

Robot Sound Localisation Neural Network Inspired by the Inferior Colliculus

Jindong Liu¹, David Perez-Gonzalez², Adrian Rees² Harry Erwin¹, Stefan Wermter¹

(1) School of Computing and Technology, University of Sunderland, SR6 0DD

(2) Institute of Neuroscience, The Medical School, Newcastle University, NE2 4HH

Modelling sound source localisation (SSL) for mobile robots is important for the development of human-machine interfaces and sensor-motor robot control. Present technologies for SSL, such as the cross-correlation method or the multiple-microphone array schema, have limitations when the robot operates in a noisy environment. The mammalian auditory system provides an efficient SSL system by using a pair of “microphones” that feed information into the system. The inferior colliculus (IC), the midbrain nucleus of the auditory pathway is a centre of convergence and integration of several brainstem pathways, including those devoted to the sound localisation. In the MiCRAM project we are studying the functional structure of the IC and applying it to a mobile robot in order to improve its SSL performance in a cluttered environment.

Our system models two ascending pathways to the IC using a spiking neural network: the ITD (Interaural Time Difference) and ILD (Interaural Level Difference) pathways. The calculations resulting from ITD and ILD are finally merged in an IC model to take advantage of ITD on the low frequency spectrum and of ILD on the high frequency spectrum. The structure and computational procedures of the system are briefly introduced as follows:

- To simulate the cochlea, the sound from two microphones placed on the sides of the robot head is filtered by a Gammatone filterbank and split into a number of frequency channels. Each channel is then encoded into a spike train. The mechanism of encoding is based on biological evidence of phase-locked spikes feeding the MSO (medial superior olive).
- In modelling the ITD pathway, we assume that no delay line exists in the ITD processing for the ipsilateral ear, while multiple delay lines exist for the contralateral ear. The phase-locked spike trains from both ears follow the delay line structure, and feed into a spiking neural network which uses a leaky integrate-and-fire model to simulate the sustained-regular cells that simulate ITD processing.
- ILD spikes are calculated directly from the Gammatone filterbank output based on the logarithmic ratio of sound level at both ears.
- The ITD and ILD spikes are counted along both frequency and ITD/ILD channels. The azimuth angle of the sound source can be estimated using Bayes' theorem or Dempster-Shafer theory.

Experimental results on pure tone, click, noise and voice show that our model performs sound localisation that approaches biological performance. Our approach demonstrates a practical application of biologically based sound localisation for robots.

Acknowledgements

Work supported by EPSRC.