Analog VLSI-Based, Neuromorphic Sensorimotor Systems: Modeling the Primate Oculomotor System

Thesis by

Timothy K. Horiuchi

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1997

(Submitted February 28, 1997)

Acknowledgements

First and foremost, I would like to thank my parents, Harvey and Akiko Horiuchi, and my sister Sophie Horiuchi for all their support, financial and emotional, throughout the many years I have been in school. I attribute my interest in analog electronics to my father who often brought home discarded electronics from work.

To Prof. Christof Koch I also give many thanks for giving me the freedom and encouragement to pursue the project of my dreams, for the tolerance he showed towards my messy, toy-laden office, and for being an inspiring example of enthusiasm and dedication to science. At a time when I was most uncertain about my career in science, he took me into the lab and offered me a chance.

I thank my thesis committee: Prof. Mark Konishi, Prof. Pietro Perona, Prof. Steve DeWeerth, and Prof. Rodney Goodman, for reading my dissertation and providing a supportive environment over the years. I also thank Prof. Gilles Laurent, Prof. Joel Burdick, and Prof. Carver Mead for serving on my candidacy committee. My work is, of course, based on much of the pioneering work of Prof. Mead's laboratory and without his support, very little of this work would have come to fruition. I thank the Office of Naval Research (ONR), the National Institute of Mental Health (NIMH), and the National Science Foundation's Engineering Research Center for Neuromorphic Systems Engineering at Caltech, for funding my research these many years.

I have had the pleasure to work with many talented and inspiring individuals at Caltech (too many to list!) who have made my graduate life rich with experiences. Of these, I would especially like to thank: John Harris who first got me interested in analog VLSI design and has provided a ton of advice throughout the years, Tobi Delbrück for much-needed circuit advice and for encouraging me at a time when I was really discouraged, Buster Boahen for taking the time to teach me about current-mode circuits and the address-event protocol, John Lazzaro for teaching me about owl audition and chip layout, and Paul Hasler for helping me get started with floating-gates. A special thanks goes to Misha Mahowald who inspired us all with her passion for science and her vision of where analog VLSI could make a difference. Many thanks also go to my good friend and collaborator, Tonia Morris, for her support through many rough times. I have enjoyed many nights this last year discussing and designing attentional circuits over the phone. I would also like to thank Brooks Bishofberger for his fastidiousness about documentation and for keeping the lab alive with wackiness and creativity, Candi Hochenedel for her patience and for always being there to listen to my musings, and Kate Finnigan in the Registrar's Office who has always shown heroic patience with my paperwork avoidance behavior.

I have come to know so many people in the Koch Lab, past and present, that I could not list them all. There are a few people with whom I have spent a lot of time that I would like to thank for their part in making the lab an inviting place to work, eat, sleep, and play: Matthew Avalos, Andrew Hsu, Walter Yamada, Dan Kammen, Ernst Niebur, Florentin Wörgötter, Wyeth Bair, Frank Perez, Giacomo Indiveri, Jörg Kramer, Grace Chang, Rainer Deutschmann, Chuck Higgins, Pam Reinagel, Gary Holt, and Reid Harrison.

Thanks also to other members of the Caltech community who have made a truly positive impact on my life here at Caltech: Bruno Olshausen for endless discussions about neurophysiology on long hikes, Doug Kerns for his expert electronics advice and his undying enthusiasm for cool toys, Erik Winfree for his insatiable need to explore and his eye for cool things, Luis Goncalves for his many fun accents, Shih-chii Liu and Tobi Delbrück for their many desert camping invitations, and Kate MacLeod, my favorite neurophysiology TA. Thanks to everyone in the computer graphics group for always making me feel at home and for putting up with the late-night antics .

And finally, I thank my best friend Bena Currin for always believing in me, for giving me the strength to carry on when things looked hopeless, and for releasing the kid in me that I thought I had lost.

Abstract

Using neuromorphic analog VLSI techniques for modeling large neural systems has several advantages over software techniques. By designing massively-parallel analog circuit arrays which are ubiquitous in neural systems, analog VLSI models are extremely fast, particularly when local interactions are important in the computation. While analog VLSI circuits are not as flexible as software methods, the constraints posed by this approach are often very similar to the constraints faced by biological systems. As a result, these constraints can offer many insights into the solutions found by evolution. This dissertation describes a hardware modeling effort to mimic the primate oculomotor system which requires both fast sensory processing and fast motor control. A one-dimensional hardware model of the primate eye has been built which simulates the physical dynamics of the biological system. It is driven by analog VLSI circuits mimicking brainstem and cortical circuits that control eye movements. In this framework, a visually-triggered saccadic system is demonstrated which generates averaging saccades. In addition, an auditory localization system, based on the neural circuits of the barn owl, is used to trigger saccades to acoustic targets in parallel with visual targets. Two different types of learning are also demonstrated on the saccadic system using floating-gate technology allowing the non-volatile storage of analog parameters directly on the chip. Finally, a model of visual attention is used to select and track moving targets against textured backgrounds, driving both saccadic and smooth pursuit eye movements to maintain the image of the target in the center of the field of view. This system represents one of the few efforts in this field to integrate both neuromorphic sensory processing and motor control in a closed-loop fashion.

Contents

List of Figures

Chapter 1 Introduction

The dream of building intelligent machines has been with us since the dawn of technology. The pursuit of understanding how the brain works has been with us nearly as long. Every new technological age has tried to view the brain in the dominant technology of the day, from hydraulics, to gears, to telephones, to digital computers. Perhaps for the first time in history, however, we are beginning to understand the operation of many different neural circuits and their functional significance in behavior. As such, we are less dependent on the dominant technology as a metaphor. It is our hope that the rapidly expanding knowledge of neuroscience will drive the creation of an entirely new technology based on computational principles utilized by the brain. Digital computers have excelled in the problems of mathematical logic, database manipulation, and repetitive number crunching with extremely high precision; problems difficult to perform by biological systems. On the other hand, visual perception, navigation, and sensorimotor control, problems easily solved by biological systems, have not been satisfactorily solved by engineering systems. Notably, the successful computing architectures in these two problem domains are markedly different.

As more neurobiological details emerge and models become significantly more complicated, computational modeling will become an invaluable tool, providing an interactive test of the many assumptions we put into our models. Using traditional software methods to model neural systems is often difficult because most neural systems are composed of very large numbers of interconnected elements with non-linear characteristics and a wide range of time-constants. Their mathematical behavior can rarely be solved analytically and simulations slow dramatically as the number and coupling of elements increases, especially in cases where capturing the details of fast dynamics is important. While quite a few user-friendly, neural-simulation software packages are now available (e.g. Neuron, Genesis), they are not very appropriate for fast, efficient simulation of very large systems. In particular, modeling sensorimotor systems to understand their interaction with the natural world often requires simulating the real world, which can often be more difficult or time-consuming than simulating the model itself. Modeling some aspects of the world is not *always* intractable; some research groups have begun experimenting with virtual-reality environments developed by others for developing robotic sensorimotor control models. These virtual environments, however, still fall short of the realism needed for certain problem domains to be certain of a model's ability to operate in the real-world.

1.1 Neuromorphic Analog VLSI

In the early 1980s, Prof. Carver Mead at the California Institute of Technology began a research effort to investigate the use of the latest silicon fabrication technology to emulate and understand neurobiological circuits. While parallel, analog computers have been used before to simulate retinal processing and other neural circuits (Fukushima et al., 1970), the significant difference was the use of the latest silicon technology which allowed engineers to cheaply integrate a few million transistors onto a centimeter of silicon. Meanwhile, the rapidly growing body of knowledge in neuroscience was helping to fuel the field of computational neuroscience, which highlighted the need for faster simulation engines.

The approach taken by Mead's group was to model the analog parallel processing of neural circuits by building analog parallel processing circuits on silicon using commercial digital VLSI (very-large-scale integration) processes. By utilizing analog techniques which exploit elementary physical phenomena as computational primitives, the efficiency of computation promises to be up to 100 times more efficient in silicon area and up to 10,000 times less power hungry (Mead, 1990). This approach of designing computational systems which emulate the function and morphology of nervous systems has come to be known as *neuromorphic engineering.* Neuromorphic analog VLSI systems thus utilize the analog VLSI medium to implement these computational architectures.

This approach has incorporated many of the features of neural circuitry such as: processing strategies that are fast and redundant, circuits that are tolerant to noise and component variability, and adaptive behavior using local parameter storage. All of these advantages do not come without cost; there are many constraints in building such systems. These constraints are, however, well matched to those of biological systems and may provide greater insight into the design of neural systems than alternative methods of modeling. This effort has been growing and considerable progress has been made in the development of circuits well-matched to this type of circuit design. Examples include models of: cortical pyramidal neurons, the cochlea, the outer plexiform layer of the retina, auditory localization, long-term sensorimotor adaptation, and the primate oculomotor system (Douglas et al., 1995).

1.1.1 Benefits: Neural Modeling

Perhaps the most common question encountered in this field of research is: "Why use analog VLSI for modeling when more flexible software approaches can simulate anything?" The two main arguments for utilizing analog VLSI for modeling are its speed and the potential benefits of working within the design constraints.

The human brain contains on the order of 10^{11} neurons; while most simulations do not yet attempt to simulate whole brains, even of the common house-fly which has a modest 340,000 neurons

(Strausfeld, 1976), eventually such simulations will be desirable. Nearly all neural models are fine-grained parallel-processing arrays and implementing such models on a serial machine can result in very low simulation speeds. Slow simulations are difficult to work with and difficult to gain any intuition from. Sensorimotor systems are particularly difficult to investigate since the interaction with the real world must either be adequately simulated inside the computer, or the simulation must run fast enough to interact with the time-constants of the real world. Typically, a sufficiently realistic simulation of the real world is impractical. Spike-based circuit modeling can be particularly slow since large, fast swings in voltage are frequent. Mixing widely disparate time scales within the same simulation is even slower (e.g. learning in spiking networks). Provided the analog VLSI circuitry can provide the proper level of detail and is configurable for the types of models under investigation, the neuromorphic analog VLSI models can provide the speed desirable for large-scale simulations. In addition, the growth of a hardware-based simulation does not affect the speed of operation and thus encourages the implementation of detail wherever necessary.

While the speed issue is perhaps the most compelling argument, there is another interesting argument: similarity in constraints. The constraints for designs in analog VLSI and biological neurons are similar and force the modeler to consider and solve similar problems. Although software modelers can include these constraints into their model, it is often the case that these details are left out in an effort to simplify and speed up the simulation. The design constraints in analog VLSI circuitry are many, however, most of them raise interesting questions about the solutions that biology has employed.

As in biological systems, each component in the fabricated chip, while nominally identical, is not fabricated identically. Differences in transistor parameters can be significant (Pavasovic et al., 1994) and can create large errors if the algorithm does not account for them or is not robust enough to tolerate them. By utilizing redundant representations, such as population encoding or place encoding, or by using adaptation based on performance, many of these offset problems can be alleviated.

All neural systems are adaptive to some degree, either exhibiting habituation, synaptic modification, or morphological change. Analog VLSI devices have also begun to demonstrate adaptation and have recently attained the capability of compact, local, non-volatile parameter storage (Diorio et al., 1995).

Noise in physical systems occurs in every device and often becomes a design issue that affects both neural circuitry and analog VLSI design. Most components in an analog VLSI circuit exhibit thermal noise (Sarpeshkar et al., 1993). While still quite a controversial topic, noise in neural systems is believed to be considerable, although to what extent this precludes the use of precise timing in cortical processing is still hotly debated (Shadlen and Newsome, 1994). Noise is known to arise from sources such as: temperature-sensitive channel kinetics, low synaptic reliability, channel current

noise, extrasensory input activity, and for vision, photon noise (Hessler et al., 1993; Longtin and Hinzer, 1996). Noise is not necessarily detrimental and may play an important role in the detection of subthreshold signals (Stemmler, 1996).

While not utilized in the work described here, the address-event-representation communications protocol (AER) (Sivilotti, 1991; Lazzaro et al., 1993; Mahowald, 1994; Mortara et al., 1995; Boahen, 1996), has been developed for communication between chips where large numbers of units must communicate with each other across chip boundaries. The AER protocol is based on the idea that units communicate *events* (spikes) asynchronously to their downstream partners. By communicating the *address* of the sending-unit on a common output bus, the downstream units decode the address from which they receive events and acknowledge the receipt of the event. With the current generation of circuits, this cycle of event communication is performed fast enough to communicate more than 2 million events per second (Boahen, 1996). For spike-based simulations where the neurons' peak firing rates are about 1 kHz, this protocol could handle about 2000 units at maximum output with some loss of temporal accuracy. Aside from the obvious purpose of enabling high-bandwidth communication between chips, the AER protocol enforces a powerful constraint on the method of communications, namely, events. While the actual technique for communication involves more traditional high-speed digital techniques, this system nominally delivers spikes asynchronously with minimal temporal jitter. By requiring long-distance communication to travel via spikes, there are now real advantages to sparse representations and temporal encoding of information.

Finally, wiring is a serious issue in neural circuits, occupying a large volume in the brain. Long distance wiring is expensive metabolically and requires significant volume. Brains tend to minimize long-distance communication and emphasize local connectivity. Wiring is also expensive in VLSI and local connectivity is thus favored.

There are, of course, some constraints in analog VLSI which are not found in the biological substrate (e.g. ability to grow and extend processes) and some constraints in the biological substrate which are not found in analog VLSI (e.g. low-resistance wiring). By understanding these differences and utilizing them carefully, it is possible to maintain the relevance of these circuits to biological modeling.

1.1.2 Benefits: Engineering

Aside from raw computational speed, the energy per unit of computation has become a critical issue for engineering systems, driving technological advances in low-power techniques for portable electronics. While most of the portable devices in the market today are small versions of the traditional digital computer, efficient analog solutions which are better-matched to the problems of real-world interface are still being sought.

Very few of these commercial devices currently attempt to solve the perceptual and sensorimotor

tasks that portable biological computers commonly perform (e.g., rats, owls, and penguins). This is primarily due to the lack of truly successful algorithms and the lack of computational resources to make them practical. The neuromorphic analog VLSI approach attempts to match the computational architecture with the computational function, which leads to a considerably more efficient implementation. While sensory perception and sensorimotor control are presently not part of the repetoire of commercial products, they may yet be. Because this research is currently carried out on commercial fabrication processes, the final product of this research is already close to a marketable implementation.

From a mobile robotics standpoint, the development of analog VLSI sensory processing chips represents a tremendous advance, moving from large, bulky, power-hungry sensory systems to cheaper, smaller, and lower-power replacements. While the ultimate goal might be to completely control a robot with neuromorphic analog VLSI processors, even modest advances will significantly contribute to mobile robotics.

Another important engineering aspect of neuromorphic analog VLSI research involves adaptive behavior and the development of self-calibrating components as part of an integrated system. At the circuit level, adaptation can be used to remove the effects of component offsets; at the system level, adaptation can be used to modify behavior based on experience. While one motivation to have adaptation in a system is that components age, drift, and become damaged, perhaps a more immediate concern is that the optimal calibration of a system or component that is intended to work in the real-world is often best calibrated in the real-world, and not on the lab bench. This is particularly important in mobile robotics where the environment changes frequently as the robot moves. In addition to basic behavioral adaptivity, the robustness of redundant, analog processing algorithms combined with smart adaptation techniques may someday allow circuits to learn how to bypass fabrication defects, increasing yield and facilitating successful wafer-scale circuits.

1.2 Oculomotor Systems

In the neurophysiological literature most neural modeling efforts have focused on the explanation of the phenomena of small, manageable systems. The simulation of the larger system in which these models operate are typically "beyond the scope" of the work. This dissertation focuses on building a large-scale, systems-level model of the primate oculomotor system with the goal of uniting the many neural models into a functioning whole. In particular, active vision is a realm of research where high-speed visual processing is important. By working with analog VLSI techniques, the implementation will be allowed to grow in complexity and scope without sacrificing the speed necessary for interaction with the natural world.

The orientation response in most animals can be triggered by a sudden change in the environment

such as a sharp sound, a sudden movement, or an unexpected tactile sensation. The movement takes on many different forms depending on the particular animal and situation: from body turns to head turns to eye or pinna movements. What they all have in common is the goal of optimizing their sensory inputs by active positioning of their eyes or ears. Studies of active positioning of cameras in machine vision have shown that this can significantly reduce the computational cost of scene analysis as well as reduce the need for huge, high-resolution sensors (Ballard 1991). On an even faster time scale, in some animals, this same task is performed without movement by attentional processes, enhancing only a selected portion of the world, constituting an internalized form of orientation. In both bottom-up, reflexive movements and in top-down, volitional movements, orientation is intimately tied to attention.

The study of attention may prove to be extremely fruitful in our search for a basic understanding of the brain since it appears to lie on the interface between bottom-up, reflexive, sensorimotor control, which has been extensively studied via neurophysiology, and top-down, intentional control of perception, which has been studied primarily via psychophysics. In addition, attention extends beyond vision to all of our sensory modalities, tying the neural activity of different sensory maps into a single, multi-modal, perceptual experience.

While we typically think of brains as quintessential examples of massively parallel information processing, there are many areas within these systems where the sheer volume of sensory information is overwhelming for a system with limited computational resources and limited time in which to react. In the human visual system, it is estimated that there is between $10^8 - 10^9$ bits per second of information flowing down the optic nerve – far more than what the brain is capable of processing and integrating into our conscious experience. The strategy that many animals have adopted for this problem is one of spatially-focused, serial processing, moving quickly from one part of the image to the next, fusing together our perception of the world over time. While our fovea limits our high-resolution vision to the direction of gaze, visual attention further limits what we actively scrutinize. The metaphor most frequently used to describe visual attention is a *spotlight* which enhances processing of information from a subregion of the visual field. The information within the subregion is also thought to be passed on to higher visual areas for further processing. This attentional spotlight is guided by early sensory processing (bottom-up) and by cognitive (top-down) influences.

This strategy of spatially-focused, serial processing is seen in two forms in the primate: *covert* and *overt* attentional shifts. Covert attentional shifting refers to the movement of the attentional spotlight around the field of view in the absence of eye movements. Overt attentional shifting refers to eye movements, shifting the fovea to center and view objects with higher spatial resolution. Given both methods of extracting visual information about our environment, how we decide when and where to saccade next is still an open question. It is the interplay of these two systems which

is the focus of this work.

1.3 Eye Movements: Behavior and Metrics

Although our eye has a field of view of about 170° (Dusenbery, 1992), we see best in the central 1°, or *fovea*, where the density of photoreceptors is the greatest. As we move around in the world and as the world moves around us, our view of the world is constantly in motion. Because our photoreceptors have relatively long impulse-response durations (200 msec for cones, 500 msec for rods; Schnapf and Baylor, 1987), it has been argued that our eye movements are concerned with keeping retinal image motions below 2-3 deg/sec at all times in order to allow our photoreceptors to fully respond and preserve our visual acuity (Westheimer and McKee, 1975). When motion is necessary, however, fast eye movements are employed to reduce the duration of blurred vision.

There are generally considered to be five types of eye movements: the vestibulo-ocular reflex (VOR), the opto-kinetic reflex (OKR), saccades, smooth pursuit, and vergence eye movements. The first two, VOR and OKR, are general stabilization eye movements which are activated when the head is moved. Saccades, smooth pursuit, and vergence eye movements are considered to be target acquisition and stabilization movements more often tied to the target's motion. In all of these eye movements, both eyes move at the same time and with the same timecourse. While the first four types are *conjugate* eye movements, meaning that both eyes move in the same direction, vergence eye movements are *disconjugate*, driving the two eyes in opposite directions to compensate for changes in fixation distance.

1.3.1 VOR and OKR

The vestibulo-ocular and opto-kinetic reflexes are, phylogenetically, the oldest eye movements. Their function in the primate is to detect movements of the head and rotate the eyes to compensate for the movement, prevent image blur on the retina (See Figure 1.1). The vestibulo-ocular reflex operates by sensing rotational velocity and acceleration using the vestibular organs in the inner ear. In contrast, the opto-kinetic reflex senses rotational velocity visually. These two reflexes complement each other by operating in different velocity ranges and at different latencies.

The vestibulo-ocular reflex has a short latency of 14 msec and is most effective for large, transient head movements. For stimulus frequencies from 0.1 and 1.0 Hz, the phase of the VOR does not deviate by more than a few degrees from perfect compensation. The VOR is able to accurately drive the eyes to stabilize head turns up to 300 deg/sec (Keller, 1978).

The opto-kinetic reflex has a longer latency of about 80 msec and is most effective for small, sustained head movements. The opto-kinetic reflex is known to have two components: a fast response (early-component OKR) and a slower response (delayed OKR). The fast response drives the eyes

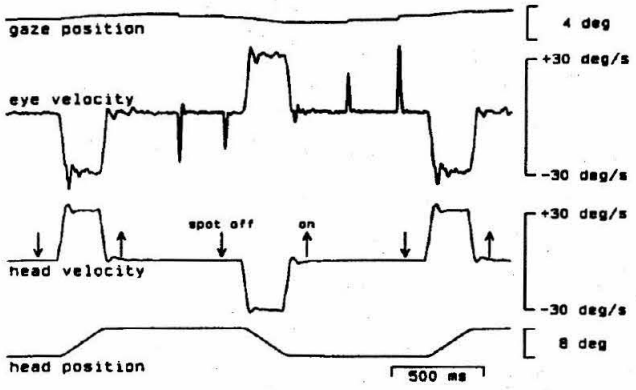

Figure 1.1: Vestibulo-Ocular Reflex: Rapid changes in head velocity were presented in darkness with the eyes and head initially stationary. The monkey fixated a stationary target which was extinguished before the head began to move. The small, brief deflections in the eye velocity are caused by small corrective saccades. (From Lisberger and Pavelko ,1986)

first, allowing the stabilization of motions up to 100 deg/sec. The delayed component builds slowly over seconds and allows the stabilization of much higher speeds.

1.3.2 Smooth Pursuit

Similar to the opto-kinetic reflex, the smooth pursuit eye movement uses visual motion detection to stabilize the image; unlike the opto-kinetic reflex, motion extraction is performed on a subregion of the field of view. Smooth pursuit allows the primate to track small objects even across patterned backgrounds. Interestingly, smooth pursuit ability, while not strictly limited to primates, is uncommon in other vertebrates. This ability is believed to have evolved to serve the special visual stabilization needs of frontally-eyed animals with a fovea manipulating objects in the world (Land, 1995).

Smooth pursuit eye movements cannot be voluntarily generated in the absence of a moving visual stimulus(Kowler, 1990). In addition to the need for a visual stimulus, smooth pursuit movements are believed to require attentive fixation of the target (Khurana and Kowler, 1987; Tam and Ono, 1994).

Smooth pursuit eye movements occur with a latency of 80-130 msec after the onset of smooth target motion and can attain speeds as high as 180 deg/sec, however, accurate tracking is limited to target speeds below 30 deg/sec. The gain of the smooth pursuit movement, while ideally equal to one, is typically below that value, triggering saccades as the target leaves the vicinity of the fovea. Values often range from 0.6 to 0.95 with the gain decreasing with target speed (Collewijn and Tamminga, 1984). Also, the presence of a textured background, typical for natural situations, produced a 10% reduction of the gain during smooth pursuit (Collewijn and Tamminga, 1984;

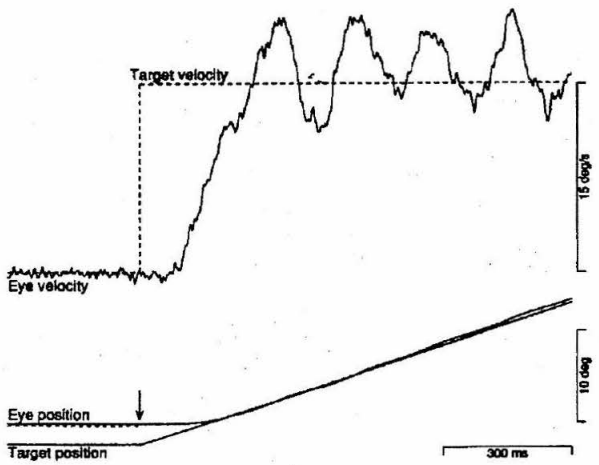

Figure 1.2: Example of monkey smooth pursuit of a constant-velocity target at 15 deg/sec (From Goldreich et al., 1992).

Masson et al., 1995). This interaction of the background is thought to be a result of incomplete suppression of the opto-kinetic reflex.

Due to the visually-driven nature of the smooth pursuit movement, delays in the visual system and the phase lag of the pursuit controller support the possibility of oscillations in the eye velocity during tracking. Figure 1.2 shows an example of this oscillation during pursuit of a constant velocity target. This oscillation is seen in primates and is approximately at 6 Hz (Goldreich et al., 1992).

1.3.3 Saccades

Saccades are rapid, ballistic eye movements which can reach peak velocities of 600 deg/sec and last between 25 and 200 msec dependent upon the saccade amplitude (For examples, see Figure 1.3). Saccades are the only conjugate eye movement that humans can generate as voluntary acts (Becker, 1989). We are able to trigger saccades to visual targets, auditory targets, as well as memorized targets. While saccades are fast, they have a relatively long latency, requiring on average 180-220 msec from the onset of a visual trigger to the beginning of the observed movement (Becker, 1989).

The peak velocity during the saccade initially increases with saccade amplitude until it begins to saturate for saccades greater than about 15 degs. The saturation is complete for saccades greater than 50 degs. In contrast, the saccade duration shows a linear relationship with amplitude for saccades from a range of 5 to 60 degs (Becker, 1989).

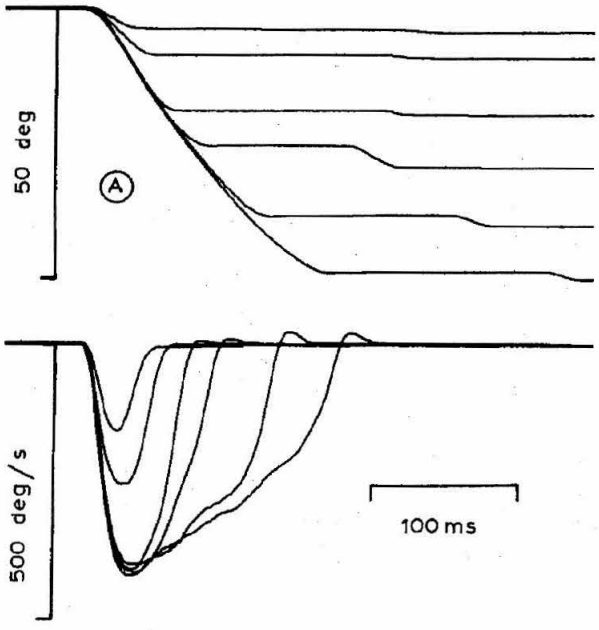

Figure 1.3: Family of saccades of sizes 5°, 10°, 20°, 30°, 40°, 50°. (From Becker, 1989)

1.3.4 Vergence

Vergence eye movements are smooth eye rotations used to maintain fixation of a target with both eyes as a target's distance changes. As the target distance increases, the eyes diverge; when the target distance decreases, the eyes converge.

Vergence eye movements are known to be strongly driven by retinal disparity. Since stereopsis has a limited range of disparity in which fusion can be maintained, the role of vergence is to find and maintain the angle of the eyes to ensure stereo fusion. The range of retinal disparity that drives the vergence system is much larger than that which drives stereopsis. Presumably because there are such strong correlations between image focus and object distance and between image size and object distance, vergence can also be stimulated by blurring of the retinal image and by dynamic changes in size of a target shape (Collewijn and Erkelens, 1990).

1.3.5 Eye Movement Coordination

In comparison to the behavior of the VOR, OKR, vergence, saccadic and smooth pursuit movements, relatively little is known about the neural mechanisms used to coordinate these different eye movements. Some of these movements have goals which are directly in conflict with each other, such as saccades which reorient the eye and the OKR, smooth pursuit, or VOR which all seek to stabilize an image. In addition to the interactions between eye movements, there are many influences of the

eye movement on the visual processing itself.

Perhaps the best studied interactions involve the combination of saccades and smooth pursuit movements which are both involved in visual tracking, (MacKenzie and Lisberger, 1986) and the combination of VOR and saccadic eye movements which seem to work together in parallel (Guitton et al., 1984). Because the smooth pursuit system is stabilizing only a part of the image, it must suppress the OKR, which is attempting to stabilize the entire scene. In animals where the pursuit system is not well-developed, the OKR prevents smooth movement and saccadic tracking is observed (Kirshfeld, 1993; Ilg and Hoffman, 1993; Masson et al., 1993). Behaviorally also, smooth pursuit movements reduce the perceived blurring of stationary background objects, compared to the equivalent retinal blur of a moving object (Bedell and Lott, 1996).

Saccadic suppression of the visual processing system has been found in the superficial layers of superior colliculus, pulvinar Robinson et al., 1991) Psychophysically, it seems that the magno system is selectively suppressed (Ilg and Hoffman, 1993). This, of course, impacts motion processing which could account for suppression of OKR and smooth pursuit during saccades.

1.4 What Has Been Done

The work presented here focuses on using neuromorphic analog VLSI techniques to model the primate oculomotor system to solve specific oculomotor tasks, such as performing saccades to visual and auditory targets, learning saccadic motor parameters, and using attention to control smooth pursuit eye movements. The goal was to construct a large system of interconnected sensory and motor processes to solve a task in the physical world.

- The central focus to date has been the construction of a one-dimensional model of the primate oculomotor system. The mechanical system described in Chapter 2 was designed to emulate the physical dynamics of the primate oculomotor plant. This mechanical system is controlled by an analog VLSI model of the primate brainstem burst generator circuits and is used to control both saccadic and smooth pursuit eye movements (Horiuchi et al., 1994).

- Chapter 3 describes a model of auditory localization developed from experimental work with barn owls to compute the direction of a sound source (Horiuchi, 1995). This system was also designed to trigger saccades to auditory targets after performing the required coordinate transformation from head to motor coordinates. This work builds off of the pioneering work of Lazzaro (1990) in analog VLSI models of the cochlea and auditory localization.

- Chapter 4 describes several schemes for extracting motion-related information from an image. One is a pulse-based, correlation method which utilizes delay lines and results in a population of velocity-tuned units (Horiuchi et al., 1991). Two other gradient-based methods for extracting

the direction of motion are presented (Horiuchi and Koch, 1996; Horiuchi et al., 1997). These motion circuits are used in the circuits and systems of later chapters.

- In Chapter 5 two different projects utilizing floating-gate circuits for non-volatile, on-chip analog memories is presented. These devices are used to demonstrate learning in the saccadic eye movement system. In the first project, we demonstrated the use of a sensitive direction-of-motion chip from Chapter 4 to provide error signals for tuning a motor control parameter to reduce post-saccadic drift in our system (Horiuchi and Koch, 1996). A second system focuses on the learning of a look-up-table mapping from retinal positions to motor commands. In this system an array of floating-gate storage circuits were trained in a supervised-learning paradigm. This system can learn to compensate for non-linearities in the optics or offsets within the saccadic burst generator circuit (Horiuchi and Koch, 1997).

- Chapters 6 and 7 jointly describe the use of an attentional model to perform target selection and extract the position and motion of the target to control saccadic and smooth pursuit eye movements (Horiuchi et al., 1997).

- Finally, in Chapter 8, having explored the design and performance of this multi-faceted oculomotor system, the contributions of this work to both neuroscience and to engineering will be discussed. In addition, the future of both this project and the neuromorphic analog VLSI field will be discussed.

Chapter 2 Modeling Saccades

The keystone of this dissertation is the hardware model of the oculomotor plant. Understanding its implementation and performance are important to understanding where some of the limitations in our system lie. While a large number of variations of both the oculomotor plant and the burst generator have been tested, only the most successful version will be described.

The first section discusses the oculomotor plant implementation and its performance, the second section discusses the saccadic burst generator and its performance in driving the plant, and the third section describes a vision-based, saccadic targeting chip.

2.1 The Oculomotor Plant

The primate eye is driven by 3 sets of muscles: the horizontal recti, the vertical recti, and the superior and inferior oblique muscles. These muscles and other suspensory tissues hold the orb in the eye socket, producing mechanical dynamics which the control circuits in the brainstem must control. Both the muscles and the suspensory tissues are elastic and, in the absence of motorneuron activation, return the orb to the primary, or center, position. Both muscles and suspensory tissues provide significant viscosity, creating an over-damped mechanical system.

The measured time-constant has been found to be approximately 250 msec (Robinson, 1964). The force-length relationship of the eye muscles was first measured by Collins et al. (1975) by measuring the tension on the antagonistic eye muscles as a function of eye position. While this relationship was fit well with a parabolic function, the combination of the two antagonistic muscles tend to cancel the non-linearity, producing a much more linear relationship.

The oculomotor plant model we have constructed is a one degree-of-freedom turntable which is driven by a pair of antagonistically-pulling DC (direct-current) brush motors. In the biological system where the agonist muscle pulls against the passive viscoelastic force of the antagonist muscle and suspensory tissues, the fixation position is determined by the balance point of these two forces. In order to maintain fixation away from the center position, a tonic signal to the motor controller must be maintained. In our system, the viscoelastic properties of the oculomotor plant are simulated electronically and the fixation point is set by the shifting equilibrium point of these forces. As the biological dynamics are not too far from linear (Collins et al., 1975), the system's dynamics have been modeled as linear, to simplify analysis and construction. The circuit schematics for the electronically-simulated mechanical dynamics are given in Appendix A.

The DC-brush motors are used to generate torque on the eye, by pulling on a thread attached

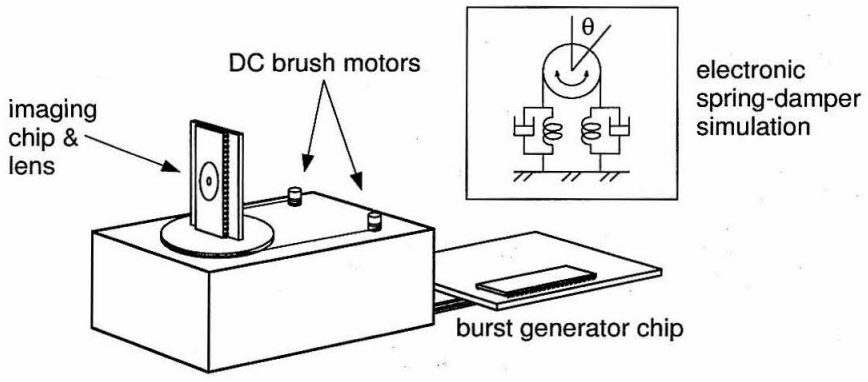

Figure 2.1: Diagram of the basic saccadic eye movement system. The physical dynamics of the primate oculomotor plant are simulated in electronics, implementing a heavily over-damped spring-mass system.

at the front of the turntable. Since the eye is stationary most of the time, static friction is a major problem, requiring a minimum amount of power to overcome the static frictional force. Once the eye is in motion, however, this minimum amount of power is considerably larger than the force needed to overcome sliding friction. Due to this problem, the motors are driven with a pulse-frequency-modulation (PFM) system in which every pulse to the motors is capable of moving the eye a tiny bit. In this way, arbitrarily-low average speeds can be achieved. At the slow end of the speed range, the eye is operating more like a stepper motor, with each pulse moving the eye a tiny fixed angle. As the frequency increases, the pulses begin to come before the eye has come to a complete stop, thus avoiding the effects of static friction. Each pulse is now more effective in driving the eye and the torque vs. input frequency curve follows a steeper slope. Note also that as the motor begins to spin, internally it acts as a generator, reducing the effective driving voltage of the controller. This, in turn, reduces the torque. A reduction in driving force due to velocity is also seen muscles (Kandel et al., 1991). The one-dimensional eye position is measured by a potentiometer which also serves as the bearing for the turntable. This position signal is used by the electronic dynamics simulation to simulate the following torque (T) relationship:

$$T(\theta, \dot{\theta}) = T_{ext} - k\theta - m\dot{\theta} = I \cdot \ddot{\theta} \tag{2.1}$$

where k is the spring constant, m is the damping coefficient, I is the rotational inertia, and T_{ext} is the externally-applied torque. Using the theory of Laplace transforms, equation 2.2 can be transformed into:

$$\frac{\theta(s)}{T_{ext}(s)} = \frac{1}{(1 + s\alpha_1)(1 + s\alpha_2)} \tag{2.2}$$

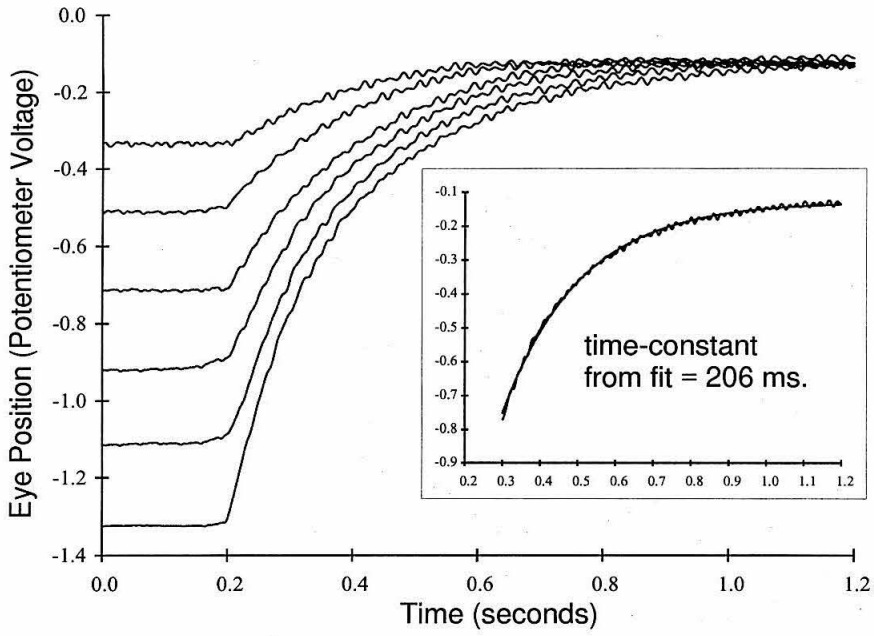

Figure 2.2: Passive decay and time constant measurement. The eye was moved from its equilibrium point to six different positions and released. The eye position is linear with the potentiometer voltage.

$$\alpha_1, \alpha_2 = \frac{m}{2I} \pm \frac{1}{2I}\sqrt{m^2 - 4kI} \tag{2.3}$$

For this passive case with no external torques (i.e. $T_{ext} = 0$), this equation results in the equations of motion for the eyeball as a function of two exponentials:

$$\theta(t) = c_1 e^{-t/\tau_1} + c_2 e^{-t/\tau_2} \tag{2.4}$$

The two time-constants, τ_1 and τ_2, are functions of I, k, and m:

$$\tau_1 = \frac{1}{\alpha_1}, \tau_2 = \frac{1}{\alpha_2}$$

For the case of the heavily over-damped eyeball, where $m^2 \gg 4kI$, one of the two time-constants is much larger than the other. By using the initial conditions,

$$\theta(0) = c_1 + c_2$$

$$\dot{\theta(0)} = \frac{-c_1}{\tau_1} + \frac{-c_2}{\tau_2}$$

the variables, c_1 and c_2 can be determined. When the eye is released from an off-center position (i.e.

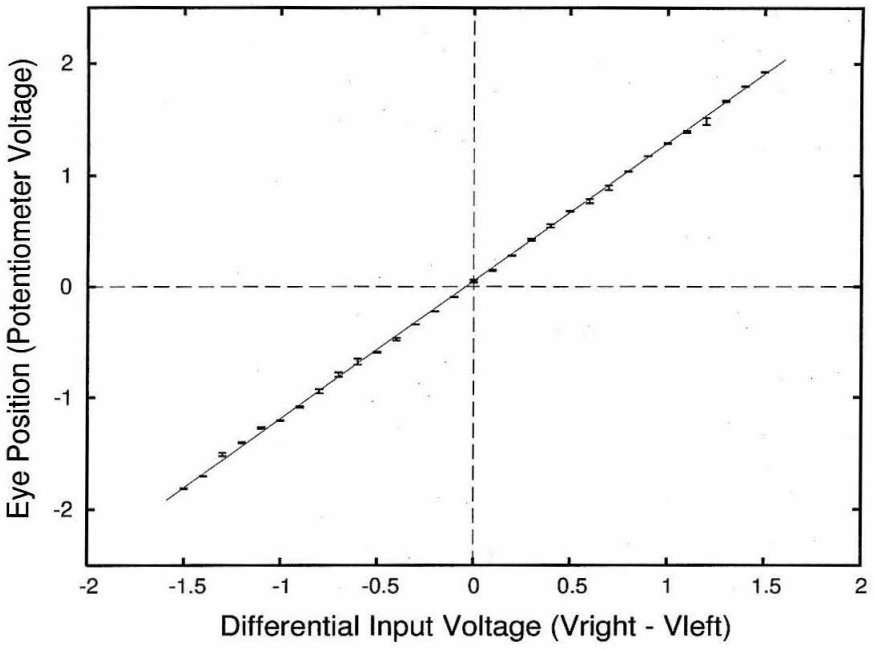

Figure 2.3: Linearity and repeatability of the oculomotor plant system. The eye position was perturbed away from the equilibrium point between each measurement. N = 11

$\theta(0) = A, \dot{\theta(0)} = 0$), the eye initially accelerates (both time-constants) and then slowly decelerates as it approaches the center position (long time-constant). Figure 2.2 shows the temporal relaxation response of the oculomotor plant from six different initial positions and the inset shows an exponential fit, late in the response, for the larger time-constant. The mechanical system has been tuned such that the larger time-constant has been set to be approximately 200 msec.

The oculomotor plant model is driven by two signals, V_{left} and V_{right}, which represent the torques generated by the two antagonistically-pulling lateral rectus muscles of the eye. The difference in these two inputs represents the net torque delivered to the eye. Figure 2.3 shows the position linearity of oculomotor plant as a function of the difference in the two command voltages.

One consequence of generating the dynamics electronically is that in steady-state, the eye is constantly vibrating. The damping forces are implemented by adjusting the motor control signals before they drive the mechanical system. Since the motors are ultimately driven with current pulses, the torque generated by the motors are not damped except by the compliance of the final mechanical system. This pulsing creates a low-amplitude oscillation around the steady-state eye position.

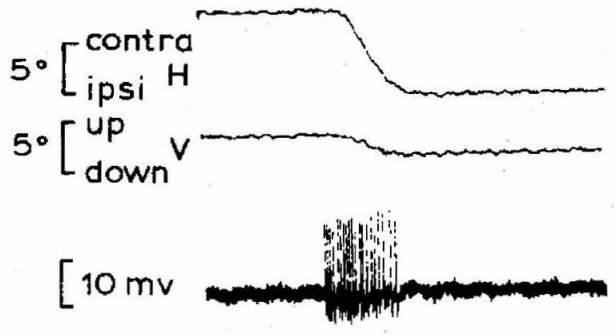

Figure 2.4: Excitatory burst neuron during a saccadic eye movement. The upper trace is the horizontal eye position, the middle trace is the vertical eye position and the lower trace is the intraaxonal recording of the membrane potential. (From Strassman et al., 1986)

2.2 The Saccadic Burst Generator Model

To drive the oculomotor plant described in the previous section, the control circuitry must provide the proper signals to overcome the tissue elasticity and viscosity. In order to maintain fixation off-center, a sustained pulling force must be generated. Also, in order to complete an eye movement faster than the eye's natural time-constant, a large, transient, acceleration force must also be generated (see Figure 2.4). Both the transient and sustained component signals can be found in brainstem areas which drive the motorneuron pools. Accurate balancing of these two signals is necessary and is observed in the motorneurons driving the eye muscles (See Figure 2.5). If the transient pulse is too large, the eye will overshoot and if the pulse is too small, the eye drifts onward after the saccade.

The analog VLSI burst generator model (Figure 2.6) receives, as its input, the desired change in eye position and creates a two-component signal, a pulse (signal (A), Figure 2.7) and a step (signal (B), Figure 2.7) in muscle innervation. A pair of these pulse/step signals drive the two muscles of the eye. The burst generator model is a double integrator model based on the experimental and modeling work of Jürgens et al., (1981), McKenzie and Lisberger (1986), and Nichols and Sparks (1995). The first integrator, which we call the burst integrator, is used to control the burst duration; the second integrator, known as the neural integrator, holds a dynamic memory of the current eye position. Unlike the original Robinson saccadic burst generator model (Robinson, 1975), this model uses initial motor error as the input to the system. The motor error is compared to the output of the burst integrator which integrates the burst unit's spike train. The burst neuron keeps bursting until the difference is zero. This arrangement has the effect of firing a number of spikes proportional to the initial value of motor error, consistent with the behavior of short-lead burst neurons found in the saccade-related areas of the brainstem (Hepp et al., 1989). In the circuit, this integrator

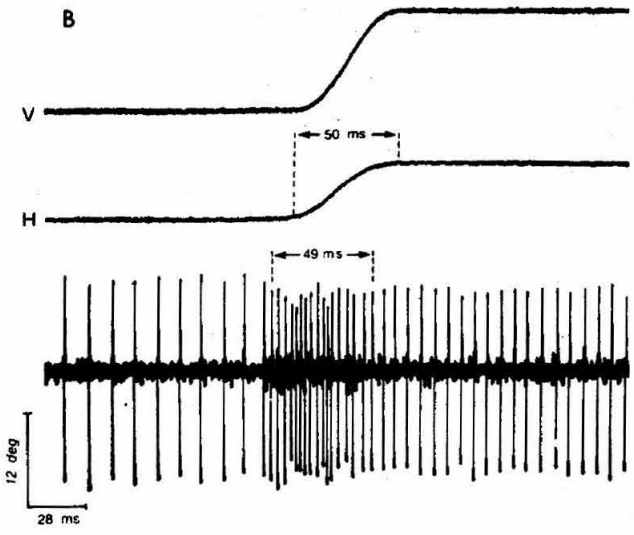

Figure 2.5: Motor neuron spike train with horizontal and vertical eye position shown during an oblique saccade. (From King et al., 1986)

is implemented by a 1.9 pF capacitor. After the burst is over, the integrator is reset. This burst of spikes serves to drive the eye rapidly against the viscosity. The burst is also integrated by the "neural integrator" (another 1.9 pF capacitor) which holds the local estimate of the current eye position from which the tonic signal is generated. The neural integrator provides two output spike trains, driving the left and right sustained components of the motor command. The motor units receive inputs from both the burst units and the neural integrator and outputs the sum of these two signals. Figure 2.7 shows output data from the burst generator chip which is qualitatively similar to spike trains seen in the motor neurons of the abducens nucleus of the brainstem (King et al., 1986).

Figure 2.8 shows the circuit used to control the burst generator activity. The inputs to the burst generator are an analog voltage representing the initial motor error (Vin) and a digital trigger signal (active-low) to initiate the burst. The Vin voltage level must be maintained throughout the duration of the burst. A control logic circuit is used to suppress bursting unless a trigger signal has been received. Once triggered, however, the control logic circuit keeps the burst going until the difference between Vin and Vbi is zero. The core of the control logic circuit is a latch circuit with SET and RESET inputs. In the quiescent state (no burst), the latch circuit is in the RESET mode and the burst units are prevented from firing spikes by keeping the "GO-L" and "GO-R" signals logic-low. Also in this state, the burst-integrator reset is held high, resetting the integrator voltage (Vbi) to the zero reference level. When the trigger signal goes low, the latch goes active and the appropriate "GO" line is driven high. As the burst proceeds, Vbi approaches and then passes the Vin voltage.

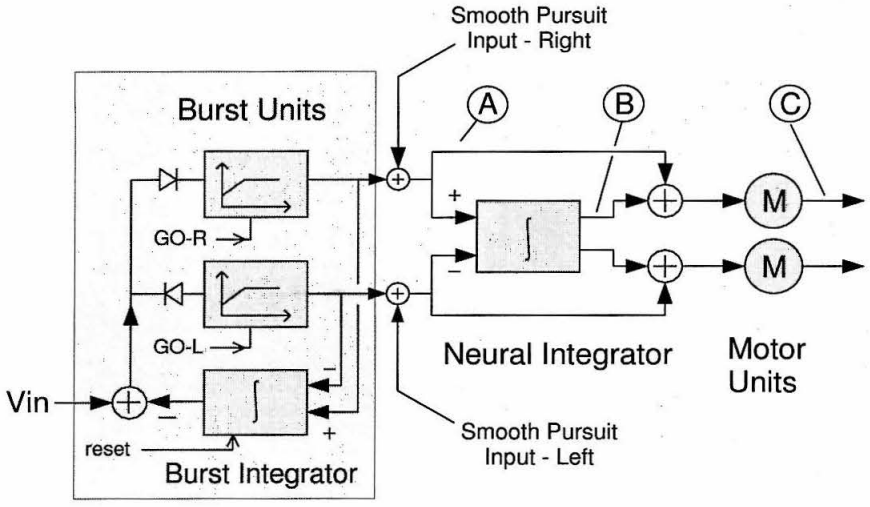

Figure 2.6: System block diagram of the burst generator, neural integrator, and motor units. The input to the system is an analog voltage representing initial motor error. This input is compared against the output of an integrator, and the difference signal drives either the left or right burst unit. The burst unit, however, is inhibited until triggered by the GO-L or GO-R signals. Once triggered, the burst spikes are integrated, and as the integrator value approaches the input value, the burst is reduced and then shut down. This "pulse" signal (A) drives the eye against the viscosity. This signal is also integrated by the neural integrator which contributes the "step" portion of the motor command (B) to hold the eye in its final position. See Figure 2.7 for example signals. The neural integrator has additional velocity inputs for other oculomotor behavior such as smooth pursuit, VOR and OKR. This circuitry has been implemented on a single silicon die.

As it passes, the comparator changes state and the zero-crossing detector shuts the latch off.

Figure 2.9 shows the burst generator circuit. The input to the burst unit is the initial motor error voltage Vin which is compared to the burst integrator voltage V_{bi}. If V_{in} is larger than V_{bi}, then the lower burst circuit is activated when the GO-L line is activated by the burst control circuitry. If V_{in} is less than V_{bi}, the upper circuit is active. During periods of no burst activity, the burst integrator reset line is high, driving V_{bi} towards the reset value, which is the zero-reference voltage. When one of the "GO" lines go high, the corresponding differential pair drives a bias current, determined by the burstbias input, into its spike-generating circuit (the current labelled Ibu in Figure 2.9).

$$B(t) = f(V_{in} - V_{bi}) \tag{2.5}$$

where

$$f(x) = \frac{\gamma e^{Kx}}{e^{Kx} + 1} \tag{2.6}$$

$$V_{bi} = V_{bi}(0) + \int_0^t B(t)dt \tag{2.7}$$

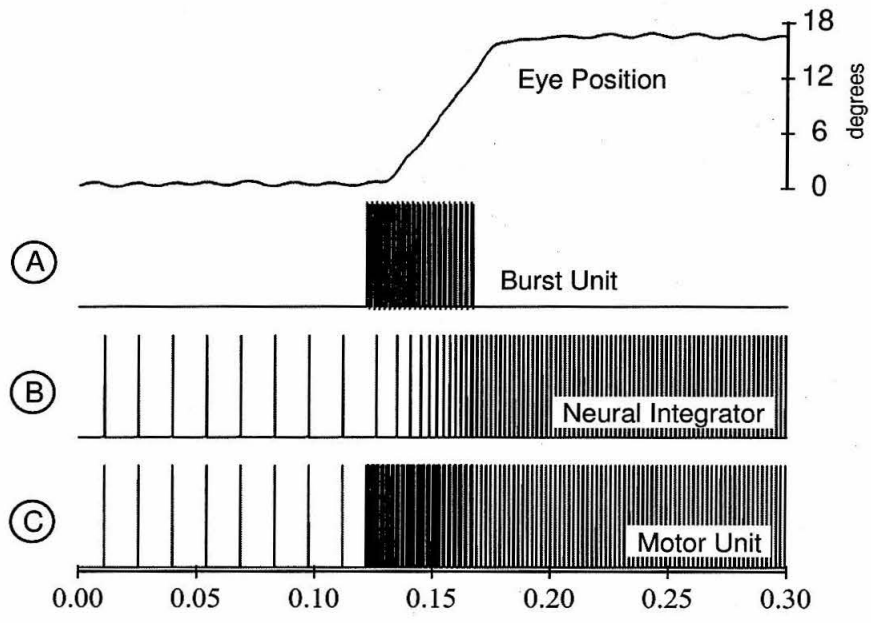

Figure 2.7: The eye position, burst unit, neural integrator output, and motor unit spike trains during a small saccade. The initial eye position is off-center with the neural integrator providing the tonic holding activity. All three spike trains are digital outputs (0 to 5 volts). Note that only the motor unit output (C) is used externally. The small oscillation seen in the eye position trace is due to the pulse-frequency modulation technique used to drive the eye.

The output of the spiking unit is sent back to the burst integrator as well as to the neural integrator and motor stages. The pulses generated by the spiking unit are converted to current pulses which are then integrated. As the burst integrator is driven towards Vin, the differential pair supplying the spiking unit reduces its current and the spike frequency drops. If the initial difference in voltage between Vin and Vbi is larger than 200 mV, the differential pair is saturated and supplies its maximum current. When Vbi equals Vin (the state in which the burst is shut down by the control circuitry) the differential pair supplies half of its maximum current. Figure 2.10 shows the simulated burst current (Ibu) vs. time for six different saccades. Figures 2.7 and 2.12 both show this decrease in firing rate. This arrangement creates increasing spike frequencies with increasing amplitude saccades, with the frequency saturating as Vin exceeds 200 mV. This is ultimately reflected in the peak-velocity of the eye during saccades, increasing with amplitude until it saturates near |Vin - Vbi| > 200 mV. (See Figure 2.14.)

Experiments by Nichols and Sparks (1995) support the resettable-integrator model by demonstrating the effect of triggering two saccades in quick succession such that the burst integrator does not have sufficient time to completely reset. In their experiments, the integrator was shown to gradually reset with a time-constant of about 250 milliseconds. Figure 2.12 shows a similar experiment

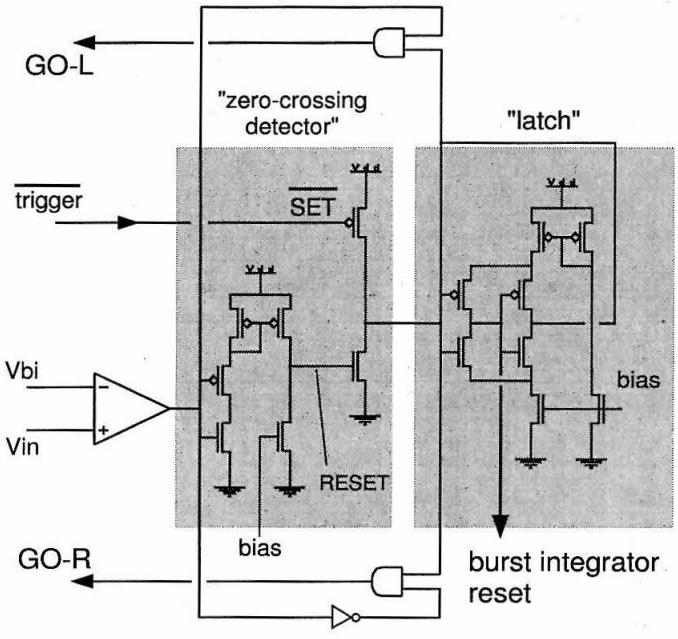

Figure 2.8: Logic circuitry for controlling the operation of the burst generator. Before a saccade, the burst integrator (Vbi) is reset to the reference voltage and the comparator uses the initial motor error (Vin) to determine whether a leftward or rightward saccade is required. The latch circuit, constructed from a pair of current-limited inverters, is initially set with the output logic-low. When the trigger signal goes low, the latch becomes active and the burst-integrator reset line goes low. One of the "GO" lines is active and the burst proceeds, driving Vbi towards Vin. When the comparator crosses zero and the output flips state, the zero-crossing detector circuit drives the latch to the "off" state. This occurs as the comparator's output voltage passes through the mid-range of the inverter, drawing a large current through the p-type mirror, activating the n-type transistor, and resetting the latch.

with the analog VLSI burst generator.

Figure 2.11 shows the neural integrator circuit. The two inputs to the circuit are the digital signals in-R and in-L, which are the spike trains from the burst generator circuit. The spikes from in-L increase the integrator value and the spikes from in-R decrease the integrator value. The input spike trains are converted to current pulses by the transistors biased by voltages synR and synL. The remainder of the circuit linearly converts this analog voltage into two currents which drive the left and right spike generator circuits. This function is performed by a population of differential pairs comparing the neural integrator voltage (Vni) with the local tap voltages along a biased, resistive line. If Vni is at the voltage midway between Vres-high and Vres-low, both currents are zero. If Vni rises, the differential pairs in the upper half sequentially turn on and saturate, adding their currents together. This recruitment of units generates a staircase function of current if the Vres-high and Vres-low voltages are set too far apart, but more modest settings of Vres-high and Vres-low generate a more linearly changing current as a function of Vni. As the current increases, the output spiking

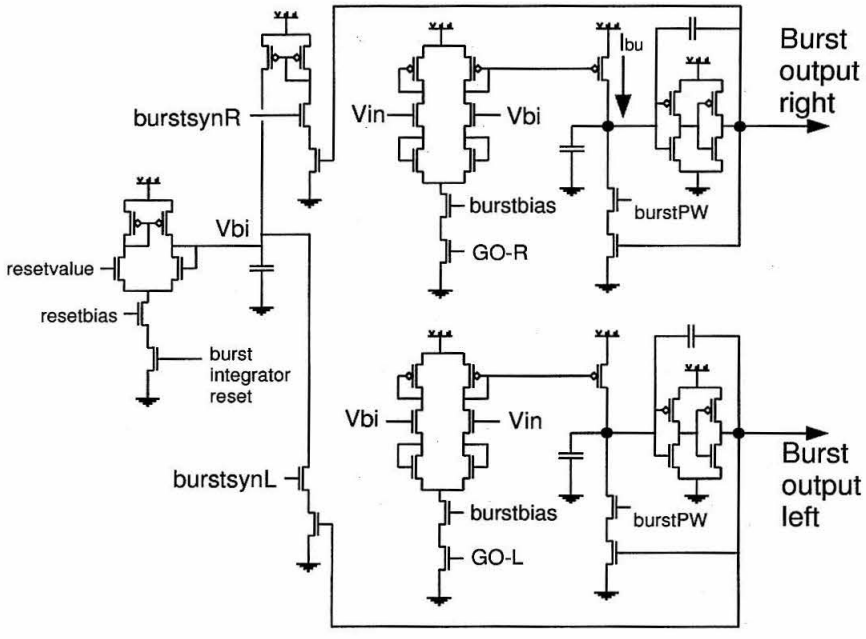

Figure 2.9: Circuitry of the resettable-integrator burst generator.

unit increases its firing frequency. Note that at high firing rates where the duty cycle approaches 50%, the pulse width begins increasing, but the effective duty cycle continues to rise linearly until the duty cycle reaches 100%. Not shown are several circuits used to reset the neural integrator voltage and to keep the integrator voltage between user-definable upper and lower limits.

The burst and neural integrator digital outputs are connected to the motor output units by a pair of transistors, similar to the inputs to the neural integrator circuit, one acting as a switch for a current defined by the other, converting the digital voltage pulse trains into current pulse trains. The transistors are biased such that each input spike triggers an output spike from the motor unit.

2.2.1 System Integration

By connecting the burst generator output to the oculomotor plant, saccadic eye movements were generated. Figure 2.13 shows an overlay of 20 saccadic trajectories. Figure 2.14 shows the peak velocity of these 20 saccades as a function of the input command. Similar to the peak-velocity vs. amplitude relationship in primate saccades, the peak velocity increases for increasing saccade amplitude and then saturates. As the peak velocity of the saccades saturate, the duration of the saccades increase linearly with amplitude. These characteristics are qualitatively consistent with primate saccades (Becker, 1989).

Careful tuning is necessary to match the relative strengths of the burst component and neu-

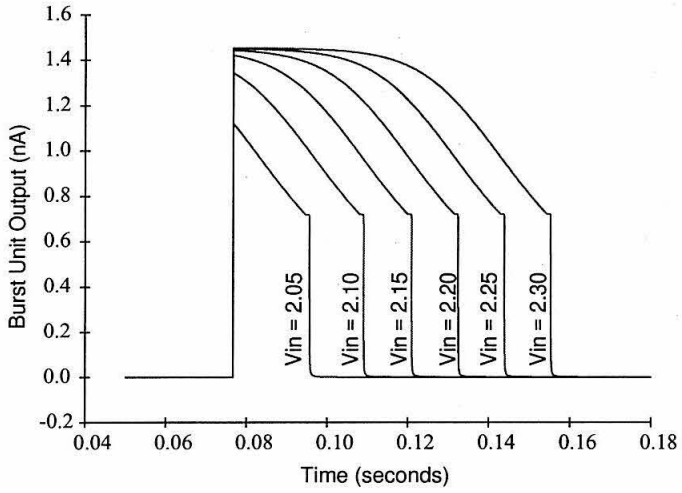

Figure 2.10: Circuit simulation of the burst current for 6 different saccade amplitudes. Using the circuit in Figure 2.9, the output current which is used to drive the burst neuron circuit is shown for 6 different Vin values. Vbi is initially reset to 2.0 volts. Note that with increasing saccade amplitudes, the initial burst current increases and then saturates. The burst current is converted to burst frequency by the burst neuron and then to eye velocity by the oculomotor plant.

ral integrator component in the motor output. While the data shown here were manually-tuned, Chapter 5 presents an example of self-adaptation possible within this system.

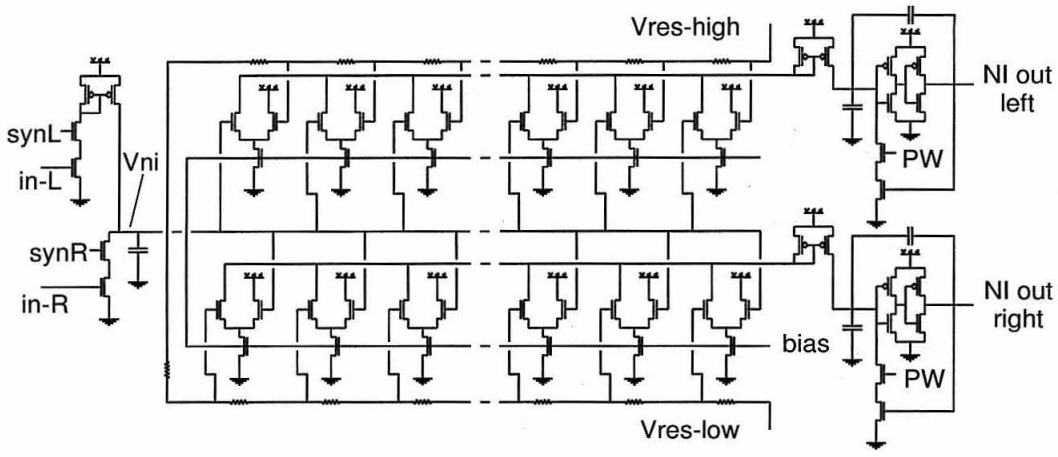

Figure 2.11: The neural integrator circuit consists of two pulse inputs, in-L and in-R, which generate current pulses to increase or decrease the neural integrator voltage Vni. Vni is linearly converted to current via a population of differential pairs which are sequentially turned on such that their output currents add together. The actual circuit contains ten differential pairs for each direction. PW controls the pulse width of the neural integator output pulses.

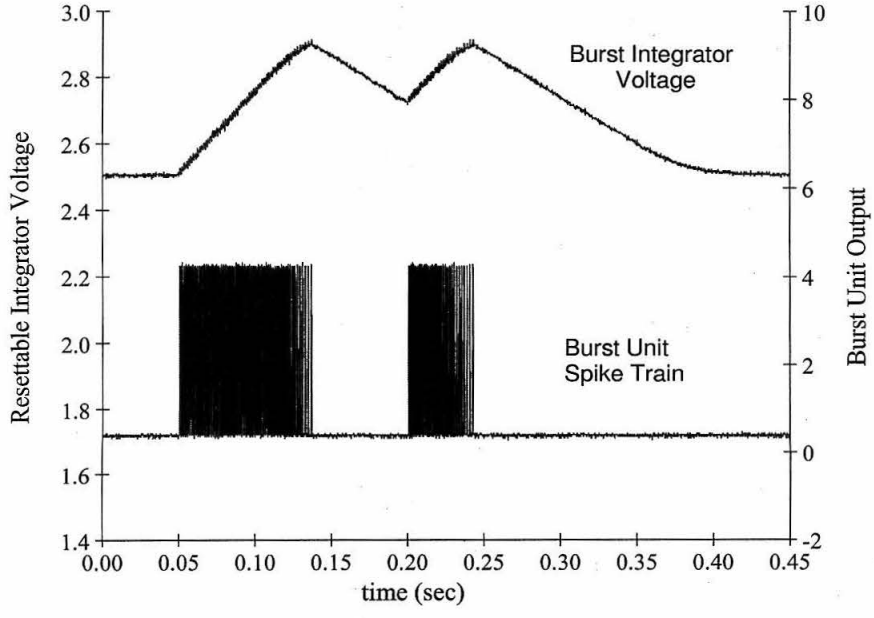

Figure 2.12: Due to the reset latency of the burst integrator (upper trace), burst commands issued before the integrator has fully reset result in reduced amplitude saccades which vary systematically with command latency.

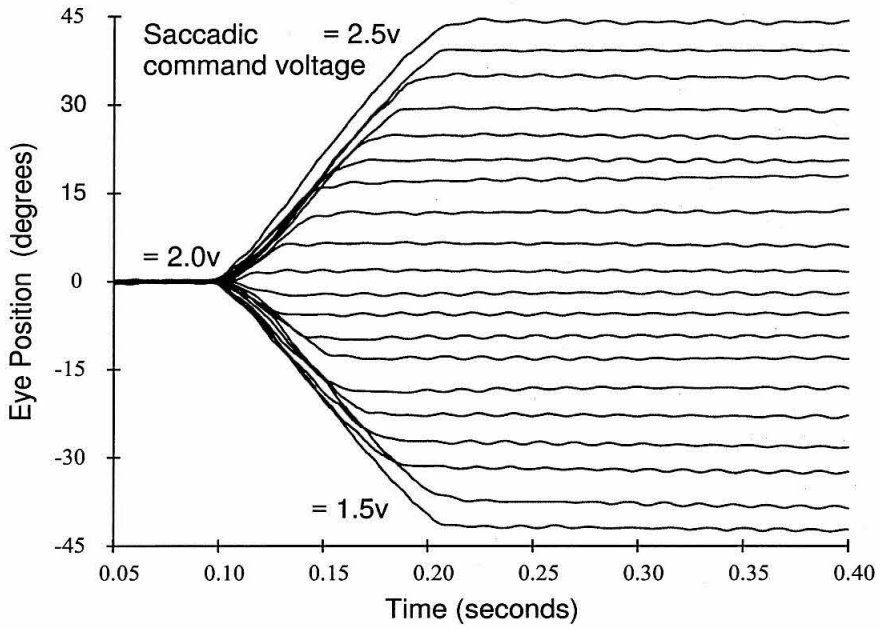

Figure 2.13: Angular eye position vs. time for 20 different saccades triggered from the center position. The differential-inputs were swept uniformly for different saccade amplitudes from leftward to rightward. Peak angular velocity achieved for the approximately 44 degree saccade to the left was approximately 450 degrees per second. Peak velocities of up to 870 degrees/sec have been recorded on this system with different parameter settings than used here. The small oscillations in the eye position are due to the discrete pulses used to drive the eye motors; at rest, the pulses are at their lowest frequency and thus most visible.

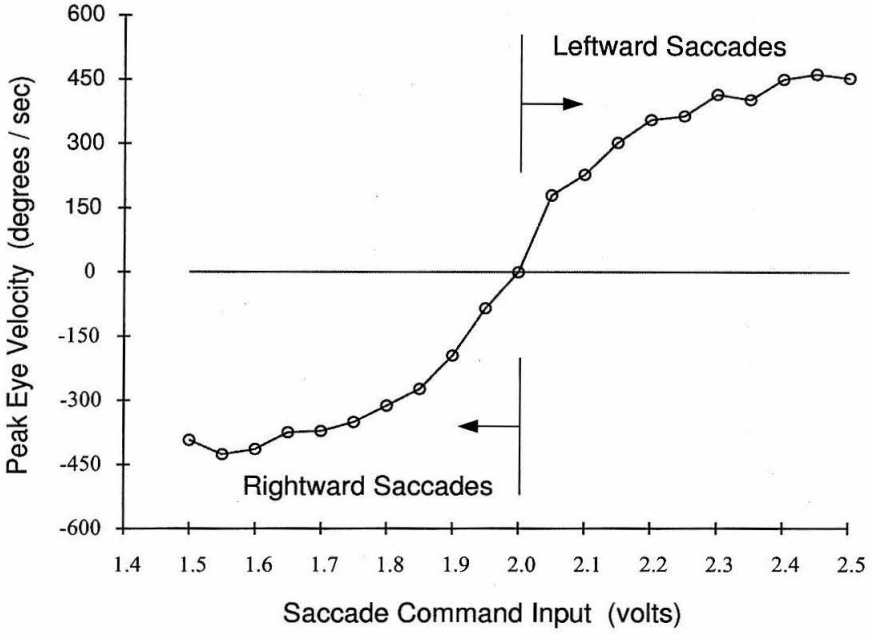

Figure 2.14: Plot of the peak velocity during each of the saccades in Figure 2.13. These velocities were computed by performing a least-squares fit to the center region of the saccade trace. The input voltage of 2.0 is the internal voltage reference for zero amplitude saccades.

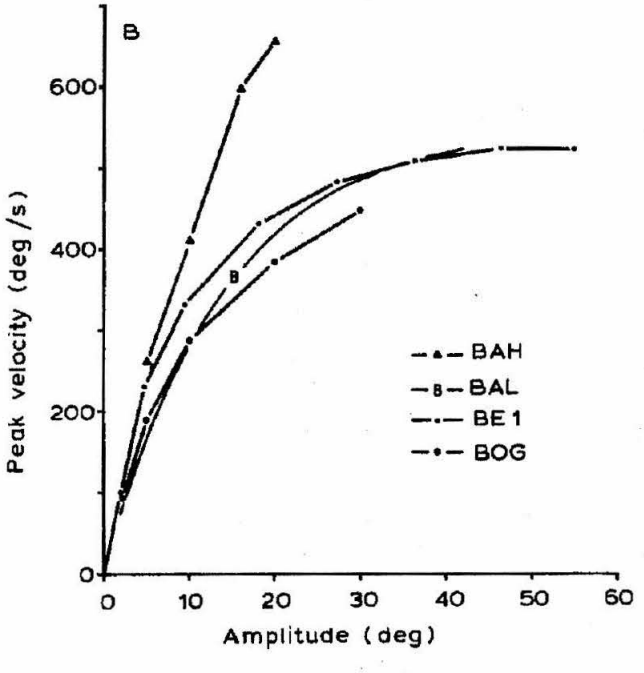

Figure 2.15: Plot of the peak velocity vs. saccade amplitude for human saccades. The different plots are different subjects. (Modified from Becker, 1989)

2.3 Visually-Guided Saccades

The superior colliculus, (SC) located on the dorsal surface of the midbrain, is a key area in the behavioral orientation system of mammals. The superficial layers have a topographic map of visual space and the deeper layers contain a motor map of saccadic vectors. Microstimulation in this area initiates saccades whose metrics are related to the location stimulated. During saccades, a large region of the superior colliculus is active; the activity is centered such that the average of the saccade vectors for each active cell generates the correct saccade. This type of representation is known as population coding. In addition, many neurons in the deeper layers of the superior colliculus are multisensory and will generate saccades to auditory and somatosensory targets as well as visual targets.

2.3.1 The Change-Detection Chip

In order to trigger saccades to visual stimuli, a temporally-sensitive vision chip was designed to model the visually-sensitive, deep-layer neurons of the primate superior colliculus (Sparks and Hartwich-Young, 1989). While it is clear that the superior colliculus performs a multitude of integrative functions between sensory modalities and attentional processes, the initial model of the SC simply computes the centroid of activity from the population of active photoreceptors.

Figure 2.16 shows the circuit for one cell in the one-dimensional array. Each cell has an adaptive photoreceptor circuit (Delbrück, 1989) which moves its operating point to remain sensitive to small-amplitude temporal changes in intensity. The photoreceptor adapts rapidly (milliseconds) for large amplitude swings and slowly (seconds) for small amplitude swings. The temporal-derivative of the photoreceptor voltage is passed through a full-wave rectifying circuit to detect changes in image intensity. The output drives a centroid circuit (DeWeerth, 1992) to output a voltage unique to each pixel position. When a group of pixels are activated, as expected with any realistic stimulus, the output voltage becomes a weighted-average of the different pixels' position, where the weighting is given by the amplitude of the temporal-derivative. Figure 2.17 shows the output of circuit in response to a flashed LED at different angular positions in front of the chip.

This temporal-derivative photoreceptor array has been mounted on the oculomotor system to drive saccades. Since the resistive line providing the centroid circuit with local reference voltages is linear, the mapping between retinal positions and motor commands is linear. Chapter 5 describes a learning system using this triggering circuit in which non-linear mappings can be trained. A close-up photograph of the photoreceptor array mounted on the eye movement system is shown in Figure 2.18.

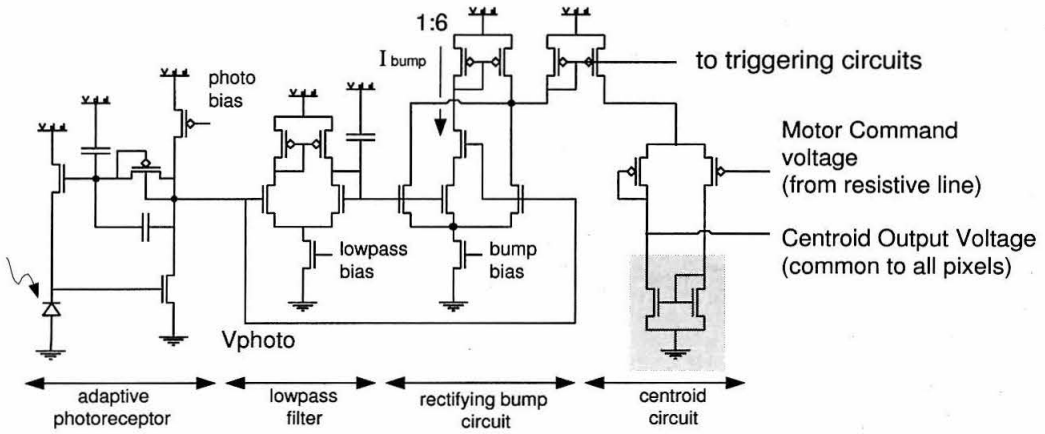

Figure 2.16: Schematic of the temporal-derivative photoreceptor cell. On the far left, the adaptive photoreceptor circuit converts photodiode current to an output voltage, Vphoto. The rectifying bump circuit compares Vphoto with the lowpassed version of Vphoto. The circuit generates zero current when the two inputs are the same and increasing currents as the difference in the inputs increase. The centroid circuit uses the input current to drive the motor command voltage onto the output line.

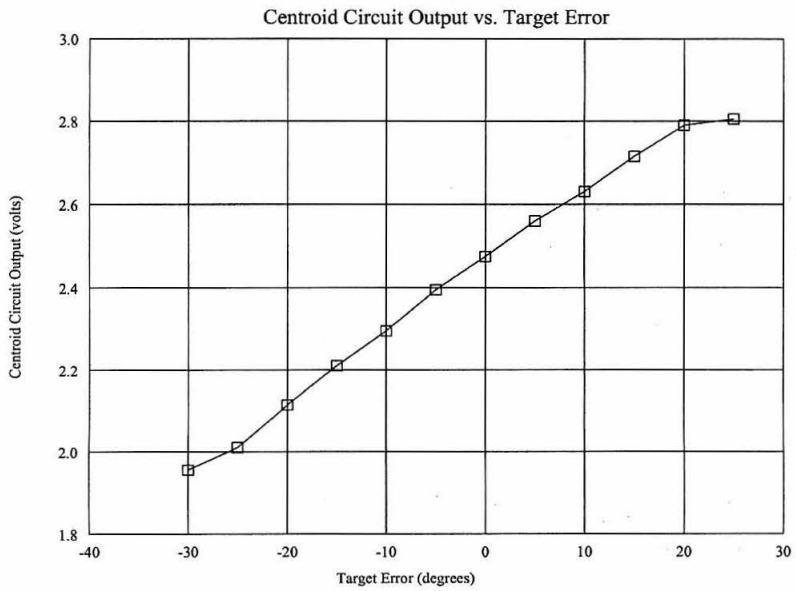

Figure 2.17: Centroid output voltage as a function stimulus angle. A red LED was flashed at different angles in the field of view and the centroid output voltage was sampled 1 msec after stimulus onset to account for capacitive delays.

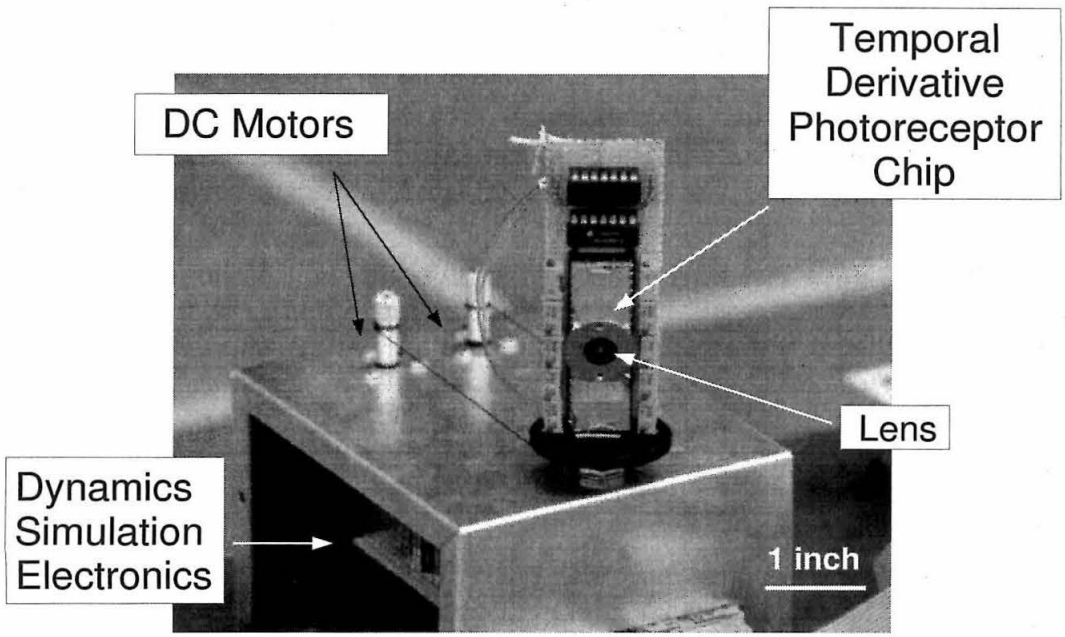

Figure 2.18: Photoreceptor array mounted on the saccadic eye movement system. Potentiometers for setting various chip parameters are visible on the board next to the photoreceptor chip. Note: only the pulleys mounted on the motor shafts are visible on the box. The motor casings are inside the box.

Chapter 3 Auditory Localization

While the visual localization system described in Chapter 2 enables a creature to detect potential predators or prey efficiently within the field of view, many realistic environmental conditions can confound this ability. The visual sense is most useful during the day, whereas the auditory localization system is perhaps most useful at night. When vision becomes ineffective due to darkness, camouflage, or a limited field of view, the auditory localization system can continue to locate potential predators or prey, albeit at lower resolution. Barn owls, dolphins, humans, and many echo-locating bats are notable examples of creatures that can successfully use auditory cues to locate sound sources quite accurately. While most birds can resolve the location of sounds to only 10 to 20 degrees, the barn owl is able to orient to sound sources with an accuracy of 1 to 2 degrees which is comparable with humans (Dusenbery, 1992). Neurophysiologically, auditory localization has been studied perhaps most extensively in barn owls.

There are two main acoustic cues for localizing sounds: interaural intensity differences (IID) at higher frequencies and interaural time-of-arrival differences (ITD) at lower frequencies (see Figure 3.1). For the measurement of ITD, a neural delay-line architecture first proposed by Jeffress (1948) has been shown to exist in the barn owl auditory localization system (Konishi, 1986). In addition, a topographic direction map of auditory space has been mapped in the inferior colliculus (IC).

In comparison to the barn owl, the neurophysiology of auditory localization in primates is not as well understood and a clear map of auditory space does not appear to be present in the inferior colliculus as it is in the owl. However, like the owl, auditory localization information is combined with visual information at the level of intermediate and deep layers of the superior colliculus. For primates, there is growing evidence that cortical auditory regions may hold the topographic map of auditory space (Groh and Sparks, 1992).

In primates, where much of the early visual and oculomotor system is based in retinotopic coordinates, any information in head-coordinates must ultimately be transformed before it can be used. Several models of coordinate transformation have been proposed for visual information (e.g., Zipser and Andersen, 1988; Krommenhoek et al., 1993) and for auditory information (Groh and Sparks, 1992) which can be used for this purpose.

This chapter describes an auditory localization system based on a subset of the analog VLSI modelling work of Lazzaro (1990) which has been extended to include a transformation of auditory target information from head coordinates into retinal coordinates. The system is then used to

$$\Delta t = t1 - t2$$

$$\Theta \doteq \arcsin\left(\frac{\Delta t \cdot V_{sound}}{D}\right)$$

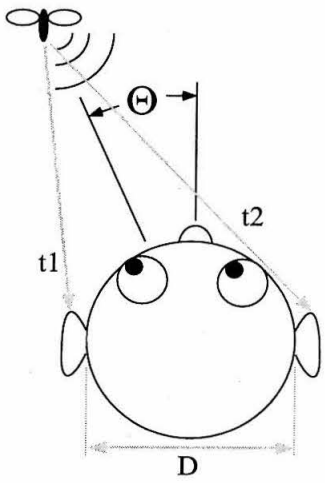

Figure 3.1: Angular localization of sound sources in the horizontal plane can be estimated by measuring the difference in arrival times of the acoustic waveforms to the two ears.

provide audition-based, target information to the analog VLSI-based saccadic eye movement system described earlier in Chapter 2 for multimodal operation.

3.1 The Localization System

The localization system consists of an analog preprocessing stage and an analog VLSI-based localization system (Figure 3.3). The discrete-component, analog preprocessing of the system consists of three basic components, the microphones, the filter stage, and the thresholded, zero-crossing stage. Two microphones are placed with their centers about 2 inches apart. For any given time difference in arrival of acoustic stimuli, there are many possible locations from which the sound could have originated. These points describe a hyperbola with the two microphones as the two foci. If the sound source is distant enough, we can estimate the angle since the hyperbola approaches an asymptote. In contrast to the full spectrum cochlear model used by Lazzaro, the current system operates on a single frequency and the inter-microphone distance has been chosen to be just under one wavelength apart at the filter frequency. The filter frequency chosen was 3.2 kHz because the author's finger snap, used extensively during development, contained a large component at that frequency. Figure 3.4 shows examples of the delay in the filtered waveforms. The next step in the computation consists of triggering a digital pulse at the moment of zero-crossing if the acoustic signal is large enough (See Figure 3.5). The circuit schematics of the preprocessing used in this system are described in Appendix B.

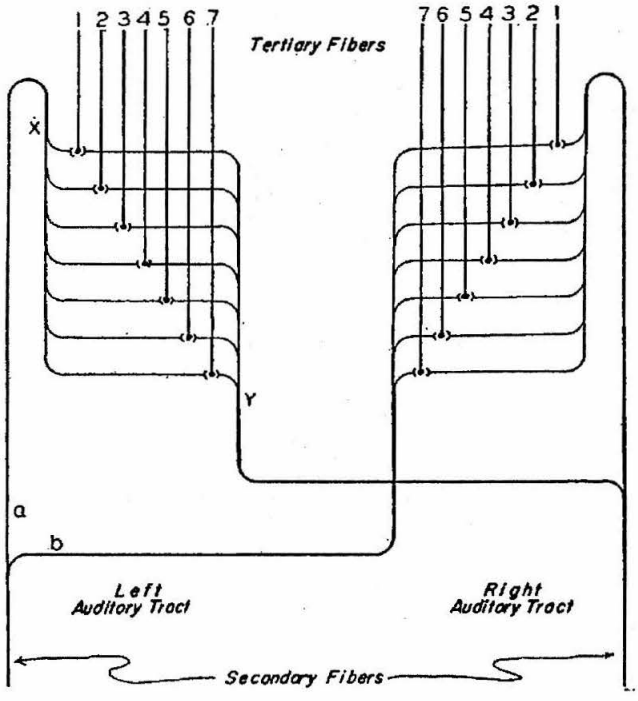

Figure 3.2: From Jeffress (1947): Hypothetical mid-brain mechanism for the localization of low-frequency tones. By using the slow rate of conduction of small nerve fibers, a difference in time can be represented as a difference in space. If impulses arrive at the left and right auditory tracts at the same moment in time, they will arrive at X and Y respectively. Traveling further, they arrive at the same moment at the neuron which sends out tertiary fiber #4, but arrive non-coincidentally at all other neurons. If the neurons require coincident stimulation from both fibers, then only neuron #4 will be active. Differences in impulse arrival times will selectively stimulate different neurons.

3.1.1 Phase Detection and Coordinate Transform

The analog VLSI component of the system consists of two axon delay lines (Mead, 1989) which propagate the left and right microphone pulse signals in opposing directions in order to compute the cross correlation (see Figure 3.6.) The location of the peak in this correlation technique represents the relative phase of the two signals. This technique is described in more detail and with more biological justification by Lazzaro (1990). The current implementation contains 15 axon circuits in each delay line. This is shown in Figure 3.6. At each position in the correlation delay line are logical AND circuits which output a logic one when there are two active axon units at that location. Since these units only turn on for specific time delays, they define auditory "receptive fields." The output of this subsystem are 15 digital lines which are passed on in parallel to the coordinate transform circuitry.

For the one-dimensional case described in this project, the appropriate transform from head to retinal coordinates is a rotation which subtracts the eye position. The eye position information

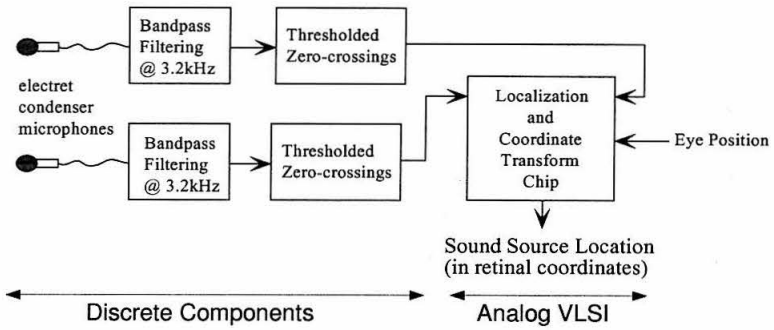

Figure 3.3: Block diagram of the auditory localization system. The analog front end consists of external discrete analog electronics.

on the chip is represented as a voltage which activates one of the eye position units. The spatial pattern of activation from the auditory output units is then "steered" to the output stage with the appropriate shift (See Figure 3.7). This is similar to a shift scheme proposed by Pitts and McCulloch (1947) for obtaining pitch invariance for chord recognition. The eye position units are constructed from an array of "bump" circuits (Delbrück, 1991) which compare the eye position voltage with its local voltage reference and output a current when the two voltages are similar. The two-dimensional array of intermediate units take the digital signal from the auditory units and switch the "bump" currents onto the output lines. The output current lines drive the inputs of a centroid circuit.

This implementation of the shift can be viewed as a basis function approach where a population of intermediate units respond to limited circular regions in the two-dimensional space of horizontal eye position and sound source azimuth (head-coordinates). The output units integrate the outputs of only those intermediate units which represent the same retinal location. It should be noted that this coordinate transformation is closely related to the "dendrite model" proposed for the projection of cortical auditory information to the deep SC by Groh and Sparks (1992) and to the basis function approach of Pouget and Sejnowski (1995).

The final output stage converts this spatial array of current-carrying lines into a single output voltage which represents the centroid of the stimulus in retinal coordinates. This centroid circuit (DeWeerth, 1991) is the same circuit used in Chapter 2 for the model of visual processing in the superior colliculus.

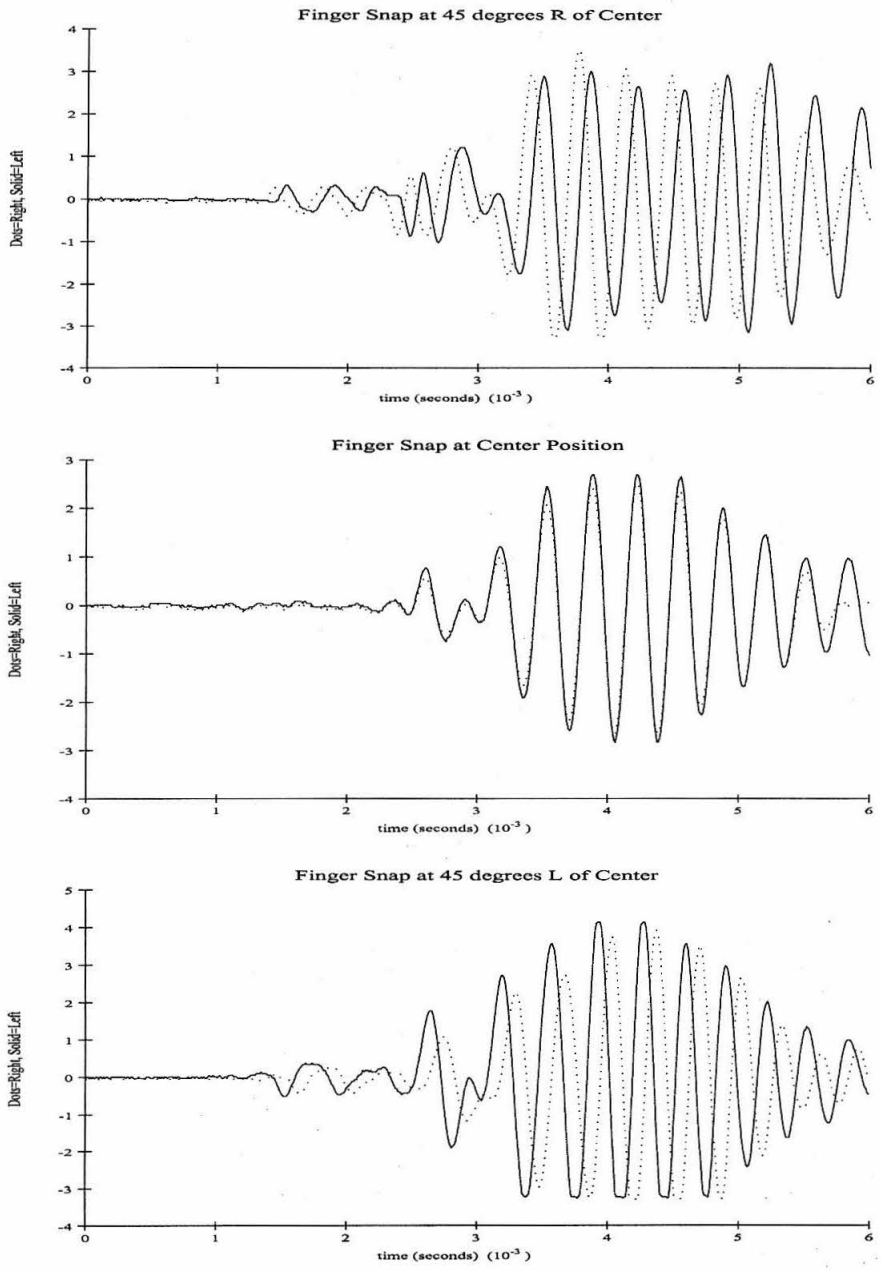

Figure 3.4: Bandpass filtered (3.2 kHz) signals of the left and right microphones from three different angles.

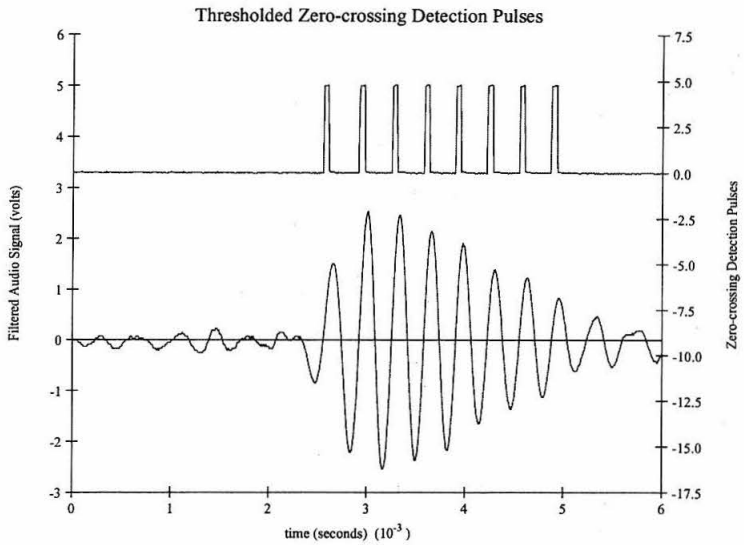

Figure 3.5: Example of output pulses from the external circuitry. Zero phase is chosen to be the positive-slope zero-crossing. Top: Digital pulses are generated at the time of zero phase for signals whose derivative is larger than a preset threshold. Bottom: 3.2 kHz Bandpass filtered signal for a finger snap.

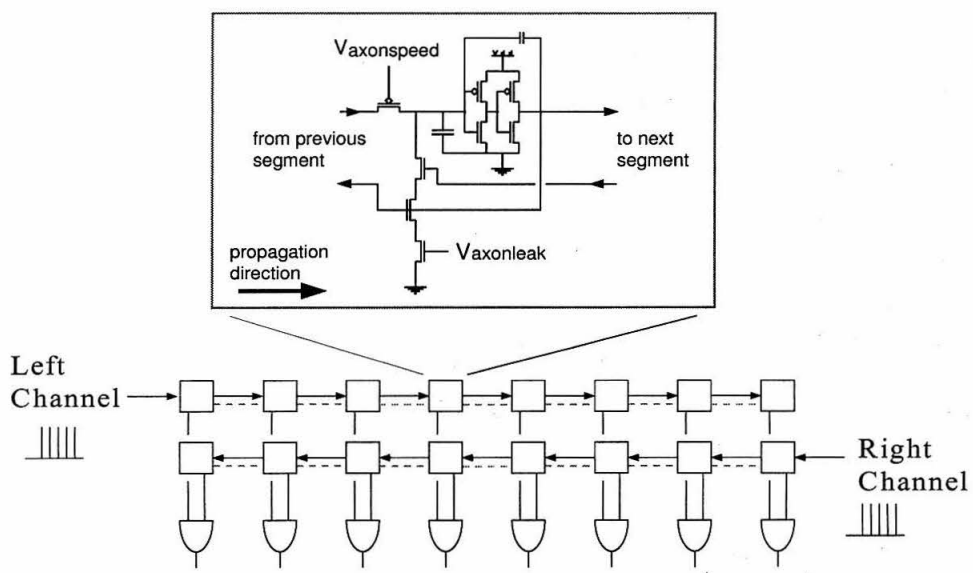

Figure 3.6: Diagram of the double axon delay line which accepts digital spikes on the left and right inputs and propagates them across the array. Whenever two spikes meet, a pulse is generated on the output AND units. The position of the activated AND circuit indicates the relative time of arrival of the left and right inputs. NOTE: the actual circuit contains 15 axon units.

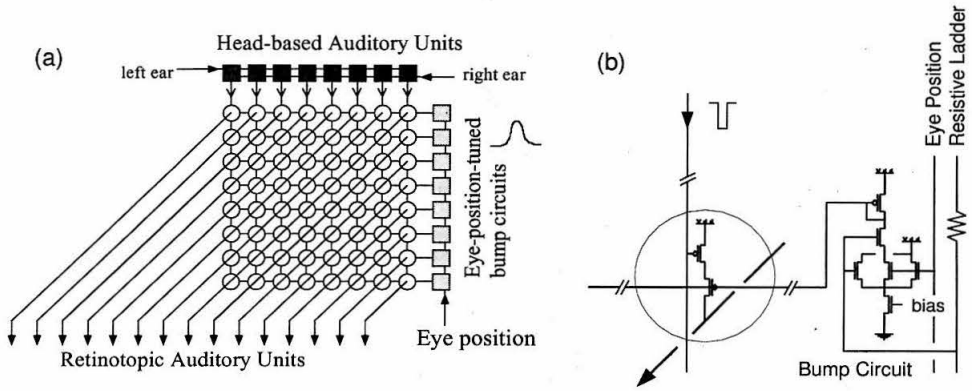

Figure 3.7: (a) Coordinate Transform Block Diagram: The cross-bar switching network shifts informaton coming from the axon delay-line correlator left or right, dependent upon which eye-position-tuned bump circuit is active. (b) Bump circuits on the right-hand edge of the array, each tuned for different eye positions, encode currents along its row which are switched onto the diagonal summing wire only at columns receiving correlation pulses from the delay line above.

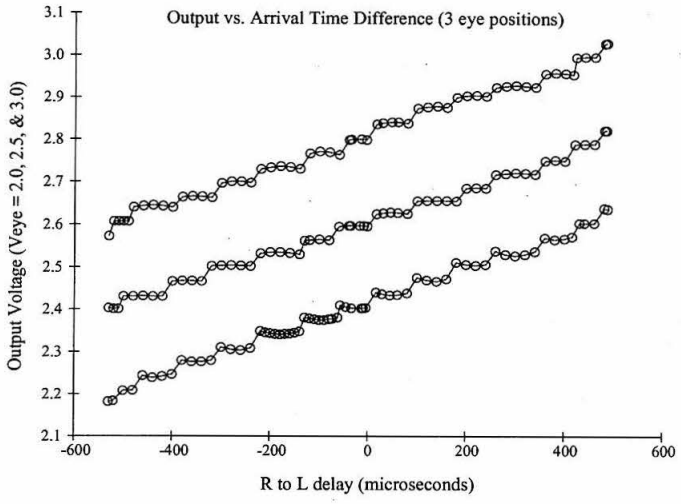

Figure 3.8: Chip output vs. input pulse timing: The chip was driven with a signal generator and the output voltage was plotted for three different eye position voltages. Due to the discretized nature of the axon, there are only 15 axon locations at which pulses can meet. This creates the staircase response.

3.2 Results and Conclusions

Figure 3.8 shows the chip's output voltage as a function of the inter-pulse time interval for three different eye positions (left, center, and right). This data was taken using artificially-generated pulses to test the chip's performance under ideal conditions. Figure 3.9 shows the performance of the system using real signals (a signal generator driving a speaker which was moved to different angles) processed through the microphones. The full system's output voltage are again shown for three different eye positions. To test the linearity of the coordinate transform, a fixed temporal difference, generated by a pulse-generator, was fed to the chip and the eye position voltage was varied. The output voltage was recorded for two different ITD values (Figure 3.10). The output is roughly linear with both azimuth and eye position (Figure 3.10).

The auditory localization system described here is currently in use with the analog VLSI-based model of the primate saccadic system to expand its operation into the auditory domain. The output voltage of the centroid circuit provides the burst generator input and a logical OR is performed on the two thresholded zero-crossing signals to provide the saccadic trigger.

By using an analog multiplexing chip, the burst generator chip operates on either the visual target, the auditory target, or the average position of the two. One of these three options is chosen by the state of the visual and auditory trigger inputs. Figure 3.11 shows a photo of the fully-integrated system.

While this particular implementation has focused on the integration of acoustic information with the saccadic system, this project demonstrates the practicality of neurobiologically-plausible models

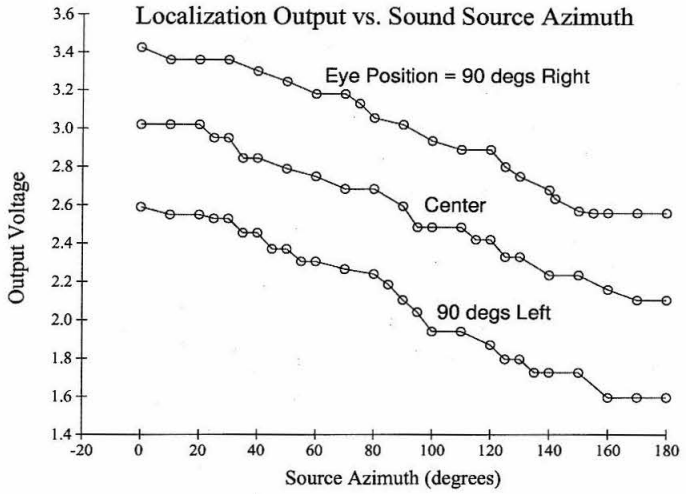

Figure 3.9: Performance of the full system on continuous sinusoidal input delivered by a speaker from different angles. Note that 90 degrees denotes the center position. Output plots for three different eye positions are shown.

of coordinate transform and multi-modal sensory convergence.

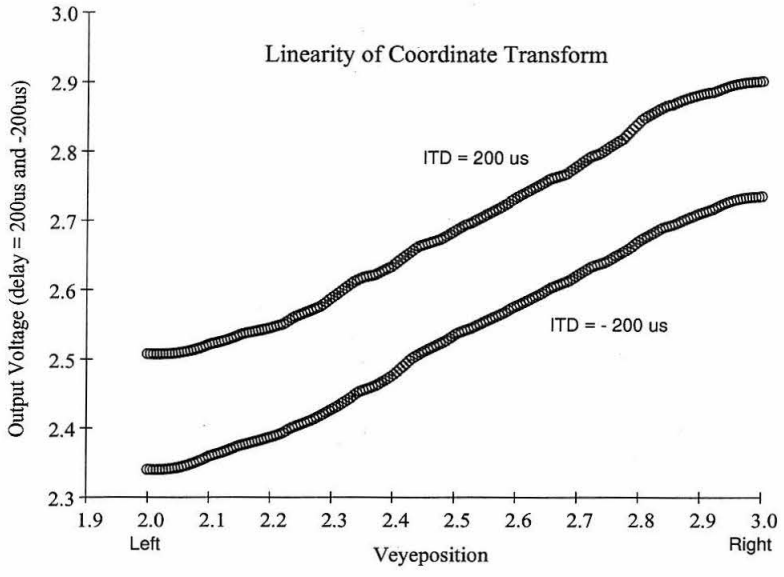

Figure 3.10: Coordinate Transform Linearity: The linearity of the shifting network was tested by using a fixed ITD input and sweeping the eye-position input voltage. Sweeps for two different ITD values are shown.

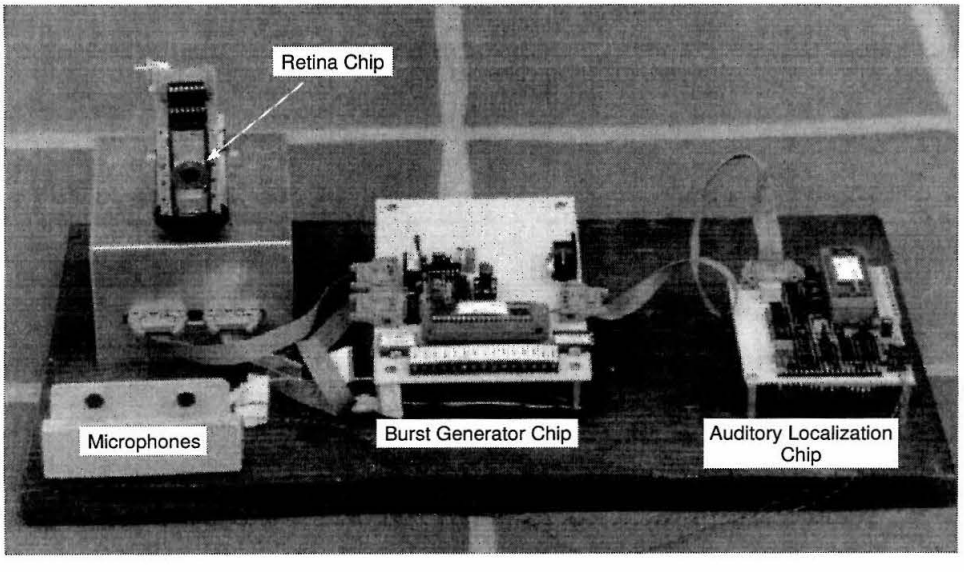

Figure 3.11: Photo of the visually and acoustically-triggered saccadic system.

Chapter 4 Motion Chips

Image motion is one of the most directly useful visual features in our environment. From the humblest insect to the not-so-humble human, visual motion detection systems are used to provide information about how our environment is changing around us, sometimes in response to our own actions and sometimes by the actions of other objects. We use it for detection of moving objects, guidance of our movements and stabilization, and even for discerning form and distance. Another use which is addressed in Chapter 5 is the use of motion as a training signal for learning motor control parameters.

While visual motion detection in biological systems is ubiquitous, successful analog VLSI implementations are not. Many attempts have been made to build them into integrated, smart sensors; only a few integrated motion detection chips have been generally successful. (For a full review, see Sarpeshkar et al., 1996.)

Most proposed short-range intensity-based motion detection schemes fall into two major categories: gradient models and correlation models. In gradient models, computation begins from local image qualities such as spatial gradients and temporal derivatives that can be vulnerable to noise or limited resolution. Correlation models, on the other hand, use a filtered version of the input intensity multiplied with the temporally delayed and filtered version of the intensity at a neighboring receptor. Many biological motion detection systems have been shown to use some type of correlation model (Grzywacz and Poggio, 1990).

Both gradient models and correlation models for computing short-range visual motion have been explored in analog VLSI technology with many designs focused upon computing the velocity at each pixel (Tanner and Mead, 1986; Sarpeshkar et al., 1993; Etienne-Cummings et al., 1993; Kramer et al., 1995; Kramer, 1996; Arreguit et al., 1996; Arias et al., 1996). In certain applications, however, the direction of the motion alone can be used, often resulting in a simpler circuit (Andreou et al., 1991; Benson and Delbrück, 1992; Delbrück, 1993; Moore and Koch, 1991).

In this chapter, three different motion circuits are presented: a pulse-correlation model, and two similar image-gradient models. The pulse-correlation method detects the time-of-passage of an image feature (edges) from one pixel to the next. The two gradient circuits utilize image intensity gradients and differ mainly in their implementation of a sensitive, direction-selective, motion model.

4.1 A Delay-Line Based Motion Chip

Inspired by biological motion detection models in the rabbit (Barlow and Levick, 1965) and by a computational architecture found in early audition (Konishi, 1986), a full-field motion detection chip has been designed that contains a large array of velocity-tuned "cells" that correlate two temporal events, using a delay-line structure. The chip contains a linear photoreceptor array with 28 elements and reports the dominant velocity detected in the field of view.

4.1.1 System Architecture

Figure 4.1 shows the block diagram of the chip. The input to the chip is a real-world image, focused directly onto the silicon via a lens mounted over the chip. The one-dimensional array of on-chip hysteretic photoreceptors (Delbrück and Mead, 1989) receives the light and reports rapid changes in the image intensity for both large and small changes. Each photoreceptor is connected to a half-wave rectifying neuron circuit (Lazzaro and Mead, 1989) that fires a single pulse of constant voltage amplitude and duration when it receives a quickly rising (but not falling) light-intensity signal.

This rising light intensity signal is interpreted to be a moving edge in the image passing over the photoreceptor. It is this signal that is the "feature" to be correlated. Note that the choice of the rising or falling intensity as a feature, from an algorithmic point of view, is arbitrary. Each neuron circuit is in turn connected to an axon circuit (Mead, 1989) that propagates the pulse down its length. By orienting the axons in alternating directions, as shown in Figure 4.1, any two adjacent receptors will generate pulses that will "race" toward each other and meet at some point along the axon delay-line. Correlators between the axons detect where pulses pass each other, indicating the detection of a specific time difference. The width of the pulses in the axon circuits is adjustable and determines the axon's speed (i.e., detectable velocity range). From the summing of "votes" for different velocities from correlators across the entire chip, a winner-take-all circuit (Lazzaro et al., 1989) determines the velocity.

4.1.2 Reading Between the Lines

The basic signal quantity that we are measuring is the time a "feature" takes to travel from one photoreceptor to one of its neighbors. By placing two delay lines in parallel that propagate signals in opposing directions, a temporal difference in signal start times from opposite ends will manifest itself as a difference in the location where the two signals will meet. Between the axons, correlation units perform a logical AND with the axon signals on both sides. If pulses start down adjacent axons with zero difference in start times (i.e., infinite velocity), they will meet in the center and activate a correlator in the center of the axon. If the time difference is small (i.e., the velocity is large), correlations occur near the center. As the time difference increases, correlations occur further out

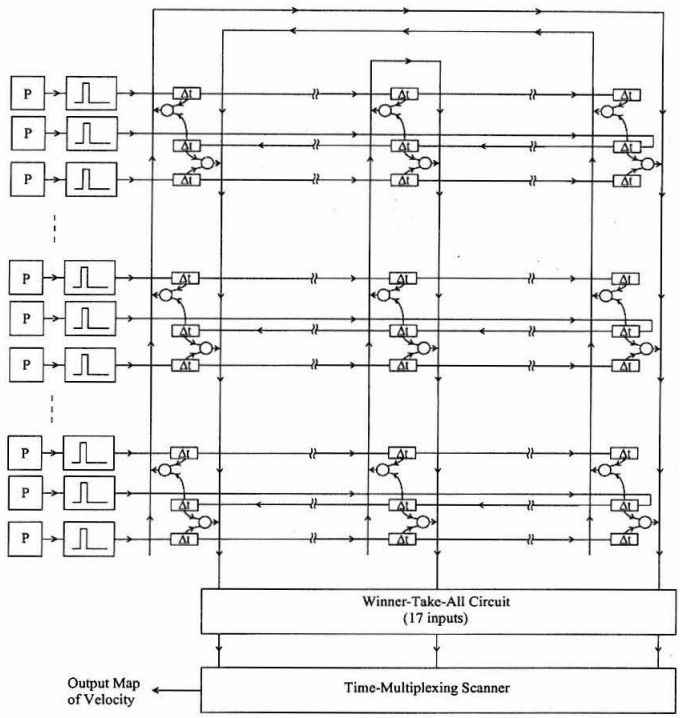

Figure 4.1: Block diagram of the chip, showing information flow from the photoreceptors (P), to the time-multiplexed winner-take-all output. Rising light signals are converted to pulses that propagate down the axons. Correlators are drawn as circles and axons are piecewise denoted by Δt boxes. See the text for explanation.

toward the edges. The two halves of the axon with respect to the center represent different directions of motion. When a single stimulus (e.g., a step edge) is passed over the length of the photoreceptor array with a constant velocity, a specific subset of correlators will be activated that all represent the same velocity. A current summing line is connected to each of these correlators which is passed to a winner-take-all circuit. The winner of the winner-take-all computation corresponds to the line that is receiving the largest number of correlation inputs. The output of the winner-take-all is scanned off the chip using an external input clock. Because the frequency of correlation affects the confidence of the data, scenes that are denser in edges provide more confident data as well as a quicker response.

4.1.3 Single Vs. Bursting Mode

Until now, the circuit described uses a single pulse to indicate a passing edge. Due to the statistical nature of this system, a large number of samples are needed to make a confident statement of the detected time difference, or velocity. By externally increasing the amplitude of the signal passed to the neuron during each event, the neuron can fire multiple pulses in quick succession. With an increased number of pulses travelling down the axon, the number of correlations increase, but with

a decrease in accuracy, due to the multiple incorrect correlations. The incorrect correlations are not random, however, but occur closely around the correct velocity. The end result is a net decrease in resolution in order to achieve increased confidence in the final data.

4.1.4 Velocity Range

The chip output is the measured time difference of two events in multiples of τ, the time-constant of a single axon section. The time difference (measured in seconds/pixel) is translated into velocity, by the equation $V = 1/\Delta t$, where V is velocity in pixels/sec and Δt can be positive or negative. Thus the linear measurement of time difference gives a non-linear velocity interpretation with the highest resolution at the slower speeds. At the slower speeds, however, we tend to have decreased confidence in the data due to the relatively smaller correlation frequency. This is expected to be less troublesome as larger photoreceptor arrays are used. The variable velocity resolution in the computation is often an acceptable feature for control of robotic motion systems since high velocity motions are often ballistic or at least coarse, whereas fine control is needed at lower velocities.

4.1.5 Performance

The chip has 17 velocity channels, and an input array of 28 photoreceptors. The output voltages from the winner-take-all circuit are scanned out sequentially by on-chip scanners, the only clocked circuitry on the chip.

In testing the chip, gratings of varying spatial frequencies and natural images from newspaper photos and advertisements were mounted on a rotating drum in front of the lens. Although the most stable data was collected using the gratings, both images sources provided satisfactory data. Figure 4.2 shows oscilloscope traces of scanned winner-take-all channels for 12 different negative and positive velocities within a specific velocity range setting. The values to the right indicate the approximate center of the velocity range. Figure 4.3(a) shows the winning time interval channel vs. actual time delay. The response is linear as expected. Figure 4.3(b) shows the data from Figure 4.3(a) converted to the interpreted velocity channel vs. velocity. The horizontal bars indicate the range of velocity inside of which each channel responds. As described above, at the lower velocities, correlations occur at a lower rate, thus some of the lowest velocity channels do not respond. By increasing the number of parallel photoreceptor channels, it is expected that this situation will improve. The circuit, currently with only eight velocity channels per direction, is able to reliably measure, over different parameter settings, velocities from 2.9 pixels/sec up to 50 pixels/sec.

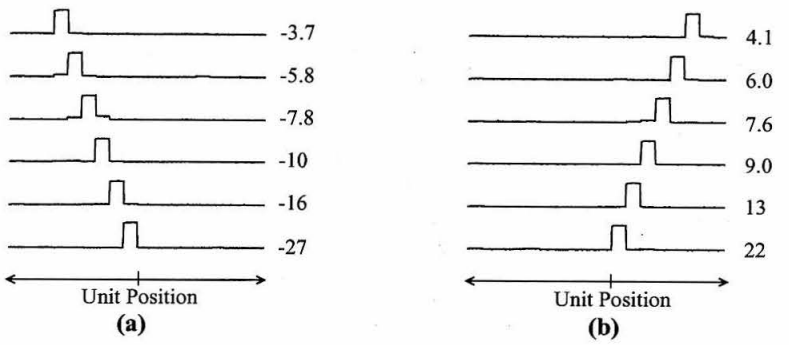

Figure 4.2: Winner-take-all oscilloscope traces for 12 different positive (a) and negative (b) velocities. The value to the right of each plot represents the approximate center of the velocity range.

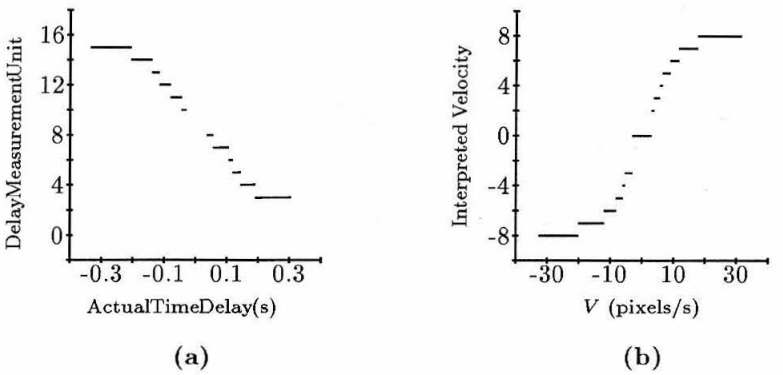

Figure 4.3: (a) Plot of winning time interval channel vs. actual time delay. (b) Plot of interpreted velocity channel vs. velocity (same data as in (a)).

4.2 Direction-Selective Gradient Models

In some tasks where the visual motion is used to determine an error signal for some adaptive sensorimotor control system, the direction of motion is often sufficient for guiding a learning process. The problem of reducing post-saccadic drift falls into this category. The particular task of correcting post-saccadic drift of the eyeball requires a visual motion signal that is sensitive to very low velocities and accurately reports the sign. The lowest detectable velocity will thus determine the lower limit of correctable drift.

If direction and sensitivity are the most important features of the motion detector required, then the *product* of the temporal derivative (I_t) and spatial derivative (I_x) gives an appropriate signal, since (for $I_x \neq 0$) $\text{sign}(\frac{I_t}{I_x}) = \text{sign}(I_t \cdot I_x)$.

In this section, two implementations of this multiplicative computation are presented. The first circuit is a voltage-mode circuit which computes the product directly and the second circuit is a

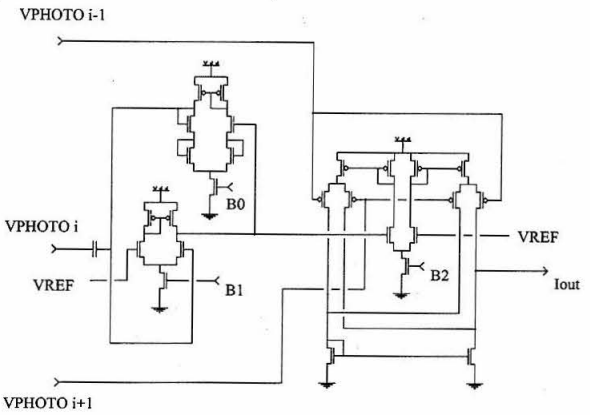

Figure 4.4: Voltage-mode circuit for computing the direction of visual motion. Iout = (temporal derivative) · (spatial derivative). The two transconductance amplifiers on the left are configured as a current-sense amplifier, providing the temporal derivative w.r.t. a reference voltage. This differential signal is fed into the Gilbert multiplier on the right. The other differential input comes from the two neighboring photoreceptors to provide the spatial derivative. Biases B1 and B2 are strong while B0 is set weakly and determines the sensitivity of the circuit.

current-mode circuit which computes a normalized product utilizing low-offset derivative circuits.

4.2.1 Voltage-Mode Direction Selectivity

Similar in aim to the analog VLSI-based circuit implemented by Moore and Koch (1991), the result is an output which is $I_x \cdot I_t$, or velocity$\cdot I_x^2$, for non-zero contrast and rigid motion in a one-dimensional image. Two benefits of this circuit are: 1) the lower limit of velocity is limited by the sensitivity and noise of the temporal derivative circuit, and 2) the output goes to zero as the image's spatial derivative goes to zero, thus avoiding the divide-by-zero problem in the gradient-model.

Figure 4.4 shows the circuit used to compute $I_t \cdot I_x$. See the caption for a description. A linear array of 20 elements has been fabricated on a TinyChip (2220μm x 2250μm) in a 2μm n-well CMOS process through the MOSIS Service. The motion circuit width on the chip is 63 μm per cell.

Since the circuit reports velocity$\cdot I_x^2$ the output depends on the stimulus used. Although the circuit computes the direction of motion at each pixel, it is possible to integrate information from across the entire array to obtain a better estimate of a full-field motion stimulus such as post-saccadic drift. In order to quantify the performance of the circuit, the motion outputs from across the array were summed and compared for different speeds. Figure 4.5 shows this data for a drifting grating stimulus which had three full cycles visible on the array.

There are two main problems with this detector: offsets and dynamic range. Offsets in the temporal derivative computation, generated in both amplifiers, create a non-zero output current even when there is no motion. The upper amplifier in the current-sense amplifier circuit nominally

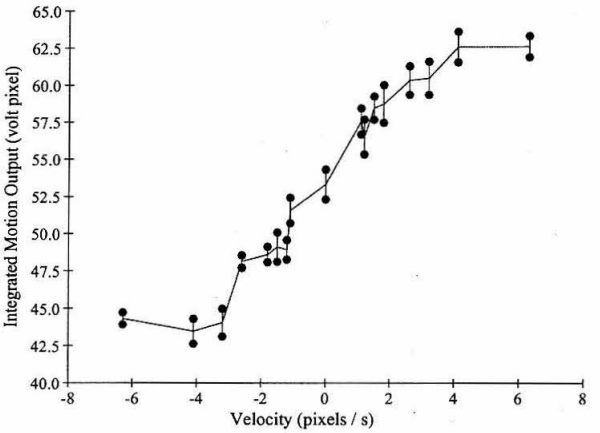

Figure 4.5: Voltage-mode Gradient Model: Spatially-integrated motion output signal vs. stimulus velocity. Note that the specific value of the integrated output is also dependent on the contrast and spatial frequency of the stimulus. The stimulus used was a high-contrast grating which fit 3 bars onto the 20 pixels of the array. The lowest discriminable speed was at +1.1 pixels/s. The vertical bars represent one standard deviation.

acts as a resistor when the voltage difference between the input and the output stays within 200 mV. If this range is exceeded, the input capacitor on the sense-amplifier is not properly held at the reference voltage Vref, creating a lag in the temporal derivative computation. This produces an incorrect direction output when the input's temporal derivative suddenly changes sign.

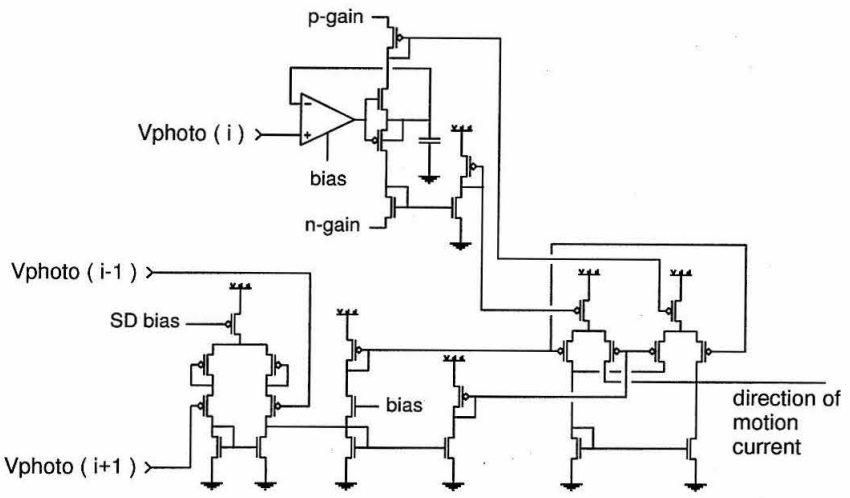

Figure 4.6: Current-mode circuit for computing the direction of visual motion.

4.2.2 Current-Mode Direction Selectivity

Adaptive photoreceptors (Delbrück, 1989) transduce the incoming pattern of light into an array of voltages. These appear in Figure 4.6 on the left side as the three voltages Vphoto(i-1, i, i+1). The temporal (TD) and spatial (SD) derivatives are computed from these voltages and are used to compute the direction of motion. In this chip, a more compact circuit than in section 4.2.1 was used in place of the multiplier. The direction-of-motion (DM) signal is now a normalized product of the two derivatives: $\frac{TD \cdot SD}{|TD|+|SD|}$. An example stimulus is shown in Figure 4.7.

Figure 4.8 shows the response of a single detector to the bidirectional passage of a grating. Each bar generates two pulses of output indicating the direction of each edge. The fast oscillations seen on the peaks of each output are due to the weak 120 Hz oscillations in the incandescent light-source used. To test the contrast-sensitivity of the detector, gratings of different contrasts were presented at a fixed speed and the peak output current during a 260ms window (4 output pulses) was plotted. Figure 4.9 shows these results.

As a comparison to the voltage-mode motion detector in the previous section, the output currents were integrated from across the entire chip and the same stimulus conditions were tested. Three cycles of a moving high-contrast grating were presented to the chip and the output current was averaged over a 500 ms. window. Figure 4.10 shows the data for the range from -6 pixels/s to 4 pixels/s. A close-up view of the data near zero motion reveals the system's ability to discriminate leftward and rightward motions from zero for velocities as slow as 0.06 pixels/s. This motion detector is considerably more sensitive than the voltage-mode gradient model due to its low-offset temporal derivative circuit. The signal-to-noise ratio is therefore considerably better when summing the outputs from across the chip. It should be noted that the stimulus used is nearly optimal for this

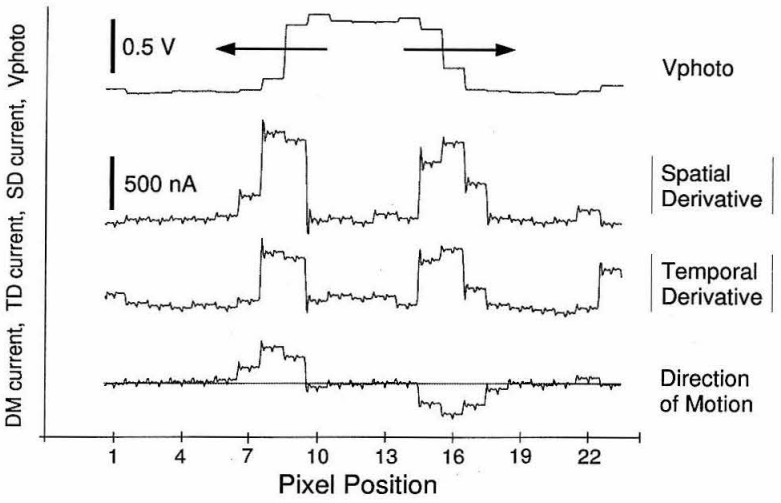

Figure 4.7: Current-mode Gradient Model: Example stimulus - Traces from top to bottom: Photoreceptor voltage, absolute value of the spatial derivative, absolute value of the temporal derivative, and direction-of-motion. The stimulus is a high-contrast, expanding bar, which provides two edges moving in opposite directions. The signed, temporal and spatial derivative signals (TD+, TD-, SD+, SD-) are used to compute the direction-of-motion shown in the bottom trace.

type of measurement and natural scenes would not be expected to generate such a low threshold of detection. Note also that at the lowest speeds tested, the adaptive photoreceptor circuit tended to adapt to the image, effectively reducing the contrast.

This motion detector has proven to be quite successful in determining the direction-of-motion and has been integrated into the smooth pursuit system described in Chapter 6.

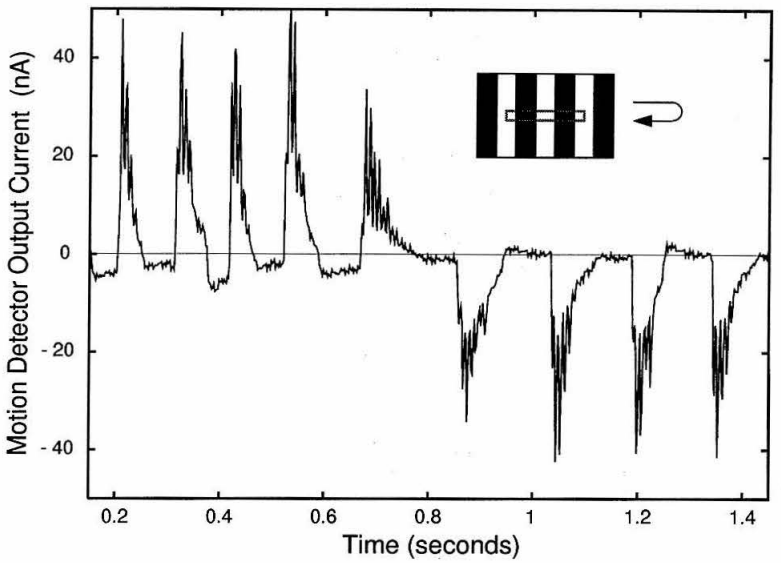

Figure 4.8: Current-mode Gradient Model: A single motion detector's response to the bi-directional passage of a medium-contrast grating.

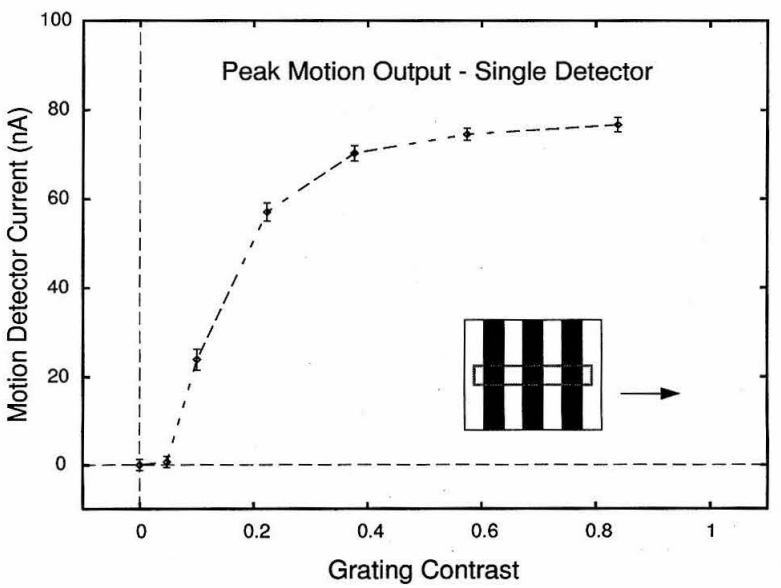

Figure 4.9: Current-mode Gradient Model: The contrast sensitivity of the detector was measured for a fixed velocity using a drifting grating stimulus. In this detector, for a fixed velocity, the contrast dependence should roughly be linear before the spatial derivative computation saturates.

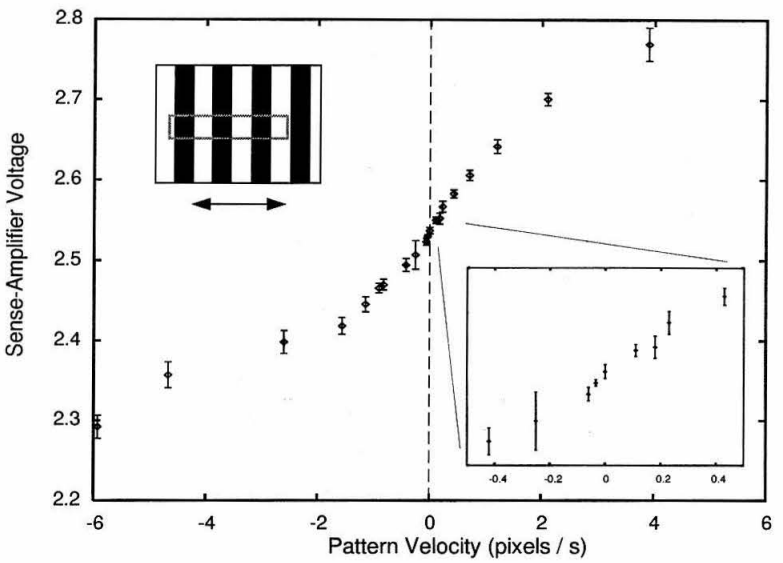

Figure 4.10: Current-mode Gradient Model: The full-field motion measurement was computed by summing the output currents from all of the motion detectors in the array. The lowest discriminable velocity was at 0.06 pixels/s. At the lowest speeds, only certain velocities were obtainable due to the stochastic nature of the stimulus generator used. The low-offset temporal derivative circuit in this model has improved the signal-to-noise ratio of the motion measurement.

Chapter 5 Adaptive Circuits

5.1 Introduction

Adaptation and learning are ubiquitous in neural systems. Through repeated experience in the world, nearly all animals modify their behavior on the basis of some type of memory. Unlike digital memory, neural memory is distributed throughout the computational architecture by many different mechanisms and on many different time scales. Memory is found in the charge stored on a neuron's membrane capacitance; in the concentration of free calcium in a neuron's dendrites, soma, and presynaptic terminals; in the density of available neurotransmitter vesicles; and, in the morphology of the dendritic tree and its connections to other neurons. While the design and management of such an extensive network of adaptive mechanisms makes its analysis difficult, the resulting system is able to operate under a wide range of conditions while maintaining optimal performance. For this reason, how and where adaptation is used in neural systems is of great interest to both biologists and engineers.

Neural systems are fast and efficient in both power and space because they use dedicated analog hardware which is well matched to the task. This approach, however, suffers from the vagaries of component mismatch and changes over time. The neuromorphic analog VLSI community has been able to mimic some of the power and space efficiency, but unlike the biological system, it has suffered from the inability to compensate for these inaccuracies with learning.

The ability to adapt is not simply an added feature of neural systems, it is a fundamental mechanism which drives the development of the brain. The use or disuse of a neuron leads either to physical growth or degeneration of neural connectivity. The keys to understanding why neurons organize in a certain way or which features are important to the computation may lie in understanding the role that adaptation and learning plays in neural systems.

From an engineering viewpoint, the analog VLSI model of the primate saccadic system described in previous chapters is a multi-chip system which continues to grow as we improve its capabilities. The number of parameters in this model has been growing rapidly, highlighting the need for automatic parameter adjustment. Additionally, component mismatch and real-world nonlinearities which have always plagued analog VLSI processing are ripe for adaptive techniques to eliminate their effects on signal precision. Utilizing known neural representations and physical embodiment, our system is ideal for investigating the system-level effects of integrating biologically-realistic adaptation at multiple locations and time scales.

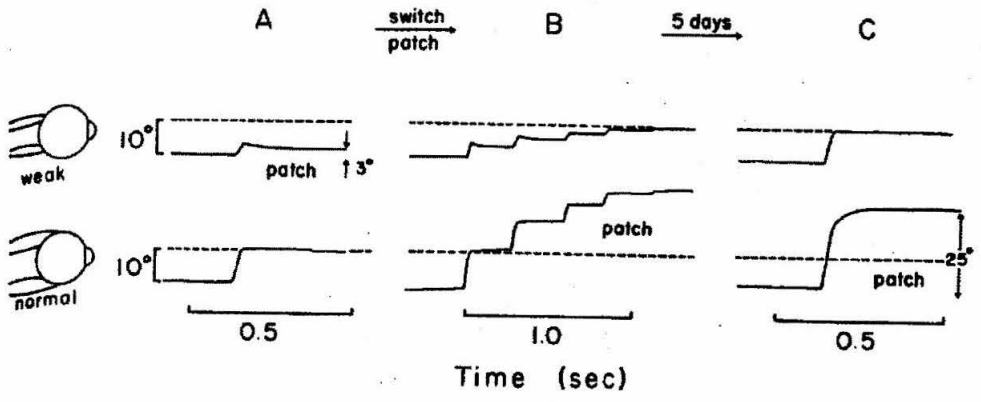

Figure 5.1: From Optican and Robinson (1980): Saccadic adaptation in the monkey GS. The effect of tenectomy is shown in A. The normal eye makes a 10 deg saccade to the target (dashed lines). The weak eye makes a hypometric saccade with postsaccadic drift. Immediately after the patch is switched to cover the normal eye in B, saccades in the weak eye are still hypometric with drift and the monkey must make a staircase of saccades to get on target. C: effects of five days of visual experience with the weak eye. The saccade in the weak eye now is essentially accurate and without drift. Note also that in this experiment the adaptation was expressed in both eyes because the normal eye was patched and did not have any visual feedback.

The first section describes two types of adaptation seen in the primate saccadic system that have been implemented. In the second section, previous work on adaptation in neuromorphic analog VLSI systems is discussed. The third and fourth sections describe in detail, two adaptation projects which address the issues of eliminating post-saccadic drift and learning to trigger saccades accurately.

5.1.1 Adaptation in the Primate Saccadic System

During saccades our eyes often reach speeds up to 750 deg/sec during which our visual acuity is severely impaired. In fact, retinal velocities exceeding even 3 deg/sec can significantly degrade our performance. It is therefore important to minimize the time during which the eye is moving. While typical human saccades last between 40 msec and 150 msec, changes or damage to either the oculomotor plant or the underlying neural circuitry can cause significant errors in accuracy and eye drift following a saccade, lengthening the period of poor acuity. Changes in the optics can also create saccadic inaccuracies, requiring multiple saccades to acquire the target. Primates demonstrate an impressive ability to adapt to such physical changes and restore their original performance. This chapter deals with several examples of how this adaptation might take place within the neurobiological models we have been describing

There have been many types of adaptive behavior identified in the primate oculomotor system in response to various induced deficits. For example, Optican and Robinson (1980) showed that weakening of the horizontal recti muscles in the rhesus monkey caused saccades which fell short and

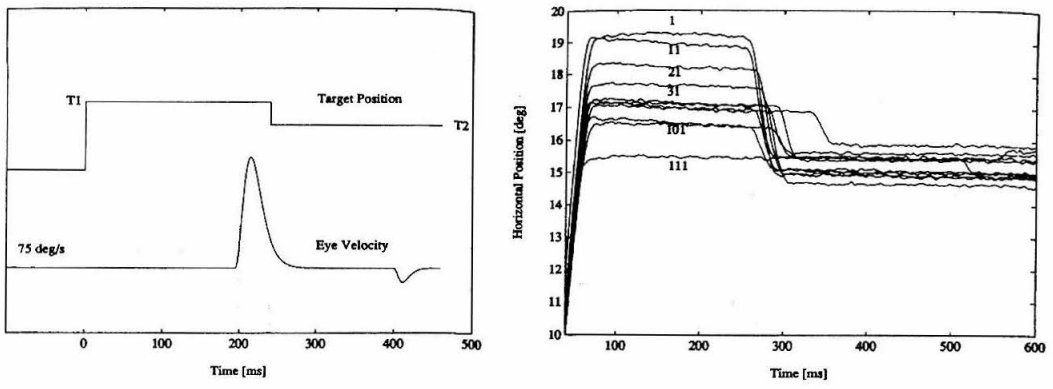

Figure 5.2: From Frens and van Opstal (1994): (Left) The adaptation paradigm. While a human subject fixates the central position, a target appears in the periphery at T1. The goal-directed eye movement serves as a trigger for the displacement of this target to position T2, 40 msec after the velocity of the saccade exceeds a threshold of 75 deg/sec. (Right) Gradual decrease of saccadic amplitude. Typical eye position traces measured during the gain-shortening paradigm. Numbers indicate trial number in which the saccades were made.

exhibited post-saccadic drift of the eyeball (See Figure 5.1). Recovery from this type of damage, which affects all saccades in a given direction, requires about 3-5 days of practice with visual feedback.

In contrast to this long adaptation period, which involves hundreds of thousands of saccades, experiments in which saccadic targets within one part of the visual field are moved a short distance during the saccade show that adaptation to this type of perturbation can occur within several hundreds of trials. See Figure 5.2. This type of visually-induced modification of the saccade amplitude is known as short-term adaptation. Further, experiments by Frens and van Opstal (1994) show this adaptation to be confined to a limited range of saccade vectors around the adaptation target.

5.1.2 Adaptation in Neuromorphic Analog VLSI

Previous work on analog VLSI modeling of neural adaptation has focused on the use of integrated capacitors due to their ease of use. The junction-leakage from the connected circuitry, however, is significant and limits its retention time to a few seconds, particularly for analog parameters. Models of spike rate adaptation in neurons (Mahowald and Douglas, 1995; Boahen, 1996; Lazzaro, 1992; Lazzaro et al., 1991), photoreceptor adaptation (Mahowald, 1992; Delbrück and Mead, 1989), time-constant adaptation (Liu and Mead, 1996), and habituation (Wang and Akers, 1995) have all been modeled using capacitors for analog memory. While integrated capacitors are capable of holding their charge for many seconds, adaptation on longer timescales (minutes and upwards) requires the use of a longer term analog memory.

Floating-gate structures in VLSI (a MOS transistor gate completely insulated from the circuit by

silicon dioxide) provide an extremely effective charge storage technique with retention measured in years. Until recently, however, the use of floating-gates required the use of either ultra-violet (UV) radiation (Glasser, 1985; Mead, 1989; Kerns, 1993) or bidirectional tunneling processes (Carley, 1989; Lande et al., 1996) to modify the charge on the floating node, both of which have significant drawbacks, impeding their widespread use. The recent development of a complementary strategy of tunneling and hot-electron injection (Hasler et al., 1995; Diorio et al., 1995) in a commercially-available BiCMOS process has alleviated some of these difficulties. In addition, adding and removing electrons from the floating-gate can be performed at extremely low rates, making it possible to create very long time constants with small capacitances.

One of the first system-level uses of the floating-gate structure for analog parameter storage in a neuromorphic design was in the adaptive retina (Mead, 1989) where the goal was offset removal. This system used UV light to adapt offsets during an explicit calibration step. Since then, a number of researchers have also used the floating-gate structures to correct for fixed pattern noise in imaging chips (Devos et al., 1993). Lazzaro et al. (1994) used floating-gate circuits to store analog parameters for his cochlear modeling. Parameters were loaded onto the chip via a computer. In addition, many artificial neural network chips have also been constructed which utilize the floating-gate storage technique for analog weight storage (Castro et al., 1993; Kramer et al., 1996; Berg et al., 1996). More recently, Hasler et al., 1995 have demonstrated arrays of very small, single-transistor Hebbian-learning synapse circuits.

5.2 Reduction of Post-saccadic Drift

In the biological oculomotor plant, the eye is suspended in the eye socket both by suspensory tissues and the eye muscles which create an overdamped spring-mass system. In order to hold the eyes off-center, it is therefore necessary to provide a sustained pulling force on the muscles to counter the elastic properties of the muscles which would passively return them to the center position. In addition, quick eye movements require a large, transient, driving force to overcome the long time-constant of the system due to the viscosity in both the muscles and the eye socket.

Since the sustained (or tonic) component of the command determines the final eye position, the ideal transient (or burst) component should bring the eye to exactly that position by the end of the burst. Mismatch of the burst and tonic components leads to either an onwards or backwards drift following an undershoot or overshoot of the final eye position. This motion is known as *post-saccadic drift*. Studies in both humans and monkeys show that these deficits can be compensated for by some type of learning process which have time constants on the order of 1.5 days. Ablation studies (Optican and Robinson, 1980) have further shown that control of the burst and tonic gains is independent and that their control depends on different areas of the cerebellum. In addition, retinal

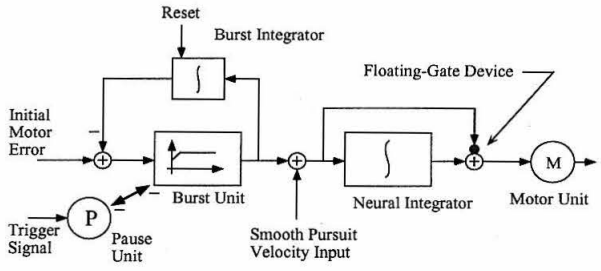

Figure 5.3: System block diagram of the burst generator, neural integrator, motor unit, and pause units. This system generates the motor signals used to drive the motors in the model oculomotor plant. Initially, the neural integrator is providing a tonic signal to the motor unit to hold the eye in its starting position. The initial motor error input to the system (far left) is applied and a trigger signal disinhibits the burst unit via the pause unit. The burst integrator activity (starting from zero) rises and is compared against the input activity to drive the burst unit. When the difference goes to zero, the pause unit shuts down the burst. Downstream, the burst unit provides the burst component and is also integrated via the neural integrator to provide the tonic component. These two signals are combined at the input to the motor unit. The parameter being trained in this system is burst gain. The burst gain is the weighting between the (burst + smooth pursuit) output and the input to the motor unit. The gain parameter is stored in the floating-gate device.

slip was shown to be necessary and sufficient to elicit these adaptive changes (Optican and Miles, 1985).

This section describes a system which reduces the post-saccadic drift in the hardware model based on detected post-saccadic visual motion. In the first section, the floating-gate device and how it is used for on-chip storage in the burst generator model is presented. The second section demonstrates the system's capability of improving its performance through repeated trials.

5.2.1 The Resettable-Integrator Burst Generator Chip

As described earlier in Chapter 2, the burst generator model used in this system (See Figure 5.3) is a double integrator model based on the work by Jürgens et al. (1981), MacKenzie and Lisberger (1986), and Nichols and Sparks (1995) that uses initial motor error as input to the system and requires a reset period following the saccade. The basic operation of the model is described in Figure 5.3. The saccadic motor command consists of the tonic component (that determines the final eye position) and the burst component (that provides the initial acceleration). These two components must be balanced to provide the shortest duration eye movement. Figure 5.4 shows eye movement traces for three example saccades on the analog VLSI system with different burst gains.

5.2.2 Non-volatile Analog Voltage Storage

In this system, the burst gain parameter (the connection weight between the burst unit and the motor unit) is stored in a floating-gate device acting as a current reference (Hasler et al., 1995). The

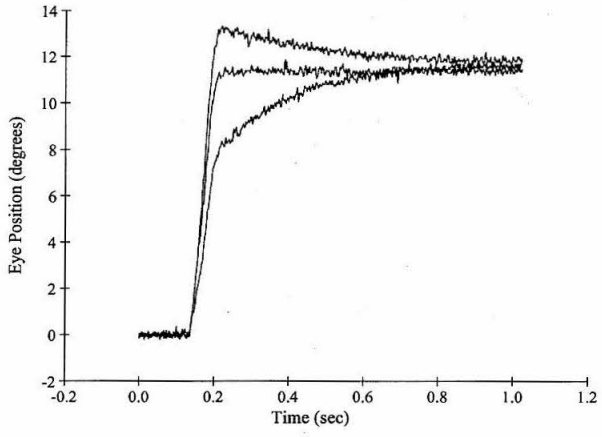

Figure 5.4: Eye movement traces for three small saccades with the same input command and tonic component gain but with different hand-tuned burst component gains. The upper trace shows an example of an overshoot condition caused by excessive gain and the lower trace shows an example of an undershoot condition caused by insufficient gain. Note that the overshoot case generates a backward post-saccadic drift and the undershoot case generates an onward post-saccadic drift. The middle trace is an example of a properly set gain. This step shape is what the adaptation system is trying to achieve.

storage circuit used in this system is shown in Figure 5.5. This device utilizes high-voltage tunneling in order to remove electrons from the floating node and hot-electron injection to put electrons onto the floating node. The tunneling occurs through a high-quality gate oxide ($\sim$ 420 Åthickness) between a well-region and the PolyI layer floating node. Usable tunneling currents for this circuit occur with approximately 27 volts across these two nodes. The hot-electron injection process is performed using a p-base substrate transistor. This heavily doped n-type MOSFET has a threshold voltage of approximately 6 volts, allowing the gate voltage to be higher than the drain voltage for subthreshold current levels. As the drain voltage is increased, the field in the drain-to-channel depletion layer increases, accelerating more electrons in the channel to the required 3.1 eV energy level necessary to jump through the gate oxide onto the gate.

The tunneling was operated by driving the high-voltage node from 25 volts up to 33 volts, generating a tiny pulse of current across the tunneling oxide. The injection was controlled by moving the source-follower transistor gate from 1.50 volts up to 4.28 volts. The source-follower technique allows control over the drain voltage (VD) of the injection transistor which, in turn, modulates the rate of injection. The standard pulse length of both the injection and tunneling processes was 10 seconds. Adaptation was performed between saccades, following evaluation of the motion circuit.

5.2.3 Training Results

To demonstrate the learning, the direction-selective motion detector chip described in Chapter 4,

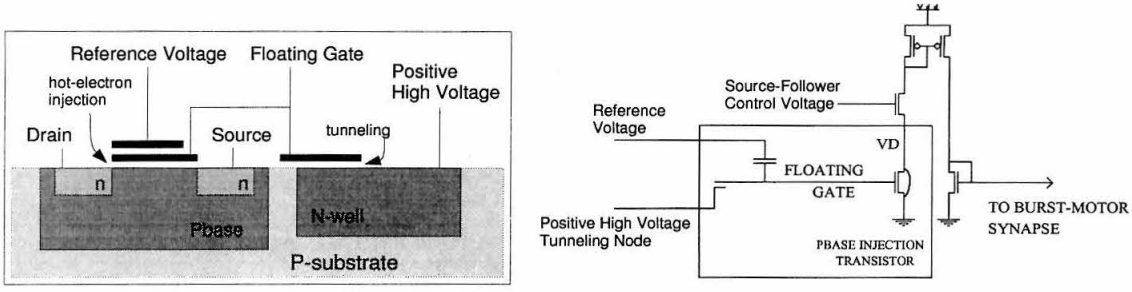

Figure 5.5: The floating-gate circuit used for storage of an analog current reference. The floating node potential is raised by tunneling electrons off of the node with high voltage and lowered by injecting electrons onto the node via the pbase injection transistor. The hot-electron injection process is modulated by controlling the drain voltage (VD) with a source-follower transistor. The current flowing through the mirror is used as the output.

section 4.2.1, was mounted on the one-dimensional eye and motion information was read from the chip 100 msec after the end of the saccadic burst activity. The burst activity period (lower trace in Figures 5.6 and 5.7) is detected by reading a signal representing the suppression of the pause circuitry on the burst generator chip. A standard leftward saccade amplitude of about 23 degrees was programmed into the burst generator input and a saccade was repeatedly triggered. The motion sensor was facing a stationary stripe stimulus which would elicit a motion signal during and after the saccade burst.

The direction-of-motion information was summed from across the motion detector array and a simple, fixed-learning-rate algorithm was used to determine which direction to change the gain. One hundred msec after each trial saccade, the motion detector output current was compared against two threshold values. If the output value was greater than the rightward motion threshold, indicating overshoot, a unit hot-electron injection pulse was issued which would reduce the floating-gate voltage and thus reduce the burst gain. If the integrated value was less than the leftward motion threshold, indicating undershoot, a unit tunneling pulse was issued which would increase the floating-gate voltage and thus increase the pulse gain. At present, the activation of both tunneling and injection pulses is performed manually; the criterion for the direction of the adaptation pulse is strictly based on the scalar value reported by the motion detection chip and could be automated.

Figure 5.6 shows an experiment where the pulse gain was initialized to a value near zero. Within eight trials (not all traces are shown), the pulse gain was raised sufficiently to eliminate the post-saccadic drift. Figure 5.7 shows a similar experiment where the pulse gain was initialized to a large value. In this case, 41 trials (not all traces are shown) were required before the gain was reduced sufficiently to eliminate the post-saccadic drift. The difference in learning rates is not important and can easily be balanced.

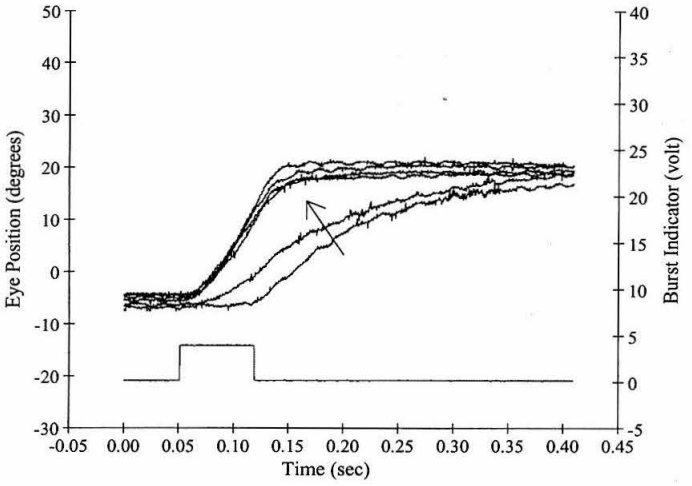

Figure 5.6: Saccade trajectories showing the reduction of an onward post-saccadic drift by increasing the burst gain via a tunneling process. See text for description. The lower digital trace indicates the time of burst unit activity. The arrow indicates the progression from early saccade trials to later saccade trials where performance has improved.

5.3 Learning Saccadic Accuracy

In contrast to the systematic changes in saccadic accuracy across the entire visual field discussed above, position-specific offsets can also be tested. In these experiments, the target is moved *during* the subjects' saccadic eye movement. In humans, Frens and van Opstal (1994) have shown that the adaptation timecourse to learn the offset is short (requiring only a few hundred presentations) and that the adaptation is confined to a limited range of saccade vectors around the target. This type of learning can be explained by a mapping similar to that of a look-up table.

5.3.1 Vector-Specific Adaptation

In Chapter 2, visual stimuli were mapped linearly from pixel position to motor command in a functional model of the deep layers of superior colliculus. Any non-linearities in the optics, photoreceptor triggering circuit, burst generator, or motor plant would create errors in proper programming of the saccade. A modified visual-triggering circuit from Chapter 2 has been used to use the output of a floating-gate circuit to determine the proper motor command for each pixel.

As shown in Figure 5.8, an array of adaptive photoreceptor circuits (P) are used to drive a temporal-derivative circuit (TD), activating circuits where the image intensity is changing. These temporal derivative signals trigger three circuits: one which activates a slowly decaying memory of which units have been active (U/D), another which drives a centroid circuit (C) to map the pixel's position to a motor command voltage, and finally a triggering circuit which compares the total activity on the chip to a threshold (not shown). The trigger circuit provides a binary output signal

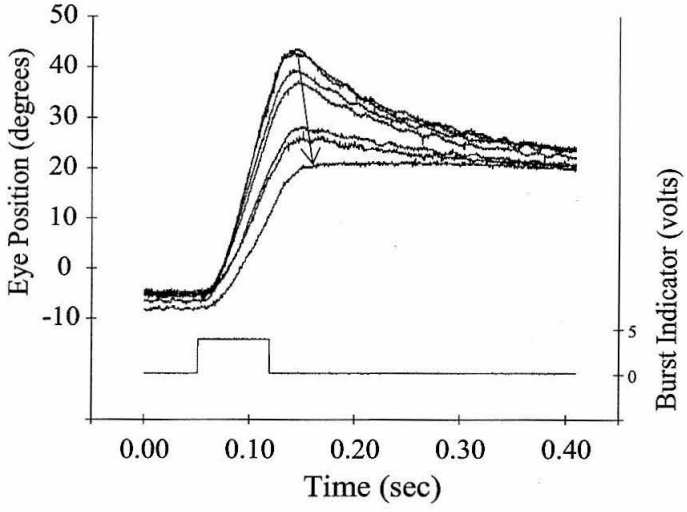

Figure 5.7: Saccade trajectories showing the reduction of a backward post-saccadic drift by decreasing the burst gain via a hot-electron injection process. See text for description. The lower digital trace indicates the time of burst unit activity. The arrow indicates the progression from early saccade trials to later saccade trials where performance has improved.

from the chip, indicating that something has occurred in the image and that the centroid output information is "valid." The centroid circuits (DeWeerth 1992) require reference voltages (motor command voltages) at each pixel which represent the saccade vector required to center the stimulus on the array.

In previous versions of this visually-based, triggering circuit, the motor command voltages were provided by a resistive line running across the array. Each end of the resistive line was held at a different voltage, providing each pixel in the array with a unique voltage reference, which changed linearly across the array. In contrast, the pixels in this new system are provided with the output voltage of a floating-gate circuit, each of which can, in principle, be set to arbitrary values, making it similar to a programmable look-up table.

The training input to the system is a global signal indicating whether the system's output was too high or too low. Pixel locations which contributed to the output remain active for a short amount of time (about 3 sec) via the U/D circuit. When the training signal becomes active, after evaluating the centroid output voltage, only those units which contributed to the output are trained in the appropriate direction. Since the triggering stimuli may activate a neighborhood of pixels, the learning is similar to Kohonen's stochastic learning algorithm where the topology of the network is preserved by training a node and its neighborhood at the same time. This technique has been explored in software in the context of saccadic learning by both Ritter et al. (1992) and by Rao and Ballard (1995).

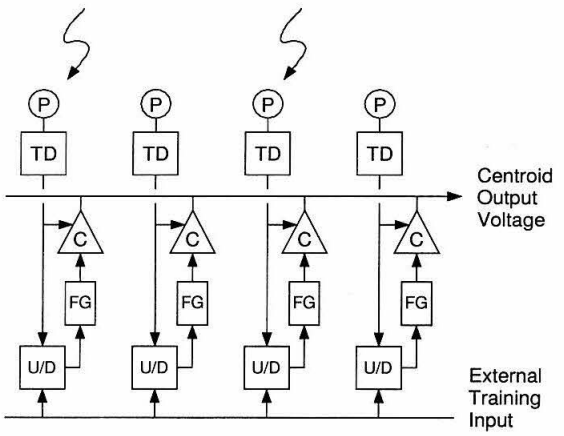

Figure 5.8: System Block Diagram: This chip consists of an array of 32 pixels which consist of an adaptive photoreceptor (P), a temporal derivative circuit (TD), a centroid circuit (C), a floating-gate circuit (FG) which provides reference voltages to the centroid circuit, and a control circuit (U/D = "up/down") for training the floating-gate.

The training system consists of a workstation which flashes visual stimuli (bars) at different locations on its monitor. The chip, with a lens, is positioned to image the stimuli on its photoreceptor array. The centroid output voltage is measured after each flash using a computer-controlled oscilloscope. Each stimulus position on the monitor is assigned a target centroid output value. If the measured value is lower than the target value, the training input voltage (driven by a computer-controlled voltage source) is lowered to a pre-determined training voltage for a fixed amount of time to increase that stored value by increasing the tunneling rate. Similarly, if the measured value is higher than the target value, the training input is raised. After repeated trials, a target function can be learned to a level of accuracy limited primarily by the system noise.

5.3.2 The Temporal-Derivative Triggering Circuit

The implementation of the architecture described above was fabricated on a TinyChip (2.25mm x 2.22mm) using a 2.0 μm, n-well, double-poly, BiCMOS process. The chip contains a one-dimensional array of 32 pixel elements.

Figure 5.9 shows the combined circuit schematics for the adaptive photoreceptor (P) (left), the temporal-derivative (TD) (middle), and the centroid circuit (C) (right). The adaptive photoreceptor (Delbrück and Mead, 1989) is a high-gain photoreceptor circuit which slowly adapts to the average light level to prevent saturation. The temporal derivative circuit combines a lowpass filter with a "bump" circuit (Delbrück, 1991) to signal the absolute-value of the temporal-derivative. The centroid circuit (DeWeerth, 1992) computes the weighted-average, motor command voltage. Since every cell in the array would connect to an n-type mirror, the gray box in the figure denotes the

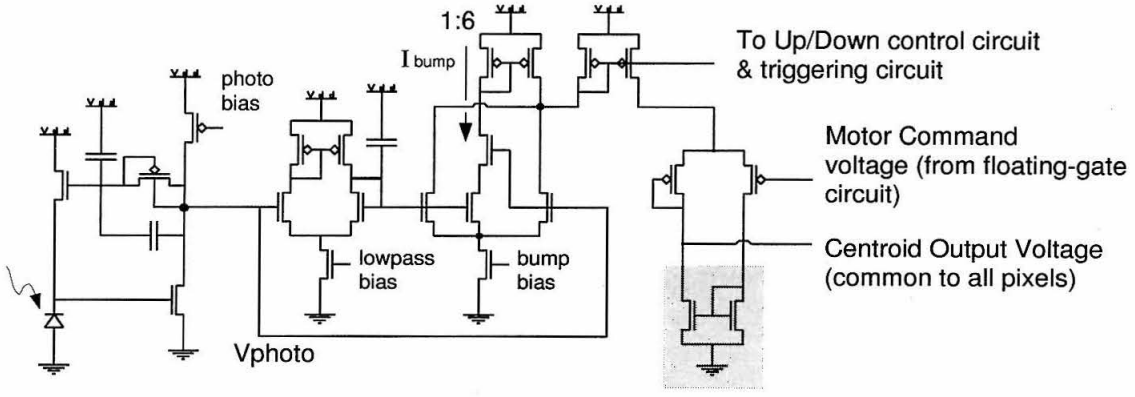

Figure 5.9: Temporal Triggering Circuit (P+TD+C): On the far left, the adaptive photoreceptor circuit amplifies temporal change in the light intensity while slowly adapting to the mean light level. The temporal-derivative (TD) circuit acts as a high-pass filter by measuring the difference in voltage between the original photoreceptor value and a low-passed version of it. The signal is then full-wave rectified and mirrored to the U/D, centroid, and thresholding circuits. The centroid circuit (on the right) operates as a follower powered by the current from the temporal derivative circuit. The motor command reference voltage is received from the floating-gate amplifier circuit (FG).

use of a single, common mirror on the edge of the array to reduce capacitance on the output node. An amplifying ratio of 6 to 1 was used on the mirror for inverting the bump current to cancel the tail currents of the differential pair. Overall, these circuits map the retinotopic location of temporal change to a motor command voltage.

5.3.3 Non-Volatile Analog Voltage Storage

The floating-gate circuit (Figure 5.10) is a modification of the circuit used by Hasler et al. (1995) to train a 2x2 array of floating-gate synaptic elements. A tunneling process is used to remove electrons from the floating node and a hot-electron injection process is used to put electrons onto the floating node. The tunneling current is controlled by manipulating the difference in voltage between the floating-node and the high-voltage tunneling line. Larger voltage differences produce larger tunneling rates. Injection of electrons is performed in an n-type transistor fabricated in the Pbase layer provided for the construction of bipolar transistors. Due to the heavier doping, the threshold voltage for this type of transistor is near 6 volts, which allows the gate to capture high-energy electrons flowing through the drain while the transistor is still operating in the subthreshold. Since the injection current is the product of the injection efficiency (controlled by the drain voltage) and the source current, injection current can be adjusted by manipulating the source current in the Pbase transistor.

The floating-gate circuit (Figure 5.10) uses two Pbase transistors, one used as an electron injector (PB1) and the other used as the current source for the amplifier (PB2). Since PB2 is only setting

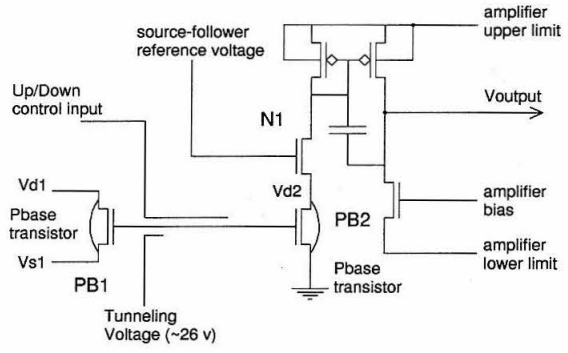

Figure 5.10: Floating-gate amplifier circuit (FG): The floating node defines a subthreshold current in transistor PB2 which is mirrored and used in a high-gain amplifier stage which has variable output limits. Source-follower transistor N1 defines PB2's drain voltage to prevent hot-electron injection. Nodes Vd1, Vs1, and the high-voltage tunneling node are fixed global values which define an equilibrium floating-gate value, and a decay rate towards this value. Modification of the floating-gate voltage is performed by capacitively moving the floating-gate up or down transiently to either increase injection or increase tunneling.

the amplifier current (and not injecting), its drain voltage Vd2 can be set to a low voltage allowing the upper limit of the amplifier's output range to be fairly large. Modification of the floating-gate charge is performed by transiently increasing the rate of either the tunneling or injection. This is performed by capacitively raising or lowering the floating-gate using the Up/Down control input. Raising the floating-node both increases the source current in PB1 and reduces the floating-gate to tunneling voltage. Likewise, lowering the floating-node both increases the floating-node to tunneling voltage and decreases the source current in PB1.

As in the system described by Hasler et al. (1995), the tunneling and hot-electron injection currents are both active, but extremely low and in opposite directions. Since both processes operate in a negative-feedback fashion (e.g., the tunneling process raises the floating-gate which tends to reduce the rate of tunneling), the system reaches an equilibrium value when the tunneling current equals the injection current. When the floating-gate voltage is larger than the equilibrium voltage, the hot-electron injection current dominates the tunneling current and the floating-gate voltage drops. Conversely, when the floating-gate voltage is lower than the equilibrium voltage, tunneling dominates and the voltage rises.

While this technique avoids high-voltage switching circuits, it suffers (or possibly benefits) from the eventual loss of stored information as the floating-gate decays back to its equilibrium voltage. This decay rate, however, can be set to be extremely slow by using a low Vd1 (transistor PB1) and a low tunneling voltage. Since the tunneling and injection parameters are kept constant, the equilibrium voltage should not depend on the stored value and the memory should decay towards an equilibrium determined solely by these parameters. Memory decay tests of the floating-gate circuits,

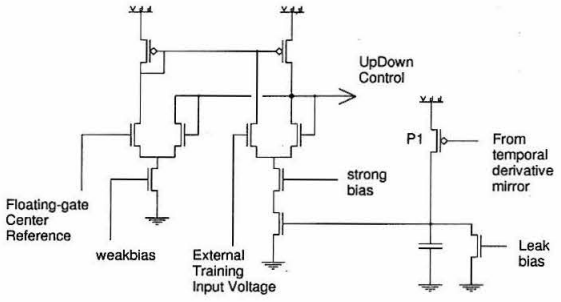

Figure 5.11: Up/Down learning control circuit (U/D): This circuit consists of two competing voltage followers: a weak follower carrying the center reference voltage and a stronger follower which receives the training voltage from off-chip. When a given pixel in the array generates a pulse of current in the TD circuit, this current is mirrored onto transistor P1, charging the capacitor node up towards Vdd. A small leak current discharges the capacitor slowly. This node acts as a switch to turn on the strong amplifier to drive the floating-gate control node towards the globally-received, training voltage. In this fashion, only those circuits which participated in generating the output centroid voltage receive the training signal.

with the parameters used for data collection, exhibited low tunneling-dominant rates (less than 0.07 mV/hour), while the injection-dominant rates showed a decay of about 1.0 mV/hour. For more details of the physics of these floating-gate devices, see Hasler et al. (1995) and Diorio et al. (1995)

The "learning" can also be turned off by bringing Vd1, Vs1, and the tunneling voltage down to zero. Unfortunately, the absolute voltage level of all the floating-gates will be DC-shifted downwards as the tunneling voltage drops due to capacitive coupling. This shift can easily be countered by increasing the U/D circuit's center reference voltage until the values have returned to their trained state. This step, however, may introduce a DC shift error since it is done manually.

To train the chip for a certain mapping, pixels are stimulated and the resultant centroid output voltage is determined to be either too high, too low, or inside a window of tolerance around the target value. Since the pixels which contributed to the output value are the ones that need to be modified, some mechanism is required to remember those pixels. The Up/Down circuit shown in Figure 5.11 performs this function by storing charge at each pixel location that contributed to the centroid output. If the pixel has not been active, the circuit holds the output to a global reference voltage. If the pixel was just used to drive the centroid output, the U/D circuit drives the output to an externally-provided voltage level for approximately five seconds (with our current leak settings). This external signal is the training voltage which is used to increase or decrease the floating-gate voltages at those locations which contributed to the previous output.

Figure 5.12 shows some of the relevant signals during a pulse of light at one pixel on the array. Although not visible in this plot, the rising centroid output reaches a stable value approximately 2 msec after the beginning of the temporal change.

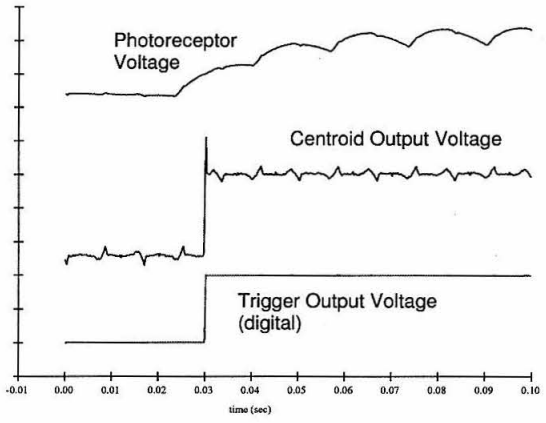

Figure 5.12: Top trace: Photoreceptor voltage, Middle trace: Centroid output voltage (analog), Bottom trace: trigger signal (digital). The photoreceptor output voltage jumps from 0.96 volts to 1.30 volts during the flash of the stimulus. The oscillation riding on the step response of the photoreceptor is due to the flicker induced by the monitor. The centroid circuit also shows some 60 Hz noise, resulting from feed-through of noise from the high-gain floating-gate circuits.

The data was taken using a tunneling voltage of about 26 volts, Vd1 = 3.1 volts, Vs1 = 0.2 volts with the floating-gate values centered around 5.5 volts. The Up/Down control line was moved from 4.0 to 7.0 volts for increased hot-electron injection and from 4.0 to 0.0 volts for increased tunneling. The coupling coefficient between the U/D control line and the floating-gate was measured to be about 0.6. In order to scan off the floating-gate values, we operated the chip using a Vdd of 8 volts.

5.3.4 Training Results

In training, the chip is aimed at a computer monitor which flashes vertical bars at different positions in the field of view. While the current chip has only 32 pixels, the training system flashes stimuli at the maximum line resolution of the screen. Our current optics configuration allows for approximately 75 different locations at which we can stimulate the array of 32 pixels. This is done both to map the subpixel behavior as the stimulus moves from one pixel location to the next and to train the pixels individually rather than as groups of pixels.

In real-world situations, however, the pixels will be activated in groups and the subsequent output will be an appropriate average of the individual pixel values. Although training the system with large stimuli does work, the training time dramatically increases since the training must rely on the uniform statistics of the training set to sort the proper values out. The training stimulus size also sets the minimum size for which the array will report the proper value. For this reason it is important to also train at the appropriate resolution. A multi-resolution training schedule may be the best strategy since training can occur in parallel, yet the smaller stimuli can fill in the details at

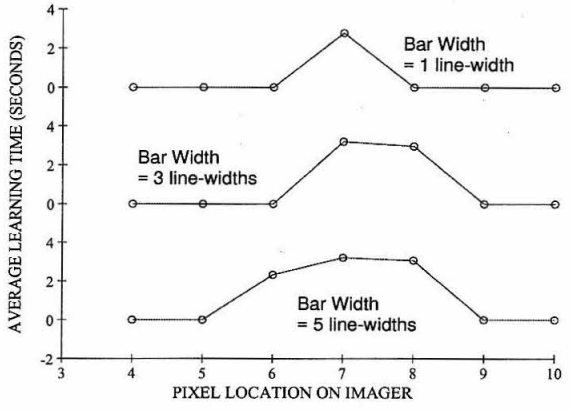

Figure 5.13: When a bar of one line-width (defined by the graphics board) is flashed at the chip, it stimulates a single photoreceptor as shown in the top plot and the one pixel is trained for a mean duration of 2.75 seconds. This timing is primarily determined by the leak bias (see Figure 5.11). When the bar is widened to three line-widths (middle plot), two adjacent pixels are stimulated and they are trained together in the same direction. A bar width of five line-widths stimulates three pixels as shown in the bottom plot. In the multi-scale training regime, all three types of bars were used randomly interleaved in the training set. The bar of five line-widths was also used to generate Figure 5.17. These plots show the results of measuring the mean time each pixel spent training for bars of different widths flashed at a position on the monitor near pixel #7. The mean was computed over seven trials.

each position. The training positions are chosen by shuffling a list of positions and selecting them from the list without replacement. Once the list is exhausted, the whole list is reshuffled. This sets an upper bound on the inter-example training time and guarantees a uniform distribution.

After training, the array can be "probed" with either a bar of one line-width or a bar of five line-widths to stimulate output values. The one line-width bar will stimulate individual pixels and the 5 line-width bar will stimulate the average of a group of three pixels. (See Figure 5.13.) The effects of averaging can be seen in Figure 5.17 for the case of the sinewave mapping, which is a particularly difficult case to learn, since individual pixels cannot satisfy the wide range of values occurring on a steep part of the function.

The first test of system level operation we discuss is an experiment in which we attempt to load a flat target function. With this function it is easiest to see the accuracy with which the system can learn a specific value. Figure 5.14 shows the results after extended training. From initial conditions where the floating-gate amplifier outputs were sitting at fairly random voltages, the system was presented approximately 20,000 examples at 75 different stimulus locations (approximately 625 examples per pixel) and then the system was probed at the 75 stimulus locations to evaluate the mapping. Noise in the chip and in the testing system contribute to the variations seen in repeated trials. It should be noted that the floating-gate amplifiers are non-linear and the highest gain occurs in the center of the range. Since the target value for the flat function in Figure 5.14 is in the center

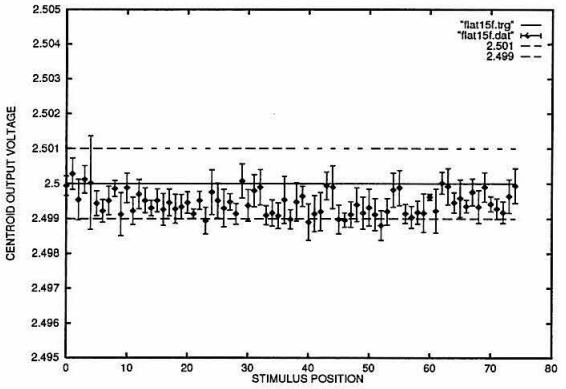

Figure 5.14: Flat Target Function - In this case, all stimulus positions were trained to lie at 2.500 volts. This plot shows the performance of the chip after approximately 20,000 presentations spread over 75 positions. The floating-gate outputs were initially spread between 2.4 and 2.6 volts. After training, the centroid array was "queried" sequentially from left to right five times without training. The error bars represent one standard deviation. The training procedure continued to modify the floating-gate until the voltage was within 1 mV of the target voltage.

of the range, we expect the largest reporting variance here due to noise. The error tolerance of the training system for this mapping was 1 mV.

The linear target function (Figure 5.15) is the mapping which was previously used to map retinal position to motor command, where 2.60 volts represented a full-scale saccade to the right and 2.40 volts represented a full-scale saccade to the left. In this case and in the following mappings, the error tolerance for learning was 2.5 mV.

In order to challenge the system, we also tested a sinewave target function (Figures 5.16 and 5.17) whose spatial derivative was difficult to match with the resolution of the current system. The expected final value in this situation when training with a uniform distribution of examples and balanced step sizes is the average of the different target values associated with the same pixel. This behavior is seen most clearly in Figure 5.16. Convergence of this mapping function takes much longer due to the statistical nature of the equilibrium, and the final value is not very stable since nearly all the training examples drive the pixel away from its current value.

During the testing process, it was determined that modifications should be made to reduce the gain of the floating-gate output amplifier. The measured DC gain from the floating-gate to the output of the amplifier was found to be approximately 60. This created many problems with noise, particularly at 60 Hz due to electrical noise in the laboratory and the 60 Hz light flicker coming from the monitor. This problem was partially solved by using a considerably smaller output voltage range (2.4 volts to 2.6 volts) to push the amplifier's output transistors partially out of saturation for the subthreshold current regime. This had the effect of reducing the gain down to about 2.0, but left a very small signal range with which to work.

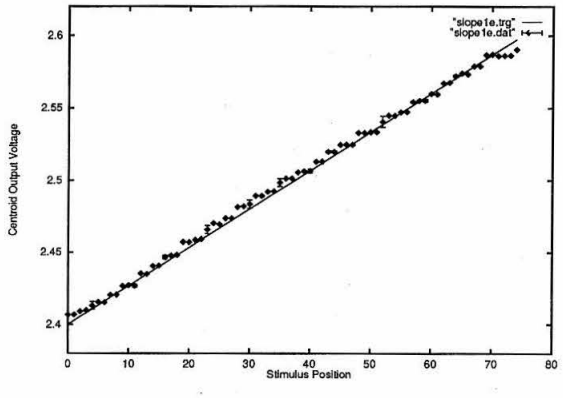

Figure 5.15: Linear Target Function - This function most closely represents a realistic sensorimotor mapping function for triggering saccades to a visual target. The training and testing procedure is the same as in the previous graph. The error bars represent one standard deviation.

By storing information locally about which units contributed to a computation, the distribution of the training signal back through the system has been made simpler. The hardware approach to this problem of delayed assignment-of-error may provide a valuable testbed in which to consider how this problem is solved in biological systems.

The neurobiological substrate for this adaptation is still unknown. Both the superior colliculus and the frontal eye fields are attractive areas for investigation of this adaptation due to their vector-specific organization for driving saccadic eye movements. While both areas are capable of driving of saccadic eye movements, the frontal eye fields are implicated in the generation of "volitional" saccades and the superior colliculus has been implicated in the generation of reflexive, visually-guided saccades. Experiments by Deubel (1995) indicate that there are context-dependent differences in vector-specific, short-term adaptation. Adaptation performed during reflexive, visually-guided saccades was not expressed during volitionally-driven saccades. The converse has also been found to be true. Frens and van Opstal (1994) also demonstrated the transfer of vector-specific adaptation to saccades triggered by auditory cues. These experiments together point to the interpretation that the adaptation is occurring at a stage after integration of these different sensory modalities, but before the parallel streams of information from the superior colliculus and frontal eye fields have converged.

5.3.5 Discussion

In the coming years, we will be switching to a 1.2 μm process for prototyping circuits. Initial tests (C. Diorio, personal communication) have indicated that the pbase injection transistor does not operate with the same desirable characteristics as the same structure in the 2.0 μm process. A

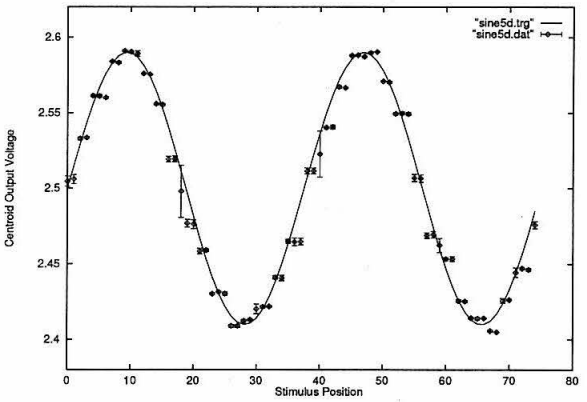

Figure 5.16: Sinewave Target Function - In this case, the target values followed a sinewave. Photoreceptor granularity is evident by the "staircasing" seen in the plot. Stimulus locations where the flashed bar occurs on the boundary of two pixels exhibit large variations in output voltage due to the narrow (one line-width) stimuli being used. Figure 5.17 shows the same pattern being probed with a much wider stimulus (three line-widths). The training and testing procedure is the same as in the previous graphs. The error bars represent one standard deviation.

good alternative to using the pbase transistor for injection is to use the standard p-type MOSFET transistor for injection.

The intermingling of memory and computation is an important and powerful aspect of neural architectures which has not yet been well exploited in neuromorphic analog VLSI designs. Smaller designs have been manageable by the use of external sources of parameters or by array structures which share global parameters. With the advent of large, multi-chip, neural systems, however, the automatic selection, storage, and maintenance of these parameters will become an unavoidable issue as it is in biological systems.

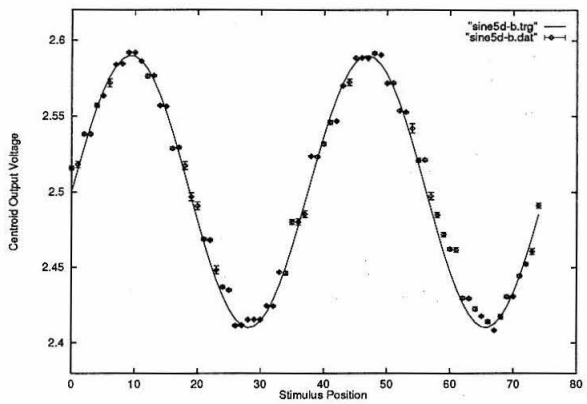

Figure 5.17: Sinewave Target Function - In this case, the evaluation of the pattern in Figure 5.16 was performed using a bar which spanned three pixels. The training and testing procedure is the same as in the previous graphs. The error bars represent one standard deviation.

Chapter 6 Attention-Based Tracking

Given that we see our world through foveated retinas, it is remarkable that our perception of the world is both stable and in sharp focus. Our eyes are constantly moving, gathering information relevant to our current task. When we examine static scenes, we typically trigger saccades, fixating sequentially on different parts of the image. When we search for a familiar face in a moving crowd, we select candidate faces and track them smoothly to stabilize their image for recognition. Eye movements are an integral part of the way we perceive objects in the world and are more than just another way to shift our gaze. For this reason it is important to understand how eye movements are controlled and how they contribute to perception.

This chapter describes a one-dimensional tracking chip which incorporates a saliency-based model of attention to select a target with the goal of driving eye movements. Chapter 7 describes the use of this chip in performing eye movements.

6.1 Attention and the Control of Eye Movements

Visual attention is the process of focusing sensory and computational resources on a spatially-limited part of the visual field. Visual attention is often discussed as two components, the *overt* and the *covert* systems. The overt visual attention system physically moves the position of highest acuity (eye movements) and the covert visual attention system internally controls the location of spatially-focused image analysis. While we typically keep the two attentional systems together at the center of gaze, our covert attention can be moved anywhere within the field of view at the cost of reduced resolution but with much shorter latency than with eye movements.

While visual attention is best understood in the context of visual search during fixation, relatively little research has been devoted to the understanding of how attentional selection is involved in the control of eye movements. While these are typically studied separately, in practice it is often difficult to treat tracking independently from the selection process since the tracking problem involves the continuous selection of the target.

Visual search has been heavily studied using experiments designed to discriminate between search tasks that require attention and those that do not. Search times which are fast and independent of the number of distractors ("pop-out") generally occur in tasks where there are differences in only a single feature type, while linearly increasing search times as a function of the number of distractors ("serial search") appear most often in tasks where the conjunction of two features is sought. The linear search time is believed to be due to the requirement of sequential attentional fixations. It

has been suggested by Treisman and Gelade (1980) that attention serves to bind together various features of an object, such as color, form, and motion as in the conjunction task. In addition, experiments by Posner (1980) showed that the performance of an observer in the simple detection of a signal changes as a function of where the observer is instructed to attend. These experiments have suggested that we are able to preferentially enhance the processing of a spatially-limited region of the scene, at the cost of perceiving the full scene one piece at a time.

In recent years, many studies have also indicated that selective visual attention is involved in the generation of saccadic (Kowler et al., 1995; Hoffman and Subramaniam, 1995; Rafal et al., 1989; Shimojo et al., 1995) and smooth pursuit eye movements (Khurana and Kowler, 1987; Ferrera and Lisberger, 1995; Tam and Ono, 1994). These studies have shown that attentional enhancement occurs at the target location just before a saccade as well as at the target location during smooth pursuit. In the case of saccades, attempts to dissociate attention from the target location has been shown to disrupt their accuracy and latency. It has been proposed that attention is involved in programming the next saccade by highlighting the target location. For smooth pursuit, attention is believed to be involved in the extraction of the target's motion. Studies have shown that the smooth pursuit system is driven by visual motion in a negative feedback loop (Rashbass, 1961). If the motion-sensitive, middle temporal area of the cortex (MT) is involved in driving pursuit, as several studies have shown (Lisberger et al., 1987), some mechanism must exist to preferentially extract the activity of neurons associated with the target at the correct place and time. Indeed, strong attentional modulation of motion information has recently been demonstrated in areas MT and MST (Treue and Maunsell, 1996).

Neurophysiological studies have revealed a wide variety of brain areas underlying visual attention. Many areas in cortex, superior colliculus, and a large part of the thalamus called the pulvinar all seem to be involved in various aspects of visual attention, as discerned by single cell, lesion, and imaging studies. Several studies have shown modulation of neural activity in favor of attended stimuli (Moran and Desimone, 1985; Motter, 1993; Treue and Maunsell, 1996). In these studies, attentional cues can significantly modulate a neuron's response to a stimulus in one part of its classical receptive field over another.

6.1.1 Models

Many neurally-plausible computational models have been proposed for the various mechanisms of selective attention. Koch and Ullman (1985) proposed a model of attentional selection using an image saliency map by combining the activity of elementary feature maps in a topographic manner. The most salient locations are where activity from many different feature maps coincide, or at locations where activity from a preferentially-weighted feature map, such as temporal change, occurs. A winner-take-all (WTA) mechanism, acting as the center of the attentional "spotlight," selects the

location with the highest saliency. Olshausen et al. (1993) proposed a neurobiological model for the dynamic routing of visual information from the attended location. This model also provided a mechanism for position and size invariance. Niebur et al. (1993) proposed a different mechanism for routing information in a spike-based model using a "temporal-tagging" strategy which modifies the timing of the spikes in the spike train, without changing the average spike rate. This entrainment allows the attended neurons' spike trains to preferentially pass through subsequent attentional filter stages. In contrast, Desimone and Duncan (1995) have proposed an attentional model which does not utilize any explicit saliency map, rather a distributed one.

6.1.2 Hardware Modelling

While the machine vision and robotics communities have implemented many systems which utilize moveable cameras, very few systems have been used to model biological active vision. None have attempted to replicate the mechanical plant or the neural control of these eye movements.

Recently, however, Etienne-Cummings et al. (1996) have built a two-dimensional analog VLSI-based visual tracking chip which performs pursuit movements using focal-plane motion detectors in the central region to drive stepper-motors for tracking. The periphery of the chip is used to trigger target acquisition. This system does not have a dynamic mechanism for figure-ground discrimination of the target.

In addition to the overt form of attentional shifting, covert attentional selection and tracking has been modelled by DeWeerth and Morris (1995) using analog VLSI circuits based on the Koch and Ullman WTA model. These circuits also demonstrate the use of delayed, long-lasting inhibition in the saliency map at the selected location to model the phenomenon of inhibition-of-return (Shimojo et al., 1995).

In order to extend the visually-triggered oculomotor system beyond the reflexive, averaging saccades described in Chapter 2, an attentional model was incorporated into the imaging array to allow the selection of a single target for saccades as well as to extract motion information for smooth pursuit.

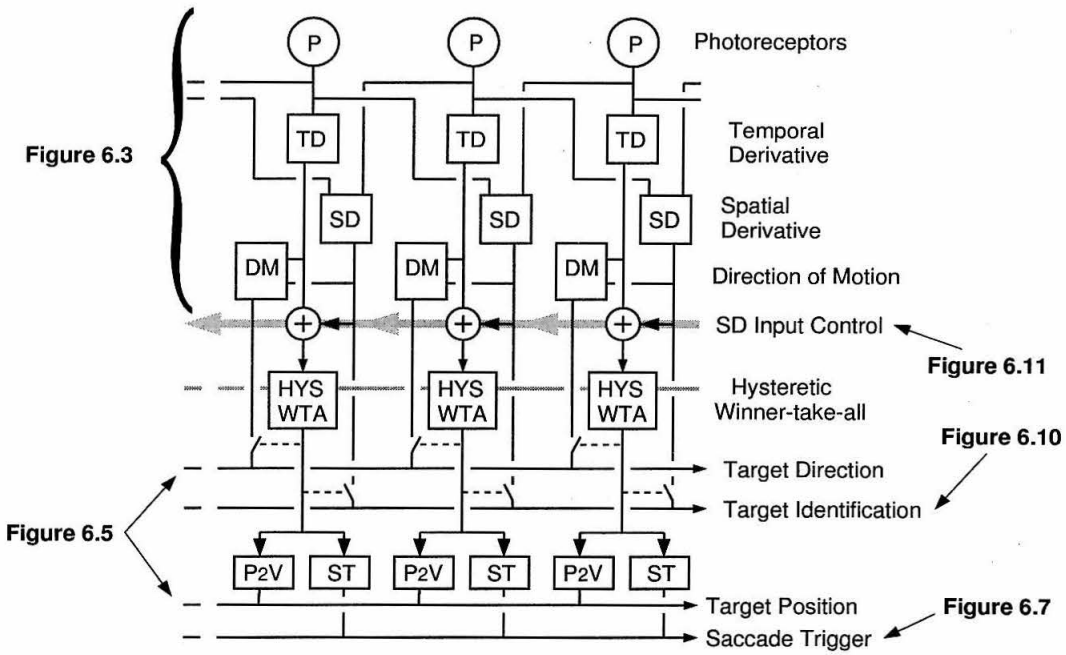

Figure 6.1: System Block Diagram: P = adaptive photoreceptor circuit, TD = temporal derivative circuit, SD = spatial derivative, DM = direction of motion, HYS WTA = hysteretic winner-take-all, P2V = position to voltage, ST = saccade trigger. The TD and SD signals are summed to form the saliency map from which the HYS WTA finds the maximum. The output of the HYS WTA steers both the direction-of-motion and the SD information onto global output lines. The HYS WTA also drives the P2V and ST circuits to convert the winning position to a voltage and to indicate when the selected pixel is outside a specified window located at the center of the array. The SD input control modulates the relative gain of the positive and negative spatial derivatives used in the saliency map. See the text for details.

6.2 An Attention-based, Visual Tracking Chip

Utilizing an analog VLSI model of selective visual attention (Morris and DeWeerth, 1996), this chip incorporates focal-plane processing to compute image saliency and to select a target feature for tracking. The target position and direction of motion are reported as the target moves across the array, providing control signals for tracking eye movements.

The computational goal of the attentional tracking chip is the selection of a target, based on a given measure of saliency, and the extraction of its retinal position and direction of motion. Figure 6.1 shows a block diagram of this computation. The first few stages of processing compute the saliency map from simple feature maps which drive the WTA-based selection of a target to track. The circuits at the selected location signal their position, the computed direction-of-motion, and the type of target being tracked. The saccadic system uses the position information to foveate the target and the smooth pursuit system uses the motion information to match the speed of the target.

Chip Name	Motion Detector	Hysteresis	Target ID & SD Input Control	WTA Input Smoothing	Pixel Count
TRACK1	No	Resistive	No	No	23
TRACK2	Yes	Nearest Neighbor	No	No	21
TRACK3	Yes	Nearest Neighbor	Yes	No	21
TRACK4	Yes	Nearest Neighbor	Yes	Yes	23

Table 6.1: Tracking chip versions and their capabilities.

All of the chips appearing in this chapter were fabricated as TinyChips (2220 μm x 2250 μm die size) through the MOSIS silicon brokerage service using the Orbit 2.0 μm CMOS, n-well, double-poly process. Four versions of the tracking chip have been fabricated (see Table 6.1) and their data appear in this chapter. At the end of each figure caption, the chip from which the data was taken is indicated. While TRACK4 is clearly the most sophisticated and includes nearly all features from the earlier versions, some of the earlier data was not retaken for presentation.

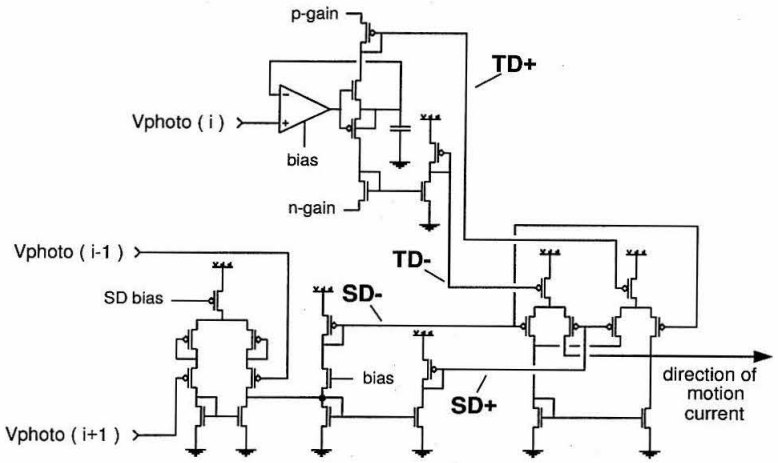

Figure 6.2: Circuits used to compute the spatial and temporal derivatives. The signals TD+, TD-, SD+, SD- are used in Figure 6.4 to generate the saliency map. Note: This figure is the same as Figure 4.6 in Chapter 2.

6.2.1 Chip Description

Adaptive photoreceptors (Delbrück 1993) (at the top of Figure 6.1) transduce the incoming pattern of light into an array of voltages. The temporal (TD) and spatial (SD) derivatives are computed from these voltages (Figure 6.2) and are used to generate the saliency map and direction of motion. Figure 6.3 shows an example stimulus and the computed features. The saliency map is formed by summing the absolute-value of each derivative (|TD| + |SD|). The direction-of-motion (DM) circuit, as described in Chapter 4, computes a normalized product of the two derivatives: $\frac{TD \cdot SD}{|TD|+|SD|}$.

Figure 6.4 shows the circuitry used at each pixel for creating the saliency map and performing the attentional selection. The inputs TD+, TD-, SD+, and SD- encode the temporal and spatial derivative currents computed in Figure 6.2 which are regenerated and summed to produce the saliency map. In the saliency map, the signed spatial derivatives can be differentially weighted to emphasize one type of edge over another (SD input control). This weighting is performed by the two differential pairs controlled by the *Vref* and *Vctrl* voltage inputs. The saliency map provides the input to a winner-take-all (WTA) circuit (transistors M1 and M2) which finds the maximum in the saliency map across all the pixels in the array. Spatially-distributed hysteresis is incorporated in this winner-take-all computation (DeWeerth and Morris, 1995) by adding a fixed current to the winner's input node (transistors M3 and M4) and to its neighbors (transistors M5 and M6). This nearest-neighbor, distributed hysteresis (NN-hysteresis) is motivated by the following two ideas: 1) Once a target has been selected, it should continue to be attended even if another equally interesting target comes along, and 2) targets will typically move smoothly across the array. Hysteresis reduces oscillation of

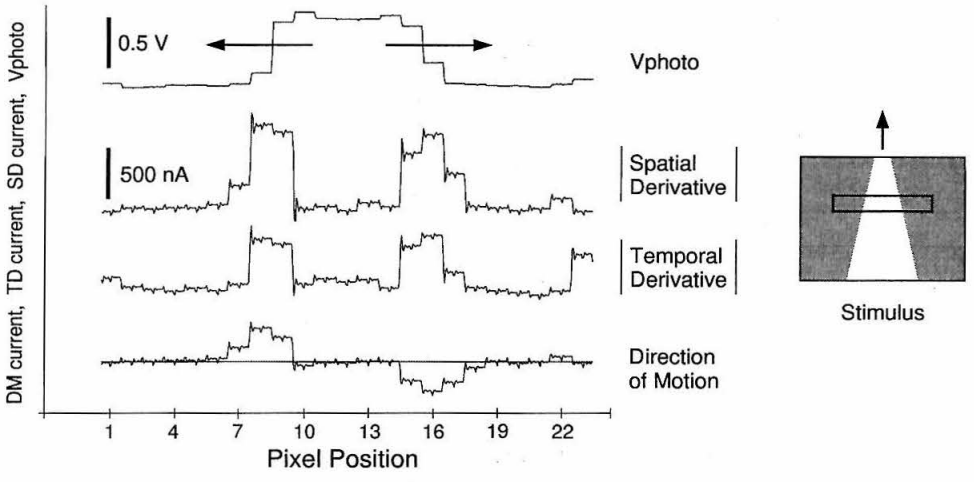

Figure 6.3: Example stimulus - Traces from top to bottom: Photoreceptor voltage, absolute value of the spatial derivative, absolute value of the temporal derivative, and direction-of-motion. The stimulus is a high-contrast, expanding bar (shown on the right), which provides two edges moving in opposite directions. The signed, temporal and spatial derivative signals are used to compute the direction-of-motion shown in the bottom trace. The three lower traces were current measurements which shows some clocking noise from the scanners used to obtain the data. Note: This figure is the same as Figure 4.7 [Chip: TRACK2].

the winning status in the case where two or more inputs are very close to the winning input level and the local distribution of hysteresis allows the winning status to freely shift to neighboring pixels rather than to another location further away.

The WTA output voltage is used to drive three different circuits: the position-to-voltage (P2V) circuit (DeWeerth, 1992), the DM-current-steering circuit (for the outputs of these circuits, see Figure 6.5.), and the saccadic triggering (ST) circuit (see Figure 6.6). The only circuits that are active are those at the winning pixel locations. The P2V circuit drives the common position-output line to a voltage representing its position in the array, the DM-steering circuit puts the local DM circuit's current onto the common motion output line, and the ST circuit (Figure 6.6) drives a position-specific current onto a common line to be compared against an externally-set threshold value. By creating an inverted-"V" shaped profile of ST currents centered on the array, winning pixels away from the center will generate currents below the threshold and send saccade requests off-chip. Figure 6.7 shows the saccade request outputs as a function of winning position for four different thresholds.

6.2.2 Performance

Figure 6.5 shows the direction of motion of the target (DM, upper trace) and the position of the target (P2V, lower trace) for a swinging edge stimulus. As the target moves across the array, different

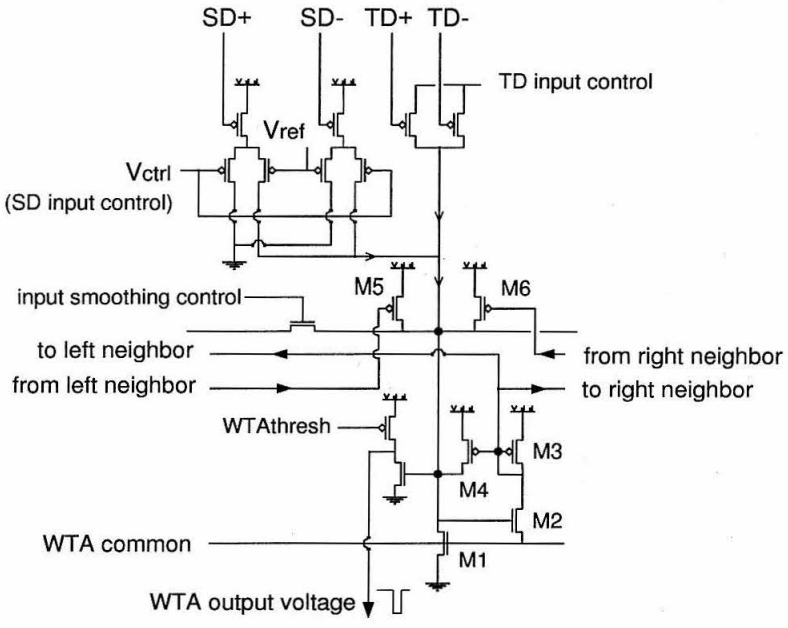

Figure 6.4: Schematic of the hysteretic winner-take-all circuit (HYS WTA). The signed spatial and temporal derivative signals from Figure 6.2 are shown at the top as inputs. The spatial derivative currents are modulated by *Vctrl* such that when *Vctrl* is equal to *Vref*, the SD+ and SD- currents are equally represented in the saliency map. When *Vctrl* is larger than *Vref*, the SD+ current is amplified and the SD- current is attenuated. The TD input control voltage modulates the contribution of the temporal derivative to the saliency map. This can be used to disable the temporal derivative during smooth pursuit eye movements. The input smoothing transistor is used to spread both the input and hysteresis currents. This smoothing deemphasizes smaller regions and helps to reduce the effects of circuit offsets.

direction-of-motion circuit outputs are switched onto the common output line. This switching is the primary cause of the noise seen on the motion output trace. At the end of the trace, the target slipped off the array and the winning-status jumped to a pixel centered on a small, stationary background edge.

To demonstrate the chip's ability to ignore the background, a test stimulus with distractors was presented (Figure 6.8). A high-contrast stimulus with a low-contrast background moving in the opposite direction was tested. Once the high-contrast target is being tracked, the lower-contrast background stripes are ignored. When the high-contrast target is no longer visible, one of the background stripes is chosen and tracked.

To test the speed of the tracking circuit, we reduced the time-constant of the photoreceptor circuit and a single edge was passed in front of the array at varying speeds. Figure 6.9 shows the ability of the tracking chip to follow even fast stimuli moving in excess of 2300 pixels/sec.

The "target identification" signal available on the later versions of this chip allows external signalling of the target edge's polarity. Figure 6.10 shows the output for a stimulus containing

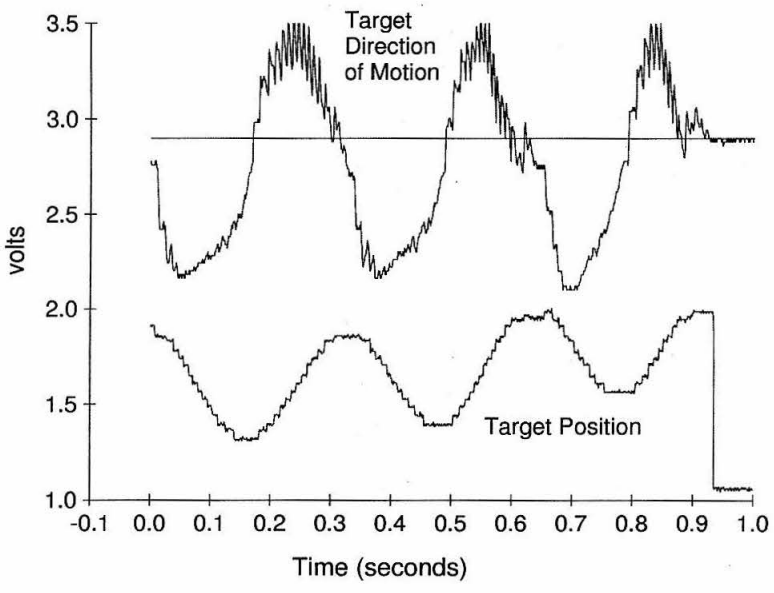

Figure 6.5: Extracting the target's position and direction of motion: The WTA output voltage is used to switch the DM current onto a common current-sensing line. The output of this signal is seen in the top trace. The zero-motion level is indicated by the flat line shown at 2.9 volts. The lower trace shows the target's position from the position-to-voltage encoding circuits. The target's position and direction of motion are used to drive saccades and smooth pursuit eye movement during tracking. [Chip: TRACK3, NN-hysteresis]

both types of edges. In addition to the identification signal, the saliency map can be controlled by an input voltage (SD input control) to differentially modify the gains of the positive and negative spatial-derivative inputs to the saliency map (see signals *Vctrl* and *Vref* in Figure 6.4). Figure 6.11 shows the effect of switching this control input in the presence of an input image. By controlling the saliency map to favor the edge polarity being tracked, a simple type of feature-based hysteresis (F-hysteresis) is created. This significantly improves the stability of tracking performance. In the layout of the chip, the target identification signal is sent off-chip as a current on the same physical wire as the SD input control voltage (*Vctrl*). By connecting the control wire to the reference voltage (*Vref*

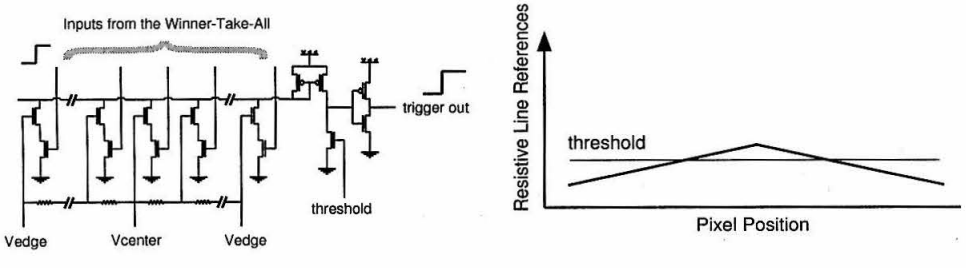

Figure 6.6: Schematic of the saccadic triggering (ST) circuit.

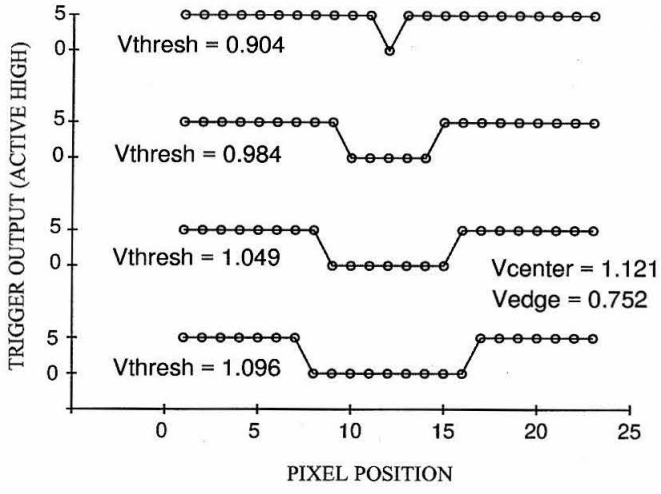

Figure 6.7: Saccadic triggering (ST) circuit output as a function of winning position. Plots have been shown for four different threshold levels. [Chip: TRACK2]

) through a resistor, the target identification current can directly modify the saliency map control voltage and induce the feature-based hysteresis (see Figure 6.12). Further, by placing a capacitor on the control voltage, it is possible to create a short-term memory of target type to improve target reacquisition across saccades or large jumps of the target.

The power consumption of the chip (23 pixels and support circuits) varies between 0.35 mW and 0.60 mW at a supply voltage of 5 volts. This measurement was taken with no clock signal driving the scanners since this is not essential to the operation of the circuit. It is expected that the power requirements of this one-dimensional circuit will scale linearly with size.

6.2.3 Discussion

An analog VLSI model of visual attention has been successfully used for the selection and tracking of a target against a non-uniform background using a saliency map model for target selection. This selection mechanism allows the chip to dynamically isolate information about the target from information about the background.

The selection circuit is based on the winner-take-all principle, selecting the most salient location. At first glance, this appears sufficient to solve the problem, even for moving targets. Experience with real moving stimuli, however, shows that as edges move across the imaging array, the saliency measurement on a short-timescale is highly variable due to temporal noise, circuit offsets, image discretization, and lighting fluctuation. This causes the winning status to jump freely between different targets in the image. The combination of locally-distributed hysteresis (Morris and Deweerth, 1996), input smoothing, and feature-based hysteresis has improved the tracking performance to the

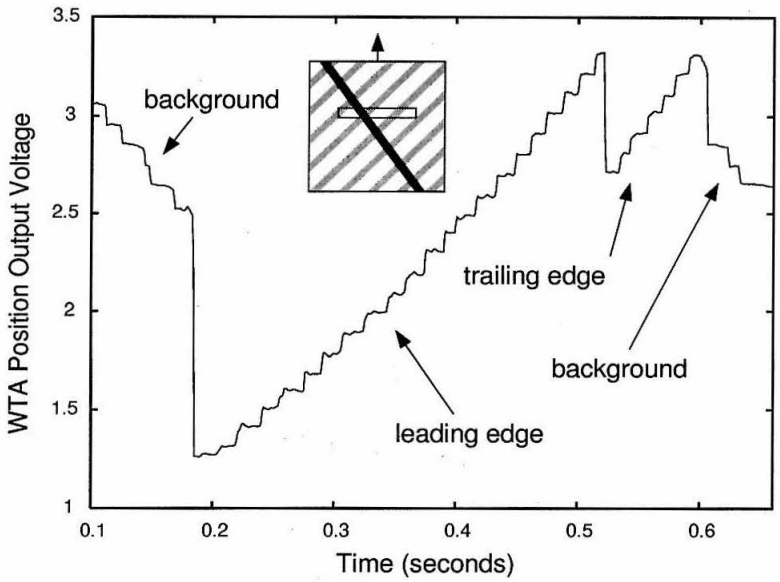

Figure 6.8: The HYS WTA's ability to ignore background activity is demonstrated in this test where a low-contrast background is moving in the opposite direction as the high-contrast target. At the beginning of the trace, a background edge is being tracked until the high-contrast edge appears. The edge is tracked to the end of the array and the trailing edge is selected next. After the trailing edge has left the image, another background edge captures the winning status. [Chip: TRACK4, no feature-based hysteresis]

point where reliable tracking can be achieved.

From a biological perspective a number of issues are raised by the analog VLSI model. First, the speed of attentional shifts in the analog VLSI model are extremely fast, allowing it to follow rapidly moving objects or jump across the array. Psychophysical experiments, while certainly not in complete agreement as to the actual speed if such a thing can be truly be defined, suggest a modest speed of 33 msec per shift (Saarinen and Julesz, 1991). Presumably, a neural implementation of the winner-take-all function will be slower.

Finally, the method of extracting the target motion on the chip is a simple gating mechanism controlled by the attentional process. On the chip, the direction of target motion is computed everywhere and only one output is gated off-chip, whereas in the biological system, as has been suggested by the Treue and Maunsell results, the motion output (MT) itself is modulated by the attentional system, and not just gated downstream. The saccadic suppression of the inputs to the pursuit integrator is a similar situation.

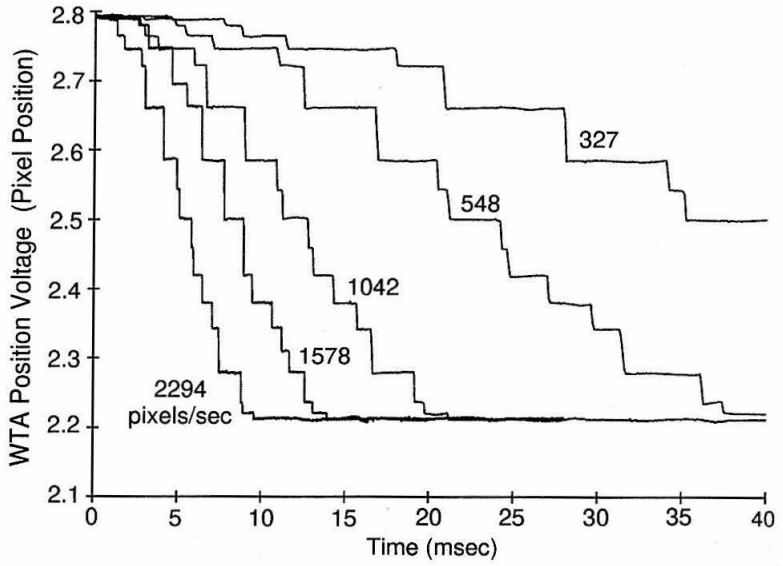

Figure 6.9: Position vs. time traces for the passage of a strong edge across the array at five different speeds. The speeds shown correspond to 327, 548, 1042, 1578, 2294 pixels/sec. [Chip: TRACK1, R-hysteresis]

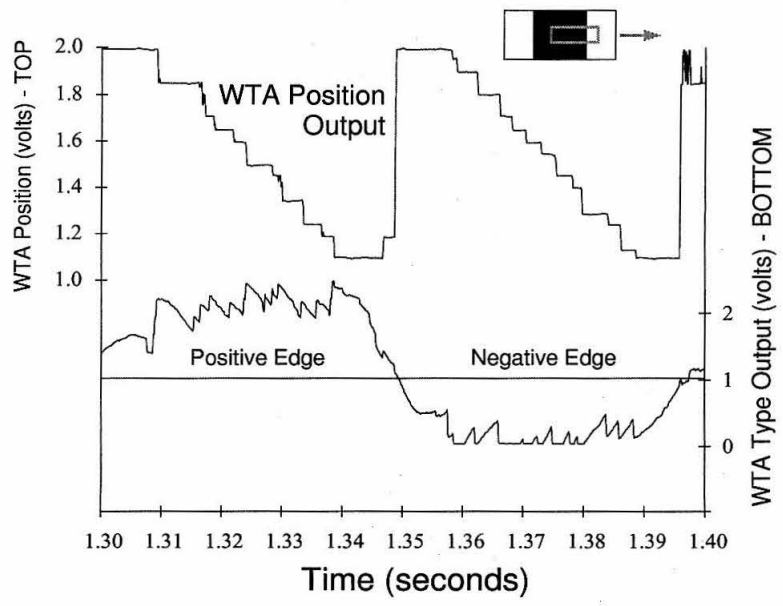

Figure 6.10: Target Identification: The chip is also capable of reporting the polarity of the edge being tracked. The upper trace shows the passage of two edges of different polarity moving in the same direction and the lower trace shows the chip output indicating the polarity. By adjusting the weights of the different-polarity SD inputs to the saliency map, the system can improve tracking by reducing the effects of opposite-polarity edges. This can also be used to improve reacquisition of the target after saccades. [Chip: TRACK3, NN-hysteresis]

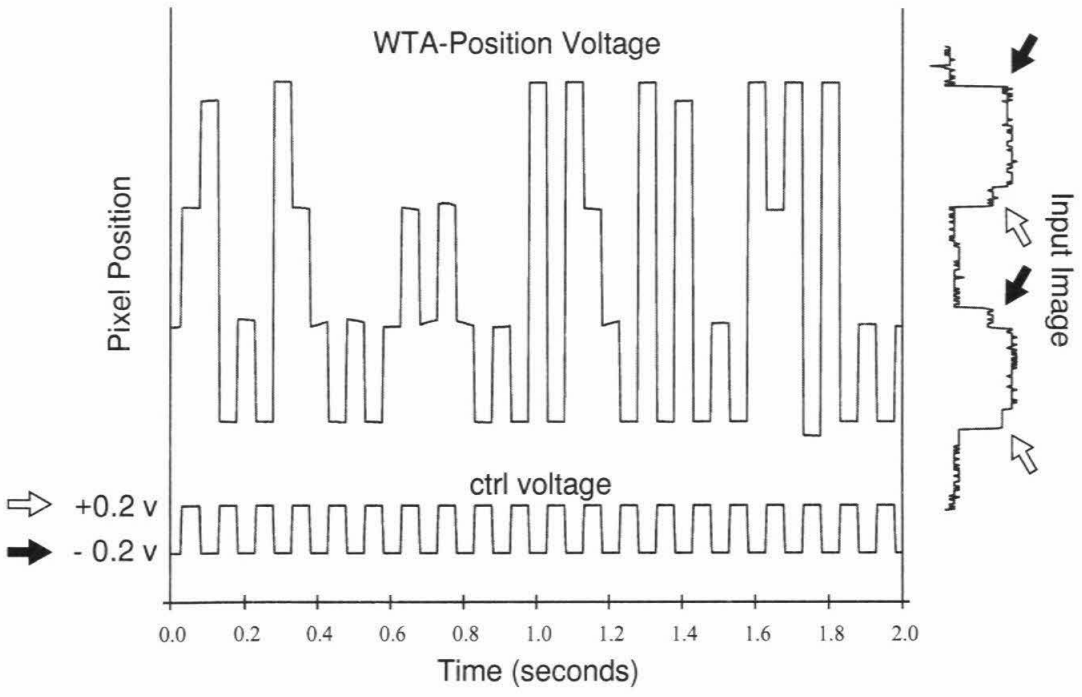

Figure 6.11: The control input for the saliency map (lower trace) is switched between allowing only positive spatial derivatives and only negative spatial derivatives. The WTA position voltage (upper trace) is seen to switch between the two positive-derivative edges and the two negative-derivative edges. In the example image, after the photoreceptor circuits begin to adapt, the strengths of the four edges become similar and the switching pattern begins to appear random. The input image is displayed on the right edge of the plot. The control voltage is given relative to the reference voltage (*Vref*) in this case equal to 2.6 volts. [Chip: TRACK4, no feature-based hysteresis]

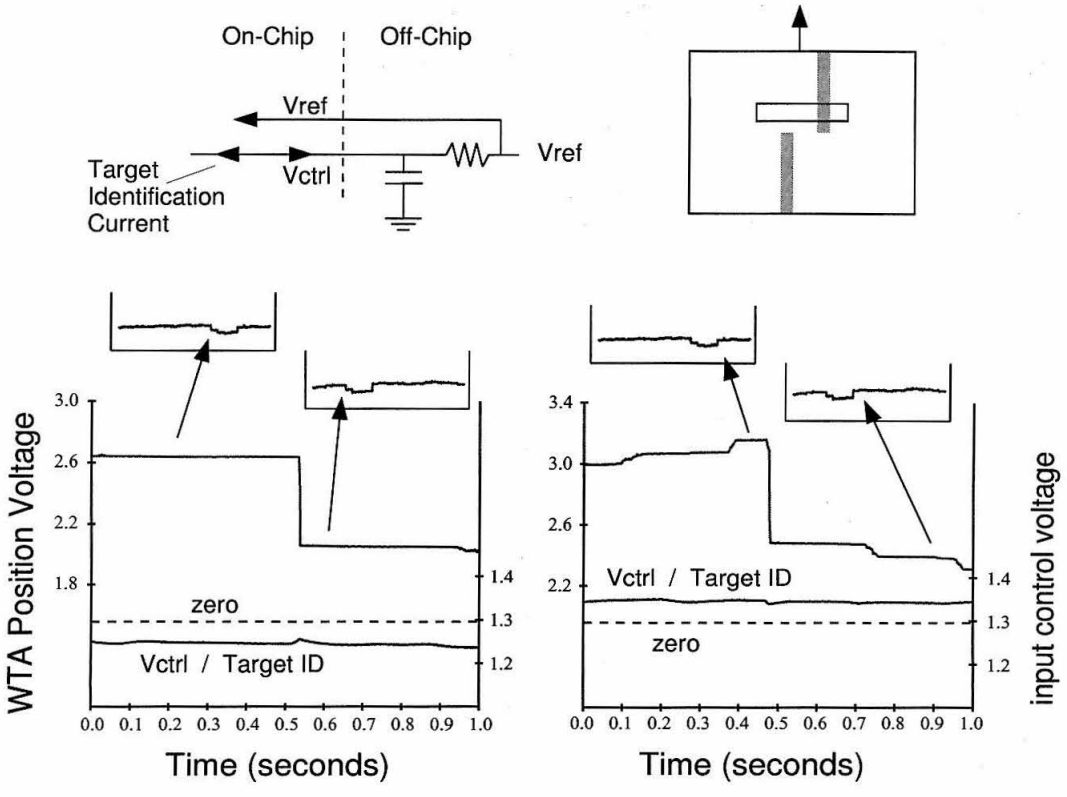

Figure 6.12: In this experiment, a tracked blob jumps from one location on the array to another. In the left panel, the chip begins tracking the negative-derivative edge (left edge of the blob) and the target identification current is inward, pulling the control input (*Vctrl*, or SD input control) below Vref. This differentially-increases the gain on the negative-derivative inputs over the positive-derivative inputs. When the target jumps, the capacitor on *Vctrl* serves as a short-term memory and the chip finds another negative-derivative edge. The right panel shows a similar experiment for a positive-derivative edge. [Chip: TRACK4, feature-based hysteresis]

Chapter 7 System Integration

Normally, when we track moving objects in the world, our eyes respond with a combination of saccadic and smooth eye movements, dependent upon the background, speed, trajectory, and size of the target. The integration of these two behaviors gives rise to a number of interesting interactions between their respective control systems and the retina. Because our visual acuity is impaired when the retinal image is moving, the control of eye movements must be orchestrated to maximize the ability to extract information from the scene.

In this chapter the interactions between eye movements and the visual processing which control them are studied by mounting the attention-based tracking chip described in Chapter 6 on the eye movement system.

7.1 System Integration

The attention-based tracking chip from Chapter 6 has been mounted on the hardware model of the primate oculomotor system described in Chapter 2 and is being used to track moving visual targets. Special boards for mounting have been constructed to carry the chip, a lens, and potentiometers for biasing. The lens currently in use is a 2.5 mm focal-length lens (2.5 mm in diameter) which is secured in front of the chip by a light-weight clamping system. Approximately ten wires are fed off the back of the board to provide information to the saccadic and smooth pursuit controller boards. These wires provide additional mass and mechanical elasticity to the oculomotor plant.

Although the desired field of view for the system is closer to 140 degrees, the field of view allowed by the current lens is approximately 35 degrees. See Figure 2.18, page 29, for a photo of the lens system.

7.1.1 Saccadic Tracking

Figure 7.1 shows the behavior of the system when the WTA-position output voltage of the tracking chip is used as the input to the saccadic burst generator for recentering saccades. Saccades are triggered when the selected pixel is outside of a specified window centered on the array (i.e. $V_{higherthreshold} < V_{WTA}$, or $V_{WTA} < V_{lowerthreshold}$). The visual target, a single vertical edge, is mounted to a swinging apparatus to generate an oscillatory motion.

Unlike the reflexive, saccadic triggering system (Chapter 2) which saccades to the centroid of change, this attention-based saccadic system selects a single target even in a field of distractors.

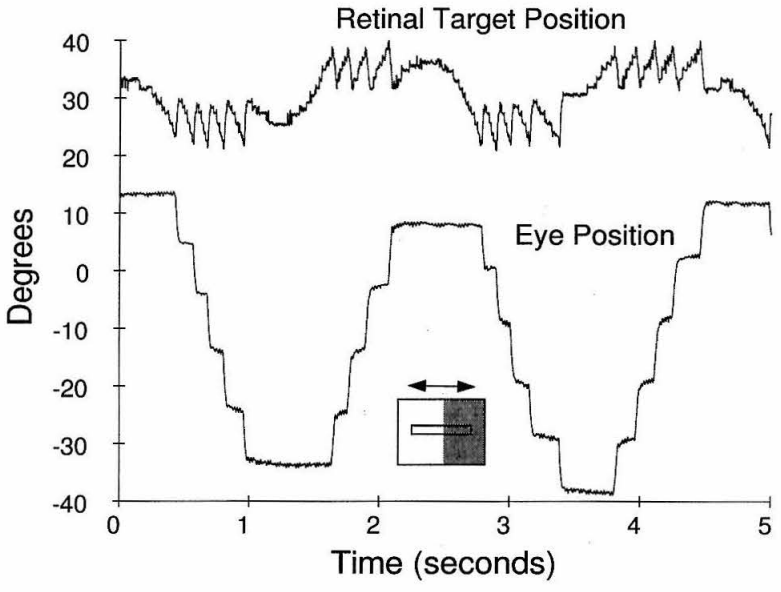

Figure 7.1: Saccadic Tracking: In this example, a swinging target – consisting of a single edge with no distractors – is tracked over a few cycles. The top trace shows the WTA position voltage which encodes the position of the tracked target on the retina and the bottom trace shows the eye position. Re-centering saccades are triggered when the target leaves a specified window centered on the array. [Chip: TRACK2, no hysteresis]

7.1.2 Smooth Pursuit Tracking

In contrast to saccadic tracking, smooth pursuit attempts to match the eye's velocity with the target velocity in order to hold the image stationary for further scrutiny. In humans, smooth pursuit is most effective when the target speed is less than 30 deg/sec and does not change directions quickly. The latency of smooth pursuit is between 80 msec and 130 msec. Smooth pursuit eye movements have two notable characteristics: they are visually-controlled and they display velocity memory when the target is stabilized or disappears.

The model used to drive smooth pursuit is based on the model proposed by McKenzie and Lisberger (1986) using only the target velocity input. This model mathematically integrates visual motion over time and drives the oculomotor plant in parallel with the saccadic burst generator.

In order to implement smooth pursuit in this system, an external circuit was constructed to model the pursuit integrators. Figure 7.2 shows the basic information flow in this circuit. The actual circuit is described in Appendix C. The transconductance amplifiers drive current into the capacitors to perform the integration from eye acceleration to eye velocity and the capacitor voltage drives the pulse generator circuits. The integrator leak time constant has been set to 1 sec. The input switch is driven by the saccadic burst generator to suppress the integrator input during saccades, allowing the integrator value to continue unmodified during and after the saccade. The output of the smooth

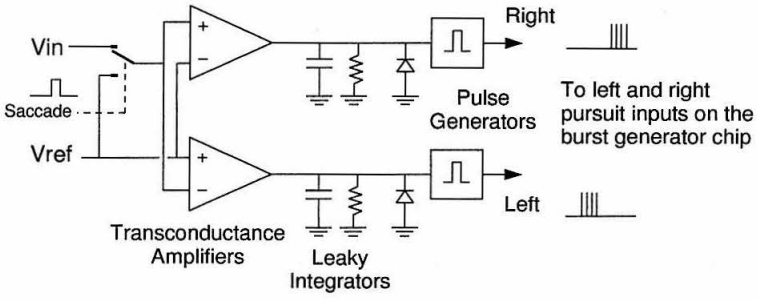

Figure 7.2: Smooth Pursuit Integrators: This integrator circuit is an external circuit constructed from discrete parts, taking the target motion signal (eye acceleration) from the tracking chip (*Vin*) and integrating the positive and negative components (eye velocity) before generating spikes to drive the burst generator chip. The input switch prevents the motion induced by the saccade from being integrated.

pursuit integrators drive the two inputs on the burst generator chip (Figure 2.6 on page 19), driving both the motor units and neural integrator in parallel with the saccadic burst signal. A test of the pursuit system is seen in Figure 7.3.

Figure 7.4 shows the behavior of the smooth pursuit system (chip: TRACK2) successfully using only the motion output to drive eye movements. The hysteresis spreading circuit on TRACK2 was disabled by a design error and shows oscillation in the winner's location on the array (top trace). Figure 7.5 shows the improved TRACK4 smooth pursuit system which utilizes nearest-neighbor hysteretic spreading and feature-based hysteresis.

Although not readily visible in Figures 7.4 and 7.5, oscillation in the pursuit velocity around the target velocity is a common occurrence in primates (see Figure 7.9) and is also seen in this system. By modelling the control of smooth pursuit as a negative-feedback system with a delay in the visual system, an analysis of the stability can be performed. In the simple case where no delay is assumed, the control-systems model of the pursuit system (see Appendix C) lends itself to oscillation under conditions of large gain in the integrator stage. This is caused by the lag in response by the inertia of the eyeball. If we assume an eye with a rotational inertia, I, a damping coefficient, m, and a pursuit integration constant, C, then to prevent oscillation in the case of no delay:

$$C \leq \frac{m^2}{4I}$$

If we introduce a delay in the visual system of t_d, the criterion for preventing oscillation becomes:

$$C \leq \frac{(m - Ct_d)^2}{4(I + \frac{Ct_d^2}{2})}$$

which is similar to the delay-free condition, but C must be smaller as t_d increases. Satisfying

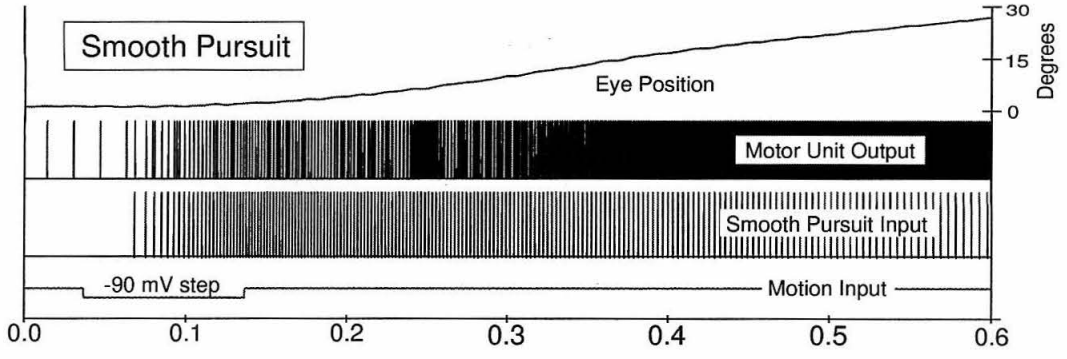

Figure 7.3: Example Smooth Pursuit: The downward-going pulse on the lowest trace is a fictitious motion input placed on the input to the pursuit system by a pulse generator. The motion input is mathematically integrated by the pursuit integrator and the resultant spiking output is shown in the second trace, labelled "Smooth Pursuit Input." This signal represents the desired eye velocity and is used as an input to the burst generator chip. Note the visible decay in the output of the pursuit integrator. The "Motor Unit Output" trace shows the steady increase in innervation as the eye (eye position: top trace) moves steadily onward, long after the motion input has ceased.

the condition for avoiding oscillation by keeping C small, however, means that the pursuit system response will be sluggish. The time to integrate the eye velocity up to the target velocity is inversely proportional to C. Any delay in the visual sensing will make the situation worse. Perhaps for this reason, the biological system sacrifices a little stability for responsiveness.

7.1.3 Integration of Saccades and Smooth Pursuit

When we view natural scenes mixed with both stationary and moving objects, we combine *saccades* and smooth pursuit in an attempt to quickly take in a scene and scrutinize moving objects. While it is known that the circuitry that mediates these two movements are largely separate, some interaction, particularly in the form of saccadic suppresion, is required. How these two eye movements are behaviorly combined is still unclear. Figure 7.6 shows an example eye position trace of a monkey pursuing a sinusoidally-moving target. The eye movement is occasionally punctuated with catch-up saccades as well as saccades which take the eye away from accurate tracking. Although primates have a strong smooth pursuit system (Lisberger et al., 1987), it is not true that smooth pursuit is always employed. Experiments with primates have shown that tracking a small target on an unstructured (uniform) background elicits smooth eye movements while tracking the same target on a structured (textured) background can sometimes elicit saccadic tracking instead of smooth pursuit behavior (Ilg et al., 1992).

In this project, the saccadic system is driven by a retinal position error and the pursuit system is driven by a retinal velocity error. While these two motor control systems operate essentially

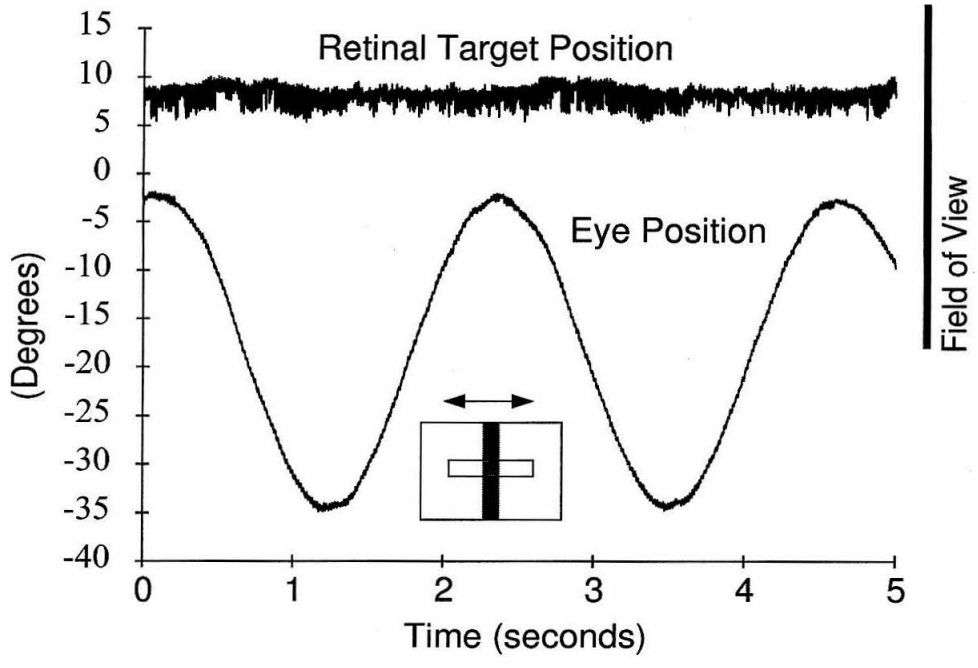

Figure 7.4: Smooth Pursuit Tracking: In this example, a swinging target consisting of a bar with no distractors is tracked over a few cycles. The top trace shows the WTA position voltage which encodes the position of the tracked target on the retina. The bottom trace shows the eye position. In this particular chip the hysteresis function was disabled. Although the attention jumps back and forth between the left and right edges, the reported direction of motion is consistent. [Chip: TRACK2, no hysteresis, TD active in the saliency map.]

independently, the target motion generated by the saccade must be suppressed at the input of the smooth pursuit integrator in order to maintain the smooth pursuit eye velocity between saccades.

Figure 7.7 shows an example of the integration of both saccadic and smooth pursuit eye movements during the tracking of a sinusoidally swinging target. When the velocity of the target exceeds the peak velocity of the pursuit system, the target slips out of the central region of the chip's field of view and saccades are triggered to recenter the target.

In a second experiment used to show the separation of the saccadic and smooth pursuit systems, a step-ramp stimulus was used to drive both systems at once, but in different directions. Figure 7.8 shows some of these results. In the top panel, the integrated system, as described above, is used to track the jumping target. As there is no explicit delay in the saccadic triggering system, the saccade is triggered immediately and the pursuit system's influence is only seen afterwards due to the saccadic suppression. In primates, visually-triggered saccades show latencies from 150 to 250 msec, revealing the shorter latency (80 to 130 msec) of the pursuit system (Lisberger et al., 1987). The second panel of Figure 7.8 shows the effect of adding an artificial delay of 100 msec in the saccadic trigger. This has the effect of separating the responses of the saccadic and smooth pursuit

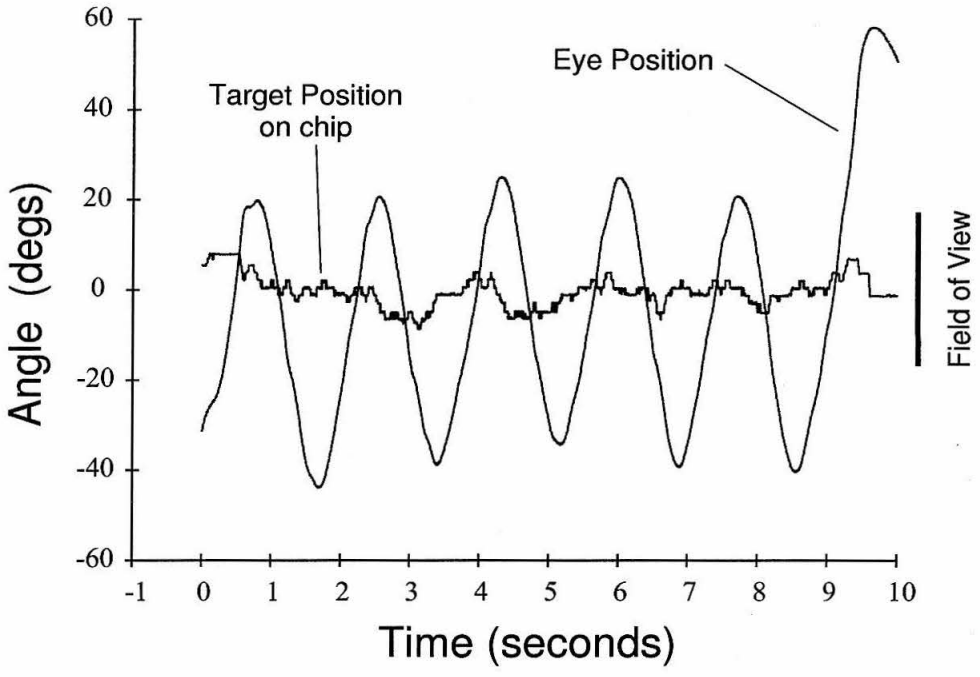

Figure 7.5: Smooth Pursuit Tracking: This experiment is similar to that of Figure 7.4, except that the TRACK4 chip is being used, utilizing both spatially-distributed and feature-based hysteresis. Also, in this case, the target was moved in an oscillating trajectory by hand. The temporal derivative has been disabled to allow tracking across non-uniform backgrounds. [Chip: TRACK4, NN-hysteresis, F-hysteresis, TD disabled in saliency map]

systems. The apparent improvement in tracking in the lower panel is believed to be coincidental. Increasing the latency of the saccadic system has the general effect of increasing the saccade size and decreasing the reliability of tracking since the target typically moves further away from the center of the array during the delay period and sometimes escapes off the array before the saccade occurs. The latency of the pursuit system without adding additional delays is approximately 50 to 60 msec. Figure 7.9 shows a comparison with primate step-ramp experiment data.

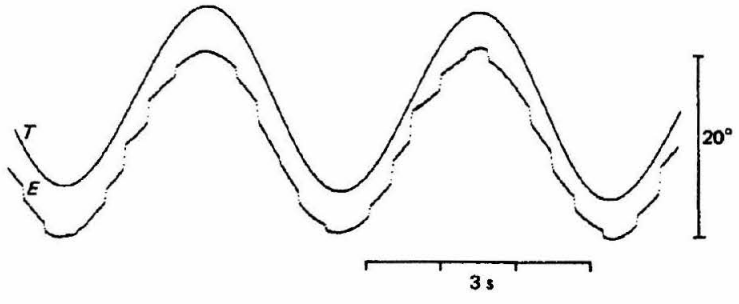

Figure 7.6: Pursuit eye movements of a monkey in reponse to sinusoidal target motion at approximately 0.27 Hz, peak-to-peak amplitude 20 deg. The target and eye position traces have been displayed with an offset for clarity. (From Collewijn and Tamminga, 1984).

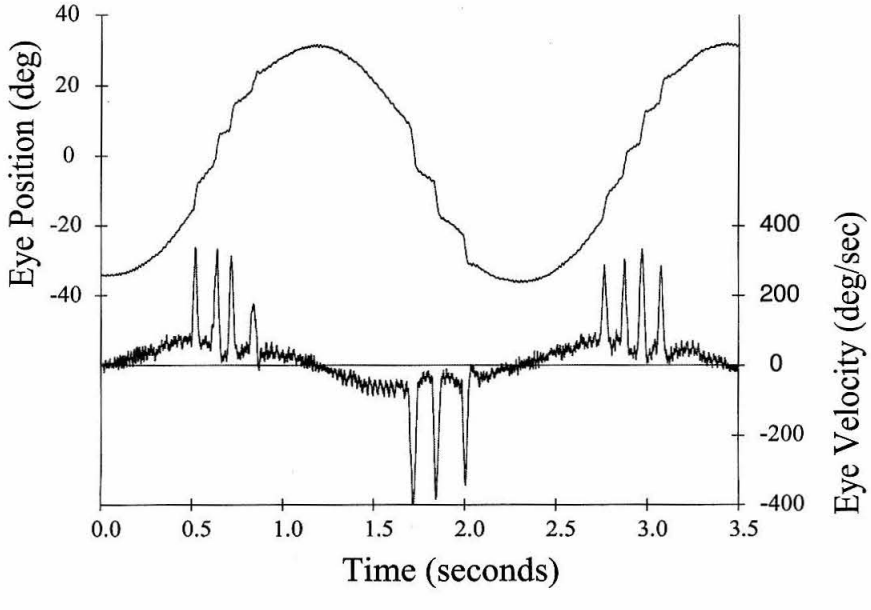

Figure 7.7: Integration of saccadic and smooth pursuit tracking: In this example, a swinging target consisting of a bar with no distractors is tracked over a few cycles. The top trace shows the eye position over time and the bottom trace shows the eye velocity. [Chip: TRACK2, no hysteresis]

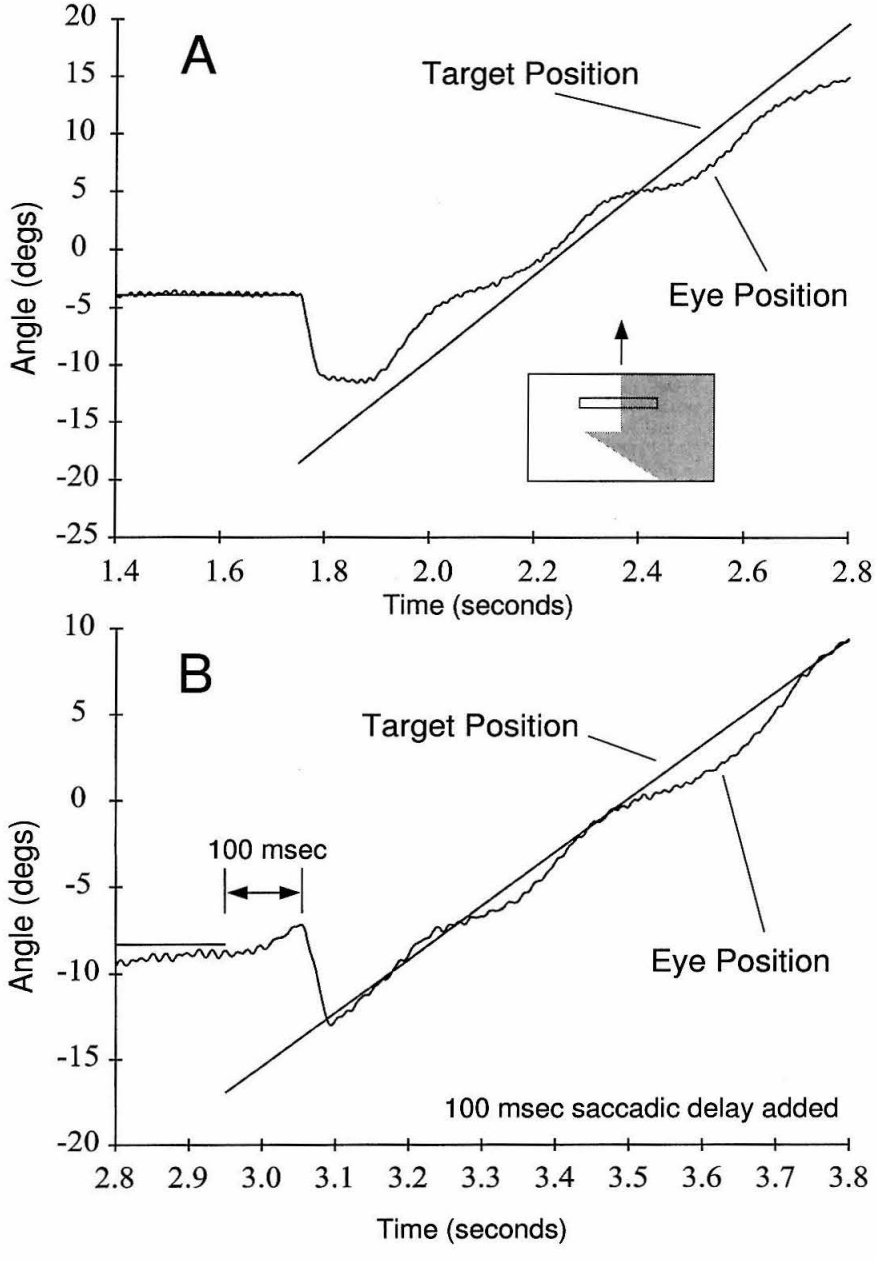

Figure 7.8: Step-Ramp Experiment: (A) The target jumps from the fixation point to a new location and begins moving with constant velocity. (B) Same experiment as in (A), but with a 100 msec delay in the saccadic trigger. Note: only the saccadic trigger is delayed; the target information is current. [Chip: TRACK4, NN-hysteresis, F-hysteresis]

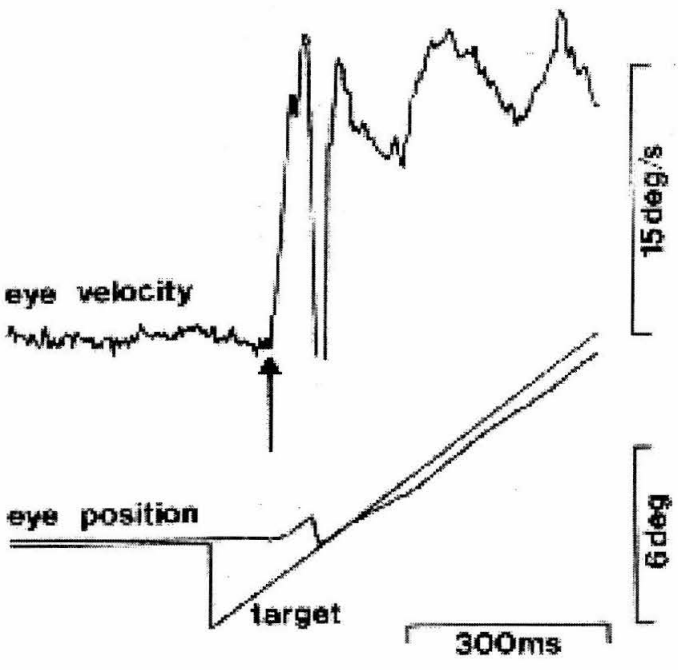

Figure 7.9: Pursuit eye movements of a monkey in response to a step-ramp target motion. Eye velocity records were obtained by electronic differentiation of eye position records. The upward arrow indicates the intiation of pursuit, which preceded the first saccade. (From Lisberger, Morris, and Tychsen, 1987).

7.1.4 Discussion

While attentional selection and tracking is fairly straightforward in the absence of eye movements as shown in Chapter 6, saccades and smooth pursuit eye movements violate the assumptions used to select targets in the first place.

When the eyes are stationary, moving objects draw our attention and thus the temporal derivative is an integral part of the saliency map. During successful smooth pursuit movements, however, the target is relatively stationary and the background is moving. If the temporal derivative is not disabled, the background will immediately draw the attention off the target. For this reason, the temporal derivative input to the saliency map must be attenuated or disabled in response to eye movements. For most of the data shown, however, it was left disabled. Ideally, the temporal derivative should be suppressed only during eye movements, but this requires a distinction between pursuit and static fixation. Currently, in the hardware model, the pursuit system is always active. This suggests the need for a fixation state during which the pursuit system is disabled and the saliency map can reinstate the temporal derivative input. Although fixation and pursuit share many characteristics (e.g. saccadic reaction time and gap-overlap paradigm effects, Boman et al., 1996), pursuit velocity oscillations are not commonly seen around a stationary fixation point. The current hardware model, without a fixation system, pursues the fixation point and exhibits these oscillations.

During saccades, our eyes often reach speeds up to 750 deg/sec where vision is severely impaired due to blurring. This can cause a loss of proper tracking. To deal with this problem, the tracking chip has two predictive mechanisms to improve the successful reacquisition of the correct target: a memory of the feature type being tracked (feature-based hysteresis), and a saliency map input to draw the attention to the center of the array. The latter input has not yet been utilized in the system due to the ability of the tracking circuit to follow targets even during saccades. In addition to this attentional prediction across saccades, saccadic suppression has successfully been used to prevent the pursuit integrators from integrating the saccade-induced motion transient. This is critical for the smooth operation of the pursuit system.

Having completed a functional saccadic and smooth pursuit system, a number of discrepancies with the known neurophysiology need to be mentioned. Delays in the oculomotor control system are an important factor in understanding the stability of the closed-loop biological system. While the long saccadic latency is not inherent in the VLSI model, an artificial delay of 100 msec has recently been added to simulate the saccadic latency seen in primates. This delay is believed to be due to the decision-making process and the mechanisms of disinhibition in the saccadic system. Initial experiments with the model have shown that this delay creates a noticeable impairment in its saccadic tracking ability. Part of the problem in the current system may lie in the fact that during this 100 msec, the target can escape the chip's small field-of-view, triggering incorrect saccades. A similar latency exists in the pursuit system which, together with the pursuit integrator gain, creates

the oscillations seen in the eye velocity.

Although the motion detection circuit used for driving smooth pursuit is intended to be a model of the magnocellular-dominated cortical motion areas (MT and MST), the circuit only extracts the direction of motion and is more like the shorter-latency direction-selective responses from V1. This type of input is believed to drive the initiation of smooth pursuit, while the responses of MT are believed to drive the maintenance of pursuit (Lisberger et al., 1987). This is believed to be one source of the exaggerated oscillations in eye velocity seen with this system.

In spite of the many improvements and details which need refinement, the basic integrated tracking system performs well, using only the direction of target motion to guide the smooth pursuit eye movements.

Chapter 8 Conclusions

The primate oculomotor system model presented in this dissertation is an exploration into the issues of systems-level modeling using neuromorphic analog VLSI. This work focuses primarily on feedback systems which involve sensory and motor interaction with natural environments. Within the analog VLSI framework, it has touched on various examples of sensorimotor control, learning, sensory fusion, coordinate transformation, and the integration of different eye movements.

There are clearly two goals within the neuromorphic analog VLSI community: 1) to demonstrate useful modeling systems which can offer an advantage over current software techniques and 2) to develop novel computational architectures useful for industrial or commercial applications. While both goals have been a consideration in this work, the primary goal of the research presented here has been to develop a useful tool for modeling biological systems.

The primate oculomotor system is an excellent sensorimotor system to study with analog VLSI for many different reasons. From the motor control perspective, the primate oculomotor system is a good system to investigate due to its relatively simple musculature and the extensive knowledge of the neural substrate driving it. Behaviorally, the primate eye shows a diversity of movements involving: largely open-loop movements (saccades), mechano-sensory-driven feedback movements (VOR), and visually-driven feedback movements (smooth pursuit and the opto-kinetic reflex). Finally, from a visual processing point of view, the visual information needed for the control of eye movements is relatively simple. Perhaps the most important aspect of eye movements is that the input and output representations are both relatively well known and their purpose is fairly clear.

The contributions of this research to the biological modeling and engineering communities will be presented first, then the future of this line of research and the field of neuromorphic analog VLSI will be discussed.

8.1 Contributions

As with most scientific endeavors, this work stands proudly on the research of those who came before. This work has not focused on the development of new circuits; it has focused on the large-scale integration of many different neural models into a functional system.

8.1.1 Contributions to Biological Modeling

Computational modeling is a quantitative tool for exploring the consequences of the assumptions and details we place into a model system. As such, our measure of success lies in producing a system which can lend insight into previously observed or unobserved phenomena and can guide future experiments.

The main contribution of this system thus far has been the demonstration of a real-time modeling system that brings together many different neural models to solve real-world tasks. To date, there are no other oculomotor modeling systems that use realistic burst generator circuits to drive an analog oculomotor plant with similar dynamics to the biological system. While other research groups have built biologically-inspired, visual tracking systems, the problems they encounter are generally not the problems faced by biological systems because they do not solve the task with hardware that has similar properties.

By building circuits which compute with the representations of information found in the brain, the modeling system presented here is capable of replicating many of the behavioral, lesion, stimulation, and adaptation experiments performed on the primate oculomotor system. Saccadic and smooth pursuit eye movements have been integrated in the system, which has raised many questions about how to model their interaction. Auditory localization based on neural models has been demonstrated in conjunction with the saccadic system to trigger saccades to auditory targets. This work has raised many questions about computations across different coordinate frames and sensory modalities. Adaptation of saccadic parameters based on biologically-constrained error measures has been demonstrated. A crude model of visual attention for target selection and selective motion extraction has been demonstrated, raising many questions about the interaction between reflexive (collicular) and volitional (cortical) eye movement systems.

Additionally, as a physical embodiment of many of the neural models developed for the primate oculomotor system, it serves as an excellent hands-on demonstration system for the various principles involved in the control of eye movements.

While the system has not yet been involved in any modeling efforts with specific experiments in mind, it is our hope that further development of the system will lead to progress in that direction.

8.1.2 Contributions to Engineering

While one large category of engineering is the design of general computational engines, another important category of engineering is the design of efficient, dedicated systems. Neuromorphic analog VLSI circuits fit well into this latter category.

Much of the work presented here has centered on visuo-motor functions which directly affect an animal's ability to survive. The chips designed for modeling active vision are directly applicable to

a wide range of security, surveillance, and robotics applications.

The visual-triggering chip described in Chapter 2 and the tracking chip described in Chapter 6 are both well-suited for intruder or occupancy detection systems. Unlike most sensors in use today, these chips can report the direction of a disturbance. Additionally, the silicon photodetectors make it suitable for operation under near-infrared illumination. The auditory localization system described in Chapter 3, while not yet fully integrated onto a single chip, also holds much promise in detecting speaker direction for improving voice-recognition systems, teleconferencing trackers. This localization system could be used to construct a sonar module capable of reporting angle as well as distance.

Due to the use of a commercially-available VLSI process and the integration of both sensor and computational circuitry, these chips are fast, low-power (e.g., 0.6 mW in the case of the tracking chip, TRACK4), and potentially very cheap to produce in quantity.

8.2 The Future

There are two futures to consider with regard to this dissertation: the future of continued modeling of the primate oculomotor system, and the future of neuromorphic analog VLSI. In the first section, a short list of potential projects based on this work are suggested. In the second section, a number of reasons why neuromorphic analog VLSI will continue to thrive are given.

8.2.1 Modeling the Primate Oculomotor System

As the oculomotor system becomes more sophisticated, there are many more interesting experiments to try. While the first design focused on creating a functional system based on the basic representations found in the brain, many of the details can now be refined.

There are a number of interesting projects within the framework of the existing system which were not addressed in this dissertation, but can be considered without too much modification or development: 1) Exploring various models of fixation, attentional engagement and disengagement, and how express saccades can be performed with the system (Fischer and Weber, 1993). This can also involve the integration of the reflexive saccadic system (Chapter 2) and the attentive saccadic system (Chapter 7) 2) Exploring how delays in both the saccadic and smooth pursuit systems affect tracking performance and how it compares with biological data (Wolpert et al., 1993; Ringach, 1995). 3) Exploring the issues of saccadic suppression, predictive shifting of attention across saccades, and interference of the opto-kinetic response with smooth pursuit (Masson et al., 1995). 4) Using the tracking chip in conjunction with the floating-gate look-up-table circuits to achieve on-line learning based on real world experience.

In considering the longer-term expansion of the current system, there are many intriguing projects

to consider, such as: 1) The integration of the inhibition-of-return circuits developed by Morris and DeWeerth (1996) for implementing conjunction search with the target identification function described in Chapter 6. 2) The integration of the object-based selection circuits which select contiguous salient regions instead of single pixels (Morris, 1996). 3) Conversion of the existing circuitry to a multi-chip system using the address-event-representation communication protocol (Boahen, 1996) to facilitate the continued growth of the system. 4) Perceptual grouping of pixels based on neighborhoods of similar directions of motion. 5) Construction of the attentional selection mechanism in a higher coordinate frame.

8.2.2 Neuromorphic Analog VLSI

The use of neuromorphic analog VLSI in both neural modeling and engineering seems very promising. While we typically view these two fields as separate endeavors, the path to understanding how the brain computes will probably involve the goals of both fields. It seems clear that the simulation of large neural systems will require considerably more powerful machines than exist today. If we consider that the human brain does its job with only a few watts of power, it becomes clear that engineering has something to learn from biology about power efficiency.

Vision, audition, olfaction, and somatosensation, all involve massive amounts of parallel processing, particularly near the sensory surfaces. What we know about the cortical architecture and known processing streams also point to massively parallel computation. While it is possible to simulate parallel computation with a serial machine, if we consider that the human brain contains approximately 10^{11} neurons and that we require only about 30 msec to identify complex visual stimuli (Tovée, 1994), the computational power of the required computer is so high, that a serial implementation doesn't make sense. For this reason massively-parallel computing architectures are likely to be the most appropriate for this task. Efficiency is another extremely important issue to understanding the brain's construction. It has been estimated that the brain only consumes a few watts, yet it performs on the order of 10^{16} operations per second (Mead, 1990). Analog circuitry, while limited in computational precision compared to digital circuits, is extremely efficient in both space and power. Many studies support the idea that neurons do not compute at signal precisions higher than 7-8 bits. For this reason, analog computing architectures are likely to be the most appropriate approach for these new computers.

Much of the utility in the analog VLSI approach for biological modeling is in its speed of computation. Speed is necessary for real-world sensorimotor interactions and for developing intuitions about models. The speed of computation is particularly important for simulations where behavior at multiple time scales is being simulated. Tuning the system to properly capture the behavior at sub-millisecond time scales makes modeling adaptation time scales of days extremely slow. It is our hope that as these systems continue to grow, the ability to simulate behavior on a larger scale in real

time will exceed the fastest digital systems available. The challenge to the designer is not simply to make such systems work, but to make them detailed enough to provide insight into biological problems and easy enough to experiment with.

There are many lessons about biological systems to be learned from attempting to build hardware to operate in the real world. From the very start, one must deal with real noise and real images. An intangible contribution of analog VLSI design to modeling neural circuits has been the constraints under which the designer must work. Manufacturing offsets, limited silicon real estate, and noise are perhaps the main considerations in the design process. This has resulted in the development of circuits which utilize adaptation wherever possible, operate on signal differences instead of absolutes, and incorporate spatial averaging to reduce the effects of offsets. With the incorporation of the address-event representation (AER) communication protocol, spike-based circuitry will become commonplace, tying designs even closer to the details of the neurons we are attempting to simulate.

Another lesson learned from building large systems is that adaptation at every level is very important if we hope to continue building larger and larger systems. Many of the circuits presented here require tens of bias inputs each, with perhaps five critically-adjusted parameters. As these systems evolve into large, multi-chip projects, managing these hand-tuned parameters will become increasingly impractical, particularly with mobile systems which frequently change locations in their environment and need to optimize their sensors accordingly.

Software modelers typically begin with complex models using simple stimuli and struggle to make them simulate faster by removing detail or by waiting for faster machines to be available. Neuromorphic analog VLSI designers typically work with simple models, real-world stimuli, and do not worry about speed; rather they struggle to increase the complexity of the model to accomplish the given task. The latter approach is perhaps closer to the evolutionary process which created the brain and may hold the key to progress in our understanding of intelligence and its development in the years to come.

Appendix A Oculomotor Plant Circuits

Figure A.1 shows the external motor control circuits used in the oculomotor system. The two inputs to the system, LEFT and RIGHT, are shown on the far left at the J1 connector. These inputs are intended to be positive analog voltages between zero and 3.0 volts. The two spike trains from the burst generator are low-passed before driving these inputs. The right input is inverted and then summed to the left input to create a bipolar signal (Command Voltage).

The eye position is measured by the potentiometer on the left side of the circuit. This is first buffered and then passed into a filter summing the eye position and the derivative of the eye position (Modified Eye Position Voltage). Potentiometer #1 provides an offset current to allow adjustment of the output. Potentiometer #2 controls the gain of the reported eye position.

The difference between these two voltages is used to drive the left and right motors. The two differential amplifiers, with the Command Voltage and Modified Eye Position Voltage as their inputs, act as transconductance amplifiers which supply current to the spike generator circuits. Potentiometers #3 and #4 control the amount of transconductance and affect the gain of the spiking output.

The spike generator circuit consists of a spike capacitor, which receives current from the transconductance circuit, a threshold comparator, and a pulse generator chip (MC14538, monostable multivibrator chip). The bleed-resistors, potentiometers #5 and #6, and the input current charge up the spike capacitor until the voltage exceeds the threshold at 2.5 volts, generating a pulse in the output and discharging the capacitor via the n-channel FETs (VN10KM). Two pulses are triggered at the same time, a long pulse (1 msec) for driving the motor, and a shorter pulse (10 μsec) to discharge the capacitor.

The bleed-resistors are used to provide a tonic input to the system to keep both motors slightly active and to keep tension in the string connecting the motors and the eye. The longer pulse is critically adjusted to slightly move the eye with each pulse. This keeps the motor commands above the static friction limit where the motor control signals are lost. In equilibrium, both motors are "idling" sufficiently to continuously keep the eye moving. The potentiometer was specially chosen due to its extremely low friction and bearing strength (10 kOhm "Green Pot," Model CP-2FB(b) Midori Precision).

Figure A.2 shows the mechanical drawings of the box containing the dynamics simulation electronics (Figure A.1) and motors. Figure A.3 is a description of the lens-mount parts.

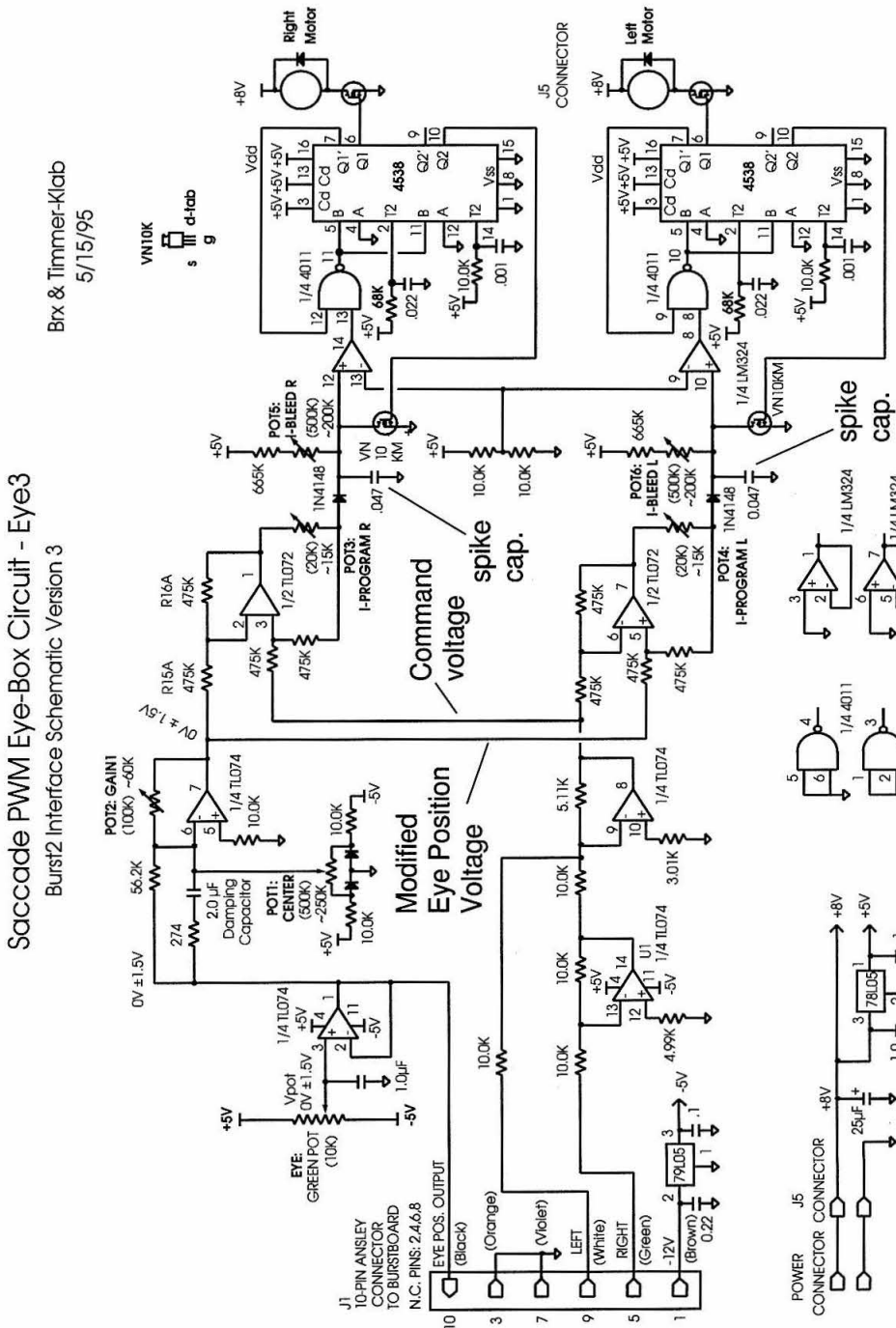

Figure A.1: Motor control circuit used in Chapter 2.

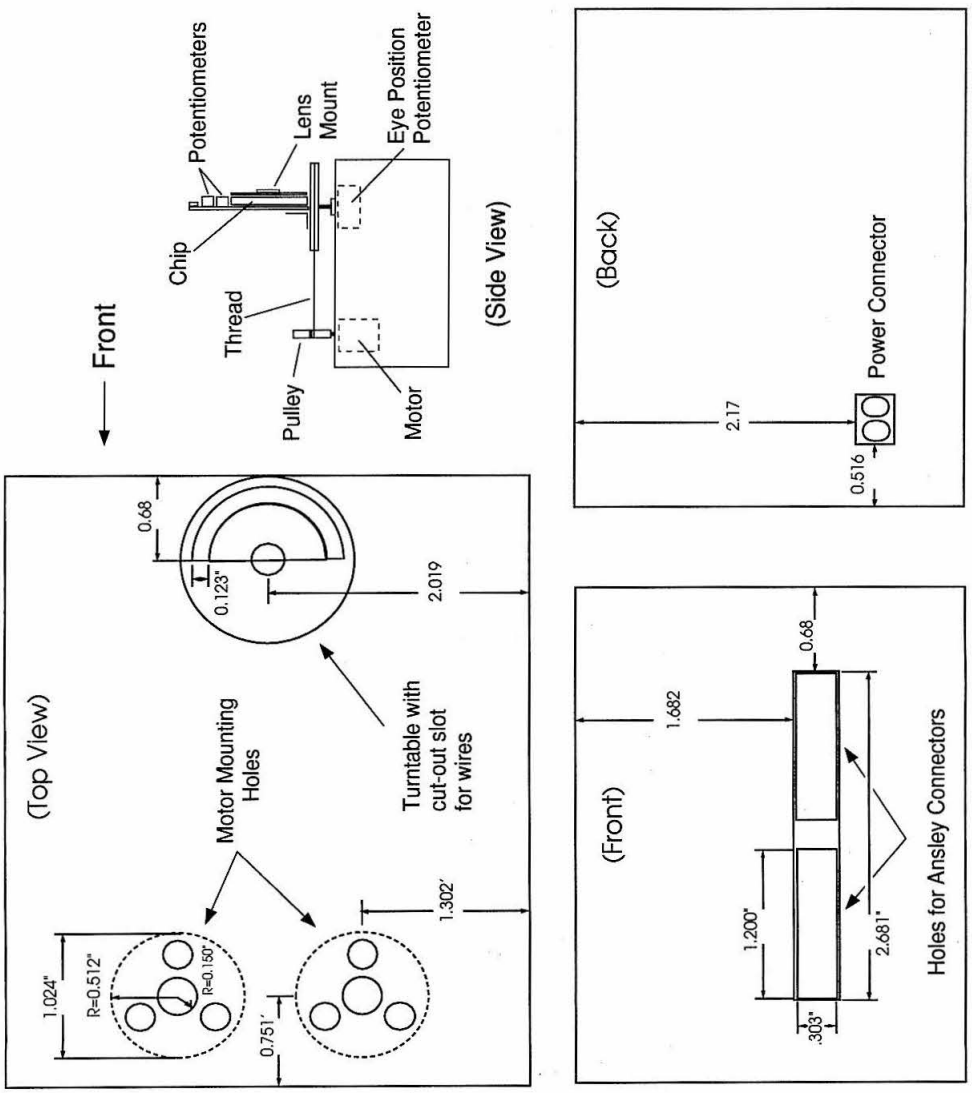

Figure A.2: Mechanical drawing of the oculomotor plant box. Drawing by Brooks Bishofberger.

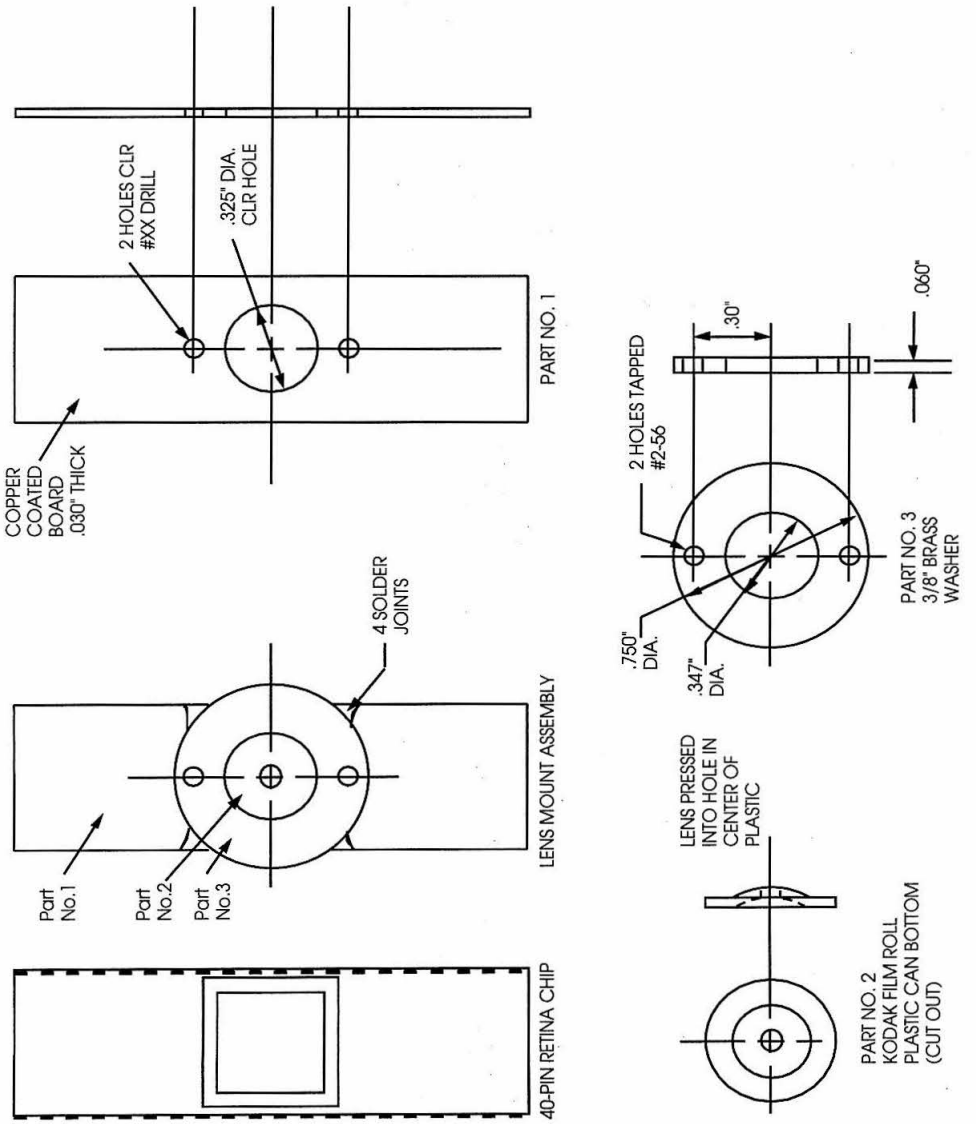

Figure A.3: Mechanical drawing of the lens mount. Drawing by Brooks Bishofberger.

Appendix B Auditory Circuits

Figure B.1 shows the auditory pre-filtering circuit for the localization system described in Chapter 3. Two copies of this circuit are used to provide the left and right input channels to the chip. This circuit receives acoustic input from a microphone, performs a bandpass filtering operation at 3.2 kHz, a thresholded zero-crossing detection stage and generates digital pulses as its output.

The microphone used is a Panasonic electret condenser microphone which is buffered, amplified, and then high-pass filtered at 100 Hz to remove the DC component then sent to the 3.2 kHz 2nd order bandpass active filter. The filter has a gain of 10 and a "Q" of 10.

The goal of the next stage is to trigger output pulses on the positive-going zero-crossing of the filtered waveform when the signal amplitude is large enough. The signal strength is determined by effectively measuring the time-derivative of the signal. Since the frequency is known, a high-pass filter with a sufficiently high cutoff frequency will produce a 90 degree phase-shifted signal. The simple RC filter chosen for this purpose was centered at 100 kHz, phase-shifting frequencies below 10 kHz by 90 degrees. This signal will have a positive peak on the rising phase of the original waveform. By comparing this "derivative" with a threshold, zero-crossing-triggered output pulses are only allowed when the original amplitude is large enough.

The output of the circuit is a pulse generator (MC14538) which connects to the analog VLSI chip using the reset function to shut off the output immediately after the chip has acknowledged receipt of the input pulse. Note that for every cycle of the waveform above threshold, a single pulse is injected into the localization chip.

The temporal derivative circuitry presented here was originally chosen to obtain the 90 degree phase-shift and limit the high-frequency gain due to noise in one step. The low-gain from the filter ultimately resulted in another gain stage which nullified any advantage. This could clearly be simplified by the appropriate use of high-pass filtering before an explicit temporal derivative circuit.

Differences in the left and right circuit gains were heuristically determined to balance the differences in the left and right microphones.

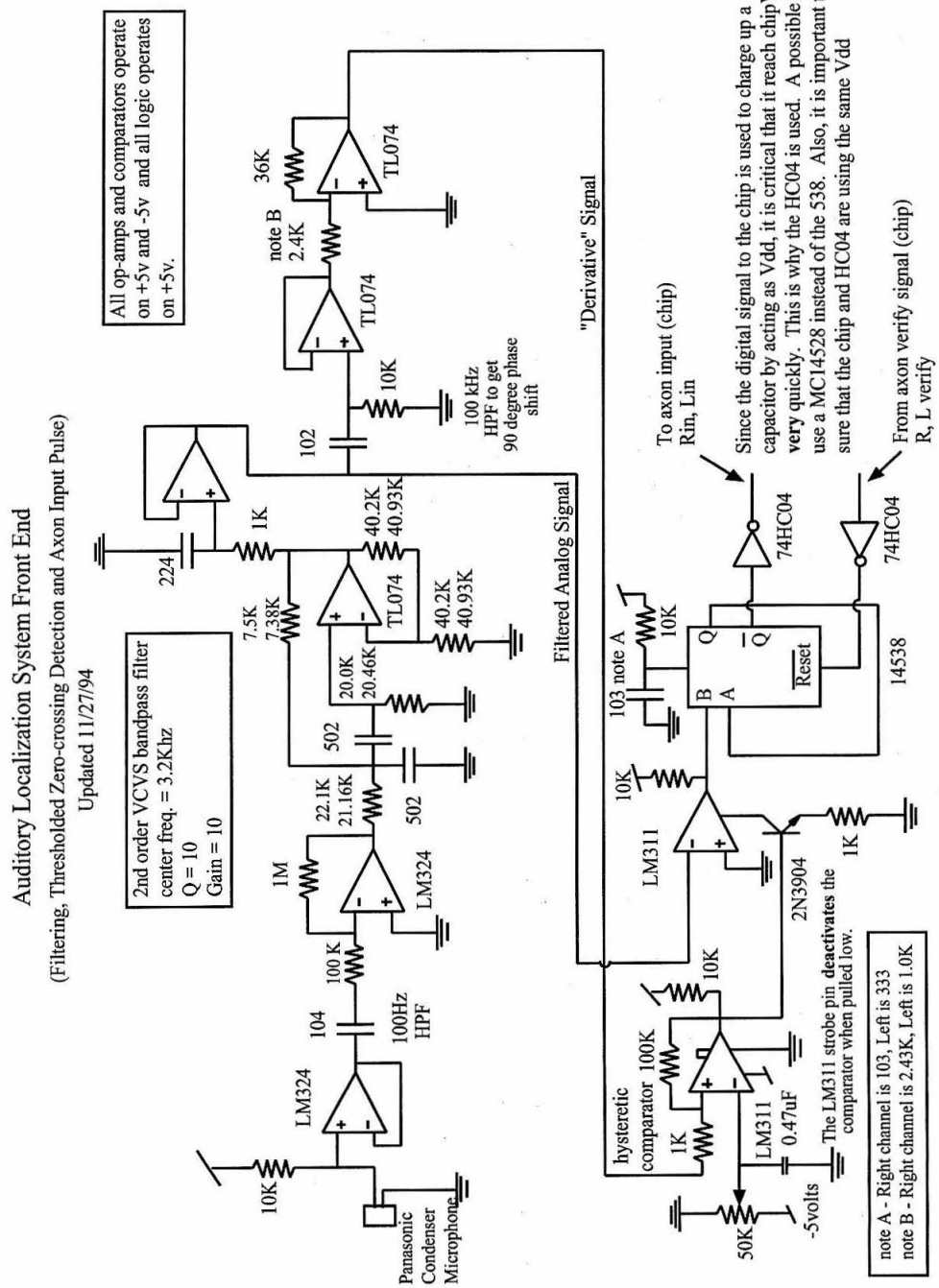

Figure B.1: Audio filtering circuit used in the auditory localization system of Chapter 3.

Appendix C Smooth Pursuit

C.0.3 Circuits

Figure C.1 shows the circuit used to model the smooth pursuit integrators in Chapter 7. The two inputs (far left) are *Vin* and $\overline{GO}$. The $\overline{GO}$ signal is received from the burst generator chip and indicates the duration of the out-going saccadic motor activity. The effective $\overline{GO}$ signal is extended by a fixed-duration pulse (MC14538, monostable multivibrator chip) triggered by the end of the $\overline{GO}$ pulse. During normal operation, *Vin* is compared to an internal reference voltage, *Vref*, by the upper and lower diffferential amplifiers. These amplifiers split the bipolar direction-of-motion signal into the leftward and rightward signal streams. The differential amplifier acts as a transconductance amplifier charging the 1 uF capacitor. During saccades, the $\overline{GO}$ signal switches the analog multiplexer (MC4051) such that both differential amplifiers have *Vref* on both inputs and the integrator remains at its current value.

The integration capacitor voltage is buffered and then compared to ground by a second differential amplifier which also acts as a transconductance amplifier, charging the 0.01 uF spike-generator capacitor. As the 0.01 uF capacitor charges up and exceeds 2.5 volts (determined by the 10K resistor-divider), the LM324 (operational amplifier), acting as a comparator, triggers a pulse in the output MC14538. The output pulse is fed back and discharges the 0.01 uF capacitor through the VN10K (n-channel FET) and the cycle begins again.

It is interesting to note that this circuitry would have been considerably easier to implement directly on-chip.

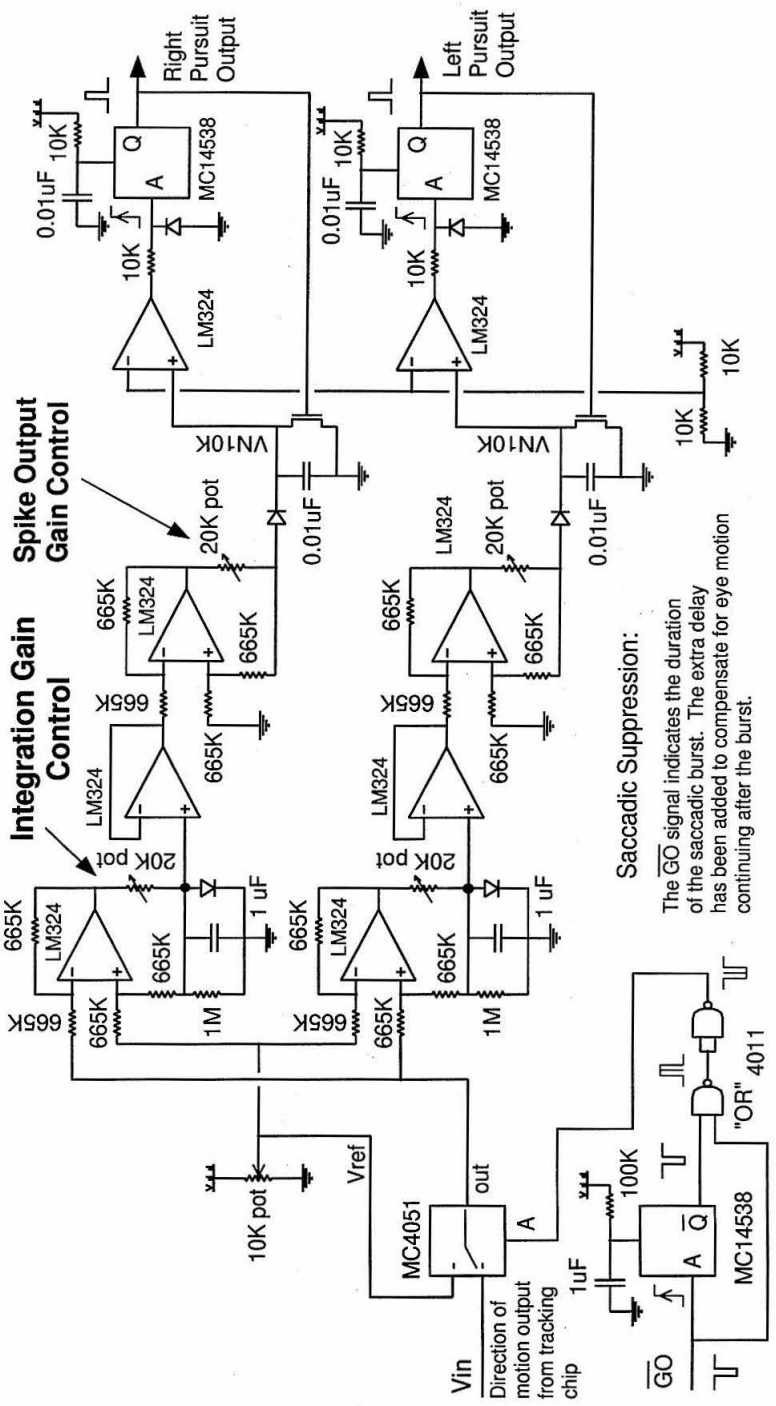

Figure C.1: Smooth Pursuit Integrator Circuit (See text for description)

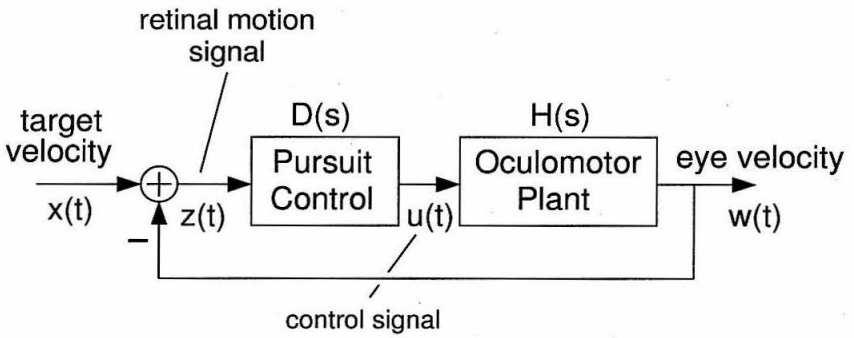

Figure C.2: The smooth pursuit control model is a negative feedback system which attempts to match eye velocity with target velocity. The visual motion system is assumed to be ideal, converting retinal motion to a motion signal linearly with unity gain. This is represented by the signal $z(t)$.

C.0.4 Control Systems Analysis

If we consider the case of an eye velocity controller (Figure C.2), the relevant terms in the equations of motion (assuming that the position-dependence is being compensated by the neural integrator) are the damping and the control signals. When we consider the system of an eye with rotational inertia, I, a damping coefficient, m, and a forcing function, $u(t)$, the equation of motion is:

$$I\dot{w}(t) = -mw(t) + u(t) \tag{C.1}$$

where $w(t)$ is the eye velocity. Computing the Laplace transform of this equation, we have:

$$IsW(s) = -mW(s) + U(s) \tag{C.2}$$

Solving for the transfer function of the oculomotor plant, $\frac{W(s)}{U(s)}$:

$$H(s) = \frac{W(s)}{U(s)} = \frac{1}{sI + m} \tag{C.3}$$

The integral control law used in the smooth pursuit system is:

$$u(t) = C \int_{-\infty}^{t} z(t)dt \tag{C.4}$$

leading to the Laplace transform:

$$U(s) = C(\frac{1}{s})Z(s) \tag{C.5}$$

Solving for the control law transfer function, we obtain:

$$D(s) = \frac{U(s)}{Z(s)} = \frac{C}{s} \tag{C.6}$$

From Figure C.2, we solve for the overall transfer function for the eye velocity as a function of the target's velocity, $\frac{W(s)}{X(s)}$.

$$W(s) = D(s)H(s)[X(s) - W(s)] \tag{C.7}$$

$$\frac{W(s)}{X(s)} = \frac{D(s)H(s)}{1 + D(s)H(s)} \tag{C.8}$$

Plugging equations C.3 and C.6 into C.8, to solve for the overall transfer function, we obtain:

$$\frac{W(s)}{X(s)} = \frac{\frac{C}{s}(\frac{1}{sI+m})}{1 + \frac{C}{s}(\frac{1}{sI+m})} = \frac{C}{s^2I + sm + C} \tag{C.9}$$

Solving for the poles of this transfer function, we find:

$$s = \frac{-m \pm \sqrt{m^2 - 4IC}}{2I} \tag{C.10}$$

In order to avoid complex poles,

$$m^2 - 4IC \geq 0 \tag{C.11}$$

$$m^2 \geq 4IC \tag{C.12}$$

$$\frac{m^2}{4I} \geq C \tag{C.13}$$

So in order to prevent oscillation, the integration constant must be kept low. This of course means that the integrator will take much longer to reach the desired target velocity. However, even if the system is operated with complex poles, the real part of the poles are negative, providing a damped oscillation.

Now, considering the case of a time-delay in the visual processing, we place a delay in the control law:

$$u(t) = C \int_{-\infty}^{t} z(t - t_d)dt \tag{C.14}$$

leading to the Laplace transform:

$$U(s) = C(\frac{1}{s})Z(s)e^{-st_d} \tag{C.15}$$

Solving for the new control law transfer function, we obtain:

$$D(s) = \frac{U(s)}{Z(s)} = \frac{Ce^{-st_d}}{s} \tag{C.16}$$

Solving again for the overall transfer function,

$$\frac{W(s)}{X(s)} = \frac{Ce^{-st_d}}{s^2I + sm + Ce^{-st_d}} \tag{C.17}$$

Solving for the poles of this new equation, we use the approximation: $e^x \approx 1 + x + \frac{x^2}{2}$ and obtain:

$$s = \frac{-(m - Ct_d) \pm \sqrt{(m - Ct_d)^2 - 4(I + \frac{Ct_d^2}{2})C}}{2(I + \frac{Ct_d^2}{2})} \tag{C.18}$$

In order to avoid complex poles,

$$\frac{(m - Ct_d)^2}{4(I + \frac{Ct_d^2}{2})} \geq C \tag{C.19}$$

This result means that C must be kept even smaller than in the delay-less case to avoid oscillation. Note also that in the case of complex poles, the real part of the poles have become less negative and therefore the damping of the oscillation will be smaller.

Bibliography

[Andreou et al., 1991] Andreou, A. G., Strohbehn, K., and Jenkins, R. E. (1991). Silicon retina for motion computation. In *Proc. of the 1991 IEEE Int. Symp. Circuits and Systems Conference*, pages 1373–1376. Singapore.

[Ballard, 1991] Ballard, D. H. (1991). Animate vision. *Artificial Intelligence*, 48:57–86.

[Barlow and Levick, 1965] Barlow, H. B. and Levick, W. R. (1965). The mechanism of directionally sensitive units in rabbit's retina. *J. Physiol.*, 178:477–504.

[Becker, 1989] Becker, W. (1989). Metrics. In Wurtz, R. H. and Goldberg, M. E., editors, *The Neurobiology of Saccadic Eye Movements*, pages 13–67. Elsevier Science Publishers B. V.

[Bedell and Lott, 1996] Bedell, H. E. and Lott, L. A. (1996). Suppression of motion-produced smear during smooth pursuit eye movements. *Current Biology*, 6(8):1032–1034.

[Benson and Delbrück, 1992] Benson, R. G. and Delbrück, T. (1992). Direction selective silicon retina that uses null inhibition. In Moody, Lippman, and Hanson, editors, *Advances in Neural Information Processing Systems 4*, pages 756–763. Morgan Kaufman, San Mateo, CA.

[Berg et al., 1996] Berg, Y., Sigvartsen, R. L., Lande, T. S., and Abusland, A. (1996). An analog feedforward neural-network with on-chip learning. *Analog Integrated Circuits*, 9(1):65–75.

[Boahen, 1996] Boahen, K. (1996). Retinomorphic vision systems. In *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems - MicroNeuro96*, pages 2–14. Feb. 12-14, 1996, Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA.

[Boman et al., 1996] Boman, D., Braun, D., and Hotson, J. (1996). Stationary and pursuit visual fixation share similar behavior. *Vision Research*, 36(5):751–763.

[Carley, 1947] Carley, L. R. (1947). Trimming analog circuits using floating-gate analog MOS memory. *IEEE J. Solid State Circ.*, 24(6):1569–1575.

[Castro et al., 1993] Castro, H., Tam, S., and Holler, M. (1993). Implementation and performance of an analog non-volatile neural-network. *Analog Integrated Circuit and Signal Processing*, 4(2):97–113.

[Collewijn and Erkelens, 1990] Collewijn, H. and Erkelens, C. J. (1990). Binocular eye movements and the perception of depth. In Kowler, E., editor, *Eye movements and their role in visual and cognitive processes*, pages 213–261. Elsevier Science Publishers B. V.

[Collewijn and Tamminga, 1984] Collewijn, H. and Tamminga, E. P. (1984). Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *J. Physiol.*, 351:217–250.

[Collins et al., 1975] Collins, C. C., O'Meara, D., and Scott, A. B. (1975). Muscle tension during unrestrained human eye movements. *J. Physiol.*, 245:351–369.

[Delbrück, 1991] Delbrück, T. (1991). Bump circuits for computing similarity and dissimilarity of analog voltages. In *Proc. of Intl. Joint Conf. on Neural Networks, Vol. 1*, pages I–475–479.

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

[Delbrück and Mead, 1989] Delbrück, T. and Mead, C. (1989). An electronic photoreceptor sensitive to small changes in intensity. In Touretzky, editor, *Neural Information Processing Systems 1*, pages 720–727. Morgan Kaufmann, San Mateo, CA.

[Desimone and Duncan, 1995] Desimone, R. and Duncan, J. (1995). Neural mechanisms of selective visual attention. In Cowan, W. M., Shooter, E. M., Stevens, C. F., and Thompson, R. F., editors, *Annual Reviews in Neuroscience, Vol. 18*, pages 193–222. Annual Reviews Inc., Palo Alto, CA.

[Deubel, 1995] Deubel, H. (1995). Separate adaptive mechanisms for the control of reactive and volitional saccadic eye movements. *Vision Res.*, 35(23/24):3529–3540.

[Devos et al., 1986] Devos, F., Zhang, M., Ni, Y., and Pône, J.-F. (1986). Centrally synthesized maps of sensory space. *TINS*, 4:163–168.

[DeWeerth, 1992] DeWeerth, S. P. (1992). Analog VLSI circuits for stimulus localization and centroid computation. *Intl. J. Comp. Vis.*, 8(3):191–202.

[DeWeerth and Morris, 1995] DeWeerth, S. P. and Morris, T. G. (1995). CMOS current mode winner-take-all with distributed hysteresis. *Electronics Letters*, 31(13):1051–1053.

[Diorio et al., 1995] Diorio, C., Mahajan, S., Hasler, P., Minch, B., and Mead, C. (1995). A high-resolution non-volatile analog memory cell. In *Proc. of the Intl. Symp. on Circuits and Systems*, pages 2233–2236. Seattle, WA.

[Douglas et al., 1995] Douglas, R., Mahowald, M., and Mead, C. (1995). Neuromorphic analogue VLSI. In Cowan, W. M., Shooter, E. M., Stevens, C. F., and Thompson, R. F., editors, *Annual Reviews in Neuroscience, Vol. 18*, pages 255–281. Annual Reviews Inc., Palo Alto, CA.

[Dusenbery, 1992] Dusenbery, D. B. (1992). *Sensory Ecology.* W. H. Freeman and Co., New York, NY.

[Etienne-Cummings et al., 1993] Etienne-Cummings, R., Fernando, S., Takahashi, N., Shtonov, V., Van der Spiegel, J., and Mueller, P. (1993). A new temporal domain optical flow measurement technique for focal plane VLSI implementation. In *Proc. Comp. Arch. Machine Perception Conference*, pages 241–250.

[Etienne-Cummings et al., 1996] Etienne-Cummings, R., Van der Spiegel, J., and Mueller, P. (1996). A visual smooth pursuit tracking chip. In Touretzky, Mozer, and Hasselmo, editors, *Advances in Neural Information Processing Systems 8*, pages 706–712. MIT Press, Cambridge, MA.

[Ferrara and Lisberger, 1995] Ferrara, V. and Lisberger, S. (1995). Attention and target selection for smooth pursuit eye movements. *J. Neurosci.*, 15(11):7472–7484.

[Fischer and Weber, 1993] Fischer, B. and Weber, H. (1993). Express saccades and visual-attention. *Behavioral and Brain Sciences*, 16(3):553–567.

[Frens and van Opstal, 1994] Frens, M. A. and van Opstal, A. J. (1994). Transfer of short-term adaptation in human saccadic eye movements. *Exp. Brain Res.*, 100:293–306.

[Glasser, 1985] Glasser, L. A. (1985). A UV write-enabled PROM. In Fuchs, H., editor, *Chapel Hill Conference on VLSI (1985)*, pages 61–65. Computer Science Press, Rockville, MD.

[Goldreich et al., 1992] Goldreich, D., Krauzlis, R. J., and Lisberger, S. G. (1992). Effect of changing feedback delay on spontaneous oscillations in smooth pursuit eye movements of monkeys. *J. Neurophysiol.*, 67(3):625–638.

[Groh and Sparks, 1992] Groh, J. and Sparks, D. (1992). 2 models for transforming auditory signals from head-centered to eye-centered coordinates. *Biol. Cybern.*, 67(4):291–302.

[Grzywacz and Poggio, 1990] Grzywacz, N. and Poggio, T. (1990). Computation of motion by real neurons. In Zornetzer, S. F., editor, *An Introduction to Neural and Electronic Networks*, pages 379–401. Academic Press, San Diego, CA.

[Guitton et al., 1984] Guitton, D., Douglas, R. M., and Volle, M. (1984). Eye-head coordination in cats. *J. Neurophysiol.*, 52(6):1030–1050.

[Hasler et al., 1995] Hasler, P., Diorio, C., Minch, B. A., and Mead, C. (1995). Single transistor learning synapses. In Tesauro, Touretzky, and Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 817–824. MIT Press, Cambridge, MA.

[Hepp et al., 1989] Hepp, K., Henn, V., Vilis, T., and Cohen, B. (1989). Brainstem regions related to saccade generation. In Wurtz, R. H. and Goldberg, M. E., editors, *The Neurobiology of Saccadic Eye Movements*, pages 105–212. Elsevier Science Publishers B. V.

[Hessler et al., 1993] Hessler, N. A., Shirke, A. M., and Malinow, R. (1993). The probability of transmitter release at a mammalian central synapse. *Nature*, 366:569–572.

[Hoffman and Subramaniam, 1995] Hoffman, J. and Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics*, 57(6):787–795.

[Horiuchi, 1995] Horiuchi, T. (1995). An auditory localization and coordinate transform chip. In Tesauro, Touretzky, and Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 787–794. MIT Press, Cambridge, MA.

[Horiuchi et al., 1994] Horiuchi, T., Bishofberger, B., and Koch, C. (1994). An analog VLSI saccadic system. In Cowan, Tesauro, and Alspector, editors, *Advances in Neural Information Processing Systems 6*, pages 582–589. Morgan Kaufmann, San Mateo, CA.

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

[Horiuchi and Koch, 1996] Horiuchi, T. K. and Koch, C. (1996). Analog VLSI circuits for visual motion-based adaptation of post-saccadic drift. In *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems - MicroNeuro96*, pages 60–66. Feb. 12-14, 1996, Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA.

[Horiuchi and Koch, 1997] Horiuchi, T. K. and Koch, C. (1997). Floating-gate circuits for adaptation of saccadic eye movement accuracy. In press.

[Horiuchi et al., 1997] Horiuchi, T. K., Morris, T. G., Koch, C., and DeWeerth, S. P. (1997). Analog VLSI circuits for attention-based visual tracking. In *Advances in Neural Processing Systems 9 (in press)*. MIT Press.

[Ilg et al., 1992] Ilg, U. J., Brenner, F., Thiele, A., and Hoffmann, K. P. (1992). Neuronal coding of retinal slip during smooth pursuit eye movements. *Eur. J. Neurosci. (Suppl)*, 5:253.

[Ilg and Hoffman, 1993] Ilg, U. J. and Hoffman, K.-P. (1993). Motion perception during saccades. *Vision Research*, 33(2):211–220.

[Jeffress, 1948] Jeffress, L. A. (1948). A place theory of sound localization. *J. Comp. Physiol. Psychol.*, 41:35–39.

[Jurgens et al., 1981] Jurgens, R., Becker, W., and Kornhuber, H. H. (1981). Natural and drug-induced variations of velocity and duration of human saccadic eye movements: Evidence for a control of the neural pulse generator by local feedback. *Biol. Cybern.*, 39:87–96.

[Kandel et al., 1991] Kandel, E. R., Schwartz, J. H., and Jessell, T. M. (1991). *Principles of Neuroscience, 3rd Edition*, chapter Muscles: Effectors of the Motor Systems, pages 548–563. Elsevier.

[Keller, 1978] Keller, E. L. (1978). Gain of the vestibulo-ocular reflex in monkey at high rotational frequencies. *Vision Research*, 20:535–538.

[Kerns, 1993] Kerns, D. A. (1993). *Experiments in Very Large-Scale Analog Computation.* PhD thesis, Electrical Engineering, California Institute of Technology.

[Khurana and Kowler, 1987] Khurana, B. and Kowler, E. (1987). Shared attentional control of smooth eye movement and perception. *Vision Research*, 27(9):1603–1618.

[King et al., 1986] King, W. M., Lisberger, S. G., and Fuchs, A. F. (1986). Oblique saccadic eye movements. *J. Neurophysiol.*, 56:769–784.

[Kirshfeld, 1994] Kirshfeld, K. (1994). Tracking of small objects in front of a textured background by insects and vertebrates: phenomena and neuronal basis. *Biol. Cybernetics*, 70:407–415.

[Koch and Ullman, 1985] Koch, C. and Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, 4:219–227.

[Konishi, 1986] Konishi, M. (1986). Centrally synthesized maps of sensory space. *TINS*, 4:163–168.

[Kowler, 1990] Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In Kowler, E., editor, *Eye movements and their role in visual and cognitive processes*, pages 1–70. Elsevier Science Publishers B. V.

[Kowler et al., 1995] Kowler, E., Anderson, E., Dosher, B., and Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research*, 35(13):1897–1916.

[Kramer, 1996a] Kramer, A. H. (1996a). Array-based analog computation: Principles, advantages and limitations. In *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems - MicroNeuro96*, pages 69–79. Feb. 12-14, 1996, Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA.

[Kramer, 1996b] Kramer, J. (1996b). Compact integrated motion sensor with three-pixel interaction. *IEEE Trans. Patt. Anal. and Mach. Intell.*, 18(4):455–460.

[Kramer et al., 1996] Kramer, J., Sarpeshkar, R., and Koch, C. (1996). Pulse-based analog VLSI velocity sensors. Submitted for publication to IEEE Trans. Circuits and Systems II, April 1995.

[Krommenhoek et al., 1993] Krommenhoek, K. P., Van Opstal, A. J., Gielen, C. C. A. M., and Van Gisbergen, J. A. M. (1993). Remapping of neural activity in the motor colliculus: A neural network study. *Vision Research*, 33(9):1287–1298.

[Land, 1995] Land, M. F. (1995). The functions of eye movements in animals remote from man. In Findlay, Walker, and Kentridge, editors, *Eye Movement Research: Mechanisms, Processes, and Applications*, pages 63–76. Elsevier Science, B. V.

[Lande et al., 1996] Lande, T. S., Ranjbar, H., Ismail, M., and Berg, Y. (1996). An analog floating-gate memory in a standard digital technology. In *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems - MicroNeuro96*, pages 271–276. Feb. 12-14, 1996, Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA.

[Lazzaro, 1990] Lazzaro, J. (1990). *Silicon Models of Early Audition.* PhD thesis, Computer Science Department, California Institute of Technology.

[Lazzaro, 1992] Lazzaro, J. (1992). Temporal adaptation in a silicon auditory nerve. In Moody, Lippman, and Hanson, editors, *Advances in Neural Information Processing Systems 4*, pages 813–820. Morgan Kaufman, San Mateo, CA.

[Lazzaro and Mead, 1989] Lazzaro, J. and Mead, C. (1989). Circuit models of sensory transduction in the cochlea. In Mead, C. and Ismail, M., editors, *Analog VLSI Implementations of Neural Networks*, pages 85–101. Kluwer Academic Publishers, Norwell, MA.

[Lazzaro et al., 1989] Lazzaro, J., Ryckebusch, S., Mahowald, M. A., and Mead, C. (1989). Winner-take-all networks of o(n) complexity. In Touretzky, editor, *Advances in Neural Information Processing Systems 1*, pages 703–711. Morgan Kaufmann, San Mateo, CA.

[Lazzaro et al., 1991] Lazzaro, J., Schwaber, J. S., and Rogers, W. T. (1991). Silicon baroreceptors: Modeling cardiovascular pressure transduction in analog VLSI. In *Advanced Research in VLSI, Proc. of the 1991 Santa Cruz Conference*, pages 163–177. MIT Press, Cambridge, MA.

[Lazzaro et al., 1994] Lazzaro, J., Wawrzynek, J., and Kramer, A. (1994). Systems technologies for silicon auditory models. *IEEE Micro*, 14(3):7–15.

[Lazzaro et al., 1993] Lazzaro, J., Wawrzynek, J., Mahowald, M., Sivilotti, M., and Gillespie, D. (1993). Silicon auditory processors as computer peripherals. *IEEE Trans. on Neural Networks*, 4(3):523–528.

[Lisberger et al., 1987] Lisberger, S. G., Morris, E. J., and Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. In Cowan, W. M., Shooter, E. M., Stevens, C. F., and Thompson, R. F., editors, *Annual Reviews in Neuroscience, Vol. 10*, pages 97–129. Annual Reviews Inc., Palo Alto, CA.

[Liu and Mead, 1994] Liu, S. and Mead, C. (1994). Continuous-time adaptive delay system. *IEEE Circuits and Systems II*, pages 744–751.

[Longtin and Hinzer, 1996] Longtin, A. and Hinzer, K. (1996). Encoding with bursting, subthreshold oscillations, and noise in mammalian cold receptors. *Neural Computation*, 8(2):215–255.

[Mahowald, 1994] Mahowald, M. (1994). *An Analog VLSI Stereoscopic Vision System.* Kluwer Academic Pub., Boston, MA.

[Mahowald, 1996] Mahowald, M. (1996). *VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function.* PhD thesis, Department of Computer Science, California Institute of Technology.

[Mahowald and Douglas, 1991] Mahowald, M. and Douglas, R. (1991). A silicon neuron. *Nature*, 354:515–518.

[Masson et al., 1995] Masson, G., Proteau, L., and Mestre, D. (1995). Effects of stationary and moving textured backgrounds on the visuo-oculo-manual tracking in humans. *Vision Research*, 35(6):837–852.

[McKenzie and Lisberger, 1986] McKenzie, A. and Lisberger, S. G. (1986). Properties of signals that determine the amplitude and direction of saccadic eye movements in monkeys. *J. Neurophysiol.*, 56(1):196–207.

[Mead, 1989a] Mead, C. (1989a). Adaptive retina. In Mead, C. and Ismail, M., editors, *Analog VLSI Implementation of Neural Systems*, pages 239–246. Kluwer Academic Publishers, Boston, MA.

[Mead, 1989b] Mead, C. (1989b). *Analog VLSI and Neural Systems*, chapter Axons, pages 193–204. Addison-Wesley, Menlo Park, CA.

[Mead, 1990] Mead, C. (1990). Neuromorphic electronic systems. *Proc. IEEE*, 78(10):1629–1636.

[Moore and Koch, 1991] Moore, A. and Koch, C. (1991). A multiplication based analog motion detection chip. *SPIE Visual Information Procesing: From Neurons to Chips*, 1473:66–75.

[Moran and Desimone, 1985] Moran, J. and Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229:782–784.

[Morris, 1996] Morris, T. G. (1996). *Analog VLSI Visual Attention Systems.* PhD thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology.

[Morris and DeWeerth, 1996] Morris, T. G. and DeWeerth, S. P. (1996). Analog VLSI circuits for covert attentional shifts. In *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems - MicroNeuro96*, pages 30–37. Feb. 12-14, 1996, Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA.

[Mortara et al., 1995] Mortara, A., Vittoz, E., and Venier, P. (1995). A communication scheme for analog vlsi perceptive systems. *IEEE Trans. Solid-State Circ.*, 30(6):660–669.

[Motter, 1993] Motter, B. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.*, 70:909–919.

[Nichols and Sparks, 1995] Nichols, M. J. and Sparks, D. L. (1995). Non-stationary properties of the saccadic system - new constraints on models of saccadic control. *J. Neurophysiol.*, 73(1):431–435.

[Niebur et al., 1993] Niebur, E., Koch, C., and Rosin, C. (1993). An oscillation-based model for the neuronal basis of attention. *Vision Research*, 33(18):2789–2802.

[Olshausen et al., 1993] Olshausen, B. A., Anderson, C. H., and Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *J. Neurosci.*, 13(11):4700–4719.

[Optican and Miles, 1985] Optican, L. M. and Miles, F. A. (1985). Visually induced adaptive changes in primate saccadic oculomotor control signals. *J. Neurophysiol.*, 54(4):940–958.

[Optican and Robinson, 1980] Optican, L. M. and Robinson, D. A. (1980). Cerebellar-dependent adaptive control of the primate saccadic system. *J. Neurophysiol.*, 44:1058–1076.

[Pavasovic et al., 1994] Pavasovic, A., Andreou, A. G., and Westgate, C. R. (1994). Characterization of subthreshold MOS mismatch in transistors for VLSI systems. *Analog Integrated Circuits and Signal Processing*, 6:75–85.

[Pitts and McCulloch, 1947] Pitts, W. and McCulloch, W. S. (1947). How we know universals: the perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9:127–147.

[Posner, 1980] Posner, M. (1980). Orienting of attention. *Quart. J. Exp. Psychol.*, 32:3–25.

[Pouget and Sejnowski, 1995] Pouget, A. and Sejnowski, T. J. (1995). Spatial representations in the parietal cortex may use basis functions. In Tesauro, Touretzky, and Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 157–164. MIT Press, Cambridge, MA.

[Rafal et al., 1989] Rafal, R., Calabresi, P., Brennan, C., and Scioltio, T. (1989). Saccade preparation inhibits reorienting to recently attended locations. *J. Exp. Psych: Hum. Percep. and Perf.*, 15:673–685.

[Rao and Ballard, 1995] Rao, R. and Ballard, D. (1995). Learning saccadic eye movements using multiscale spatial filters. In Tesauro, Touretzky, and Leen, editors, *Advances in Neural Information Processing Systems 7*, pages 893–900. MIT Press, Cambridge, MA.

[Rashbass, 1961] Rashbass, C. (1961). The relationship between saccadic and smooth tracking eye movements. *J. Physiol.*, 159:326–338.

[Ringach, 1995] Ringach, D. (1995). A 'tachometer' feedback model of smooth pursuit eye movements. *Biol. Cybern.*, 73:561–568.

[Ritter et al., 1992] Ritter, H., Martinetz, T., and Schulten, K. (1992). *Neural Computation and Self Organizing Maps: An Introduction.* Addison-Wesley, Reading, MA.

[Robinson, 1964] Robinson, D. A. (1964). The mechanics of human saccadic eye movement. *J. Physiol.*, 174:245–264.

[Robinson, 1975] Robinson, D. A. (1975). *Basic Mechanisms of Ocular Motility and Their Clinical Implications*, chapter Oculomotor Control Signals, pages 337–374. Pergamon Press.

[Robinson et al., 1991] Robinson, D. L., McClurkin, J. W., Kertzman, C., and Petersen, S. E. (1991). Visual responses of pulvinar and collicular neurons during eye-movements of awake, trained macaques. *J. Neurophysiol.*, 66(2):485–496.

[Saarinen and Julesz, 1991] Saarinen, J. and Julesz, B. (1991). The speed of attentional shifts in the visual field. *Proc. Natl. Acad. Sci.*, 88:1812–1814.

[Sarpeshkar et al., 1993] Sarpeshkar, R., Bair, W., and Koch, C. (1993). Visual motion computation in analog VLSI using pulses. In Hanson, Cowan, and Giles, editors, *Advances in Neural Information Processing Systems 5*, pages 781–788. Morgan Kaufman.

[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. *Proc. of the IEEE*, 84(7):969–987.

[Schnapf and Baylor, 1987] Schnapf, J. L. and Baylor, D. A. (1987). How photoreceptor cells respond to light. *Scientific American*, 256(4):40–47.

[Shadlen and Newsome, 1994] Shadlen, M. N. and Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.*, 4:569–579.

[Shimojo et al., 1995] Shimojo, S., Tanaka, Y., Hikosaka, O., and Miyauchi, S. (1995). Vision, attention, and action – inhibition and facilitation in sensory motor links revealed by the reaction time and the line-motion. In Inui, T. and McClelland, J. L., editors, *Attention and Performance XVI.* MIT Press.

[Sivilotti, 1991] Sivilotti, M. (1991). *Wiring Considerations in Analog VLSI Systems, with Application to Field-Programmable Networks.* PhD thesis, Dept. of Computer Science, California Institute of Technology.

[Sparks and Hartwich-Young, 1989] Sparks, D. L. and Hartwich-Young, R. (1989). The deep layers of the superior colliculus. In Wurtz, R. H. and Goldberg, M. E., editors, *The Neurobiology of Saccadic Eye Movements*, pages 213–255. Elsevier Science Publishers B. V.

[Stemmler, 1996] Stemmler, M. (1996). A single spike suffices: the simplest form of stochastic resonance in model neurons. *Network: Computation in Neural Systems*, 7:1–29.

[Strausfeld, 1976] Strausfeld, N. J. (1976). *Atlas of an Insect Brain.* Springer-Verlag, New York, NY.

[Tam and Ono, 1994] Tam, W. J. and Ono, H. (1994). Fixation disengagement and eye-movement latency. *Perception and Psychophysics*, 56(3):251–260.

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

[Tovée, 1994] Tovée, M. J. (1994). How fast is the speed of thought? *Current Biology*, 4(12):1125–1127.

[Treisman and Gelade, 1980] Treisman, A. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136.

[Treue and Maunsell, 1996] Treue, S. and Maunsell, J. (1996). Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*, 382:539–541.

[Wang and Akers, 1995] Wang, T. and Akers, L. (1995). Hardware implementation of habituation. In *Proc. of the 1995 IEEE Int. Conf. on Circuits and Systems, Vol 3.*, page 2092. Seattle.

[Westheimer and McKee, 1975] Westheimer, G. A. and McKee, S. (1975). Visual acuity in the presence of retinal image motion. *J. Opt. Soc. Am*, 65:847–850.

[Wolpert et al., 1993] Wolpert, D. M., Miall, R. C., Kerr, G. K., and Stein, J. F. (1993). Ocular limit cycles induced by delayed retinal feedback. *Exp. Brain Res.*, 96:173–180.

[Zipser and Andersen, 1988] Zipser, D. and Andersen, R. A. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331:679–684.