A Biologically Plausible Spiking Model of Human Auditory Pathways for Sound Localization

Xiao FENG[1,2 a], Weibei DOU[1,2 b]

[1]Department of Electronic Engineering, Tsinghua University, Beijing, China

[2]Center for Brain Inspired Computing Research, Tsinghua University, Beijing, China

[a]fx14@mail.tsinghua.edu.cn, [b]douwb@mail.tsinghua.edu.cn

Keywords: Spiking model, Sound Localization, Medial Superior Olive, Lateral Superior Olive, Inferior Colliculus.

Abstract. A comprehensive human auditory neural pathway spiking (HANPS) model from cochlea to inferior colliculus is proposed in this paper. It consists of medial superior olive (MSO) pathway for interaural time difference processing, lateral superior olive (LSO) pathway for interaural level difference processing, and the inferior colliculus (IC) module which integrates results. Architecture and design are provided which is strictly based on the physiological organization of human auditory pathway. The MSO and LSO pathway are both designed in a biological plausible way by using all neuron and no mathematical functions. Traditional binaural probabilistic model for localization is applied to the IC module in a Bayesian way. Quantities of simulations are conducted, analyzed which demonstrates the good ability to extract the localization cues and ability to integrate them, thus the ability to localize. Moreover, good ability to distinguish different sounds is also shown by simulation with multiple sound input. A brain-like circuits structure for the proposed HANPS model and some designing requirements are suggested.

Introduction

Human ears performs a remarkable ability to localize sound and distinguish multiple sounds. Researchers have been trying to simulate the way how human ears work, models built by whom could be applied to many applications such as hearing aid, robot ears and virtual reality. Human using both interaural time difference (ITD) and interaural level difference (ILD) to localize sounds is better known as the ``duplex theory of sound localization" and was first devised by Thompson and Rayleigh [1,2]. The auditory pathways for processing these cues have been extensively studied during the last decades [3,4]. Both cues have different workbands: (i) interaural time difference (ITD) for low frequencies (20Hz-1kHz), (ii) interaural level difference (ILD) for mid and high frequencies (>1kHz). The ITD and ILD cues are respectively processed in the medial and lateral superior olive (MSO and LSO), which together project to the inferior colliculus (IC) that combines these cues to estimate the azimuth of the sound in the midbrain [5].

In the field of auditory pathways spiking models for sound localization, most of the spiking models are focused on the ITD processing pathway including the MSO [6,7], due to the simple physiological structure and the fact that it is easier to implement using IF neurons. However, the localization precision of these models are usually low, such as 30 and 50 degrees.

Spiking model by Liu *et al.* for sound localization outlined an auditory processing system utilized both ITD and ILD cues [8]. The ITD processing uses an IF neuron model. However, the ILD processing does not use a neuron model, rather a logarithmic ratio computes the level difference and a spike is generated when a certain ratio is detected at a given frequency.

Wall *et al.* described another model for the ITD and ILD processing pathways in [9] where they showed that sound sources from different angles cause different spiking rates of the LSO. While pointing out that it is possible to judge the angle from spiking rates of the LSO, this model did not provide a module that implements this function. And in [10], Wall *et al.* remedied this deficiency. By adding a ``Receptive-field layer" to screen particular LSO spiking rates and training the connection weights, they successfully mapped different LSO spiking rates to different coming angles. Nonetheless, the screening process of the receptive-field layer uses a complex Gaussian function, a

function unable for neurons and synapses to perform, to render those spikes with particular spiking rates get through particular receptive-field neurons.

However, none of the former LSO pathways models works in the manner similar to an actual auditory pathway in human brain that implements the localization function based on only regular neurons which simply integrates membrane potential and generate spikes which leads to bad speed and energy performance. What's more, the MSO models works in a low localization precision. Thus the former models are characterized as low biological plausibility and bad performance. In this paper, we successfully designed a comprehensive model including both MSO and LSO pathway and a IC to integrate cues with only Integrate-and-fire (IF) neurons and no mathematical functions. The architecture of the proposed model is also faithfully based on physiology organization.

This paper is organized as follows. Section II presents the neurophysiological organization of the human auditory pathway and describes the architecture and design the proposed HANPS model. In Section III, experimental result is presented to show the feasibility and performance of the proposed model, and characteristics are listed in comparison to former models. Section IV discusses the structure and requirements of the brain-like circuit for the proposed model. Finally, conclusions are made in Section V.

Model Description

Physiological Organization. The auditory pathways in human brain, along which the nuclei are all organized in a tonotopical way [11], is shown in Fig. 1. Neural spikes are first generated at the anteroventral cochlear nucleus (AVCN) and sent to the MSO pathway which extracts the ITD from binaural sound based on a series of coincidence detectors (CD). The most important nucleus to process the ILD defined as the difference in sound pressure levels (SPL) is the LSO, a folded sheet of excited-inhibited neurons. Responses that the LSO outputs then interact with those from the MSO at the IC to form an estimation of coming angles [12]. Specifically, in response to a sound on the left, the left LSO receives excitation from the small spherical bushy cells of the ipsilateral AVCN, and inhibition from the contralateral side, meditated by the medial nucleus of the trapezoid body (MNTB), which transforms excitation from the contralateral AVCN to inhibition with no delay [12].
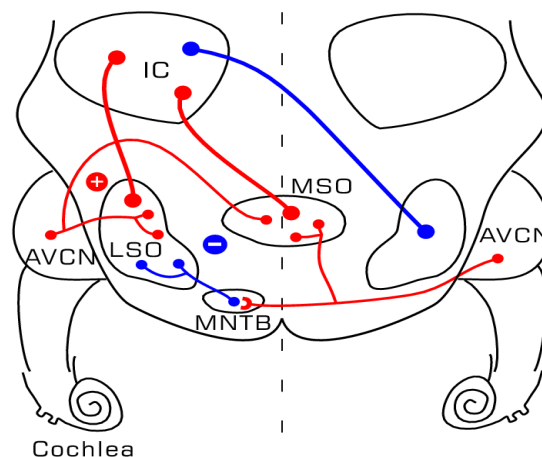

Fig. 1 Schematic diagram of the auditory pathways from the cochlea to the inferior colliculus (IC). AVCN, anteroventral cochlear nucleus; MSO, medial superior olive; LSO, lateral superior olive; MNTB, medial nucleus of the trapezoid body.

Model Architecture. The architecture of the whole proposed model is shown in Fig. 2. It shows that the whole model structure is based on the real physiological organization which consists both MSO and LSO pathways in an excited-inhibited manner. The cochlea module is the start point of frequency separation of N_{cf} subbands representing different frequencies. It works as a filterbank and send the processed result into the AVCN module which generates excitatory spikes. Theses spikes are sent to the MSO module, the ipsilateral MNTB and LSO module.
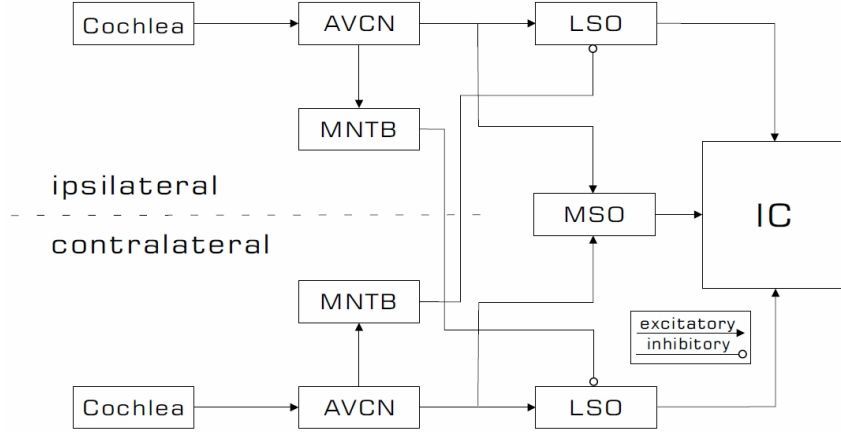
Fig. 2 The architecture of the whole proposed model.

MSO Pathway. The structure of MSO pathway is shown in Fig. 3(a). For each suband, there are N_{cd} coincidence detectors (CD) to extract the ITD from the binaural sounds. The synapse delay from AVCN to each CD is theoretically calculated to react to specific azimuthal angles.

LSO Pathway. Fig. 3(b) shows the topology of LSO pathway that contains MNTB modules to transform excitatory spikes from contralateral AVCN into inhibitory ones. In this excited-inhibited manner, level difference can be detected from binaural sound. In each suband of the IC module, each neuron represents a specific azimuthal angle thus the neuron of the maximum spiking rate could represent the coming angle of the sound. The weight of the synapses from LSO to the IC is constructed from a learning algorithm to get rid of the mathematical functions such as *log* and *Gaussian* that used to be used in the LSO pathway in former models.
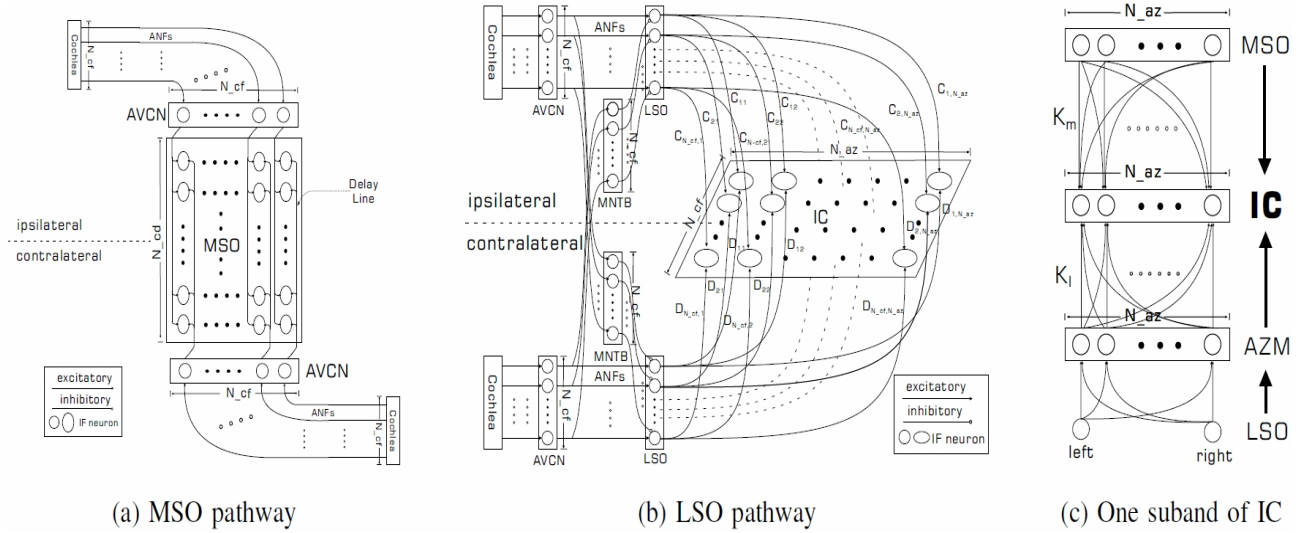

(a) MSO pathway   (b) LSO pathway   (c) One suband of IC

Fig. 3 Model Architecture.

Bayesian IC. Inspired by the binaural probabilistic model proposed in [13], we designed the IC module in our model in a Bayesian manner, shown in Fig. 3(c). We treat the IC module in the single LSO pathway as a part of the whole IC, labeled as AZM module. The weight of synapses from MSO to IC K_M is represented in a probabilistic way as eqn. (1) and we define K_L similarly.

$$K_{kij}^m = \begin{cases} R_m \cdot P(\theta_j | m_{ki}, f_k), & f_k < f_h \cap P > 0.5\max(\theta_j | m_{ki}, f_k) \\ 0, & otherwise \end{cases} \tag{1}$$

where f_k represents the character frequency of suband k, m_{ki} means the event that M_{ki} generates a spike, and R_m is general coefficient, f_h=1.5kHz. Thus $P(\theta_j | m_{ki}, f_k)$ denotes the probability that under the frequency of f_k, when M_{ki} spikes, the sound comes from θ_j.

From a series of mathematical transforms, we can present the equations above with all data that we can acquire from experiments. Then K_M and K_L will be settled.

Simulation and Comparison.

Fig. 4 shows the simulation results when input 1kHz pure tones with 10 different angles. Figures at the left side are the spike number of different angle-representing neurons, such as CD in MSO, at different angles of the input sound (θ). Figures at the right side are the probabilities that the estimated coming angle calculated from the spikes. It shows that MSO pathway gives a clearer and different estimation between the correct and the incorrect angles yet the angle with the maximum probability is usually 10 degree deviated from the correct one. On the other side, the LSO pathway gives a more precise estimation for the maximum probability yet the angles far away from the correct angle still get close probabilities to the maximum. After the integration of the two kinds of localization cues from the IC module, it combines the advantages and gets rid of the shortages of both MSO and LSO pathway. The probability estimation not only seem clear and has a great difference between correct and incorrect angles, but also maintains the correct angles get the greatest possibility.

When -50 degrees whitenoise is input into the model, the spike number and probability for each suband between 1kHz and 1.5kHz is shown in Fig. 5. It demonstrates good localization ability of the proposed model. Moreover, Fig. 6 shows the simulation results when 3 different sounds from different angle (f_s=1.02kHz, -10 degrees; f_s=1.22kHz, -40 degrees; f_s=1.47kHz, -70 degrees) are input together, which means the proposed model has a good ability to distinguish multiple sounds.

In comparison to former models, our model achieves the highest precision and the highest speed and energy performance under similar localization accuracy (63.9% and 85.7% for ±10 degrees), since the LSO pathways in our model is the first model that uses only neurons and no complex functions which also gives our model possess the highest biological plausibility among similar models. Moreover, besides localizing, our model demonstrates a great ability to distinguish multiple sounds. It is an ability hardly found in former models.

Brain-like Circuits

The suggested structure of the brain-like circuits for the proposed model is shown in Fig. 7. Notice that the suggested structure is very close to the physiological organization of human auditory pathway. This design is intended to make the performance and characteristics of the circuits as close to real human ears as possible.

The weight of the synapses in the circuits, to be specific, synapses from LSO to AZM and synapses from MSO and AZM to IC need to be trained before the whole model begins operating. We suggest the implemented circuits could operate the learning process both by online and offline procedure, which will make the circuits easy to adapt new sound environment where it is working. Moreover, neuron circuit and synapse circuit should better be made of simple circuit components, and the circuit parameters should be controllable to simulate different kind of neurons. In the MSO module, synapses work with precise different delays, which also requires the circuit to be able to implement this function.

Concluding Remarks

A comprehensive spiking model and its brain-like circuits structure for human auditory pathways have been proposed and tested by simulation and comparison. It can be used for sound localization, due to it contains not only both MSO and LSO pathway, but also the IC which integrates the information from the pathways and produces a decision result. It is called HANPS model because of the high biological plausibility that comes from of pure IF neurons, the structure and topology
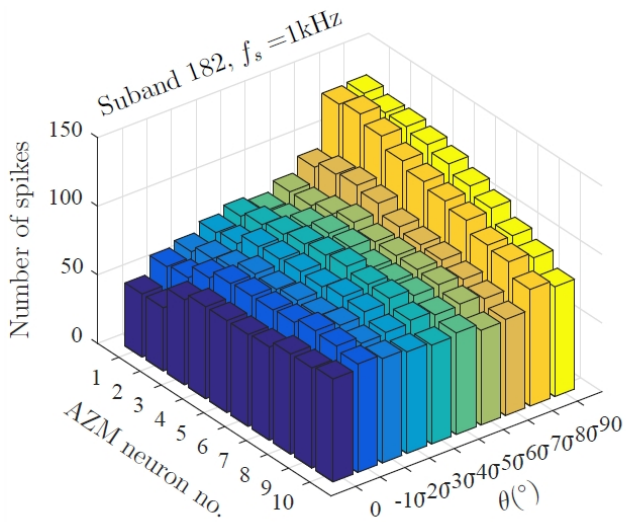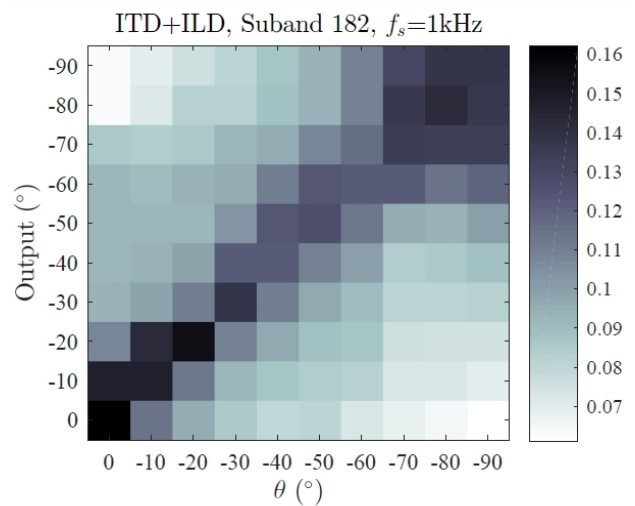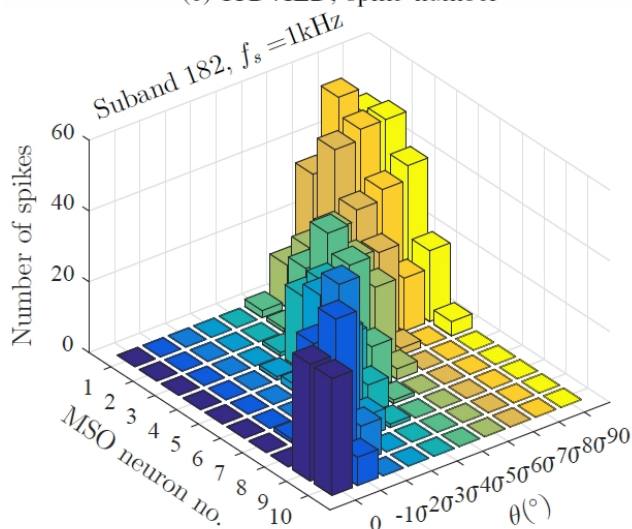
(c) ILD, spike number
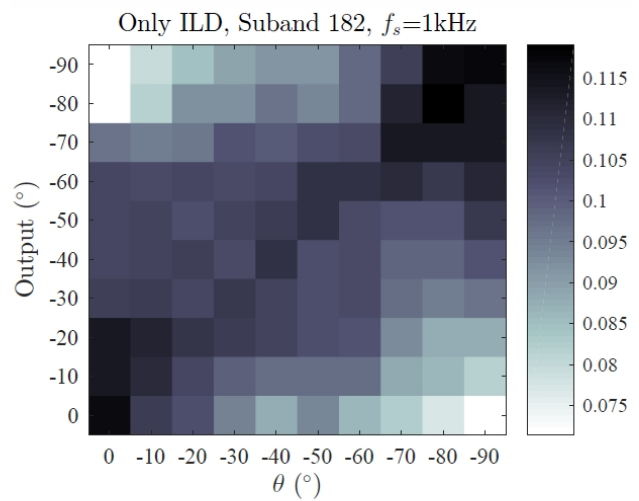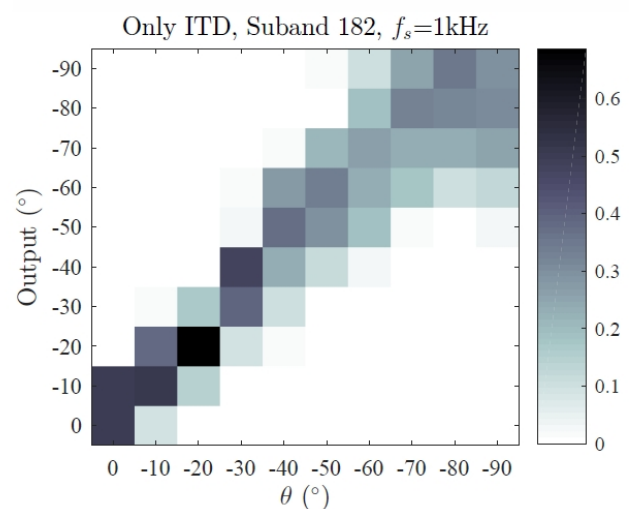
(d) ILD, probability
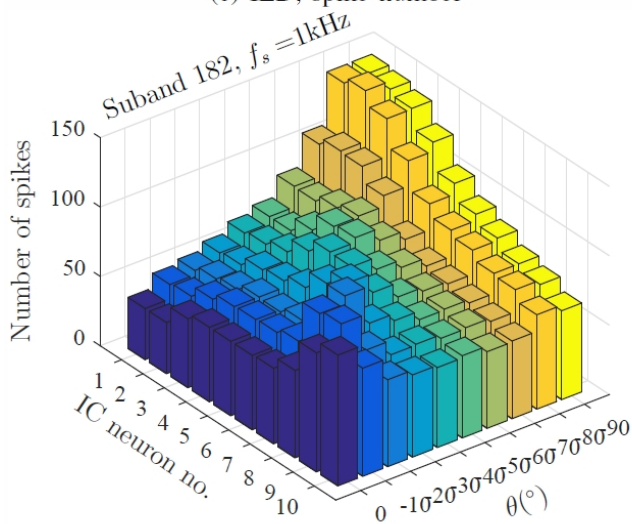
(e) ITD+ILD, spike number

(f) ITD+ILD, probability

(a) ITD, spike number

(b) ITD, probability

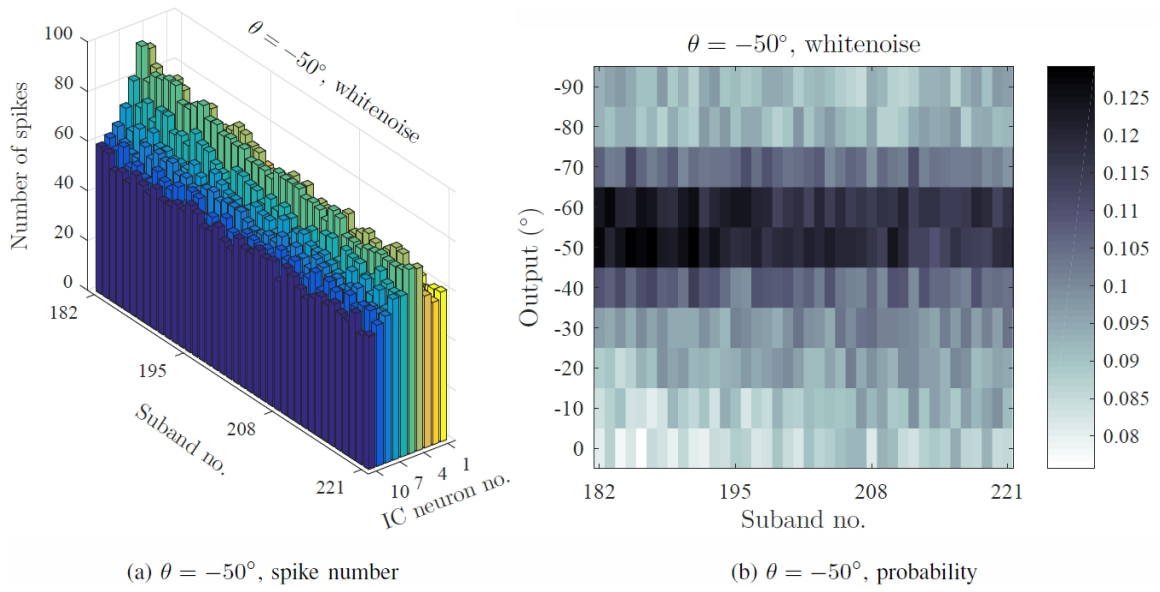Fig. 4 Spike number and probability when input 1kHz tones at suband 182.

(a) $\theta = -50°$, spike number

(b) $\theta = -50°$, probability

Fig. 5 Spike number and probability when input whitenoise

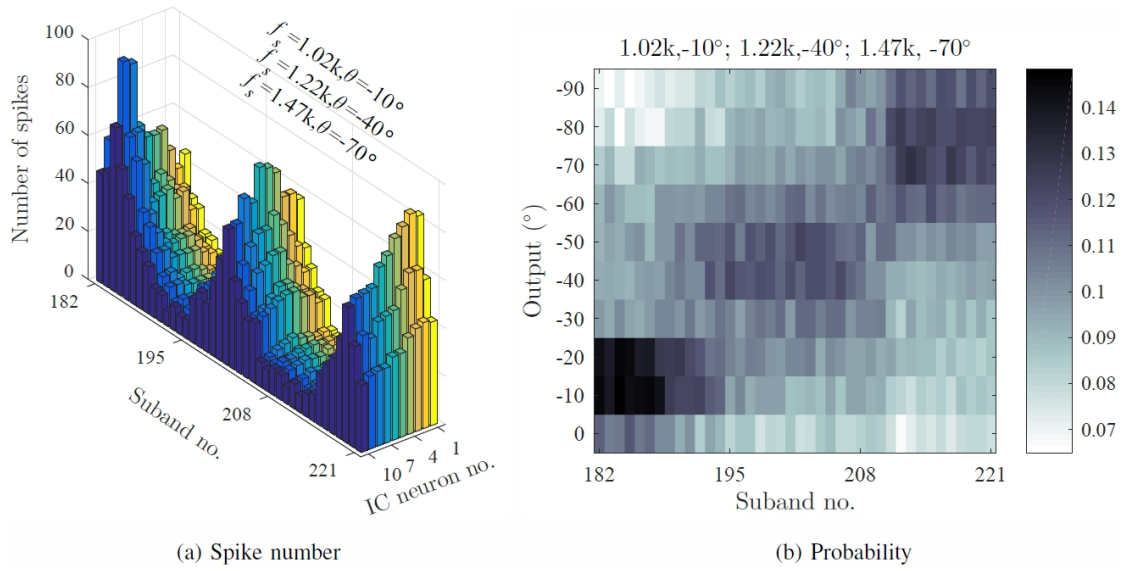(a) Spike number

(b) Probability

Fig. 6 Spike number and probability when input multiple sounds
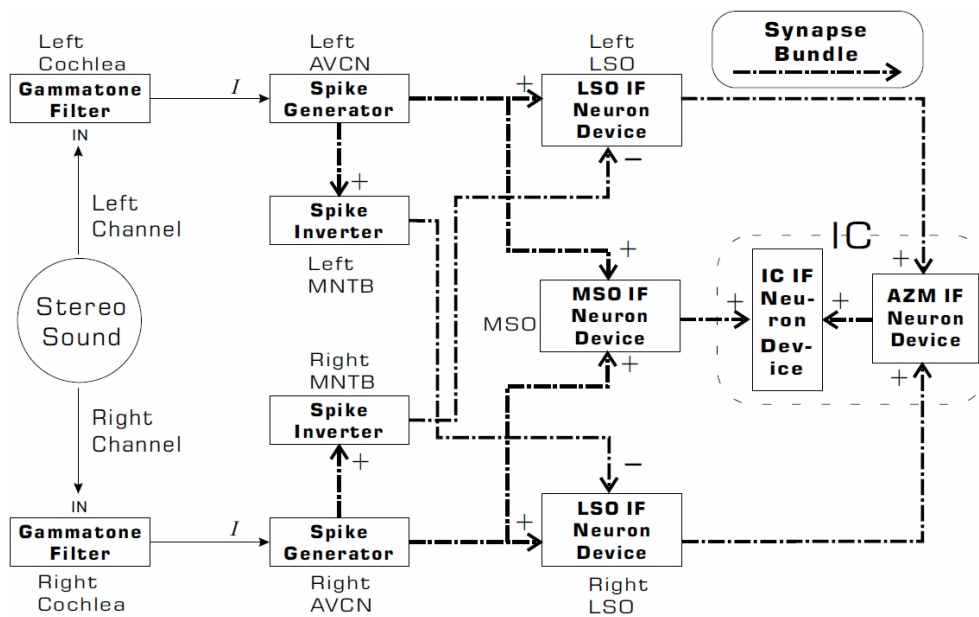

Fig. 7 The suggested circuit structure.

faithfully based on physiological architecture, and no complex mathematical functions. Among similar spiking models, the proposed HANPS model is the first one that is made from all neurons without complex functions and contains both MSO, LSO pathways and a Bayesian IC, which renders the proposed model to possess better performances at running speed, energy consumption, accuracy and precision. Some requirements have been suggested for implementing the brain-like HANPS circuits. These implementations should be a hard work in the future.

Acknowledgment

References

[1] S. P. Thompson, "On the function of the two ears in the perception of space," The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, vol. 13, no. 83, pp. 406–416, 1882.

[2] L. Rayleigh, "On our perception of sound direction," Philosophical Magazine, vol. 13, no. 74, pp. 214–232, 1907.

[3] D. J. Tollin, "The lateral superior olive: a functional role in sound source localization," The neuroscientist, vol. 9, no. 2, pp. 127–143, 2003.

[4] B. C. Moore, An introduction to the psychology of hearing. Brill, 2012.

[5] T. C. Yin, "Neural mechanisms of encoding binaural localization cues in the auditory brainstem," in Integrative functions in the mammalian auditory pathway. Springer, 2002, pp. 99–159.

[6] K. Voutsas and J. Adamy, "A biologically inspired spiking neural network for sound source lateralization," Neural Networks, IEEE Transactions on, vol. 18, no. 6, pp. 1785–1799, 2007.

[7] D. Goodman and R. Brette, "Learning to localise sounds with spiking neural networks," in Advances in Neural Information Processing Systems, 2010, pp. 784–792.

[8] J. Liu, D. Perez-Gonzalez, A. Rees, H. Erwin, and S. Wermter, "A biologically inspired spiking neural network model of the auditory midbrain for sound source localisation," Neurocomputing, vol. 74, no. 1, pp. 129–139, 2010.

[9] J. A. Wall, L. J. McDaid, L. P. Maguire, T. M. McGinnity et al., "Spiking neuron models of the medial and lateral superior olive for sound localisation," in Neural Networks, 2008. IEEE International Joint Conference on. IEEE, 2008, pp. 2641–2647.

[10] J. A. Wall, L. J. McDaid, L. P. Maguire, and T. M. McGinnity, "Spiking neural network model of sound localization using the interaural intensity difference," Neural Networks and Learning Systems, IEEE Transactions on, vol. 23, no. 4, pp. 574–586, 2012.

[11] D. Purves, G. J. Augustine, D. Fitzpatrick, L. C. Katz, A.-S. LaMantia, J. O. McNamara, S. M. Williams et al., Neuroscience. Sinauer Associates, 2001.

[12] B. Grothe, M. Pecka, and D. McAlpine, "Mechanisms of sound localization in mammals," Physiological Reviews, vol. 90, no. 3, pp. 983–1012, 2010.

[13] V. Willert, J. Eggert, J. Adamy, R. Stahl, and E. Körner, "A probabilistic model for binaural sound localization," Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, vol. 36, no. 5, pp. 982–994, 2006.