Chapter 6
Future Research

The projects described in this dissertation demonstrate the utility of analog VLSI technology as a modeling tool in auditory neuroscience. The dissertation does not, however, propose any new auditory theories. The auditory-localization and pitch-perception chips implement theories that were proposed by Jeffress (Jeffress, 1948) and Licklider (Licklider, 1951) several decades ago. The inner-hair-cell and spiral-ganglion-cell models presented in Chapter 2 are a codification of known properties of these cells.

This conservative approach to the dissertation was deliberate, and is in harmony with a well-known computer-science dictum: Do not let your research project depend on another research project. Through the projects described in this dissertation, my co-workers and I wanted to answer the question: Is analog VLSI technology an appropriate medium for auditory modeling? This dissertation answers this question affirmitively; if we had attempted, and failed, to invent new models of auditory function, we would have left our primary question unanswered. There is another question for the future, which this chapter explores: What are the interesting, yet tractable, auditory questions to ask using analog VLSI technology? We devote each section of the chapter to a different topic.

6.1 The Cochlea: Frequency-Specific Automatic Gain Control

As detailed in Chapter 2, the cochlea contains outer hair cells, which are active electromechanical elements. Outer hair cells reduce the damping of the passive basilar membrane, and allow weaker signals to be heard. Axons from higher brain centers innervate the outer hair cells; these centers might dynamically vary the local damping of the cochlea, providing frequency-specific automatic gain control (Kim, 1984).

The circuit model of auditory-nerve response, described in Chapter 2, includes outer-hair-cell function, but does not model dynamic variation of outer-hair-cell motility by higher brain centers. Adding a frequency-specific automatic-gain-control system to the silicon model of auditory-nerve response is a tractable project in analog VLSI technology. The control system should independently vary the damping of local regions of basilar-membrane circuitry, in response to sound input. The control circuit must maintain the stability of the basilar-membrane circuit; although human ears are occasionally oscillatory (ringing ears), chronic oscillations are classified as a symptom (tinnitus) of auditory diseases.

6.2 The Barn Owl: Learning Elevational Space

As described in Chapter 4, the barn owl uses interaural time differences to encode azimuthal position in auditory space, and interaural intensity differences to encode elevational position. The interaural intensity difference that corresponds to a particular spatial position is frequency dependent; computing a spatial map from interaural intensity difference requires knowledge about the spectral characteristics of the head of the animal.

A silicon model of elevation coding in the barn owl would require several levels of processing. A possible model would begin with two circuit models of auditory-nerve response (Chapter 2), followed by processing to enhance the intensity information in the auditory-nerve response; this enhancement processing would model the function of the nucleus angularis (Sullivan and Konishi, 1984). Binaural information would converge onto neurons that compute interaural intensity differences in different frequency regions, modeling neuronal responses in the nucleus ventralis lemnisci lateralis pons posterior (VLVp) (Manley *et al.*, 1988). To complete the computation, a learning network would correlate patterns of interaural intensity difference at different frequencies, to form a representation of elevation space. The output of this network would model the elevational response of the computed map of auditory space in the ICx, shown in Figure 4.1 in Chapter 4.

6.3 The Mustache Bat: Active Sonar Processing

The mustache bat, *Pteronotus parnelli rubiginosus*, uses active sonar for capturing flying insects, and for short-term navigation. The bat emits ultrasonic signals, and processes the resulting echos, to sense the environment. Suga and co-workers (Suga, 1984) have characterized a variety of neural maps in the bat auditory cortex, which systematically process different aspects of the sonar signal.

Analog VLSI technology is an appropriate tool for modeling these maps. The bat processes sound frequencies in excess of 100 kHz; analog VLSI circuits can process sounds at these frequencies. The peripheral auditory system of the bat is specialized for sonar processing; these specializations could be incorporated into the silicon model of auditory-nerve response (Chapter 2). Time-correlation techniques used in the pitch-perception (Chapter 5) and auditory-localization (Chapter 4) chips, in modified form, could form the basis for silicon models of certain maps in the bat auditory cortex.

6.4 Bats and Owls: Sensorimotor Feedback

When localizing a sound, the barn owl aligns its head to the source of the sound in a stereotyped manner (Knudsen and Konishi, 1979). Designing an electromechanical model of the alignment motion of the owl, using a motor-driven head with two microphones and the silicon model of auditory localization (Chapter 4), is a tractable project. Similar electromechanical systems modeling alignment to a visual stimuli have been built (DeWeerth and Mead, personal communication).

A more challenging project involves electromechanical modeling of bat insect capture. A flying moth-chaser, linking a sonar transceiver, a radio-controlled airplane, and silicon models of neural maps in the bat auditory cortex, would be a spectacular display of analog VLSI processing. Radio control would be used for takeoff and landing; once in flight, the airplane would seek out a flying object of a specified shape, size, and speed, and would follow the object at a safe distance. This project clearly would not be a trivial undertaking, but it would be possible to complete with our current technology.

6.5 Human Perception: Modeling Binaural Phenomena

Human binaural perception includes phenomena other than auditory localization of sounds. Binaural processing enhances our ability to detect signals in noise: for example, binaural hearing aids the comprehension of speech against a background of other speech (the cocktail-party effect). Several binaural signal-detection theories have been advanced (Jeffress, 1972; Durlach, 1972). Constructing analog VLSI models of these theories is a tractable project.

Binaural perception includes a wealth of interesting auditory illusions, which raise interesting psychological and physiological questions about audition

(Tobias, 1972; Deutsch, 1981). The systems described in Chapters 2, 4, and 5 are appropriate substrates for analog VLSI projects that explore these curious binaural phenomena.

6.6 Speech Recognition: The Final Frontier

Analog VLSI systems have two different missions in the speech-recognition community. The projects described in this dissertation offer real-time biological models of early auditory processing, for integration into existing speech-recognition architectures. In addition, analog VLSI technology is an appropriate implementation medium for neurally inspired speech-recognition algorithms. The projects described in this dissertation form a substrate for speech-oriented neural maps in analog VLSI technology.

Such maps could represent interesting characteristics of the speech signal; alternatively, these maps could represent the muscle movements necessary to produce the input speech signal (Liberman and Mattingly, 1989). Using the techniques of Chapter 4 and Chapter 5, these maps could process temporal as well as spectral aspects of the speech signal, as encoded by the silicon auditory-nerve model (Chapter 2). The pitch-perception chip (Chapter 5) computes the fundamental frequency of a speech signal in real time; the chip could serve as a component in a speech-recognition system that uses fundamental-frequency knowledge to create robust neural maps of interesting characteristics of the speech signal.

6.7 Conclusions

This chapter has presented interesting, tractable projects in auditory modeling, using analog VLSI technology. The previous chapters detail completed VLSI auditory models. Transforming project ideas into completed systems is a rewarding experience; we invite interested researchers to verify this claim.

Building analog VLSI chips has never been easier. The MOS Implementation Service (MOSIS), administered by the University of Southern California Information Sciences Institute (USC ISI), offers the fabrication of prototype VLSI chips to researchers in North America, for as little as $500 per project. Computer-aided design tools for personal computers and workstations are available at nominal costs from universities, and from innovative companies. Practical texts exist for learning the craft of analog VLSI design (Mead, 1989).

To become a new branch of computational neuroscience, analog VLSI modeling needs more participants, from both the engineering and the scientific communities. The territory is vast, and largely unexplored. The rewards are great for those who simply press forward.