

# Mobile Robot Broadband Sound Localisation Using a Biologically Inspired Spiking Neural Network

Jindong Liu, Harry Erwin and Stefan Wermter

**Abstract**—A biologically inspired azimuthal broadband sound localisation system is introduced to simulate the functional organisation of the human auditory midbrain up to the inferior colliculus (IC). Supported by recent neurophysiological studies on the role of the IC and superior olivary complex (SOC) in sound processing, our system models two ascending pathways of the auditory midbrain: the ITD (Interaural Time Difference) pathway and ILD (Interaural Level Difference) pathway. In our approach to modelling the ITD pathway, we take account of Yin's finding that only a single delay line exists in the ITD processing from cochlea to SOC for the ipsilateral ear while multiple delay lines exist for the contralateral ear. The ILD pathway is modelled without varied delay lines because of neurophysiological evidence that indicates the delays along that pathway are minimal and constant. First, two-dimensional (2D) tonotopical ITD and ILD spike maps over frequency and ITD/ILD are calculated by a spiking neural network which follows the biological delay structure. Then these maps are weighted considering the advance of ITD in low frequency and ILD in middle and high frequency. Finally, ITD and ILD maps are merged together to find out the best estimation of the sound source. Experimental results involving noise and voice show that our model performs sound localisation that approaches biological performance. Our approach brings not only new insight into the brain mechanism of the auditory system, but also demonstrates a practical application of sound localisation for mobile robots.

**Index Terms**—Broadband sound localisation, inferior colliculus, interaural time difference, interaural level difference, interaural phase difference, auditory robotics

## I. INTRODUCTION

Sound source localisation is an important task for mobile robots operating in social environments, such as tour guide robots and rescue robots. The traditional approaches [1] for this task have some limitations such as requiring four or more microphones, restricting processing to one or a few input channels and requiring pure tone sound source. Recently, the known performance of animals using two ears in the sound localisation task has inspired researchers to work on new computational auditory models to understand the mechanisms of auditory neural networks. During the last twenty-five years, the structure and function of a number of pathways in the auditory brainstem have been well studied and become better understood [2]. For example, multiple spectral representations [3] are known to exist both in the early stages of sound processing, in the cochlea and cochlear

nucleus, and at later stage, from SOC to IC. Excitatory and inhibitory connections in the neural network of the midbrain sound processing pathways have been clarified [4]. Modelling these networks can help us to understand the brain mechanisms and provide a robust approach of broadband sound understanding to mobile robots.

The project discussed in this paper aims to study sound processing in the mammalian brain and to build a computational model that can be tested on a biomimetic mobile robots to validate and refine the neuroscience models for focused hearing. These models are expected first to clarify how the mammalian system achieves good sound localisation in noisy environments. Binaural sound localisation systems take advantage of two important cues [5] of the arriving sound signals in two ears: (i) interaural time difference (ITD) or interaural phase difference (IPD), and (ii) interaural level difference (ILD).

Assuming a pure tone sound source is closer to the left ear, the sound signal at the left ear can be represented by  $a \sin 2\pi ft$  while the delayed sound at the right ear is represented by  $(a - \Delta a) \sin 2\pi f(t + \Delta t)$ , where  $\Delta a$ , i.e. ILD, and  $\Delta t$ , i.e. ITD, represents respectively the sound fading and extra time for the sound wave to propagate the distance to the right ear. Using these two cues, sound source position can be estimated in the horizontal or azimuthal plane. ITD is believed to be more efficient at localising low frequency sounds (50Hz ~1.2kHz) while ILD is better for mid- and high- frequency sound (>1.0 kHz) [5].

Models of ITD and ILD processing have been developed by several researchers [5][6]. Jeffress [5] originally proposed a widely used model to detect ITDs, in which sound from each ear passes different delay lines before reaching a coincidence neuron which fires with a maximum rate when two specific delay times are present for the sound. Yin [6] improved Jeffress's model in response to biological evidence by introducing a single delay line for the ipsilateral ear while retaining multiple delay lines for the contralateral ear. For ILD, Jeffress suggested a so-called "latency hypothesis" to explain the processing mechanism, in which the latency of inhibitory input is delayed relative to the excitatory input from the ipsilateral ear. Evidence for this idea was provided by Hirsch [7], however, the mechanism of ILD processing remains unclear. In an engineering study, Bhadkamkar [8] proposed a system to process ITD using a CMOS circuit. Willert [9] built a probabilistic model which separately measures ITDs and ILDs at a number of frequencies for binaural sound localisation. Recently, Voutsas realised a single delay line Jeffress model by using spiking neural networks which

This work is supported by EPSRC (EP/D055466)

J. Liu, H. Erwin and S. Wermter are with Faculty of School of Computing and Technology, University of Sunderland, Sunderland, SR6 0DD, United Kingdom [jindong.liu](mailto:jindong.liu), [harry.erwin](mailto:harry.erwin), [stefan.wermter@sunderland.ac.uk](mailto:stefan.wermter@sunderland.ac.uk)

incorporate realistic neuronal models. He takes into account ITDs and gets excellent results. However, ILDs have not been utilised in his model so that his system is only sensitive to low frequency sound (50 Hz  $\sim$  1.2 kHz). Furthermore, the sound processing time of his model is too long to implement real-time applications.

This paper presents a new auditory processing system designed to provide live sound source positions via a spiking neural network. It is the first example of applying both ITD and ILD into a spiking neural network to localise a broadband sound. In this system, the sound from two microphones is filtered by a Gammatone filterbank and splitted into a number of frequency channels. Each channel is then encoded into phase-locking spikes to simulate the biological spiking feeding into the MSO (medial superior olive). Then, a two-dimensional (2D) tonotopical ITD map over frequency and ITD is calculated by a spiking neural network which processes the phase-locking spikes by following the biological delay structure. Meanwhile, 2D ILD spike map is calculated directly from the Gammatone filterbank results by computing the logarithmic ratio of two sound levels. Finally, these maps are weighted considering the advance of ITD in low frequency and ILD in middle and high frequency. ITD and ILD maps are merged together to find out the best estimation of sound source.

The rest of this paper is organised as the follows. Section II presents the neurophysiological data of human auditory pathway. Section III proposes a system model using a spike neural network which combines the ITD pathway and the ILD pathway. In Section IV, experimental results involving noise and speech are presented to show the feasibility and performance of the sound localisation system. Finally, conclusions and future work are given in Section V.

## II. BIOLOGICAL FUNDAMENTALS AND ASSUMPTIONS

When sound arrives at the external ear, it is directed through the outer auditory canal to the ear drum and then propagates to the middle ear, where it triggers the vibration of auditory ossicles. As a result of the vibration, the sound pressure is amplified and transmitted from the medium of air into a liquid medium in the inner ear. The sound pressure in the inner ear propagates down the basilar membrane in cochlea which is sensitive to the sound signal in a tonotopical way, i.e. different tone frequencies are transmitted separately along specific parts of the structure. Higher frequency sound is sensed at the beginning of the cochlea while lower frequency sound at the end of it. The inner hair cells along the cochlea respond to the motion of the basilar membrane to generate spikes that are sent through the auditory nerve (AN) up to the central nervous system.

The temporal and amplitude information of the sound wave are encoded by the hair cells up to auditory nerve [5]. Two properties of the hair cells are important for this encoding. First, the cochlea is tonotopically organised so that each inner hair cell is maximally excited by stimulation at a characteristic frequency (CF) [10]. In other words, each hair cell has a specific frequency with a highest sensitivity.

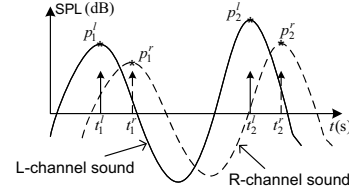


Fig. 1: An example of phase locking spike signal ( $t_1^l$ ,  $t_1^r$ ,  $t_2^l$  and  $t_2^r$ ).  $p_i^{l/r}$  is the sound level value when phase locking spikes are generated.

Second, the hair cells are polarised so that their movement is excited during 90 degree phase of the sinusoidal sound wave while inhibited during other phases. This phase locking occurring at frequency of 50  $\sim$  1.2 kHz is the basis for the encoding of timing information for sound. Figure 1 shows an example of phase locking spiking signals, where the hair cells generate spike when the sound wave has the phase of 90 degree ( $t_1^l$ ,  $t_1^r$ ,  $t_2^l$  and  $t_2^r$ ).

As the sound pressure level (SPL) increases, auditory nerve fibers increase their rate of discharge with a sigmoidal relationship to the decibel of sound over a relative range of 30 dB. In order to cover a wide SPL range, e.g. 120 dB, the relative range is adaptively changed according to the background sound level. However, the mechanism of this adaptivity is still not clear. In this paper, we do not model the biological amplitude encoding and directly generate ILD spikes. See details in later paragraph.

After encoding the temporal and amplitude information, AN fibers pass the spiking sequence through the superior olivary complex (SOC) up to the inferior colliculus (IC) to calculate the ITDs and ILDs. Two separate circuits, the ITD pathway and the ILD pathway, are involved in the calculation. The ITDs [10] are proved to be processed in the medial superior olive (MSO), which is one part of the SOC, and the ILDs are processed in another part of the SOC, i.e. the lateral superior olive (LSO). The MSO in one side receives excitation from the AVCN (anteroventral cochlear nucleus) from both the ipsilateral and contralateral ears. Figure 2 shows Yin's model of the ITD. The neurons in the MSO are assumed to be distributed along two dimensions: characteristic frequency (CF) and ITD. In the figure, the length of the delay line represents the delay duration in the propagation of the spiking signal from the AVCN to the MSO, so that the ipsilateral spiking signals have a fixed delay time (the single delay line in the figure), while the contralateral spiking signals have variable delays (the multi-delay lines in the figure). When spikes arrive from both sides at the same time, the ITD neuron fires. After the ITD processing in the MSO, the results are projected to the IC.

For the ILDs processing, cells in the LSO are excited by sounds that are greater in level at the ipsilateral ear than the contralateral ear and inhibited by sounds that are greater in level at the contralateral ear [10]. Therefore, the LSO receives excitation from the ipsilateral AVCN, but inhibition from the MNTB (medial nucleus of the trapezoid body)

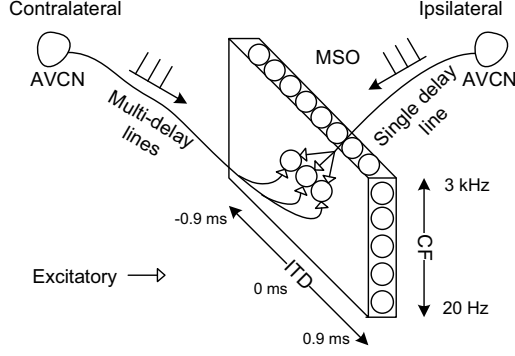


Fig. 2: Schematic diagram the ITD processing of the MSO in the human.

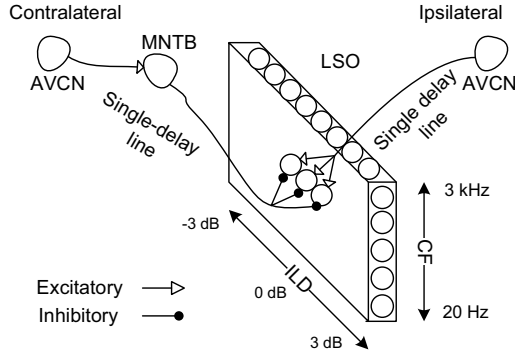


Fig. 3: Schematic diagram the ILD processing of the LSO in the human. Note, the ILD range in human is actually 30 dB. 3dB is chosen for our robot head and it is adjustable.

which converts the excitation from the contralateral AVCN to inhibition. In the LSO, the differences in timing between the ipsilateral and contralateral spiking inputs have been studied by neurophysiologists. Joris [11][12] found that there is only one fixed delay line on each side rather than variable delay lines in the MSO. Although the model of the ILD processing is unclear yet, the spike rate of LSO is found to be proportional to the sound level difference between two sides. In this paper, we assume there is a ILD map (please see Figure 3) in LSO, similar to the ITD map in MSO. Each cell in the ILD map generates ILD spikes according to the sound level difference. For example, for the sound signal in Figure 1, the left side LSO cell with characteristic ILD of  $\log(p_1^l/p_1^r)$  will generate a spike at time  $t_1^l$ , while the right side LSO cell with characteristic ILD of  $\log(p_1^r/p_1^l)$  will generate a spike at time  $t_1^r$ . Note that the LSO cell also has characteristic frequency similar to the MSO and the spiking output of the LSO is tonotopically projected to IC too.

### III. SYSTEM MODEL OF SOUND LOCALISATION

Inspired by the neurophysiological data and the proposed ILD pathway assumption presented in Section II, we designed a system model of sound localisation by using spiking

neural networks (SNNs). The sound is first encoded into spikes as inputs to the SNN. The synaptic response  $I(t)$  to a spike, occurred at  $t = t_s$ , is modelled as a constant square current with an amplitude (also called weight) of  $w_s$ , a latency  $l_s$  relative to the timing of the spike, and a lasting time  $\tau_s$ . The sign of the the amplitude indicates whether the synapse inhibits (negative) or excites (positive) the following soma. The soma response to  $I(t)$  can be modelled based on the leaky integrate-and-fire model in [13].

$$u(t) = u_r \exp\left(-\frac{t-t_s}{\tau_m}\right) + \frac{1}{C} \int_0^{t-t_s} \exp\left(-\frac{s}{\tau_m}\right) I(t-s) ds \quad (1)$$

where  $u_r$  is the initial membrane potential, and  $\tau_m$  is a time constant. In this paper, a typical value for  $\tau_m$  is 1.6 ms.  $C$  is the capacitor which is charged by  $I(t)$ , in order to simulate the procedure of the postsynaptic current charging the soma. The soma model has one more parameter, the action potential  $\varphi$ . When  $u(t) = \varphi$ , the soma will fire a spike, then  $u(t)$  is reset to 0.

In contrast to other sound localisation systems, e.g. [14], which only applied an ITD pathway, our model utilises both the ITD and ILD pathways for both sides of ear. This feature provides a broadband frequency localisation ability to the model as we will see in Section IV. Furthermore, the SNN of the model simplifies the model of the synapse, and the parameters of the synapse and soma are independent to sound frequencies in contrast to the model in [14]. This feature enables our system to have a real-time computation ability and potential optimisation possibility. Figure 4 shows a schematic structure for the sound localisation procedure. In the figure, the tonotopical sound encoding feature of the cochlea is simulated by a bandpass filterbank consisting of a series of discrete second-order Gammatone filters [15]. The filterbank decomposes the sound into several frequency channels in a similar way as the cochlea processing the sound.

After the Gammatone filterbank, the temporal information of the sound in each frequency channel is encoded into a spike sequence by the phase locking module in Figure 4, which simulates the phase locking character of the inner hair cell in the cochlea. At every positive peak of each frequency channel, a phase locking spike is triggered and then fed into the ITD model later. Meanwhile, the sound level is detected at the time of phase-locking spike. It is used for the following ILD model.

Following the ITD pathway in Figure 2, the spike sequence of the contralateral ear passes variable delay lines  $\Delta t_i$ . We denote the delayed spike sequence as  $S_{CP}(\Delta t_i, f_j)$ , where  $C$  stands for the contralateral,  $P$  for phase locking,  $\Delta t_i$  for the delay time,  $f_j$  for the frequency channel  $j$ . Similarly,  $S_{IP}(\Delta T, f_j)$  represents the delayed spike sequence of the ipsilateral ear with a fixed delay time  $\Delta T$ .  $S_{CP}(\Delta t_i, f_j)$  and  $S_{IP}(\Delta T, f_j)$  are then input to the ITD coincidence model (please see Figure 5-(a) for details) to calculate the ITD. The calculated output of the ITD coincidence model is a new spike sequence represented as  $S_{ITD}((\Delta T - \Delta t_i), f_j)$ .

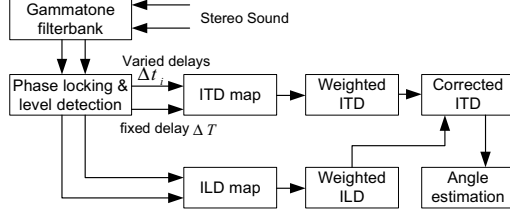


Fig. 4: Schematic structure of biologically inspired sound localisation system. This example assumes the left ear is the ipsilateral ear.

If there are spikes in  $S_{ITD}((\Delta T - \Delta t_i), f_j)$ , it means the sound arriving to the ipsilateral ear is earlier than that to the contralateral ear by  $ITD = \Delta T - \Delta t_i$  second. Once the ITD calculation is implemented for all frequency channels, a two dimension ITD distribution map can be drawn, where the x-axis is for ITD, y-axis for the frequency channel and z-axis for the mean spike number in a unit time. Figure 6 shows an ITD calculation results for a speech of 'fish' from left side at 60 degree in 1.28m. A negative ITD value means the sound arriving at the right ear is later than that at the left ear.

The ILD pathway is modelled without using the leaky integrate and fire model. Instead, the detected sound levels of two sides are compared to calculate the level difference and the corresponding ILD cell is fired with a spike. The level difference can be calculated as  $\Delta p^j = \log(p_I^j/p_C^j)$ , where  $p_I^j$  and  $p_C^j$  stand for the ipsilateral and contralateral sound level for the frequency channel  $j$  respectively. Similar to the ITD model, the output of ILD model can be represented as 2D spiking map,  $S_{ILD}(\Delta p^j, f_j)$ . Figure 7 shows an ILD calculation results for the same sound in Figure 6. A negative ILD value means the level of sound going to the right ear is lower than that to the left ear.

Before merging ITD and ILD maps for broadband sound localisation, we have to consider the advance of ITD and ILD in different frequency band. From biological evidences, ITD can provide accurate sound source azimuth information up to 1.2kHz and starts to include more and more ambiguity when frequency is above 1.2kHz (e.g. see Figure 6). In contrast, ILD plays a trivial role in low frequency (<1kHz) localisation and begin to be meaningful when high frequency sound arrives. Therefore, we build two weight arrays (Equation 2),  $ITD_w$  and  $ILD_w$ , over frequency and compute a weighted ITD or ILD map by multiplying the weight array with the matrix of 2D ITD/ILD map. Note that the spiking number has been summed up over frequency during matrix multiplication.

$$\begin{aligned} ITD_w^j &= \frac{\sum_j (\max(f_j/1200, 1))}{\max(f_j/1200, 1)} \\ ILD_w^j &= \frac{\max(\log(f_j/1000), 0)}{\sum_j (\max(\log(f_j/1000), 0))} \end{aligned} \quad (2)$$

where  $j$  is the frequency channel index.

The weighted ITD and ILD maps are finally merged together as shown in the "corrected ITD" module of Figure 4.

Considering the complex head transfer function between the ITDs and source source angles [3], we use the ILD maps to roughly identify the sound source direction with  $\pm 30$  degree tolerance. Then we use this information to remove the ambiguity in the ITD map. For example, if peak point of ILD map indicates the sound source from 45 degree, we can conclude that the sound came from a range of  $[15, 75]$  degree and then all the ITD spike out of the range can be ignored. After correcting ITD spikes, we can choose the significant ITD to estimate the sound source angle. The relationship between ITD or ILD with the sound source azimuth angle is obtained by sampling a broadband noise over -90 to 90 degree range and measuring the ITD or ILD in each angle.

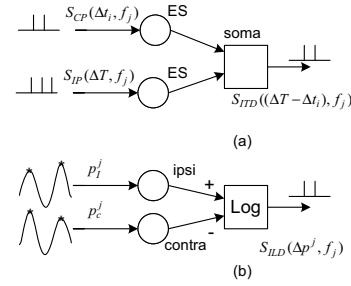


Fig. 5: (a) ITD coincidence model (b) ILD model. ES stands for excitatory synapse. Ipsi for ipsilateral and contra for contralateral Gammatone frequency channel.

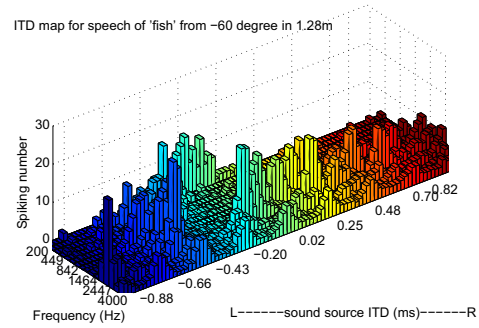


Fig. 6: ITD calculation illustration. The sound source is a speech of 'fish' from left side at 60 degree in 1.28m. The sound is recorded with a real robot head in a general quiet room. The correct ITD spike should be around -0.5ms. However, there are many ambiguity spikes for frequency channel above 1.2kHz.

#### IV. EXPERIMENTAL RESULTS

To justify the feasibility and performance of our sound localisation model, we tested our system on a real robot head (Figure 8) which has the similar dimension as an adult human. One pair of binaural cardioid microphones (from Core Sound) are equipped at the ear positions together with

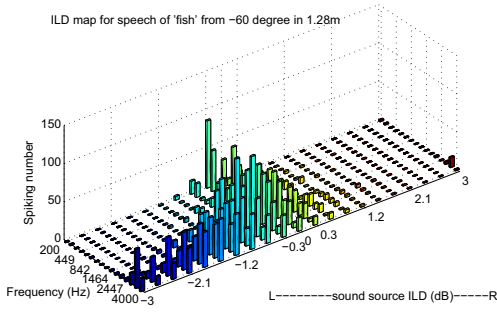


Fig. 7: ILD calculation illustration. The sound source condition is the same as in Figure 6

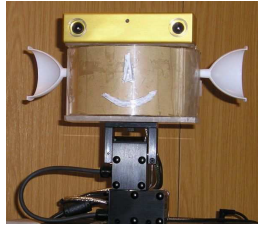


Fig. 8: The head of our robot for sound localisation experiments. It is equipped with two microphones and a stereo vision.

half-cone cups acting as pinnae. Two microphones are placed 15 cm apart each other. The whole robot head sits on a pan-tilt unit to have the ability to turn around to face the sound source. Two types of broadband sound sources are used: noise and speech. The speech includes five words, hello, look, fish, coffee and tea, which cover a wide frequency pattern in our daily speech.

The sound in real pure tone localisation is recorded in a low noisy environment with 5 dB background noise. The distance of the sound source to the robot is 128 cm and the sound pressure at the speaker side is controlled to be  $90 \pm 5$  dB. The sample rate is 44100 Hz and the duration is 1.5s with 10 ms silence at the beginning. One speaker is placed at the positions of -90, -60, -30, 0, 30, 60 and 90 degrees. The parameters of the spiking neural network are listed in Table I. The Gammatone filterbank is set with 16 channels over 200 to 4000 Hz. The relationship between ITD/ILD with sound source angle  $\theta$  used in our experiments calculated in advance by sampling the noise in the same environment. See Equation 3 for details.

$$\begin{aligned} \theta(ITD) &= \arcsin((ITD + 0.03)/0.684)/0.9155 * 180/\pi \\ \theta(ILD) &= \arcsin((ILD + 0.1)/0.8304)/0.7 * 180/\pi \end{aligned} \quad (3)$$

Figure 9 shows the spike distributions in the sound localisation of three broadband sounds. In these distribution figures, the x-axis is the sound source azimuth angle and

TABLE I: Parameters for ITD model

	Synapse			Soma			$\Delta t_i$	$\Delta t_i$
	$l_s$	$\tau_s$	$w_s$	$\theta$	$t_m$	$C$	step	range
ITD	2.1	0.08	0.1	$8e-4$	1.6	10	$2.26e-2$	$[-0.9 \ 0.9]$

\*Note:  $\Delta T = 0$ . The unit of  $l_s$ ,  $\tau_s$ ,  $w_s$ ,  $t_m$  and  $\Delta t_i$  step/range, is ms and the unit of  $C$  is mF.

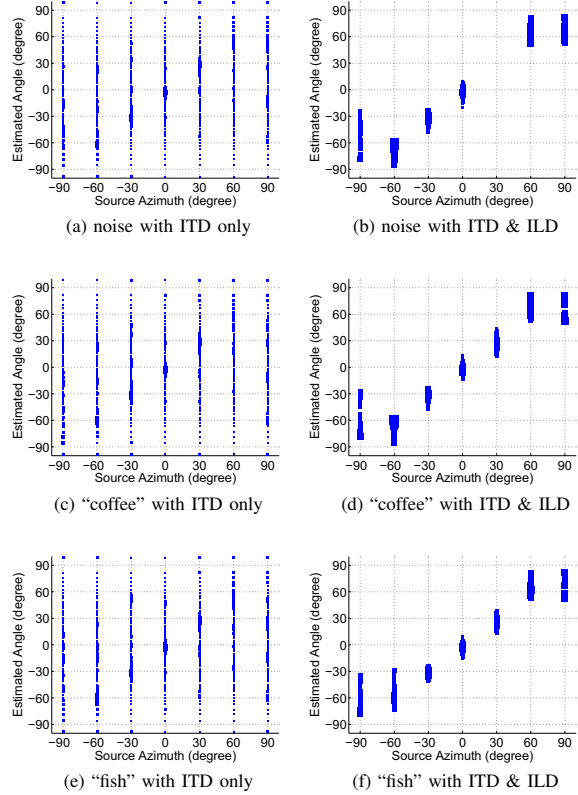


Fig. 9: The artificial sound localisation results for noise, “coffee” and “fish”. The size of square represents the spiking proportion in the corresponding angle estimation.

the y-axis is the estimated angle after calculation. The size of square is proportional to the ITD spiking number in the corresponding angle estimation. For example, in Figure 9b, a big square at (-60 60) means that when the sound comes from -60 degree the ITD spiking number of estimated angle of -60 degree is the majority of spikes of all estimated angles. The localisation efficiency of the system is defined as a percentage of the spiking number at the correct estimation point, such as (-30 -30) and (60 60), to the total spiking number. In the figure, we compare the performance of two methods, i.e. the localisation using (i) the ITD only and (ii) the corrected ITD which is amended by both weight array and the ILD.

In the left column of Figure 9, the estimation is only based on ITD information. The estimated angle spreads over the

whole range of azimuth angles and it is hard to see where the significant value is. This is because the incoming sound is broadband sound including the frequency from 200 to 4000Hz. The ITD only works well in low frequency and high frequency component introduces many ambiguities. The localisation efficiency across all sound types and angles is only about 25%.

After using weights on ITD and ILD over frequency and merging the ILD with ITD, the experiments results shown in the right column of Figure 9 demonstrate that the spiking distribution is much more concentrated than that of the results by only using the ITDs. The overall localisation efficiency is increased to 80% and 90% of the sound signals across all frequencies were recognised correctly between -45 to 45 degree. The highest localisation efficiency occurred at the 0 degree. The efficiency decreases when the sound moves to the sides.

Our experimental results match the fact [10] that (i) the ITDs cue has the highest efficiency for sound localisation when the sound source is in front of the observer, (ii) the ITD cue effect on sound localisation fades down over 1.2 kHz, and (iii) the ILDs is the main cue for the high frequency sound localisation.

The time cost for processing a 100 ms sound signal is less than 50 ms on a 2.6 GHz CPU. Comparing the 9.1 s time cost when using the model in [14], our system performs much quicker without losing localisation efficiency. It is feasible to apply our system for real time sound localisation.

## V. CONCLUSION AND FUTURE WORK

In this paper, we designed and implemented a sound localisation system using spiking neural network inspired by mammalian auditory processing. In this system, both the ITD and ILD pathway were adopted and modelled based on recent neurophysiologic findings. ITD/ILD spikes were weighted according to the advance of ITD/ILD in difference frequency bands localisation, and they were merged together to achieve broadband sound localisation. The experimental results showed that our system can localise a broadband sound source from the azimuth -90 to 90 degree. The sound frequency can vary from 200 to 4000 Hz. The effect of frequency and sound source position on the localisation efficiency had a high correspondence with neurophysiologic data.

In the future, active sound localisation, which can specify the feature frequencies of an interesting object, will be the next step of our research. For the application of our system to a mobile robot, we are planning to implement a self-calibration sound localisation system which can adaptively adjust the synapse and soma parameters according to the environment or electrical hardware change.

## ACKNOWLEDGMENT

Our thanks go to Dr. Adrian Rees and Dr. David Perez Gonzalez at University of Newcastle for their contribution on neurophysiologic support toward the project. This project, MiCRAM, is a cooperative project with them and they are

mainly specialising in the neurophysiologic structure on the inferior colliculus for sound processing. We would also like to thank Chris Rowan for helping with building the robot head.

## REFERENCES

- [1] J. Huang, T. Supaongprapa, I. Terakura, F. Wang, N. Ohnishi, and N. Sugie. A model-based sound localization system and its application to robot navigation. *Robotics and Autonomous System*, vol. 27(4), 1999;pp. 199–209.
- [2] D. Oertel, R. Fay, and A. Popper, eds. *Integrative Functions in the Mammalian Auditory Pathway*. Springer: New York, 2002.
- [3] E. Young and K. Davis. *Integrative Functions in the Mammalian Auditory Pathway*, chap. Circuitry and Function of the Dorsal Cochlear Nucleus. Springer: New York, 2002, pp. 160–206.
- [4] D. Fitzpatrick, S. Kuwada, and R. Batra. Transformations in processing interaural time differences between the superior olivary complex and inferior colliculus: Beyond the Jeffress model. *Hear. Res.*, vol. 168(1-2), 2002;pp. 79–89. ISSN 03785955.
- [5] L. Jeffress. A place theory of sound localization. *J. Comp. Physiol. Psychol.*, vol. 41, 1948;pp. 35–39.
- [6] P. Smith, P. Joris, and T. Yin. Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *J. Comp. Neurol.*, vol. 331, 1993;pp. 245–260.
- [7] J. A. Hirsch, J. C. Chan, and T. C. Yin. Responses of neurons in the cat's superior colliculus to acoustic stimuli. I. Monaural and binaural response properties. *J. Neurophysiol.*, vol. 53, 1985;pp. 726–745.
- [8] N. A. Bhadkamkar. Binaural source localizer chip using subthreshold analog CMOS. In *Proceedings of the 1994 IEEE International Conference on Neural Networks. Part 1 (of 7)*, vol. 3. IEEE, Orlando, FL, USA, 1994, pp. 1866–1870.
- [9] V. Willert, J. Eggert, J. Adamy, R. Stahl, and K. E. A probabilistic model for binaural sound localization. *IEEE Trans Syst Man Cybern Part B Cybern.*, vol. 36(5), 2006;pp. 982–994. ISSN 10834419.
- [10] T. Yin. Neural mechanisms of encoding binaural localization cues in the auditory brainstem. *Integrative Functions in the Mammalian Auditory Pathway*, 2002;pp. 99–159.
- [11] P. Joris. Envelope coding in the lateral superior olive. II. Characteristic delays and comparison with responses in the medial superior olive. *J. Neurophysiol.*, vol. 76, 1996;pp. 2137–2156.
- [12] P. Joris and T. Yin. Envelope coding in the lateral superior olive. III. Comparison with afferent pathways. *J. Neurophysiol.*, vol. 79, 1998;pp. 1253–1269.
- [13] W. Gerstner and W. M. Kistler. *Spiking Neuron Models, Single Neurons, Populations, Plasticity*. Cambridge University Press, 2002.
- [14] K. Voutsas and J. Adamy. A biologically inspired spiking neural network for sound source lateralization. *IEEE Trans Neural Networks*, vol. 18(6), 2007;pp. 1785–1799. ISSN 10459227.
- [15] R. Meddis, M. Hewitt, and T. Shackleton. Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse. *J. Acoust. Soc. Am.*, vol. 87(4), 1990;pp. 1813–1816.