Autonomous Robots 7, 211–224 (1999)

A Robust Analog VLSI Motion Sensor Based on the Visual System of the Fly

REID R. HARRISON AND CHRISTOF KOCH
Computation and Neural Systems Program, MS 139-74, California Institute of Technology, Pasadena CA 91125, USA
harrison@klab.caltech.edu
koch@klab.caltech.edu

Abstract. Sensing visual motion gives a creature valuable information about its interactions with the environment. Flies in particular use visual motion information to navigate through turbulent air, avoid obstacles, and land safely. Mobile robots are ideal candidates for using this sensory modality to enhance their performance, but so far have been limited by the computational expense of processing video. Also, the complex structure of natural visual scenes poses an algorithmic challenge for extracting useful information in a robust manner. We address both issues by creating a small, low-power visual sensor with integrated analog parallel processing to extract motion in real-time. Because our architecture is based on biological motion detectors, we gain the advantages of this highly evolved system: A design that robustly and continuously extracts relevant information from its visual environment. We show that this sensor is suitable for use in the real world, and demonstrate its ability to compensate for an imperfect motor system in the control of an autonomous robot. The sensor attenuates open-loop rotation by a factor of 31 with less than 1 mW power dissipation.

Keywords: optomotor system, analog VLSI, optic flow, motion detector, insect vision

1. Introduction

Optic flow patterns produced by self-motion are one of the richest sources of navigation information available to a mobile creature (Gibson, 1950). As an animal moves through its environment, images of the outside world move across its retina in predictable ways. Objects being approached grow larger; objects left behind grow smaller. When moving forward, images of nearby objects move across the retina faster than images of distant objects. If a creature rotates in place, the entire visual scene moves across its retina at a rate that is independent of object distance. Much information can be gained from patterns of visual motion, even if no explicit object recognition is performed (Duchon et al., 1998). Indeed, motion parallax information is immune to camouflage that can defeat even the most sophisticated static pattern recognition scheme when object and background have similar textures. Humans have no difficulty detecting the structure of randomly patterned objects against identically patterned backgrounds from motion cues alone.

Using egomotion-induced optic flow for robot navigation is a computationally demanding sensory task. By its very nature it must be done in real time. Most object recognition tasks are performed on static images, and often one can tolerate latencies of several seconds. But optic flow is available only while the robot is moving, and relevant information must be extracted in real time and fed back to the motor control system to steer the robot in the right direction. The rate of computation needed depends on the rate of robot motion, but typical real-world situations require times on the order of tens or hundreds of milliseconds.

Optic flow is also computationally demanding because, like other early vision tasks, it involves operations that must be performed identically on every pixel of an image. Local estimates of motion must be laboriously computed before the overall pattern is analyzed. This is a task that is ideally suited for parallel computation.

Measuring optic flow also involves large amounts of data. While audition involves one time-varying signal (two in the case of binaural audition), vision involves

many time-varying signals. Rapid navigation requires many frames to be analyzed each second. This can tax even the most sophisticated microprocessor because it must deal with all the signals at once. If we divide the job to many processors, each dealing with one pixel and communicating with its immediate neighbors, the task becomes much easier. Human photoreceptors have bandwidths of 55 Hz (Kandel et al., 1991).

Vision is a vitally important sense for flying insects. In the housefly's brain, over half of the 350,000 neurons are believed to have some role in visual processing. The fly's optic lobes contain wide-field motion-sensitive neurons which respond to moving stimuli over large portions of the visual field. Many of these neurons have been linked to specific visually-guided behaviors that help the animal navigate through a complex environment in a robust manner (Egelhaaf and Borst, 1993). For example, the well-studied optomotor system estimates rotation from optic flow and uses this information to produce a stabilizing torque with the wings (Götz, 1975; Warzecha and Egelhaaf, 1996).

In the past decade, researchers have been endowing mobile robots with biologically-inspired (more specifically, insect-inspired) visual systems (Franceschini et al., 1992; Srinivansan et al., 1997; Huber, 1997; Lewis, 1998). These efforts have yielded promising results, but many problems still exist. Vision is a computationally intensive task, so powerful hardware is required to operate in real time. From an algorithmic viewpoint, the structure of visual scenes is often very complex, and it can be difficult to extract relevant information robustly.

Biological visual systems are highly parallel structures that compute in very different ways from traditional digital computers. There is no global clock, and the architecture specifies the algorithm. A recent approach is to attempt to build analog electronic circuits that emulate neural systems (Mead, 1989; Douglas et al., 1995). By taking advantage of VLSI (Very Large Scale Integration) technology, we can fabricate silicon chips with thousands of transistors in an automated fashion, allowing reliable construction of medium-scale neural models. These chips process information in parallel at a very high power efficiency.

Silicon is a flexible medium, and it allows us to construct photodetectors in the same substrate as our computational elements. Many different analog VLSI imagers with integrated motion detectors have been built in the past decade, and functionality is slowly increasing (Delbrück, 1993; Etienne-Cummings and Van der Spiegel, 1996; Sarpeshkar et al., 1996; Moini et al., 1997; Higgins et al., in press). Here we report on one such sensor, and its applications to autonomous systems.

2. Biological Motion Detection

2.1. Motion Processing in the Housefly Brain

Insects process visual motion information in a local, hierarchical manner. Despite the multi-lens construction of the compound eye, the pattern projected onto the underlying retina is a single image of the visual scene. Photoreceptors in the retina adapt to the ambient light level, and signal deviations from this level. These signals are passed on to the next layer of cells, the lamina. Lamina cells generally show transient or highpass responses, emphasizing temporal change (Weckström et al., 1992). The next stage of processing is the medulla, a layer of cells that are extremely difficult to study directly due to their small size. Indirect evidence suggests that local measures of motion (i.e., between adjacent photoreceptors) are computed here. These local, direction-selective motion estimates are integrated by huge tangential cells in the lobular plate (Hausen and Egelhaaf, 1989). The housefly has 50–60 tangential cells in each half of its brain. These are the best-studied cells in the fly visual system, and much is known about their properties.

Lobular plate cells generally respond to motion over large parts of the visual field. Some of these cells seem to be matched filters for the optic flow patterns produced by rotation or translation along particular axes (Krapp and Hengstenberg, 1996). Some of these cells most likely control compensatory motor reflexes that prevent the fly from rotating during flight. Others are sensitive only to small objects moving across the visual field (Egelhaaf, 1985). It is believed that these "figure detection" cells allow the fly to locate nearby objects through motion parallax (Kimmerle et al., 1997). All of these sensory abilities require that motion first be detected locally between every pair of photoreceptors.

2.2. Elementary Motion Detectors

We now turn to a relatively well-understood biological computational element that is used to explain how the brains of flying insects process dynamic visual input. This element is the elementary motion detector

(EMD), sometimes called the Hassenstein-Reichardt motion detector after the men who first proposed its architecture in 1956 while modeling beetle vision (Hassenstein and Reichardt, 1956). The Hassenstein-Reichardt EMD belongs to the class of correlation-based motion detectors; a measure of image motion between two pixels is computed by correlating the output of one photoreceptor with the delayed output of an adjacent photoreceptor (Borst and Egelhaaf, 1989). A closely related class of models, known as spatiotemporal energy models (Adelson and Bergen, 1985) mimics many aspects of motion perception in primates.

Figure 1(a) shows the architecture of an EMD. Two adjacent photoreceptors send their outputs to temporal bandpass filters which remove constant illumination (containing no motion information) and high temporal frequencies (having low signal-to-noise ratios). These signals are then "delayed" by exploiting the phase lag inherent in a first-order temporal lowpass filter. Delayed channels are then correlated with adjacent, non-delayed channels by means of a multiplication operation. If an image is moving across the retina, it produces a strong response when it first passes over the delayed photoreceptor, then over the non-delayed photoreceptor. The maximum response is elicited when the transit time between two photoreceptors equals the EMD time delay. Thus the EMD has a preferred velocity at which it responds maximally; slower or faster velocities produce weaker responses. For a detailed analysis of the Reichardt motion detector's response to sinusoidal gratings, see Egelhaaf et al. (1989).

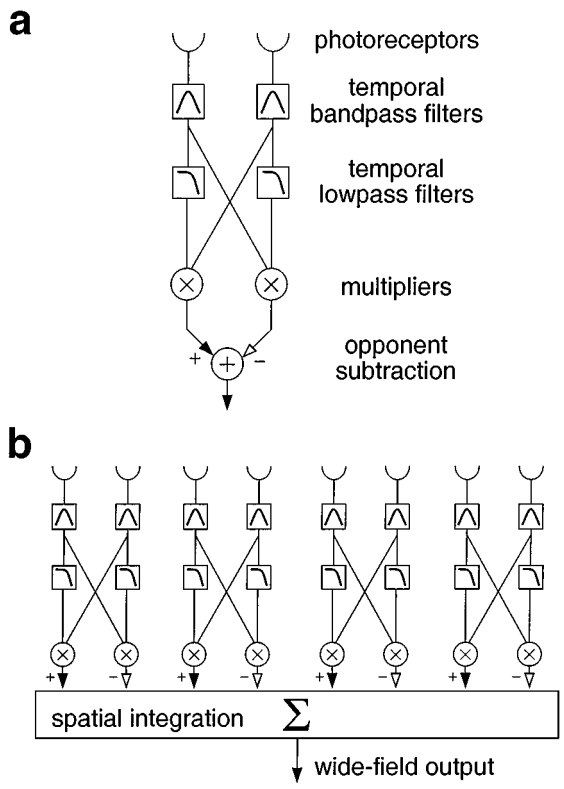

Figure 1. Elementary motion detector (EMD) architecture. (a) An opponent pair of elementary motion detectors. Lowpass filters delay the incoming signals, which are correlated with non-delayed signals from the adjacent photoreceptor. Outputs from left- and right-oriented channels are subtracted to give a strongly directionally selective response. (b) An array of EMDs with linear spatial integration. Spatial integration can reduce the pattern dependence associated with single EMD pairs (Reichardt and Egelhaaf, 1988).

Individual delay-and-correlate units are only weakly directionally selective. We subtract the outputs of two opponent units to yield a strongly direction-selective response (i.e., positive for leftward motion and negative for rightward motion) (Borst and Egelhaaf, 1990).

2.3. Integrating Local Motion Information

By spatially integrating the response of many EMDs, we can detect full-field motion, such as the kind produced by rotation. Flies use visual motion information to estimate self-rotation and generate a compensatory torque response to maintain stability during flight. This well-studied behavior is known as the optomotor response. It is interesting from an engineering point of view because it extracts relevant information from a dynamic, unstructured environment using passive sensors and uses this information to generate appropriate motor commands in real time.

Spatial integration has another advantage. The response of the EMD pair is not a static measure of velocity. Rather, there is a pattern dependence where the output fluctuates in time about a mean value. We can reduce this pattern dependence by integrating over a larger portion of the visual scene, a trick exploited by wide-field motion-sensitive neurons in the fly (see Fig. 1(b)) (Reichardt and Egelhaaf, 1988).

3. Silicon Motion Detection

We constructed a silicon implementation of the Hassenstein-Reichardt motion detector in a standard CMOS VLSI process. We built an array of photoreceptors with integrated motion processing circuitry on

the same chip, resulting in a monolithic visual motion sensor. A lens was mounted over the chip, focusing an image onto the photoreceptors. Motion information was read from the chip as a time-varying current, which was digitized by an external computer.

3.1. Circuit Architecture

We chose to implement the EMD architecture in analog, continuous-time circuitry using neuromorphic engineering techniques pioneered by Mead and colleagues (Mead, 1989). There is no software; the algorithm is entirely specified by the circuit architecture. By exploiting the physics of transistors and capacitors, we achieve very efficient silicon implementations of basic computations. Each computational "building block" represented by a box in Fig. 1(a), such as temporal filtering and correlation, is implemented with a few transistors (see Fig. 2).

An array of photodiodes was used to transduce light into electric currents. We used a small, four-transistor circuit with each photodiode to convert the photocurrent into a logarithmically encoded voltage (Delbrück and Mead, 1996). The circuit also provided local intensity-based gain control which allowed individual photoreceptors to adapt to the local constant illumination at each point in the image. This adaptation allowed the photoreceptor to adjust to large changes in ambient lighting without external parameter changes. The signal from each adaptive photoreceptor was an analog, continuous-time voltage corresponding to the image intensity.

The filter time constants were set by off-chip bias voltages, so we could tune the EMD to a particular range of temporal frequencies. One entire EMD (including adaptive photoreceptor) measures 61×199 μm, and consists of 34 transistors and five capacitors (1.9 pF of total capacitance), and consumes 25 nW of power. We gain power efficiency by staying in the analog domain for computations that are better performed in parallel across the chip, and thus need only a small bandwidth ($<$100 Hz).

All data shown was measured from an analog VLSI chip fabricated in a standard, commercially available 1.2 μm CMOS process. The 2.2×2.2 mm chip contained six parallel one-dimensional arrays of 24 EMD opponent pairs each with integrated photoreceptors (see Fig. 3). Multiple rows of motion detectors are useful in practical applications because some rows may be focused on featureless parts of a scene. The outputs of all EMD pairs were summed to simulate the wide-field motion-sensitive neurons found in flies (see Fig. 1(b)). We mounted a 2.6 mm lens over the chip, which gave the photoreceptors an angular spacing of 1.3° (similar to the 1°–2° angular spacing observed in fly eyes), and a total field of view of 30° (much less than the fly's eye, which sees almost an entire visual hemifield). The lowpass filter time contant was set to 50 ms, and the

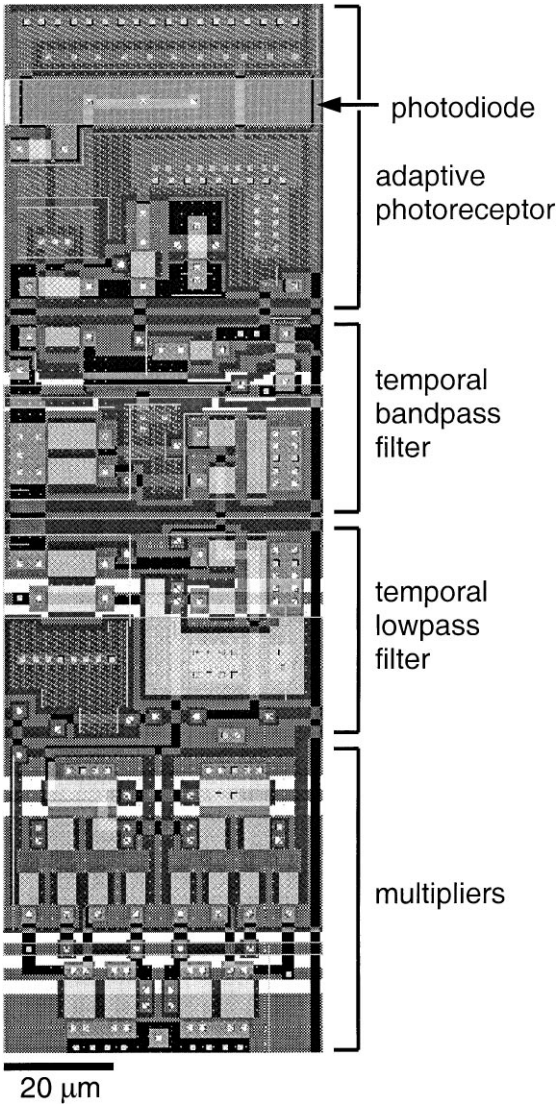

Figure 2. Elementary motion detector (EMD) circuit layout. This shows one EMD cell as it was fabricated in a 1.2 μm CMOS process. The small white dots are vertical connections between coplanar layers of metal and silicon. Larger gray squares are transistors, and the large dark gray structures are capacitors. The entire cell measures 61×199 μm.

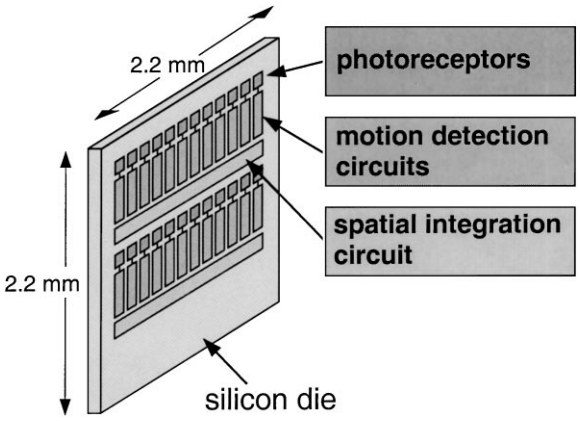

Figure 3. Layout of EMD circuits on a chip. This schematic shows how EMDs were arranged in parallel rows covering the chip. The actual chip contained seven rows with 24 EMD circuits in each row. One row was used only for circuit testing, so a 24×6 array of motion detectors was used in the optomotor experiments.

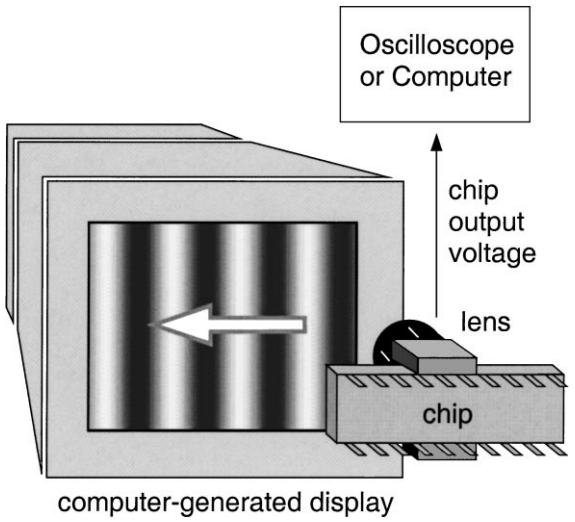

Figure 4. Chip testing methodology. We mounted a lens directly over the chip to focus an image on the photoreceptor array. Moving patterns were generated on a standard computer monitor. The temporal bandpass filters in each EMD blocked the 72 Hz refresh rate signal from the monitor.

bandpass filters were set to pass frequecies between 0.5 and 8 Hz.

3.2. Response to Simple Stimuli

To demonstrate the direction selectivity of the silicon EMD, we repeated an experiment often performed on flies when investigating the behavior of motion-sensitive neurons: we presented a drifting 1-D sinusoidal grating (see Fig. 4). The visual motion stimuli were generated by computer and displayed on a monitor (Sony Multiscan 17seII). We were able to achieve frame rates of 70 Hz, and the screen resolution far exceeded the photoreceptor spacing.

The summed chip response (see Fig. 5) was similar to the membrane potential of HS and VS cells, nonspiking wide-field motion-sensitive neurons in flies in three ways:

- **Direction selectivity.** The sign of the response indicates motion direction (Haag and Borst, 1997).
- **Transient oscillations.** At the onset of stimulus motion, a large transient response oscillates with the temporal frequency of the stimulus pattern. This transient decays to a steady state level at a rate given by the time constant of the EMD lowpass filter (approximately 50 ms) (Egelhaaf and Borst, 1989).
- **Pattern dependence.** Due to imperfect spatial summation and device mismatch, the steady-state output shows some residual pattern dependence, oscillating at the temporal frequency of the stimulus pattern (Single and Borst, 1998).

In previous experiments with an older version of our silicon EMD, we varied the spatial and temporal frequency of the sinusoidal stimulus. The response of the chip was consistent with the observed behavior of motion-sensitive cells from various flying insects during in similar experiments (Harrison and Koch, 1998; O'Carroll et al., 1996).

3.3. Response to Stimuli with Natural Image Statistics

While useful for initial evaluations, sinusoidal gratings are simple artificial stimuli that a creature is rather unlikely to encounter while navigating through the real world. (We have also repeated the above experiments with square-wave gratings, and the results are very similar.) We would like to characterize the performance of our sensor with real-world stimuli to test its robustness in the face of more complex visual scenes.

One of the difficulties in measuring "real-world robustness" is that complex stimuli may be hard to define and standardize. If we use a "cluttered office environment" for a visual stimulus, how does another group in a different lab reproduce this stimulus to evaluate the

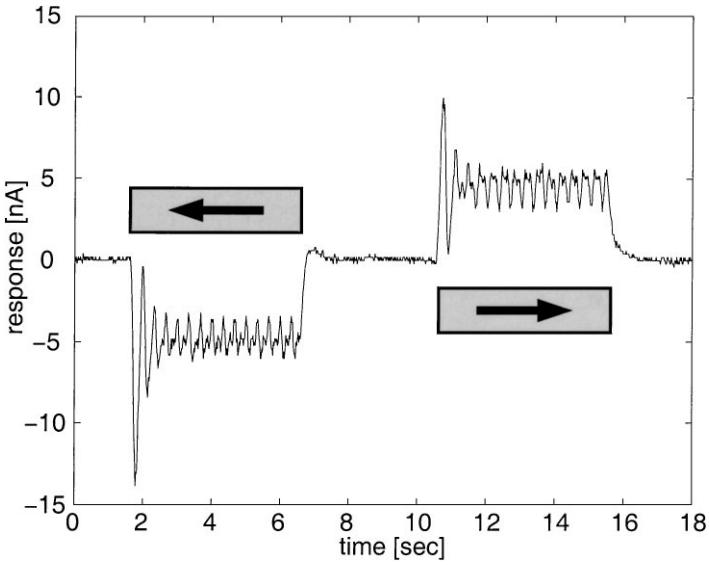

Figure 5. Transient response of silicon 24-EMD array. A drifting sinusoidal grating ($v = 25$ deg/s) evoked this response, which resembles the response of HS cells, wide-field motion-sensitive interneurons in the optic lobe of the fly. These non-spiking neurons also show direction selectivity (Haag and Borst, 1997), transient oscillations at the onset of motion (Egelhaaf and Borst, 1989), and residual pattern dependence at the pattern temporal frequency (Single and Borst, 1998).

relative robustness of their sensor? Of course the real world is always the ultimate acid test for robustness, but we propose a useful middle ground: generating random stimuli that conform to the statistics observed in the natural environment, and using these to test sensors.

In the set of all possible images a computer monitor can display, the subset of these images that do not look like random noise is vanishingly small. It has been found that natural images exhibit a predictable statistical structure (Field, 1987; Ruderman and Bialek, 1994; Dong and Atick, 1995). These statistics hold for images of natural as well as man-made objects. Static natural scenes exhibit a Fourier power spectrum R that goes as

$$R(f_s) \sim \frac{1}{f_s^m} \tag{1}$$

where f_s is spatial frequency, and $m \approx 2.3$. Low spatial frequencies are "over-represented", indicating that pixels in natural images are highly correlated with neighboring pixels. (The power spectrum is simply the Fourier transform of the autocorrelation.)

We generated one-dimensional patterns with a $1/f_s^{2.3}$ power spectrum and random phase. By randomizing the phase, we could generate any number of distinct stimuli with natural-scene statistics. Figure 6 shows four examples of these stimuli. Of course, natural scenes certainly have higher-order statistics that are not present in our stimuli, but we argue that these random $1/f_s^m$ patterns are significantly more complex than traditional visual stimuli (e.g., bars, dots, sinusoids)

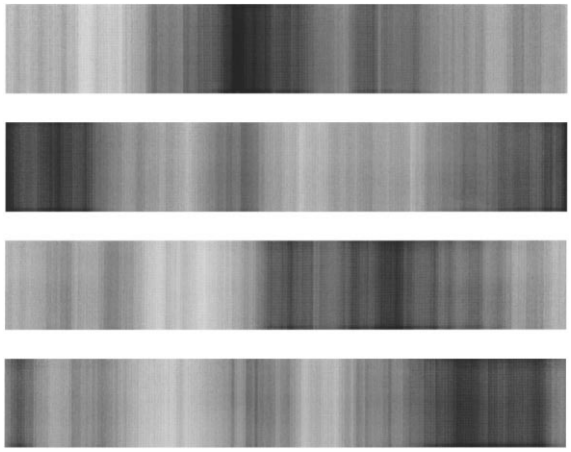

Figure 6. Example of random "natural" patterns. These one-dimensional stimuli have a spatial frequency power spectrum $R(f_s) = 1/f_s^{2.3}$ which is observed in natural images. Each pattern has a different, randomly-generated phase spectrum.

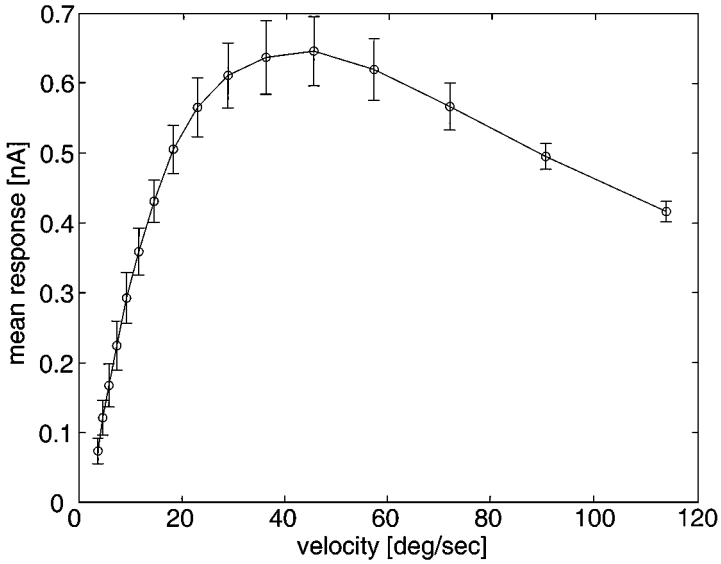

Figure 7. Chip response to $1/f_s^{2.3}$ patterns. Ten random patterns with natural image statistics (see Fig. 6) were presented to the EMD array. The mean response of the chip to these stimuli is shown as a function of pattern velocity. Error bars denote standard deviation across the mean response from all ten patterns.

and constitute a valuable step towards evaluating the robustness of a vision sensor.

We presented ten different $1/f_s^m$ stimuli to our silicon EMD array. We varied the velocity of each pattern and measured the steady-state chip response for all ten random pattern presentations. In Fig. 7, the mean and standard deviation over all ten random patterns are plotted versus pattern velocity. The mean response is roughly proportional to pattern velocity for angular velocities less than 20 deg/s. The response is monotonic up to 45 deg/s and then begins to decrease, a consequence of the correlation motion detection algorithm.

As indicated by the error bars, the response varies little as we change the phase characteristics of the pattern. The Hassenstein-Reichardt EMD does not require explicit image features in order to estimate velocity. By duplicating this architecture in silicon, we create a robust sensor that is sensitive to natural stimuli over a broad range of velocities.

4. Optomotor System

As mentioned above, the fly uses visual motion information to stabilize its flight. Mismatch of body components or environmental disturbances may impart rotation on the animal, but sensory feedback is used to produce compensatory torque responses. This sensorimotor feedback is known as the optomotor system, and is one of the best-studied behaviors of the fly (Götz, 1975; Warzecha and Egelhaaf, 1996).

4.1. Hardware Implementation

We constructed a hardware implementation of the optomotor system using a two-wheeled vehicle. We chose to build a physical motor system so we could evaluate our sensor's performance in the real world. Figure 8 shows a schematic of the system. Our wide-field motion sensor estimates self-rotation, and this signal is used to produce a compensatory rotation by the drive motors.

We constructed a simple robotic platform on which we mounted the wide-field motion sensor (see Fig. 9). The robot had two large wheels driven independently by two DC motors, and a free-turning wheel in the back to maintain balance. Each drive motor was controlled with a pulse-width modulation circuit that varied the duty cycle of a constant-amplitude square wave voltage. By changing the duty cycle of the waveform, each motor could be driven at any speed up to a maximum. If the motors were driven at different speeds, the robot would drive in a curved trajectory.

A large asymmetry was introduced into the robot's mechanics by connecting the left and right motors to

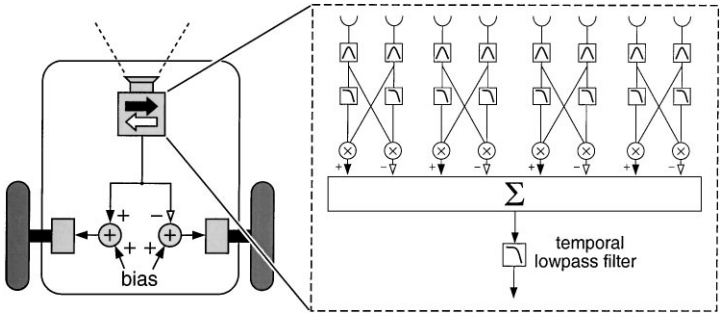

Figure 8. Schematic of our optomotor system. A motion sensor chip is mounted facing forward on a two-wheeled robotic platform. Forward-facing motion sensors are largely blind to optic flow produced by forward translation, so we only measure rotation. The chip's wide-field output is used as an estimate of self rotation, then lowpass filtered ($\tau = 750$ ms) to stabilize the control loop. This signal is added to one motor and subtracted from the other, producing a compensatory rotation. A constant motor bias produces forward translatory motion.

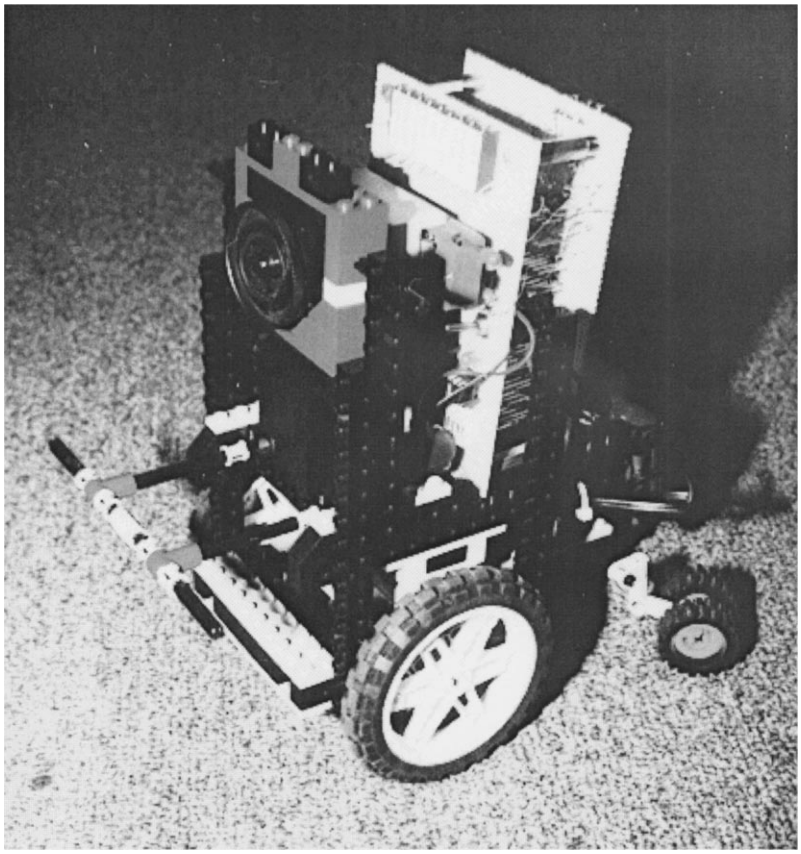

Figure 9. Photograph of the optomotor system. The lens is covering the motion sensor chip. Additional off-chip electronics have been constructed to drive the DC motors. The back wheels turn freely, and merely prevent the robot from falling over. The robot measures $13 \times 19 \times 22$ cm, and is powered by on-board batteries.

their respective wheels with different gear ratios. The left motor was connected to the left wheel with a $1:5$ gear ratio, while the right motor was connected to the right wheel with a $1:1$ gear ratio (see Fig. 10). This caused the robot to drive in tight circles if both motors were driven at the same speed. This asymmetry was made extreme for the sake of experiment, but perfect symmetry is impossible to achieve in any physical

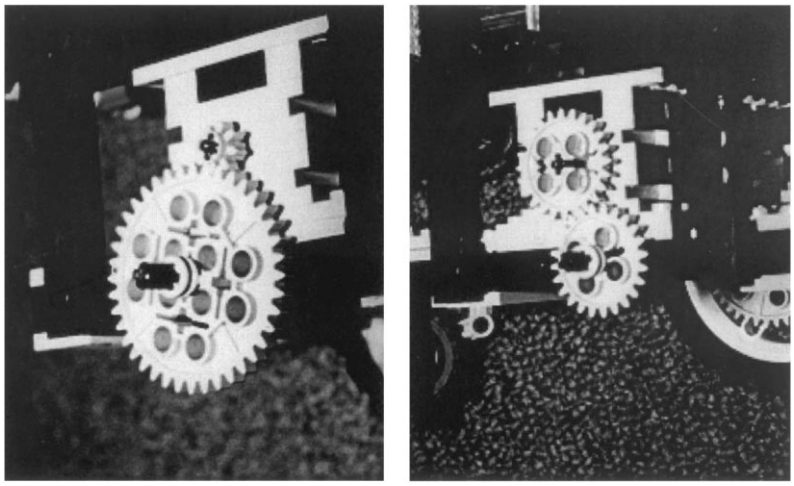

Figure 10. Asymmetrical gear ratios. The left and right motors drove their wheels with 1 : 5 and 1 : 1 gear ratios, respectively. This caused the robot to drive in tight circles if both motors were driven at the same speed.

robot. While two actuators may match perfectly in simulation, they will never match when built and tested in the real world. This difficulty is especially pronounced in outdoor terrain, where wheels or feet may slip in sand or mud. Legged robots are especially prone to walking in curved lines due to footslip or terrain differences, even if they have been designed and constructed with high precision (R. Quinn, personal communication).

When open-loop control falls short, we must introduce sensory feedback to further improve performance. Optic flow information has the potential to guide a robot in a straight path, because any deviation involves a yaw rotation, however slight. If yaw rotation can be estimated from optic flow reliably, we can use this as an error signal in a negative feedback loop in which the motors execute a compensatory rotation to null the sensory error signal.

We constructed a feedback loop of this type using our VLSI wide-field motion sensor. The sensor was mounted facing forward on the robot, oriented so it was sensitive to horizontal motion. We oriented the sensor facing straight ahead since translatory motion by the robot produces little optic flow in the direction of travel, while rotatory (yaw) motion produces uniform optic flow around the visual field parallel to the ground. Thus the optic flow in the forward region of the visual field will be dominated by the rotatory component. The hoverfly *Syritta pipiens* uses this strategy to stabilize its flight. When moving forward, the animal uses optic flow from the forward region of the visual field to estimate self rotation. This agile creature is also capable of flying sideways, and when doing so it uses optic flow from the lateral visual fields to estimate self rotation (Collett, 1980). Presumably, it is attempting to measure optic flow in the regions least contaminated with optic flow produced by its own translation.

The output of our motion sensor was a continuous, time-varying voltage. This signal was filtered by a first-order lowpass filter with a time constant of 750 ms. This is a simple model of the relationship between the output of a wide-field motion-sensitive neuron in the fly and the torque response produced by the wings (Egelhaaf, 1987; Warzecha and Egelhaaf, 1996). The filtered output of the motion sensor was added to the left motor command and subtracted from the right motor command (see Fig. 8). This has the effect of adding a rotatory component to the robot's trajectory. In the absence of visual feedback, both motors turned at the same rate (so one wheel turns five times faster than the other). Visual feedback slowed one wheel and sped up the other.

4.2. Experiments

Experiments were performed indoors in our laboratory, but the visual nature of the room was not altered in any way to accommodate the motion sensor. The room was a typical cluttered laboratory environment with many shady areas under tables (see Fig. 11). The robot's position was recorded 10–20 times per second with a magnetic field tracking system that returned location and orientation in three dimensions (Polhemus,

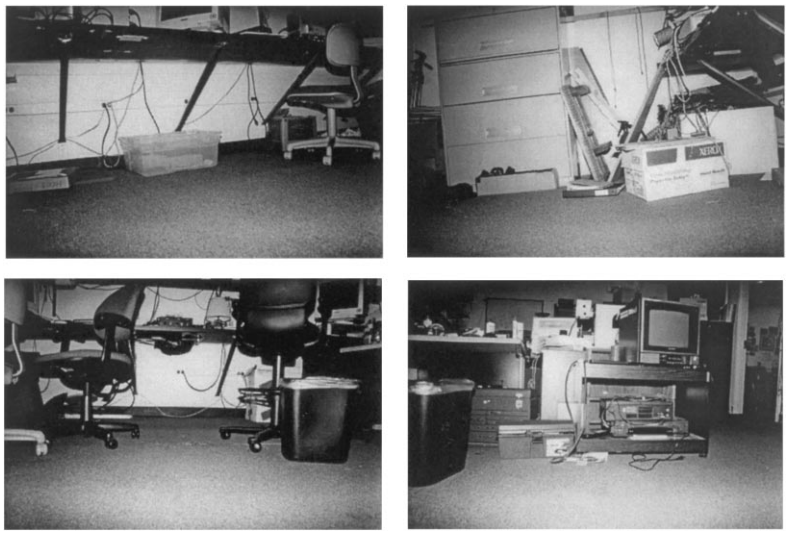

Figure 11. Visual environment during the experiments. The robot saw a typical, cluttered laboratory environment which was not changed in any way to accommodate the motion sensor.

Colchester, VT). The scale of experiments was limited by the range of this system, approximately a 70 × 140 cm area for highest accuracy.

The optic flow feedback proved capable of nearly eliminating the effect of physical asymmetry. Figure 12 shows one trial without visual feedback. The line shows the robot's path, and the circle indicates the ending position. The robot is turning in tight circles. Figure 13 shows ten trials where visual feedback has been enabled. In general, the robot travels in straight lines. We purposely started the robot at different orientations to demonstrate that the sensor works well for general visual scenes around a room. When moving in straight lines, the robot traveled at a speed of approximately

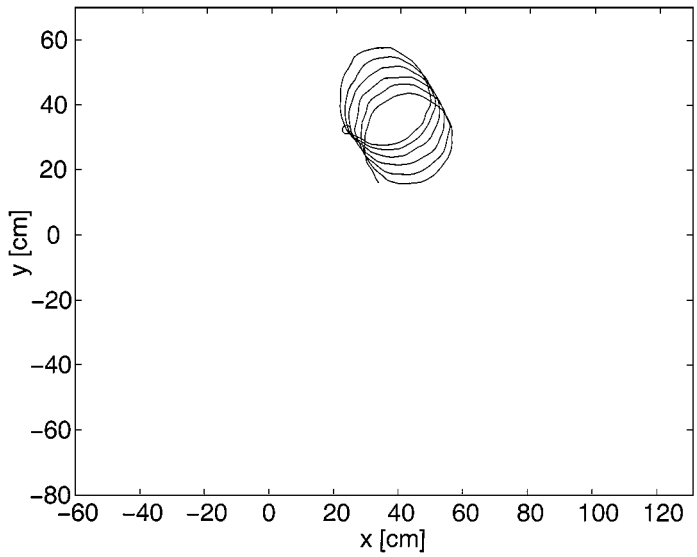

Figure 12. Robot path with no sensory feedback. With the motion sensor disabled, the robot turns in circles due to the asymmetry in its mechanics. The circle denotes the ending location of the robot.

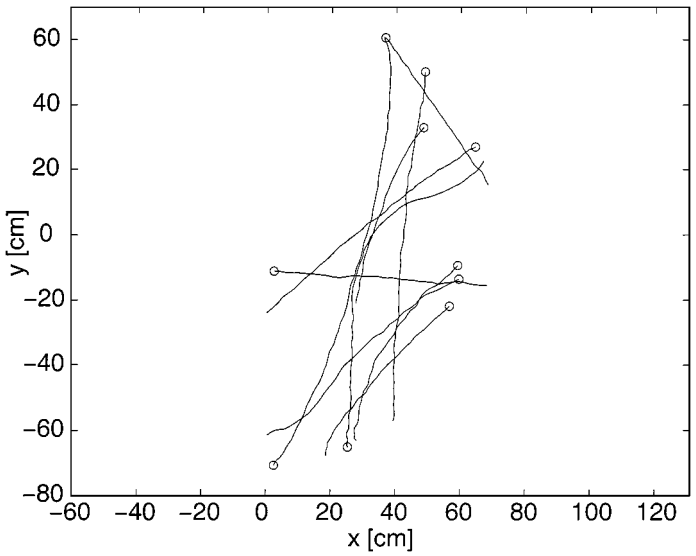

Figure 13. Robot path with sensory feedback. Ten trials are shown where the motion sensor provided optomotor control to straighten the course of the robot despite an extreme mechanical asymmetry. Circles denote the ending location of the robot in each trial. The robot was exposed to different visual scenes during the trials.

20 cm/s. Objects and walls were typically 0.2 to 1.5 meters away from the robot, depending on the direction.

The angular velocity of the robot (yaw rate) was computed along each path by differentiating the robot's heading as recorded by the tracking system. Figure 14 shows a histogram of angular velocities for the trials without feedback and all ten trials with visual feedback. The mean angular velocity in the open-loop case is −116 deg/s, while for the closed-loop case this decreased to −3.7 deg/s, an improvement by a factor of 31.

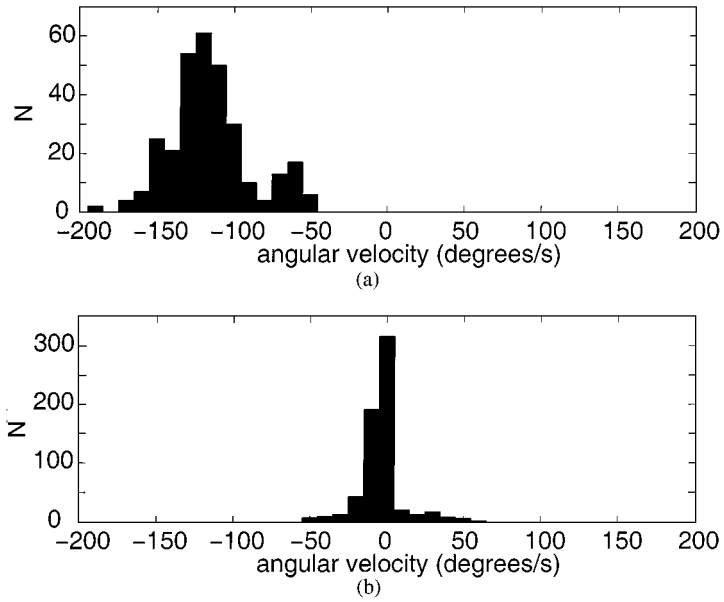

Figure 14. Histogram of angular velocities. (a) No visual feedback. The turning behavior of the robot is obvious. The mean angular velocity was −116 deg/s. (b) Compilation of all ten trials with visual feedback. The mean angular velocity was greatly reduced to −3.7 deg/s.

Occasionally, the feedback did fail to keep the course straight. A 45° turn is visible in Fig. 13, most likely caused by the sensor being oriented toward a relatively featureless part of the room, where no motion information is available. A larger field of view would reduce the likelihood of such occurrences. Also, the magnitude of the error depends on the degree of asymmetry in the gear ratios. In a more realistic situation with higher open-loop precision, it is likely that large closed-loop errors would be rare.

5. Discussion

We have evaluated the performance and robustness of a single-chip silicon motion sensor and demonstrated its use for an autonomous mobile robot. By modeling biological motion detectors, we gain many of the advantages that have been "built in" by evolution and optimized the sensor for real-time operation in complex environments. Roboticists have long realized that when placed in the real world, their artificial creatures lack a robustness that biological creatures possess. Much of this robustness is likely achieved by sensing the world *in the right way.* We believe that our silicon EMD array represents a significant step toward building hardware sensors based on algorithms and architectures found in biological systems.

Our sensor is small and extremely low power, making it easily adaptable to mobile robot applications. While the power budget on a robot is usually dominated by motors, traditional CCD imagers consume significant amounts of power, and digital microprocessors or DSPs capable of processing real time video consume even more. On the Sojourner rover of the recent Pathfinder Mars mission, the CCD imagers alone consumed 0.75 W, 5% of the total power budget at peak solar cell output. The CPU system consumed an additional 24%, and much of the CPU's time was devoted to processing static images while the rover was not moving (Matthies et al., 1995). By comparison, our EMD array consumed less than 5 μW of power. Traditional imaging and image processing is expensive in terms of time, size, and power. Biologically-inspired analog VLSI approaches to this problem can bring down the cost and make robot vision more practical.

One consequence of building an integrated motion sensor that includes photoreceptors and motion processing on the same chip is low resolution. Modern CCD imagers with resolutions of 640×480 or higher are common. By placing motion processing circuitry next to each photoreceptor, we greatly increase the size of our pixels (see Fig. 2). (Multichip systems are capable of overcoming this problem, but the high-bandwidth communication between imagers and motion processors negate many of the power and size advantages mentioned above.) However, we argue that while high-resolution imagers may be necessary for face recognition, low-resolution imagers are sufficient for many sophisticated visually-guided behaviors. If we look across fly species we find eyes with 600–8000 photoreceptors, equivalent to two-dimensional cameras with 25×25 to 90×90 pixels. We are able to build motion sensor arrays within this range today. Consider also that each eye covers almost one entire visual hemifield. The visual acuity of flies is 100–500 times lower than in humans, yet these animals are capable of extraordinary maneuvers as they navigate through a dynamic, unstructured environment.

We are hopeful that our silicon EMD circuit can be applied to more sophisticated visual capabilities exhibited by flying insects. EMD-based models have been proposed to explain the landing response of flies, where legs are extended milliseconds before contacting the ground (Borst, 1990). Other EMD-based models explain small-object detection by relative motion in the fly (Reichardt et al., 1983), and the ability of honeybees to detect and land on textural edges (Kern et al., 1997). Motion-sensitive cells in the locust respond strongest to objects moving on a collision course with the animal, and trigger evasive behaviors (Hatsopoulos et al., 1995). These models could be implemented in analog VLSI using the EMD circuit demonstrated here, and would provide useful visual capabilities for minimal power and size costs.

Acknowledgments

This work was supported by the Center for Neuromorphic Systems Engineering as a part of NSF's Engineering Research Center program, and by ONR. R.H. was supported by an NDSEG fellowship from ONR and by the NIMH.

References

Adelson, E.H. and Bergen, J.R. 1985. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am.*, A 2:284–299.

Borst, A. 1990. How do flies land? *BioScience*, 40:292–299.

Borst, A. and Egelhaaf, M. 1989. Principles of visual motion detection. *Trends Neurosci.*, 12:297–306.

Borst, A. and Egelhaaf, M. 1990. Direction selectivity of blowfly motion-sensitive neurons is computed in a two-stage process. *Proc. Natl. Acad. Sci. USA*, 87:9363–9367.

Collett, T.S. 1980. Some operating rules for the optomotor system of a hoverfly during voluntary flight. *J. Comp. Physiol.*, 138:271–282.

Delbrück, T. 1993. Silicon retina with correlation-based, velocity-tuned pixels. *IEEE Trans. Neural Networks*, 4:529–541.

Delbrück, T. and Mead, C.A. 1996. Analog VLSI phototransduction by continuous-time, adaptive, logarithmic photoreceptor circuits. CNS Memo 30, California Institute of Technology.

Dong, D.W. and Atick, J.J. 1995. Statistics of natural time-varying images. *Network*, 6:345–358.

Douglas, R., Mahowald, M., and Mead, C. 1995. Neuromorphic analogue VLSI. *Ann. Rev. Neurosci.*, 18:255–281.

Duchon, A.P., Kaelbling, L.P., and Warren, W.H. 1998. Ecological robotics. *Adaptive Behavior*, 6:473–507.

Egelhaaf, M. 1985. On the neuronal basis of figure-ground discrimination by relative motion in the visual system of the fly: II. figure-detection cells, a new class of visual interneurones. *Biol. Cybern.*, 52:195–209.

Egelhaaf, M. 1987. Dynamic properties of two control-systems underlying visually guided turning in house-flies. *J. Comp. Physiol.*, A 161:777–783.

Egelhaaf, M. and Borst, A. 1989. Transient and steady-state response properties of movement detectors. *J. Opt. Soc. Am.*, A 6:116–127.

Egelhaaf, M. and Borst, A. 1993. A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons. *J. Neurosci.*, 13:4563–4574.

Egelhaaf, M., Borst, A., and Reichardt, W. 1989. Computational structure of a biological motion-detection system as revealed by local detector analysis in the fly's nervous system. *J. Opt. Soc. Am.*, A 6:1070–1087.

Etienne-Cummings, R. and Van der Spiegel, J. 1996. Neuromorphic vision sensors. *Sensors and Actuators*, A 56:19–29.

Field, D.J. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am.*, A 4:2379–2394.

Franceschini, N., Pichon, J.M., and Blanes, C. 1992. From insect vision to robot vision. *Phil. Trans. R. Soc.*, B 337:283–294.

Gibson, J.J. 1950. *The Perception of the Visual World.*, Houghton-Mifflin: Boston.

Götz, K.G. 1975. The optomotor equilibrium of the *Drosophila* navigation system. *J. Comp. Physiol.*, 99:187–210.

Haag, J. and Borst, A. 1997. Encoding of visual motion information and reliability in spiking and graded potential neurons. *J. Neurosci.*, 17:4809–4819.

Harrison, R.R. and Koch, C. 1998. 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), MIT Press: Cambridge, Mass., pp. 880–886.

Hassenstein, B. and Reichardt, W. 1956. Systemtheoretische Analyse der Zeit-, Reihenfolgen-, und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers, *Chlorophanus*. *Z. Naturforsch.*, 11b:513–524.

Hatsopoulos, N., Gabbiani, F., and Laurent, G. 1995. Elementary computation of object approach by a wide-field visual neuron. *Science*, 270:1000–1003.

Hausen, K. and Egelhaaf, M. 1989. Neural mechanisms of visual course control in insects. In *Facets of Vision*, D.G. Stavenga and R.C. Hardie (Eds.), Berlin: Springer-Verlag.

Higgins, C., Deutschmann, R., and Koch, C. 1999. Pulse-based 2D motion sensors. *IEEE Trans. Circuits and Systems* II, 46:677–687.

Huber, S.A. 1997. Studies of the visual orientation behavior in flies using the artificial life approach. Ph.D. Thesis, Universität Tübingen, Germany.

Kandel, E.R., Schwartz, J.H., Jessell, T.M. 1991. *Principles of Neural Science*, Appleton & Lange: Norwalk, Conn.

Kern, R., Egelhaaf, M., and Srinivasan, M.V. 1997. Edge detection by landing honeybees: behavioural analysis and model simulations of the underlying mechanism. *Vis. Research*, 37:2103–2117.

Kimmerle, B., Warzecha, A.-K., and Egelhaaf, M. 1997. Object detection in the fly during simulated translatory flight. *J. Comp. Physiol.*, A 181:247–255.

Krapp, H.G. and Hengstenberg, R. 1996. Estimation of self-motion by optic flow processing in single visual interneurons. *Nature*, 384:463–466.

Lewis, M.A. 1998. Visual navigation in a robot using zig-zag behavior. In *Advances in Neural Information Processing Systems 10*, M.I. Jordan, M.J. Kearns, and S.A. Solla (Eds.), MIT Press: Cambridge, Mass., pp. 822–828.

Matthies, L., Gat, E., Harrison, R., Wilcox, B., Volpe, R., and Litwin, T. 1995. Mars microrover navigation: performance evaluation and enhancement. *Autonomous Robots*, 2:291–311.

Mead, C. 1989. *Analog VLSI and Neural Systems*, Reading, Mass.: Addison-Wesley.

Moini, A., Bouzerdoum, A., Eshraghian, K., Yakovleff, A., Nguyen, X.T., Blanksby, A., Beare, R., Abbott, D., and Bonger, R.E. 1997. An insect vision-based motion detection chip. *IEEE J. Solid State Circuits*, 32:279–284.

O'Carroll, D.C., Bidwell, N.J., Laughlin, S.B., and Warrant, E.J. 1996. Insect motion detectors matched to visual ecology. *Nature*, 382:63–66.

Reichardt, W., Poggio, T., and Hausen, K. 1983. Figure-ground discrimination by relative movement in the visual system of the fly, part II: towards the neural circuitry. *Biol. Cybern.*, 46:1–30.

Reichardt, W. and Egelhaaf, M. 1988. Properties of individual movement detectors as derived from behavioral experiments on the visual system of the fly. *Biol. Cybern.*, 58:287–294.

Ruderman, D.L. and Bialek, W. 1994. Statistics of natural images: scaling in the woods. *Phys. Rev. Lett.*, 73:814–817.

Sarpeshkar, R., Kramer, J., Indiveri, G., and Koch, C. 1996. Analog VLSI architectures for motion processing: from fundamental limits to system applications. *Proc. of the IEEE*, 84:969–987.

Single, S. and Borst, A. 1998. Dendritic integration and its role in computing image velocity. *Science*, 281:1848–1850.

Srinivasan, M.V., Chahl, J.S., and Zhang, S.W. 1997. Robot navigation by visual dead-reckoning: inspiration from insects. *Intl. J. Patt. Recog. and Art. Intelligence*, 11:35–47.

Warzecha, A.-K. and Egelhaaf, M. 1996. Intrinsic properties of biological motion detectors prevent the optomotor control system from getting unstable. *Phil. Trans. R. Soc.*, B 351:1579–1591.

Weckström, M., Juusola, M., and Laughlin, S.B. 1992. Presynaptic enhancement of signal transients in photoreceptor terminals in the compound eye. *Proc. R. Soc. Lond.*, B 250:83–89.

Reid R. Harrison was born in DeFuniak Springs, Florida, in 1972. He received the B.S. degree in electrical engineering from the University of Florida in 1994. After working for a short time at the Jet Propulsion Laboratory and Los Alamos National Laboratory, he joined the Computation and Neural Systems program at the California Institute of Technology, where he is currently a Ph.D. candidate. His research interests include silicon implementations of biological sensory systems, low-power analog VLSI, and floating gate CMOS circuits.

Christof Koch was born in Kansas City, Missouri, in 1956 and studied physics and philosophy at the University of Tübingen in West Germany. He was awarded his Ph.D. in biophysics from the Max Planck Institute for Biological Cybernetics in Tübingen in 1982. He worked until 1986 at MIT's Artificial Intelligence Laboratory before joining the Computation and Neural Systems (CNS) program at the California Institute of Technology where he is now a Full Professor and Executive Officer of the CNS department. His research interests include the biophysical mechanisms underlying information storage and processing in individual neurons, neuromorphic vision chips, and the neuronal basis of visual awareness and consciousness.