A VLSI-Based Model of Azimuthal Echolocation in the Big Brown Bat

TIMOTHY HORIUCHI

Electrical and Computer Engineering Department, and the Institute for Systems Research, A.V. Williams Bldg., University of Maryland, College Park, MD 20742, USA

KAI M. HYNNA

Department of Bioengineering, 120 Hayden Hall, 3320 Smith Walk, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract. The azimuthal localization of objects by echolocating bats is based on the difference of echo intensity received at the two ears, known as the interaural level difference (ILD). Mimicking the neural computation of ILD in bats, we have constructed a spike-based VLSI model of the lateral superior olive (LSO) that can successfully produce direction-dependent responses. This simple algorithm, while studied in the acoustic domain, is applicable to any localization based on direction-dependent signal attenuation differences.

Au.:Are Keywords OK as typeset?

1. Introduction

The development of low-power, autonomous, mobile robotics is a growing concern in the planning of future space exploration due to the long latency of communication with Earth and the decreasing availability of solar energy as we consider missions at Mars and beyond. Although propulsion will typically have the largest power budget, many missions will require continuous sensory capabilities that can add up to significant power consumption. Organisms here on Earth have evolved successfully under similar pressures of power and speed and can provide inspiration in the design of future devices whether we look at them mechanically, neurally, or behaviorally. In this paper we consider the example of sound localization in bats and the implementation of the beginning steps of neural processing in VLSI.

Sound Localization and Bats

While the bat's carefully controlled ultrasonic sonar pulse is not a feature common to most mammals, the neural mechanisms for determining the direction of the reflected sound are thought to be very similar. The azimuthal direction of a sound source can be extracted using two basic concepts: measuring the difference in arrival times of a sound at spatially-separated receivers and measuring the difference in intensity level at spatially-separated receivers at different frequencies. Practical constraints often limit the usefulness of either or both of these cues (e.g. small receiver separation or insufficient attenuation between receivers).

The echolocation systems of bats primarily utilize ultrasonic frequencies to detect objects in three-dimensions. The use of high frequencies is critical for producing echoes from tiny objects (insects), for limiting the range of echoes and for producing a direction-dependent intensity level difference between the two ears. Given the head size of most echolocating bats ($\sim$1 cm), the maximum interaural time difference cue ($\sim$50 μsec) is very small and arguably unusable. Additionally, at ultrasonic frequencies, the hair cells of the cochlea in bats are unable to transmit phase information about the waveform, leaving only envelope modulation to provide interaural time difference cues. The interaural level difference (ILD) cue is, therefore,

242   *Horiuchi and Hynna*

the dominant acoustic feature for azimuthal echo localization in bats.

Sound waves reaching the head of the animal attenuate as they wrap around towards the back of the head. Wavelengths smaller than the size of the head are significantly attenuated and wavelengths larger than the size of the head are not. This frequency-dependence interacting with various "shadow producing" features of the head, snout, body, and pinnae produce a complex direction-dependent transfer function that can be used to determine the direction of sound by measuring the attenuation factor at many different frequencies. By subtracting the logarithm of the signal amplitudes (i.e. intensity level) at each ear, we produce a measure of this attenuation factor.

2. The Interaural Level Difference Pathway in the Bat

The sensitivity to ILD found in many cells of the midbrain of the bat are of the excitatory-inhibitory (EI) type. EI-type cells are distinguished by their excitatory response to sounds in one ear and inhibitory response to sounds in the opposite ear. Pollak and Park (1995), using known anatomy and physiology, have identified five different circuits that can produce this behavior in the inferior colliculus. In this paper, we discuss the neural circuit and test of one of these circuits in a hardware implementation. In particular, we mimic the double target experiments of Pollak and Burger (2000) to expose the effects of dynamic synapses on such circuits.

The earliest binaural comparison of sounds from the two ears occurs at the lateral superior olive, or LSO. The LSO receives excitation from the ipsilateral ear (same side of the body) and inhibition from the contralateral ear (opposite side of the body), performing a (left-right) operation. This excitatory-inhibitory arrangement is referred to as EI and creates a spike threshold at a particular interaural intensity level difference. When the sound is ipsilateral to a particular LSO cell, the received excitation will, in general, be stronger than the inhibition coming from the contralateral ear, resulting in a spike. If, however, the sound is contralateral to the LSO cell, then the resulting inhibition can prevent the cell from firing. (See Fig. 1) Recall that echoes are short in duration (0.5 ms to 3 ms) and will generally elicit only one or two spikes when maximally driven.

Using an artificial sonar system, we have demonstrated the functionality of the LSO circuit. Our working circuit architecture is shown in Fig. 2. We modeled
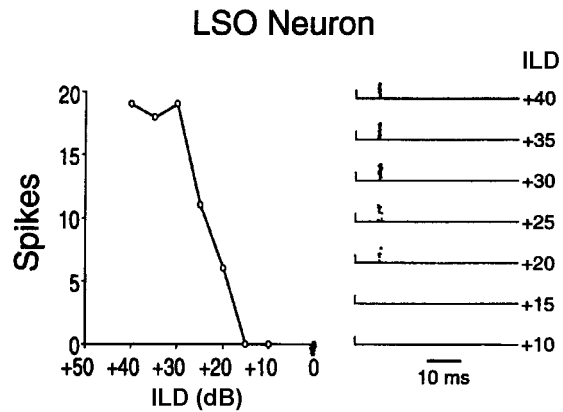

Figure 1. Left: spike response (per 20 trials) of an LSO cell as a function of interaural level difference to a brief stimulus. Note that the interaural level difference corresponds to the direction-dependent attenuation factor. The spike rasters on the right of the figure show the same data but highlight the fact that cells typically fire one or no spikes per trial. (Mexican free-tailed bat) Modified from Park (1998).
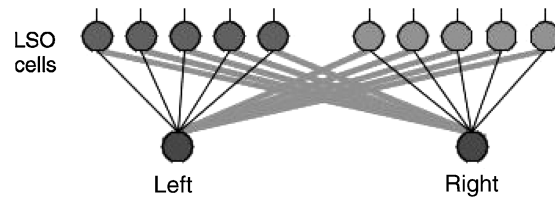

Figure 2. The LSO cells receive ipsilateral excitatory connections (dark lines) and contralateral inhibitory connections (gray lines). For example, the five cells on the left respond to activity coming from the left side and not the right side. The relative strengths of excitation vs. inhibition determine the cells' direction threshold for firing.

the connectivity of five LSO cells, receiving excitation from the ipsilateral side and inhibition from the contralateral side. The multiple units shown here represent cells with different weighted connections to produce different inhibitory thresholds. While we will show the basic functionality, we also demonstrate the effect of time-dependent inhibition on multiple sonar targets. In particular we will show "masking" behavior in the LSO, the suppressive effect of a preceding stimulus on subsequent stimuli.

3. The Narrow-Band Sonar Front-End

The experimental front-end of the current system is a narrow-band (39–41 kHz) ultrasonic speaker and microphone pair. The microphones are oriented at nearly 90 deg from each other to produce sensitivity patterns
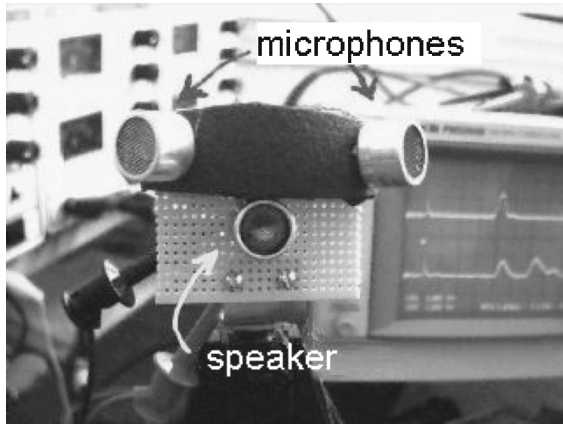
Figure 3. Photo of the narrowband sonar front-end. The oscillo-scope in the background on the right shows an example echo return from the 3 legs of a tripod. The sonar head is approximately 2.5″ tall and is mounted on a model airplane servo for tracking experiments (not described).

such that the relative amplitude of signal received in the two microphones is consistent with a particular direc-tion (see Fig. 3). While this sonar head is larger than a typical bat head and we are able to detect the interaural

time difference, this information is not used because it is not available to the bat. This narrow-band sonar front-end will soon be replaced by a tiny (1 cm³) broad-band system using miniature MEMS microphones and speaker.

A burst of 40 kHz sound (∼0.5 msec duration) is transmitted from the speaker periodically (∼20 Hz) to generate echoes from targets placed at various points in the room. Each microphone signal is amplified and broadly bandpass filtered before being half-wave rec-tified. The microphones (Murata) are very narrowly tuned (39–41 kHz), which provides most of the fre-quency tuning.

The half-wave rectification drives a peak-detector circuit with a decay (∼1 ms time-constant) to simu-late the dynamics of the hair cells in the cochlea. We will often refer to this as the envelope. At ultrasonic frequencies, this results in a signal proportional to the echo amplitude with no carrier phase information. This envelope (an example of which is visible in the photo of the sonar head and in Fig. 4) is used to drive a spike-generation circuit that simulates the spikes seen on the auditory nerve.
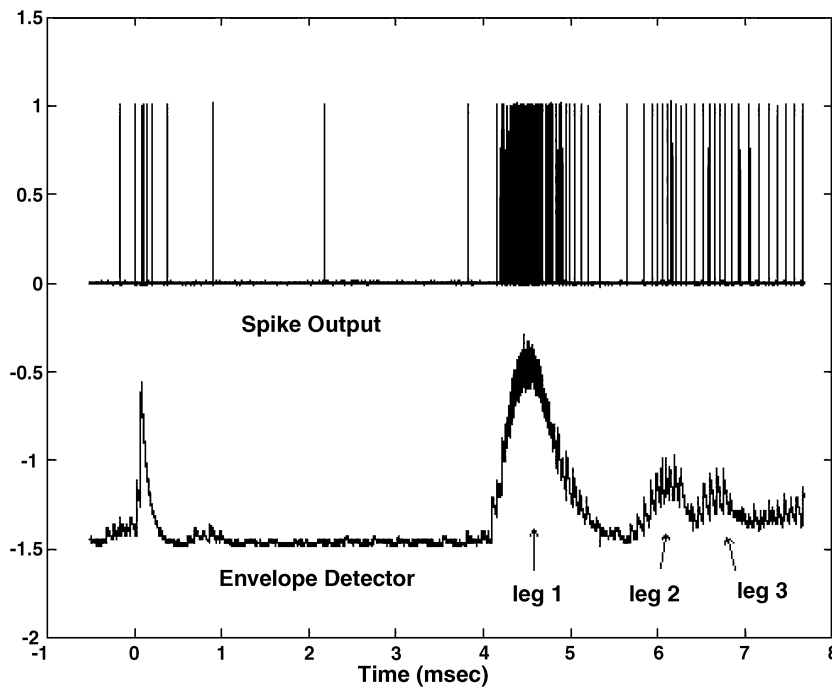

Figure 4. Bottom: Echo returns off of the three legs of a camera tripod. The sonar pulse occurs at time = 0. The filtered echo waveform is half-wave rectified and peak detected (with a decay). This output (lower trace) is then used to generate spikes (upper trace) at a rate proportional to the envelope height. The single fast spiking neuron is intended to represent the firing of a pool of neurons with different firing thresholds. Note: the first pulse in the envelope and the corresponding spikes are due to the incomplete sensory suppression of the outgoing sonar pulse. These initial spikes will be suppressed prior to driving the echolocation system.
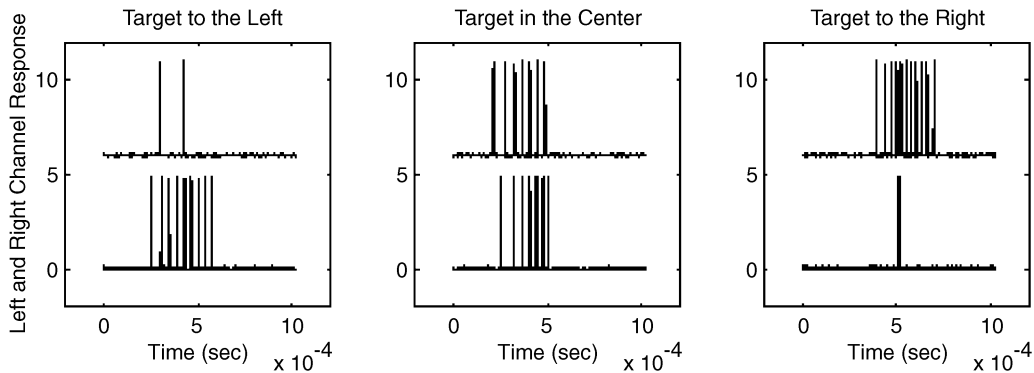
244   *Horiuchi and Hynna*

Figure 5. Left and right ear spike outputs for a vertical cylinder presented to the left, center, and right of the sonar's midline. Only the time interval of the echo is shown to retain spike detail.

The envelope signal is then used to drive a spike generating circuit that produces voltage pulses at a rate proportional to the envelope amplitude. This produces a very high spike rate that is intended to reproduce the *population* response of a pool of cells with different thresholds. Figure 4 shows an example of this using the echo returns off of the legs of a camera tripod.

Figure 5 shows the direction-dependent attenuation produced by the sonar head for echoes reflecting off of a vertical cylinder (2" diameter) placed approximately 1 meter away. These are representative signals that will be used to drive the neurons in the neuron chip described in the next section.

4.   Circuit Models and Network Connectivity

The neurons are represented by a single neuromorphic transceiver chip called the MAX chip. The MAX chip contains a one-dimensional array of 36 silicon neurons. The array is divided into two sections, with each section sharing a common set of circuit biases. All neural connectivity and activity is handled off-chip using the *Address Event Representation* system for communication. The MAX chip was fabricated using 0.8 micron HP technology, with a total area of 3.2 mm^2.

Silicon Neuron

Our silicon neuron model (Fig. 6) can be divided into three distinct circuits: a spike generator, Ca^{++} dependent K$^+$ channels, and two synapses (not shown). The spike generator is a modified version of the axon hillock circuit described by Mead (1989). Input current, I_{in}, from the synapses integrates onto the membrane

node capacitance $C_m (= C_1 + C_2)$. When the membrane voltage V_m reaches threshold (controlled by V_{pu}), the cascaded inverters transition, causing V_{spk} to switch from Gnd (resting state) to V_{dd}. Due to the coupling capacitor C_2, the transition of V_{spk} increases V_m above the threshold by $V_{dd} C_2 / (C_1 + C_2)$. With V_{spk} now high, the reset current through the two series transistors begins to remove charge from V_m, at a rate controlled by the gate voltage V_{mpd}. Once V_m decreases below threshold, the cascaded inverters transition back to resting state and V_{spk} becomes low, shutting off the reset current and decreasing V_m from threshold by a similar amount $V_{dd}(C_2/C_1 + C_2)$. The neuron is now again ready to
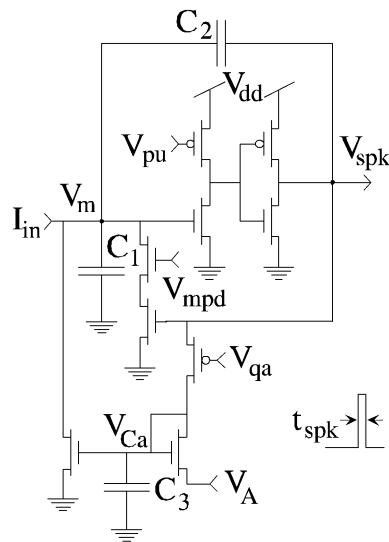

Figure 6. Schematic of silicon neuron. Not shown here are the synapses (see Fig. 8). See text for details.

integrate the input. The width of the output pulse (t_{spk}) is controlled by the reset current I_{mpd}. For $I_{mpd} \gg I_{in}$, the pulse width is negligible to the interspike interval. In this state, the spike generator behaves like an integrate and fire neuron, with the firing rate dependant on the input current (see Mead, 1989).

The dynamics of Ca^{++} dependent K^+ channels are modeled using a current-mirror integrator circuit. The voltage, V_{Ca}, at the integrator node is the circuit analog of intracellular calcium concentration. Each time the neuron spikes, a small quanta of charge—controlled by V_{qa}—is dumped onto the capacitor C_3. The diode-connected transistor leaks the accumulated charge away at a rate dependent on the source voltage V_A. If charge enters the integrator node faster than it can decay, then the voltage at that node, V_{Ca}, rises. Equilibrium is reached when amount of charge removed from C_3 during the interspike interval matches the amount added by each spike. A single transistor, whose drain is attached to V_m and whose gate voltage is V_{Ca}, represents the Ca^{++} dependent K^+ channels. This negative-feedback effect, also found in biology, is called spike-rate adaptation.

Two synapses—one excitatory the other inhibitory—provide input into the silicon neuron. These synapses have the same form of the Ca^{++} dependent K^+ channels described above, namely a current-mirror integrator (Fig. 7). An incoming presynaptic spike (Spk_{In}) dumps a quanta of charge onto the integrator node through a transistor. The gate voltage of the transistor, V_w, controls the amount of charge that is added with each incoming pulse and therefore acts as the synaptic strength. The charge on the integrator node decays via the diode connected transistor at a rate depen-

Figure 7. Inhibitory synapse. An input pulse dumps a quanta of charge (controlled by V_w) onto a current-mirror integrator. The decay of the charge depends on the source voltage V_A. The output is connected to V_m, and thus inhibits the cell. The excitatory synapse is identical except the output passes through a p-type current mirror to reverse the current.

dent on the source voltage V_S. The output current I_{out} is controlled by the node voltage and is either inhibitory (directly connected to V_m of the neuron) or excitatory (passing through a P current mirror to reverse the current). It is this finite duration (set to about 1–2 msec) of inhibitory current that we utilize in our circuit operation described in Section 5. For more information on the dynamics of the synapses, see Indiveri (2000).

Address Event Representation and Neural Connectivity

Input and output activity for the chip is conveyed over a digital bus using *address-event representation* (Boahen, 1998). Each silicon neuron is given a unique identifying address so when it fires, an encoder transmits its address onto the output address-event bus. The time of the spike is implicitly encoded in the occurrence of the address on the bus.

On the input side, each synapse, rather than each neuron, is given a unique address. A decoder reads the new address-event and sends a pulse to the appropriate synaptic location. Once completed, the receiver circuitry acknowledges the now expired data on the bus and waits for a new event.

To increase the fan-out of the neuron (as well as to provide flexibility in the circuits), a lookup table (LUT) is placed onto the digital bus between the two chips. This LUT acts as a translator for new events from chip 1 en route to chip 2, essentially storing the connectivity pattern between the layers. Going back to the previous example, when the LUT receives address 2, it can send out addresses {2, 4, 6} (a fan-out of 3) to chip 2, exciting the central neuron and inhibiting the neighbors.

We used a Microchip PIC16C55 microcontroller (PIC) as the lookup table for each neural array. The PIC16C55 has a 200 ns instruction cycle, with 24 bytes of RAM and 768 bytes of program memory. An arbiter circuit provided the input to each PIC. The arbiter is a simple circuit that merges two address-event streams into a single bus, providing all the appropriate handshaking to communicate with the various streams. When there is a new event on one of the buses, the merge circuit allows the data through to the output bus, usually adding an extra bit to the address indicating the source. Should an event arrive at the other input, it waits until the first has been transmitted before it can pass through. The arbiter is a simplified version of the arbiter circuitry found within the MAX chip (see Boahen, 1998).
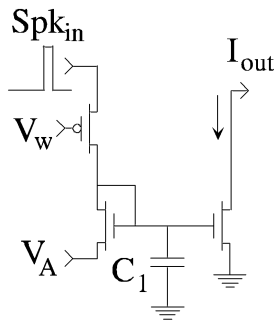
246   *Horiuchi and Hynna*

For the LSO circuit, the activity from each ear was translated into excitatory inputs for ipsilateral neurons and inhibitory inputs for contralateral neurons (as in Fig. 2). Since the MAX chips provide the same synaptic weight for all neurons, an attempt to create variable weights was performed in the PIC using pseudo-probabilistic synapses. Given the complexity of the PIC code (for these synapses), the fan-out of each connection and the firing rate of the ear neurons, we were limited to using only 10 neurons (5 for each side) in order stay within bandwidth limits of the bus. Only the excitatory synapses (5 for each ear) were made probabilistic, each allowing only a certain percentage of the activity through (100, 90, 80, 67 and 50 percent).

5. Experimental Results

In Fig. 5, we showed the responses of various units to echoes returning to the sonar system from different directions. These spiking inputs were used to drive the LSO circuitry defined by the address-event infrastructure. The EI circuit arrangement for the LSO units produces the expected threshold in ILD (Fig. 8), however, the expected variation in threshold position due to the differential weightings is extremely weak.

An interesting aspect of the circuit is revealed when we observe the response of the LSO cells to multiple targets. Figure 9 shows two experiments where a distant target on the opposite side of the sonar's midline is
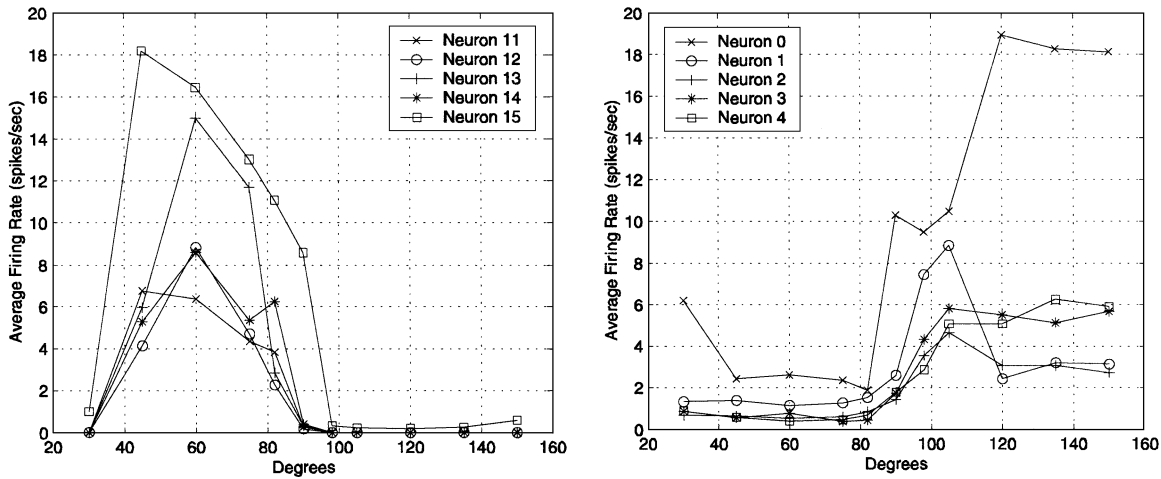

Figure 8. Responses of the VLSI LSO neurons to echoes from different directions. Left panel: EI neurons with excitation coming from the left and inhibition from the right. These cells fire for echoes coming from the left side. Note that the drop in response for echoes from the far left is a result of the left microphone's drop in sensitivity. Right panel: EI neurons with excitation coming from the right and inhibition from the left. This cells fires for echoes coming from the right side.
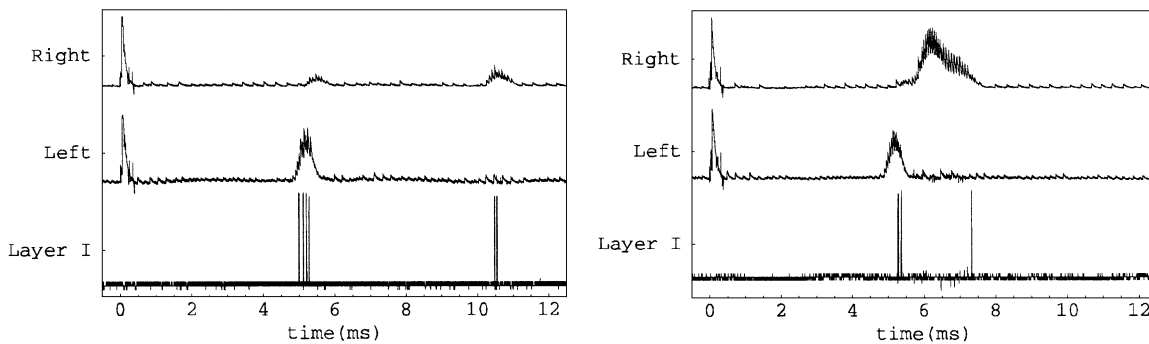

Figure 9. Masking behavior of inhibition. Two targets on opposite sides of the midline are moved close to each other. The top trace is the envelope output of the right channel, the middle trace is the envelope of the left channel and the bottom trace is a digital output that fires whenever any neuron on the chip fires. As the target on the right approaches the first target (even though the echo signal is much larger) the neuronal response is significantly lower.

moved closer until the spiking response that it normally elicits (when presented alone) is diminished. This is a consequence of the inhibition produced on the LSO cells by the closer target. This masking effect produced a shadow behind the first target in spite of a clear, strong echo visible in the envelope trace. It should be noted that not all LSO neurons would be disabled by this inhibition; rather a limited population would provide a differential response for this type of target configuration.

6. Discussion

In this paper, we describe our ongoing efforts to construct a large-scale model of the neural processing in the echolocation system of the bat. To date, we have focused on the processing of a single frequency band, however, our expansion into broadband processing has already begun. Using a spike-based representation for echo amplitude, we have demonstrated the EI processing in an integrate-and-fire type of neuron to produce units that only respond to targets from a limited range of directions, consistent with the types of responses seen in LSO cells of the bat. We investigate the dynamic effects of inhibition (1–2 ms) on the response of the various cells and suggest a possible utility for these dynamics.

While we have not duplicated all of the known elements of the bat midbrain pathway in our silicon model, we have been able to demonstrate similar phenomena that are only present in models that preserve the dynamics of the original neural circuit. This type of modeling hints that the EI cells in the bat IC has much more diversity in their spatial response fields that is not well described by experiments using only single targets and static models of computation.

Our efforts in neuromorphic circuit implementation are intended to go beyond simply mimicking neurobiology to developing fast processing strategies for transforming temporal signals into spatial patterns appropriate for effective survival behavior.

References

Boahen, K.A. 1998. Communicating neuronal ensembles between neuromorphic chips. In *Neuromorphic Systems Engineering*, T.S. Lande (Ed.), Kluwer Academic Publishers: Boston, MA, Ch. 11, pp. 229–261.

Indiveri, G. 2000. Modeling selective attention using a neuromorphic analog VLSI device. *Neural Computation*, 12(12):2857–2880.

Mead, C. 1989. *Analog VLSI and Neural Systems*, Addison-Wesley: Menlo Park, pp. 193–204.

Park, T. 1998. IID Sensitivity differs between two principal centers in the interaural intensity difference pathway: The LSO and the IC. *J. Neurophysiol.*, 79(5):2416–2431.

Pollak, G. and Burger, R.M. 2000. Reversible inactivation of the dorsal nucleus of the lateral lemniscus reveals its role for processing multiple sound sources in the inferior colliculus. Poster Presentation at the Symposium on Synaptic Function in Hearing and Balance, Johns Hopkins University, June 8, 2000.

Pollak, G. and Park, T. 1995. Inferior colliculus. In *Hearing by Bats*, A. Popper and R.R. Fay (Eds.), Springer-Verlag: New York, pp. 296–367.