??, ??, 1–9 (??)

Floating-Gate Circuits for Adaptation of Saccadic Eye Movement Accuracy

TIMOTHY K. HORIUCHI & CHRISTOF KOCH

Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 91125

timmer@klab.caltech.edu

Received ??. Revised ??.

Abstract.

In this paper we describe an analog VLSI circuit, fabricated using a standard 2 μm, n-well, BiCMOS process, which utilizes floating-gate structures for non-volatile, on-chip, analog parameter storage. This circuit is designed to operate in the context of a hardware model of the primate oculomotor system and performs visually-guided, saccadic adaptation. The chip contains a one-dimensional array of photoreceptors and floating-gate circuits which are used to map retinal positions to motor output commands. The system's functionality is demonstrated by training the chip with several different mapping functions using a supervised-learning technique.

Keywords: neuromorphic analog VLSI, saccadic eye movements, learning, floating-gate, short-term adaptation

1. Introduction

The most common eye movements in primates are the quick reorienting movements known as saccades. Our eyes often reach speeds up to 750 degs/s during a saccade which severely impairs our visual acuity. It is therefore important to minimize the time during which the eyes are moving. While typical human saccades have a duration of 40ms to 150 ms, changes in the optics, the oculomotor plant, or the underlying neural circuitry can cause deficits which delay optimal viewing conditions.

There have been many types of adaptation behavior identified in the primate oculomotor system in response to different induced deficits. For example, Optican and Robinson [11] showed that weakening of the horizontal recti muscles in the rhesus monkey initially caused saccades which fell short and exhibited post-saccadic drift of the eyeball. Recovery from this type of damage, which affects all saccades in a given direction, requires about 3-5 days. In contrast to this long adaptation period, which involves hundreds of thousands of saccades, experiments where saccadic targets are moved a short distance during the saccade, require only hundreds of trials for the adaptation to reach steady-state. This type of visually-induced modification of saccade amplitude is known as short-term adaptation. Experiments by Frens and van Opstal [6] show this adaptation to be confined to a limited range of saccade vectors around the adaptation target.

Our laboratory is involved in building a hardware model of the primate oculomotor system [10] using analog VLSI circuitry. The model oculomotor plant simulates linear dynamics and provides a good foundation upon which to build biologically-realistic eye movement systems. The current system can be triggered using both visual and auditory stimuli [9] and begins to model the convergence of multi-modal spatial information at the level of the superior colliculus. Three aspects of this modelling system that have become important are the compensation for non-linearities, the need for self-calibration, and on-chip storage of these parameters.

The intermingling of memory and computation is an important and powerful aspect of neu-

ral architectures which has not yet been well exploited in neuromorphic 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. The majority of circuit designs which have attempted to use on-chip storage of parameters have used digital RAM or externally-refreshed, capacitive storage, both of which are generally bulky and low-precision. Until recently, the use of floating-gates (a MOS transistor gate completely isolated from the circuit by silicon dioxide) required the use of ultra-violet radiation or bidirectional tunneling processes which have also been fraught with difficulties, impeding their widespread use. The development of a complementary strategy of tunneling and hot-electron injection [7] in a commercially-available BiCMOS process has alleviated some of these difficulties.

In previous work, we demonstrated the use of floating-gate devices in a model of the saccadic burst generator to reduce post-saccadic drift using visual motion as an error signal [8]. In this paper we present a chip which uses floating-gate structures to store a mapping of retinal position to motor command voltage. In the beginning of the next section we will discuss the architecture of the chip, in section 3 we present the circuits and their behavior, and in section 4 we discuss the chip's performance within the training system.

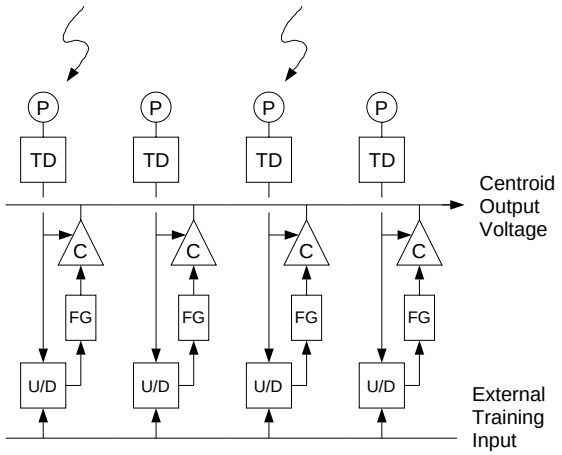

Fig. 1. System Block Diagram: This chip consists of an array of 32 pixels which consist of an adaptive photoreceptor (P), a temporal derivative circuit (TD), a centroid circuit (C), a floating-gate circuit (FG) which provides reference voltages to the centroid circuit, and a control circuit (U/D = "up/down") for training the floating-gate.

2. Vector-Specific Adaptation

In an experiment where human subjects are performing saccades from a fixation point to a visible target, if the target of a specific retinotopic position consistently moves to a new location during the saccade, Frens and van Opstal (1994) have shown results indicating that the adaptation time-course 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 [6]. This type of learning can be explained by a mapping similar to that of a look-up table.

In the previous implementation of our analog VLSI-based saccadic system [10], 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. We have modified the visual-triggering circuit [10] (also in figure 2) to use the output of a floating-gate circuit to determine the proper motor command for each pixel.

As shown in Figure 1, an array of adaptive photoreceptor circuits (P) are used to drive a temporal-derivative circuit (TD), activating regions in an image where the 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 an output signal from the chip, indicating that something has occurred in the image and that the centroid output information is "valid". The centroid circuits [4] require reference voltages (motor command voltages) at each pixel which represent the saccade vector required to center the stimulus on the center of the array.

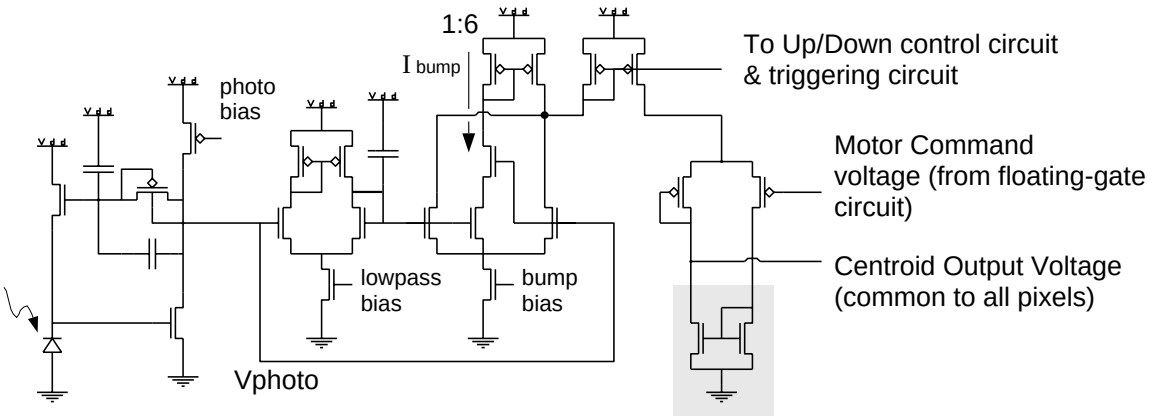

Fig. 2. Temporal Triggering Circuit (P+TD+C): On the far left, the adaptive photoreceptor circuit amplifies temporal change in the light intensity while slowly adapting to the mean light level. The temporal-derivative (TD) circuit acts as a high-pass filter by measuring the difference in voltage between 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).

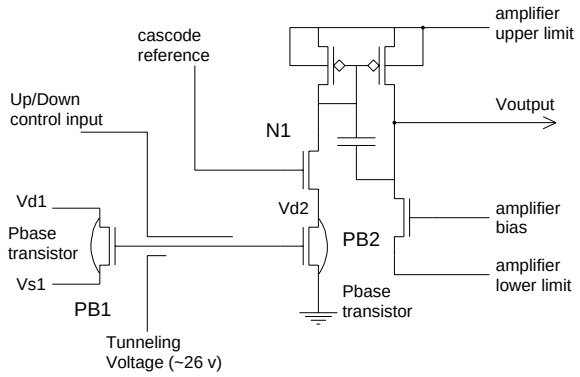

Fig. 3. Floating-gate amplifier circuit (FG): The floating node defines a subthreshold current in transistor PB2 which is mirrored and used in a high-gain amplifier stage which has variable output limits. Cascode transistor N1 defines PB2's drain voltage to prevent hot-electron injection. Nodes Vd1, 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.

In previous versions of this visually-based, triggering circuit [10], 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. [13] and by Rao and Ballard [12].

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 GPIB-equipped (General Purpose Interface Bus) 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 in-

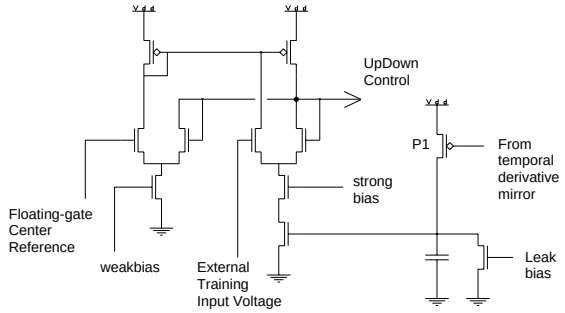

Fig. 4. Up/Down learning control circuit (U/D): This circuit consists of two competing followers, a weak follower carrying the center reference voltage and a stronger follower which receives the training voltage from off-chip. When a given pixel in the array generates a pulse of current in the TD circuit, 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 ampifier 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.

put voltage, (driven by a GPIB-equipped 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.

3. Circuits

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 we discuss in this paper is a one-dimensional array of 32 pixel elements.

Figure 2 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 [1] 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 [2] to signal the absolute-value of the temporal-derivative. The centroid circuit [4] 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 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.

The floating-gate circuit (figure 3), is a modification of the circuit used by Hasler et al. [7] 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. 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 3) 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 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

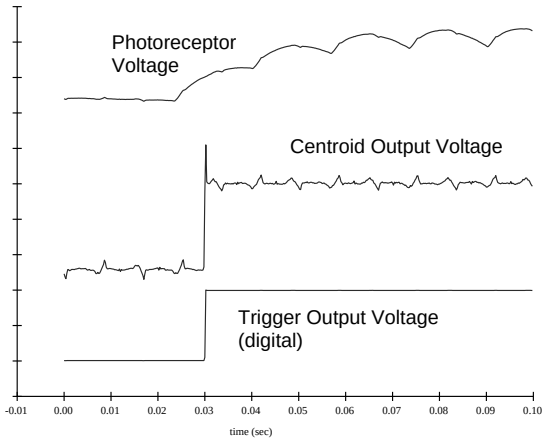

Fig. 5. Top trace: Photoreceptor voltage, Middle trace: Centroid output voltage (analog), Bottom trace: trigger signal (digital). The photoreceptor output voltage jumps from 0.96 volts to 1.30 volts during the flash of the stimulus. The oscillation riding on the step response of the photoreceptor is 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.

tunneling voltage and decreases the source current in PB1.

As in the system described by Hasler et al. [7], 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 our floating-gates exhibited extremely 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. [7] and Diorio et al. [5]

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 4 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 shows some of the relevant signals during a pulse of light on the array. Although not visible, the centroid output rises to a stable value approximately 2ms after the beginning of the temporal change.

The data presented in this paper 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

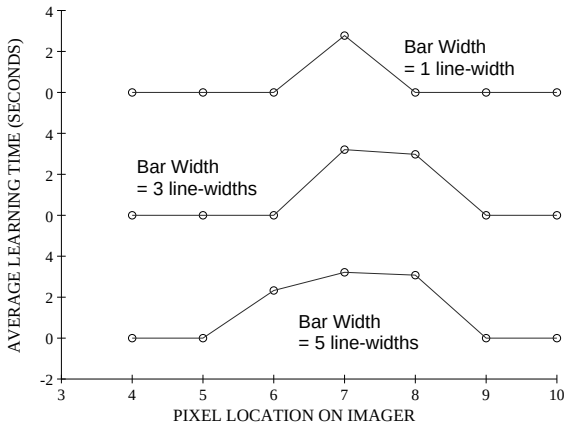

Fig. 6. When a bar of one line-width (defined by the graphics board) is flashed at the chip, it stimulates a single photoreceptor as shown in the top plot and the one pixel is trained for a mean duration of 2.75 seconds. This timing is primarily determined by the leak bias (see Figure 4). When the bar is widened to 3 line-widths (middle plot), 2 adjacent pixels are stimulated and they are trained together in the same direction. A bar width of 5 line-widths stimlates 3 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 5 line-widths was also used to generate Figure 10. 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 7 trials.

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.

4. System Performance

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

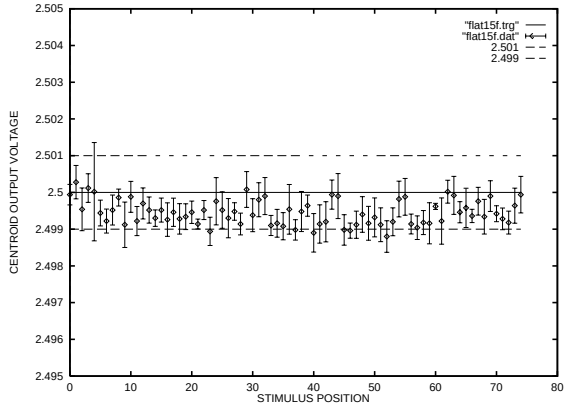

Fig. 7. Flat Target Function - In this case, all stimulus positions were trained to lie at 2.500 volts. This plot shows the performance of the chip after approximately 20,000 presentations spread over 75 positions. The floating-gate outputs were initially spread between 2.4 and 2.6 volts. After training, 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.

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 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 5 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 3 pixels. (See Figure 6) The effects of averaging can be seen in Figure 10 for the case of the sinewave mapping, which is a particularly difficult case to learn, since individual pixels can-

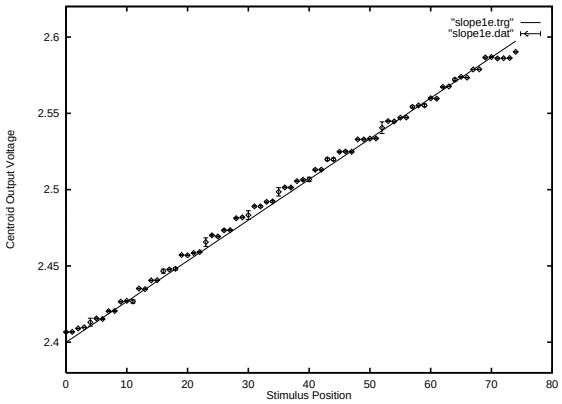

Fig. 8. Linear Target Function - This function most closely represents a realistic sensorimotor mapping function for triggering saccades to a visual target. The training and testing procedure is the same as in the previous graph. The error bars represent one standard deviation.

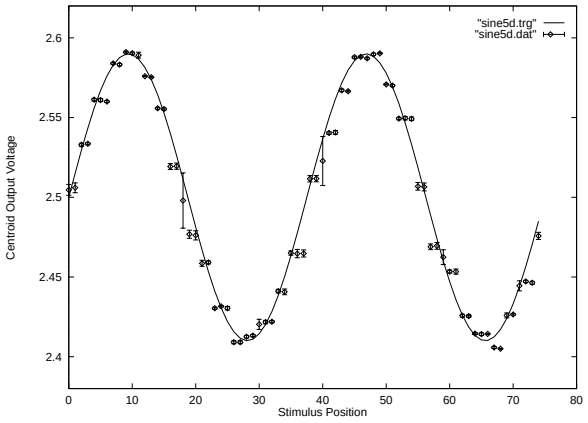

Fig. 9. Sinewave Target Function - In this case, the target values followed a sinewave. Photoreceptor granularity is evident by the "staircasing" seen in the plot. Stimulus locations where the flashed bar occurs on the boundary of two pixels exhibit large variations in output voltage due to the narrow (one line-width) stimuli being used. Figure 10 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.

not satisfy the wide range of values occuring 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 7 shows the results after extended training. From initial conditions where the floating-gate amplifier outputs

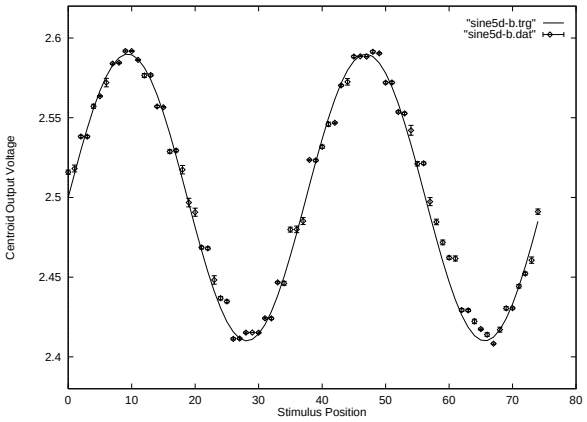

Fig. 10. Sinewave Target Function - In this case, the evaluation of the pattern in Figure 9 was performed using a bar which spanned 3 pixels. The training and testing procedure is the same as in the previous graphs. The error bars represent one standard deviation.

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 7 is in the center 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 8) 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 9 and 10) 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 9. 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, we 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. We partially solved this problem 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.

5. Discussion

We have successfully fabricated and tested a trainable array of floating-gate memories whose operation and modification is integrally related to a specific visual task. 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 [3] 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 [6] 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. Following these constraints, it is our hope to also demonstrate this transfer of adaptation with our VLSI-based auditory localization system.

The investigation of neural information processing architectures in analog VLSI can provide insight into the issues that biological nervous systems face. Analog VLSI architectures share many of the advantageous properties with neural systems such as speed, space-efficiency, and lower power consumption. In addition, analog VLSI must face similar constraints such as real-world noise, component variability or failure, and interconnection limitations. With the development of reliable floating-gate circuits, the powerful ability of neural systems to modify and store their parameters locally can finally be realized in analog VLSI.

Beyond our effort to understand neural systems by building large-scale, physically-embodied biological models, adaptive analog VLSI sensorimotor systems can be applied to many commercial and industrial applications involving self-calibrating actuation systems. In particular, we believe that for real-world tasks such as mobile robotics or remote sensing, these circuits will be invaluable for systems trying to keep up with the ever-changing world.

6. Acknowledgements

The authors would like to thank Reid Harrison for constructing one of the test boards and generally motivating the completion of this project, Paul Hasler for valuable technical advice on the design and analysis of the floating-gate cell, Tonia Morris and Giacomo Indiveri for assistance in the careful editing of this document. T.H. is supported by an Office of Naval Research AASERT grant and by the NSF Center for Neuromorphic Systems Engineering at Caltech.

References

1. T. Delbrück, Ph.D. Thesis, Computation and Neural Systems Program California Institute of Technology, 1993.
2. T. Delbrück, "Bump circuits for computing similarity and dissimilarity of analog voltages" in *Proc. of Intl. Joint Conf. on Neural Networks*, vol 1. pp. I-475-479, 1991.
3. H. Deubel, "Separate Adaptive Mechanisms for the Control of Reactive and Volitional Saccadic Eye Movements" *Vision Res.* Vol. 35, No. 23/24, pp. 3529-3540, 1995.
4. S. P. DeWeerth, "Analog VLSI Circuits for Stimulus Localization and Centroid Computation" *Intl. J. Comp. Vis.* 8(3), pp. 191-202, 1992.
5. C. Diorio, S. Mahajan, P. Hasler, B. Minch, C. Mead, "A High-Resolution Non-Volatile Analog Memory Cell" *Proc. of the Intl. Symp. on Circuits and Systems*, Seattle, WA, USA, pp. 2233-2236, 1995.
6. M. A. Frens, and A. J. van Opstal "Transfer of short-term adaptation in human saccadic eye movements" *Exp. Brain Res.* 100:293-306, 1994.
7. P. Hasler, C. Diorio, B. A. Minch, C. Mead, "Single Transistor Learning Synapses" *Advances in Neural Information Processing Systems 7*, MIT Press, pp. 817-824, 1995.
8. T. K. Horiuchi, C. Koch, "Analog VLSI Circuits for Visual Motion-Based Adaptation of Post-Saccadic Drift" *Proc. 5th Intl. Conf. on Microelectronics for Neural Networks and Fuzzy Systems* - MicroNeuro96, Feb 12-14, 1996. Lausanne, Switzerland, IEEE Computer Society Press, Los Alamitos, CA, pp. 60-66, 1996.
9. T. K. Horiuchi, "An Auditory Localization and Coordinate Transform Chip" *Advances in Neural Information Processing Systems 7*, MIT Press, pp. 787-794, 1995.
10. T. K. Horiuchi, B. Bishofberger, and C. Koch, "An Analog VLSI Saccadic System" *Advances in Neural Information Processing Systems 6*, Morgan Kaufmann, pp. 582-589, 1994.
11. L. M. Optican and D. A. Robinson, "Cerebellar-Dependent Adaptive Control of the Primate Saccadic System" *J. Neurophysiol.* Vol. 44, pp. 1058-1076, 1980.
12. R. Rao, and D. Ballard "Learning Saccadic Eye Movements Using Multiscale Spatial Filters" *Advances in Neural Information Processing Systems 7*, MIT Press, pp. 893-900, 1995.
13. H. Ritter, T. Martinetz, and K. Schulten, *Neural Computation and Self Organizing Maps: An Introduction* Addison-Wesley: Reading, MA, 1992.