sensor output over ten temporal periods of a moving sinusoidal grating with a particular combination of spatial and temporal frequency. Note that negative temporal (or spatial) frequency corresponds to a change in stimulus direction. The asymmetric shape of the HR response is due to light-induced leakage (see text).

response at all to orthogonal stimuli; any response of the sensor is due to transistor mismatch. Compared to the AB and HR sensors, the BL sensor has little response to orthogonal stimuli due to symmetry between its left and right subunits, resulting from improved photoreceptor matching at an above-threshold bias current (see Section V).

A series of moving sinusoidal stimuli with varying spatial and temporal frequency was presented to each chip, and the mean output of the sensor over ten temporal periods of each sinusoid computed. The results are plotted in Fig. 5. It can clearly be seen that each sensor is tuned in spatial and temporal frequency, and has responses of opposite sign for stimuli moving in different directions. The nominal tuning of each sensor in spatial frequency should occur at 0.25 cycles per chip pixel, and this is significantly violated only for the HR sensor which suffers a light-induced leakage problem. This light leakage also leads to the asymmetry of the HR's plot (see Discussion).

The optimal temporal frequency of each sensor can be varied electronically, as demonstrated for the AB sensor in Fig. 6. At each setting, the temporal frequency tuning curve is bandpass, as predicted by (3). Note the very low temporal frequencies in this plot: the graph tuned to the lowest frequency peaks at about 1 Hz. While maintaining subthreshold bias conditions, the peak temporal frequency can be biased to more than 100 Hz, which is already more bandwidth than many biological visual systems utilize.

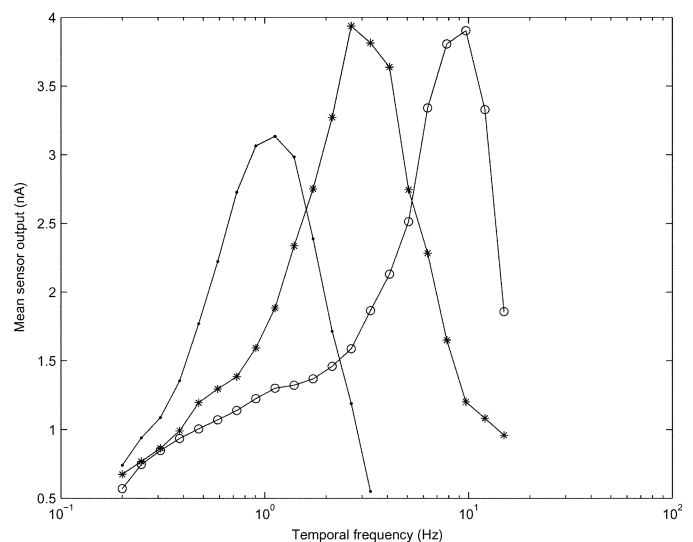


Fig. 6. Electronic variation of temporal frequency response, demonstrated for the AB motion sensor. Bias V_{tau1} was set at 4.46 V (plotted with dots), 4.42 V (asterisks), and 4.38 V (circles), all of which were subthreshold. Note tunability to extremely low-speed stimuli (very low temporal frequencies). Peak response is modulated by the bandpass frequency characteristic of the photoreceptor circuit, operated in subthreshold.

Using the optimal spatial and temporal frequency for each chip, the contrast was varied from zero to the maximum possible, resulting in the curves of Fig. 7. Responses to stimuli

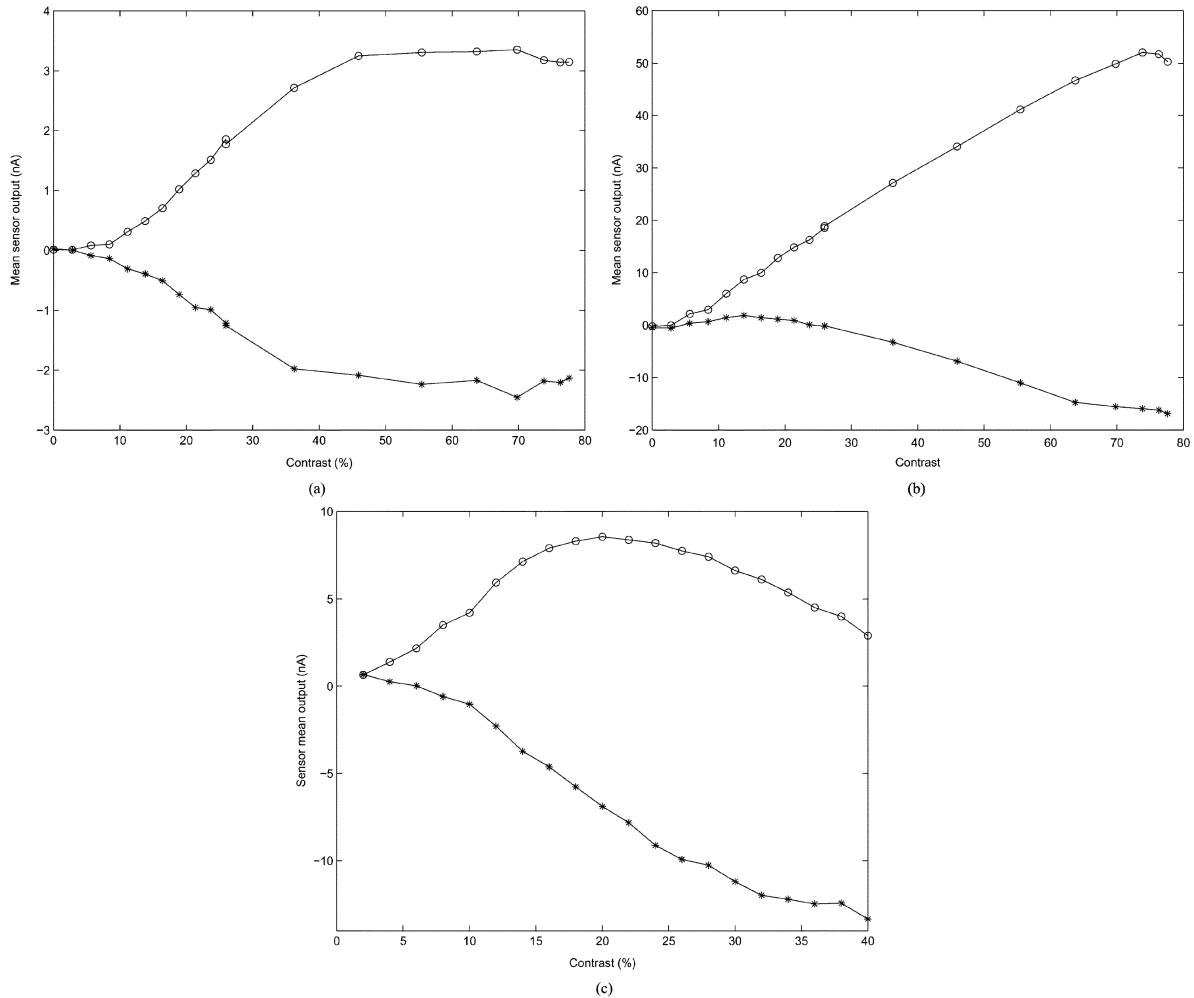


Fig. 7. Contrast response of analog VLSI motion implementations (mean output shown in nanoamperes). (a) AB. (b) BL. (c) HR (note smaller contrast scale due to light-induced leakage). Each point shown is the mean output over ten temporal stimulus periods. Circles indicate responses to motion in the preferred direction; asterisks in the null direction. The shape of the preferred-direction response of the HR sensor at higher contrast is due to light-induced leakage (see Discussion).

moving in the preferred and null directions are compared at a range of contrasts. The linear response of the AB and BL sensors to contrast is apparent at contrasts less than 30%, after which the AB curve saturates. Despite the weaker response of the BL sensor in the null direction relative to the preferred direction [also visible in the raw data of Fig. 4(b)], its response is monotonic in contrast at contrasts greater than about 15%. At high contrasts, both the AB and BL sensors show a response slightly less than their maximum due to light-induced leakage. The HR sensor's theoretically predicted square-law response to contrast is masked by the effects of light-induced leakage (see Discussion), reaching a maximum difference between preferred and null mean outputs at 20% contrast, after which the preferred direction response increasingly approaches zero. At low contrasts, all sensors were able to discriminate leftward from rightward motion at contrasts less than 5%.

It is often true that the characterization with sinusoidal stimuli of an analog VLSI motion sensor (which in truth is not a linear system, even in the mean, given photoreceptor adaptation, light-induced leakage, and circuit nonidealities) is not predictive of its performance in real-world scenarios. In order to demonstrate the extreme sensitivity of the present motion sensors to real-

world motion, the BL sensor was presented with a real-world stimulus with medium contrast and complex spectral content: a hand was waved back and forth over the sensor, resulting in the trace shown in Fig. 8.

Finally, power consumption was measured for each chip during operation. Biases were optimized for performance, not for power consumption, and especially in the case of the BL sensor bias currents could be reduced significantly at the cost of performance. In all cases, due to the design of the analog circuits utilized, there was no measurable dependence of power consumption upon normal visual stimulation. Power consumption figures, along with other metrics of each sensor, are summarized in Table I.

V. DISCUSSION

We have presented analog VLSI implementations of the three most prominent spatio-temporal frequency tuned visual motion algorithms. Each algorithm shows a linear or sublinear response to increasing contrast, is able to discriminate the direction of motion at less than 5% contrast, and can be tuned for optimal response down to temporal frequencies less than 1 Hz.

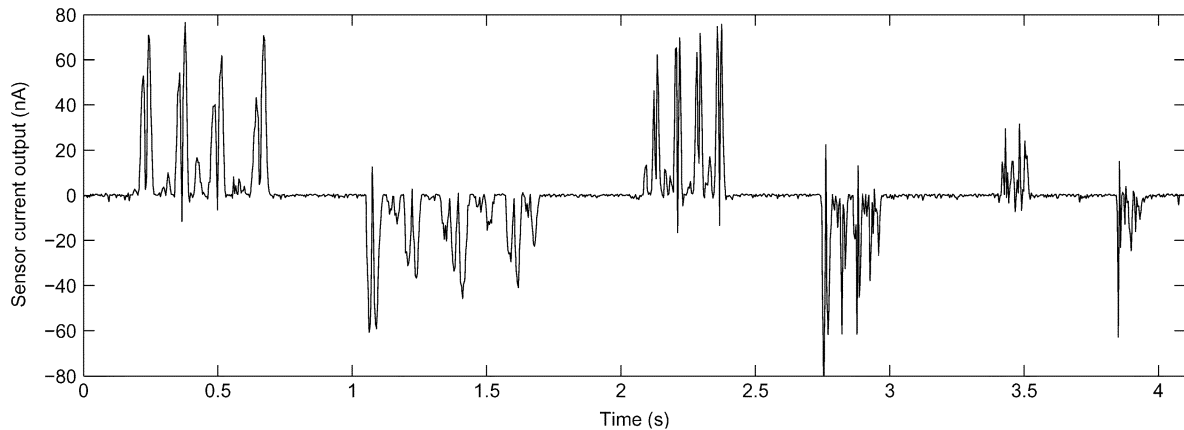


Fig. 8. Raw output of the BL sensor to the motion of a hand in its visual field. The hand was waved back and forth three times, at first slowly and then faster. It is possible not only to easily discriminate the direction of motion of the hand, but also to see the two edges of each of the four fingers in the response of the sensor, which is tuned to low frequencies and thus responds less as the hand moves faster.

Although the three algorithms are quite mathematically similar, the implementations differ greatly. The AB sensor, with the highest transistor count by far, also consumes the smallest amount of power because most of the circuitry simply controls the pathway of current flow from earlier stages and requires no bias currents. In terms of raw response characteristics (Fig. 4), the best overall performance is seen from the BL sensor. This sensor achieves a large separation between mean responses to preferred and null direction stimuli, and a minimal response to orthogonal stimuli, by operating its photoreceptor above threshold and thus allowing better transistor matching. This performance comes at the cost of power consumption, which is 26 times greater than the AB sensor. The HR sensor, with the lowest transistor count and power consumption only three times that of the AB sensor, suffers from a light-induced leakage current problem which limits its useful operating regime to low mean light intensities.

Since any drain/source region exposed to light in the focal plane becomes a vertical photodiode, light-induced “leakage” currents are an issue to be dealt with at the layout stage in all focal plane vision processors. Typically, the area intended to act as a photodiode is exposed, while all other drain/source regions are “covered” with one or more metal layers. Since such thin metal layers can be translucent, this leakage is virtually always present in such sensors, so the problem is not so much light-induced leakage *per se*, but rather an *imbalance* of light leakage leading to undesired circuit operation. The HR sensor, which was actually fabricated as part of a tracking experiment [49], suffers a loss of directional selectivity at high light intensity levels due to an imbalance of light leakage between the two Gilbert multipliers in each pixel caused by exposed structures. This leakage is seen strongly in the output of the Gilbert multiplier in the negative pathway of the HR output. The HR sensor was thus characterized at the lowest possible mean light intensity. Despite the use of a low mean light intensity, as contrast increases the highest absolute light level required to produce a visual stimulus increases, and thus leakage imbalance effects are still visible in the HR characterization results. The HR contrast curve [Fig. 7(c)] shows this effect most strongly. Due to the fact that the leakage is more pronounced in one Gilbert multiplier output than the other, higher brightness levels cause sensor

output currents to become more negative regardless of stimulus direction. Thus, the null direction (negative) response is apparently strong, while the preferred direction (positive) response approaches zero at medium contrasts. Also due to leakage effects, the optimal spatial frequency for the HR sensor [Fig. 5(c)] is shifted from 0.25 down to 0.15 cycles per chip pixel. This phenomenon is caused by the fact that higher spatial frequencies lead to higher spatial gradients of local luminance. Since the two Gilbert multipliers in each HR sensor are located side by side along the preferred-null direction axis, this increases the instantaneous difference in illumination between the two Gilbert multipliers, worsening the leakage imbalance and leading to weaker mean responses. This leakage could be reduced by devoting slightly more layout area to allow better light shielding. In the case of the HR sensor, rerouting a single wire in the second metal layer would ameliorate the leakage problem while adding only a few percent to the pixel area. The AB and BL designs suffer much less from light leakage, but this effect can be seen as a “droop” in the response of these chips at high contrast (see Fig. 7).

The spatio-temporal frequency characteristic of each sensor, shown in Fig. 5, can be compared with the theoretical prediction of Fig. 9. While each sensor is clearly spatio-temporal frequency tuned, the shapes of these tunings for the VLSI implementations are much more complex than predicted due to frequency filtering from circuit components taken in the derivation to have wide frequency bandwidth. The tuning of the AB sensor is most like its prediction, particularly in the sharp transition between positive and negative responses at low temporal frequency relative to the transition at low spatial frequency. The tuning of the BL sensor is much narrower in temporal frequency than predicted, likely due to temporal frequency filtering by the two operational transconductance amplifiers (OTAs), operated in subthreshold, which follow the explicit low-pass filters. The HR sensor tuning is much affected by light-induced leakage, with its optimal spatial frequency shifted down to 0.15 cycles per chip pixel, and some overlap of negative responses into preferred-direction stimuli.

Fig. 7(a) shows that the response of the AB sensor to increasing contrast becomes sublinear at about 30% and is virtually constant from 50% to maximum. This *contrast saturation*

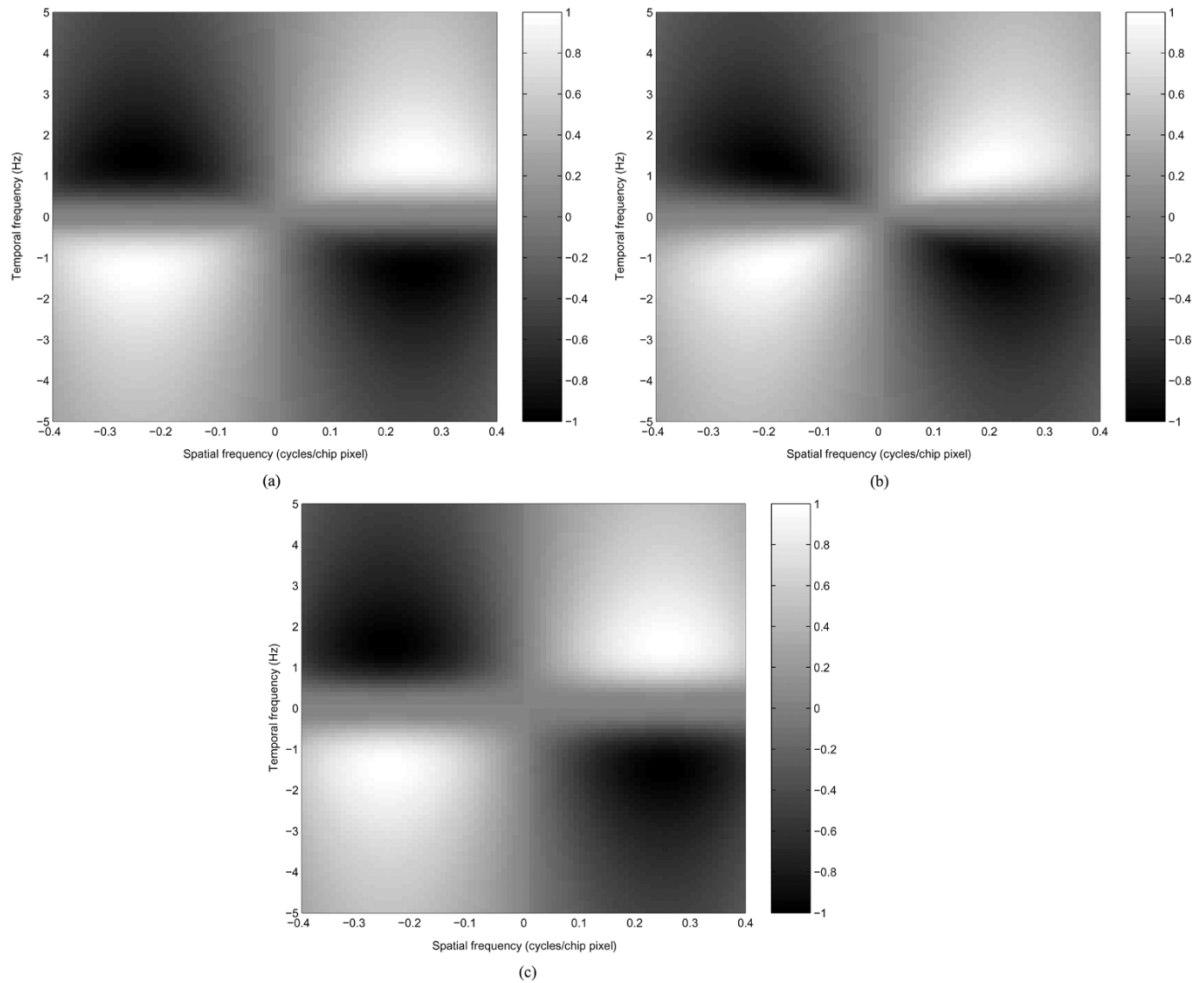


Fig. 9. Theoretically predicted mean output of analog VLSI motion sensor implementations to variation of spatial and temporal frequency, normalized to a maximum of unity. (a) AB. (b) BL. (c) HR. To match the optimal temporal frequency of the empirical spatio-temporal frequency data taken from the sensors, time constants for both high-pass and low-pass filters have been chosen to be 180 ms. Brighter gray levels indicate stronger responses. Each point in the plot corresponds to the theoretical mean sensor output to a moving sinusoidal grating with a particular combination of spatial and temporal frequency.

is due to saturation of the output current of the differential pairs in the AB circuit pathway. This is a highly desirable property for real-world motion sensors, allowing changes in contrast to be ignored while remaining sensitive to changes in stimulus motion, and in fact is a well-known property of biological motion sensors [50]. The BL sensor, due both to its lower photoreceptor gain relative to the AB sensor and the intentional wide linear range of its OTAs, does not exhibit contrast saturation. At contrasts less than 15%, the BL sensor's mean response to null-direction stimuli is slightly positive, likely due to mismatch in the subthreshold OTAs used to compare filtered to unfiltered photoreceptor signals. The HR sensor's contrast response is masked by light-induced leakage.

The AB pixel has the smallest pixel power consumption, at $1.04 \mu\text{W}$ per pixel; at this power consumption a 300×300 array would consume only a modest 93.6 mW. Despite the large number of structures in the AB implementation, past the low-pass filter stage all circuits simply perform current routing functions and thus do not require bias currents. The HR implementation consumes somewhat more power due to the necessity of Gilbert multipliers, which require their own bias currents, late in the computation. The much larger power consumption of the

BL sensor is due to a large bias current in the photoreceptor circuit, which is necessary to raise the dc level of the photoreceptor voltage signal high enough to maintain the bias transistor of the N-type g_m -C low-pass filter in saturation. If refabricated with P-type low-pass filters, the photoreceptor bias current could be greatly reduced. However, due to this large bias current, the BL photoreceptor frequency bandwidth is very broad, leading to the highly sensitive responses highlighted in Fig. 8.

The sensitivity of these sensors to low speed and contrast, along with their minimal power consumption, make them excellent alternatives to a CCD/DSP combination in cases such as that of small mobile robots where power consumption, size, and weight are major design considerations.

APPENDIX

In this appendix, we present a detailed analysis of the theoretical mean output of all three motion algorithms shown in Fig. 1, particularly to account for the effect of the absolute-value non-linearity used in the AB and BL sensors.

All three sensors in common compute their output from the four signals shown in Fig. 1: S_{11} , S_{12} , S_{21} , and S_{22} . We analyze

the output of each sensor in response to a sinusoidal grating stimulus

$$I(x, t) = m \cdot (1 + C \cdot \sin(\omega_t \cdot t + \omega_x \cdot x)) \quad (\text{A.1})$$

where m is the stimulus mean value (taken as $1/2$ for this derivation), C is the Michelson contrast, ω_t is the temporal frequency, and ω_x the spatial frequency. The two photoreceptor inputs S_1 and S_2 to each motion detector can then be expressed as

$$S_1 = 1/2 \cdot (1 + C \cdot \sin(\omega_t \cdot t)) \quad (\text{A.2})$$

$$S_2 = 1/2 \cdot (1 + C \cdot \sin(\omega_t \cdot t + \phi_s)) \quad (\text{A.3})$$

where ϕ_s is the spatial phase factor (4).

The first step in each model passes each input signal through a high-pass filter. We assume that the HPF completely removes the sustained component of the signal. The effect of the linear HPF on these sinusoids is simply to multiply the amplitude by a frequency-dependent term $h_1(\omega_t)$ and to add a phase $\phi_1(\omega_t)$. Thus, the two signals after being high-pass filtered become

$$S_{12} = C/2 \cdot h_1 \cdot \sin(\omega_t \cdot t + \phi_1) \quad (\text{A.4})$$

$$S_{22} = C/2 \cdot h_1 \cdot \sin(\omega_t \cdot t + \phi_s + \phi_1). \quad (\text{A.5})$$

These two signals are then passed through a low-pass filter with magnitude response $h_2(\omega_t)$ and phase response $\phi_2(\omega_t)$ to produce

$$S_{11} = C/2 \cdot h_2 \cdot h_1 \cdot \sin(\omega_t \cdot t + \phi_1 + \phi_2) \quad (\text{A.6})$$

$$S_{21} = C/2 \cdot h_2 \cdot h_1 \cdot \sin(\omega_t \cdot t + \phi_s + \phi_1 + \phi_2). \quad (\text{A.7})$$

The implementation of each algorithm uses a saturating circuit (a differential pair alone, in an OTA, or in a Gilbert multiplier) in further processing of these four signals. This effect could be modeled by passing appropriate signals through a saturating hyperbolic tangent function

$$\hat{S} = \tanh(g \cdot S) \quad (\text{A.8})$$

where g is an empirical gain factor. However, the introduction of the hyperbolic tangent term at this stage makes a compact closed-form derivation impossible, so we omit it and carry on the analysis for unsaturated signals in each algorithm.

I. ADELSON–BERGEN ALGORITHM

The AB output using the absolute value operation [Fig. 1(a)] may be computed as

$$O_{\text{AB}} = |S_{11} - S_{22}| + |S_{12} + S_{21}| \\ - |S_{11} + S_{22}| - |S_{12} - S_{21}|. \quad (\text{A.9})$$

The sum of scaled and shifted sinusoids at the same frequency is always still a sinusoid. By use of common trigonometric identities, the amplitude of the sinusoid inside each absolute value operator can be evaluated. Then, using the empirical approximation

$$|A \cdot \sin(\omega_t \cdot t + \phi)| \approx \frac{A}{2} \cdot (1 - \cos(2 \cdot \omega_t \cdot t + 2 \cdot \phi)) \quad (\text{A.10})$$

it is possible to compute the mean output of the AB algorithm as

$$\bar{O}_{\text{AB}} = \frac{C}{4} \cdot h_1 \cdot \left(\sqrt{1 + h_2^2 - 2h_2 \cos(\phi_2 - \phi_s)} \right. \\ - \sqrt{1 + h_2^2 + 2h_2 \cos(\phi_2 - \phi_s)} \\ + \sqrt{1 + h_2^2 + 2h_2 \cos(\phi_2 + \phi_s)} \\ \left. - \sqrt{1 + h_2^2 - 2h_2 \cos(\phi_2 + \phi_s)} \right). \quad (\text{A.11})$$

Owing to the use of the absolute value operation instead of the square, this response is directly proportional to contrast rather than the square of contrast [compare (A.15)]. The value of (A.11) is shown as spatial and temporal frequency are varied in Fig. 9(a). This sensor is tuned in spatial and temporal frequency, but the shape of the tuning is slightly distorted relative to O_{STF} [Fig. 9(c)].

II. BARLOW–LEVICK ALGORITHM

The BL output using the absolute value operation [Fig. 1(b)] may be computed as

$$O_{\text{BL}} = |S_{22} - S_{11}| - |S_{12} - S_{21}|. \quad (\text{A.12})$$

Using similar identities and approximations as for the AB algorithm, the mean output of the BL algorithm can be shown to be

$$\bar{O}_{\text{BL}} = \frac{C}{4} \cdot h_1 \cdot \left(\sqrt{1 + h_2^2 - 2h_2 \cos(\phi_2 - \phi_s)} \right. \\ \left. - \sqrt{1 + h_2^2 - 2h_2 \cos(\phi_2 + \phi_s)} \right). \quad (\text{A.13})$$

Again, owing to the use of the absolute value operation, this response is directly proportional to contrast rather than the square of contrast. The value of (A.13) is shown as spatial and temporal frequency are varied in Fig. 9(b). Like the AB sensor, this sensor is tuned in spatial and temporal frequency, but the shape of the tuning is slightly distorted relative to O_{STF} [Fig. 9(c)].

III. HASENSTEIN–REICHARDT ALGORITHM

The HR algorithm [Fig. 1(c)] is implemented in canonical form. This sensor computes

$$O_{\text{HR}} = S_{21} \cdot S_{12} - S_{11} \cdot S_{22} = O_{\text{STF}} \quad (\text{A.14})$$

the value of which can be directly derived from (A.4)–(A.7) by the use of trigonometric identities to be

$$O_{\text{STF}} = -\frac{C^2}{4} \cdot \sin(\phi_s) \cdot h_1^2 \cdot h_2 \cdot \sin(\phi_2). \quad (\text{A.15})$$

This response is independent of time, and proportional to the square of contrast. The value of (A.15) is shown as spatial and temporal frequency are varied in Fig. 9(c). This sensor is separably tuned in spatial and temporal frequency.

ACKNOWLEDGMENT

The authors would like to acknowledge the skillful technical assistance of University of Arizona undergraduate students W. Sheung, who performed a detailed hardware characterization of the Barlow–Levick motion sensor, and J. Halimeh, who assisted in the testing of the Adelson–Bergen motion sensor; and the support of the Laboratory of Christof Koch, California Institute of Technology, Pasadena.

REFERENCES

- [1] W. Reichardt, "Autocorrelation, a principle for the evaluation of sensory information by the central nervous system," in *Sensory Communication*, W. A. Rosenblith, Ed. New York: MIT Press, 1961, pp. 303–317.
- [2] P. H. Schiller, B. L. Finlay, and S. F. Volman, "Qualitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields," *J. Neurophysiol.*, vol. 39, no. 6, pp. 1288–1319, 1976.
- [3] J. P. H. Van Santen and G. Sperling, "Elaborated Reichardt detectors," *J. Opt. Soc. Amer. A*, vol. 2, pp. 300–320, 1985.
- [4] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Amer. A*, vol. 2, no. 2, pp. 284–299, 1985.
- [5] E. C. Hildreth and C. Koch, "The analysis of visual-motion-from computational theory to neuronal mechanisms," *Ann. Rev. Neurosci.*, vol. 10, pp. 477–533, 1987.
- [6] N. M. Grzywacz and A. L. Yuille, "A model for the estimate of local image velocity by cells in the visual cortex," *Proc. Royal Soc. London B: Biol. Sci.*, vol. 239, pp. 129–161, 1990.
- [7] N. J. Strausfeld, *Atlas of an Insect Brain*. New York: Springer, 1976.
- [8] C. M. Higgins, J. K. Douglass, and N. J. Strausfeld, "The computational basis of an identified neuronal circuit for elementary motion detection in dipterous insects," *Vis. Neurosci.*, vol. 21, no. 4, pp. 567–586, 2004.
- [9] C. M. Higgins, "Non-directional motion may underlie insect behavioral use of image speed," *Biol. Cybern.*, vol. 91, no. 5, pp. 326–332, 2004.
- [10] J. Tanner and C. Mead, "An integrated analog optical motion sensor," in *VLSI Signal Processing II*, S.-Y. Kung, R. Owen, and G. Nash, Eds. New York: IEEE Press, 1986, pp. 59–76.
- [11] J. Hutchinson, C. Koch, J. Luo, and C. Mead, "Computing motion using analog and binary resistive networks," *IEEE Comput.*, vol. 21, no. 3, pp. 52–63, Mar. 1988.
- [12] A. G. Andreou and K. Strohhahn, "Analog VLSI implementation of the Hassenstein–Reichardt–Poggio model for vision computation," in *Proc. IEEE Int. Conf. Systems, Man, Cybernetics*, Los Angeles, CA, Nov. 1990, pp. 708–710.
- [13] A. Moore and C. Koch, "A multiplication based analog motion detection chip," *Proc. SPIE Vis. Inform. Process.*, vol. 1473, pp. 66–75, 1991.
- [14] 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. San Mateo, CA: Morgan Kaufmann, 1991, pp. 406–412.
- [15] A. G. Andreou, K. Strohhahn, and R. E. Jenkins, "Silicon retina for motion computation," in *Proc. Int. Symp. Circuits Systems (ISCAS'91)*, vol. , Singapore, Jun. 1991, pp. 1373–1376.
- [16] R. G. Benson and T. Delbrück, "Direction selective silicon retina that uses null inhibition," in *Advances in Neural Information Processing Systems 4*, D. S. Touretzky, Ed. San Mateo, CA: Morgan Kaufmann, 1991, pp. 756–763.
- [17] R. Sarpeshkar, W. Bair, and C. Koch, "An analog VLSI chip for local velocity estimation based on Reichardt's motion algorithm," in *Advances in Neural Information Processing Systems 5*, S. J. Hanson, J. Cowan, and L. Giles, Eds. San Mateo, CA: Morgan Kaufmann, 1993, pp. 781–788.
- [18] T. Delbrück, "Silicon retina with correlation-based, velocity-tuned pixels," *IEEE Trans. Neural Netw.*, vol. 4, no. 5, pp. 529–541, May 1993.
- [19] A. Bilgin, H. Abut, and M. W. Marcellin, "A novel vector quantizer based architecture for block-matching motion estimation," in *Proc. 29th Asilomar Conf. Signals Computers*, vol. 2, Pacific Grove, CA, Oct. 1995.
- [20] R. C. Meitzler, K. Strohhahn, and A. G. Andreou, "A silicon retina for 2-D position and motion computation," in *Proc. Int. Symp. Circuits Systems (ISCAS'95)*, vol. 3, Apr./May 1995, pp. 2096–2099.
- [21] J. Kramer, "Compact integrated motion sensor with three-pixel interaction," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 4, pp. 455–460, Apr. 1996.
- [22] X. Arreguit, F. A. van Schaik, F. Bauduin, M. Bidiville, and E. Raeber, "A CMOS motion detector system for pointing devices," *IEEE J. Solid-State Circuits*, vol. 31, no. 12, pp. 1916–1921, Dec. 1996.
- [23] R. Sarpeshkar, J. Kramer, G. Indiveri, and C. Koch, "Analog VLSI architectures for motion processing: From fundamental limits to system applications," *Proc. IEEE*, vol. 84, no. 7, pp. 969–987, 1996.
- [24] A. Moini, A. Bouzardoum, K. Eshraghian, A. Yakovlev, X. T. Nguyen, A. Blanksby, R. Beare, D. Abbott, and R. E. Bogner, "An insect vision-based motion detection chip," *IEEE J. Solid-State Circuits*, vol. 32, no. 2, pp. 279–283, Feb. 1997.
- [25] R. Etienne-Cummings, J. Van der Spiegel, and P. Mueller, "A focal plane visual motion measurement sensor," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 44, no. 1, pp. 55–66, Jan. 1997.
- [26] R. A. Deutschmann and C. Koch, "Compact real-time 2-D gradient based analog VLSI motion sensor," in *Proc. Int. Conf. Advanced Focal Plane Arrays and Electronic Cameras*, Zurich, Switzerland, 1998.
- [27] —, "An analog VLSI velocity sensor using the gradient method," in *Proc. IEEE Int. Symp. Circuits Systems (ISCAS'98)*, vol. 6, Monterey, CA, 1998, pp. 649–665.
- [28] 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: MIT Press, 1999, vol. 11.
- [29] C. M. Higgins, R. A. Deutschmann, and C. Koch, "Pulse-based 2-D motion sensors," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 46, no. 6, pp. 677–687, Jun. 1999.
- [30] A. Torralba and J. Hérault, "An efficient neuromorphic analog network for motion estimation," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 46, no. 2, pp. 269–280, Feb. 1999.
- [31] R. Etienne-Cummings, J. Van der Spiegel, and P. Mueller, "Hardware implementation of a visual-motion pixel using oriented spatiotemporal neural filters," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 46, no. 9, pp. 1121–1136, Sep. 1999.
- [32] R. R. Harrison and C. Koch, "A robust analog VLSI Reichardt motion sensor," *Analog Integr. Circuits Signal Process.*, vol. 24, pp. 213–229, 2000.
- [33] C. M. Higgins and S. K. Korrapati, "An analog VLSI motion energy sensor based on the Adelson–Bergen algorithm," in *Proc. Int. Symp. Biologically-Inspired Systems*, Wollongong, NSW, Australia, Dec. 2000.
- [34] S. C. Liu, "A neuromorphic aVLSI model of global motion processing in the fly," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 47, no. 12, pp. 1458–1467, Dec. 2000.
- [35] A. A. Stocker, "Analog VLSI focal-plane array with dynamic connections for the estimation of piecewise-smooth optical flow," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 51, no. 5, pp. 963–973, May 2004.
- [36] H. B. Barlow and W. R. Levick, "The mechanism of directionally selective units in rabbit's retina," *J. Physiol. (London)*, vol. 178, pp. 477–504, 1965.
- [37] B. Hassenstein and W. Reichardt, "Systemtheoretische analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers *Chlorophanus*," *Zs. Nat.*, vol. 11b, pp. 513–524, 1956.
- [38] R. C. Miall, "The flicker fusion frequencies of six laboratory insects, and the response of the compound eye to mains fluorescent 'ripple'," *Physiol. Entomol.*, no. 3, pp. 99–106, 1978.
- [39] E. Özalpeli and C. M. Higgins, "Reconfigurable biologically-inspired visual motion systems using modular neuromorphic VLSI chips," *IEEE Transact. Circuits Syst. I, Reg. Papers*, vol. 52, no. 1, pp. 79–92, Jan. 2005.
- [40] C. W. G. Clifford and M. R. Ibbotson, "Fundamental mechanisms of visual motion detection: Models, cells and functions," *Progr. Neurobiol.*, vol. 68, pp. 409–437, 2003.
- [41] D. J. Heeger, E. P. Simoncelli, and J. A. Movshon, "Computational models of cortical visual processing," *Proc. Nat. Acad. Sci.*, vol. 93, pp. 623–627, 1996.
- [42] R. A. Harris, D. C. O'Carroll, and S. B. Laughlin, "Adaptation and the temporal delay filter of fly motion detectors," *Vis. Res.*, vol. 39, pp. 2603–2613, 1999.
- [43] C. A. Mead and T. Delbrück, "Scanners for visualizing activity of analog VLSI circuitry," *Analog Integr. Circuits Signal Process.*, vol. 1, pp. 93–106, 1991.
- [44] T. Delbrück and C. Mead, "Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits," *Prog. Computation Neural Syst.*, California Institute of Technology, Pasadena, Tech. Rep. 30, 1993.
- [45] S. C. Liu, "Silicon retina with adaptive filtering properties," *Analog Integr. Circuits Signal Process.*, vol. 18, no. 2/3, pp. 243–254, 1999.

- [46] R. Schaumann and M. E. Van Valkenburg, *Design of Analog Filters*. Oxford, U.K.: Oxford Univ. Press, 2001.
- [47] A. G. Andreou and K. A. Boahen, "Translinear circuits in subthreshold MOS," *Analog Integr. Circuits Signal Process.*, vol. 9, pp. 141–166, 1996.
- [48] R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," *Analog Integr. Circuits Signal Process.*, vol. 16, no. 3, pp. 245–274, 1998.
- [49] V. Pant and C. M. Higgins, "A biomimetic VLSI architecture for small target tracking," in *Proc. IEEE Int. Symp. Circuits Systems*, Vancouver, BC, Canada, 2004.
- [50] M. Egelhaaf and A. Borst, "Transient and steady-state response properties of movement detectors," *J. Opt. Soc. Amer. A*, vol. 6, pp. 116–127, 1989.



Charles M. Higgins (SM'04) received the Ph.D. degree in electrical engineering from the California Institute of Technology (Caltech), Pasadena, in 1993.

He worked in the Radar Systems Group at the Lincoln Laboratory, Massachusetts Institute of Technology, Cambridge, MA, until 1996. He then returned to Caltech as a Postdoctoral Research Fellow in the Division of Biology studying engineering applications of visual neuroscience. In 1999, he joined the Department of Electrical and Computer Engineering, University of Arizona, Tucson, as

Assistant Professor, where he holds a joint appointment in the Division of Neurobiology. His research focuses on investigation and engineering application of neurobiological computational and control architectures and data representations, particularly in the area of insect visual motion processing and associated motor control. Projects in this area range from simulation modeling to the construction of biologically inspired analog/digital custom very large-scale integrated vision chips and robotic systems.



Vivek Pant received the M.S. degree from the University of Arizona, Tucson. He is working toward the Ph.D. degree in electrical and computer engineering at the same university.

His Master's research focused on elaborating biological algorithms for target tracking and developing a small-target tracking custom very large-scale integration (VLSI) architecture. His research interests include analog/digital/mixed-signal VLSI, computational modeling of neural systems, and neuromorphic VLSI design.



Rainer Deutschmann received the diploma degree in physics from the California Institute of Technology, Pasadena, and the doctoral degree in semiconductor physics (*summa cum laude*) from the University of Technology, Munich, Germany, in 1997, and 2001, respectively.

His graduate work involved developing motion detection algorithms and their very large-scale integration (VLSI) implementation and his doctoral study investigated two-dimensional electron systems in atomically precise periodic potentials. Since 2001,

he has been with McKinsey & Company, Inc., Munich, Germany, primarily helping hi-tech and telecommunications clients in strategy and operations.