35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Fast Neuromorphic Sound Localization for Binaural Hearing Aids

Paul K. J. Park, Hyunsurk Ryu, Jun Haeng Lee, Chang-Woo Shin, Kyoo Bin Lee, Jooyeon Woo, Jun-Seok Kim, Byung Chang Kang, Shih-Chii Liu, and Tobi Delbruck

Abstract— **We report on the neuromorphic sound localization circuit which can enhance the perceptual sensation in a hearing aid system. All elements are simple leaky integrate-and-fire neuron circuits with different parameters optimized to suppress the impacts of synaptic circuit noises. The detection range and resolution of the proposed neuromorphic circuit are 500 us and 5 us, respectively. Our results show that, the proposed technique can localize a sound pulse with extremely narrow duration (~ 1 ms) resulting in real-time response.**

I. Introduction

There is an increasing demand for the development of real-time and low-power sound localization technique in the hearing aid industries [1]-[2]. Currently, various digital processing techniques based on Fast Fourier Transform (FFT) have been proposed to determine from where a sound signal was generated [3]-[4]. However, these techniques require considerable latency (~ 100 ms) for the sampling of sound signal. In addition, it is necessary to use power consuming devices such as Digital Signal Processor (DSP), Analog-to-Digital (AD) converter, and memory to perform complex signal processing. On the other hand, it has been recently reported that a neuromorphic silicon cochlea can be utilized for spatial audition and auditory scene analysis [5]. Because the event-based cochlea produces a sparse stream of events just like the human auditory system, post-processing for sound localization can be cheaper than digital methods. Previously, it has been reported that the binaural sound localization is accomplished by measuring the Interaural Time Difference (ITD) [6]. The ITD between sounds arriving to two silicon cochleas can be extracted simply by comparing the timing information of spike outputs [6]-[8]. These techniques employed the computational algorithm (for example, cross correlogram [7] and artificial neural networks [8]) to estimate the location of a sound source. In this paper, we devise the neuromorphic sound processing circuit by mimicking the neuronal organization of barn owl's auditory pathway. In neurobiology, the dominant model of barn-owl's sound localization has consisted of a set of delay lines with predetermined delays of expected ITDs and an array of coincidence detectors, which are located in Nucleus Magnocellularis (NM) and Nucleus Laminaris (NL) of the auditory system [9]. Even though there have been several previous efforts to use digital circuits for implementing delay lines and coincidence detectors [10]-[12], it is still desirable to design in a spiking neuron circuit level for fast response of neuromorphic sound processing. In addition, it is required to develop a spatio-temporal processing circuit for obtaining auditory space maps tonotopically organized in the Inferior Colliculus (IC) [12]-[14]. Here we present a neuromorphic sound localization circuit being made up of delay, Coincidence Detection (CD), Time Division Multiplexing (TDM), and integration neurons. All processing elements are simple Leaky Integrate-and-Fire (LIF) neurons to be implemented in VLSI circuit without elaborate effort. Because the proposed technique needs to be designed with precise timing accuracy better than sub-millisecond, whole circuits become to be vulnerable to synaptic circuit noises. Thus, we optimized the membrane and synaptic time constants of each neuron element to minimize the effects of synaptic noises of VLSI circuit. In particular, we report that the neuromorphic spatio-temporal processing of auditory space maps can be implemented by using TDM and integration neurons. The results show that the proposed neuromorphic technique can localize a sound pulse with extremely narrow duration (~ 1 ms) resulting in real-time response.

Paul K. J. Park*, Hyunsurk Ryu, Jun Haeng Lee, Chang-Woo Shin, Kyoo Bin Lee, Jooyeon Woo, Jun-Seok Kim, and Byung Chang Kang are with SAIT, Samsung Electronics, San 14-1, Nongseo-dong, Giheung-gu, Yongin, 446-712 Korea (*corresponding author to provide phone: +82-31-280-1617; fax: +82-31-280-9860; e-mail: {k.j.park, eric_ryu, junhaeng2.lee, changwoo.shin,kyoobin.lee,jy.woo,junseok7.kim,bckang}@samsung.com).

Shih-Chii Liu, and Tobi Delbruck are with the Institute of Neuroinformatics, University of Zurich/ETH, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland (e-mail: {shih, tobi}@ini.phys.ethz.ch).

II. Method and Results

Neuromorphic silicon cochlea emulates basilar membranes, inner hair cells, and ganglion cells [15]. Because the spike outputs are transmitted rapidly via an address event interface, it properly preserves the timing relationships of synaptic transmission. Fig. 1 (a) shows the experimental setup to localize the sound source using the silicon cochlea. A speaker located to the left of the cochlea speaks "ah" while moving to the right (the distance between the speaker and the cochlea board is 40 cm). The incoming sound first passes through the basilar membranes which consist of 2 cascaded filter banks and is separated into 64 different frequency bands. The filtered signals are sent to the inner hair cells which perform a half-wave rectification of the input. The output of the inner hair cell then drives a set of 4 neurons within each frequency band [15]. Thus a total of 256 spike trains from each ear are generated as shown in Fig. 1 (b).

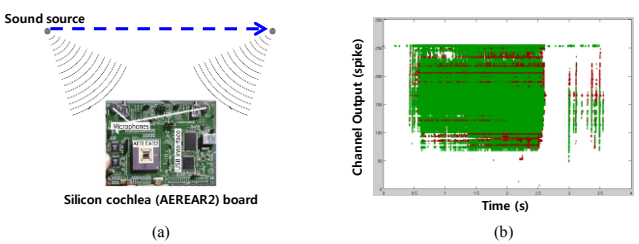

Figure 1. Silicon cochlea board and (b) spiking outputs.

Fig. 2 (a) shows the proposed neuromorphic sound localization technique based on LIF neurons. Spiral ganglion neurons in the barn owl produce action potentials in response to the sound waveform at each frequency channel. The temporal patterns of these action potentials are preserved along the time-coding pathway to NL [10]. Thus, the outputs of 256 ganglion neurons in each silicon cochlea are directly connected to CD neurons via several discrete time delays as shown in Fig. 2 (a). In this case, CD neurons receive inputs which are counter-propagating across delay lines from both silicon cochleas. Each CD neuron is sensitive to binaural sounds with a specific ITD because it fires whenever two combined Excitatory Post-Synaptic Potentials (EPSPs) present simultaneously from both ipsilateral and contralateral pathways reach a threshold. In neurobiology, the outputs of NL neurons are transmitted to IC neurons in order to produce a systematic map of auditory space [16]. To emulate the function of IC neurons, we propose a neuromorphic spatio-temporal processing method based on TDM and integration neurons. Each TDM neuron multiplexes all the outputs of CD neurons assigned to a specific ITD into one pathway. This enables us to investigate the temporal spike density generated at each ITD domain. Then the TDM output is sent to the integration neuron. This neuron integrates the temporal spikes over time window of fixed length which is determined by the membrane and synaptic time constants. Thus, the output spike of integration neuron represents the specific ITD at a certain time window. In this case, because one integration neuron also receives the TDM outputs from other ITD domains through inhibitory synapses, it prevents multiple firing of spikes at the same time window. In the numerical simulation, the number of delay neurons located in one channel was set to 200 (delay = 2.5 us). Thus, the measurement range and resolution of ITD detection are ±500 us and 5 us, respectively. Fig. 2 (b) shows simple synapse and LIF neuron circuits which can be implemented with a few transistors. The synaptic and neuron membrane time constants can be adjusted by setting capacitances.

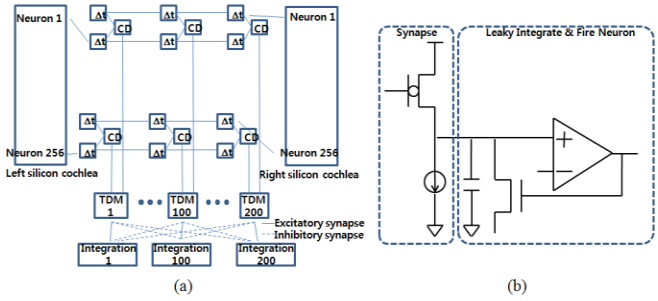

Figure 2. (a) Neuromorphic sound localization model (Δt: delay neuron, CD: coincidence detection neuron, TDM: Time Division Multiplexing neuron) and (b) synapse and neuron circuit.

It should be noted that, even though a neuromorphic chip is operating in a low-voltage environment, the electronic noise can be generated by the thermal agitation of the electrons. Thus, it is necessary to design the neuromorphic sound localization circuit considering the effects of the noise. In computational biology, a neuron is considered to contain various sources for noise. For example, ions and electrons performing a Brownian motion driven by thermal energy causes membrane noises. In addition, the spontaneous release of neurotransmitters causes synaptic noises [17]-[18]. The impact of noise on neuronal dynamics has been studied in detail in LIF neuron model. The general form of the basic equation for sub-threshold behavior of the LIF neuron can be written as

$$\tau_m dV/dt = V_r - V + RI \quad (1)$$

where τ_m is the membrane time constant, V is the membrane potential, V_r is the resting potential, R is the membrane resistance, and I is the input synaptic current. The input synaptic current including white noise can be expressed as

$$I(t) = \mu(t) + \sigma(t)\eta(t) \quad (2)$$

where μ is the mean of the synaptic input, σ is the standard deviation of the noise, and η represents white Gaussian noise. Equation (1) can be transformed to the stochastic differential equation by using Ito's Lemma and Wiener process, and then the membrane potential represented by the Markov chain V_t can be obtained as [18]

$$\Delta V_{k\Delta t} = (V_r - V_{(k-1)\Delta t} + R\mu_{(k-1)\Delta t})\Delta t/\tau_m + R\sigma_{(k-1)\Delta t} Z_k (\Delta t/\tau_m)^{1/2} \quad (3)$$

where Δt is the time step of numerical calculation and $\{Z_k\}$ are normal random numbers with mean zero and variance one. In this equation, the mean synaptic input current can be defined as $\mu(t) = w*\exp(-(t-t_0)/\tau_s)$, where w is the synaptic weight, t_0 is the onset time of synaptic input, and τ_s is the synaptic time constant [19]. Table 1 shows the neuron model parameters used in the numerical simulation. We investigated the neuronal dynamics by varying the membrane capacitance and synaptic time constant.

TABLE I. NEURON MODEL PARAMETERS

Parameter	Value	Description
V_r	-60 mV	Resting potential
V_0	-60 mV	Reset potential
R	1 MΩ	Membrane resistance

A. Delay Neuron

In the barn owl, afferent axons in the NL act as delay lines [20]. However, to use electrical delay lines, it should be required to have conductor lines with the total length of ~ 1 km. Thus, it is proper to employ a buffer for a neuromorphic delay line. In this paper, we utilized the LIF neuron as an analog buffer. Fig. 3 (a) shows membrane potentials calculated as a function of time. The membrane potential of a delay neuron starts to increase as soon as the synaptic current is injected. In this case, the membrane potential increases rapidly with faster membrane time constants. In the simulation, the maximum membrane potential was maintained to be constant at -42 mV by adjusting the synaptic weight. The delay neuron fires a spike when the membrane potential becomes beyond the threshold. Thus, the temporal spikes can be delayed with a certain amount of time. This delay can be controlled by setting the threshold properly. For example, to obtain a 2.5-us time delay, the threshold can be set to -46 mV when τ_m is 2 us. However, the delay time can be varied in the presence of the synapse and threshold noises. Fig. 3 (b) shows the delay variation simulated while varying τ_m and τ_s (number of simulation = 300). The results show that the standard

deviation of delay variation can be minimized when τ_m and τ_s are 3 ms and 4 us, respectively.

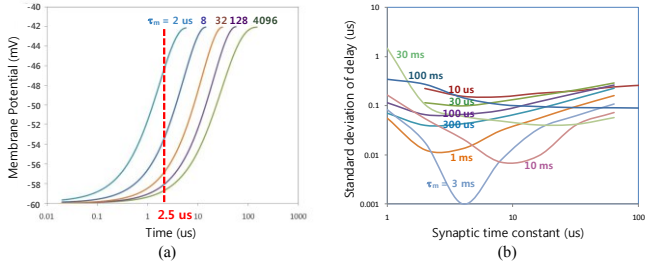

Figure 3. Membrane potentials calculated as a function of time (τ_m = 2, 8, 32, 128, 4096 us and τ_s = 30 us) and (b) standard deviations of buffer delays calculated while varying the synaptic time constant (τ_m = 10, 30, 100 us and 1, 3, 10, 30, 100 ms, σ = 1 nA, threshold noise = 10 uV, and number of simulation = 300).

B. *Coincidence Detection Neuron*

The CD neurons receive spikes conveyed from ipsilateral and contralateral pathways. In particular, these neurons respond to the spikes arrived simultaneously. For example, the maximum membrane potential decreases as the time difference between two synaptic inputs increases as shown in Fig. 4 (a). Thus, the detectable time difference can be controlled by adjusting the threshold. The detection resolution of these CD neurons is also changed by synaptic noises. Fig. 4 (b) shows the detection variation simulated while varying τ_m and τ_s when standard deviations of synapse and threshold noises are 0.3 nA and 10 uV (number of simulation = 300). The results show that the standard deviation of detection variation can be minimized when τ_m and τ_s are 1 us and 2 us, respectively. In the simulation, we assumed that the detection resolution was 3 us.

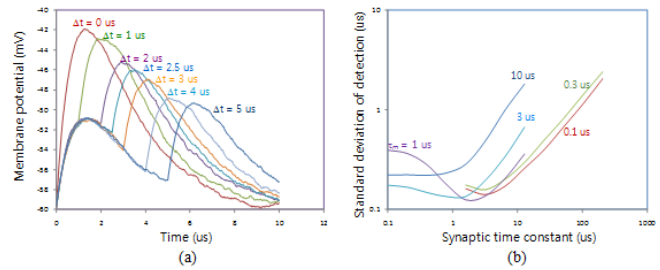

Figure 4. (a) Membrane potentials generated by two synaptic inputs (time difference = 0, 1, 2, 2.5, 3, 4, 5 us and σ = 0.3 nA) and (b) standard deviations of coincidence detections calculated while varying the synaptic time constant (τ_m = 0.1, 0.3, 1, 3, 10 us, σ = 0.3 nA, threshold noise = 10 uV, and number of simulation = 300).

Because the synaptic noise changes the actual detection resolution, the CD neuron does not respond to specific ITD, which in turn may cause a significant error. Fig. 5 (a) shows the number of full detection simulated while varying the designed detection resolution (simulation number = 300). In each simulation, ITD was varied from 0 to 5 us with a time step of 0.1 us (we assumed that ipsilateral and contralateral delays were 252.5 and 250 us, respectively). The results show that the detection resolution should be designed to be more than 3.5 us when σ is less than 0.8 nA. Using these design parameters and spikes data obtained from silicon cochleas in Fig. 1 (b), we performed the CD simulation as shown in Fig. 5 (b) (neuron channel = 136 (@ 910 Hz)). The results show that the firing CD neuron number is shifted with respect to time (i.e., measured ITD = -200 ~ 300 us @ time duration from 1 to 2.2 s).

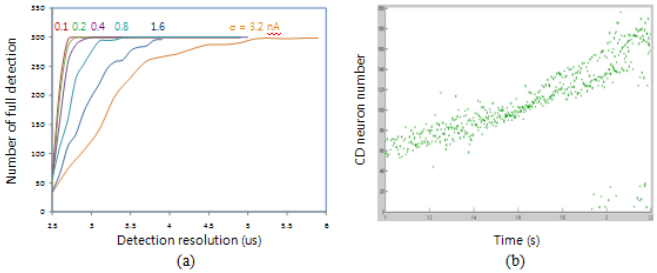

Figure 5. (a) Number of full detection simulated while varying the detection resolution (contralateral delay = 250 us, ipsilateral delay = 252.5 us, number of simulations = 300, τ_m = 1 us, τ_s = 2 us) and (b) outputs of 200 CD neurons @ 910 Hz (neuron channel = 136).

C. *Time Division Multiplexing Neuron*

Auditory neurons in IC integrate information across frequency channels to create an auditory space map. Fig. 6 shows auditory space maps measured during various time intervals. The measured channel frequencies were ranged from 846 Hz to 1411 Hz because the phase locking of cochlea is limited to frequencies less than 1.5 kHz [21]. The tonotopic distribution of spikes generated from CD neurons is shifted over to the right as time passes. The insets show the sum of spiking rates across all frequency channels. The results show that the auditory space map can be obtained even when the time interval is reduced to 1 ms (Fig. 6 (d), (e), and (f)). Thus, the proposed technique can localize a sound pulse with narrow duration (~ 1 ms) unlike the conventional digital method.

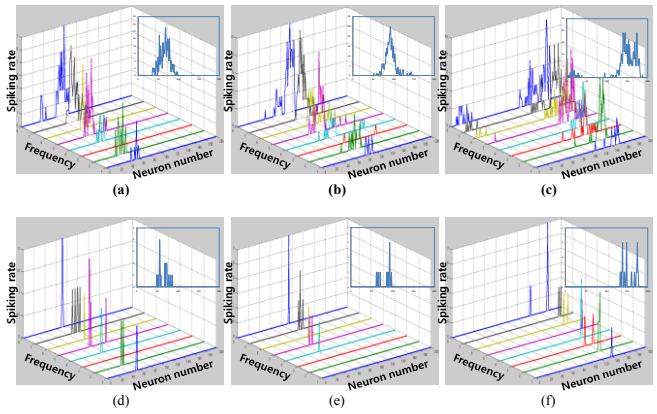

Figure 6. Measured auditory space maps (a) time duration = 50 ms @ 1.1 s, (b) time duration = 50 ms @ 1.5 s, (c) time duration = 50 ms @ 2.1 s, (d) time duration = 1 ms @ 1.1 s, (e) time duration = 1 ms @ 1.5 s, and (f) time duration = 1 ms @ 2.1 s.

To integrate all of the spikes information across whole frequency channels, it is required to integrate spikes from every pathway. To do this, we propose to combine all spikes generated from each frequency channel using the TDM neuron as shown in Fig. 2. Fig. 7 shows the output spikes measured at all TDM neurons. Here the membrane and synaptic time constants were designed to be very small to have sub-microsecond multiplexing resolution (τ_m = 1 us and τ_s = 0.01 us).

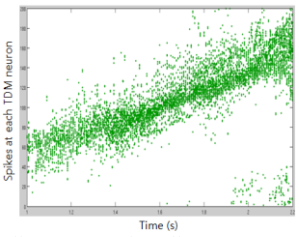

Figure 7. Output spikes measured at all TDM neurons (τ_m = 1 us and τ_s = 0.01 us).

D. Integration Neuron

Neurons can operate in two different ways, which are determined by mean inter-spikes interval of the input and the effective summative period of the neuron. For example, if the effective summative period is longer than the mean inter-spikes interval, neuron acts as temporal integrators. On the contrary, if the effective summative period is shorter, then it is expected to act like a coincidence detector. Thus, to integrate sub-microsecond input spikes in the proposed technique, the membrane and synaptic time constants were set to be longer than 1 us. It should be noted that each integration neuron receives spikes from all TDM neurons via either excitatory or inhibitory synapses. In this case, the synaptic weight of inhibitory synapse is increased as the channel distance between the integration and TDM neurons increases. This weighted synaptic connection prevents a simultaneous firing of integration neurons located at a long distance. Fig. 8 (a) shows the spikes of integration neurons when spikes trains generated with normal distribution are injected. By averaging the excitatory and inhibitory synaptic inputs, the output of integration neurons is not fluctuated sharply. Fig. 8 (b) shows the final output spikes of 200 integration neurons. In this figure, the membrane and synapse time constants were set to 1 us and 1 us, respectively. The integrating time window can be adjusted properly in accordance with a sound duration.

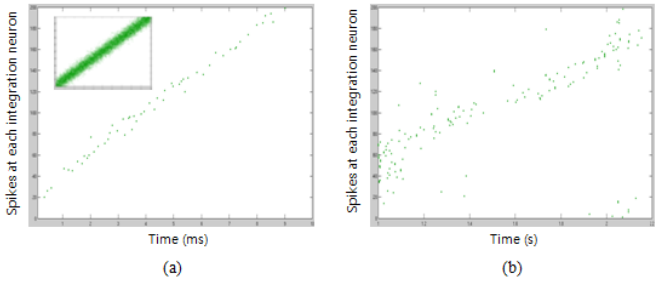

Figure 8. (a) Integrating example of integration neurons (inset: input spikes generated with normal distribution) and (b) final outputs of all integration neurons (τ_m = 1 us and τ_s = 1 us).

III. Conclusion

We have reported on the neuromorphic sound localization technique using spiking neuron circuits. Unlike previous techniques, the proposed technique can be easily implemented by analog silicon circuits. The results show that the proposed neuromorphic technique can localize a sound pulse with narrow duration (~ 1 ms) resulting in real-time response. In addition, because the whole neuromorphic sound localization circuit requires only 200,000 transistors, it can be implemented with low-power consumption. Thus, this technique holds promise for use in reliable real-time and low-power sound localization for binaural hearing aids.

References

[1] W. C. Wu, C. H. Hsieh, H.-C. Huang, O. T.-C. Chen, and Y. J. Fang, "Hearing aid system with 3D sound localization," in *Proc. of IEEE Int. Conf. on TENCON*, Oct. 2007.

[2] H. J. Simon, "Bilateral amplification and sound localization: Then and now," *Journal of Rehabilitation Research & Development*, vol. 42, pp. 117-132, 2005.

[3] Y. Tamai, S. Kagami, H. Mizoguchi, K. Sakaya, K. Nagashima, and T. Takano, "Circular microphone array for meeting system," *Sensors, Proceedings of IEEE*, vol. 2, pp. 1100-1105, 2003.

[4] S. S. Yeom, J. S. Choi, Y. S. Lim, M. Park, and M. Kim, "DSP implementation of sound source localization with gain control," in *2007 International Conference on Control, Automation and Systems*, pp. 224-229.

[5] S.-C. Liu, A. V. Schaik, B. A. Mich, and T. Delbruck, "Event-based 64 channel binaural silicon cochlea with Q enhancement mechanisms," in *2010 IEEE International Conference on Circuits and Systems*, pp. 2027-2030.

[6] S.-C. Liu and T. Delburck, "Neuromorphic sensory systems," *Current Opinion in Neurobiology*, vol. 20, pp. 288-295, 2010.

[7] H. Finger and S.-C. Liu, "Estimating the location of a sound source with a spike-timing localization algorithm," *IEEE International Conference on Circuits and Systems*, pp. 2461-2464, 2011.

[8] V. Y.-S. Chan, C. T. Jin, and A. V. Schaik, "Adaptive sound localization with a silicon cochlea pair," *Frontiers in Neuroscience*, vol. 4, p. 196, 2010.

[9] D. McAlpine and B. Grothe, "Sound localization and delay lines – do mammals fit the model," *Trends in Neurosciences*, vol. 26, pp. 347-350, 2003.

[10] J. S Lazzaro and C. Mead, "A silicon model of auditory localization," *Neural Computation*, vol. 1, pp. 41-70, 1989.

[11] C. Schauer and P. Paschke, "A spike-based model of binaural sound localization," *Int. J. Neural Syst.*, vol. 9, pp. 447-452, 1999.

[12] J. Liu, H. Erwin, S. Wermter, and M. Elsaid, "A biologically inspired spiking neural network for sound localization by the inferior colliculus," *Lecture Notes in Computer Science*, vol. 5164, pp. 396-405, 2008.

[13] B. Glackin, J. A. Wall, T. M. McGinnity, L. P. Maguire, and L. J. McDaid, "A spiking neural network model of the medial superior olive using spike timing dependent plasticity for sound localization," *Frontiers in Computational Neuroscience*, vol. 4, p. 18, 2010.

[14] J. L. Pena and W. M. DeBello, "Auditory processing, plasticity, and learning in the barn owl," *Institute for Laboratory Animal Research Journal*, vol. 51, pp. 338-352, 2010.

[15] V. Chan, S.-C. Liu, and A. V. Schaik, "AER EAR: a matched silicon cochlea pair with address event representation interface," *IEEE Transactions on Circuits and Systems*, vol. 54, pp.48-59, 2007.

[16] B. Grothe, M. Pecka, and D. McAlpine, "Mechanisms of sound localization in mammals," *Physiological Reviews*, vol. 90, pp. 983-1012, 2010.

[17] N. Fourcaud and N. Brunel, "Dynamics of the firing probability of noisy integrate-and-fire neurons," *Neural Computation*, vol. 14, pp. 2057-2110, 2002.

[18] F. Patzelt, "Mean and variance coding in neocortical neurons," M.S. thesis, Department of Physics, University of Bremen, 2007.

[19] W. Maass, T. Natschlager, and H. Markram, "A model for real-time computation in generic neural microcircuits," *Neural Information Processing Systems*, vol. 15, pp. 213-220, 2002.

[20] C. E. Carr and M. Konishi, "Axonal delay lines for time measurement in the owl's brainstem," *Proceedings of the National Academy of Sciences*, vol. 85, pp. 8311-8315, 1988.

[21] Y. Du, L. Kong, Q. Wang, X. Wu, and L. Li, "Auditory frequency-following response: a neurophysiological measure for studying the "cocktail-party problem," *Neuroscience & Biobehavioral Reviews*, vol. 35, pp. 2046-2057, 2011.