Robot Sound Localisation Neural Network Inspired by the Inferior Colliculus (IC)

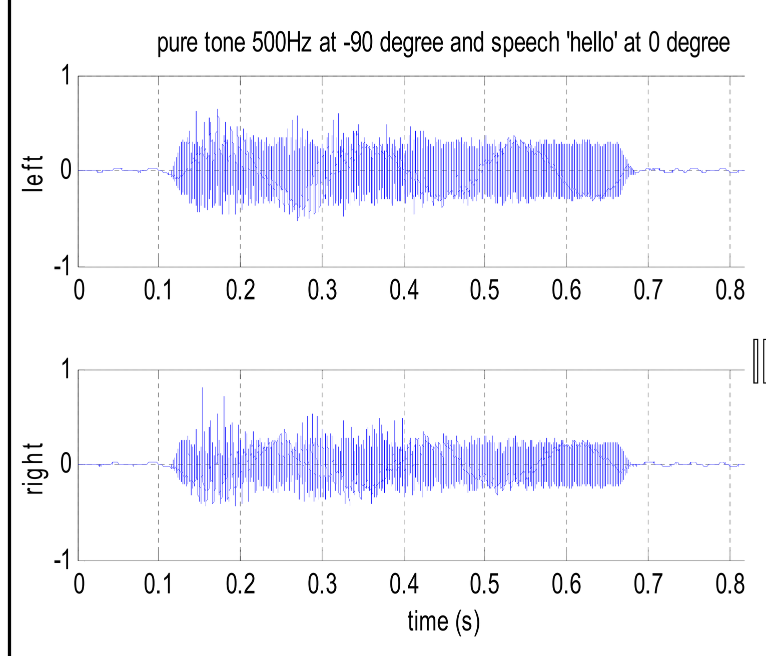
Jindong Liu[1], David Perez-Gonzalez[2], Adrian Rees[2] Harry Erwin[1] and Stefan Wermter[1]
(1)School of Computing and Technology, University of Sunderland, UK
(2)Institute of Neuroscience, The Medical School, Newcastle University, UK

University of Sunderland

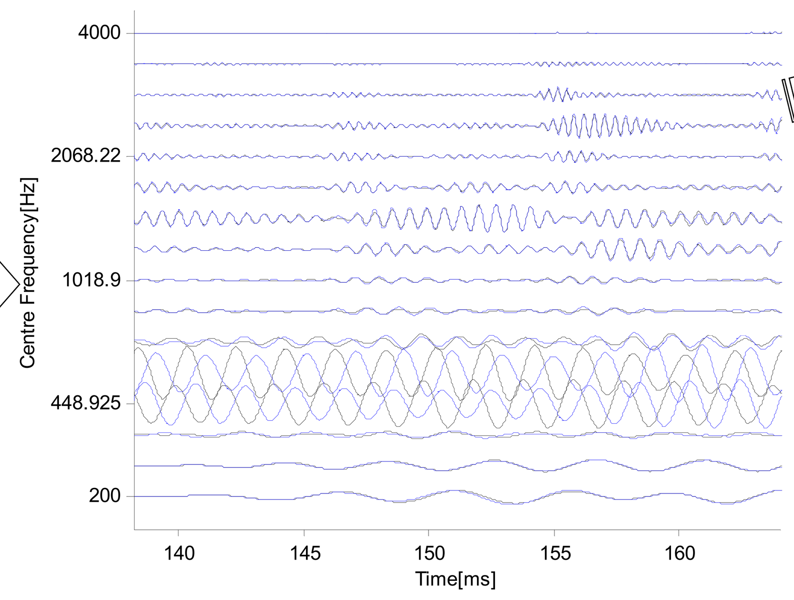
Introduction

- Our goal is to model the integration of auditory information in the inferior colliculus (IC) to aid mobile robots to work in noisy environments. The model uses information recorded by two microphones and consists of an interaural time difference (ITD) pathway and an interaural level difference (ILD) pathway. A Bayes model is used to estimate the sound source angle from the ITD and ILD information; these cues are then combined in the modelled IC.

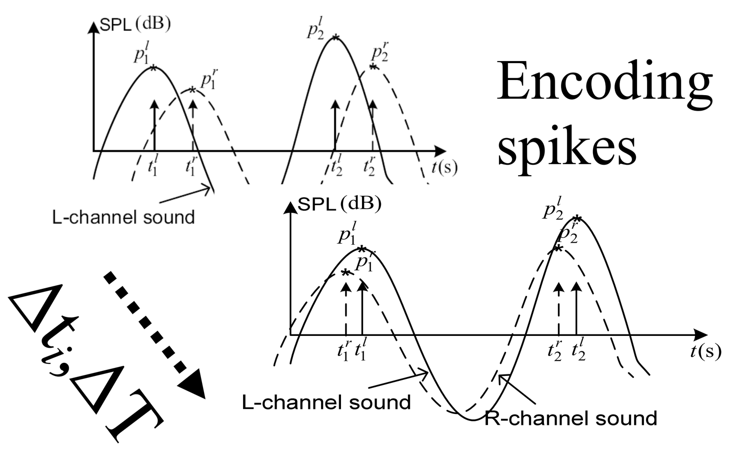
ITD & ILD pathway model



Raw sound (500Hz at –90 and 'hello' at 0 degree)



Gammatone filter bank



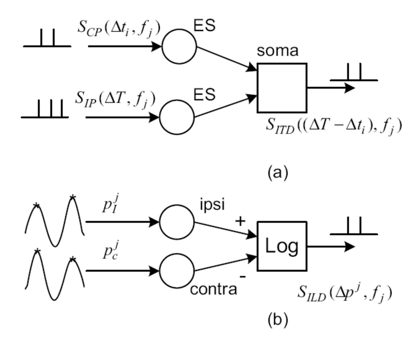
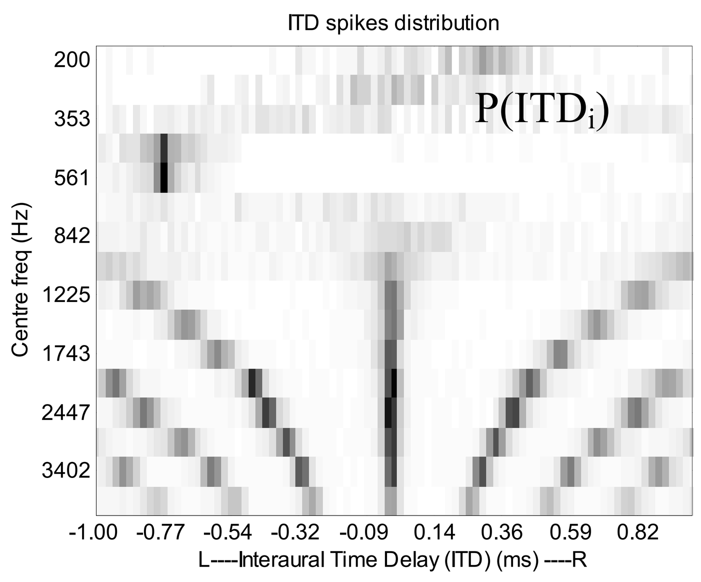
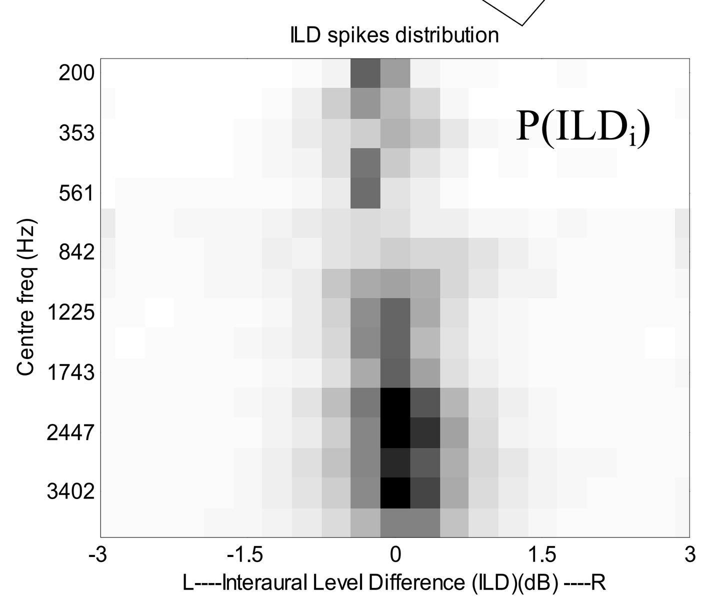


Fig. 5: (a) ITD coincidence model (b) ILD model. ES stands for excitatory synapse. Ipsi for ipsilateral and contra for contralateral Gammatone frequency channel.





Bayes model

Basic idea: Calculate Posterior Possibility P(θ|ITD) P(θ|ITD) from likelihood P(ITD|θ) P(ILD|θ) and prior possibility P(θ). Then update the P(θ) using posterior possibility and biological IC evidence.

Step 1:
Offline calculate the likelihood P(ITD|θ) and P(ILD|θ) using broadband noise sample from known directions for each centre frequency. The initial prior possibility of P(θ) is set to an uniform distribution.

Step 2:
Calculate posterior possibility using Bayes model:

$$P(\theta \mid ITD) = \frac{P(ITD \mid \theta)}{\sum_i P(ITD \mid \theta_i) * P(\theta_i)}$$ And

$$P(\theta \mid ILD) = \frac{P(ILD \mid \theta)}{\sum_i P(ILD \mid \theta_i) * P(\theta_i)}$$

Step 3:
Then update the $P_{ITD}(\theta)$ $P_{ILD}(\theta)$ using results from ITD/ILD pathway model:

$$P_{ITD}(\theta) = \sum_i P(\theta \mid ITD_i) * P(ITD_i)$$

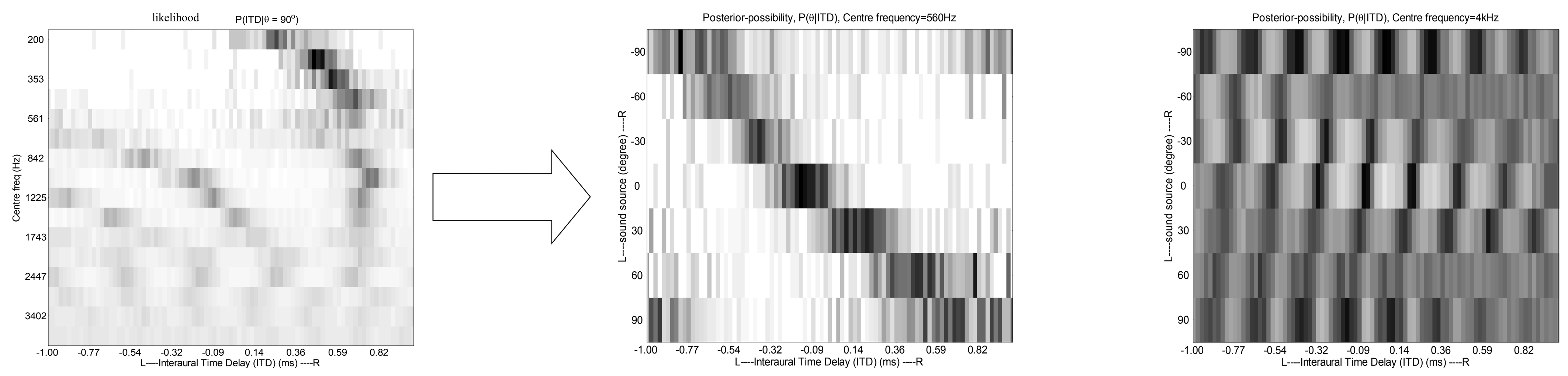
$$P_{ILD}(\theta) = \sum_i P(\theta \mid ILD_i) * P(ILD_i)$$ and

Step 4:
Combine ITD cue and ILD cue to get updated P(θ). Inspired by IC.
–For f<1.2kHz $P(\theta) = P_{ITD}(\theta)$
–For f>1.2kHz $P(\theta) = \sqrt{P_{ITD}(\theta) * P_{ILD}(\theta)}$
Then back to **step 2.**

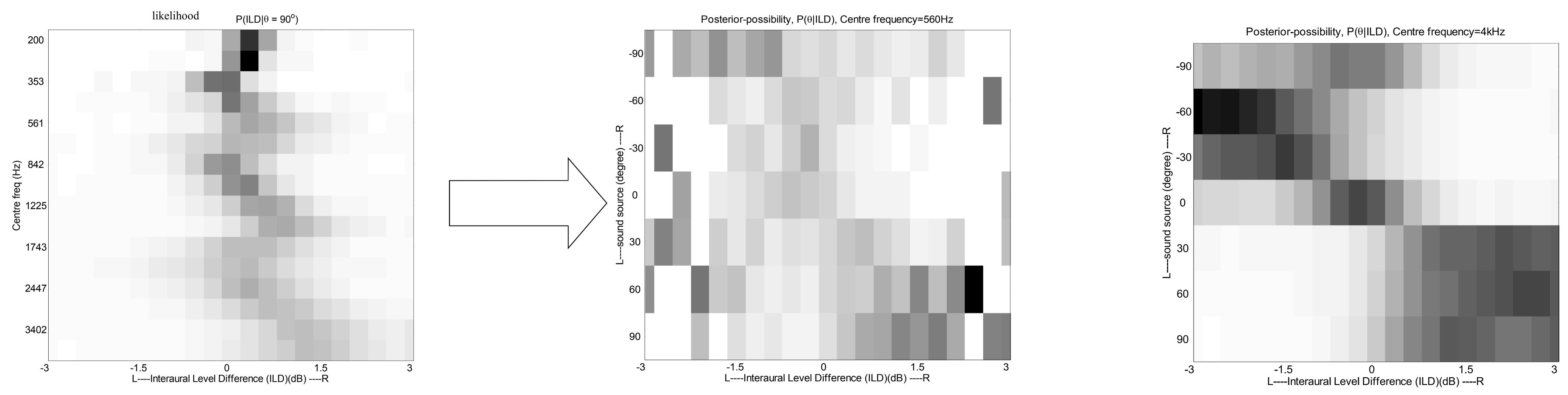
The following figures show the calculation in step 1 and step 2. The first column shows the likelihood of P(ITD|θ=90°) P(ILD|θ=90°). The black area shows high possibility. The middle and right columns show the posterior possibility for a low centre frequency 560Hz and high centre frequency 4kHz. These figures match the biological evidence, i.e. ITD is efficient in low frequency sound localisation while ILD is more efficient for high frequency sound.

Freq=560Hz Freq=4kHz

ITD



ILD



Experimental results (TWO sound sources, 500Hz & 'Hello')

The right three figures show the calculation from step3 to step4. The sound sample is a two sound mixture: a 500Hz pure tone at –90 degree and a human speech 'hello' at 0 degree. It shows that a hybrid model of ITD and ILD is better than a monolithic ITD or ILD.

Future work will focus on the function of cells in IC for sound localisation and identification. Four types of IC cells will be explored initially :
Sustained-Regular, rebound-regular, onset and pause/build.