An Analog VLSI Motion Sensor Based on the Fly Visual System

Thesis by

Reid R. Harrison

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2000

(Submitted May 9, 2000)

Acknowledgements

I would first like to thank my parents, Don and Lesa Harrison, for instilling a love of science in me at a very young age, for nurturing my intellectual interests for as long as I can remember, and for providing unconditional love and support as I explored and grew.

I owe many thanks to my advisor and *Doktorvater* Prof. Christof Koch, who gave me both freedom and support to pursue my vision. His guidance and encouragement made this work possible.

I thank my thesis committee: Prof. Michael Dickinson, Prof. Rodney Goodman, Prof. Gilles Laurent, and Prof. Pietro Perona for reading my dissertation and providing a supportive environment. I also thank Prof. Jim Bower and Prof. Mark Konishi for serving on my Ph.D. candidacy committees. I thank the Defense Advanced Research Projects Agency (DARPA), the Office of Naval Research (ONR), the National Institute of Mental Health (NIMH), and the National Science Foundation's Engineering Research Center (ERC) for Neuromorphic Systems Engineering at Caltech, for funding my research over the years.

Many people have enriched my academic experience during the past five years, and some have become good friends. Among those contributing to my growth as a scientist and engineer are: Axel Borst, Tobi Delbrück, Rainer Deutschmann, Chris Diorio, Rodney Douglas, Martin Egelhaaf, Volker Gauck, Jürgen "Bulle" Haag, Paul Hasler, Chuck Higgins, Tim Horiuchi, Giacomo Indiveri, Vincent Koosh, Jörg Kramer, Oliver Landolt, Shih-Chii Liu, Ania Mitros, Brad Minch, Alberto Pesavento, Rahul Sarpeshkar, Theron Stanford, and Barbara Webb. I thank you for sharing your knowledge and enthusiasm with me. Thanks are also due to Candi Maechtlen and Margaret Lindstrom for administrative assistance and advice.

I would also like to thank many friends at Caltech who have kept me sane and given me fond memories of my time in Pasadena: Stephen Glade, Adrian Hightower, Peter Bogdanoff, Andy Waniuk, Anthony Leonardo, Michael Gibson, Fidel Santamaria-Perez, Zsuzsi Hamburger, Alyssa Apsel, Angie Louie, Eric Slimko, Grace Chang, Sarah Laxton, Jennie Yoder, and Mika Nyström.

Finally, I would like to thank Lanny for her love, patience, and support, and for making everything worthwhile.

Abstract

Vision is a vitally important sense for flying insects. Over half of the 350,000 neurons in the housefly's brain are believed to have some role in visual processing. Flies use visual motion cues to navigate through turbulent air, avoid obstacles, chase other flies, and land safely. Much is known about the neural circuitry that extracts motion information from retinal light intensity signals. This dissertation describes the development and testing of integrated silicon sensors that estimate visual motion using architectures derived from the neurophysiology of the fly optic lobe. We built VLSI systems incorporating light sensors and information processing circuits side by side on the same chip. These continuous-time analog CMOS circuits operate in the weak inversion (subthreshold) regime to match biological time constants and achieve sub-milliwatt power dissipation. Detailed characterization showed our sensor to be an accurate implementation of the Hassenstein-Reichardt motion detector model, originally developed to describe insect visual responses. We developed a novel test paradigm using stimuli with natural image statistics and spatiotemporal noise to evaluate the sensor's robustness. Our sensors were able to discriminate motion direction using naturalistic stimuli in noisy conditions (SNR < 1). Information theoretic techniques were used to measure the ability of our sensor to encode time-varying image velocity. Coding efficiency was quantified and compared with results from motion-sensitive neurons in the fly. The silicon system was tested in the context of a visually guided behavior—the optomotor stabilization response. Direct comparisons with the fly were made in real-time, closed-loop control experiments. A circuit architecture was developed to model the biophysical properties of wide-field motion-sensitive neurons in greater detail. This gave our sensor array nonlinear spatial integration properties that decreased sensitivity to gaps in the optic flow field. Finally, we investigated the issue of sensory fusion and explored a circuit that could assist in the integration of visual motion sensors with other sensors, such as angular rate gyroscopes.

Contents

List of Figures

Chapter 1 Introduction

Engineers have long looked to nature for inspiration. The diversity of life produced by five billion years of evolution provides countless existence proofs of organic machines with abilities that far surpass those of our own relatively crude automata. We have learned how to harness large amounts of energy and thus far exceed the capabilities of biological systems in some ways (e.g., supersonic flight, space travel, and global communications). However, biological information processing systems (i.e., brains) far outperform today's most advanced computers at tasks involving real-time pattern recognition and perception in complex, uncontrolled environments. If we take energy efficiency into account, the performance gap widens. The human brain dissipates 12 W of power, independent of mental activity. A modern microprocessor dissipates around 50 W, and is equivalent to a vanishingly small fraction of our brain's functionality.

Indeed, the human brain is a daunting goal for biologists and engineers alike. Our brain takes several years to fully develop, and contains between 10^{10} and 10^{11} neurons (nerve cells), each communicating with 10^3 other cells, on average. Brains of other animals (particularly invertebrates) are much smaller but still perform remarkably complex computations. Insect brains, for example, typically contain between 10^5 and 10^6 neurons. As we shall see in the following chapter, insects perform sophisticated information-processing tasks rapidly and efficiently.

In this body of work, we have attempted to extract computational principles from the visual system of the fly and apply these principles to an engineered system—an integrated, low-power visual motion sensor. As our engineering tool we use very-large scale integration (VLSI) of silicon circuits—the most advanced information-processing substrate available today. In particular, we explore continuous-time (unclocked), continuous-value (analog) circuit architectures. This approach was pioneered by Mead and colleagues beginning the in 1980s (Mead, 1989).

We begin by reviewing the biological system of interest—the fly—in Chapter 2, paying special attention to its motion-sensitive visual system. In Chapter 3, we review motion detection algorithms and previous hardware implementations of motion detectors. We introduce our VLSI designs in Chapter 4, and present characterization results in Chapter 5. Chapter 6 describes experiments that quantify the encoding ability of our sensor and compare it to neuronal encoding ability. We continue the direct comparison of our sensor to the fly in Chapter 7, where we evaluate a motion-mediated sensorimotor system. Chapter 8 introduces a circuit architecture that improves our sensor's performance by modeling the biophysics of a motion-sensitive neuron in the fly's optic lobe. Finally in Chapter 9, we discuss how our visual sensor might be combined with other sensory modalities.

Chapter 2 Sensory Systems of the Fly

The fly is an attractive target for biologically-inspired approaches to engineering. Its brain and sensory systems have been studied for decades, so much is known about their operation. Of course, we are still decades (or centuries) away from understanding the entire system, but a wealth of behavioral and electrophysiological data has led to the development of several models of information processing.

Flies possess a diverse array of organs for sensing their environment. In addition to the familiar sense of vision, flies employ Coriolis-force "gyroscopes," polarized light sensors, and body proprioception to aid in navigation. In this chapter, we will discuss these sensory systems.

2.1 The Visual System of the Fly

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 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).

Insects process visual motion information in a local, hierarchical manner. This information processing begins at the sensor—the retina (see Figure 2.1). 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 temporal 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 large tangential cells in the lobular plate (Hausen and Egelhaaf, 1989). The housefly has 50-60 tangential cells in each hemisphere 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

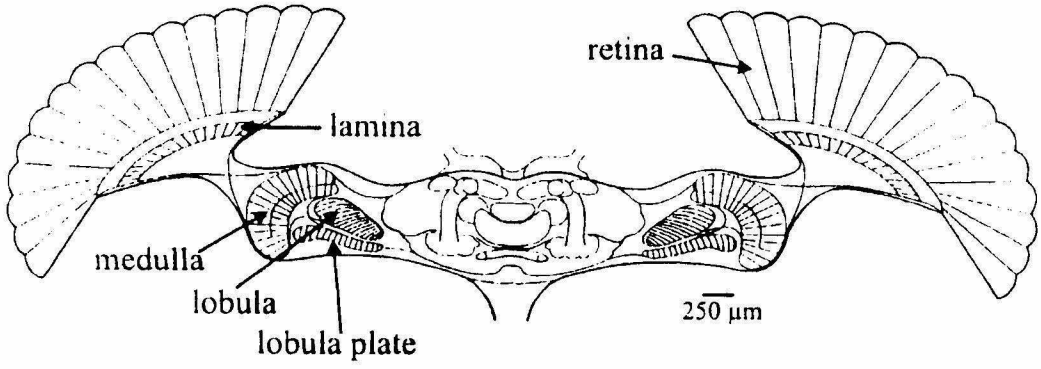

Figure 2.1: Central nervous system of the fly. Lenses in each compound eye focus light onto the retina. Photoreceptor signals are transmitted to the lamina, which emphasizes temporal change. A retinotopic arrangement is maintained through the medulla. The lobular plate contains wide-field, motion-sensitive tangential neurons that send information to the contralateral optic lobe as well as to the thoracic ganglia, which control the wings and legs. Adapted from Borst and Haag, 1996.

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.

2.1.1 Photoreception

Each eye of the blowfly *Calliphora erythrocephala* consists of approximately 6000 individual lenses. Beneath each lens is a cluster of eight light-sensitive cells. Each lens and its associated photoreceptors forms a unit called an ommatidium. Six of the eight photoreceptors are used to implement neural superposition, a technique to increase the effective lens diameter by pooling the responses of neighboring ommatidia. The other two photoreceptors do not seem to be involved in the detection of motion. Mutants with these photoreceptors impaired cannot discriminate colors, but show no motion-related deficits (Heisenberg and Buchner, 1977). From an information-processing perspective, each ommatidium records one "pixel" of the external world's image. Interommatidial angular spacing is 1.1-1.3° (Land, 1997). This angular resolution is approximately 150 times worse than the 0.008° resolution in foveated region of the human retina (Wandell, 1995). (This is roughly equivalent to having 20/3000 vision.)

Although the ommatidia are arranged in a hexagonal array, it is useful to think of the equivalent array size in terms of the standard rectangular array used by computer monitors and digital cameras. Taking the square root of the number of ommatidia (6000), we see that *Calliphora*'s eye is roughly equivalent to a 77 × 77 pixel array covering one visual hemifield. The eye of the fruit fly *Drosophila*

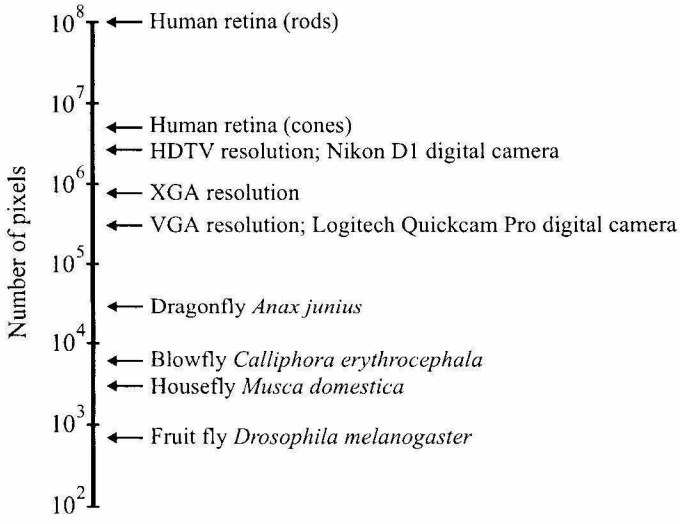

Figure 2.2: Number of pixels in biological and silicon systems. Current state-of-the-art silicon imagers have numbers of pixels approaching human retina levels. Flying insects get by with orders of magnitude fewer pixels. All data is for single eyes. Human retina data from Wandell, 1995; insect data from Land, 1997.

contains only 700 ommatidia, resulting in an equivalent array size of 26 × 26 (Land, 1997). Today's cheap digital cameras provide 640 × 480 pixel images, and emerging photo-quality digital cameras provide 1800 × 1200 pixels or more–around two orders of magnitude more photoreceptors than a fly's eye (see Figure 2.2). Typical cameras concentrate these pixels into a 40° field of view, while each fly eye sees nearly a complete hemisphere.

It is remarkable that flies are capable of such impressive navigation when one considers their low-resolution eyes. This limited spatial acuity is a consequence of the compound eye design. In order to increase spatial acuity, more ommatidia are required. However, the resolving capability of each ommatidium is limited by diffraction, so each lens must also be made larger. If we wanted to build a compound eye with the acuity of the human fovea (0.008°), it would have a radius of 11.7 meters! The visual acuity of the largest insect eye in nature (that of the aeschnid dragonfly) reaches 0.24° in its most acute zone, still 30 times coarser than the human fovea (Land, 1997).

While inferior to human eyes spatially, fly vision far exceeds ours temporally. Human vision is sensitive to temporal modulations up to 20 or 30 Hz, while fly photoreceptors respond to temporal frequencies as high as 300 Hz (Autrum, 1958).

2.1.2 Signal Processing in the Peripheral Optic Lobe

The laminar region, also called the first optic ganglion, contains cells that exhibit transient responses to step intensity changes. The large monopolar cells (LMCs) in this ganglion ignore the dc light

level but amplify temporal changes (Weckström et al., 1992). This highpass response has been shown to optimize information transfer through this region (Laughlin, 1994). Laminar cells do not exhibit motion-specific responses. There is a strong retinotopic organization from the retina through the lamina to the next layer, the medulla. Every ommatidia has an associated neural "cartridge" beneath it in these underlying ganglia, suggesting many identical processing units operating in parallel (Strausfeld, 1976).

Cells in this second optic ganglion are extremely small and difficult to record from, and little is know about their structure or function. DeVoe recorded from medullar cells in *Calliphora* and reported a wide variety of response characteristics: transient temporal responses, sustained responses, directional motion responses, and nondirectional motion responses (DeVoe and Ockleford, 1976; DeVoe, 1980).

2.1.3 The Tangential Cells of the Lobular Plate

The third optic ganglion is also known as the lobula-lobular plate complex. At this point in the optic lobe, the retinotopic organization ends with massive spatial convergence. Information from several thousand photoreceptors converges onto 50-60 tangential cells. These cells have broad dendritic trees that receive synaptic input from large regions of the medulla, resulting in large visual receptive fields (Hausen, 1982a; Hausen, 1982b; Hengstenberg, 1982; Hausen, 1984; Krapp and Hengstenberg, 1996).

A subset of these neurons were found to respond primarily to horizontal motion, and these cells were given names beginning with 'H'. H1 is a spiking neuron that responds to back-to-front optic flow. HSS, HSE, and HSN are graded potential (nonspiking) neurons covering the southern, equatorial, and northern regions of the visual field, respectively. Collectively called the HS cells, these neurons are depolarized by full-field visual motion from the front to the back of the eye, and hyperpolarized by back-to-front motion. They have been shown to encode horizontal motion as effectively as the spiking H1 cell (Haag and Borst, 1997). Each HS cell integrates signals from an ipsilateral retinotopic array of elementary motion detectors (EMDs), units in the medulla that estimate local motion in small areas of the visual field. The HS cells synapse onto descending, spiking neurons which relay information to the motor centers of the thoracic ganglion. Another class of neurons, the VS cells, respond to vertical motion. Recently, it has been shown that these HS and VS cells are not simply responsive to motion along one axis, but rather act as matched filters for complex patterns of optic flow that would be produced by body rotations (Krapp and Hengstenberg, 1996).

Four "figure detectors," or FD neurons, have been identified in the lobular plate. These cells respond more vigorously to small moving objects than to full-field motion (Egelhaaf, 1985).

2.1.4 Visually Guided Behaviors

Flies rely heavily on visual motion information to survive. In the fly, motion information is known to underlie many important behaviors including stabilization during flight, orienting towards small, rapidly-moving objects (Egelhaaf and Borst, 1993), and estimating time-to-contact for safe landings (Borst and Bahde, 1988). Some motion-related tasks like extending the legs for landing can be executed less than 70 ms after stimulus presentation. Wagner reports a 30 ms reaction time for male flies chasing prospective mates (Wagner, 1986). The computational machinery performing this sensory processing is fast, small, low-power, and robust.

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 standpoint because it extracts relevant information from a dynamic, unstructured environment on the basis of passive sensors and uses this information to generate appropriate motor commands during flight. This system is implemented in biological hardware that is many orders of magnitude smaller and more power efficient than CCD imagers coupled to a conventional digital microprocessor.

Much of the computation underlying the optomotor control system is performed by the HS cells (Geiger and Nässel, 1981; Geiger and Nässel, 1982; Egelhaaf et al., 1988; Hausen and Wehrhahn, 1990; Egelhaaf and Borst, 1993). This well-studied system estimates rotation from optic flow and uses this information to produce a stabilizing torque with the wings (Götz, 1975; Warzecha and Egelhaaf, 1996).

Flies also use visual motion information to coordinate landings. Behavioral and modeling studies indicate that such "time-to-landing" estimation could be produced by a temporal integration of HS-type neurons sensitive to expanding optic flow patterns (Borst and Bahde, 1988; Borst, 1990). A similar visual capability allows flies to avoid rapidly approaching predators. This escape response is sensitive to motion as well as to decreases in light intensity (Holmqvist and Srinivasan, 1991).

Behavioral experiments both with freely-flying and tethered flies demonstrate the ability to discriminate objects from background using relative motion (parallax) cues (Kimmerle et al., 1996; Kimmerle et al., 1996). The FD cells mentioned above are thought to underlie this capability.

2.2 The Vestibular Sense

Dipterans (true flies and mosquitos) possess a remarkable evolutionary specialization for measuring angular velocity. The hind wings of these animals evolved from flight surfaces into dedicated angular rate "gyroscopes." These halteres, as they are called, resemble small balls at the end of sticks (see Figure 2.3). The halteres beat up and down antiphase to the wings at the wingbeat frequency (about 150 Hz in *Calliphora*; over 200 Hz in the smaller *Drosophila*). They move at nearly constant velocity

during each upstroke and downstroke, covering nearly 180° (Nalbach, 1993).

While body rotations produce centrifugal forces on the halteres, these forces are oriented radially and tangentially, and for typical maneuvers are several orders of magnitude smaller than the radial centrifugal forces due to halteres oscillation. Centrifugal forces are proportional to the square of angular velocity and thus provide no information on the direction of rotation. A more useful signal is the Coriolis force, which is proportional to the cross product of the instantaneous haltere velocity and the axis of body rotation. Coriolis forces acting normal to the plane of haltere oscillation are detected by about 335 campaniform sensilla organized in five groups at the haltere base. These sensory cells are embedded in the flexible exoskeleton, and act as strain gauges.

By integrating Coriolis force information over the haltere's 180° sweep, and by combining signals from the two non-coplanar halteres, the fly can measure angular rotation about all three axes. Vestibular information from the halteres system is critical for maintaining stable flight; when a fly's halteres are removed it quickly falls to the ground. (With only one haltere removed, flight is still possible.)

Free flight seems to be controlled by a combination of visual and vestibular sensors. Little is known about the interactions between these two sensory modalities. Recent experiments by Dickinson and colleagues have shown that visual interneurons stimulate small haltere control muscles that exert force in the same direction as Coriolis forces (Chan et al., 1998). They propose a model where the halteres control flight equilibrium in a fast feedback loop and the slower visual interneurons steer the animal by tugging on the halteres to create a vestibular illusion.

2.3 Other Sensors

2.3.1 Ocelli

In addition to the two compound eyes, flies have three other photosensitive organs called ocelli. These sensors are located between the eyes on the dorsal region of the head. Each ocellus consists of a single circular lens approximately 75 μm in diameter which focuses light onto a low-resolution retina containing approximately 220 photoreceptors. The image produced on the retina is a wide-angle, underfocused view of the surroundings above and lateral to the fly. There is rapid convergence of the photoreceptors onto 4-6 interneurons that seem to measure mean brightness.

The ocelli seem to contribute to the *dorsal light response* observed behaviorally, where flies align the top of their head with the center of brightness (Schuppe and Hengstenberg, 1993). This often corresponds with the zenith in outdoor environments. Dragonflies appear to use their ocelli as horizon detectors to provide information on their head attitude relative to the ground (Stange, 1981). The ocelli seem to work in concert with the compound eyes to position the head.

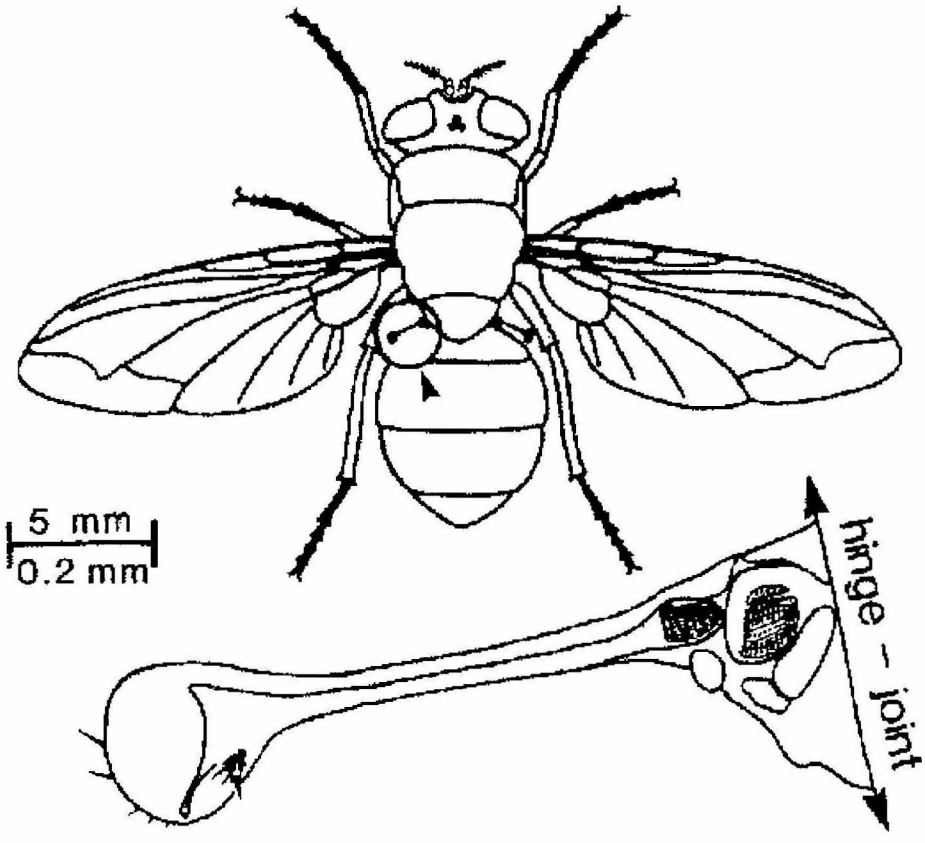

Figure 2.3: The halteres of the blowfly *Calliphora*. The halteres evolved from hindwings but no longer serve any aerodynamic function. Located in the "waist" between the thorax and the abdomen, these devices beat up and down (i.e., in and out of the plane of the page) antiphase to the wings. Groups of mechanoreceptors at the base measure Coriolis forces produces by the angular rotation of the animal. Adapted from Nalbach, 1993.

2.3.2 Polarized Light Detection

The dorsal regions of insect eyes contain polarization-sensitive photoreceptors. Bees and desert ants have been shown to use skylight polarization patterns as a compass and can infer their heading even when the sun and much of the sky is obscured by clouds. These specialized photoreceptors are most sensitive to ultraviolet light, which is more scattered and polarized by the atmosphere than longer-wavelength "visible" light (Wehner, 1987). Polarization-sensitive cells have been found in flies (Wolf et al., 1980), though polarization-sensitive behaviors have not been investigated in detail.

2.3.3 Linear Acceleration

Presumably, flying insects can also detect linear acceleration. While there are no known organs dedicated to this sense, the legs and neck of the fly are equipped with proprioceptive sensors that

measure both position and strain. Presumably, flies can sense the inertia of their head and limbs and infer acceleration.

In the remainder of this dissertation, we will focus primarily on visual motion sensing. Visual motion perception underlies many interesting behaviors in the fly and could be applied to useful engineering applications. The following chapter introduces motion detection algorithms, including the model commonly used to explain early vision in the fly.

In Chapter 9, we will revisit vestibular sensing and discuss how Coriolis-force sensors might be integrated with visual motion sensors.

Chapter 3 Motion Detection - Algorithms and VLSI Implementation

During the past 15 years, many analog, digital, and hybrid VLSI motion sensors have been developed and tested. Most of these designs incorporate photodetection and motion computation on the same chip. These focal-plane processors typically cannot achieve the high pixel density of dedicated CMOS imagers or CCDs, but rather trade off density for functionality. By extracting motion information at the level of light detection instead of using an external microprocessor, large savings in size, power, and system complexity is achieved.

Nearly every motion detection algorithm devised has been implemented in VLSI in some form. Motion detection algorithms can be divided into two broad classes: *feature-tracking* or *token-based* algorithms, and *intensity-based* algorithms. Models of motion detection in the fly represent a special case of intensity-based algorithms. In this chapter, we will review the principles of motion detection commonly used in both hardware implementations and biological models.

3.1 Feature-Based Motion Detection

Algorithms of this type use feature detectors to identify salient points in the raw image. Binary tokens indicating the absence or presence of a feature are then passed on to a velocity-estimation stage. Two types of feature detectors have been used in silicon motion sensors: spatial feature detectors and temporal feature detectors.

3.1.1 Motion Detectors Using Spatial Feature Detectors

Spatial feature detectors typically look for intensity edges—places where the image changes rapidly over a short distance. These edges are identified in successive discrete-time frames. Once features are identified in successive frames, each feature at time t must be matched up with the corresponding feature at time $t + 1$. This is known as the *correspondence problem* (see Figure 3.1). If identical features are too dense, this can be quite difficult. If features are sparse, the resulting optic flow field will have low resolution. By matching a feature from time t to the same feature at time $t + \Delta t$, velocity can be measured. This technique is commonly used in software motion detection schemes due to its compatibility with discrete-time computation, but it has also been used in hardware motion detectors as a front-end before additional continuous-time processing (Etienne-Cummings et

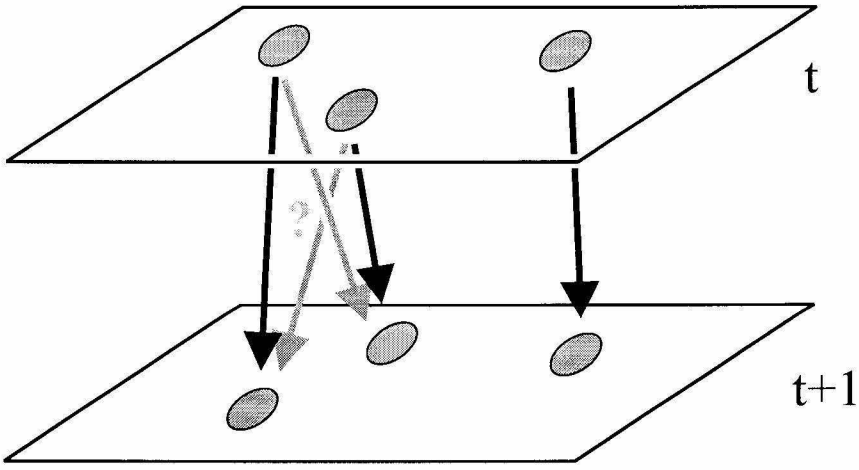

Figure 3.1: The correspondence problem. Identified features at time t must be matched with the corresponding features at time $t+1$. As denoted with the gray arrows, there can be multiple solutions if features are dense.

al., 1997; Barrows, 1998).

3.1.2 Motion Detectors Using Temporal Feature Detectors

Temporal feature detectors typically look for rapid changes in the image brightness at each pixel—temporal edges. Due to the local nature of this computation, it has been quite popular in analog VLSI approaches and has been implemented efficiently in continuous-time circuits (Kramer, 1996; Sarpeshkar et al., 1996; Higgins et al., 1999). These chips measure time-of-travel: the time it takes for an edge to pass from one pixel to an adjacent pixel. They offer the advantage of measuring true image speed over many orders of magnitude, and can operate at contrasts as low as 0.15 as long as the moving image contains sharp temporal edges (Kramer et al., 1997).

3.1.3 The Obliquity Problem

Motion detectors employing feature detectors suffer from what has been called the "obliquity problem" (Srinivasan and Zhang, 1993). A slow-moving stimulus that is nearly perpendicular to the sensor axis appears the same as a stimulus moving rapidly along the sensor axis, since both trigger two adjacent temporal feature detectors in the same way (see Figure 3.2). This can be overcome by communication with a perpendicularly-oriented sensor, but it is sensitive to errors: If one sensor registers the feature but the orthogonal sensor does not, the illusion of a fast-moving stimulus moving along one sensor axis can result.

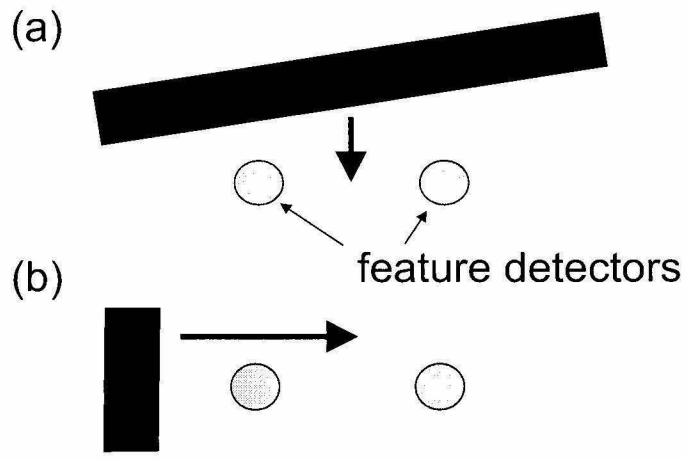

Figure 3.2: The obliquity problem. A pair of feature detectors perceives a slow-moving object nearly orthogonal to their axis (a) as equivalent to a fast-moving object aligned with the axis (b).

3.1.4 The Aperture Problem

A fundamental limitation of all local motion detectors with limited receptive fields is the ambiguity associated with measuring the trajectory of an edge. Only the component of motion orthogonal to the edge can be measured. This problem, known as the aperture problem, can be solved by spatially integrating motion information over larger regions of an image. One must identify rigid objects and analyze the motion information from several nonparallel edges of each object to resolve the ambiguity. There is good evidence that flies do *not* solve the aperture problem, at least at the level of wide-field lobular plate neurons (Borst et al., 1993).

3.1.5 Noise

Feature-tracking algorithms—especially those employing relatively simple feature detectors, as VLSI implementations must—may yield spurious responses to weak signals. In hardware motion detectors, features are typically encoded as binary entities which are either present or absent. Weak signals produce features near the threshold of detectability. Physically instantiated feature detectors have thresholds which are not perfectly matched across an array, so a weak signal may trigger one detector but not its neighbor. This makes the correspondence problem difficult and may yield inaccurate results (C. Higgins, personal correspondence).

3.2 Intensity-Based Motion Detection

Intensity-based algorithms use some filtered version of the image brightness, such as the photocurrent, for motion processing without prior binarization or classification. These algorithms can be

subdivided into two types: gradient and correlation algorithms.

3.2.1 The Gradient Model

Let $I(x, y, t)$ describe the image irradiance at the point (x, y) at time t. We can describe the optic flow at each point in terms of its x and y components: $u(x, y)$ and $v(x, y)$. From this, we would predict

$$I\left(x + u\delta t, y + v\delta t, t + \delta t\right) = I\left(x, y, t\right) \tag{3.1}$$

if the total irradiance I stays constant with time:

$$\frac{dI}{dt} = 0 \tag{3.2}$$

If we make the assumption that the image intensity varies smoothly across space and time, we can write the Taylor expansion of this equation to obtain

$$I\left(x, y, t\right) + \delta x \frac{\partial I}{\partial x} + \delta y \frac{\partial I}{\partial y} + \delta t \frac{\partial I}{\partial t} + e = I\left(x, y, t\right) \tag{3.3}$$

where e represents higher-order terms. Neglecting these terms and taking the limit as $\delta t \to 0$, we can obtain

$$\frac{\partial I}{\partial x}\frac{dx}{dt} + \frac{\partial I}{\partial y}\frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \tag{3.4}$$

Our optic flow components u and v can be written

$$u = \frac{dx}{dt} \tag{3.5}$$

$$v = \frac{dy}{dt} \tag{3.6}$$

So we can rewrite Equation 3.4 equation as

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \tag{3.7}$$

This is known as the *optical flow constraint equation*, and it tells us how to solve for the optical flow (u and v) from spatial and temporal derivatives (gradients) in the image (Klaus and Horn, 1986).

Gradient algorithms use spatial and temporal derivatives of an image to measure velocity under the assumption that the overall intensity of the image remains constant (see Equation 3.2). As long as this assumption holds, the gradient algorithm measures speed directly and independent of contrast. The use of derivatives makes this algorithm vulnerable to high-frequency noise. However, gradient-based hardware motion sensors have been developed (Tanner and Mead, 1986; Deutschmann and

Koch, 1998).

3.2.2 Correlation-Based Motion Detection

The other class of intensity-based motion detectors measure spatiotemporal correlations caused by moving objects. These algorithms include the spatiotemporal motion energy model of Adelson and Bergen (Adelson and Bergen, 1985) and the Reichardt motion detector, first proposed by Hassenstein and Reichardt in 1956 as a model of motion detection in insects (Hassenstein and Reichardt, 1956). This algorithm will be discussed in detail in the following section. A related algorithm was proposed by Barlow and Levick to explain direction-selective cells in the rabbit's retina (Barlow and Levick, 1965). This algorithm has been implemented in hardware using temporal feature detectors (Horiuchi et al., 1991) and in a purely intensity-based architecture (Benson and Delbrück, 1992).

Delbrück built a continuous-time, continuous-value correlation-based hardware motion detector (Delbrück, 1993b) that used delay lines and quadratic nonlinearities to compute a measure of spatiotemporal motion energy. There was no feature detection; rather, the raw output from the photoreceptors were used to compute motion. Due to the nature of the delay lines, an object moving across the chip's field of view caused a directional response to gradually build over space and time. The circuit exhibited a velocity-tuned response. Like many correlation-based algorithms, the response of this chip was highly contrast-dependent.

3.3 Hassenstein-Reichardt Model in Detail

This section describes in detail the model of motion detection in the fly. This model, commonly known as the Reichardt model, has been successful at explaining both detailed electrophysiological responses of motion-sensitive neurons to visual stimuli (Egelhaaf and Borst, 1989; Zanker, 1990) and visually-guided behavioral responses (Reichardt and Poggio, 1976;Reichardt and Egelhaaf, 1988; Borst, 1990; Warzecha and Egelhaaf, 1996). Modified versions of the Reichardt model have also been used to explain motion perception properties in vertebrates, including humans (Borst and Egelhaaf, 1989; Clifford et al., 1997).

3.3.1 Theoretical Analysis

We now describe the Reichardt motion detection algorithm which will underlie our hardware implementation. As mentioned above, the Reichardt motion detector is a correlation-based algorithm, whose output is equivalent to the output of the spatiotemporal motion energy model proposed by Adelson and Bergen (Adelson and Bergen, 1985) and others (van Santen and Sperling, 1985; Watson and Ahumada, 1985). The Reichardt model is in some ways a more efficient instantiation of the

motion energy model, as it requires fewer subtractions in the signal flow graph. When models are implemented in analog systems where component mismatch is inevitable, small amounts of component mismatch can result in sign reversal of the output if the inputs are of similar magnitudes.

The basic idea of the Reichardt motion detector is to correlate the signal from one photoreceptor with the delayed signal from an adjacent photoreceptor (see Figure 3.3a). This delay-and-correlate algorithm produces a velocity-tuned response that is weakly directionally selective. By subtracting the responses of two opponent half-detectors from each other, strong direction selectivity is achieved (Borst and Egelhaaf, 1990).

It is instructive to consider the case where the stimulus is a sinusoidal grating moving at velocity v. Image intensity $i(x,t)$ can be expressed as

$$i(x,t) = I + \Delta I \sin\left[2\pi f_s(x + vt)\right] \tag{3.8}$$

where I is the mean intensity, and f_s is the spatial frequency. The contrast of the grating is $\Delta I/I$. At any single photoreceptor, this moving grating produces a temporal sinusoidal signal with a frequency $f_t = vf_s$. This allows us to rewrite Equation 3.8 as

$$i(x,t) = I + \Delta I \sin(\omega_t t + \omega_s x) \tag{3.9}$$

where $\omega_t = 2\pi f_t$ and $\omega_s = 2\pi f_s$. If two photoreceptors have an angular separation of ϕ, then the signals measured by the photoreceptors can be expressed as

$$p_1(t) = |H(\omega_t)|\,\Delta I \sin\left(\omega_t t - \omega_s \frac{\phi}{2}\right) \tag{3.10}$$

$$p_2(t) = |H(\omega_t)|\,\Delta I \sin\left(\omega_t t + \omega_s \frac{\phi}{2}\right) \tag{3.11}$$

We introduce $H(\omega_t)$ as the temporal frequency response of the photoreceptors. For simplicity we ignore the phase contribution of $H(\omega_t)$ as it will be identical in $p_1(t)$ and $p_2(t)$, and thus have no effect on perceived motion. We also assume that the photoreceptors have a highpass behavior which eliminates the dc component of illumination I. We model the photoreceptor response as

$$H(\omega_t) = \frac{K j\omega_t \tau_H}{(j\omega_t \tau_H + 1)(j\omega_t \tau_{photo} + 1)} \tag{3.12}$$

where τ_H is the time constant of the dc-blocking highpass filter, τ_{photo} is the time constant defining the photoreceptor bandwidth, and K is a constant of proportionality.

The delay required by the delay-and-correlate motion detector architecture is implemented using the phase lag inherent in a first-order lowpass filter. Lowpass filtering each photoreceptor signal

yields

$$l_1(t) = \frac{|H(\omega_t)|\,\Delta I}{\sqrt{\tau^2{\omega_t}^2+1}} \sin\left(\omega_t t - \omega_s \frac{\phi}{2} - \tan^{-1}\tau\omega_t\right) \tag{3.13}$$

$$l_2(t) = \frac{|H(\omega_t)|\,\Delta I}{\sqrt{\tau^2{\omega_t}^2+1}} \sin\left(\omega_t t + \omega_s \frac{\phi}{2} - \tan^{-1}\tau\omega_t\right) \tag{3.14}$$

Correlation is accomplished by multiplying the phase lagged signals with adjacent, non-delayed signals. The results are two "half-detector" responses:

$$m_1(t) = G\left[\cos(\omega_s\phi + P) - \cos(2\omega_t t - P)\right] \tag{3.15}$$

$$m_2(t) = G\left[\cos(\omega_s\phi - P) - \cos(2\omega_t t - P)\right] \tag{3.16}$$

where

$$G = \frac{\left(|H(\omega_t)|\,\Delta I\right)^2}{2\sqrt{\tau^2{\omega_t}^2+1}} \tag{3.17}$$

$$P = \tan^{-1}\tau\omega_t \tag{3.18}$$

Once these signals are subtracted in opponency, the final output becomes

$$o(t) = (\Delta I)^2\,|H(\omega_t)|^2\,\frac{\tau\omega_t}{\tau^2{\omega_t}^2+1}\sin\phi\omega_s \tag{3.19}$$

This describes the sensitivity of a Reichardt motion detector to a sinusoidal grating with a particular contrast, temporal frequency, and spatial frequency. Notice that the response is a separable function of these three parameters. We can rewrite this equation to make the dependency on the grating velocity v explicit:

$$o(t) = (\Delta I)^2\,|H(\omega_s v)|^2\,\frac{\tau\omega_s v}{\tau^2{\omega_s}^2v^2+1}\sin\phi\omega_s \tag{3.20}$$

Although this response is direction selective [*i.e.*, the sign of $o(t)$ is equal to the sign of v], it does not encode velocity independent of spatial frequency and contrast. Notice that the $\sin\phi\omega_s$ term predicts spatial aliasing, as it becomes negative for $1/2\phi < f_s < 1/\phi$.

There are no time-dependent terms in this equation. This indicates a dc response to moving patterns. However, if the mean intensity of the image is not completely removed by prefiltering, or if the opponent subtraction is not perfectly balanced, oscillations at harmonics of the stimulus temporal frequency will be superimposed on the dc response. (For a complete analysis of the Reichardt motion detector in these nonideal cases, see (Egelhaaf et al., 1989).) These oscillations may be reduced by using an array of Reichardt motion detectors (see Figure 3.3b) and summing their responses. This has the effect of integrating over different phases of the stimulus and canceling pattern-dependent oscillations, and has been proposed as a single model of tangential cell integration (Reichardt and

Egelhaaf, 1988; Single and Borst, 1998).

3.3.2 Hardware Implementations

Early attempts to implement the intensity-based Reichardt architecture in silicon used translinear, current-mode circuits (Andreou et al., 1991; Harrison and Koch, 1998). As we showed in Section 3.3.1, the response of these traditional Reichardt motion sensors is affected strongly by contrast. Attempting to build contrast-independent Reichardt motion sensors, some have designed circuits that perform an initial binarization of the image based on temporal edges and then delay and correlate these digital signals (Moini et al., 1997; Jiang and Wu, 1999). These circuits would not be expected to perform well in noisy, low-contrast environments without additional image preprocessing. Another VLSI implementation involved continuous-level signal processing after the photoreceptors, but the final motion detector output was a binary value (Liu, 1997). Reichardt-inspired sensors have also been built in discrete hardware and used on mobile robots, although the particular implementation more closely resembled a feature-tracking, time-of-travel scheme (Pichon et al., 1989; Franceschini et al., 1992).

In the next chapter, we describe two circuit architectures we developed to implement Reichardt motion detectors in VLSI hardware.

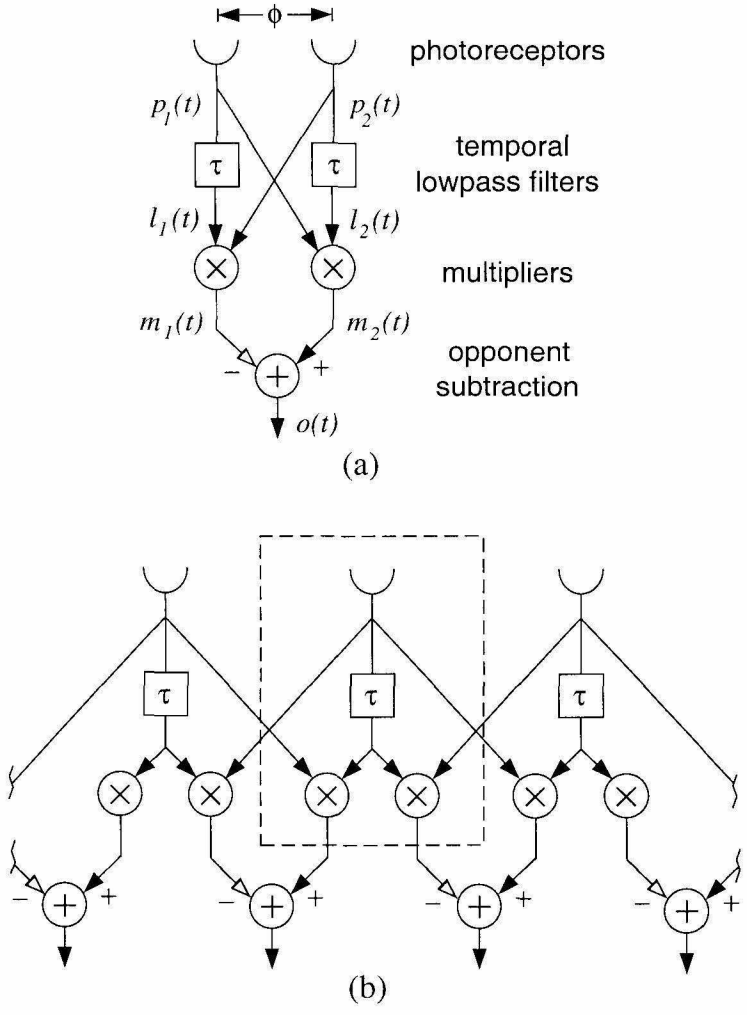

Figure 3.3: Reichardt motion detector architecture. (a) The signal from one photoreceptor is correlated with the delayed signal from an adjacent photoreceptor. Direction selectivity is increased by subtracting the responses of two half-detectors in opponency. (b) A 1-D array of Reichardt motion detectors, illustrating the repeated computational element. Subtraction currently performed off-chip for flexibility, but would be easy to implement given the current-mode outputs of the multipliers. A 2-D arrangement is possible with two additional multipliers in each cell.

Chapter 4 VLSI Reichardt Detector Design

We developed two distinct circuit architectures for VLSI Reichardt motion detectors. Both circuits operate in continuous time with analog signals, and incorporate light sensing and information processing on the same chip. The first circuit described is largely a *current-mode* design. That is, signals are represented as currents throughout the majority of the circuit. The second circuits is called the *voltage-mode* design since signals are represented a voltages, although the final output is produced as a current. Both circuits make use of the weak-inversion, or subthreshold region of operation of the MOS transistor for micropower operation.

4.1 Current-Mode Design

Each elementary motion detector (EMD) uses photodiodes as light sensors. We use a four-transistor adaptive photoreceptor circuit developed by Delbrück (Delbrück and Mead, 1996) that produces a continuous-time output voltage proportional to the logarithm of light intensity (Figure 4.1a). This circuit has a temporal lowpass characteristic with a cutoff frequency that can be set with a bias voltage. The photoreceptor is connected to a temporal derivative circuit (Mead, 1989) (Figure 4.1b), which has a highpass behavior. Transient firing, characteristic of a temporal highpass response, has been observed in fly laminar cells that receive input from retinal photoreceptors (Weckström et al., 1992). Together, the lowpass filtering of the photoreceptor and the highpass filtering of the temporal derivative circuit form a bandpass filter which improves performance by eliminating dc illumination (which contains no motion information), and attenuating high-frequency noise such as the 120 Hz flicker of ac incandescent lighting. These bandpass filters were set to attenuate frequencies below 2.8 Hz and above 10 Hz.

The temporal derivative circuit relies on a high-gain differential amplifier in a negative feedback configuration to keep the voltage on the capacitor equal to the input voltage. As the capacitor charges and discharges to maintain this equality, the currents through the two source follower transistors (labeled "sf" in Figure 4.1b) may be measured. The outputs of the temporal derivative circuit are these two unidirectional currents, which are proportional to the positive and negative components of temporal derivative of the input voltage. This resembles the ON and OFF channels found in many biological visual systems. One study suggests ON and OFF channels are present in the fly (Franceschini et al., 1989), but the evidence is mixed (Egelhaaf and Borst, 1992). This two-channel representation is useful for current-mode circuits, since the following translinear circuits work only with unidirectional currents. It should be noted that the use of ON and OFF channels introduces

nonlinearities into the circuit that are not accounted for in the simple model described in Chapter 3.

We use the phase lag inherent in a first-order lowpass filter as a time delay. The currents from the temporal derivative circuit are passed to current-mode first-order lowpass filter circuits (Figure 4.1c) (Himmelbauer, 1996). These are log-domain filters that take advantage of the exponential behavior of field-effect transistors (FETs) in the subthreshold (weak inversion) region of operation. Note that two filters are needed for each EMD—one for the ON channel, and one for the OFF channel, which are processed in parallel. The time constant of the filters is controlled with a bias current that can be set externally. This time constant can be changed to tune the EMD to a specific optimal temporal frequency. We fixed this time constant to 40 ms, which gave our chip a maximum temporal frequency sensitivity of 4 Hz, similar to motion-sensitive neurons in flies (O'Carroll et al., 1996).

To correlate the delayed and non-delayed signals for motion computation, we use a current-mode multiplier circuit (Figure 4.1d). This circuit also takes advantage of the exponential behavior of subthreshold FETs to perform a computation. Two diode-connected FETs convert the input currents into log-encoded voltages. The weighted sum of these voltages is computed with the capacitive divider on the floating gate of the output transistor, and this transistor exponentiates the summed voltages into the output current, completing the multiplication. Any trapped charge remaining on the floating gates from fabrication is eliminated by exposing the chip to ultraviolet light, which imparts sufficient energy to the trapped electrons to allow passage through the surrounding insulator. This circuit represents one of a family of floating-gate MOS translinear circuits developed by Minch that are capable of computing arbitrary power laws with current-mode signals (Minch et al., 1996b).

After the multiplication stage, the currents from the ON and OFF channels are summed, and the final subtraction of the left and right channels was done off-chip. Due to transistor mismatch, there was a gain error of approximately 2.5 between the left and right channels that was compensated for manually. It is interesting to note that there is no significant offset error in the output currents from each channel. This is a consequence of using translinear circuits which typically have gain errors due to transistor mismatch, but no fixed offset errors.

One entire EMD (left and right channels) consists of 31 transistors and 25 capacitors with 8.0 pF of total capacitance. Most of the capacitors were small devices (8 μm $\times$ 8 μm or less) associated with the floating-gate multiplier circuits. Each EMD takes 0.044 mm^2 of silicon area in a 2.0-μm CMOS process, including the integrated photoreceptors. By operating most of the transistors in the subthreshold regime, we achieve extremely low power dissipation (approximately 7.5 μW per elementary motion detector).

We built a simple model of the HS cell by constructing a one-dimensional array of 13 complete EMDs and linearly summing their outputs. This is easily achieved due to the current-mode nature of the EMD output signals; we simply tie all the wires together.

4.2 Voltage-Mode Design

Our voltage-mode version of the Reichardt motion detector offers several advantages over the current-mode design, including superior matching characteristics and reduced contrast dependence. To the best of our knowledge, this is the closest approximation to this biological motion sensor that has been built.

4.2.1 Circuit Architecture

As in the current-mode design, we measure light intensity with an adaptive photoreceptor circuit developed by Delbrück and Mead (Delbrück and Mead, 1996). This four-transistor circuit uses a substrate photodiode and source follower (M_1) to convert incident light into a logarithmically encoded voltage (see Figure 4.2a). A high gain amplifier (M_2 and M_3) and feedback network (C_1 and C_2) amplify the voltage signal by a factor of 18. The adaptive element (M_4) acts as a nonlinear feedback element that conducts only if the voltage across it exceeds several hundred millivolts. This allows the photoreceptor to adapt to large changes in illumination. Thus we maintain a large dynamic range over a wide operating range. At low bias current levels, the bandwidth of the photoreceptor is limited by the parasitic output capacitance C_p. For a detailed discussion of this circuit, see (Delbrück and Mead, 1996).

The adaptive photoreceptor signal is sent to a g_mC highpass filter (see Figure 4.2a). We use a source follower to provide a low-impedance driver, but in future designs we will leave this out and compensate for the increased output capacitance by increasing the photoreceptor bias current I_{pr}. We use a highpass filter for two reasons. First, the ac coupling eliminates any systematic offsets caused by device variation in the adaptive photoreceptor. Second, by fixing the dc component of the signal to V_a, we can eliminate any common-mode effects later in the circuit.

The delay is accomplished with a first-order g_mC lowpass filter (see Figure 4.2b). The bias transistor in the circuit was made several times minimum size to improve time constant matching across the chip. By operating this circuit at low current levels, we can achieve time constants useful for motion detection (10–100 ms) with reasonably sized capacitors (on the order of 1 pF).

Correlation is approximated by a Gilbert multiplier (see Figure 4.2c). The input V_2 comes from the lowpass filter, and V_1 comes from the highpass filtered photoreceptor from an adjacent pixel (see Figure 3.3b). The voltage V_a is the reference voltage used by the highpass filter, and V_b is another dc bias voltage set a few tens of millivolts below V_a. We operate these field-effect transistors (FETs) in subthreshold, where their drain current I_d, ignoring channel-length modulation effects, is given by

$$I_d = I_0 e^{\frac{\kappa V_g}{U_T}} \left(e^{-\frac{V_s}{U_T}} - e^{-\frac{V_d}{U_T}} \right) \tag{4.1}$$

where I_0 is a process-dependent constant, V_g, V_s and V_d are the gate, source, and drain voltages

referenced to the bulk potential, κ is the gate efficiency factor (typically around 0.7), and U_T is the thermal voltage kT/q (Mead, 1989). Subthreshold FETs exhibit exponential behavior, much like the BJTs with which the Gilbert multiplier was originally built. We take a single-ended current-mode output from the circuit, which gives us

$$I_{out} = \frac{I_b}{2} + \frac{I_b}{2} \tanh\left[\frac{\kappa(V_1 - V_a)}{2U_T}\right] \tanh\left[\frac{\kappa(V_2 - V_b)}{2U_T}\right] \tag{4.2}$$

where I_b is the bias current. For small-signal inputs, this can be approximated as

$$I_{out} = \frac{I_b}{2} + \frac{\kappa^2 I_b}{8{U_T}^2}(V_1 - V_a)(V_2 - V_b) \tag{4.3}$$

An older version of this circuit used a pFET mirror to eliminate the dc component $I_b/2$ and double the signal amplitude. Device mismatch in the mirror introduced additional offsets, so it was not used in this design. We could in principle use both single-ended outputs, but that would require more wires routed through each pixel, which would consume more area.

For the multiplier to work properly, the common-mode voltage of the lower inputs (V_2 and V_b) must be lower than the common-mode voltage of the upper inputs (V_1 and V_a). Simulation results show that acceptable behavior is obtained with a difference of only 50 mV. In order to lower the dc level of the lowpass filter output, we lowered the source voltage of the output FET in the current mirror of the g_mC filter (see Figure 4.2b). By placing the V_{tilt} bias a few tens of millivolts below V_{dd}, we lower the dc output level by $(V_{dd} - V_{tilt})/\kappa$. This source voltage "tilt" increases the time constant of the lowpass filter, but we can compensate by raising I_τ. The difference in source voltages also creates an asymmetry in the up-going and down-going slew rates of the filter, but in practice this does not seem to have a significant effect on the overall circuit performance.

It can be shown from Equation 4.2 that the circuit output saturates for differential inputs greater than about $4U_T \approx 100$ mV. Rather than restrict our signals to this small linear region, we exploit the nonlinear behavior of the circuit to improve our motion detection algorithm. It has been shown that by adding saturating nonlinearities before the correlation stage, the contrast dependence of a Reichardt detector can be reduced (Egelhaaf and Borst, 1989). To understand this effect, consider Equation 4.2 in the extreme case where both differential inputs are much greater than U_T:

$$\begin{aligned} I_{out} &\approx \frac{I_b}{2} + \frac{\kappa^2 I_b}{8{U_T}^2} \operatorname{sgn}(V_1 - V_a) \operatorname{sgn}(V_2 - V_b) \\ &\approx \frac{I_b}{2} + \frac{\kappa^2 I_b}{8{U_T}^2} \operatorname{sgn}[(V_1 - V_a)(V_2 - V_b)] \end{aligned} \tag{4.4}$$

where the sign function $\mathrm{sgn}(x)$ is defined as

$$\mathrm{sgn}(x) = \begin{cases} 1, & x \geq 0 \\ -1, & x < 0 \end{cases} \tag{4.5}$$

Incorporating this "saturated multiplier" model into our analysis from Chapter 3, we can rewrite Equations 3.15 and 3.16 as

$$m_1(t) = \mathrm{sgn}\left[\cos(\omega_s \phi + P) - \cos(2\omega_t t - P)\right] \tag{4.6}$$

$$m_2(t) = \mathrm{sgn}\left[\cos(\omega_s \phi - P) - \cos(2\omega_t t - P)\right] \tag{4.7}$$

where P is given by Equation 3.18. We normalize for the constant prefactor in Equation 4.4 and neglect the constant dc component since the opponency subtraction will cancel this current.

We can continue the analysis of an EMD array with nonlinear multipliers by using the time averages of $m_1(t)$ and $m_2(t)$, which are given by

$$\langle m_1(t) \rangle = 1 - \frac{2}{\pi}\left(\omega_s \phi + \tan^{-1} \tau \omega_t\right) \tag{4.8}$$

$$\langle m_2(t) \rangle = 1 - \frac{2}{\pi}\left(\omega_s \phi - \tan^{-1} \tau \omega_t\right) \tag{4.9}$$

When these signals are subtracted in opponency, the time average of the output is given by

$$\begin{aligned} \langle o(t) \rangle &= \frac{4}{\pi} \tan^{-1} \tau \omega_t \\ &= \frac{4}{\pi} \tan^{-1} \tau v \omega_s \end{aligned} \tag{4.10}$$

Spatial integration across a small array of motion detectors will have the effect of integrating over different phases of the stimulus. This will remove the time-dependent components of the motion detector output.

We see in Equation 4.10 that in the limit of full multiplier saturation the output is sensitive to pattern velocity and spatial frequency but is independent of pattern contrast. Compare this to the original Reichardt motion detector response given by Equation 3.20, where the contrast dependence is quadratic. We use the inherent saturation in the Gilbert multiplier to achieve this behavior without adding additional hardware. We shall demonstrate this reduced contrast dependence in the following chapter.

Note that while we call this circuit a voltage-mode design, the final outputs of the EMD are two currents.

Figure 4.3 shows the layout for one 1-D motion sensor, corresponding to the circuit element

outlined in Figure 3.3b. All experimental results shown below were measured from arrays of this circuit, which was fabricated in a 1.2 μm double-poly, double-metal n-well CMOS process, yielding a pixel size of 61 μm $\times$ 199 μm with 32 transistors and a 4 capacitors totaling 3.0 pF. In order to build a 2-D motion sensor, we need add only two more multiplier circuits and additional interpixel and output wiring. Only two wires in each direction are required for nearest-neighbor communication, making 2-D layout practical. An additional interpixel wire may be required if opponent subtraction is performed locally.

While our fill factor is low (3.3%), we argue that fill factor is not an important factor for motion sensors. Defocused optics can be used to eliminate spatial aliasing. If low-light operation is required and catching every photon is essential, microlens technology currently used in CCDs could be used.

All data shown was measured from an analog VLSI chip fabricated in a standard, commercially available 1.2 μm CMOS process. The 2.2 mm $\times$ 2.2 mm chip contained six parallel one-dimensional arrays of 24 EMD opponent pairs each with integrated photoreceptors (see Figure 4.4). 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. 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 constant was set to 50 ms, and the bandpass filters were set to pass frequencies between 0.5 Hz and 8 Hz.

4.2.2 Supply Voltage Requirements

By computing motion in parallel, we do not need time constants less than a millisecond at any pixel. The fastest known visual systems (those of houseflies) have bandwidths of less than 200 Hz, and humans can barely perceive the flicker of a 60 Hz monitor. This low bandwidth requirement allowed us to operate the entire circuit in subthreshold (drain currents typically less than 1 μA). Subthreshold operation allowed us to operate at V_{dd} = 2.5 V despite the Gilbert multiplier circuit, where three transistors are in series between the power supply and ground.

In order to operate an above-threshold transistor in the saturation region (where the drain voltage has no effect on the drain current, neglecting channel-length modulation), the drain-to-source voltage V_{ds} must exceed $V_{gs} - V_t$, where V_t is the threshold voltage. In subthreshold operation, the condition for saturation is $V_{ds} \geq 4U_T$, where U_T is the thermal voltage, and $4U_T$ is about 100 mV at room temperature. This low voltage overhead requirement allows us to operate with a power supply voltage of 2.5 V even though the circuit was fabricated in a 5 V process. The source follower buffer for the photoreceptor is the subcircuit limiting our voltage ceiling. Once this is removed from later designs, operation with a supply voltage of less than 2 V should be possible.

4.2.3 *Matching* Considerations

Device mismatch is inherent in any physical circuit, and is an important consideration when designing analog circuits that will be repeated many times across a chip. We want every motion sensor on a die to exhibit similar performance. The large number of sensors on a single chip precludes off-chip trimming of each circuit to achieve matching. Some floating-gate circuits are capable of storing correction factors locally, but these would add to the size and complexity of each circuit (Harrison et al., 1998).

We use two types of devices in our EMD: transistors and capacitors. Parallel plate capacitors in CMOS, even very small ones, match very well (Minch et al., 1996a). In order to study transistor matching, we fabricated arrays of several hundred nMOS and pMOS transistors, and measured their matching characteristics.

Transistor mismatch is most simply modeled as a voltage source in series with the gate. This gate voltage variation is particularly important in subthreshold operation due to the increased relative transconductance in this region. We modeled this source as a Gaussian distribution of mean zero and standard deviation σ. We estimated σ for nFETs and pFETs of various sizes by measuring I-V characteristics across the transistor arrays.

As expected, σ decreased with the square root of transistor area (data not shown). This behavior has been observed in other studies of subthreshold FETs (Pavasović et al., 1994). We also found that the value of σ for pFETs was 2.7 times as large as the value for nFETs of equal gate area, on average.

We used a simple model for transistor mismatch:

$$\sigma_n(A) = \frac{\sigma_{0n}}{\sqrt{A}} \tag{4.11}$$

$$\sigma_p(A) = \frac{\sigma_{0p}}{\sqrt{A}} \tag{4.12}$$

where A is the gate area, and $\sigma_{0p} = 2.7\sigma_{0n} \approx 31$ mV·μm in our technology (AMI 1.2 μm 2-poly, 2-metal BiCMOS process available through the MOSIS fabrication service).

Using this knowledge, we can apportion our limited layout area in a way that maximizes interpixel matching. We analyzed subcircuits in our motion sensor to determine what effect variation in an individual device would have on the entire subcircuit. For example, analysis of the five-transistor lowpass filter (see Figure 4.2b) reveals that mismatch in each of four transistors—the two nFETs in the differential pair (M_2 and M_3) and the two pFETs in the current mirror (M_4 and M_5)—contribute equally to the intercircuit variance of the output voltage:

$$\sigma_{out}{}^2 = \sigma_{M_2}{}^2 + \sigma_{M_3}{}^2 + \sigma_{M_4}{}^2 + \sigma_{M_5}{}^2 \tag{4.13}$$

Given a limited area for circuit layout, it follows that to minimize mismatch, we must apportion the chip area as follows:

$$\frac{A_P}{A_N} = \sqrt{\frac{N_P}{N_N}}\frac{\sigma_p}{\sigma_n} \tag{4.14}$$

where A_P and A_N are the layout areas devoted to pFETs and nFETs, N_P and N_N are the number of pFETs and nFETs that contribute equally to the total circuit variance, and $\sigma_p = 2.7\sigma_n$ in our technology. We used this consideration—allocating more area for pFETs due to their worse matching properties—when designing the layout. We also tried to reduce the total number of pFETs in the circuit (e.g., removing the pFET current mirror from the Gilbert multipliers).

Also, we drew transistors at least twice minimum width and three times minimum length to facilitate matching and reduce channel-length modulation effects.

The type of circuits used can also effect interpixel matching. In the current-mode design described in Section 4.1, we used current-mode, translinear filters and multipliers to build a Reichardt motion sensor. While individual circuits could be made to perform well, matching across a chip was very poor. It was discovered that translinear circuits were prone to *gain* errors due to transistor mismatch, while traditional voltage-mode circuits were prone to *dc offset* errors. Gain errors before a subtraction operation (such as the opponency subtraction) result in large errors, while offset errors tend to cancel.

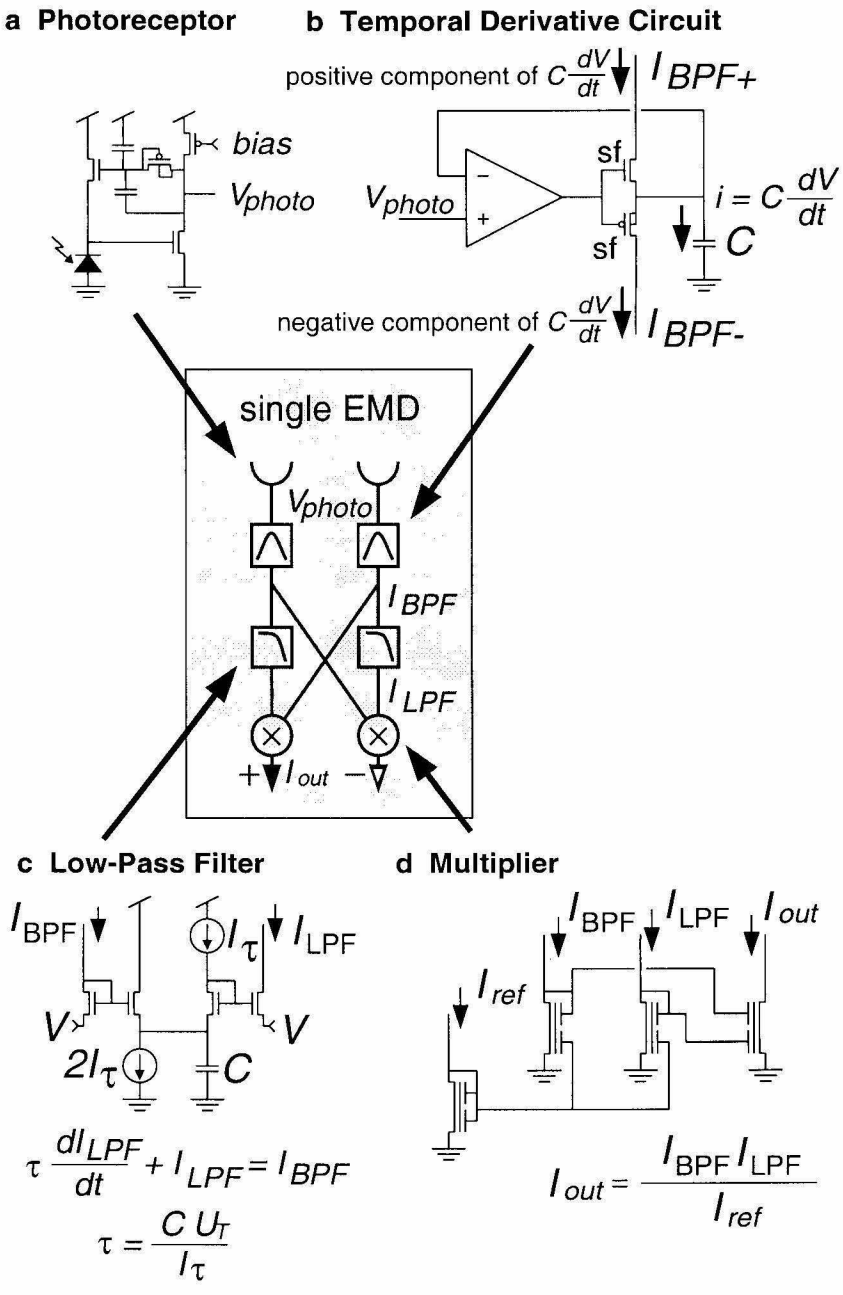

Figure 4.1: Current-mode EMD subcircuits. (a) Photoreceptor. This circuit produces an output voltage proportional to the logarithm of light intensity. (b) Temporal derivative circuit. In combination with the lowpass filter inherent to the photoreceptor, this forms a temporal bandpass filter with a current-mode output. (c) Lowpass filter. The time constant of this first-order filter is determined by the bias current I_τ (which is set by a voltage supplied from off-chip) and the capacitance C. (d) Multiplier. The devices shown are floating-gate nFET transistors with capacitive inputs. The two inputs couple to the floating gate, forming a capacitive divider. The input transistors are diode-connected, which converts the input currents into log-encoded voltages. The capacitive divider on the output transistor computes a weighted sum of these voltages. The output transistor produces a current proportional to the exponential of this sum.

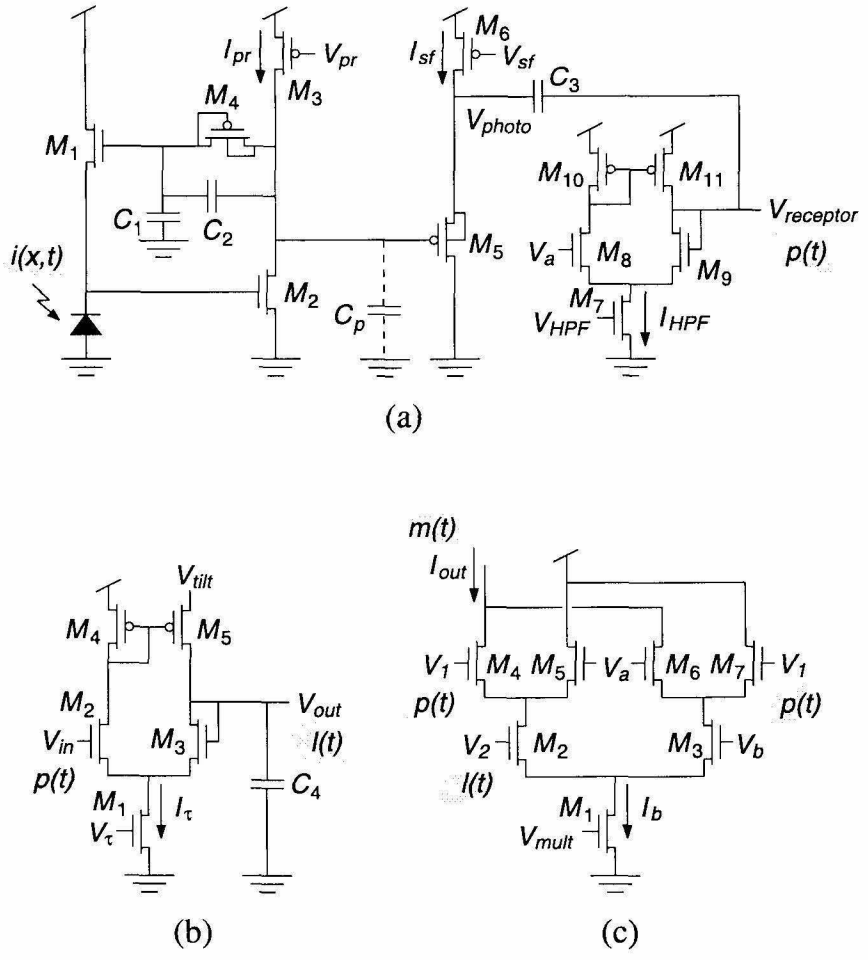

Figure 4.2: Voltage-mode EMD subcircuits. Shaded labels indicate corresponding signals from Figure 3.3a. (a) Adaptive photoreceptor (M_1–M_4, C_1–C_2) with source follower (M_5–M_6) and temporal highpass g_mC filter (M_7–M_{11}, C_3) to remove the dc component of V_{photo}. (b) Temporal lowpass g_mC filter. This circuit's phase lag acts as a delay. (c) Gilbert multiplier. This circuit multiplies delayed and nondelayed photoreceptor signals. The output is a current I_{out}, which allows for easy spatial summation.

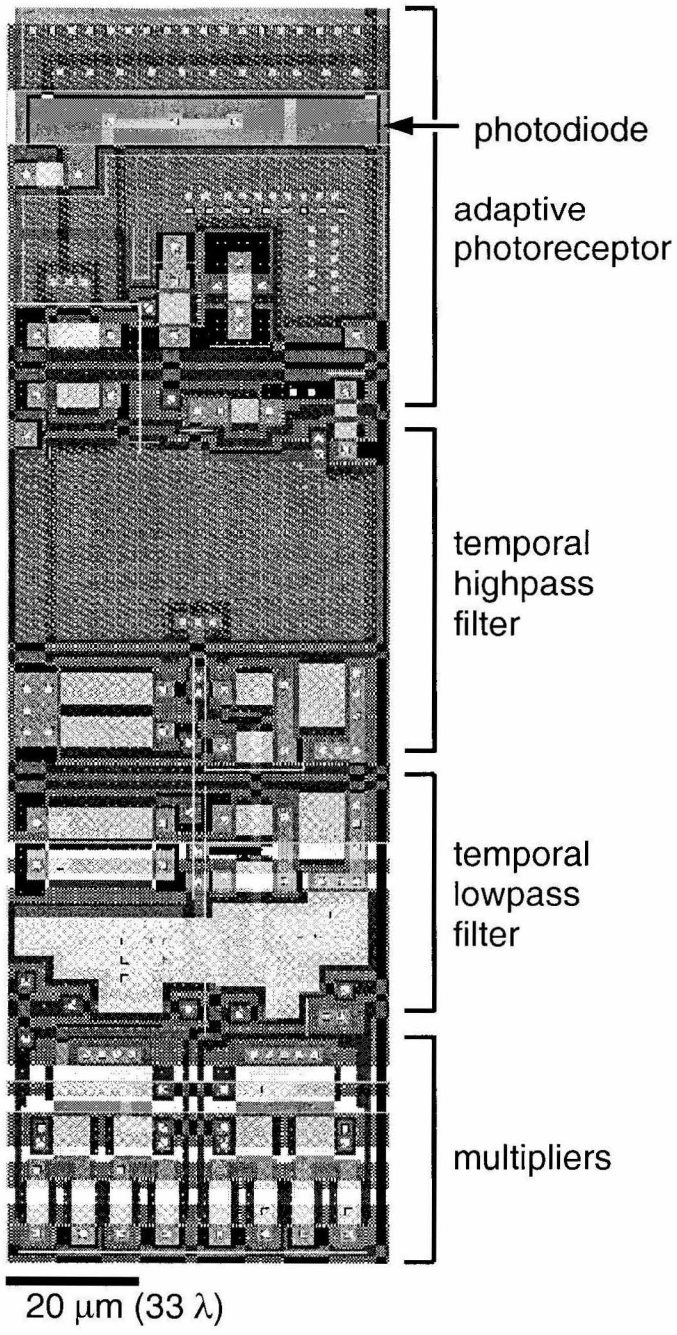

Figure 4.3: Voltage-mode motion detector layout. Cell measures 61 μm $\times$ 199 μm in a standard 1.2 μm process with 32 transistors and 4 capacitors totaling 3.0 pF. In order to build a 2-D motion sensor, we need only add two more multiplier circuits, and some additional interpixel wiring.

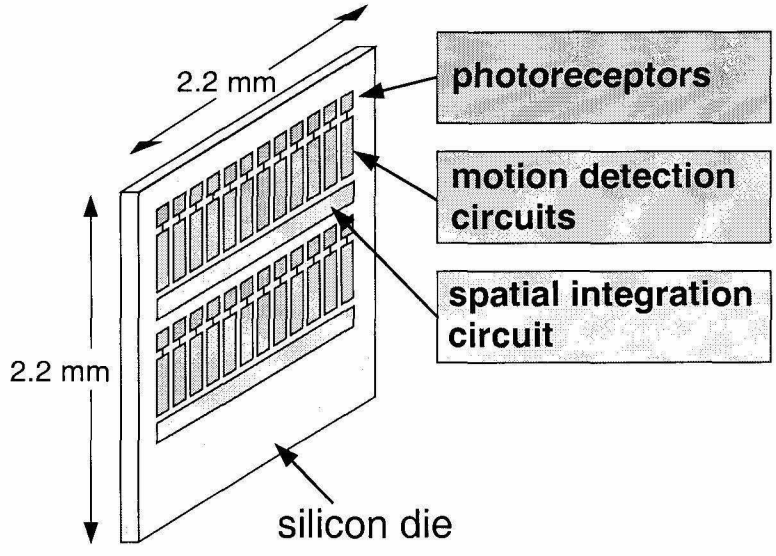

Figure 4.4: Layout of EMD circuits on a chip. This schematic shows how EMDs were arranged in parallel rows covering the chip. The actual chip contained seven rows with 24 EMD circuits in each row. One row was used only for circuit testing, so a 24 $\times$ 6 array of motion detectors was used in the optomotor experiments.

Chapter 5 VLSI Motion Detector Characterization

The voltage-mode elementary motion detectors described in the previous chapter demonstrated superior matching characteristics for similar pixel sizes. In this chapter, we characterize in detail the behavior of the voltage-mode EMD to both simple and complex visual stimuli.

5.1 Methodology

All of the experiments in this chapter were carried out on a 1×22 array of motion sensors fabricated on a 2.2 mm $\times$ 2.2 mm die in a standard 1.2 μm CMOS process. A 2.6 mm focal length lens was mounted directly over the chip, giving a 35° field of view across the entire array. The angle ϕ between adjacent photoreceptors was 1.5°, comparable to the eyes of many flying insects (Land, 1997). The chip was biased to an appropriate operating range, and the bias settings were unchanged during all experiments, except where explicitly stated.

For experiments involving spatial integration over many sensors, the individual output currents were summed on two wires, one for the rightward-facing half-receptors (i.e., the m_1 signal in Figure 3.3a), and one for the leftward-facing half-receptors (i.e., the m_2 signal in Figure 3.3a). The currents were measured with off-chip sense amplifiers. The two opponent signals were subtracted to yield a direction selective response.

We presented computer-generated visual stimuli on a standard monitor (Sony Multiscan 17se II) with a refresh rate of 72 Hz (see Figure 5.1). Our software was able to update the screen at approximately the same rate. The bandwidth of the adaptive photoreceptors was set sufficiently low to attenuate screen refresh artifacts by 20 dB. This also prevented the photoreceptors from responding to the 120 Hz signal in ac incandescent lighting.

We generated visual stimuli with spatial resolution far exceeding the motion sensor array resolution. We used a 64-value gray scale to generate sinusoidal gratings and other complex stimuli of varying contrasts.

5.2 Direction Selectivity

Figure 5.2 shows the output of a single Reichardt motion sensor and the summed output of the 22-element sensor array in response to a sinusoidal grating drifting along the sensor axis. The sensor

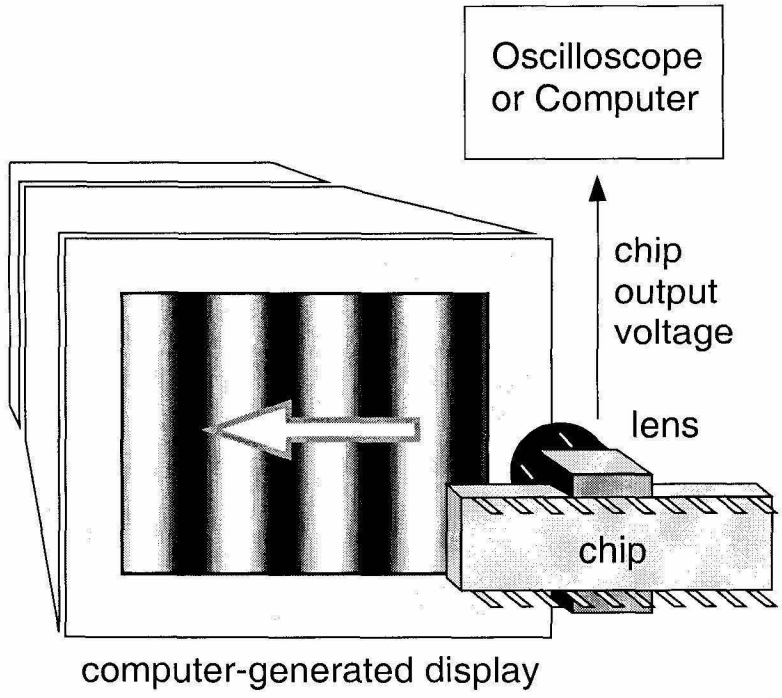

Figure 5.1: Chip testing methodology. We mounted a lens directly over the chip to focus an image on the photoreceptor array. Moving patterns were generated on a standard computer monitor. The temporal bandpass filters in each EMD blocked the 72 Hz refresh rate signal from the monitor.

array is highly direction selective, giving responses of opposite sign to motion in opposite directions. The individual sensor shows a high degree of pattern dependence superimposed on a dc direction selective response. Much of this pattern dependence is caused by device mismatch in the Gilbert multiplier. If the differential pairs are not perfectly matched, the output contains components of the raw input signals.

Pattern dependence is greatly reduced by spatially integrating over a small group of motion sensors that see different phases of the stimulus. Pattern dependence has also been observed in motion-sensitive cells in flying insects, where it is also reduced by spatial integration (Reichardt and Egelhaaf, 1988; Single and Borst, 1998). In principle, pattern dependence could also be removed through temporal integration (averaging over time), but this would limit the response time of the sensor. We chose spatial integration, which sacrifices resolution, but maintains temporal bandwidth. The transients observed at the onset of motion are also observed in biological motion-sensitive cells and have been shown to be a consequence of summing many EMDs which see different phases of a periodic stimulus (Egelhaaf and Borst, 1989).

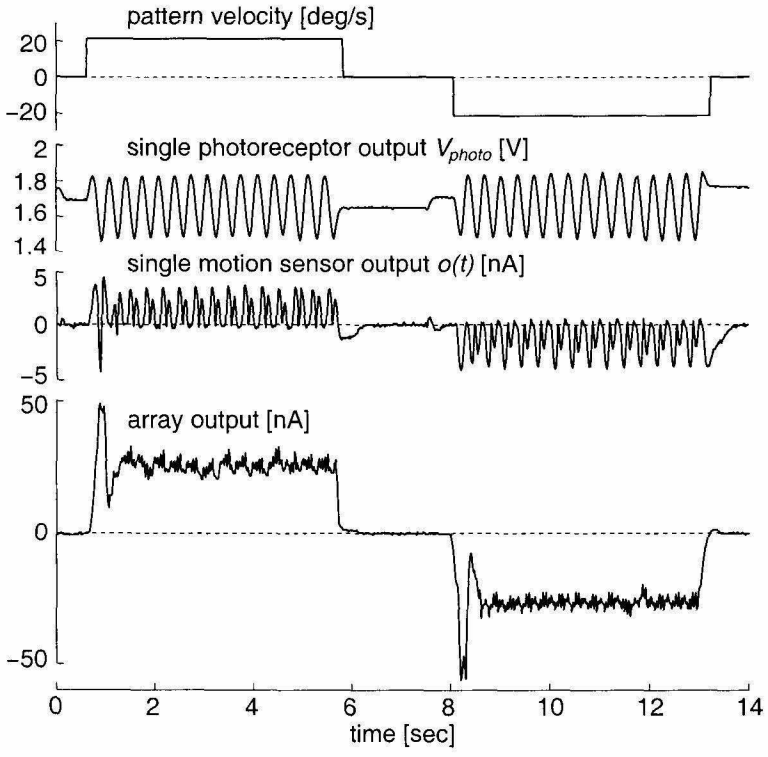

Figure 5.2: Direction selectivity. A sinusoidal grating (f_t = 3.0 Hz, f_s = 0.14 cycles/degree [cpd], contrast = 1) moved along the motion detector axis in alternating directions. Spatial integration over an array of 22 Reichardt detectors eliminates much of the pattern dependence seen in the single sensor trace.

5.3 Response to Simple Stimuli

The summed chip response (see Figure 5.2) was similar to the membrane potential of HS and VS cells, non-spiking 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 motion onset.** At the onset of stimulus motion, a large transient response oscillates with the temporal frequency of the stimulus pattern (see Figure 5.3). 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).

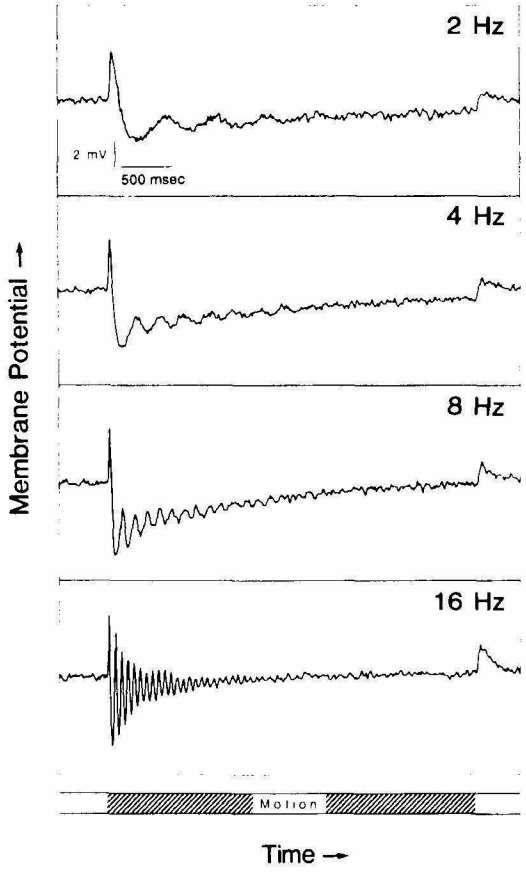

Figure 5.3: Motion onset and offset responses in the fly's nonspiking HS interneuron. At the onset of stimulus motion, large transients oscillate with the temporal frequency of the stimulus pattern (indicated above each plot). The offset of stimulus motion produces no oscillations. (From Egelhaaf and Borst, 1989)

To further investigate the nature of the onset oscillations, we derived the transient response of the Reichardt motion detector at the beginning and end of motion. The analysis in Chapter 3 considered only the steady-state case. Following is the time-domain analysis for transient motion of a sinusoidal grating. Transient Reichardt motion detector responses have been derived for a simplified EMD model with no highpass filters in the photoreceptors to adapt away dc illumination (Egelhaaf and Borst, 1989). The addition of the highpass time constant greatly complicates the analysis, but significantly affects the response.

5.3.1 Motion Onset

Consider the case where a stationary grating begins moving at a constant velocity. The grating is stationary until time $t = 0$, at which time it begins moving with velocity v_o. Its velocity profile $v(t)$

is given by

$$v(t) = \begin{cases} 0, & t < 0 \\ v_o, & t \geq 0 \end{cases} \tag{5.1}$$

So the image brightness $i(x,t)$ can be described as

$$\begin{aligned} i(x,t) &= I + \Delta I \sin\left[2\pi f_s \left[x + v_o u(t) t\right]\right] \\ &= I + \Delta I \left[\sin(\omega_t t + \omega_s x) u(t) + \sin(\omega_s x) u(-t)\right] \end{aligned} \tag{5.2}$$

where $u(t)$ is the unit step function, $\omega_t = 2\pi f_t = 2\pi v_o f_s$, $\omega_s = 2\pi f_s$, I is the mean image intensity and the contrast of the grating is given by $\Delta I / I$.

Following the signal processing in the Reichardt motion detector shown in Figure 3.3, we pass the intensity signal through a highpass filter (see Equation 3.12). For our analysis here, we will neglect the high-frequency pole at τ_{photo}. We account for the highpass response first by taking the temporal derivative, and then multiplying by τ_H:

$$\begin{aligned} i_h(x,t) &= \tau_H \frac{\partial}{\partial t} i(x,t) \\ &= \tau_H \Delta I \left[\cos(\omega_t t + \omega_s x) u(t) + \sin(\omega_t t + \omega_s x)\delta(t) - \sin(\omega_s x)\delta(t)\right] \\ &= \Delta I \tau_H \omega_t \cos(\omega_t t + \omega_s x) u(t) \end{aligned} \tag{5.3}$$

where $\delta(t)$ is the unit impulse function.

These signals are now passed through a lowpass filter with time constant τ_H to yield the highpass-filtered photoreceptor signals $p(x,t)$ shown in Figure 3.3:

$$\begin{aligned} p(x,t) &= \int_{-\infty}^{+\infty} i_h(x,\eta) h(t-\eta) d\eta \\ &= \int_0^t \Delta I \tau_H \omega_t \cos(\omega_t \eta + \omega_s x) \frac{1}{\tau_H} e^{-(t-\eta)/\tau_H} d\eta \end{aligned} \tag{5.4}$$

where η is the variable of integration, and the impulse response of the filter is given by

$$h(t) = \frac{1}{\tau_H} e^{-t/\tau_H} u(t) \tag{5.5}$$

$$p(x,t) = \Delta I \frac{\tau_H \omega_t}{\sqrt{\tau_H^2 \omega_t^2 + 1}} \left[\cos(\omega_t t + \omega_s x - \tan^{-1} \tau_H \omega_t) - \cos(\omega_s x - \tan^{-1} \tau_H \omega_t) e^{-t/\tau_H}\right] u(t) \tag{5.6}$$

$$p(x,t) = \Delta I K_H \left[\cos\left(\omega_t t + \omega_s x + \theta_H\right) - \cos\left(\omega_s x + \theta_H\right) e^{-t/\tau_H}\right] u(t) \tag{5.7}$$

where

$$K_H = \frac{\tau_H \omega_t}{\sqrt{\tau_H^2 \omega_t^2 + 1}} \tag{5.8}$$

$$\theta_H = -\tan^{-1}\tau_H\omega_t \tag{5.9}$$

If we pass the photoreceptor signals $p(x,t)$ through a first-order lowpass filter with time constant τ (the delay in our delay-and-correlate architecture), the response is given by

$$l(x,t) = \int_{-\infty}^{+\infty} p(x,\eta)h(t-\eta)d\eta \tag{5.10}$$

where η is the variable of integration, and the impulse response of the filter is given by

$$h(t) = \frac{1}{\tau}e^{-t/\tau}u(t) \tag{5.11}$$

$$\begin{aligned} l(x,t) \;=\; & \Delta I K_H \left[\frac{1}{\sqrt{\tau^2\omega_t^2+1}} \cos\left(\omega_t t + \omega_s x + \theta_H + \theta\right) + \right. \\ & -\frac{\cos\left(\omega_s x + \theta_H\right)}{1-\tau/\tau_H} e^{-t/\tau_H} + \\ & \left. \left(\frac{\cos\left(\omega_s x + \theta_H\right)}{1-\tau/\tau_H} - \frac{\cos\left(\omega_s x + \theta_H + \theta\right)}{\sqrt{\tau^2\omega_t^2+1}} \right) e^{-t/\tau} \right] u(t) \end{aligned} \tag{5.12}$$

where

$$\theta = -\tan^{-1}\tau\omega_t \tag{5.13}$$

For a particular EMD with interommatidial spacing ϕ, the left and right photoreceptor channels $p_L(x,t)$ and $p_R(x,t)$ can be described as

$$p_L(x,t) = \Delta I K_H \left[\cos\left[\omega_t t + \omega_s\left(x - \phi/2\right) + \theta_H\right] - \cos\left[\omega_s\left(x-\phi/2\right) + \theta_H\right] e^{-t/\tau_H} \right] u(t) \tag{5.14}$$

$$p_R(x,t) = \Delta I K_H \left[\cos\left[\omega_t t + \omega_s\left(x + \phi/2\right) + \theta_H\right] - \cos\left[\omega_s\left(x+\phi/2\right) + \theta_H\right] e^{-t/\tau_H} \right] u(t) \tag{5.15}$$

After passing through the EMD lowpass filters, the responses $l_L(x,t)$ and $l_R(x,t)$ are given by

$$\begin{aligned} l_L(x,t) \;=\; & \Delta I K_H \left[\frac{1}{\sqrt{\tau^2\omega_t^2+1}} \cos\left[\omega_t t + \omega_s\left(x-\phi/2\right) + \theta_H + \theta\right] + \right. \\ & -\frac{\cos\left[\omega_s\left(x-\phi/2\right) + \theta_H\right]}{1-\tau/\tau_H} e^{-t/\tau_H} + \\ & \left. \left(\frac{\cos\left[\omega_s\left(x-\phi/2\right) + \theta_H\right]}{1-\tau/\tau_H} - \frac{\cos\left[\omega_s\left(x-\phi/2\right) + \theta_H + \theta\right]}{\sqrt{\tau^2\omega_t^2+1}} \right) e^{-t/\tau} \right] u(t) \end{aligned} \tag{5.16}$$

$$l_R(x,t) \;=\; \Delta I K_H \left[\frac{1}{\sqrt{\tau^2\omega_t^2+1}} \cos\left[\omega_t t + \omega_s\left(x+\phi/2\right) + \theta_H + \theta\right] + \right.$$

$$-\frac{\cos\left[\omega_s\left(x+\phi/2\right)+\theta_H\right]}{1-\tau/\tau_H}e^{-t/\tau_H}+$$
$$\left(\frac{\cos\left[\omega_s\left(x+\phi/2\right)+\theta_H\right]}{1-\tau/\tau_H}-\frac{\cos\left[\omega_s\left(x+\phi/2\right)+\theta_H+\theta\right]}{\sqrt{\tau^2\omega_t^2+1}}\right)e^{-t/\tau}\Bigg]u(t) \tag{5.17}$$

We now perform the multiplication characteristic of Reichardt motion detectors, and subtract the left and right channels $m_L(x,t)$ and $m_R(x,t)$ to obtain the opponent response $o(x,t)$:

$$m_L(x,t) = p_R(x,t)l_L(x,t) \tag{5.18}$$

$$m_R(x,t) = p_L(x,t)l_R(x,t) \tag{5.19}$$

$$o(x,t) = m_R(x,t) - m_L(x,t) \tag{5.20}$$

The analytical form of these signals is lengthy, and is omitted here. A simpler solution emerges after spatially integrating across the spatial phase of the pattern. Integrating across one complete spatial wavelength $\lambda = 2\pi/\omega_s$ in the continuum limit, we get the EMD array output $a(t)$:

$$\begin{aligned} a(t) &= \frac{1}{\lambda}\int_0^{\lambda} o(x,t)dx \\ &= \left[(\Delta I)^2\left|H(\omega_t)\right|^2\frac{\tau\omega_t}{\tau^2\omega_t^2+1}\sin\omega_s\phi\right]f_{on}(t) \end{aligned} \tag{5.21}$$

where the magnitude of the photoreceptor transfer function $H(\omega_t)$ is given by

$$\left|H(\omega_t)\right| = \frac{\tau_H\omega_t}{\sqrt{\tau_H^2\omega_t^2+1}} \tag{5.22}$$

which is equivalent to Equation 3.12 if we neglect the lowpass roll-off at high temporal frequencies (i.e., $\tau_{photo} = 0$).

We see the transient behavior of the EMD is simply the steady-state response which we derived in Chapter 3 (see Equation 3.19), multiplied by a time-varying function $f_{on}(t)$, where

$$f_{on}(t) = 1 + e^{-t/(\tau_H\|\tau)} - \sqrt{A^2+1}\left[e^{-t/\tau}\cos(\omega_t t + \tan^{-1}A) + e^{-t/\tau_H}\cos(\omega_t t - \tan^{-1}A)\right]u(t) \tag{5.23}$$

and A is given as

$$A = \frac{\tau_H\tau\omega_t^2+1}{(\tau_H-\tau)\omega_t} \tag{5.24}$$

Notice that as $t \to \infty$, $f_{on}(t) \to 1$. This function is plotted against silicon EMD array responses for left and right motion transients in Figures 5.4 and 5.5. We used $\tau_H = 200$ ms, $\tau = 80$ ms, and

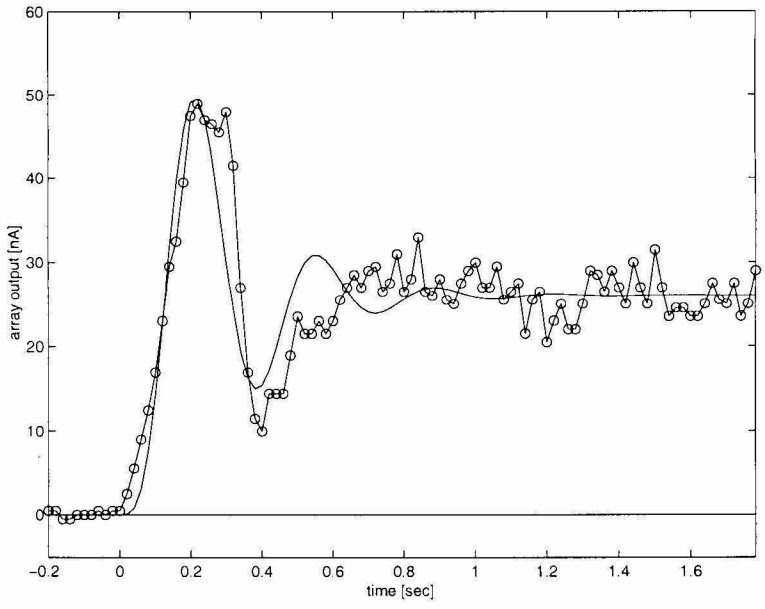

Figure 5.4: Motion onset: Leftward motion. Chip output is plotted against Equation 5.21 for $\tau_H = 200$ ms, $\tau = 80$ ms, and $\omega_t = 2\pi \cdot (3$ Hz). Only the scaling constant was varied to fit the data. Our analytical solution predicts the height of the transient peak accurately.

$\omega_t = 2\pi \cdot (3$ Hz) as parameters for the $f_{on}(t)$, all of which were taken from circuit and stimulus time constants. Only the scaling constant was varied to fit the data in steady state. Notice that the oscillations occur at the frequency of the input stimulus ($\omega_t = \omega_s v_o$). The time constant of the highpass filter τ_H dominates the oscillation decay time. Our analytical solution predicts the height of the transient peak accurately.

5.3.2 Motion Offset

We now consider the motion offset case, where a sinusoidal grating is moving with velocity v_o until time $t = 0$, at which time it stops. Pattern velocity $v(t)$ is given by

$$v(t) = \begin{cases} v_o, & t < 0 \\ 0, & t \geq 0 \end{cases} \tag{5.25}$$

So the image brightness $i(x,t)$ can be described as

$$\begin{aligned} i(x,t) &= I + \Delta I \sin\left[2\pi f_s \left[x + v_o u(-t)t\right]\right] \\ &= I + \Delta I \left[\sin(\omega_t t + \omega_s x)u(-t) + \sin(\omega_s x)u(t)\right] \end{aligned} \tag{5.26}$$

Repeating the steps from the previous section, we account for the highpass photoreceptor response

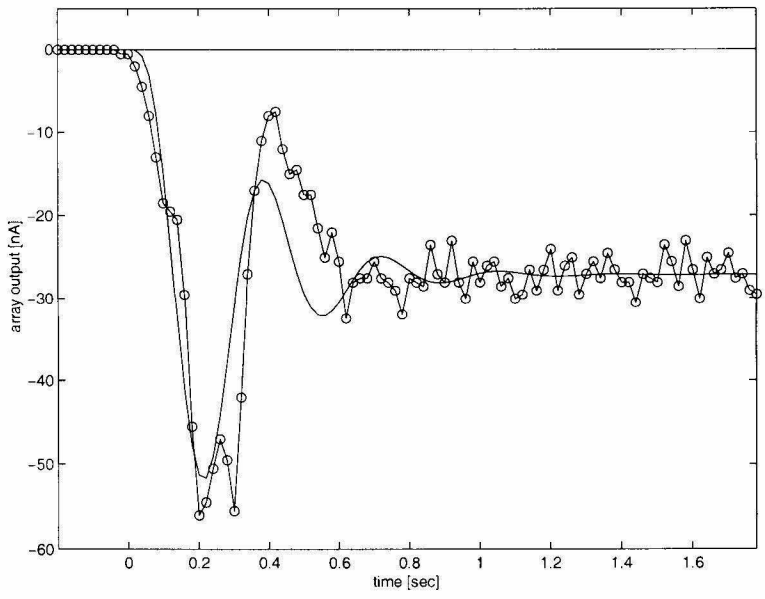

Figure 5.5: Motion onset: Rightward motion. Chip output is plotted against Equation 5.21 for $\tau_H = 200$ ms, $\tau = 80$ ms, and $\omega_t = 2\pi \cdot (3 \text{ Hz})$. Only the scaling constant was varied to fit the data. Our analytical solution predicts the height of the transient peak accurately.

first by taking the temporal derivative, and then multiplying by τ_H:

$$\begin{aligned} i_h(x,t) &= \tau_H \frac{\partial}{\partial t} i(x,t) \\ &= \tau_H \Delta I \left[\cos(\omega_t t + \omega_s x)u(-t) - \sin(\omega_t t + \omega_s x)\delta(t) + \sin(\omega_s x)\delta(t)\right] \\ &= \Delta I \tau_H \omega_t \cos(\omega_t t + \omega_s x)u(-t) \end{aligned} \tag{5.27}$$

where $\delta(t)$ is the unit impulse function.

Now passing this through a lowpass filter with time constant τ_H, we get the highpass-filtered photoreceptor signals $p(x,t)$:

$$p(x,t) = \int_{-\infty}^{+\infty} i_h(x,\eta)h(t-\eta)d\eta \tag{5.28}$$

where η is the variable of integration, and the impulse response of the filter is given by

$$h(t) = \frac{1}{\tau_H} e^{-t/\tau_H} u(t) \tag{5.29}$$

$$p(x,t) = \Delta I \frac{\tau_H \omega_t}{\sqrt{\tau_H^2 \omega_t^2 + 1}} \left[\cos(\omega_t t + \omega_s x - \tan^{-1}\tau_H\omega_t)u(-t) + \cos(\omega_s x - \tan^{-1}\tau_H\omega_t)e^{-t/\tau_H}u(t)\right] \tag{5.30}$$

$$p(x,t) = \Delta I K_H \left[\cos(\omega_t t + \omega_s x + \theta_H)u(-t) + \cos(\omega_s x + \theta_H)e^{-t/\tau_H}u(t)\right] \tag{5.31}$$

$$
\begin{aligned}
l(x,t) &= \Delta I K_H \left[\frac{1}{\sqrt{\tau^2\omega_t^2+1}} \cos\left(\omega_t t + \omega_s x + \theta_H + \theta\right) u(-t) + \right. \\
&\quad \frac{1}{\sqrt{\tau^2\omega_t^2+1}} \cos\left(\omega_s x + \theta_H + \theta\right) e^{-t/\tau} u(t) + \\
&\quad \left. \frac{\tau_H}{\tau_H - \tau} \cos(\omega_s x + \theta_H) \left(e^{-t/\tau} - e^{-t/\tau_H}\right) u(t) \right] \qquad (5.32)
\end{aligned}
$$

Integrating across one complete spatial wavelength $\lambda = 2\pi/\omega_s$ in the continuum limit, we get

$$
\begin{aligned}
a(t) &= \frac{1}{\lambda} \int_0^{\lambda} o(x,t) dx \\
&= \left[(\Delta I)^2 \left|H(\omega_t)\right|^2 \frac{\tau\omega_t}{\tau^2\omega_t^2+1} \sin\omega_s\phi \right] f_{off}(t) \qquad (5.33)
\end{aligned}
$$

We see the the transient behavior of the EMD is simply the steady-state response which we derived in Chapter 3 (see Equation 3.19), multiplied by a time-varying function $f_{off}(t)$, where

$$
f_{off}(t) = u(-t) + e^{-t/(\tau_H \| \tau)} u(t) \qquad (5.34)
$$

Notice that as $t \to \infty$, $f_{off}(t) \to 0$. Unlike f_{on}, the off response contains no oscillations and is independent of the stimulus temporal frequency. This function is plotted against silicon EMD array responses for left and right motion transients in Figures 5.6 and 5.7. We used $\tau_H = 200$ ms and $\tau = 80$ ms as parameters for $f_{off}(t)$, all of which were taken from circuit time constants. Only the scaling constant was varied to fit the data in steady state.

Notice that the offset decay is much quicker than the onset decay, and its time course is independent of the stimulus temporal frequency. In this example, the time constant $\tau_H \| \tau$ is approximately 57 ms. The offset response in fly interneurons is also quick and independent of the stimulus temporal frequency (see Figure 5.3).

5.3.3 Directional Tuning

Although the sensors are configured in a 1-D arrangement, they respond to motion in other directions. Figure 5.8 shows the mean response of the sensors to a sinusoidal grating moving in other directions. Dashed lines indicate negative responses. The sensor exhibits a cosine tuning curve, similar to motion-sensitive cells observed in flies (see Figure 5.9) (van Hateren, 1990; Zanker, 1990). This is a consequence of the spatial-frequency dependence in Equation 3.19. For spatial wavelengths much greater than the photoreceptor spacing ϕ, the $\sin\phi\omega_s$ term can be approximated as $\phi\omega_s$. The temporal frequency of the stimulus moving at an angle θ stays the same as it rotates, but the spatial frequency as seen by two adjacent photoreceptors is proportional to $\cos\theta$.

This type of directional tuning is desirable because it responds strongly to slightly off-axis mo-

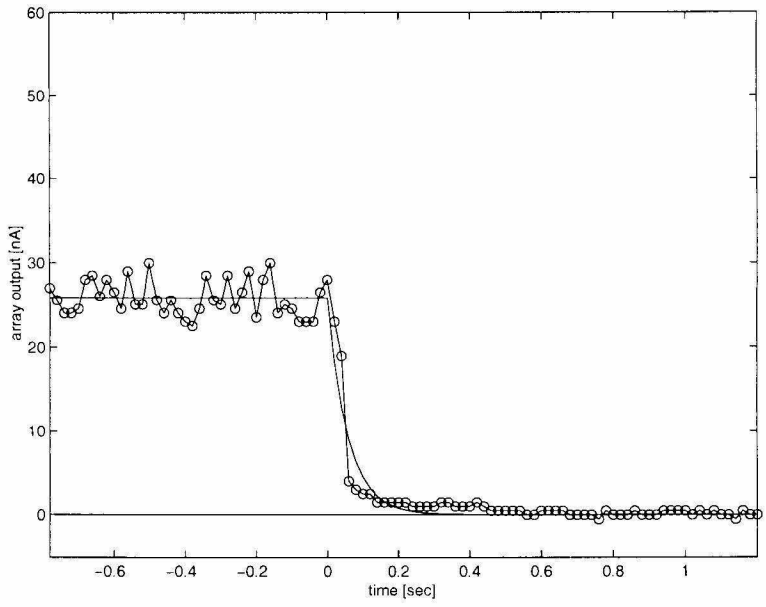

Figure 5.6: Motion offset: Leftward motion. Chip output is plotted against Equation 5.33 for $\tau_H = 200$ ms and $\tau = 80$ ms. Only the scaling constant was varied to fit the data.

tion, but tends to ignore perpendicular motion. As discussed previously, by adding two additional multipliers to each pixel, we can create an additional motion sensor that is oriented perpendicular to the original sensor. Together, these sensors would return signals proportional to the sine and cosine of image motion, so the direction of motion could be determined. Of course, this sensor only measures the normal component of local motion, and does not solve the aperture problem (Marr, 1982).

Our sensor preferentially ignores stimuli in the orthogonal direction, so if one is interested in motion only along one axis where off-axis motion is rare (such as estimating the translation or rotation of a ground vehicle), a simple 1-D sensor array might suffice, since vertical motions would be largely ignored. This is not possible with feature-trackers, which suffer from the obliquity problem discussed in Chapter 3.

5.4 Spatiotemporal Frequency Tuning

Next, we varied the temporal and spatial frequencies of the sinusoidal gratings. Figures 5.10 and 5.11 show the mean response of the sensor array, as well as the standard deviation of the signal over ten temporal cycles of the stimulus. The error bars give an indication of the magnitude of residual pattern dependence. Noise levels were far below the deterministic pattern-dependent fluctuations observed. Theoretical fits are plotted as dashed lines. These fits use Equation 3.19, including parameters from

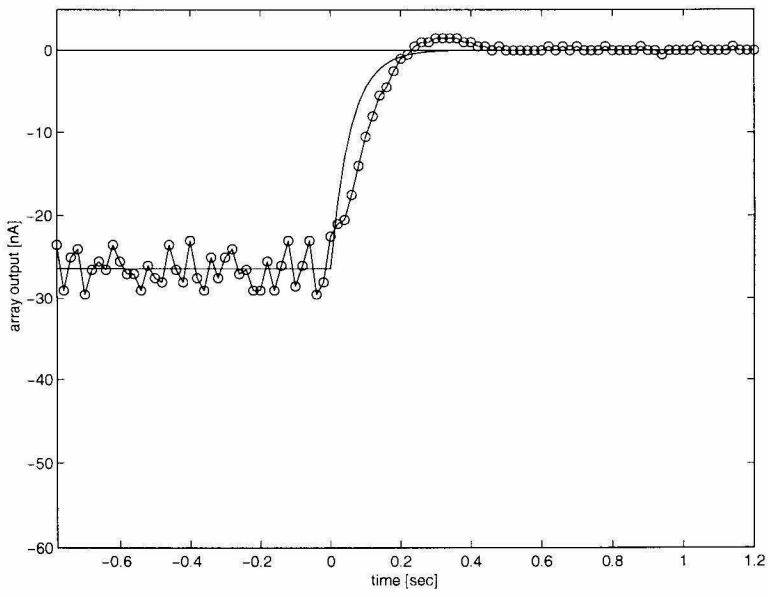

Figure 5.7: Motion offset: Rightward motion. Chip output is plotted against Equation 5.33 for $\tau_H = 200$ ms and $\tau = 80$ ms. Only the scaling constant was varied to fit the data.

Equation 3.12: the first-order temporal lowpass behavior of the adaptive photoreceptor ($\tau_{photo} = 30$ ms) and the first-order temporal highpass filter that follows ($\tau_H = 200$ ms).

The same parameters were used for all fits in Figures 5.10 and 5.11: $\tau = 80$ ms, $\phi = 1.5°$, and a fixed constant of proportionality. The circuit behaves as a Reichardt motion detector over a wide range of spatial and temporal frequencies. Similar temporal frequency tuning is observed in the motion-sensitive lobular plate neurons of flies (see Figure 5.12).

Spatial aliasing should produce response reversals at $N/2\phi$, where $N = 1, 2, 3, \ldots$. Indeed, the first reversal can be seen near $f_s = 1/2\phi \approx 0.33$ cpd. This reversal is also observed in flies, and has been used to measure their interommatidial angle ϕ (Götz, 1965). The effect of aliasing at higher spatial frequencies is reduced by the finite photoreceptor size as well as by slightly defocused optics. Both effects attenuate at high spatial frequencies.

The spatiotemporal frequency tuning of our circuit was measured in detail (see Figure 5.13). Motion-sensitive cells in flying insects exhibit similar sensitivities (see Figure 5.14).

The spatial frequency dependence of the motion sensor is determined by the angular spacing of the photoreceptors. This can be changed only by altering the layout or by changing the focal length of the lens. This would have the effect of shifting the response curve to higher or lower frequencies. Spatial prefiltering of the image could also be introduced to alter the spatial frequency response.

The location of the peak in temporal frequency response is determined by the time constant of the lowpass filter. By changing the bias V_τ (see Figure 4.2b), we can shift the area of maximum

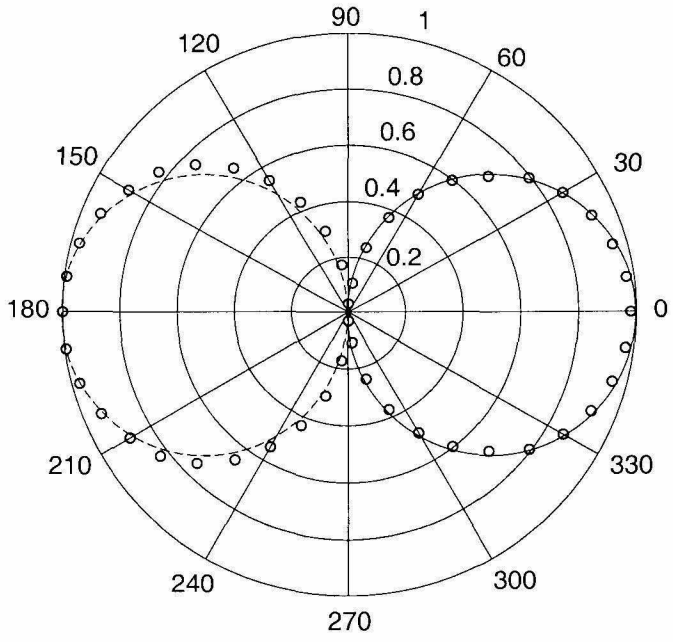

Figure 5.8: Directional tuning of the Reichardt detector array. A sinusoidal grating (f_t = 3.0 Hz, f_s = 0.05 cycles/degree [cpd], contrast = 1) moved at various directions θ relative to the motion detector axis. Magnitude of normalized mean response is shown. Line shows $\cos\theta$; dashed section indicates negative response.

sensitivity (see Figure 5.15). This might allow us, in future versions of the EMD, to introduce temporal adaptation to sustained motion stimuli that shifts the response curve toward dominant stimulus frequencies, a phenomenon observed in biological motion sensors (de Ruyter van Steveninck et al., 1986).

The above experiments were repeated with square wave gratings, and the results were similar (data not shown).

5.5 Contrast Dependence

One of the biggest disadvantages of the Reichardt motion detector is its strong (quadratic) dependence on contrast. This not only confounds contrast with spatiotemporal frequencies in the response, it also greatly amplifies the effect of high-contrast features, while attenuating low-contrast features. Studies of natural scenes have shown that low contrasts are much more common (Ruderman and Bialek, 1994), so we do not want them to be underrepresented by a motion sensor. Motion sensors that respond only to high-contrast edges may work well in the laboratory, but will have sub-optimal performance in the real world.

As suggested in Chapter 4, by exploiting the nonlinear, saturating nature of the Gilbert mul-

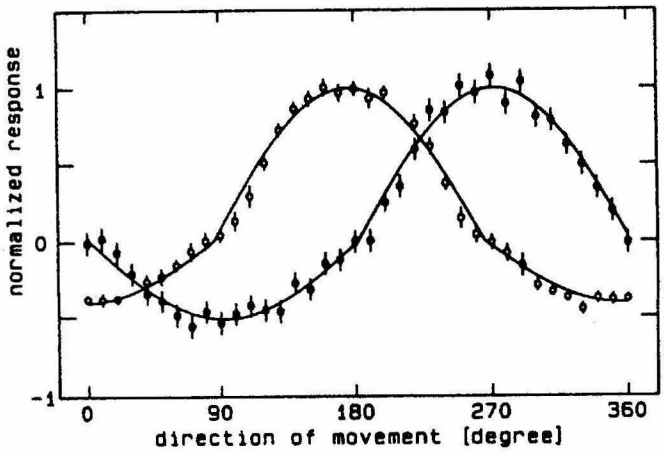

Figure 5.9: Directional tuning of fly lobular place neurons. Data shows the directional tuning curves of H1 (open circles) and V1 (filled circles) in the blowfly. These neurons exhibit sinusoidal directional sensitivity. The stimulus was a drifting square-wave grating. The asymmetry in the preferred and null directions is due to the spiking behavior of H1 and V1 in *Calliphora vicina*. The spontaneous spiking rate (i.e., zero normalized response) is less than half of the maximum spiking rate. (From van Hateren, 1990)

tiplier, we can greatly reduce this contrast dependence. Figure 5.16 shows the motion sensor array's response to drifting sinusoidal gratings of varying contrast. While the contrast dependence is quadratic for low contrasts (dashed line), it saturates at moderate contrast levels as the Gilbert multipliers begin to saturate. Other stimulus information is preserved, as is shown by the saturation to different levels for stimuli with different temporal frequencies. Thus for moderate to high contrasts, our sensor's dependence on contrast is greatly reduced. By increasing the photoreceptor gain, we should be able to further decrease contrast dependence in future designs. The response to low contrasts is weak, but direction information is still encoded.

5.6 Interpixel Variation

As discussed in Chapter 4, good interpixel matching is essential before large arrays of local motion sensors will be feasible. We measured the matching characteristics of our motion sensors across the 22-sensor 1-D array that spanned 1.3 mm. We used a sinusoidal grating of fixed spatial frequency, and measured the mean response of each sensor in the array as we varied its velocity from -200 deg/s to $+200$ deg/s. Figure 5.17 shows the mean and standard deviations of the 22 responses measured across the chip. The ratio of standard deviation to maximum mean response varied between 0.10 and 0.25, depending on the temporal frequency. The individual sensors perform similarly, indicating good matching of gains, dc levels, and time constants.

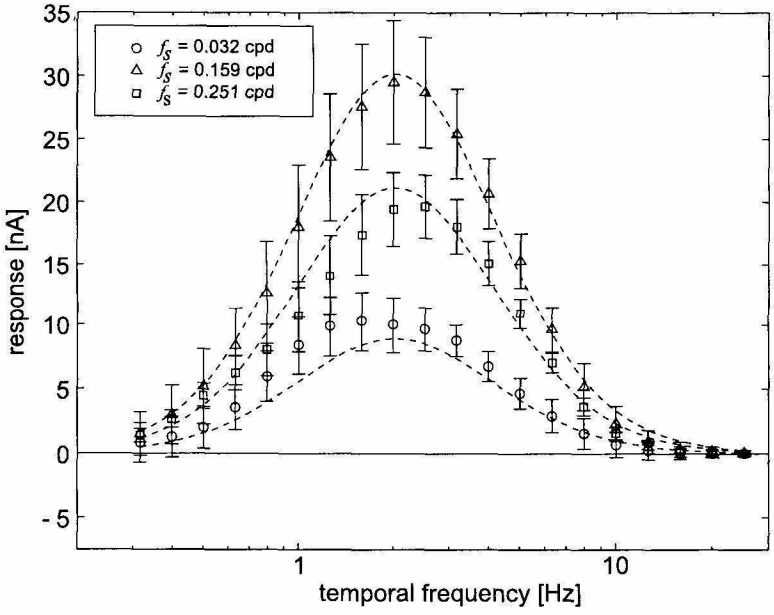

Figure 5.10: Temporal frequency sensitivity of the Reichardt detector array. In this figure and in those following, error bars show one standard deviation of the time response computed over 10 stimulus periods, and represent residual deterministic pattern dependence such as that seen in Figure 5.2. Noise levels were small by comparison. Dashed lines show fits to Equation 3.19.

5.7 Response to Naturalistic Stimuli

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 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

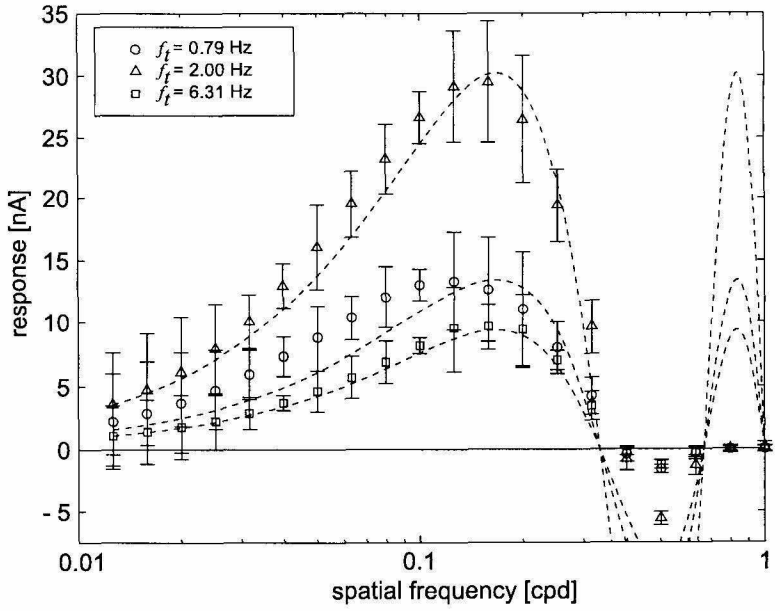

Figure 5.11: Spatial frequency sensitivity of the Reichardt detector array. In this figure and in those following, error bars show one standard deviation of the time response computed over 10 stimulus periods, and represent residual deterministic pattern dependence such as that seen in Figure 5.2. Noise levels were small by comparison. Dashed lines show fits to Equation 3.19. Spatial aliasing is reduced by the finite photoreceptor size as well as by slightly defocused optics.

a Fourier power spectrum R that goes as

$$R(f_s) \sim \frac{1}{f_s^m} \tag{5.35}$$

where f_s is spatial frequency, and m is typically between 2.0 and 2.3 (see Figure 5.18). 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 5.19 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 $1/f$ patterns are significantly more complex than traditional visual stimuli (e.g., bars, dots, sinusoids) and constitute a valuable step towards evaluating the robustness of a vision sensor.

We presented ten different $1/f$ 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 Figure 5.20, the mean and standard deviation over all ten random patterns are plotted versus pattern

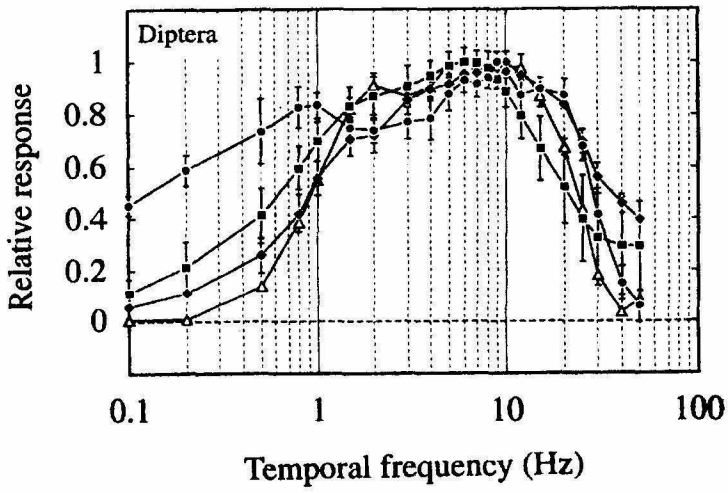

Figure 5.12: Temporal frequency tuning in fly neurons. Drifting sinusoidal gratings of varying temporal frequency were used as stimuli. These traces show the relative response of lobular plate neurons in four dipterans: bee-fly *Bombylius major* (circles), *Eristalis tenax* (squares), hoverfly *Volucella pelluscens* (diamonds), and blowfly *Calliphora erythrocephala* (triangles). Motion sensitive cells in all four species exhibit f_t-tuned responses. (From O'Carroll et al., 1996)

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.

5.8 Robustness to Noisy Stimuli

Most of the VLSI motion sensors reviewed in Chapter 3 were tested with noiseless, highly artificial stimuli—sharp intensity edges and gratings. Some were tested with low-contrast stimuli, but noise was never intentionally introduced into the image. There is shot noise in photoreceptor signals, but this is usually small compared to the signal, and is rarely quantified. [A theoretical noise analysis is shown in (Sarpeshkar et al., 1996), but noise is introduced into experiments only by lowering the overall light intensity.]

To test the robustness of our sensor, we added either spatial or temporal $1/f$ noise to a moving stimulus. We chose not to use white noise because it contains relatively little power in the low frequencies that the motion sensor array can resolve. Figure 5.21a shows a one-dimensional $1/f$ stimulus moving rightward at a constant speed as a function of position and time. The slope of the

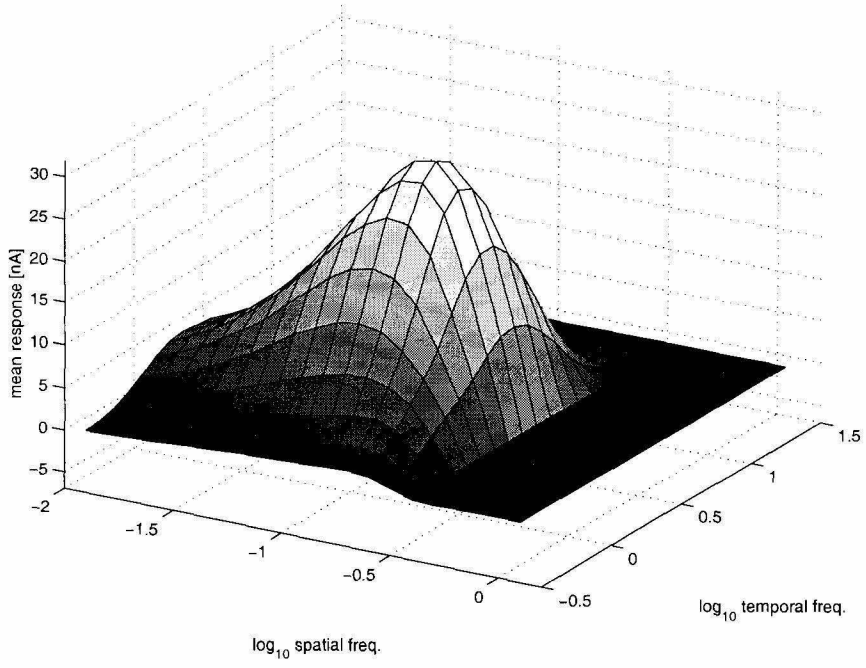

Figure 5.13: Spatiotemporal frequency tuning. By sweeping the temporal and spatial frequencies of the computer-generated stimulus, we measured the spatiotemporal sensitivity of the EMD in detail (contrast = 1).

pattern corresponds to its velocity in this spatio-temporal diagram. We introduced spatial noise by adding a fixed $1/f$ pattern (with different random phase spectrum) to the scene. The signal-to-noise ratio (SNR) of this stimulus is +25 dB, where SNR is defined as

$$SNR_{dB} = 10 \log_{10} \frac{\sigma^2{}_{signal}}{\sigma^2{}_{noise}} \tag{5.36}$$

In Figure 5.21b, we have reduced the SNR to −8 dB. After adding noise to a stimulus, we rescaled the image to cover the entire dynamic range our monitor was able to display, from white to black. This allowed us to cover a greater range of SNRs, but reduced the signal contrast as more noise was added.

Figure 5.21b also shows three "snapshots" of the stimulus at different times. Our sensor array can easily discriminate its direction of motion in this case. It is clear from looking at the snapshots that there are no obvious "features" that one could track from one frame to the next. Although human observers have no problem seeing the direction of motion at this SNR level, features seem to wax and wane in a seemingly random manner as the stimulus moves across the retina.

We also added $1/f$ temporal noise. Figure 5.22a shows the position vs. time graph of the stimulus with SNR = +25 dB. The perception is of motion in one direction, with full-field flicker

or illumination modulation added. Figure 5.22b shows the case where SNR = −8 dB, along with three time traces at fixed locations, such as three photoreceptors might see. There are no obvious temporal features that move from one trace to the next.

We tested the direction selectivity of the sensor array as a function of SNR for spatial noise (see Figure 5.23) and temporal noise (see Figure 5.24). Both figures show the mean and standard deviation of time courses from the sensor array. The increase in mean response can be accounted for by the increasing signal contrast as SNR is raised while dynamic range is kept constant. Even at very low SNR levels, the sensor is still able to discriminate between leftward and rightward motion with high accuracy. Our analog, correlation-based motion algorithm is robust against noisy inputs that would foil other feature-based motion-detection schemes. Our sensor behaves quite robustly in the face of temporal noise, thanks to the opponent subtraction that eliminates common-mode signals. Note that the gradient algorithm would fail here by reporting motion in the direction of the spatial intensity gradient, since the assumption that the image intensity is constant does not hold (see Chapter 3).

5.9 Speed Tuning

A common criticism of the Reichardt motion detector is that it is not a true speed sensor; even with contrast saturation, the output is dependent on the spatial frequency of the stimulus (see Figure 5.25a). Recently, a variation of the Reichardt motion detector has been proposed that greatly reduces the spatial frequency dependence and makes it more speed sensitive (Zanker et al., 1999).

By using half motion detectors with no opponency, we achieve a more speed-tuned response (see Figure 5.25b). However, the lack of opponency reduces directional selectivity (see Figure 5.25c and compare with Figure 5.8). Flying insects seem to use both systems: a highly directional, temporal frequency dependent optomotor system for stabilizing flight and a speed-tuned, non-directional system for navigation (Srinivasan and Zhang, 1993; Srinivasan et al., 1999). It remains unclear whether this second system uses a correlation-type of motion computation.

5.10 Power Dissipation

One entire Reichardt motion sensor consumes 50 nW of power at $V_{dd} = 2.5$ V under normal indoor illumination of 10 cd/m^2. Photocurrent contributes significantly to overall power consumption. Under increased illumination of 2500 cd/m^2, the power consumption increased to 110 nW. The circuit consumes approximately the same amount of power with or without motion present. This is the lowest power requirement of any motion sensor we are aware of. A 100 × 100 array of 2-D

Reichardt *motion sensors* would consume less than 1 mW.

Results described in this chapter also were published in Harrison and Koch, 2000a.

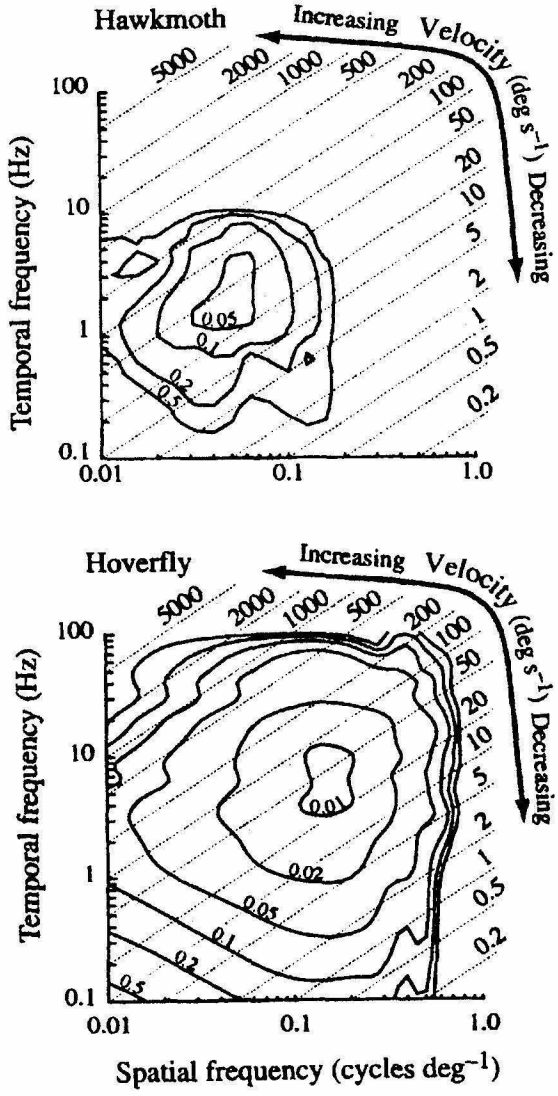

Figure 5.14: Spatiotemporal frequency tuning in fly neurons. Contour plots show the contrast sensitivity of lobular plate neurons from two flying insects. (From O'Carroll et al., 1996)

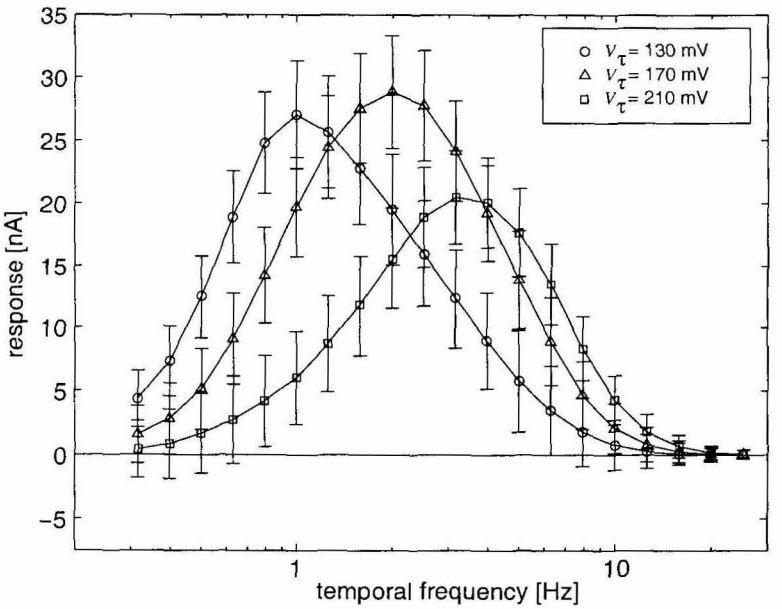

Figure 5.15: Tuning the temporal frequency response. By varying the lowpass filter bias voltage V_τ (see Figure 4.2b), we shift the sensitivity of the motion detectors ($f_s = 0.12$ cpd; contrast = 1).

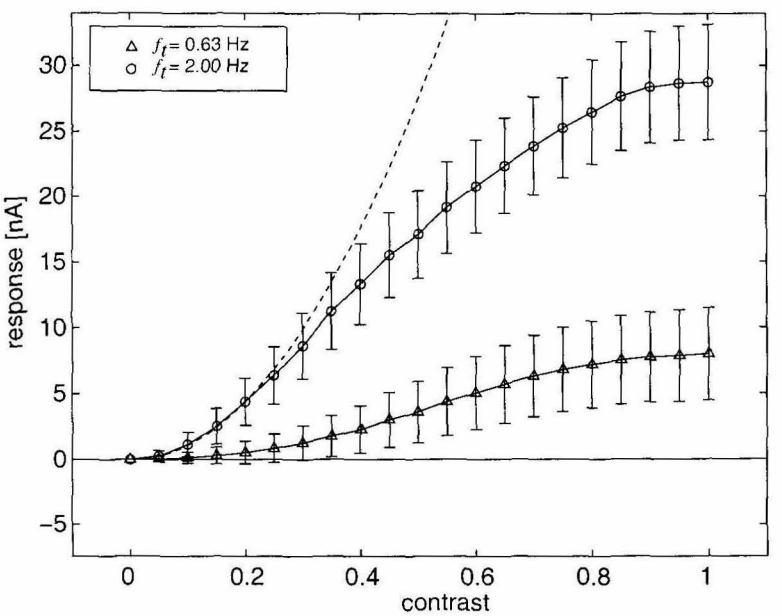

Figure 5.16: Contrast sensitivity. By operating the Gilbert multipliers outside of their linear range, we achieve reduced contrast dependence for moderate to high contrasts for two different velocities ($f_s = 0.12$ cpd). Dashed line shows quadratic dependence, as predicted by the simple Reichardt model in Equation 3.19. Despite saturating behavior, temporal frequency information is still encoded.

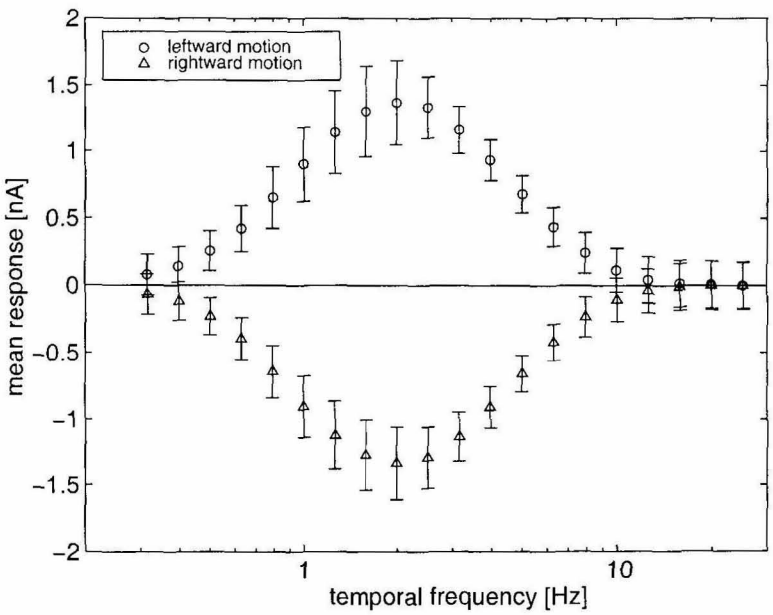

Figure 5.17: Interpixel variation. Here, error bars show standard deviation of mean single-element response across the 22-element array for leftward and rightward motion at different temporal frequencies or, equivalently, speeds ($f_s = 0.12$ cpd; contrast = 1).

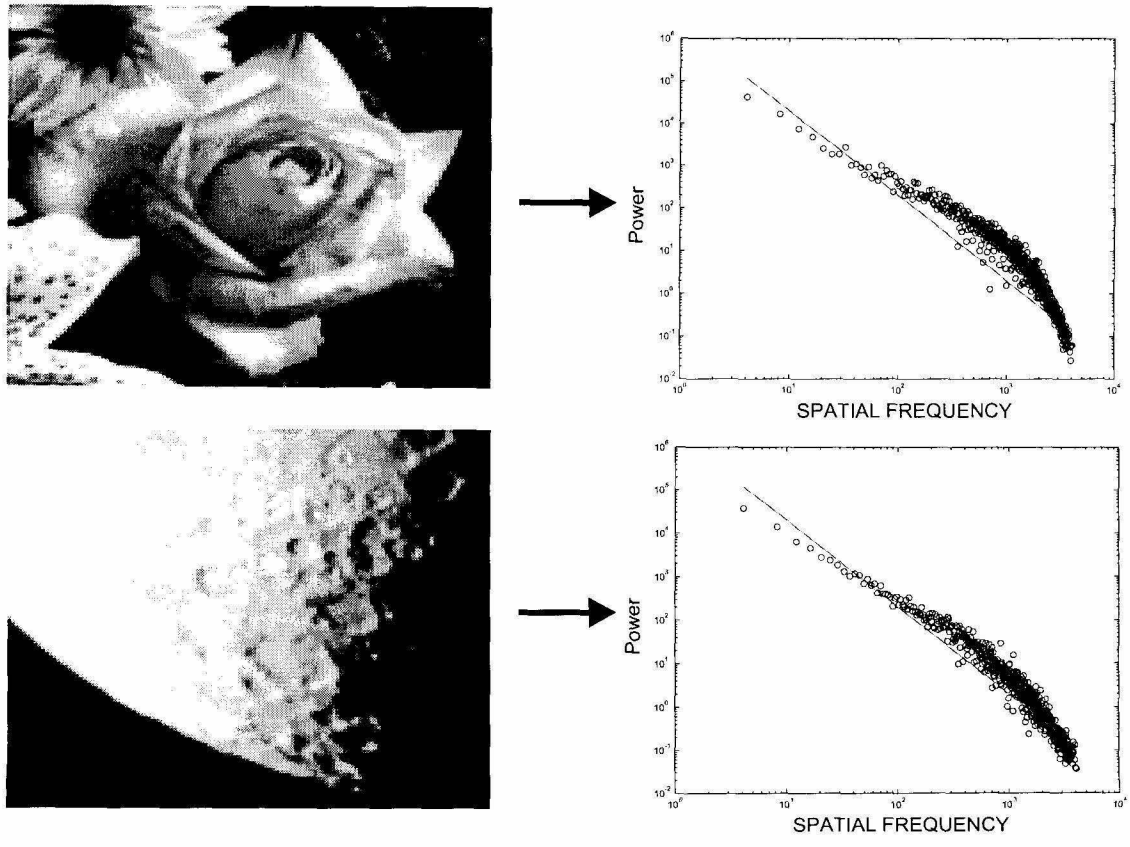

Figure 5.18: $1/f$ Spatial Frequency Spectra in Natural Scenes. Two natural images with corresponding spatial frequency power spectral densities. Straight lines indicate $1/f^2$ slope.

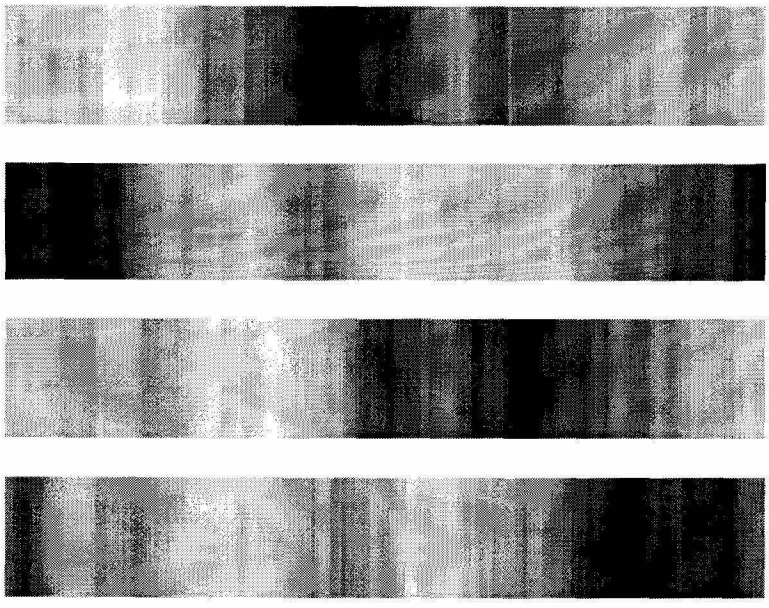

Figure 5.19: Example of random "natural" patterns. These one-dimensional stimuli have a spatial frequency power spectrum $R(f_s) = 1/f_s^{2.3}$ which is observed in natural images. Each pattern has a different, randomly-generated phase spectrum.

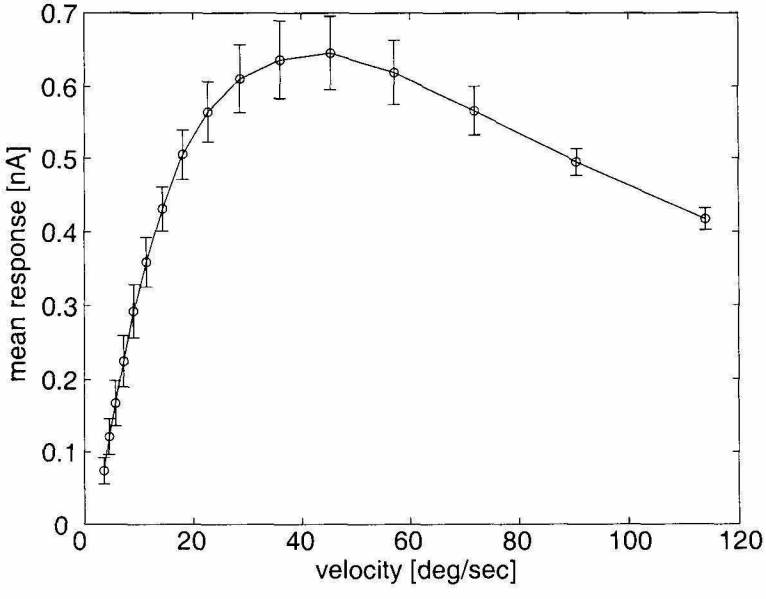

Figure 5.20: Chip response to $1/f$ patterns. Ten random patterns with natural image statistics (see Figure 5.19) were presented to the EMD array. The mean response of the chip to these stimuli is shown as a function of pattern velocity. Error bars denote standard deviation across the mean response from all ten patterns.

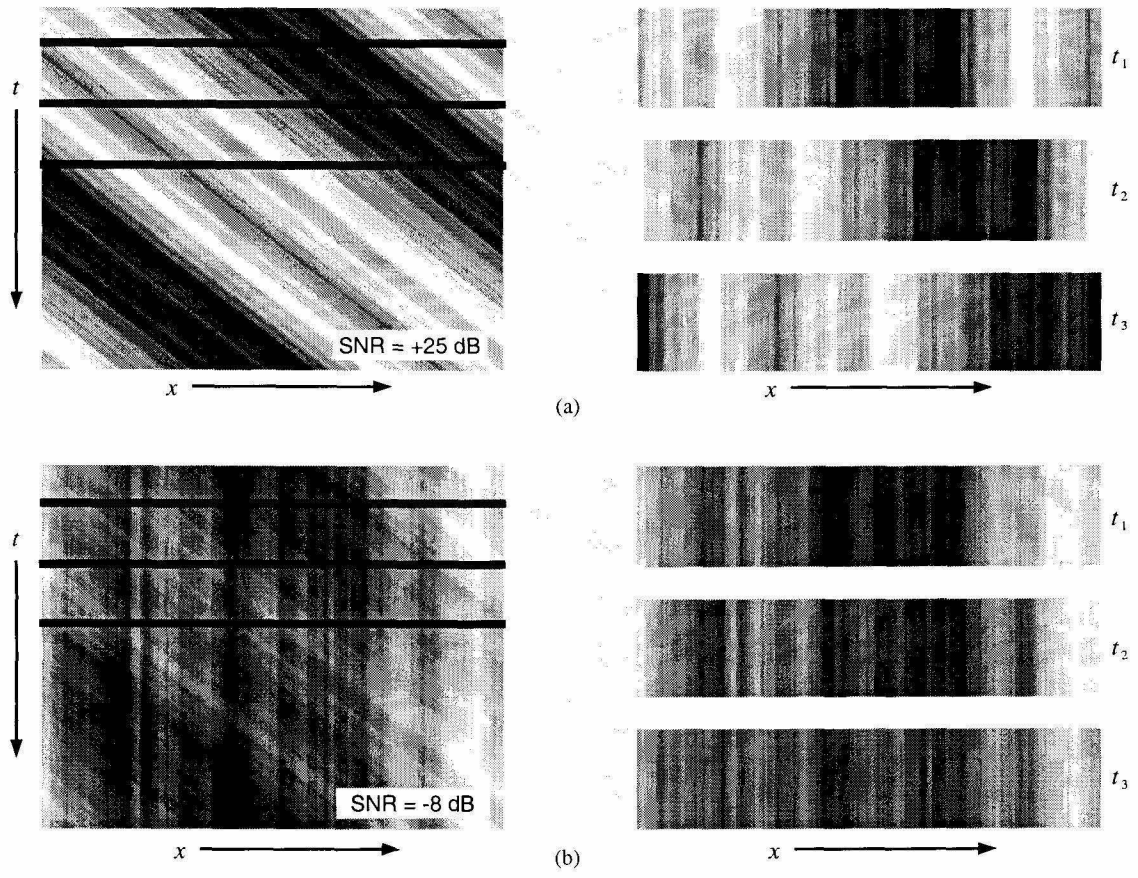

Figure 5.21: Adding spatial noise. Diagrams show a one-dimensional $1/f$ stimulus moving rightward at a constant velocity. In both cases, $1/f$ spatial noise has been added with an effective (a) SNR of +25 dB (b) or of −8 dB. Also shown are three "snapshots" of each image. Notice that under low SNR conditions, few obvious spatial features are apparent that could be reliably tracked to estimate image motion.

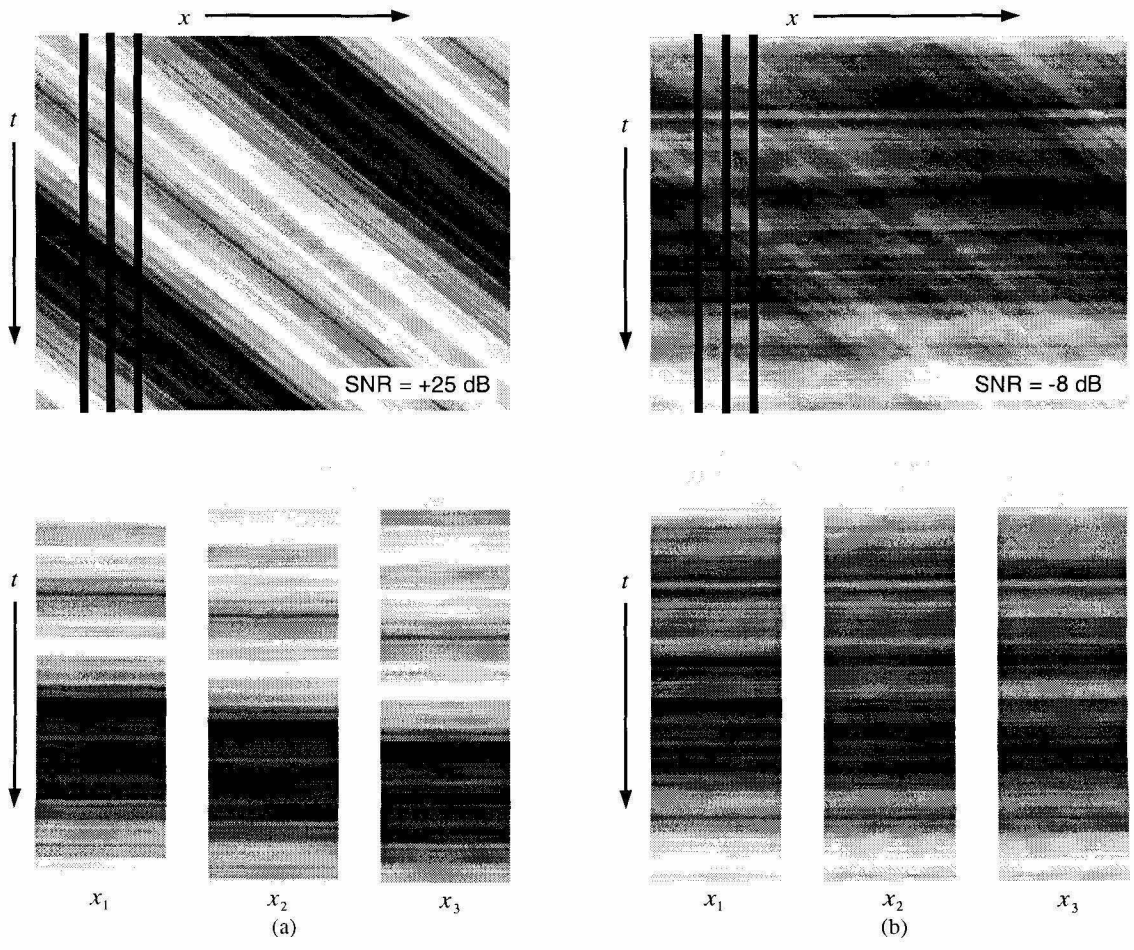

Figure 5.22: Adding temporal noise. Diagrams show a 1-D $1/f$ stimulus moving rightward at a constant velocity. In both cases, $1/f$ temporal noise has been added. (a) SNR = +25 dB. (b) SNR = −8 dB. Also shown are three "photoreceptor signals" from each image. Notice that when the SNR is low, no obvious temporal features exist that could be used to estimate image motion.

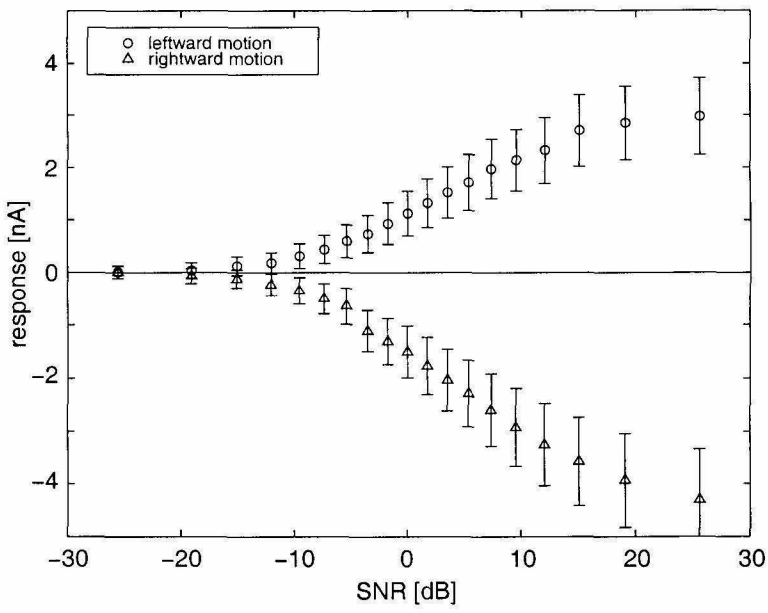

Figure 5.23: Robustness with spatial noise. Performance of Reichardt detector array on direction discrimination with a $1/f$ pattern moving at a constant velocity of 14 deg/s while $1/f$ spatial noise was added. Error bars show one standard deviation of the time response, and represent residual deterministic pattern dependence such as that seen in Figure 5.2.

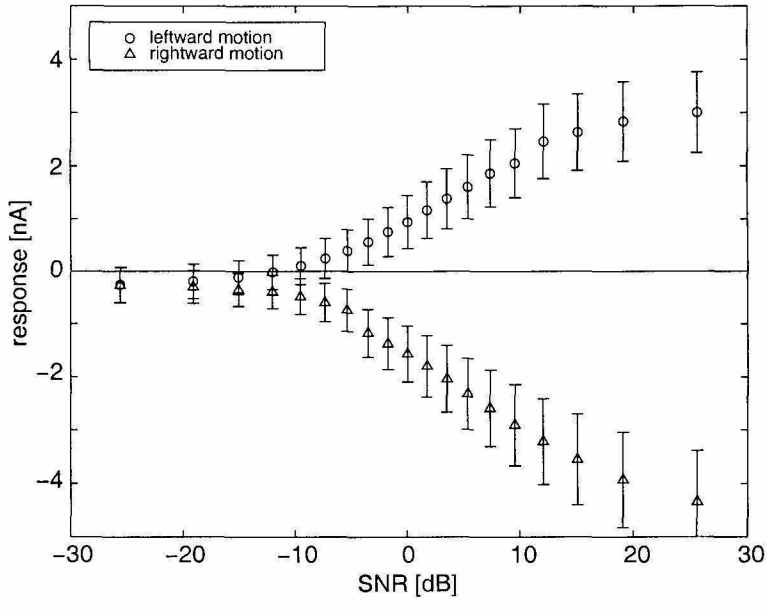

Figure 5.24: Robustness with temporal noise. Performance of Reichardt detector array on direction discrimination with a $1/f$ pattern moving at a constant velocity of 14 deg/s while $1/f$ temporal noise was added. Error bars show one standard deviation of the time response, and represent residual deterministic pattern dependence such as that seen in Figure 5.2.

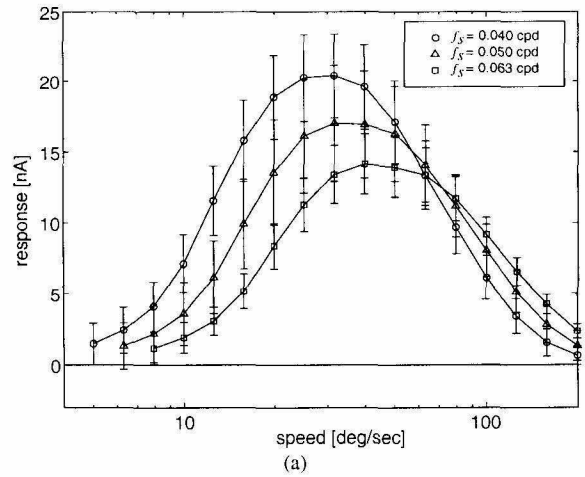

Figure 5.25: Speed tuning. Data from chip is shown. (a) Spatial frequency dependence in the standard opponent Reichardt detector is high. (b) The spatial frequency dependence can be reduced by using only half-detectors with no opponency (see inset). (c) However, direction selectivity is also reduced. (Compare to Figure 5.8).

Chapter 6 Stimulus Reconstruction

In this chapter, we evaluate our sensor's ability to encode information about the velocity of a simple stimulus. We also compare our sensor's encoding ability with the encoding ability of the HS cell in the lobular plate of the blowfly *Calliphora erythrocephala*, which has been previously measured by other researchers. We replicated the experiments of Haag and Borst, 1997, which use *stimulus reconstruction techniques* to measure encoding fidelity. We shall first describe the techniques and then discuss the experimental results.

6.1 Stimulus Reconstruction Techniques

One way of measuring how well a sensor encodes a stimulus is to determine how well we can reconstruct an unknown stimulus from the sensor's response. The stimulus reconstruction technique used in the following experiments finds the linear filter that transforms the sensor response into an estimate of the stimulus that is optimal in the least-squares sense. That is, the reconstruction filter minimizes the square error between the estimate and the actual stimulus. This linear reconstruction represents a lower bound on encoding ability. Nonlinear filters could of course generate an estimate with a lower error, but we compare the sensor's response variability under repeated experiments with identical stimulus conditions to generate an upper bound on encoding ability. A more detailed description of these techniques may be found in Borst and Theunissen, 1999.

Suppose we stimulate a time-invariant system with a set of i Gaussian stimuli $s_i(t)$ and record responses $r_i(t)$. Given the frequency-domain representation of the stimuli $S_i(f)$ and the responses $R_i(f)$, the optimal reverse reconstruction filter is given by the average cross-correlation normalized by the average autocorrelation of the responses:

$$G_{rev}(f) = \frac{\langle R_i^*(f) \cdot S_i(f) \rangle}{\langle R_i(f) \cdot R_i^*(f) \rangle} \tag{6.1}$$

Using this filter, we obtain the estimated stimulus $Sest_i(f)$ from the response:

$$Sest_i(f) = R_i(f) \cdot G_{rev}(f) \tag{6.2}$$

Taking the inverse Fourier transform of this signal gives us the time-domain stimulus estimate $sest_i(t)$.

We can also calculate a filter between the stimulus $S_i(f)$ and the stimulus estimate $Sest_i(f)$.

This filter, denoted $G(f)$, is given by

$$G(f) = \frac{\langle S_i^*(f) \cdot Sest_i(f) \rangle}{\langle S_i(f) \cdot S_i^*(f) \rangle} \tag{6.3}$$

and gives us an idea of how much information about the stimulus is transmitted at specific frequencies. It can be shown that $G(f)$ is equivalent to the coherence function, $\gamma^2(f)$:

$$G(f) = \gamma^2(f) = \frac{\langle S_i^*(f) \cdot R_i(f) \rangle}{\langle S_i(f) \cdot S_i^*(f) \rangle} \cdot \frac{\langle R_i^*(f) \cdot S_i(f) \rangle}{\langle R_i(f) \cdot R_i^*(f) \rangle} \tag{6.4}$$

The coherence function is bounded between zero and one. A coherence value of one indicates that the stimulus is reconstructed perfectly by the reverse reconstruction filter at a certain frequency.

The coherence measure is derived from stimulus reconstruction fidelity, while expected coherence measure is derived from trial-to-trial response variability. (An identical Gaussian stimulus is used for each trial, so any variability in the responses must be due to noise in the sensor, assuming the system in time invariant.) The expected coherence is given by

$$\gamma_{exp}^2(f) = \frac{snr^2(f)}{snr^2(f) + 1} \tag{6.5}$$

where *snr* is the signal-to-noise ratio. The signal is taken to be the average of all i responses, and the noise is calculated from the residual signal after subtracting the mean response from each individual response.

The coherence function is derived from a purely linear reconstruction filter technique and thus represents a *lower bound* on the information encoded by the sensor since nonlinear filters might better extract information about the stimulus from the response. The expected coherence function only takes into account trial-to-trial response variability and thus represents an *upper bound* on the information encoded by the sensor. (A broken sensor might have a constant zero-level output and achieve an expected coherence of one for all frequencies.)

6.2 Pattern Velocity Estimation by Fly Interneurons

The methods above have been applied to several biological sensory systems. Here we describe an experiment performed by Haag and Borst (Haag and Borst, 1997). A square-wave intensity grating displayed on an oscilloscope was presented to the eye of a female blowfly (*Calliphora erythrocephala*). During the experiment, the grating moved with a random velocity (with a screen update rate of 200 Hz). The velocity profile had a flat ("white") spectrum up to approximately 20 Hz.

During stimulus presentation, an HS cell was recorded from using an intracellular electrode. As discussed in Chapter 2, HS cells are nonspiking cells in the lobular plate (third optic ganglion)

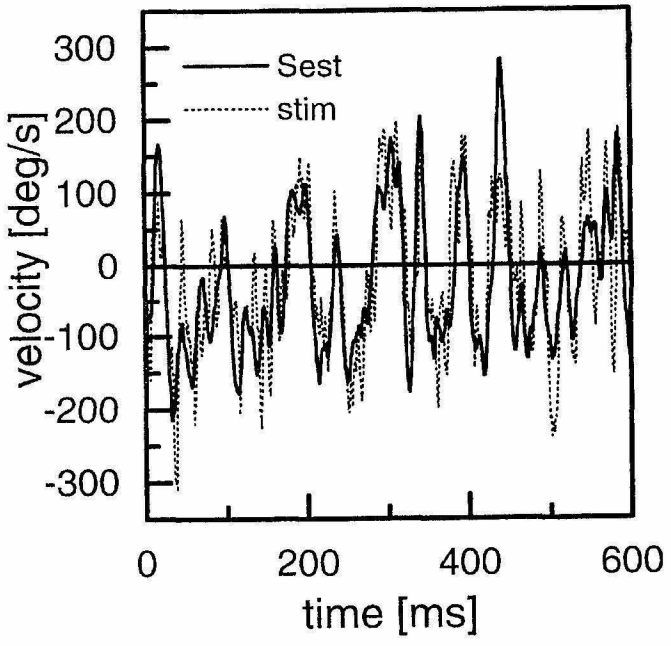

Figure 6.1: Stimulus reconstruction in the HS neuron. A segment of the random velocity stimulus is shown along with the estimated stimulus obtained by passing the cell's response through a reverse reconstruction filter. (From Haag and Borst, 1997)

that respond to horizontal motion in the visual field. These cells exhibit relatively symmetric hyperpolarization and depolarization on the order of ±10 mV to bidirectional motion. During stimulus presentation, which lasted 40 seconds, the intracellular potential was recorded using a sampling rate of 2 kHz. Each 40-second stimulus and response [$s(t)$ and $r(t)$] was broken down into 4-second segments [$s_i(t)$ and $r_i(t)$]. The 40-second stimulus was presented multiple times as well.

Using the cell's response to the random velocity stimuli, a reverse reconstruction filter was calculated using Equation 6.1. This filter was used to generate the estimated stimulus $sest(t)$. Figure 6.1 shows a 600 ms excerpt of the stimulus with the estimated stimulus superimposed. Based on the linear reconstruction derived from the intracellular response, it is clear that the HS cell encodes slow changes in pattern velocity, but fails to encode rapid changes.

This frequency-limited encoding can be quantified by computing the coherence function $\gamma^2(f)$, as shown in Figure 6.2. A coherence of 1 indicates that the stimulus is reconstructed perfectly by the reverse reconstruction filter at a certain frequency. The coherence plot indicates that the cell encodes velocity fluctuation information up to approximately 10 Hz.

There are two possible reasons for a coherence value less than one. First, there may be noise added to the signal in the cell. Second, the encoding may be nonlinear, which the linear reconstruction filter could not account for. We know from our analysis of Reichardt motion detectors in Chapter 3

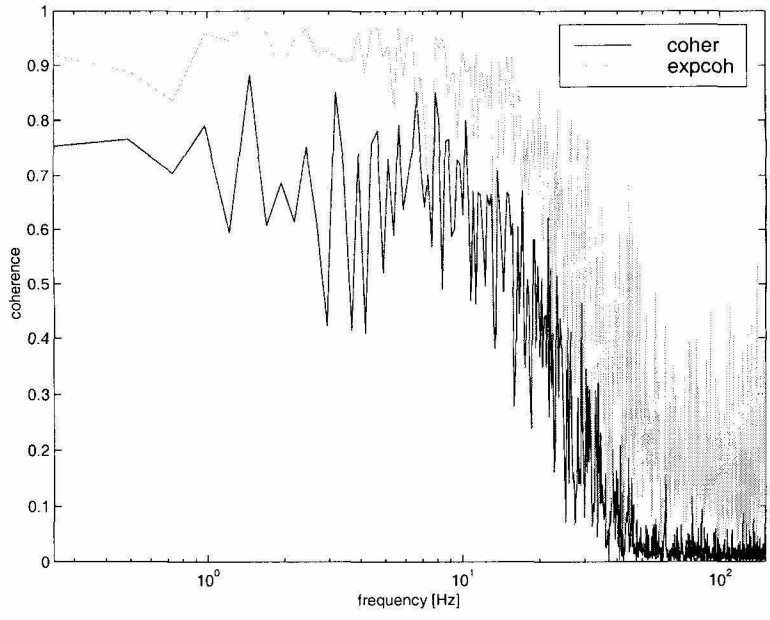

Figure 6.2: Coherence functions for fly HS cell. A coherence of 1 indicates that the stimulus is reconstructed perfectly by the reverse reconstruction filter at a certain frequency. The coherence measure is derived from stimulus reconstruction fidelity, while expected coherence measure is derived from trial-to-trial response variability. Data courtesy of Jürgen Haag and Alexander Borst (Haag and Borst, 1997).

that models of the HS cell do not predict a linear encoding of velocity. In order to determine which effect (noise or nonlinear encoding), dominates the loss of coherence, Haag and Borst estimated the noise independently. By presenting the same stimulus on repeated trials, a "mean response" could be calculated. Noise was calculated by subtracting this mean response from each individual responses. The spectra of these residuals were averaged, and signal and noise spectra were produced (see Figure 6.3).

From the frequency-dependent signal-to-noise ratio $[snr(f)]$, Equation 6.5 can be used to calculate the expected coherence assuming a linear encoder (i.e., the imperfect encoding is caused only by additive noise in the system, not nonlinearities). This expected coherence is plotted with the actual coherence in Figure 6.2. Based on this data, Haag and Borst conclude:

> Although the coherence functions expected from a linear encoding are settled at ~0.9 for frequencies <10 Hz, the neural coherence functions reach only ~0.7, similar to what has been found in the experiments described above. The consequence of this finding is that about one-third of the missing accuracy can be accounted for by the statistical fluctuations in the neural signals, whereas the remaining two-thirds have to be attributable to nonlinear encoding (Haag and Borst, 1997).

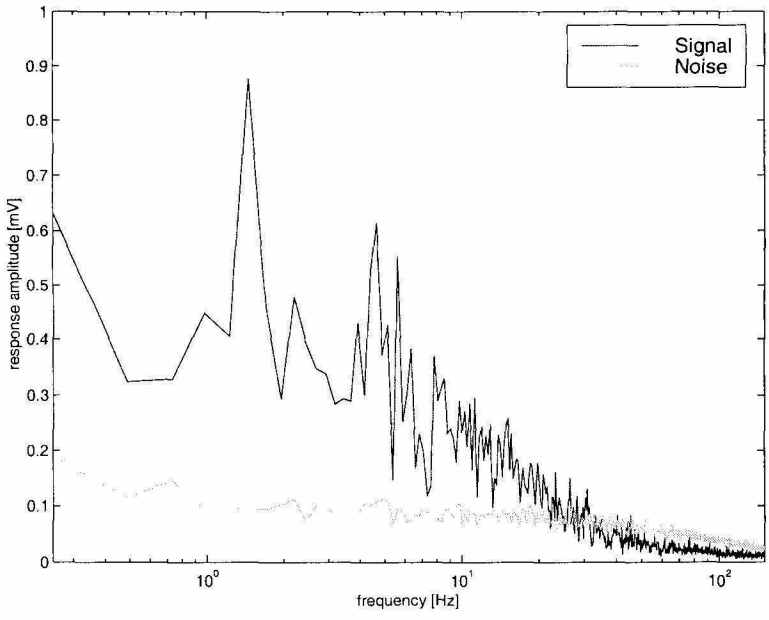

Figure 6.3: Signal and noise spectra for fly HS cells. Data courtesy of Jürgen Haag and Alexander Borst (Haag and Borst, 1997).

6.3 Pattern Velocity Estimation by Silicon EMD Array

We repeated the above experiment using a silicon 1 × 13 EMD array. The experiment was performed using the same experimental apparatus and control software as in the fly experiments. The current-mode output of the chip was converted to a voltage by a sense amplifier, and this voltage was sampled at 2 kHz during stimulus presentation.

Figure 6.4 shows a segment of the stimulus and the estimation derived from the chip response. Comparing this with Figure 6.1, we see that the chip performs similarly to the HS cell; it follows the stimulus at low frequencies, but misses high-frequency information. The coherence function for the silicon system is displayed in Figure 6.5, along with the expected coherence function. Figure 6.6 shows the signal and noise spectra for the chip. From this data, one can see that the relative noise level in our chip is several times lower than that in the fly's HS cell. Most of the missing accuracy in chip is due to coding nonlinearities, not noise.

Figure 6.7 overlays the measure coherence functions of the HS cell and the chip. The fly encodes velocity fluctuations in the 5 Hz to 20 Hz range with higher accuracy than our chip. We believe that parasitic capacitances in the current-mode lowpass filter begin introducing nonlinearities at these frequencies, resulting in a less accurate encoding.

We used information-theoretic techniques to quantify upper and lower bounds on the mutual information encoded by the chip (Borst and Theunissen, 1999). The lower bound and upper bound

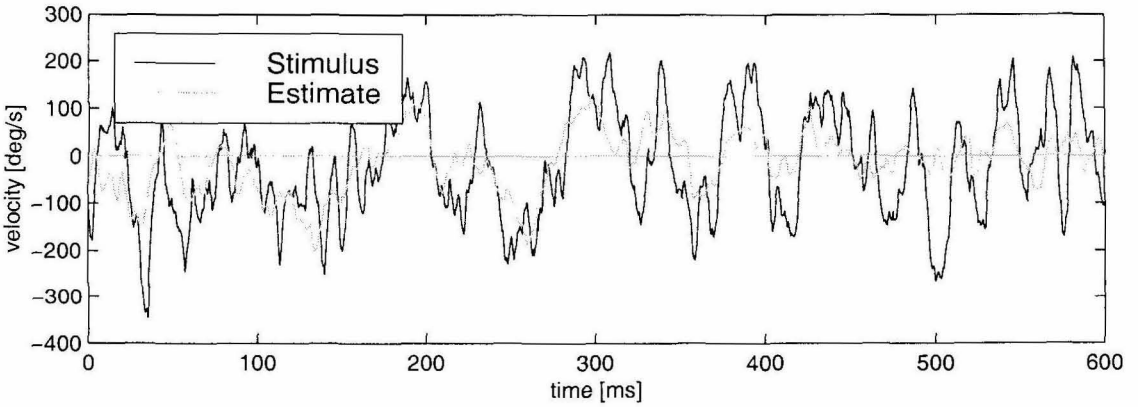

Figure 6.4: Stimulus reconstruction in the silicon EMD array. Compare to Figure 6.1; stimulus is identical.

of the transinformation T in bits/second can be defined as

$$T_{LB} = -\int_0^{f_{max}} \log_2 \left[1 - \gamma^2(f)\right] df \tag{6.6}$$

$$T_{UB} = -\int_0^{f_{max}} \log_2 \left[1 - \gamma^2_{exp}(f)\right] df \tag{6.7}$$

Using an upper frequency limit f_{max} of 50 Hz, the lower-bound information rate of the chip was calculated to be 27 bits/sec, compared with 37 bits/sec in the HS neuron using identical stimuli and upper frequency limit (Haag and Borst, 1998). The channel capacity (upper bound) of our sensor was 137 bits/sec compared with 110 bits/sec in the HS neuron. This is not surprising considering our chip dissipated many orders of magnitude more power than the fly's visual system. The true measure of transinformation probably lies closer to the lower bound. Our chip used far fewer EMDs than an HS neuron, and was thus subject to greater pattern dependence which obscured velocity information. Parasitic effects and second-order nonlinearities also prevented our sensor from encoding velocity information as efficiently as the fly's HS neuron.

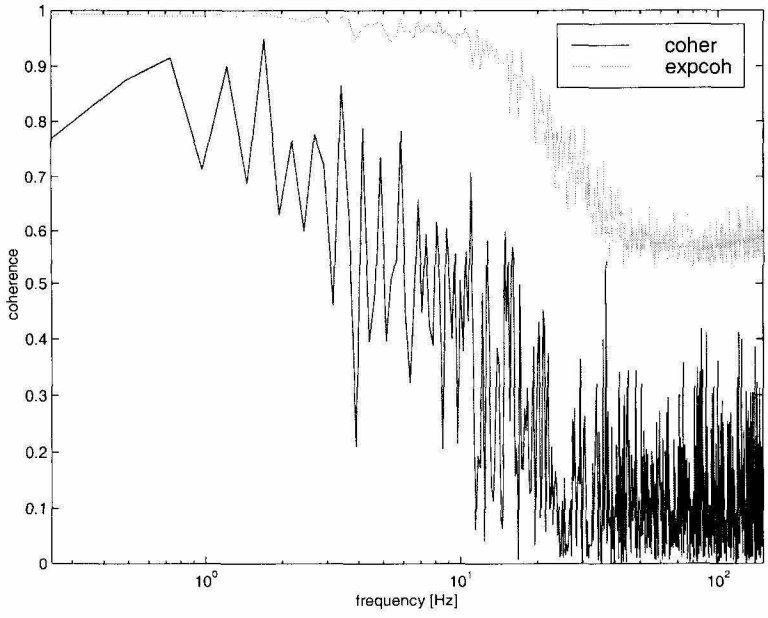

Figure 6.5: Coherence functions for silicon EMD array. The coherence measure is derived from stimulus reconstruction fidelity, while expected coherence measure is derived from trial-to-trial response variability.

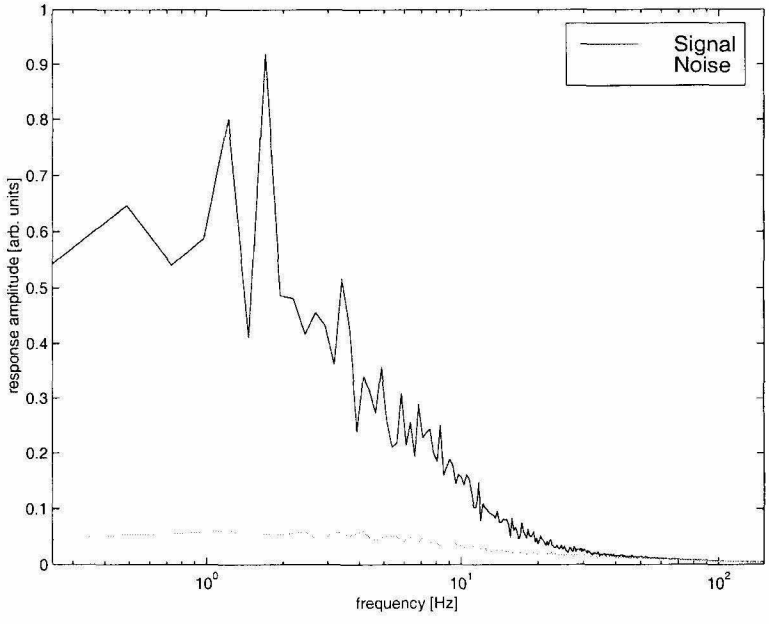

Figure 6.6: Signal and noise spectra for silicon EMD array. Our sensor's response has a higher SNR than the HS neuron (compare with Figure 6.3). This is not surprising since it dissipates many orders of magnitude more power.

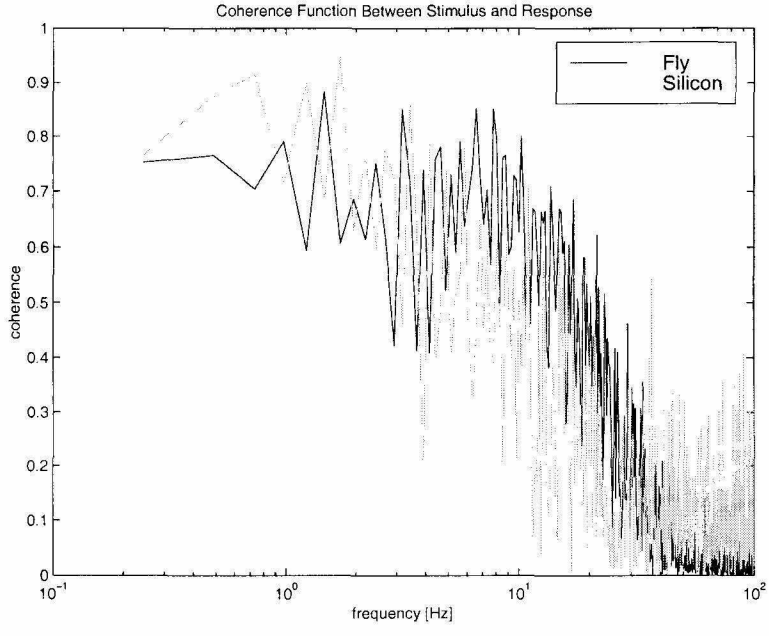

Figure 6.7: Coherence function for fly and silicon EMDs. The two systems encode stimulus information with nearly identical fidelity to around 3 Hz. The fly neuron clearly outperforms the silicon system for frequencies between 6 and 20 Hz. Information rate (lower bound) is proportional to the area under the coherence curve.

Chapter 7 Optomotor Control

In Chapter 2, we introduced the optomotor response, where visual motion information is used as a feedback control signal to estimate and cancel self-rotation. This sensorimotor loop is perhaps the best-studied visually-guided behavior of the fly. We will first describe optomotor experiments performed with flies and then describe real-time experiments where our sensor was compared directly against flies.

7.1 Measuring the Optomotor Response

7.1.1 Experiments Previously Performed on Flies

Warzecha and Egelhaaf recently characterized the optomotor behavior of the fly under closed-loop conditions (Warzecha and Egelhaaf, 1996). A female sheepfly (*Lucilia cuprina,* Calliphoridae) was rigidly attached to a meter that measured yaw torque produced while the fly attempted to turn in response to visual stimuli (Figure 7.1a), reducing the fly's behavior to a single degree of freedom. Vertical bars were presented to a large region of the fly's visual field, and could be drifted clockwise or counterclockwise. In closed-loop experiments, the fly's yaw torque was measured in real time and scaled by a constant gain term to yield angular velocity. This simulates the observed dominance of air friction in determining the instantaneous angular velocity in flies (Reichardt and Poggio, 1976). The fly's simulated angular velocity was subtracted from the angular velocity imposed by the experimenter. The resulting signal was used to control the drift rate of the visual stimulus. This simulated free-flight conditions, and allowed evaluation of the optomotor system performance.

The imposed motion schedule consisted of 3.75 s of zero imposed motion, then 7.5 s of clockwise rotation at 44 deg/s. Figure 7.2a shows the torque data and resulting stimulus position for an individual trial. Figure 7.2b shows the averaged data over 139 trials in a total of five animals. See Warzecha and Egelhaaf, 1996, for details on the experimental protocol.

The fly is able to stabilize its flight and cancel out most of the imposed motion. Simulation results suggest that the nonmonotonic temporal frequency response of Reichardt motion detectors results in greater stability for the optomotor control system (Warzecha and Egelhaaf, 1996). The individual trials show an oscillatory component to the torque response around 2 Hz. This oscillation is not phase-locked to the stimulus since it is not present in the average torque trace. Oscillations are not observed under open-loop conditions, suggesting they arise from optomotor feedback (Geiger and Poggio, 1981; Warzecha and Egelhaaf, 1996). Notice that despite the large amplitude of the

torque oscillations, the position trace is not dominated by this effect. This fluctuation amplitude, in terms of number of photoreceptors, is close to the amplitude observed in human microsaccades (Warzecha and Egelhaaf, 1996). Poggio and colleagues observed similar oscillations in closed-loop experiments and proposed that they arose from the 60–75 ms synaptic delay inherent in the fly visual system (Geiger and Poggio, 1981; Poggio and Reichardt, 1981).

7.1.2 Duplicating Experiments with the Silicon System

We were able to directly replicate these experiments with our silicon analog of the optomotor system (Figure 7.1b). To provide visual stimulation, we used an LED display with a 200 Hz refresh rate which is currently being used to test flies in closed-loop experiments. The stimulation time schedule was identical to the fly experiments, but an angular velocity of 50 deg/s was used. Our chip had a much smaller field of view (10°) than the fly, so we set the stimulus distance such that the EMD array saw approximately one wavelength of the pattern. The output signal from the silicon model of the HS cell was passed through an off-chip first-order lowpass filter with a time constant of 680 ms, modeling the behavior of the thoracic motor centers (Egelhaaf, 1987; Wolf and Heisenberg, 1990; Warzecha and Egelhaaf, 1996). The filtered output of the chip was treated exactly like the signal from the torque meter in the fly experiments, and closed-loop experiments were run in real time. Figure 7.3a shows torque and position data from the chip for an individual trial, and Figure. 7.3b shows the averaged response over 100 trials.

The silicon system shows the same ability to greatly cancel the imposed motion. The fly showed an average drift of 9.4% of the open-loop drift velocity, with position fluctuations of 7.8° (standard deviation) about this drift. The chip showed an average drift of 22% of the open-loop drift velocity, with position fluctuations of 6.2° (s.d.) about this drift. Also, we observe the same 2 Hz oscillations in the individual trials. Since we did not build any explicit delay into our system, this demonstrates that the phase lags and nonlinearities in this simple model are sufficient to produce oscillations, even in the absence of additional synaptic delays.

We believe this hardware modeling approach will prove increasingly valuable in the future, as biological models of the neural circuitry underlying more complex and sophisticated behaviors arise. To simulate a sensorimotor system in software, one must construct *two* models: a model of the biological system, and a model of the world. The physical environment is an essential element in a sensorimotor feedback loop, so this world model must increase in detail as we study more advanced behaviors. Since animals interact with their three-dimensional environment in very dynamic ways, it may not be long before software simulations of sensorimotor systems require more computational resources to model the world than to model the neural circuitry of interest.

Results described in this section also were published in Harrison and Koch, 2000b.

7.2 Application to Autonomous Vehicle Control

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.

In the past decade, researchers have been endowing mobile robots with biologically-inspired (more specifically, insect-inspired) visual systems (Franceschini et al., 1992; Srinivasan 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.

7.3 Robot 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).

7.3.1 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 7.4 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 Figure 7.5). 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 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 Figure 7.6). 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 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 Figure 7.4). 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.

7.3.2 Robot 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 Figure 7.7). 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, Colchester, VT). The scale of experiments was limited by the range of this system, approximately a 70 cm $\times$ 140 cm area for highest accuracy.

The optic flow feedback proved capable of nearly eliminating the effect of physical asymmetry. Figure 7.8 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 7.9 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 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 7.10 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.

Occasionally, the feedback did fail to keep the course straight. A 45° turn is visible in Figure 7.9, 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.

Results described in this section also were published in Harrison and Koch, 1999b.

As mentioned in Chapter 5, 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.

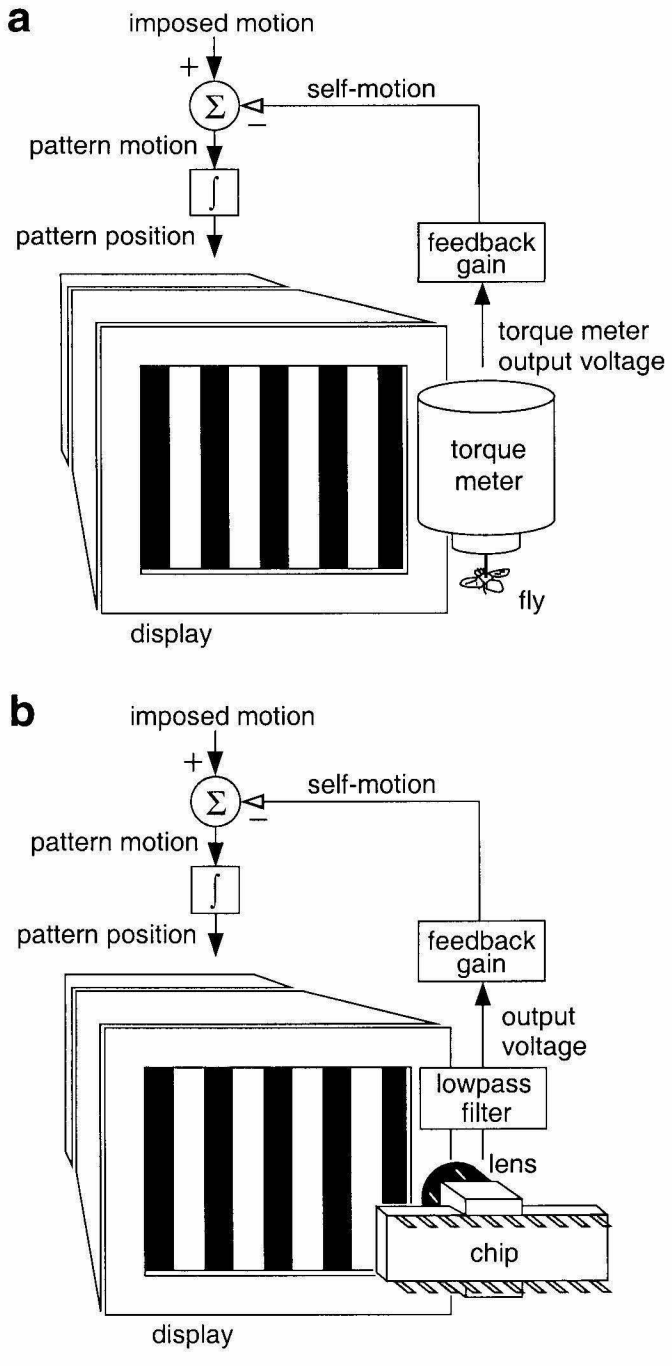

Figure 7.1: Experimental methodology. (a) Setup used by Warzecha and Egelhaaf to measure the closed-loop torque response of the sheepfly *Lucilia* (Warzecha and Egelhaaf, 1996). The torque meter output is scaled to produce a measure of what the fly's self-motion would be if it were free to rotate. This self-motion is subtracted from the imposed motion to determine the pattern motion, creating the illusion of free flight in a room with distant walls. (Only rotation, not translation, is simulated.) (b) Setup used to replicate the closed-loop experiments with the silicon model. The output voltage from the circuit is used in place of the torque meter output voltage. The rest of the system is identical to (a).

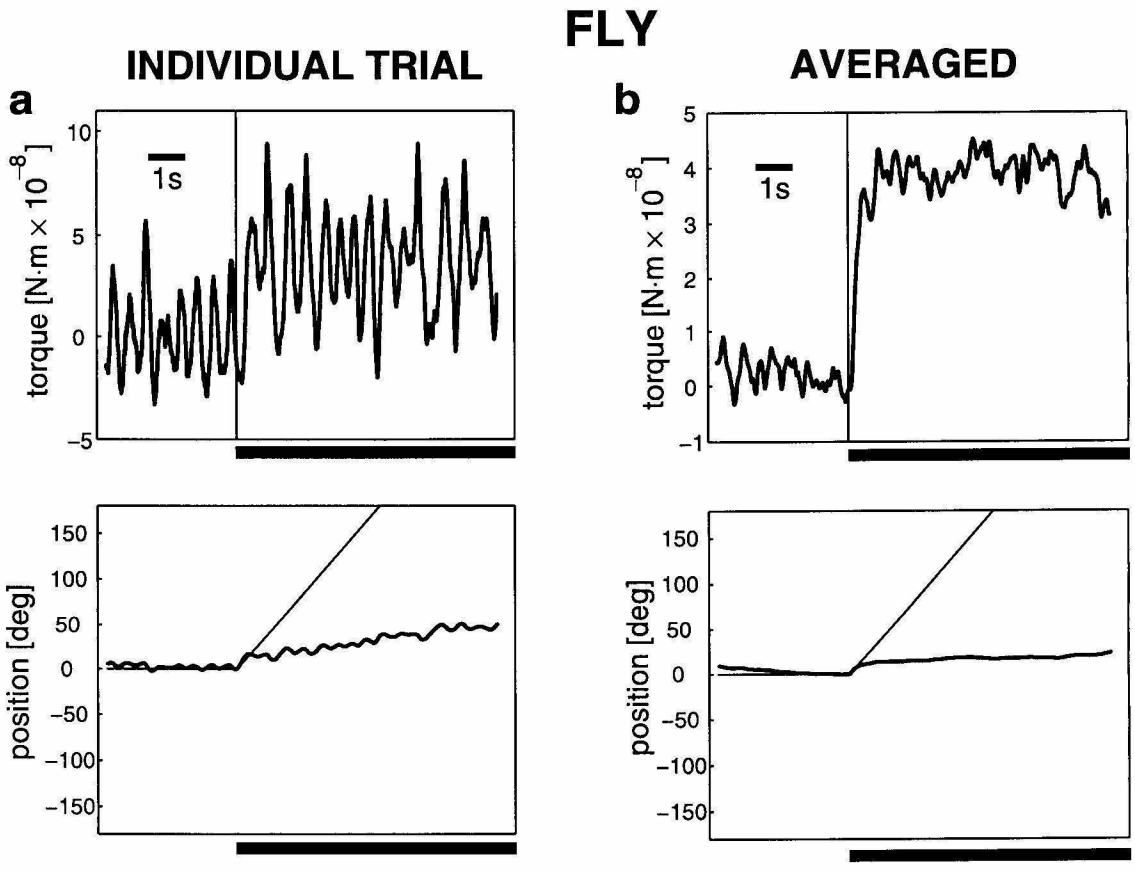

Figure 7.2: Fly's optomotor behavior. (a) Torque (top panel) and angular position (bottom panel) vs. time for an individual closed-loop trial with a fly. Dark horizontal bar indicates experimenter-imposed rotation. Thin lines on position trace indicate position in the open-loop case. Most of the imposed rotation is canceled out by the optomotor control system. Since the position is proportional to the integral of the torque (see text for details), large torque oscillations do not cause large position oscillations. (b) Averaged torque response and angular position trace for multiple trials (N = 139, 5 flies). The fly showed an average drift of 9.4% of the open-loop drift velocity, with position fluctuations of 7.8° (standard deviation) about this drift. Data in (a) and (b) redrawn from Warzecha and Egelhaaf, 1996.

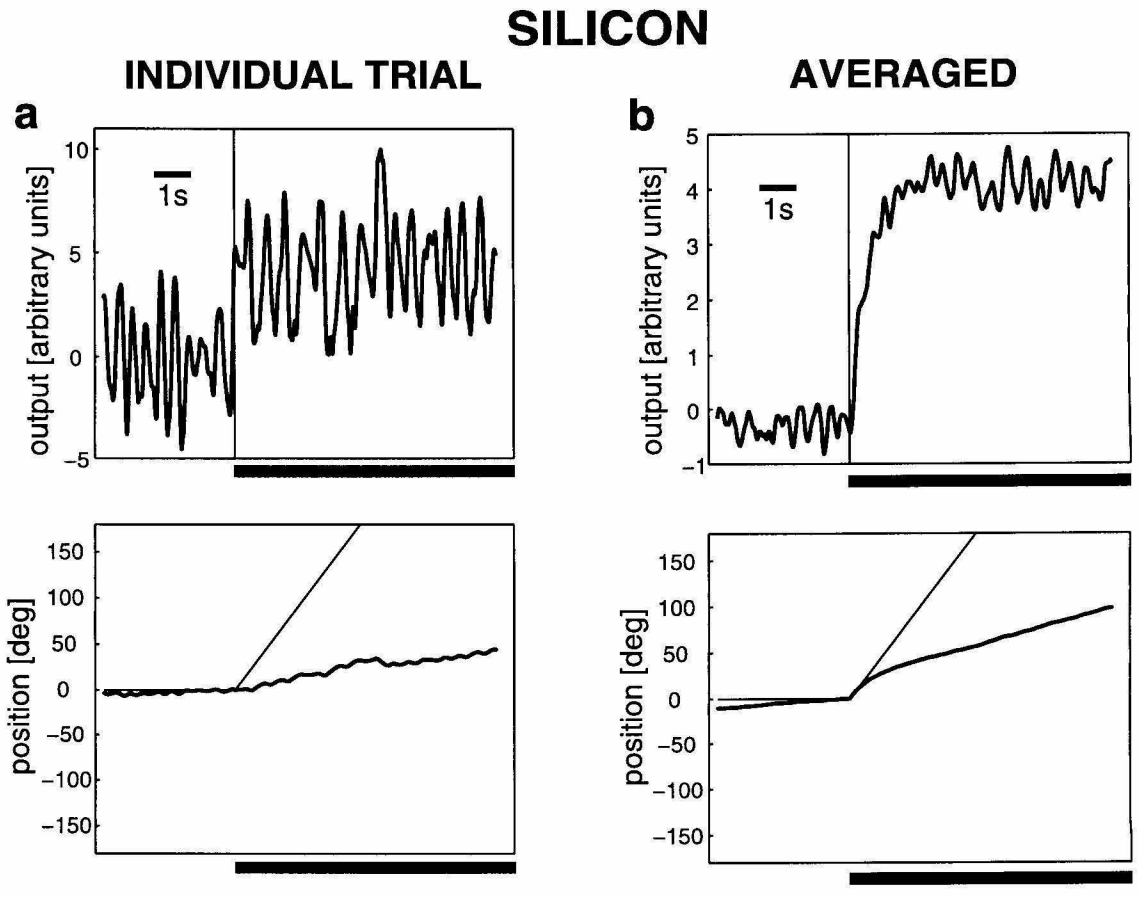

Figure 7.3: The optomotor behavior of our silicon system. (a) Chip output signal (analogous to torque) and position vs. time for the silicon system in an individual trial. (b) Averaged torque response and angular position trace for multiple trials (N = 100, 1 chip). The chip showed an average drift of 22% of the open-loop drift velocity, with position fluctuations of 6.2° (s.d.) about this drift.

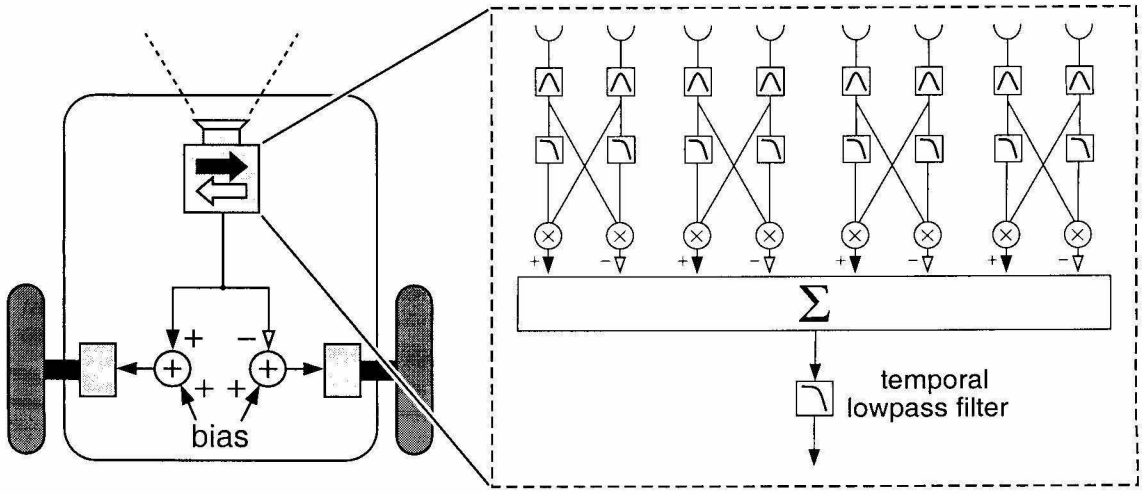

Figure 7.4: Schematic of our optomotor system. A motion sensor chip is mounted facing forward on a two-wheeled robotic platform. Forward-facing motion sensors are largely blind to optic flow produced by forward translation, so we only measure rotation. The chip's wide-field output is used as an estimate of self rotation, then lowpass filtered (τ = 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.

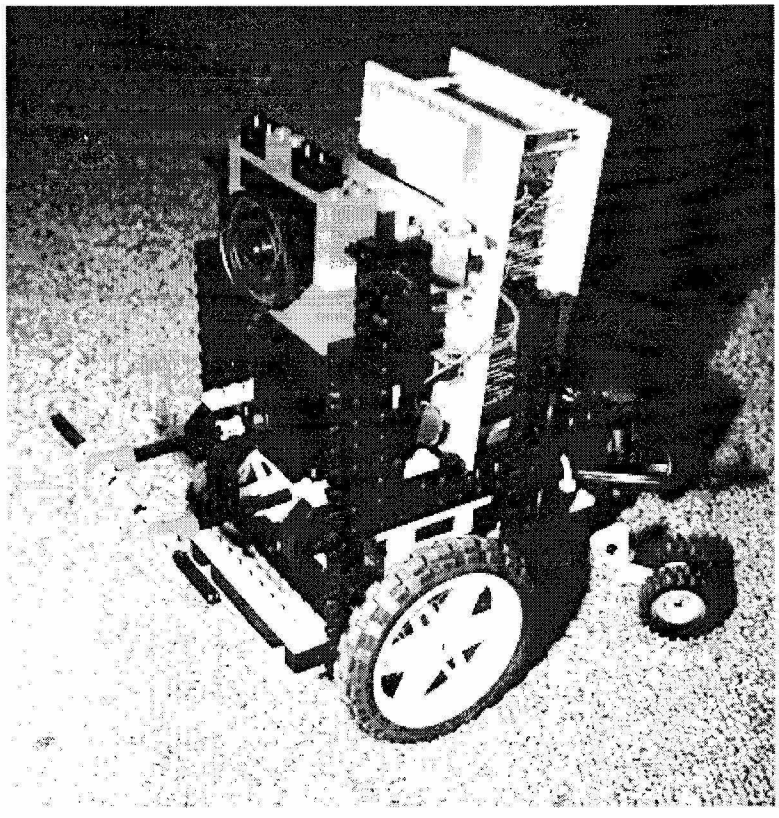

Figure 7.5: Photograph of the optomotor system. The lens is covering the motion sensor chip. Additional off-chip electronics have been constructed to drive the dc motors. The back wheels turn freely, and merely prevent the robot from falling over. The robot measures 13 cm $\times$ 19 cm $\times$ 22 cm, and is powered by on-board batteries.

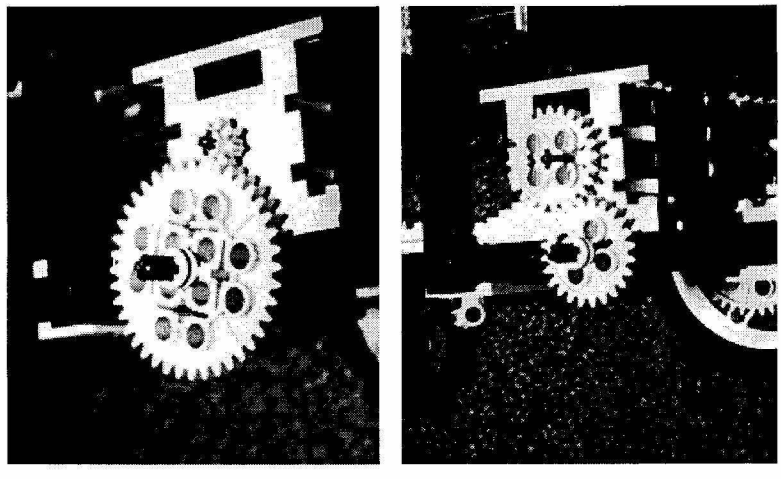

Figure 7.6: Asymmetrical gear ratios. The left and right motors drove their wheels with 1:5 and 1:1 gear ratios, respectively. This caused the robot to drive in tight circles if both motors were driven at the same speed.

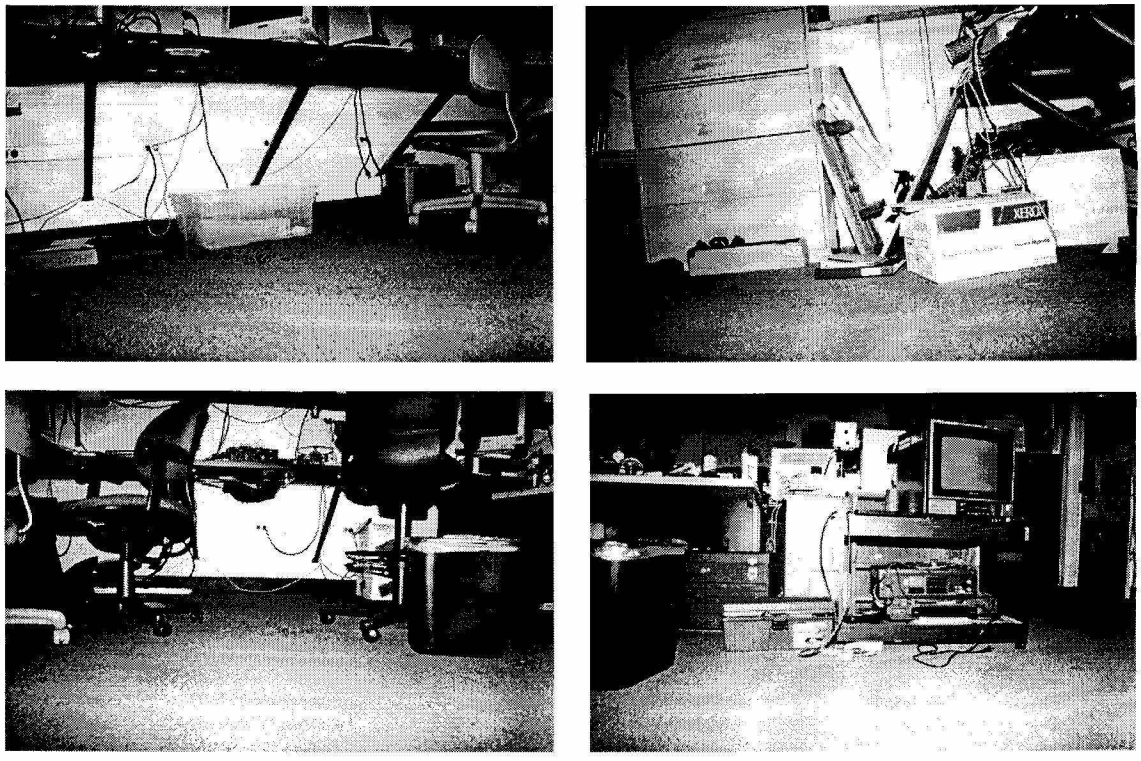

Figure 7.7: Visual environment during the experiments. The robot saw a typical, cluttered laboratory environment which was not changed in any way to accommodate the motion sensor.

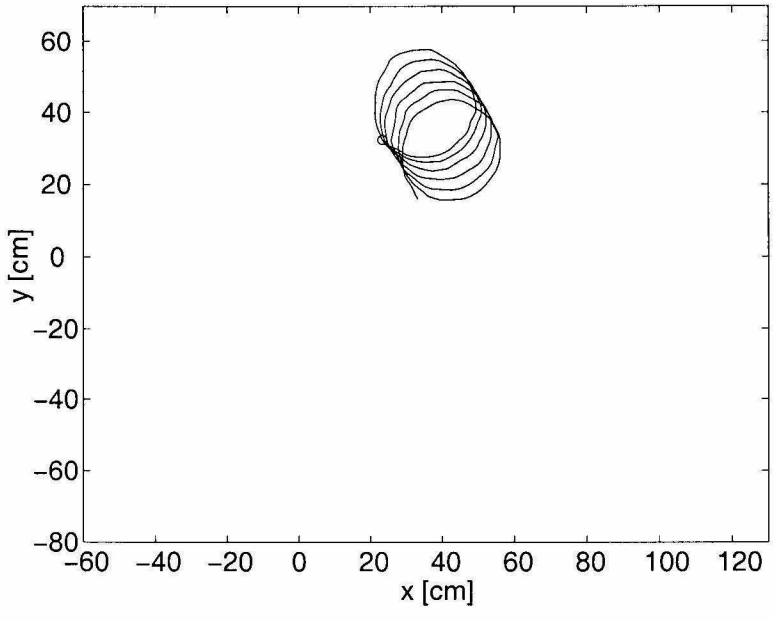

Figure 7.8: Robot path with no sensory feedback. With the motion sensor disabled, the robot turns in circles due to the asymmetry in its mechanics. The circle denotes the ending location of the robot.

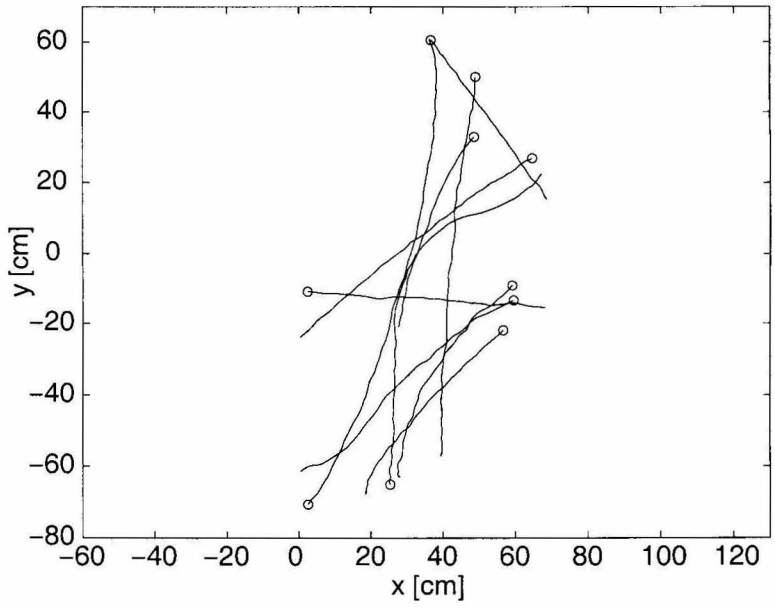

Figure 7.9: Robot path with sensory feedback. Ten trials are shown where the motion sensor provided optomotor control to straighten the course of the robot despite a 5:1 mechanical asymmetry. Circles denote the ending location of the robot in each trial. The robot was exposed to different visual scenes during the trials.

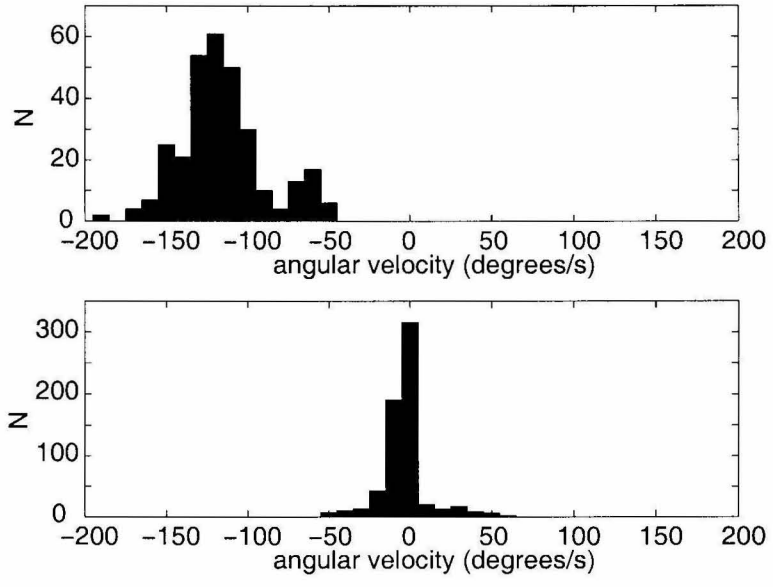

Figure 7.10: Histogram of angular velocities. (a) No visual feedback. The turning behavior of the robot is obvious. The mean angular velocity was -116 deg/s. (b) Compilation of all ten trials with visual feedback. The mean angular velocity was greatly reduced to -3.7 deg/s.

Chapter 8 Nonlinear Spatial Integration

In the previous chapters, we have described arrays of Reichardt elementary motion detectors with linear spatial integration—currents from each EMD are tied to a single wire and sum linearly. In this chapter we will explain why linear spatial integration is undesirable and introduce a circuit architecture of nonlinear integration based on the properties of wide-field lobular plate neurons in the fly's optic lobe.

Real-world optic flow fields are sparse. Natural images have areas of little or no contrast such as blank patches of sky (see Figure 8.1). The lack of spatial detail in these areas leads to "holes" in the otherwise full-field patterns of optic flow produced by rotation. When estimating self-rotation, for example, one would like to extract information based on wide-field motion that is robust against these gaps.

8.1 Gain Control in Fly Tangential Neurons

Flies have developed a remarkably elegant method for dealing with optical flow sparseness. The optic lobe in the brain of the fly contains several wide-field motion-sensitive neurons that integrate motion information from many elementary motion detectors (EMDs) in large receptive fields to produce estimations of self-rotation (Krapp and Hengstenberg, 1996). These neurons have been studied for decades, and much is known about their response properties. One property exhibited by some of these cells is called gain control, and seems to make the sensory response robust against gaps in the optical flow field.

Gain control describes the saturating response of these motion-sensitive cells with increasing stimulus size (see Figure 8.2). As the extent of the stimulating pattern across the visual receptive field increases linearly, the response of the cell saturates, but it saturates at different levels for different stimulus velocities. This cannot be explained by a simple saturating output channel. The wide-field motion-sensitive neuron is integrating motion information spatially, but this integration is nonlinear.

This size-dependent saturation assures that at reasonably high levels of stimulation, the cell is not sensitive to gaps in the optic flow field. (Featureless parts of the visual scene decrease the effective stimulus size.) The cell now encodes the stimulus velocity, which in this case may represent a measure of self-rotation, largely independent of the visual sparseness of the environment.

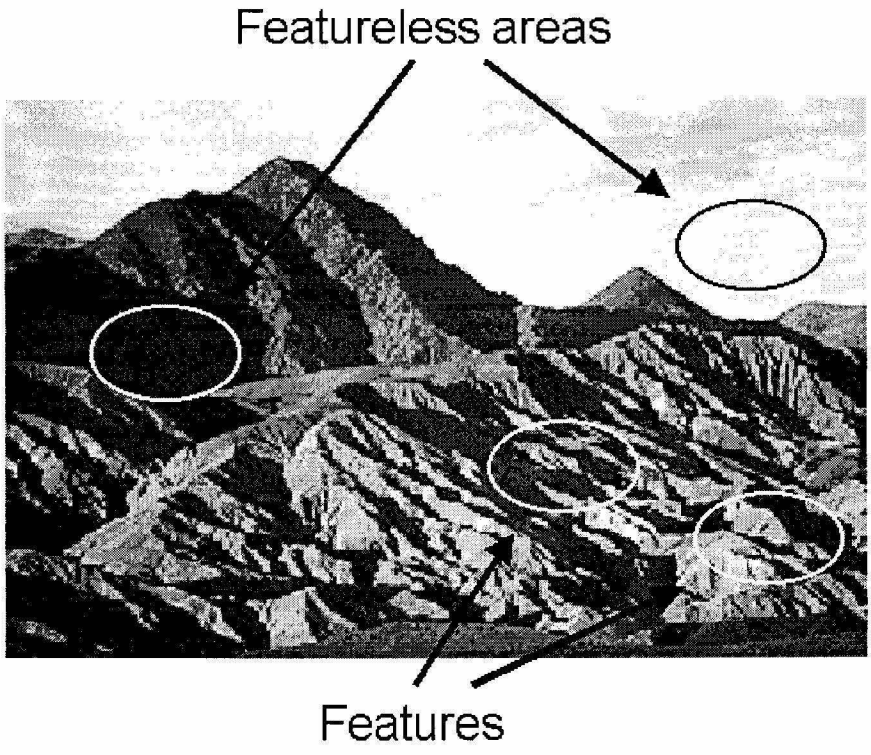

Figure 8.1: Gaps in optic flow fields. In natural images, some areas are featureless, and thus convey no local motion information. They effectively decrease the pattern size. A wide-field motion sensor ideally should integrate optic flow from a large field of view, but the presence of some featureless areas should not affect its output. ("Ignore the blank parts.")

8.2 Algorithm and Biological Architecture

The neural architecture and biophysical mechanisms underlying gain control in the fly are now understood (Borst et al., 1995; Single et al., 1997). Simple linear models of spatial integration result in an output that is linearly dependent on stimulus size (see Figure 8.3a). Size-dependent saturation comes about if we use a more accurate model of the wide-field motion-sensitive neuron (see Figure 8.3b). In this model, the EMD outputs are not directly conveyed to the wide-field neuron. Instead, the EMDs modulate synapses, which are modeled as conductances between the intracellular potential and a fixed ion reversal potential. Depending on the type of ion involved, the reversal potential can be above or below the resting potential of the cell, creating excitatory or inhibitory synapses. We connect preferred-direction EMDs to excitatory synapses and null-direction EMDs to inhibitory synapses. Each EMD modulates its corresponding ion channel conductance, which acts to pull the cell away from its resting potential, where it is held by the fixed leakage conductance g_{rest}.

If we measure all voltages relative to the cell resting potential V_{rest}, then the cell potential can

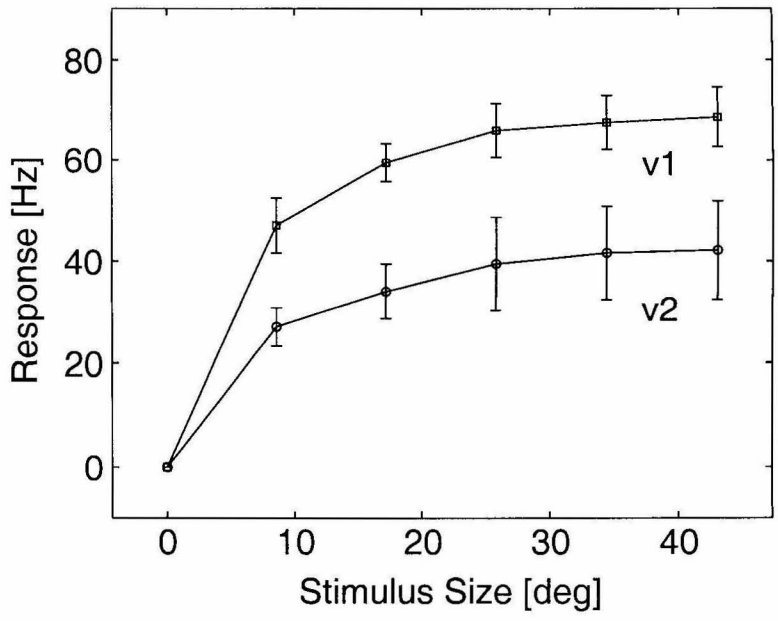

Figure 8.2: Gain control in a wide-field motion-sensitive neuron in the fly. The cell response saturates with increasing pattern size, and saturates at different levels depending on the stimulus velocity ($v1 = 72°/\text{sec}$; $v2 = 360°/\text{sec}$). Stimulus was a sinusoidal grating with spatial wavelength of 24° and 29% contrast. Data represent mean ± SEM of extracellular recordings from the lobula plate spiking neuron H1 of four different female blowflies (*Calliphora erythrocephala*). Data reprinted from (Single et al., 1997).

be expressed as:

$$V_{cell} = \frac{g_{exc}V_{exc} + g_{inh}V_{inh}}{g_{exc} + g_{inh} + g_{rest}} \tag{8.1}$$

Here, V_{exc} and V_{inh} represent the ion reversal potentials, and g_{exc} and g_{inh} represent the ion channel conductances, which are controlled by the preferred-direction and null-direction EMD outputs. Thus V_{cell} saturates at V_{exc} for $g_{exc} \gg g_{inh}, g_{rest}$, and saturates at V_{inh} for $g_{inh} \gg g_{exc}, g_{rest}$.

To generate gain control, this method of spatial integration relies on the particular architecture of the EMD. The Reichardt-type correlation-based EMD exhibits only weak direction selectivity and produces a significant but weaker response for stimuli in its null direction (see Figure 8.4). Strong directionally selectivity is achieved only after the EMD opponent pair signals are subtracted (Borst and Egelhaaf, 1990). Importantly, it has been demonstrated that the degree of directional selectivity is itself a function of stimulus velocity (Borst et al., 1995). Thus, a particular stimulus velocity will generate a characteristic ratio of preferred-direction to null-direction EMD output. This in turn will lead to a characteristic voltage that the cell is driven towards. We can rewrite Equation 8.1 as

$$V_{cell} = V_{stimulus}\left(\frac{g_{stimulus}}{g_{stimulus} + g_{rest}}\right) \tag{8.2}$$

where

$$V_{stimulus} = \frac{g_{exc}V_{exc} + g_{inh}V_{inh}}{g_{exc} + g_{inh}} \tag{8.3}$$

$$g_{stimulus} = g_{exc} + g_{inh} \tag{8.4}$$

Since conductances sum in parallel, a larger pattern size will cause more EMD opponent pairs to be active, producing a larger $g_{stimulus}$ and driving V_{cell} away from V_{rest} and towards the $V_{stimulus}$ that is characteristic of the stimulus velocity.

8.3 Silicon Implementation

The EMD circuits described in Chapter 4 produce current-mode outputs, which allowed us easily to sum the outputs of an EMD array simply by tying the wires together. But this linear integration across a receptive field does not produce a saturating size response. To produce this response, we developed a silicon implementation of the biophysical model shown in Figure 8.3b.

The silicon implementation of excitatory and inhibitory synaptic conductances consumes little silicon area or power. To achieve a variable conductance, we use a differential transconductance amplifier in a follower configuration (i.e., the output directly connected to the negative input) so that the output node is driven toward the input voltage (see Figure 8.5a) (Mead, 1989). As long as the follower is operated in the subthreshold regime ($I_b < 100$ nA) and in the linear region of operation (the input and the output voltages differ by less than 75 mV) the circuit has an output conductance g_{out}, which is given by

$$g_{out} = \frac{\partial I_{out}}{\partial V_{out}} = \frac{q\kappa}{2kT} I_b \tag{8.5}$$

where κ is the subthreshold slope, a CMOS process parameter that typically has a value close to 0.7. If we assume constant temperature, the circuit produces an output conductance proportional to the bias current I_b, approximately $g = (14\ \mathrm{V}^{-1})I_b$ at room temperature. By adding two additional diode-connected transistors in a source-degeneration configuration (Watts et al., 1992), we extend the linear range of this circuit from ± 75 mV to ± 150 mV (see Figure 8.5b).

We use this six-transistor circuit to model a population of ion channels whose ion reversal potential is specified by the input voltage. Thus each compartment of our cell has three ion channel populations modeled by three follower circuits. One channel has a constant leakage conductance g_{rest} and sets the resting potential of the cell, V_{rest}. The other two model the excitatory and inhibitory inputs from the local opponent pair of EMDs. These conductances, g_{exc} and g_{inh}, are controlled by the preferred- and null-direction EMDs and have reversal potentials above (V_{exc}) and below (V_{inh}) the resting potential, respectively. The current-mode output of each EMD directly biases a variable

conductance circuit.

8.4 Experiments

We fabricated a one-dimensional array with 28 compartments (each receiving input from an EMD) in a standard 1.2-μm CMOS VLSI process. Each compartment contained a preferred- and null-direction EMD and the corresponding variable conductance circuits described above. The chip included integrated photoreceptors, so imaging and computation were performed on a single die. We mounted a standard 4 mm CCTV camera lens over the chip, which gave the photoreceptors an angular spacing of 0.73°. We presented high-contrast drifting sinusoidal stimuli to the chip with a computer monitor (Sony Multiscan 17seII). We were able to achieve frame rates of 70 Hz, and screen resolution far exceeding the photoreceptor spacing.

8.4.1 Varying Pattern Size

We varied the extent of our stimulus across the chip's field of view. (All sizes are expressed as the fraction of the photoreceptor array that was stimulated.) First we investigated the behavior of the linear spatial summation architecture, achieved simply by tying together the output currents of all the EMDs in the array (see Figure 8.6a). The output increases linearly with pattern size. Next we investigated nonlinear spatial integration by directing the EMD output currents to the variable conductance circuits. We measured the voltage of our 28-compartment cell model as a function of pattern size (see Figure 8.6b). The cell response saturates with increasing pattern size. We also see that it saturates at different levels for different velocities.

To further quantify the degree of pattern size saturation, we fit the four traces shown in Figure 8.6 to a simple power law:

$$R = kS^{\alpha} \tag{8.6}$$

where S represents pattern size, R indicates chip response, α indicates the degree of nonlinearity, and k is a free parameter varied to fit each particular curve. The data from the linear spatial integration case (Figure 8.6a) were fit with $\alpha = 1.04$ and 1.08, indicating a linear relationship between pattern size and chip response. The data from the nonlinear spatial integration case (Figure 8.6b) were fit with $\alpha = 0.45$ and 0.51, indicating a compressive nonlinearity similar to a square root function.

8.4.2 Varying Leakage Conductance

We can control the nature of this saturation by increasing or decreasing g_{rest}, the conductance that tries to hold the cell at the resting potential. We investigated pattern size saturation for three different values of g_{rest} (see Figure 8.7). As g_{rest} decreases, it becomes easier to pull the cell away

from V_{rest} and only a few EMD opponent pairs need contribute to dominate the response. Rapid saturation decreases pattern size dependence but increases the sensitivity of the chip to small-field motion that might not be produced by self-motion.

8.4.3 Power Dissipation

This method for achieving gain control is extremely power efficient. Since the excitatory and inhibitory synapses are powered by the current-mode output of the existing EMD circuitry, the only power cost we incur for adding gain control is the power consumed by the leakage conductance g_{rest}, which is simply $V_{dd}I_b$. For the experiments shown in Figure 8.7, this power varied between 14 pW and 5.4 nW per compartment, depending on the value of g_{rest}.

The circuit implementation of this architecture can be made more efficient by noting that parallel conductances sum. Since the Reichardt detector output signals are currents, we can sum all excitatory signals and all inhibitory signals on two separate wires, and then use a single-compartment model with a global excitatory synapse and a global inhibitory synapse.

Results described in this chapter also were published in Harrison and Koch, 1999a.

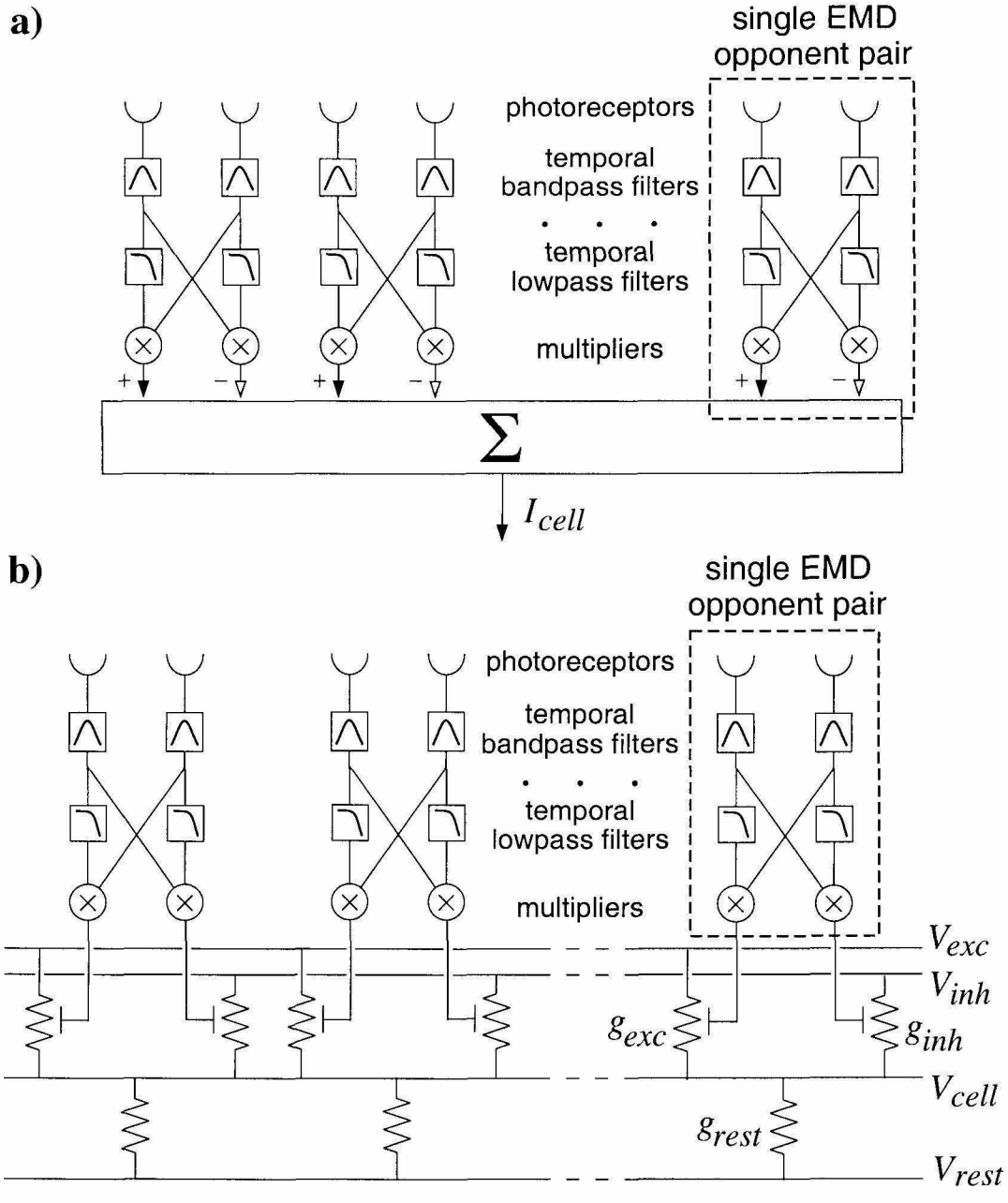

Figure 8.3: Models of wide-field motion-sensitive neurons in the fly. (a) Linear spatial integration. An array of Reichardt-type delay-and-correlate elementary motion detector (EMD) opponent pairs are subtracted locally to achieve strong direction selectivity. The spatially distributed outputs are summed. (b) Nonlinear spatial integration. EMD opponent pairs modulate excitatory and inhibitory synaptic conductances of the wide-field neuron. EMD activity can pull the neuron away from its resting potential V_{rest}.

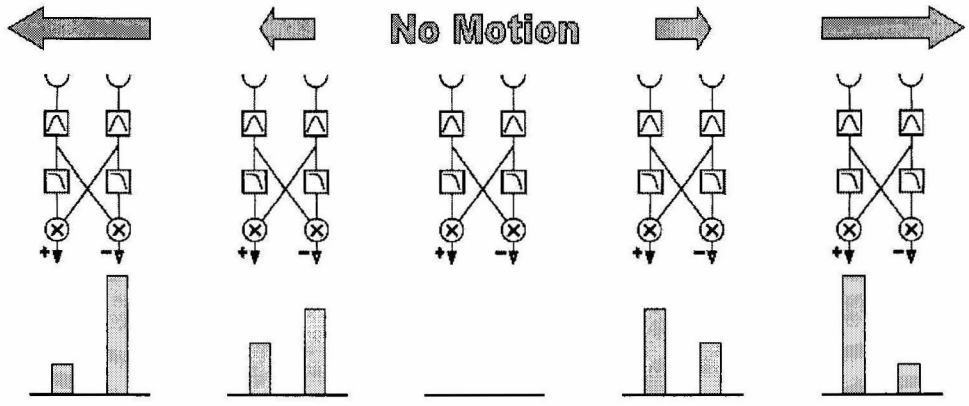

Figure 8.4: Weak direction selectivity. Because the Reichardt motion detector has weak direction selectivity before the opponent subtraction, stimuli in the null direction produce weaker but significant positive responses.

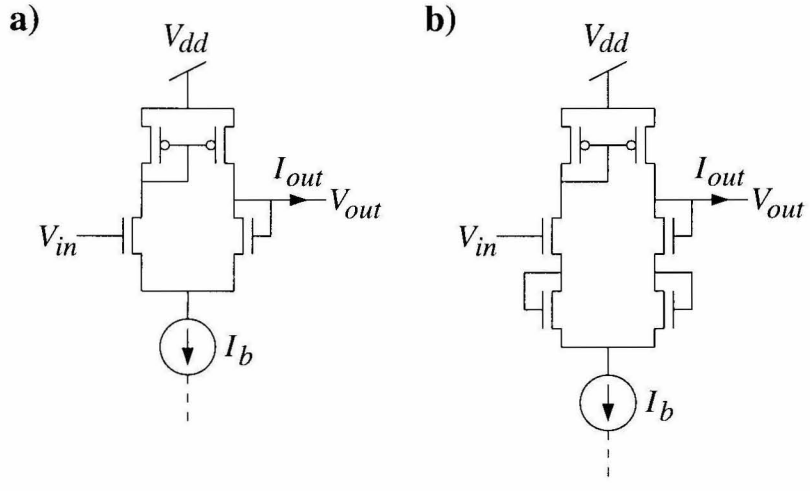

Figure 8.5: Variable conductance circuits. (a) Four-transistor follower. The output node is driven towards the potential at V_{in}. The output conductance of this circuit is proportional to the bias current I_b. (b) Six-transistor follower. Same behavior as the four-transistor circuit, but source-degeneration diode-connected transistors double the circuit's linear range.

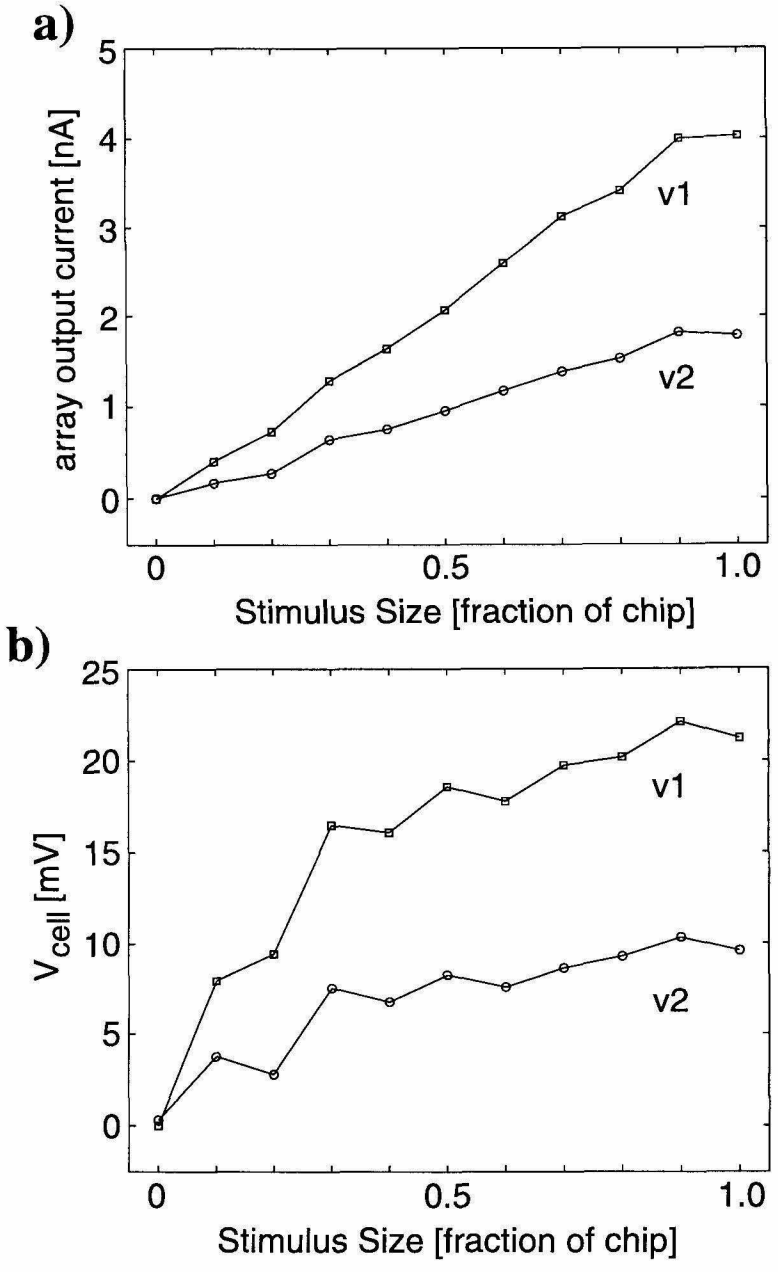

Figure 8.6: Gain control in the silicon system. (a) Linear spatial integration. The outputs from the EMD array are summed; gain control is not observed. (b) Nonlinear spatial integration with synaptic conductances. As is the biological system, the chip response saturates with increasing pattern size, and saturates at different levels depending on the stimulus velocity ($v1$ = 8.8°/sec; $v2$ = 4.4°/sec). The ion reversal potentials were V_{exc} = +150 mV; V_{inh} = −80 mV relative to V_{rest}. The asymmetry was necessary to counteract observed transistor mismatch. Stimulus was a sinusoidal grating with spatial wavelength of 2.9° and >99% contrast. Data represent mean output during stimulus presentation.

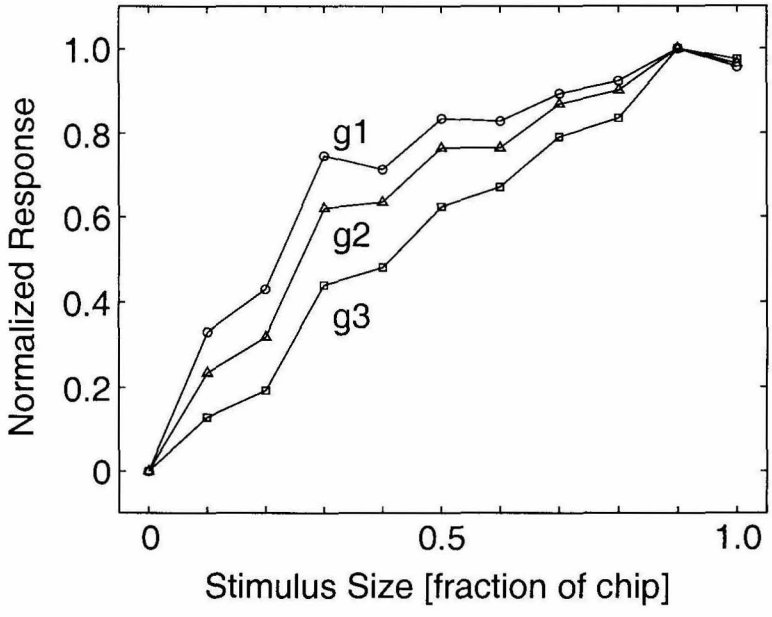

Figure 8.7: Varying leakage conductance. The degree of saturation can be tuned by changing the leakage conductance. Data are normalized for each value of g_{rest} ($g1 = (26\ \mathrm{G}\Omega)^{-1}; g2 = (1.0\ \mathrm{G}\Omega)^{-1}; g3 = (0.10\ \mathrm{G}\Omega)^{-1}$).

Chapter 9 System Integration

As discussed in Chapter 2, flies use a wide variety of sensors to execute tasks. In most autonomous systems, sensors of many types will be present. How can the motion sensors described in the previous chapters be integrated with sensors of other types, particularly vestibular sensors? This chapter addresses possible strategies for sensory fusion that use the strengths of one sensor to compensate for deficiencies in another.

9.1 Visual Motion Sensors and Their Limitations

Visual motion sensors have a significant limitation which must be taken into account when using their signals at the system level: In the absence of a patterned visual stimulus, it is impossible to compute optic flow. If the sensor is pointed towards a blank wall or if the environment is completely dark, the motion sensor will have a "zero optic flow" output regardless of the relative motion between the sensor and the external environment. In this case, we should ignore the motion sensor. On the other hand, if the sensor is looking at a high-contrast, patterned stimulus and reports "zero optic flow," that conveys a great deal of information. In order to combine visual and vestibular information reliably, we need some type of "confidence measure" from the visual motion sensor that conveys some information about the contrast of patterns in the field of view.

We propose a simple method to return a confidence measure from the motion sensor. We measure the local spatial derivative in the direction of each motion detector, then sum the absolute value of all these spatial derivatives over the receptive field of the sensor. If this value is low, we know there is no significant image contrast, and we should ignore the output of the motion sensor.

We use an "antibump" circuit to estimate the absolute value of the spatial derivative at each pixel. This compact, five-transistor circuit, originally described by Delbrück, produces a current that approximates the absolute value of the differential voltage input over a range of about $8U_T$ (Delbrück, 1993a). Figure 9.1 shows the antibump circuit. Its output current I_{out} is given by

$$I_{out} = I_b \left(1 - \frac{1}{1 + \frac{4}{S} \cosh^2 \frac{\kappa \Delta V}{2U_T}} \right) \tag{9.1}$$

where

$$\Delta V = V_1 - V_2 \tag{9.2}$$

and S is the ratio between the strength of the transistors in the middle leg to the strength of the

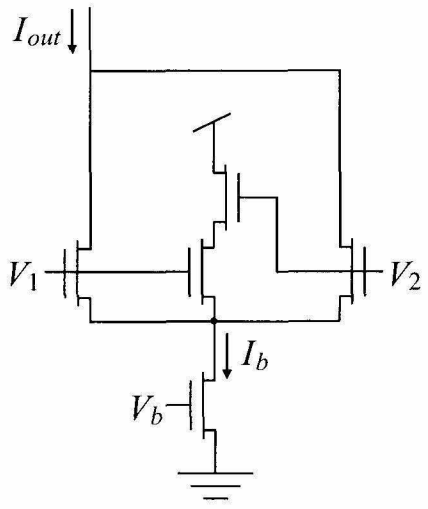

Figure 9.1: Antibump circuit. When V_1 and V_2 are similar, most of the bias current I_b flows through the middle leg. When V_1 and V_2 are very different, most of the bias current will flow through either the left or right outer leg, increasing the output current I_{out} .

transistors in the outer leg:

$$S = \frac{(W/L)_{middle}}{(W/L)_{outer}} \tag{9.3}$$

where W/L is the width-to-length ratio of each MOS transistor.

Figure 9.2 shows the measured response of such a circuit fabricated in a 1.2 μm process with $S = 16$. The theoretical fit indicates a functional value of $S = 50$. This is likely due to the narrower effective channel width also observed by Delbrück in his original experiments with this circuit (Delbrück, 1993a).

If the differential voltage to the antibump circuit comes from two adjacent photoreceptors, the output will approximate the absolute value of the spatial derivative along an axis. Since the output of these circuits is a current, it is trivial to sum the signals over the entire EMD array, and thus generate a monotonic measure of image contrast.

We placed one of these circuits between every horizontal pair of photoreceptors in a 24 × 5 array and summed the output currents on one wire. When the image of a featureless white piece of paper was focused onto the array, the summed output current was 0.7 nA. This nonzero current is due to both photoreceptor mismatch and the nonzero output of each antibump circuit for equal input voltages.

The current increased to 1.1 nA when the experimenter's face was presented to array. High-contrast vertical stripes increased the output to 2.2 nA, while horizontal stripes produced the same output as the white paper (i.e., 0.7 nA). A true 2-D implementation would include vertically-oriented antibump circuits to measure spatial derivatives along this axis.

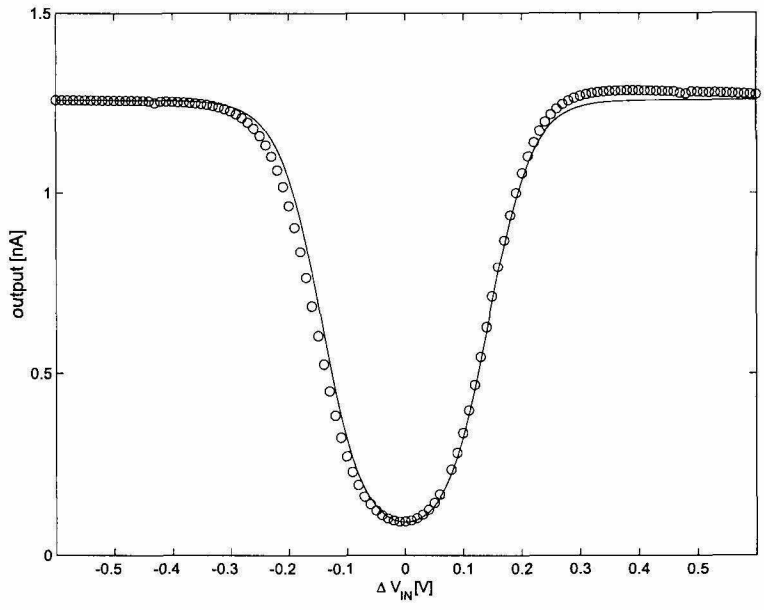

Figure 9.2: Antibump circuit data. The antibump circuit takes a differential voltage as an input and so performs an implicit subtraction. For voltage differentials less than 200 mV, the circuit approximates an absolute value function. Theoretical fit to Equation 9.1 is plotted for $S = 50$.

Clearly, this signal could be used at the system level to decide whether a zero output from the EMD array was caused by zero optic flow or zero contrast in the image.

9.2 Coriolis-Force Gyroscopes and Their Limitations

As we discussed in Section 2.2, flies rely on Coriolis-force sensors to provide angular velocity information during flight. Many commercially-available gyroscopes today operate on the same principle: A mass is driven into oscillation, and the Coriolis force normal to the axis of oscillation is measured. For a review of small Coriolis-force gyroscopes, see Söderkvist, 1994.

The SNR of a gyroscope is proportional to angular velocity. Typical small Coriolis-force gyroscopes have sensitivities of around 1 deg/s, and operate on 35 mW of power. However, as we decrease power usage in MEMS gyroscopes to the microwatt range, their sensitivities will decrease. Both signal-to-noise ratio and bandwidth are proportional to power dissipation in analog systems (Vittoz, 1994). Visual motion sensors can be designed to detect very slow velocities of optic flow simply by increasing their time constants or decreasing the angular spacing between adjacent photoreceptors. Perhaps wide-field optic flow sensors could fill in information about low angular velocities.

Coriolis-force gyroscopes suffer from a fundamental sensitivity to mechanical asymmetries. The

displacements along the driving axis are so much larger than the displacements along the sensing axis that small mechanical asymmetries can lead to crosstalk between the two axes. This produces *quadrature error*, which effectively introduces large dc offsets to output signal (Maenaka et al., 1996). This *zero rate output* (ZRO) can be as high as 500 deg/s in some designs (Söderkvist, 1994). Temperature variation or device aging may shift this offset, introducing ultra-low frequency noise. Most gyroscope outputs are highpass filtered to adapt out this dc error and low frequency noise. However, this scheme allows slow changes in angular velocity to go undetected. If offsets in a visual motion sensor are sufficiently low, optic flow could be used to fill in this low frequency velocity information.

Chapter 10 Conclusions

In this dissertation, we have presented VLSI circuits modeled after neural mechanisms in the visual system of the fly. We have characterized these sensors and compared them against their counterparts in biology. Our biologically-inspired sensors operate at sub-milliwatt power levels—the lowest of any VLSI motion sensor we are aware of—and are capable of perceptual discriminations in situations where the signal is weaker than the noise. While testing our circuits, we developed new methodologies for evaluating the robustness of visual motion sensors in natural-image conditions. We investigated several circuit architectures to improve the robustness of these motion detectors by incorporating computational strategies used by the fly. By increasing our system's biophysical fidelity, we increased its performance—an encouraging example of how biologically-inspired approaches to engineering can yield valuable results.

Have we learned anything new about the fly in the process? We have not generated any testable hypotheses, so the pessimistic answer is "no." However, our reverse engineering approach has yielded system-level insights into the information processing strategies chosen by evolution. We hypothesize that the Reichardt motion detector architecture was selected due to its ability to extract motion information from noisy stimuli. Direct comparison of silicon sensors and real flies lend further support to biological models by proving that the underlying algorithms work robustly even when implemented in mismatched, imprecise hardware.

Is neuromorphic analog VLSI a useful approach to engineering? Perhaps for applications where micropower sensors are needed, and the environment cannot be controlled. Unfortunately, in many cases the power savings earned by using subthreshold analog CMOS circuits is rendered insignificant in any system using an actuator. For example, we investigated building a small glider with an on-board motion sensor to provide an optomotor steering response. The VLSI sensor consumed less than one milliwatt, including on-chip voltage regulation and motor control signal generation. We obtained a small servomotor with a mass of 2.4 g to move the glider's rudder. This motor consumed between 300 and 1000 mW of power when active, and generated tremendous levels of noise on the power supply. The power dissipation of the intelligent sensor was dwarfed by the motor.

Custom digital implementations of neural systems could provide power levels in the milliwatt range and possibly require less design time than analog implementations. Low power digital design and efficient architectures using bit serial arithmetic show promise for real-time neural system implementations. While today's general-purpose microprocessors may be ill-suited for implementing massively parallel, neuromorphic systems, this should not rule out all digital computation as a

useful substrate for biologically-inspired designs. Indeed, biological systems have chosen a hybrid analog/digital scheme (dendritic integration uses continuous values; spiking mechanisms use discrete values) for processing information with extremely high energy efficiency.

Engineering stands to benefit greatly by learning from biological information processing strategies. Much of this progress must wait for more knowledge of how neural systems operate, but many interesting and useful models have already been developed and are awaiting implementation.

Bibliography

[Adelson and Bergen, 1985] Adelson, E. H. and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2:284–299.

[Andreou et al., 1991] Andreou, A. G., Strohbehn, K., and Jenkins, R. E. (1991). Silicon retina for motion computation. In *Proceedings of the 1991 IEEE International Symposium on Circuits and Systems*, pages 1373–1376.

[Autrum, 1958] Autrum, H. (1958). Electrophysiological analysis of the visual systems in insects. *Experimental Cell Research Suppliment*, 5:426–439.

[Barlow and Levick, 1965] Barlow, H. B. and Levick, W. R. (1965). The mechanism of directionally selective units in the rabbit's retina. *Journal of Physiology*, 178:447–504.

[Barrows, 1998] Barrows, G. L. (1998). Feature tracking linear optic flow sensor for 2-D optic flow measurement. In *Proceedings of the 1998 International Conference on Automation, Robotics, and Computer Vision (ICARCV'98), Singapore*, pages 732–736.

[Benson and Delbrück, 1991] Benson, R. G. and Delbrück, T. (1991). Direction selective silicon retina that uses null inhibition. In Lippman, R., Moody, J., and Touretzky, D. S., editors, *Advances in Neural Information Processing Systems 3*, pages 756–763. Morgan Kaufman, San Mateo, CA.

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

[Borst and Bahde, 1988] Borst, A. and Bahde, S. (1988). Spatio-temporal integration of motion. *Naturwissenschaften*, 75:265–267.

[Borst and Egelhaaf, 1989] Borst, A. and Egelhaaf, M. (1989). Principles of visual motion detection. *Trends in Neurosciences*, 12:297–306.

[Borst and Egelhaaf, 1990] Borst, A. and Egelhaaf, M. (1990). Direction selectivity of blowfly motion-sensitive neurons is computed in a two-stage process. *Proceedings of the National Academy of Sciences USA*, 87:9363–9367.

[Borst et al., 1995] Borst, A., Egelhaaf, M., and Haag, J. (1995). Mechanisms of dendritic integration underlying gain control in fly motion-sensitive interneurons. *Journal of Computational Neuroscience*, 2:5–18.

[Borst et al., 1993] Borst, A., Egelhaaf, M., and Seung, H. S. (1993). Two-dimensional motion perception in flies. *Neural Computation*, 5:856–868.

[Borst and Haag, 1996] Borst, A. and Haag, J. (1996). The intrinsic electrophysiological characteristics of fly lobula plate tangential cells: I. Passive membrane properies. *Journal of Computational Neuroscience*, 3:313–336.

[Borst and Theunissen, 1999] Borst, A. and Theunissen, F. E. (1999). Information theory and neural coding. *Nature Neuroscience*, 2:947–957.

[Chan et al., 1998] Chan, W. P., Prete, F., and Dickinson, M. H. (1998). Visual input to the efferent control system of a fly's "gyroscope". *Science*, 280:289–292.

[Clifford et al., 1997] Clifford, C. W. G., Ibbotson, M. R., and Langley, K. (1997). An adaptive Reichardt detector model of motion adaptation in insects and mammals. *Visual Neuroscience*, 14:741–749.

[Collett, 1980] Collett, T. S. (1980). Some operating rules for the optomotor system of a hoverfly during voluntary flight. *Journal of Comparative Physiology*, 138:271–282.

[de Ruyter van Steveninck et al., 1986] de Ruyter van Steveninck, R. R., Zaagman, W. H., and Mastebroek, H. A. K. (1986). Adaptation of transient responses of a movement-sensitive neuron in the visual system of the blowfly *Calliphora erythrocephala*. *Biological Cybernetics*, 54:223–236.

[Delbrück, 1993a] Delbrück, T. (1993a). Bump circuits for computing similarity and dissimilarity of analog voltages. CNS Memo 26, California Institute of Technology.

[Delbrück, 1993b] Delbrück, T. (1993b). Silicon retina with correlation-based, velocity-tuned pixels. *IEEE Transactions on Neural Networks*, 4:529–541.

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

[Deutschmann and Koch, 1998] Deutschmann, R. A. and Koch, C. (1998). Compact real-time 2-D gradient based analog VLSI motion sensor. In *Proceedings of the International Conference on Advanced Focal Plane Arrays and Electronic Cameras, Zürich.*

[DeVoe, 1980] DeVoe, R. D. (1980). Movement sensitivities of cells in the fly's medulla. *Journal of Comparative Physiology A*, 138:93–119.

[DeVoe and Ockleford, 1976] DeVoe, R. D. and Ockleford, E. M. (1976). Intracellular responses from cells in the medulla of the fly, *Calliphora erythrocephala*. *Biological Cybernetics*, 23:13–24.

[Dong and Atick, 1995] Dong, D. W. and Atick, J. J. (1995). Statistics of natural time-varying images. *Network: Computation in Neural Systems*, 6:345–358.

[Duchon et al., 1998] Duchon, A. P., Kaelbling, L. P., and Warren, W. H. (1998). Ecological robotics. *Adaptive Behavior*, 6:473–507.

[Egelhaaf, 1985] 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. *Biological Cybernetics*, 52:195–209.

[Egelhaaf, 1987] Egelhaaf, M. (1987). Dynamic properties of two control-systems underlying visually guided turning in house-flies. *Journal of Comparative Physiology A*, 161:777–783.

[Egelhaaf and Borst, 1989] Egelhaaf, M. and Borst, A. (1989). Transient and steady-state response properties of movement detectors. *Journal of the Optical Society of America A*, 6:116–127.

[Egelhaaf and Borst, 1992] Egelhaaf, M. and Borst, A. (1992). Are there separate ON and OFF channels in fly motion vision? *Visual Neuroscience*, 8:151–164.

[Egelhaaf and Borst, 1993] Egelhaaf, M. and Borst, A. (1993). A look into the cockpit of the fly: visual orientation, algorithms, and identified neurons. *Journal of Neuroscience*, 13:4563–4574.

[Egelhaaf et al., 1989] 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. *Journal of the Optical Society of America A*, 6:1070–1087.

[Egelhaaf et al., 1988] Egelhaaf, M., Hausen, K., Reichardt, W., and Wehrhahn, C. (1988). Visual course control in flies relies on neuronal computation of object and background motion. *Trends in Neurosciences*, 11:351–358.

[Etienne-Cummings et al., 1997] Etienne-Cummings, R., Van der Spiegel, J., and Mueller, P. (1997). A focal plane visual motion measurement sensor. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 44:55–66.

[Field, 1987] Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4:2379–2394.

[Franceschini et al., 1992] Franceschini, N., Pichon, J. M., and Blanes, C. (1992). From insect vision to robot vision. *Philosophical Transactions of the Royal Society of London B*, 337:283–294.

[Franceschini et al., 1989] Franceschini, N., Riehle, A., and le Nestour, A. (1989). Directionally selective motion detection by insect neurons. In Stavenga and Hardie, editors, *Facets of Vision*. Springer-Verlag, Berlin.

[Geiger and Nässel, 1981] Geiger, G. and Nässel, D. R. (1981). Visual orientation behaviour of flies after selective laser beam ablation of interneurones. *Nature*, 293:398–399.

[Geiger and Nässel, 1982] Geiger, G. and Nässel, D. R. (1982). Visual processing of moving single objects and wide-field patterns in flies: behavioural analysis after laser-surgical removal of interneurons. *Biological Cybernetics*, 44:141–149.

[Geiger and Poggio, 1981] Geiger, G. and Poggio, T. (1981). Asymptotic oscillations in the tracking behaviour of the fly *Musca domestica*. *Biological Cybernetics*, 41:197–201.

[Gibson, 1950] Gibson, J. J. (1950). *The Perception of the Visual World*. Houghton-Mifflin, Boston, MA.

[Götz, 1965] Götz, K. (1965). Die optischen Übertragungseigenschaften der Komplexaugen von *Drosophila*. *Kybernetik*, 2:215–221.

[Götz, 1975] Götz, K. G. (1975). The optomotor equilibrium of the *Drosophila* navigation system. *Journal of Comparative Physiology*, 99:187–210.

[Haag and Borst, 1997] Haag, J. and Borst, A. (1997). Encoding of visual motion information and reliability in spiking and graded potential neurons. *Journal of Neuroscience*, 17:4809–4819.

[Haag and Borst, 1998] Haag, J. and Borst, A. (1998). Active membrane properties and signal encoding in graded potential neurons. *Journal of Neuroscience*, 18:7972–7986.

[Harrison et al., 1998] Harrison, R. R., Hasler, P., and Minch, B. A. (1998). Floating-gate CMOS analog memory cell array. In *Proceedings of the 1998 IEEE International Symposium on Circuits and Systems*, volume 2, pages 204–207.

[Harrison and Koch, 1998] Harrison, R. R. and Koch, C. (1998). An analog VLSI model of the fly elementary motion detector. In Jordan, M. I., Kearns, M. J., and Solla, S. A., editors, *Advances in Neural Information Processing Systems 10*, pages 880–886. MIT Press, Cambridge, MA.

[Harrison and Koch, 1999a] Harrison, R. R. and Koch, C. (1999a). An analog VLSI implementation of a visual interneuron: enhanced sensory processing through biophysical modeling. *International Journal of Neural Systems*, 9:391–395.

[Harrison and Koch, 1999b] Harrison, R. R. and Koch, C. (1999b). A robust analog VLSI motion sensor. *Autonomous Robots*, 7:211–224.

[Harrison and Koch, 2000a] Harrison, R. R. and Koch, C. (2000a). A robust analog VLSI Reichardt motion sensor. *Analog Integrated Circuits and Signal Processing (In Press)*.

[Harrison and Koch, 2000b] Harrison, R. R. and Koch, C. (2000b). A silicon implementation of the fly's optomotor control system. *Neural Computation (In Press)*.

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

[Hausen, 1982a] Hausen, K. (1982a). Motion sensitive interneurons in the optomotor system of the fly I. The horizontal cells: structure and signals. *Biological Cybernetics*, 45:143–156.

[Hausen, 1982b] Hausen, K. (1982b). Motion sensitive interneurons in the optomotor system of the fly II. The horizontal cells: receptive field organization and response characteristics. *Biological Cybernetics*, 46:67–79.

[Hausen, 1984] Hausen, K. (1984). The lobula-complex of the fly: structure, function, and significance in behaviour. In Ali, M. A., editor, *Photoreception and Vision in Invertibrates*, pages 523–559. Plenum, New York, NY.

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

[Hausen and Wehrhahn, 1990] Hausen, K. and Wehrhahn, C. (1990). Neural circuits mediating visual flight control in flies: II. Separation of two control-systems by microsurgical brain-lesions. *Journal of Neuroscience*, 10:351–360.

[Heisenberg and Buchner, 1977] Heisenberg, M. and Buchner, E. (1977). The rôle of retinula cell types in visual behavior of *Drosophila melanogaster*. *Journal of Comparative Physiology*, 117:127–162.

[Hengstenberg, 1982] Hengstenberg, R. (1982). Common visual reponse properties of giant vertical cells in the lobula plate of the blowfly *Calliphora*. *Journal of Comparative Physiology A*, 149:179–193.

[Higgins et al., 1999] Higgins, C. M., Deutschmann, R. A., and Koch, C. (1999). Pulse-based 2-D motion sensors. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 46:677–687.

[Himmelbauer et al., 1996] Himmelbauer, W., Furth, P. M., Pouliquen, P. O., and Andreou, A. G. (1996). Log-domain filters in subthreshold MOS. Technical report, Johns Hopkins University.

[Holmqvist and Srinivasan, 1991] Holmqvist, M. H. and Srinivasan, M. V. (1991). A visually evoked escape response of the housefly. *Journal of Comparative Physiology A*, 169:451–459.

[Horiuchi et al., 1991] Horiuchi, T., Lazzaro, J., Moore, A., and Koch, C. (1991). A delay-line based motion detection chip. In R. Lippman, e. a., editor, *Advances in Neural Information Processing Systems 3*, pages 406–412. Morgan Kaufman, San Mateo, CA.

[Huber, 1997] Huber, S. A. (1997). *Studies of the Visual Orientation Behavior in Flies Using the Artificial Life Approach.* PhD thesis, Universität Tübingen, Germany.

[Jiang and Wu, 1999] Jiang, H.-C. and Wu, C.-Y. (1999). A 2-D velocity- and direction-selective sensor with BJT-based silicon retina and temporal zero-crossing detector. *IEEE Journal of Solid State Circuits*, 34:241–247.

[Kimmerle et al., 1996] Kimmerle, B., Egelhaaf, M., and Srinivasan, M. V. (1996). Object detection by relative motion in freely flying flies. *Naturwissenschaften*, 83:380–381.

[Kimmerle et al., 1997] Kimmerle, B., Warzecha, A.-K., and Egelhaaf, M. (1997). Object detection in the fly during simulated translatory flight. *Journal of Comparative Physiology A*, 181:247–255.

[Klaus and Horn, 1986] Klaus, B. and Horn, P. (1986). *Robot Vision.* MIT Press, Cambridge, MA.

[Kramer, 1996] Kramer, J. (1996). Compact integrated motion sensor with three-pixel interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18:455–460.

[Kramer et al., 1997] Kramer, J., Sarpeshkar, R., and Koch, C. (1997). Pulse-based analog VLSI velocity sensors. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 44:86–101.

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

[Land, 1997] Land, M. F. (1997). Visual acuity in insects. *Annual Review of Entomology*, 42:147–177.

[Laughlin, 1994] Laughlin, S. B. (1994). Matching coding, circuits, cells, and molecules to signals – general principles of retinal design in the fly's eye. *Progress in Ret. Eye Research*, 13:165–196.

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

[Liu, 1997] Liu, S.-C. (1997). *Neuromorphic Models of Visual and Motion Processing in the Fly Visual System.* PhD thesis, California Institute of Technology.

[Maenaka et al., 1996] Maenaka, K., Fujita, T., Yonishi, Y., and Maeda, M. (1996). Analysis of a highly sensitive silicon gyroscope with cantilever beam as vibrating mass. *Sensors and Actuators A*, 54:568–573.

[Marr, 1982] Marr, D. (1982). *Vision.* W. H. Freeman & Co., New York.

[Matthies et al., 1995] 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, 1989] Mead, C. (1989). *Analog VLSI and Neural Systems.* Addison-Wesley, Menlow Park, CA.

[Minch et al., 1996a] Minch, B. A., Diorio, C., Hasler, P., and Mead, C. (1996a). The matching of small capacitors for analog VLSI. In *Proceedings of the 1996 IEEE International Symposium on Circuits and Systems*, volume 1, pages 239–241.

[Minch et al., 1996b] Minch, B. A., Diorio, C., Hasler, P., and Mead, C. (1996b). Translinear circuits using subthreshold floating-gate MOS transistors. *Analog Integrated Circuits and Signal Processing*, 9:167–179.

[Moini et al., 1997] Moini, A., Bouzerdoum, A., Eshraghian, K., Yakovleff, A., Nguyen, X. T., Blanksby, A., Beare, R., Abbott, D., and Bogner, R. E. (1997). An insect vision-based motion detection chip. *IEEE Journal of Solid-State Circuits*, 32:279–284.

[Nalbach, 1993] Nalbach, G. (1993). The halteres of the blowfly *Calliphora* I. Kinematics and dynamics. *Journal of Comparative Physiology A*, 173:293–300.

[O'Carroll et al., 1996] 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.

[Pavasović et al., 1991] Pavasović, A., Andreou, A. G., and Westgate, C. R. (1991). Characterization of CMOS process variations by measuring subthreshold current. In Ruud, C. O. et al., editors, *Nondestructive Characterization of Materials IV*, pages 363–370. Plenum, New York, NY.

[Pichon et al., 1989] Pichon, J.-M., Blanes, C., and Franceschini, N. (1989). Visual guidance of a mobile robot equipped with a network of self-motion sensors. In *SPIE Mobile Robots IV*, volume 1195, pages 44–53.

[Poggio and Reichardt, 1981] Poggio, T. and Reichardt, W. (1981). Visual fixation and tracking by flies: mathematical properties of simple control systems. *Biological Cybernetics*, 40:101–112.

[Reichardt and Egelhaaf, 1988] Reichardt, W. and Egelhaaf, M. (1988). Properties of individual movement detectors as derived from behavioural experiments on the visual system of the fly. *Biological Cybernetics*, 58:287–294.

[Reichardt and Poggio, 1976] Reichardt, W. and Poggio, T. (1976). Visual control of orientation behaviour in the fly. *Quarterly Reviews of Biophysics*, 9:311–375.

[Ruderman and Bialek, 1994] Ruderman, D. L. and Bialek, W. (1994). Statistics of natural images: scaling in the woods. *Physical Review Letters*, 73:814–817.

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

[Schuppe and Hengstenberg, 1993] Schuppe, H. and Hengstenberg, R. (1993). Optical properties of the ocelli of *Calliphora erythrocephala* and their role in the dorsal light response. *Journal of Comparative Physiology A*, 173:143–149.

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

[Single et al., 1997] Single, S., Haag, J., and Borst, A. (1997). Dendritic computation of direction selectivity and gain control in visual interneurons. *Journal of Neuroscience*, 17:6023–6030.

[Söderqvist, 1994] Söderqvist, J. (1994). Micromachined gyroscopes. *Sensors and Actuators A*, 43:65–71.

[Srinivasan et al., 1997] Srinivasan, M. V., Chahl, J. S., and Zhang, S. W. (1997). Robot navigation by visual dead-reckoning: inspiration from insects. *International Journal of Pattern Recognition and Artificial Intelligence*, 11:35–47.

[Srinivasan et al., 1999] Srinivasan, M. V., Poteser, M., and Kral, K. (1999). Motion detection in insect orientation and navigation. *Vision Research*, 39:2749–2766.

[Srinivasan et al., 1993] Srinivasan, M. V., Zhang, S. W., and Chandrashekara, K. (1993). Evidence for two distinct movement-detecting mechanisms in insect vision. *Naturwissenschaften*, 80:38–41.

[Stange, 1981] Stange, G. (1981). The ocellar component of flight equilibrium control in dragonflies. *Journal of Comparative Physiology*, 141:335–347.

[Strausfeld, 1976] Strausfeld, N. J. (1976). *Atlas of an Insect Brain.* Springer, Berlin.

[Tanner and Mead, 1986] Tanner, J. and Mead, C. (1986). An integrated analog optical motion sensor. In *VLSI Signal Processing II*, pages 59–76. IEEE Press, New York, NY.

[van Hateren, 1990] van Hateren, J. H. (1990). Directional tuning curves, elementary movement detectors, and the estimation of the direction of visual movement. *Vision Research*, 30:603–614.

[van Santen and Sperling, 1985] van Santen, J. P. H. and Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2:300–321.

[Vittoz, 1994] Vittoz, E. A. (1994). Low-power low-voltage limitations and prospects in analog design. In van de Plassche, R. S., Sansen, W. M. C., and Huijsing, J. H., editors, *Analog Circuit Design*. Kluwer Academic Publishers, Dordrecht, The Netherlands.

[Wagner, 1986] Wagner, H. (1986). Flight performance and visual control of flight of the free-flying housefly *Musca domestica* L. II. Pursuit of targets. *Philosophical Transactions of the Royal Society of London B*, 312:553–579.

[Wandell, 1995] Wandell, B. A. (1995). *Foundations of Vision*. Sinauer Associates, Sunderland, MA.

[Warzecha and Egelhaaf, 1996] Warzecha, A.-K. and Egelhaaf, M. (1996). Instrinsic properties of biological motion detectors prevent the optomotor control system from getting unstable. *Philosophical Transactions of the Royal Society of London B*, 351:1579–1591.

[Watson and Ahumada, 1985] Watson, A. B. and Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, 2:322–342.

[Watts et al., 1992] Watts, L., Kerns, D. A., Lyon, R. F., and Mead, C. A. (1992). Improved implementation of the silicon cochlea. *IEEE Journal of Solid-State Circuits*, 27:692–700.

[Weckström et al., 1992] Weckström, M., Juusola, M., and Laughlin, S. B. (1992). Presynaptic enhancement of signal transients in photoreceptor terminals in the compound eye. *Proceedings of the Royal Society of London B*, 250:83–89.

[Wehner, 1987] Wehner, R. (1987). 'Matched filters' – neural models of the external world. *Journal of Comparative Physiology A*, 161:511–531.

[Wolf et al., 1980] Wolf, R., Gebhard, B., Gademann, R., and Heisenberg, M. (1980). Polarization sensitivity of course control in *Drosophila melanogaster*. *Journal of Comparative Physiology*, 139:177–191.

[Wolf and Heisenberg, 1990] Wolf, R. and Heisenberg, M. (1990). Visual control of straight flight in *Drosophila melanogaster*. *Journal of Comparative Physiology A*, 167:269–283.

[Zanker, 1990] Zanker, J. M. (1990). On the directional sensitivity of motion detectors. *Biological Cybernetics*, 62:177–183.

[Zanker et al., 1999] Zanker, J. M., Srinivasan, M. V., and Egelhaaf, M. (1999). Speed tuning in elementary motion detectors of the correlation type. *Biological Cybernetics*, 80:109–116.