Communicated by John Wyatt

A Silicon Model Of Auditory Localization

John Lazzaro
Carver A. Mead
Department of Computer Science, California Institute of Technology, MS 256-80, Pasadena, CA 91125, USA

The barn owl accurately localizes sounds in the azimuthal plane, using interaural time difference as a cue. The time-coding pathway in the owl's brainstem encodes a neural map of azimuth, by processing interaural timing information. We have built a silicon model of the time-coding pathway of the owl. The integrated circuit models the structure as well as the function of the pathway; most subcircuits in the chip have an anatomical correlate. The chip computes all outputs in real time, using analog, continuous-time processing.

1 Introduction

The principles of organization of neural systems arose from the combination of the performance requirements for survival and the physics of neural elements. From this perspective, the extraction of time-domain information from auditory data is a challenging computation; the system must detect changes in the data which occur in tens of microseconds, using neurons which can fire only once per several milliseconds. Neural approaches to this problem succeed by closely coupling algorithms and implementation. We are building silicon models of the auditory localization system of the barn owl, to explore the general computational principles of time-domain processing in neural systems.

The barn owl (*Tyto alba*) uses hearing to locate and catch small rodents in total darkness. The owl localizes the rustles of the prey to within one to two degrees in azimuth and elevation (Knudsen et al. 1979). The owl uses different binaural cues to determine azimuth and elevation. The elevational cue for the owl is interaural intensity difference (IID). This cue is a result of a vertical asymmetry in the placement of the owl's ear openings, as well as a slight asymmetry in the left and right halves of the owl's facial ruff (Knudsen and Konishi 1979). The azimuthal cue is interaural time difference (ITD). The ITDs are in the microsecond range, and vary as a function of azimuthal angle of the sound source (Moiseff and Konishi 1981). The external nucleus of the owl's inferior colliculus (ICx) contains the neural substrate of sound localization, a map of auditory space (Knudsen and Konishi 1978). Neurons in the ICx respond

 © 1989 Massachusetts Institute of Technology

maximally to stimuli located in a small area in space, corresponding to a specific combination of IID and ITD.

There are several stages of neural processing between the cochlea and the computed map of space in the ICx. Each primary auditory fiber initially divides into two distinct pathways. One pathway processes intensity information, encoding elevation cues, whereas the other pathway processes timing information, encoding azimuthal cues. The time-coding and intensity-coding pathways recombine in the ICx, producing a complete map of space (Takahashi and Konishi 1988).

2 A Silicon Model of the Time-Coding Pathway

We have built an integrated circuit that models the time-coding pathway of the barn owl, using analog, continuous-time processing. Figure 1 shows the floorplan of the chip. The chip receives two inputs, corresponding to the sound pressure at each ear of the owl. Each input connects to a silicon model of the cochlea, the organ that converts the sound energy present at the eardrum into the first neural representation of the auditory system. In the cochlea, sound is coupled into a traveling wave structure, the basilar membrane, which converts time-domain information into spatially-encoded information, by spreading out signals in space according to their time scale (or frequency). The cochlea circuit is a one-dimensional physical model of this traveling wave structure; in engineering terms, the model is a cascade of second-order sections, with exponentially scaled time constants (Lyon and Mead 1988).

In the owl, inner hair cells contact the basilar membrane at discrete intervals, converting basilar-membrane movement into a graded, half-wave rectified electrical signal. Spiral ganglion neurons connect to each inner hair cell, producing action potentials in response to inner-hair-cell electrical activity. The temporal pattern of action potentials encodes the shape of the sound waveform at each basilar-membrane position. Spiral ganglion neurons also reflect the properties of the cochlea; a spiral ganglion neuron is most sensitive to tones of a specific frequency, the neuron's characteristic frequency.

In our chip, inner hair cell circuits connect to taps at discrete intervals along the basilar-membrane model. These circuits compute signal processing operations (half-wave rectification and nonlinear amplitude compression) that occur during inner hair cell transduction. Each inner hair cell circuit connects to a spiral ganglion neuron circuit. This integrate-to-threshold neuron circuit converts the analog output of the inner-hair-cell model into fixed-width, fixed-height pulses. Timing information is preserved by greatly increasing the probability of firing near the zero crossings of the derivative of the neuron's input.

In the owl, the spiral ganglion neurons project to the nucleus magnocellularis (NM), the first nucleus of the time-coding pathway. The NM

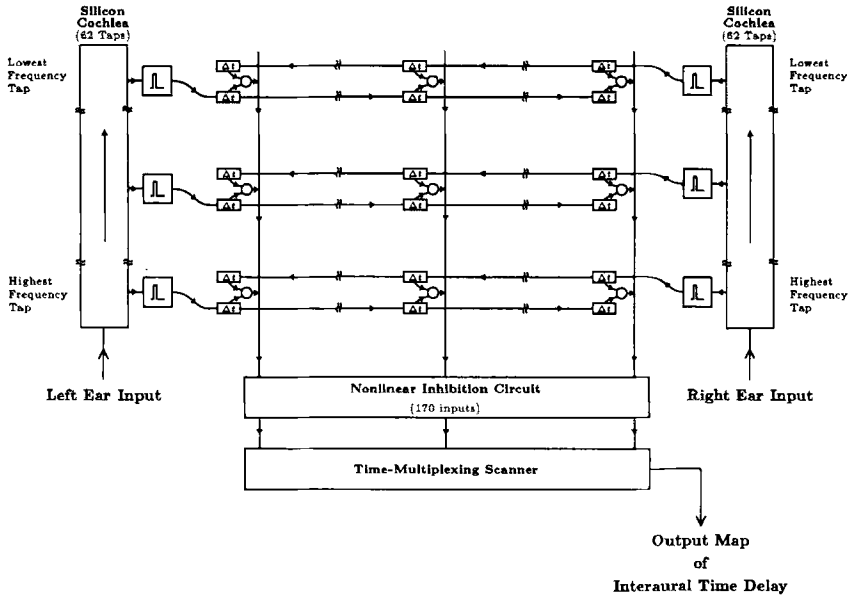

Figure 1: Floorplan of the silicon model of the time-coding pathway of the owl. Sounds for the left ear and right ear enter the respective silicon cochleas at the lower left and lower right of the figure. Inner hair cell circuits tap each silicon cochlea at 62 equally-spaced locations; each inner hair cell circuit connects directly to a spiral ganglion neuron circuit. The square box marked with a pulse represents both the inner hair cell circuit and spiral ganglion neuron circuit. Each spiral ganglion neuron circuit generates action potentials; these signals travel down silicon axons, which propagate from left to right for spiral ganglion neuron circuits from the left cochlea, and from right to left for spiral ganglion circuits from the right cochlea. The rows of small rectangular boxes, marked with the symbol Δt, represent the silicon axons. 170 NL neuron circuits, represented by small circles, lie between each pair of antiparallel silicon axons. Each NL neuron circuit connects directly to both axons, and responds maximally when action potentials present in both axons reach that particular neuron at the same time. In this way, ITDs map into a neural place code. Each vertical wire which spans the array combines the response of all NL neuron circuits which correspond to a specific ITD. These 170 vertical wires form a temporally smoothed map of ITD, which responds to a wide range of input sound frequencies. The nonlinear inhibition circuit near the bottom of the figure increases the selectivity of this map. The time-multiplexing scanner transforms this map into a signal suitable for display on an oscilloscope.

acts as a specialized relay station; neurons in the NM preserve timing information, and project bilaterally to the nucleus laminaris (NL), the first nucleus in the time-coding pathway that receives inputs from both ears. For simplicity, our chip does not model the NM; each spiral ganglion neuron circuit directly connects to a silicon NL.

Neurons in the NL are most sensitive to binaural sounds with a specific ITD. In 1948, Jeffress proposed a model to explain the encoding of ITD in neural circuits (Jeffress 1948). In the Jeffress model applied to the owl, axons from the ipsilateral and contralateral NM, with similar characteristic frequencies, enter the NL from opposite surfaces. The axons travel antiparallel, and action potentials counterpropagate across the NL; the axons act as neural delay lines. NL neurons are adjacent to both axons. Each NL neuron receives synaptic connections from both axons, and fires maximally when action potentials present in both axons reach that particular neuron at the same time. In this way, ITD is mapped into a neural place coding; the ITD that maximally excites an NL neuron depends on the position of the neuron in the NL. Anatomical and physiological evidence in the barn owl supports this theory (Carr and Konishi 1988).

The chip models the anatomy of the NL directly (Fig. 1). Two silicon cochleas lie at opposite ends of the chip; spiral ganglion neuron circuits from each cochlea, with similar characteristic frequencies, project to separate axon circuits, which travel antiparallel across the chip. The axon circuit is a discrete neural delay line; for each action potential at the axon's input, a fixed-width, fixed-height pulse travels through the axon, section by section, at a controllable velocity (Mead 1989). NL neuron circuits lie between each pair of antiparallel axons at every discrete section, and connect directly to both axons. Simultaneous action potentials at both inputs excite the NL neuron circuit; if only one input is active, the neuron generates no output. For each pair of antiparallel axons, there is a row of 170 NL neuron circuits across the chip. These neurons form a place encoding of ITD.

Our silicon NL differs from the owl's NL in several ways. The silicon NL neurons are perfect coincidence detectors; in the owl, NL neurons also respond, with reduced intensity, to monaural input. In the owl, many axons from each side converge on an NL neuron; in the chip, only two silicon axons converge on each silicon NL neuron. Finally, the brainstem of the owl contains two NLs, symmetric about the midline; each NL primarily encodes one half of the azimuthal plane. For simplicity, our integrated circuit has only one copy of the NL, which encodes all azimuthal angles.

In the owl, the NL projects to a subdivision of the central nucleus of the inferior colliculus (ICc), which in turn projects to the ICx. The ICx integrates information from the time-coding pathway and from the amplitude-coding pathway to produce a complete map of auditory space. The final output of our integrated circuit models the responses of ICx

neurons to ITDs. In response to ITDs, ICx neurons act differently from NL neurons. Experiments suggest mechanisms for these differences; our integrated circuit implements several of these mechanisms to produce a neural map of ITD.

Neurons in the NL and ICc respond to all ITDs that result in the same interaural phase difference (IPD) of the neuron's characteristic frequency; neurons in the ICx respond to only the one true ITD. This behavior suggests that ICx neurons combine information from many frequency channels in the ICc, to disambiguate ITDs from IPDs; indeed, neurons in the NL and ICc reflect the frequency characteristics of spiral ganglion neurons, whereas ICx neurons respond equally to a wide range of frequencies.

In our chip, all NL neuron outputs corresponding to a particular ITD are summed to produce a single output value. NL neuron outputs are current pulses; a single wire acts as a dendritic tree to perform the summation. In this way, a two-dimensional matrix of NL neurons reduces to a single vector; this vector is a map of ITD, for all frequencies. In the owl, inhibitory circuits between neurons tuned to the same ITD may also be present, before summation across frequency channels. Our model does not include these circuits.

Neurons in the ICc are more selective to ITDs than are neurons in the NL; in turn, ICx neurons are more selective to ITDs than are ICc neurons, for low frequency sounds. At least two separate mechanisms join to increase selectivity. The selectivity of ICc and ICx neurons increases with the duration of a sound, for sounds lasting less than 5 milliseconds, implying that the ICc and perhaps the ICx may use temporal integration to increase selectivity (Wagner and Konishi, in preparation). Our chip temporally integrates the vector that represents ITD; the time constant of integration is adjustable.

Nonlinear inhibitory connections between neurons tuned to different ITDs in the ICc and ICx also increase sensitivity to ITDs; application of an inhibitory blocker to either the ICc or ICx decreases sensitivity to ITD (Fujita and Konishi, in preparation). In our chip, a global shunting inhibition circuit (Lazzaro et al. 1988) processes the temporally integrated vector that represents ITD. This nonlinear circuit performs a winner-take-all function, producing a more selective map of ITD. The chip time-multiplexes this output map on a single wire for display on an oscilloscope.

3 Comparison of Responses

We presented periodic click stimuli to the chip (Fig. 2a), and recorded the final output of the chip, a map of ITD. Three signal-processing operations, computed in the ICx and ICc of the owl, improve the original encoding of ITDs in the NL: temporal integration, integration of infor-

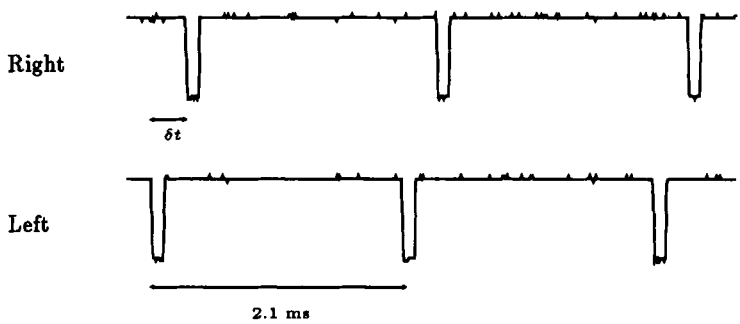

Figure 2: Input stimulus for the chip. Both left and right ears receive a periodic click waveform, at a frequency of 475 Hz. The time delay between the two signals, notated as δt, is variable.

mation over many frequency channels, and inhibition among neurons tuned to different ITDs. In our chip, we can disable the inhibition and temporal-integration operations, and observe the unprocessed map of ITD (Fig. 2b). By combining the outputs of 62 rows of NL neurons, each tuned to a separate frequency region, the maps in figure 2b correctly encode ITD, despite random variations in axonal velocity and cochlear delay. Figure 3 shows this variation in velocity of axonal propagation, due to circuit element imperfections. Figure 2c shows maps of ITD taken with inhibition and temporal integration operations enabled. Most maps show a single peak, with little activity at other positions.

Figure 4a is an alternative representation of the map of ITD computed by the chip. We recorded the map position of the neuron with maximum signal energy, for different ITDs. Carr and Konishi (1988) performed a similar experiment in the owl's NL (Fig. 4b), mapping the time delay of an axon innervating the NL, as a function of position in the NL. The linear properties of our chip map are the same as those of the owl map.

4 Conclusions

Traditionally, scientists have considered analog integrated circuits and neural systems to be two disjoint disciplines. The two media are different in detail, but the physics of computation in silicon technology and in

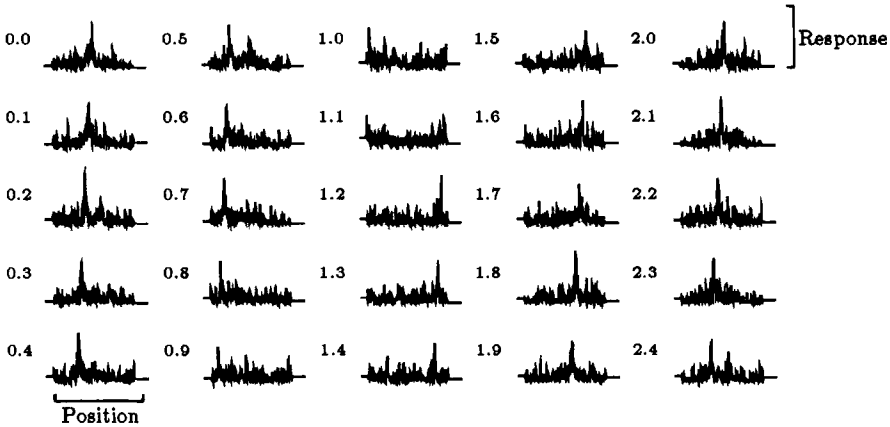

Figure 3: Map of ITD, taken from the chip. The nonlinear inhibition and temporal smoothing operations were turned off, showing the unprocessed map of ITD. The vertical axis of each map corresponds to neural activity level, whereas the horizontal axis of each map corresponds to linear position within the map. The stimulus for each plot is the periodic click waveforms of Figure 2a; δt is shown in the upper left corner of each plot, measured in milliseconds. Each map is an average of several maps recorded at 100 millisecond intervals; averaging is necessary to capture a representation of the quickly changing, temporally unsmoothed response. The encoding of ITD is present in the maps, but false correlations add unwanted noise to the desired signal. Since we are using a periodic stimulus, large time delays are interpreted as negative delays, and the map response wraps from one side to the other at an ITD of 1.2 milliseconds.

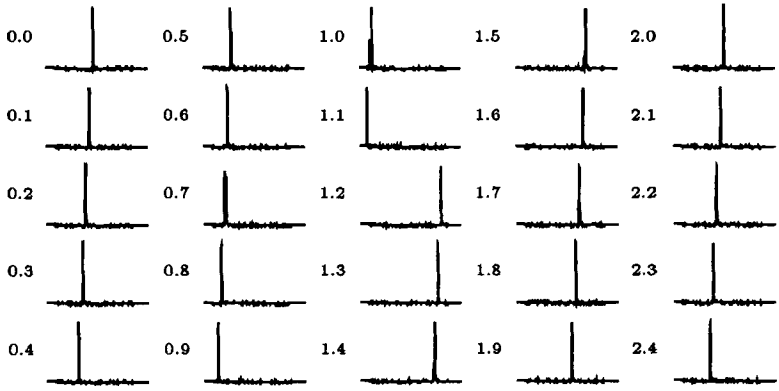

Figure 4: Map of ITD, taken from the chip. The nonlinear inhibition and temporal smoothing operations were turned on, showing the final output map of ITD. Format is identical to Figure 2b. Off-chip averaging was not used, since the chip temporally smooths the data. Most maps show a single peak, with little activity at other positions, due to nonlinear inhibition. The maps do not reflect the periodicity of the individual frequency components of the sound stimulus; additional experiments with a noise stimulus confirm the phase-disambiguation property of the chip.

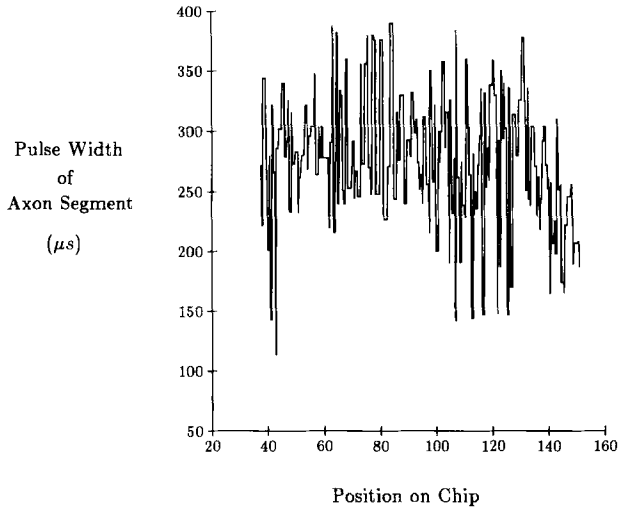

Figure 5: Variation in the pulse width of a silicon axon, over about 100 axonal sections. Axons were set to fire at a slower velocity than in the owl model, for more accurate measurement. In this circuit, a variation in axon pulse width indicates a variation in the velocity of axonal propagation; this variation is a potential source of localization error.

neural technology are remarkably similar. Both media offer a rich palette of primitives in which to build a structure; both pack a large number of imperfect computational elements into a small space; both are ultimately limited not by the density of devices, but by the density of interconnect. Modeling neural systems directly in a physical medium subjects the researcher to many of the same pressures faced by the nervous system over the course of evolutionary time.

We have built a 220,000 transistor chip that models, to a first approximation, a small but significant part of a spectacular neural system. In doing so we have faced many design problems solved by the nervous system. This experience has forced us to a high level of concreteness in specifying this demanding computation. This chip represents only the first few stages of auditory processing, and thus is only a first step in auditory modeling. Each individual circuit in the chip is only a first approximation to its physiological counterpart. In addition, there are other auditory pathways to explore: the intensity-coding localization pathway, the elevation localization pathway in mammals, and, most formidably, the sound-understanding structures that receive input from these pathways.

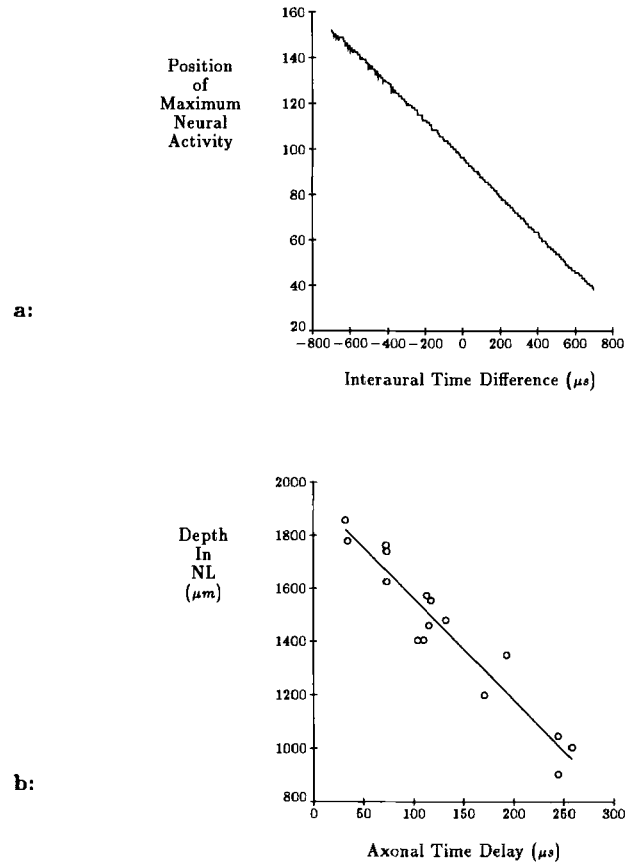

Figure 6: (a) Chip data showing the linear relationship of silicon NL neuron position and ITD. For each ITD presented to the chip, the output map position with the maximal response is plotted. The linearity shows that silicon axons have a uniform mean time delay per section. (b) Recordings of the NM axons innervating the NL in the barn owl (Carr and Konishi 1988). The figure shows the mean time delays of contralateral fibers recorded at different depths during one penetration through the 7 KHz region.

Acknowledgments

We thank M. Konishi and his entire research group, in particular S. Volman, I. Fujita, and L. Proctor, as well as D. Lyon, M. Mahowald, T. Delbruck, L. Dupré, J. Tanaka, and D. Gillespie, for critically reading and correcting the manuscript, and for consultation throughout the project.

We thank Hewlett-Packard for computing support, and DARPA and MOSIS for chip fabrication. This work was sponsored by the Office of Naval Research and the System Development Foundation.

References

Carr, C.E. and M. Konishi. 1988. Axonal Delay Lines for Time Measurement in the Owl's Brainstem. *Proc. Nat. Acad. Sci.* **85**, 8311–8315.

Fujita, I. and M. Konishi. In preparation.

Jeffress, L.A. 1948. A Place Theory of Sound Localization. *J. Comp. Physiol. Pyschol.* **41**, 35–39.

Knudsen, E.I., G.G. Blasdel, and M. Konishi. 1979. Sound Localization by the Barn Owl Measured with the Search Coil Technique. *J. Comp. Physiol.* **133**, 1–11.

Knudsen, E.I. and M. Konishi. 1979. Mechanisms of Sound Localization in the Barn Owl (Tyto alba). *J. Comp. Physiol.* **133**, 13–21.

———. 1978. A Neural Map of Auditory Space in the Owl. *Science* **200**, 795–797.

Lazzaro, J.P., S. Ryckebusch, M.A. Mahowald, and C.A. Mead. 1988. Winner-Take-All Networks of O(n) Complexity. *Proc. IEEE Conf. Neural Information Processing Systems*, Denver, CO.

Lyon, R.F. and C. Mead. 1988. An Analog Electronic Cochlea. *IEEE Trans. Acoust., Speech, Signal Processing* **36**, 1119–1134.

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

Moiseff, A. and M. Konishi. 1981. Neuronal and behavioral sensitivity to binaural time differences in the owl. *J. Neurosci.* **1**, 40–48.

Takahashi, T.T. and M. Konishi. 1988. Projections of the Nucleus Angularis and Nucleus Laminaris to the Lateral Lemniscal Nuclear Complex of the Barn Owl. *J. Compar. Neurol.* **274**, 221–238.

Wagner, H. and M. Konishi. In preparation.

Received 26 October; accepted 9 November 1988.