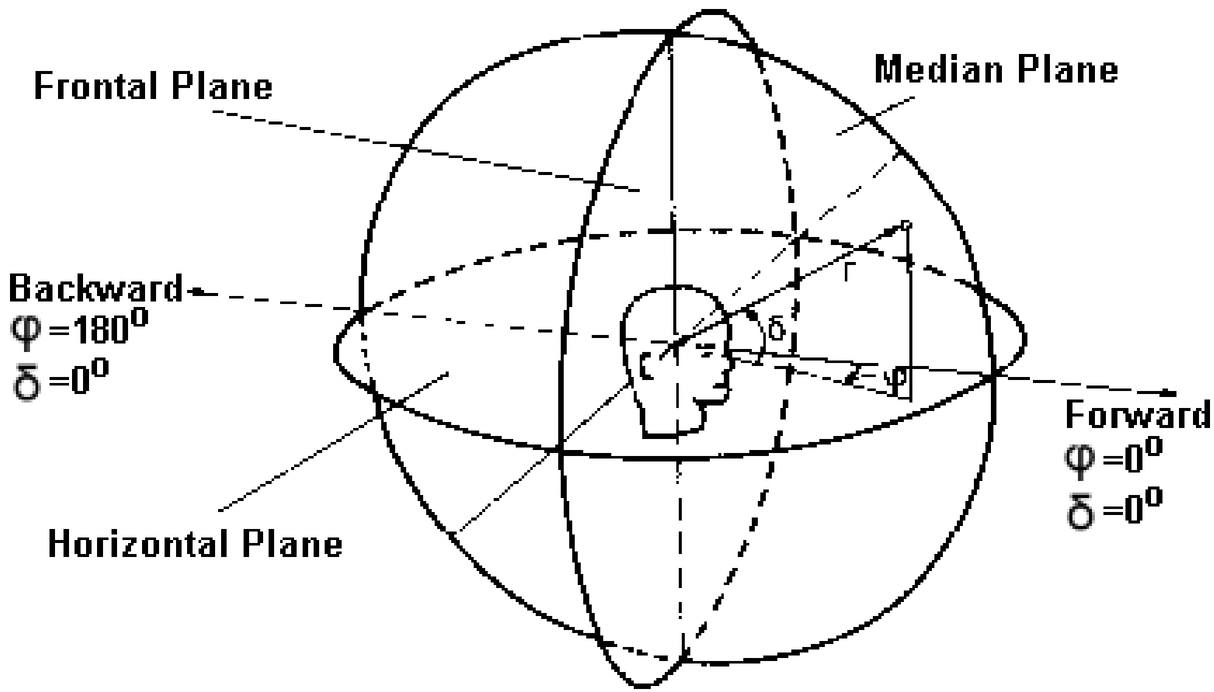
Bio-Inspired Artificial Intelligence

Lecture 10: Bioinspired Robotic Sound Localization

http://www.informatik.uni-hamburg.de/WTM/

Localisation of sound sources in biological systems

- Acoustic element or sound traverses the environment (Blauert 1997)
- Localisation of sound source has two coordinates ***azimuth*** ‘φ’ and ***elevation*** ‘δ’ in general



Interaural acoustic cues

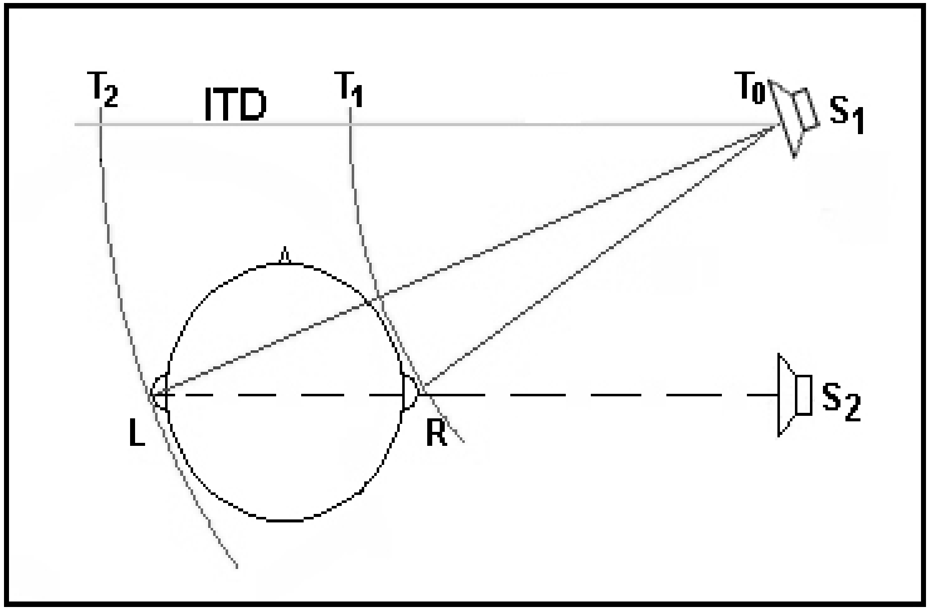
- To estimate azimuth specific auditory cues are used
- ***Interaural Time Difference*** (ITD) – Time difference between arrival of the signal at the two ears

T_0 - Sound begins

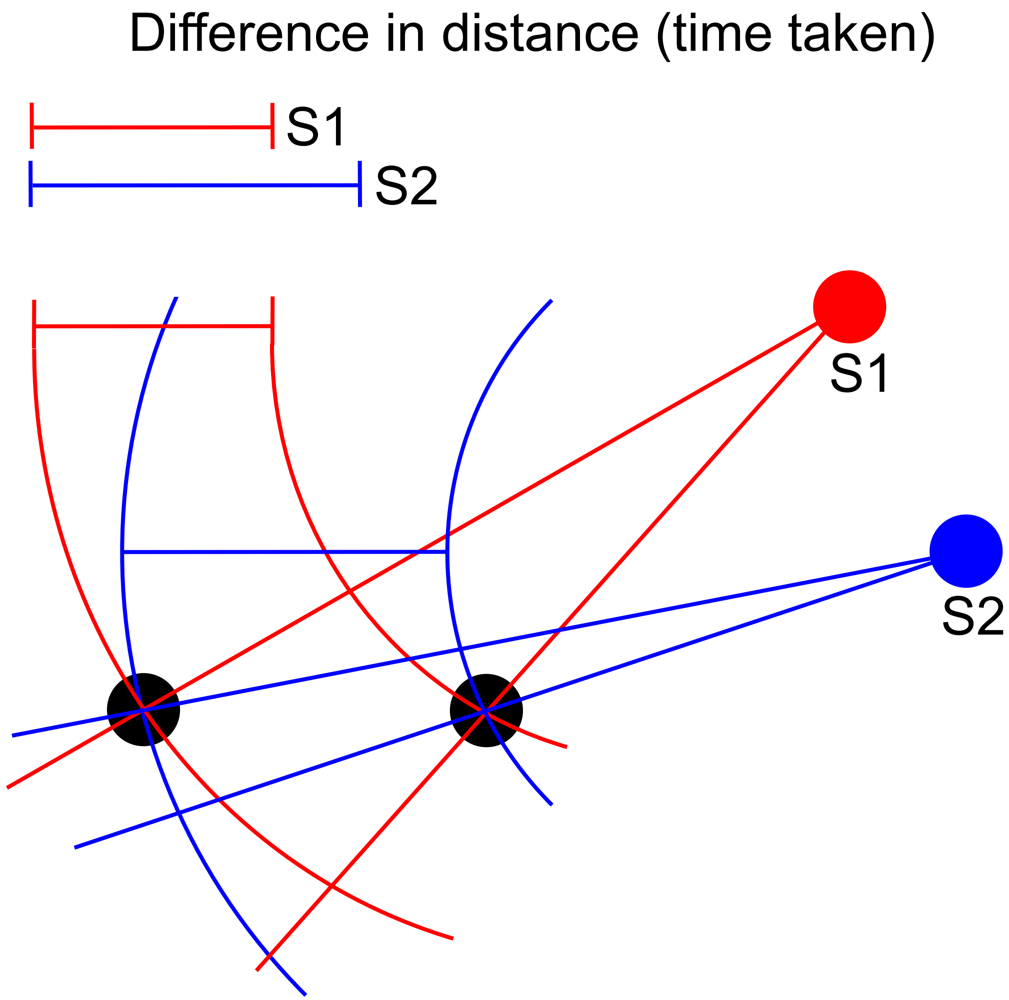
T_1 - Sound reaches ‘R’

T_2 - Sound reaches ‘L’

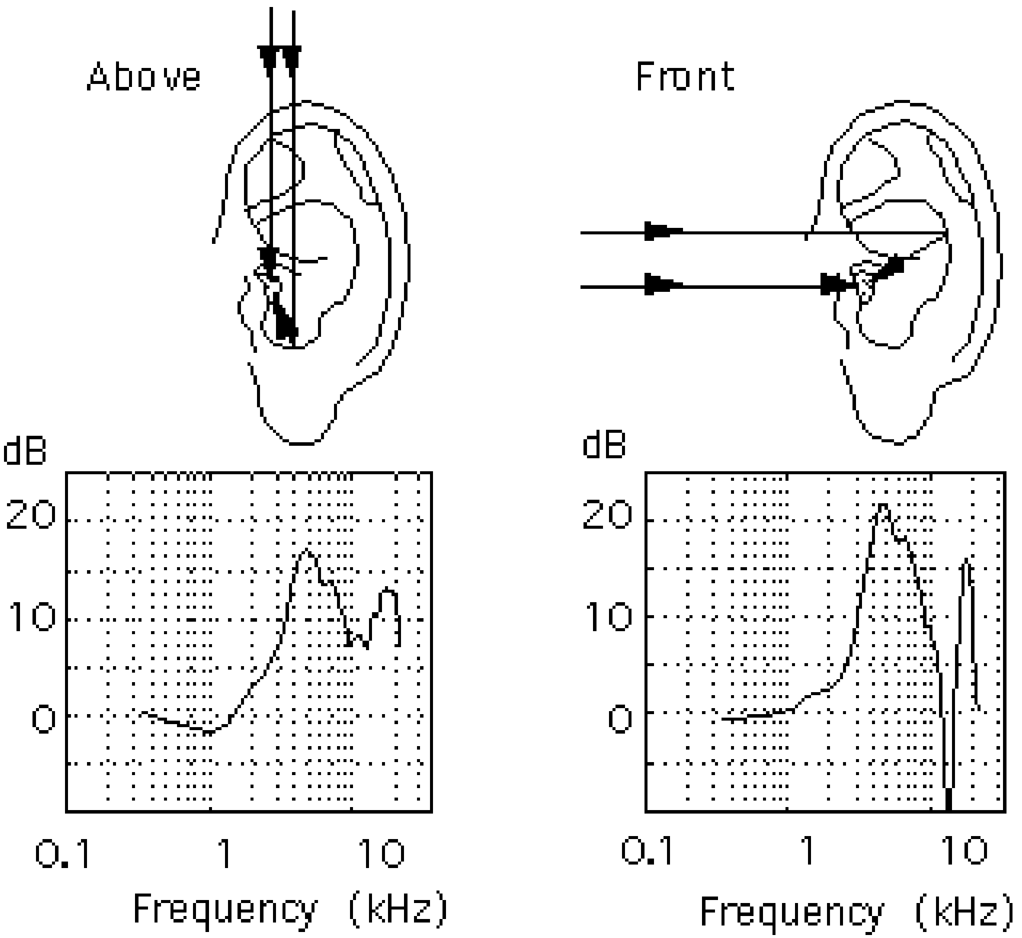
$ITD = T_2 – T_1$



Difference in azimuth changes time delay of signal arriving at the receivers



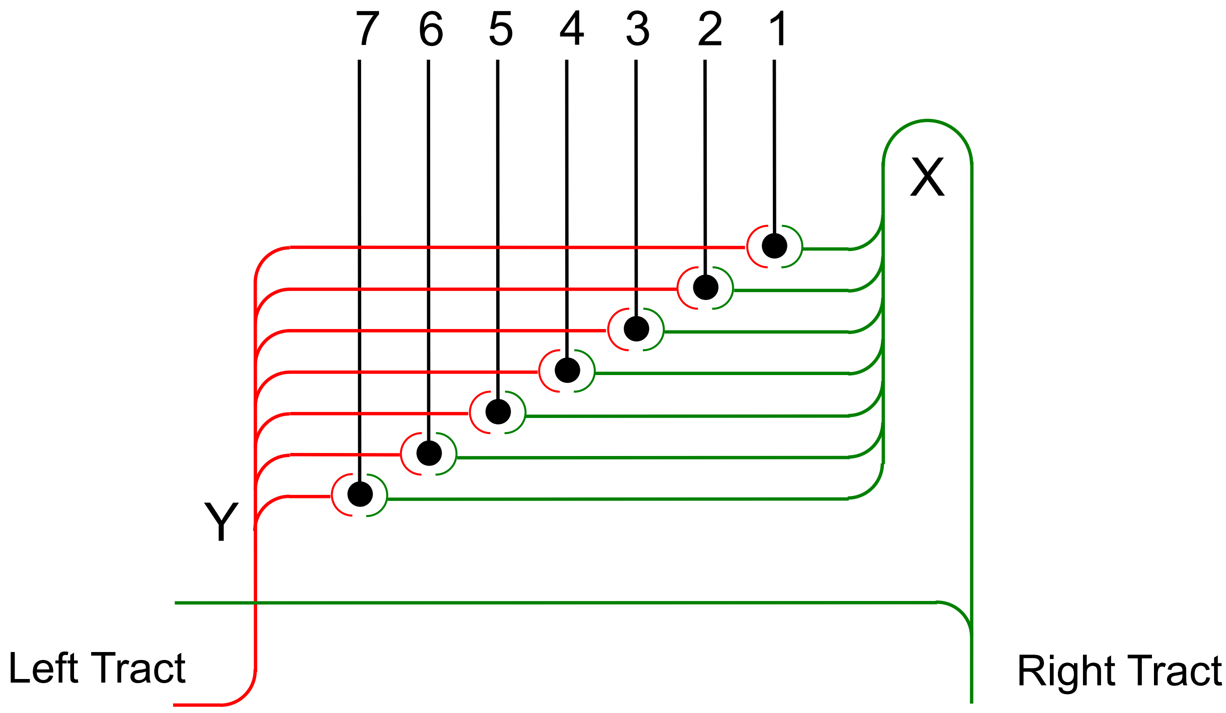
Difference in elevation changes also lead to different frequency responses (Duda 2000)



Early Jeffress neuronal coincidence array (1948)

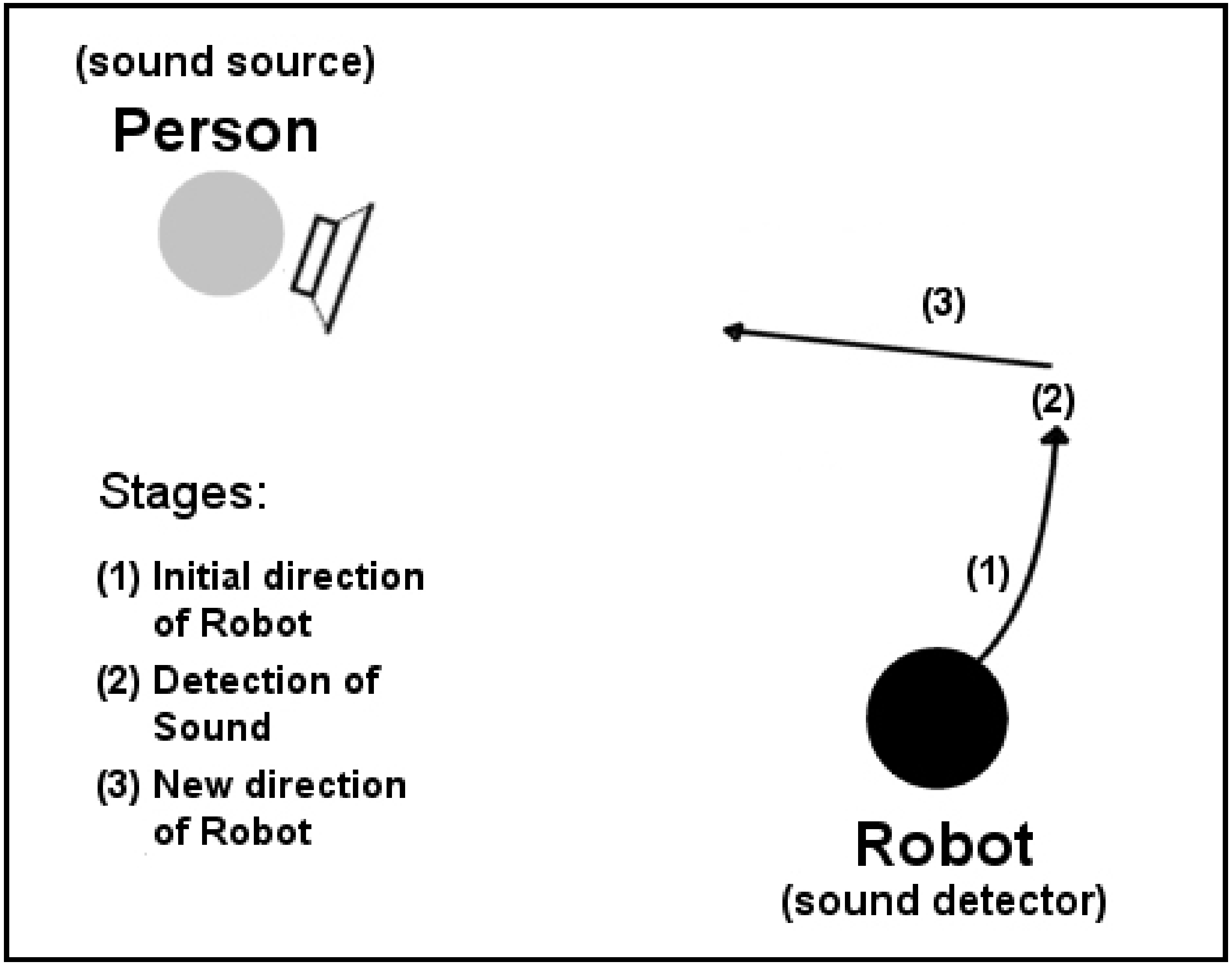
Fibres to X and Y are of equal length but branches vary in length

Therefore different delay for connected neurons 1-7



Acoustic tracking

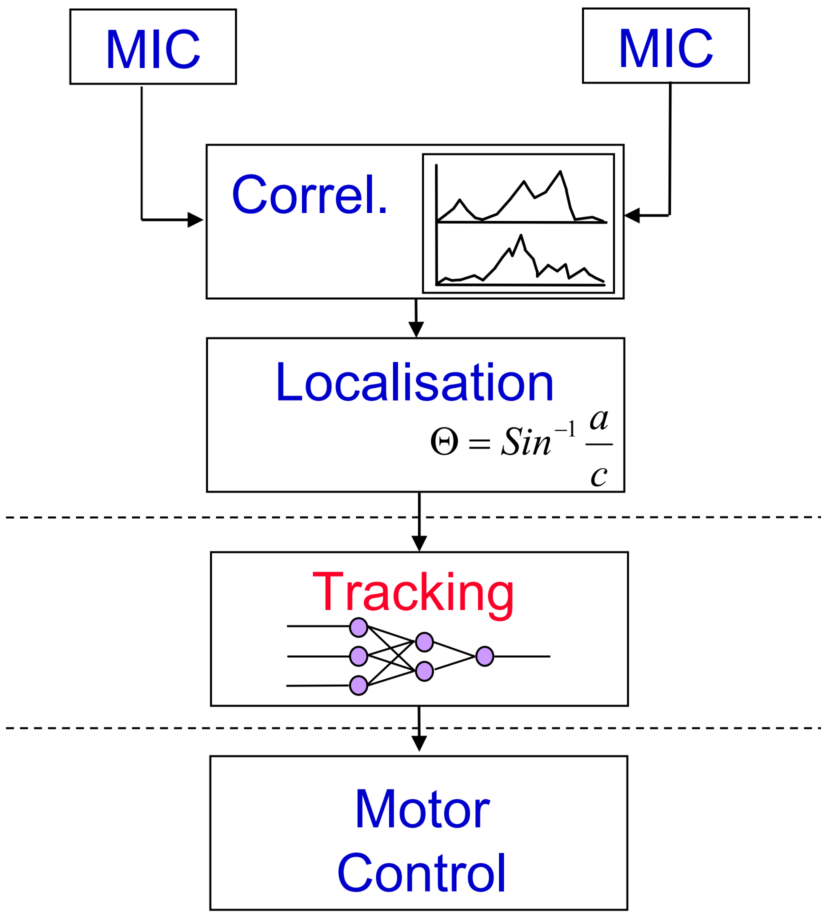
- Robot should detect and track sound source of interest



Hybrid acoustic tracking

- Inspired by some properties of mammalian system (2 ear model, Interaural Time Difference (ITD))
- Well understood ***cross correlation*** algorithm for sound localization
- Neural: **recurrent neural networks** for tracking sound source to improve the localization task
- Learning and adaptation to acceleration and deceleration

Hybrid acoustic localisation and tracking



Digital representation of detected signals

Angle position representing location of source

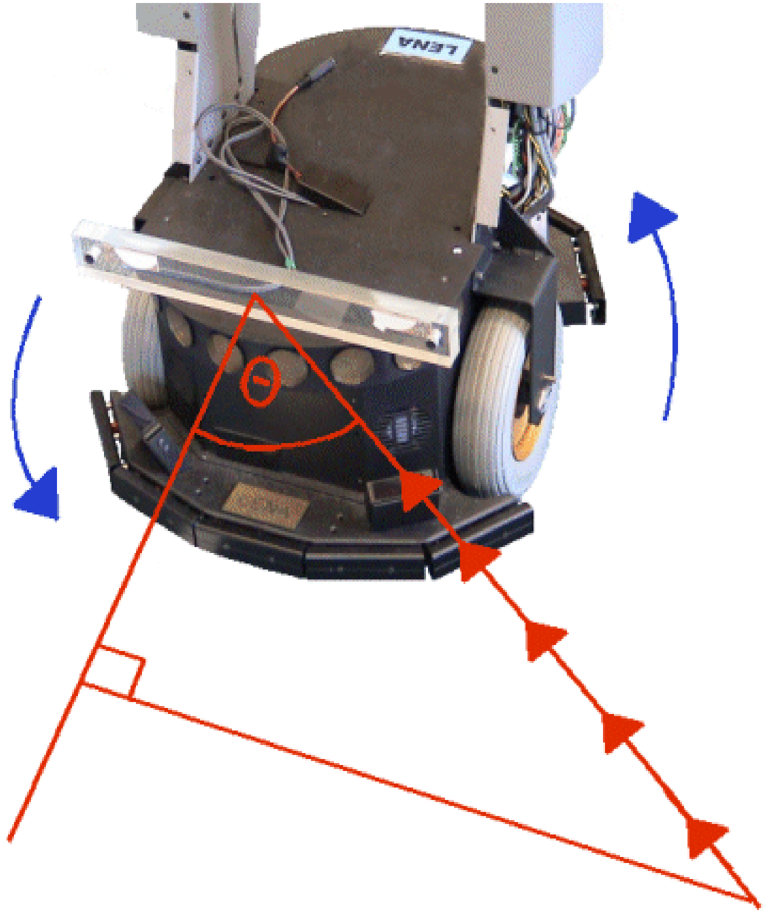
Next predicted position/angle of source

Experimental Setup

- Two microphones 30cm apart

- Sound processed by cross-correlation to compute ITD and actual angle of source position

- Recurrent neural network to improve tracking

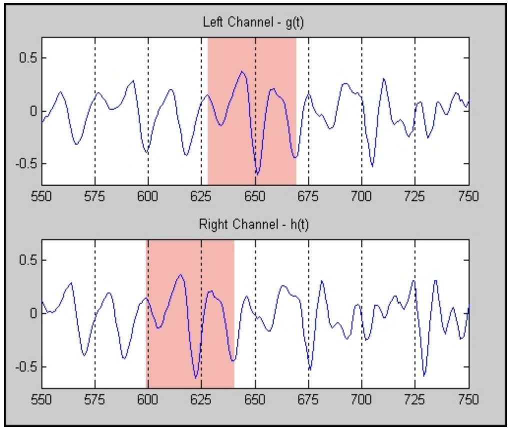


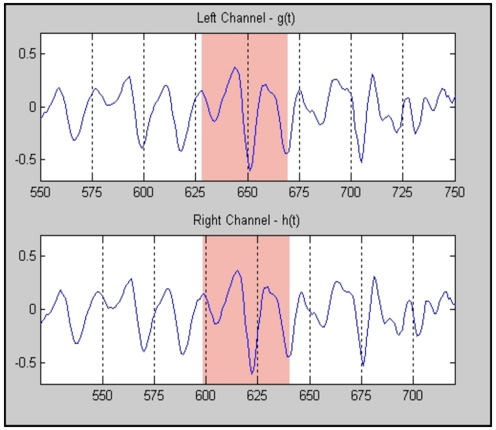
Cross correlation for localisation

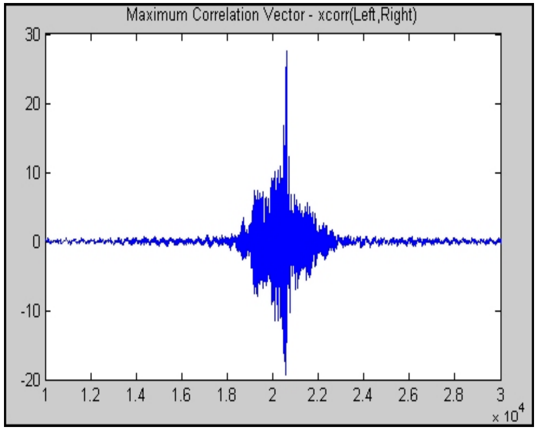
- Determines ***maximum similarity*** between two signals *g(t)* & *h(t) e.g.*

$$\begin{bmatrix} g(t) \\ h(t) \end{bmatrix} = \begin{bmatrix} 1111232111 \\ 1111112321 \end{bmatrix}$$

- Correlation vector represents ITD delay between signals
- Allows then to determine the angle of incidence of the source







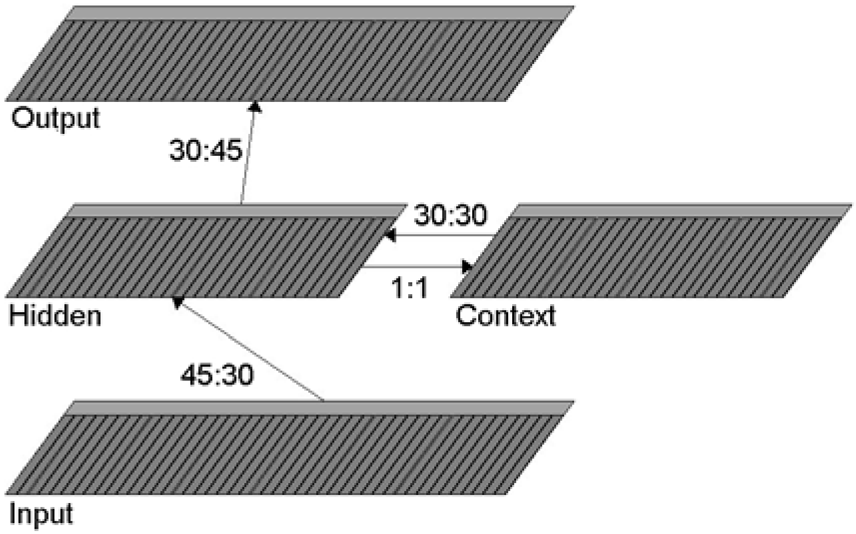
Cross correlation value creation

Correlation Vector Element *N*	Correlation of Signals *g(t)* and *h(t)*	Correlation Vector Element *N* Value
1	1111232111000000000 0000000001111112321	1
2	111123211100000000 000000001111112321	2
3	11112321110000000 00000001111112321	3
4	1111232111000000 0000001111112321	5
…..	…..	…..
11	01111232111 11111123210	21
12	001111232111 111111232100	**22**
13	0001111232111 1111112321000	19
…..	…..	…..
17	00000001111232111 11111123210000000	6
18	000000001111232111 111111232100000000	3
19	0000000001111232111 1111112321000000000	1

Simple recurrent network for prediction

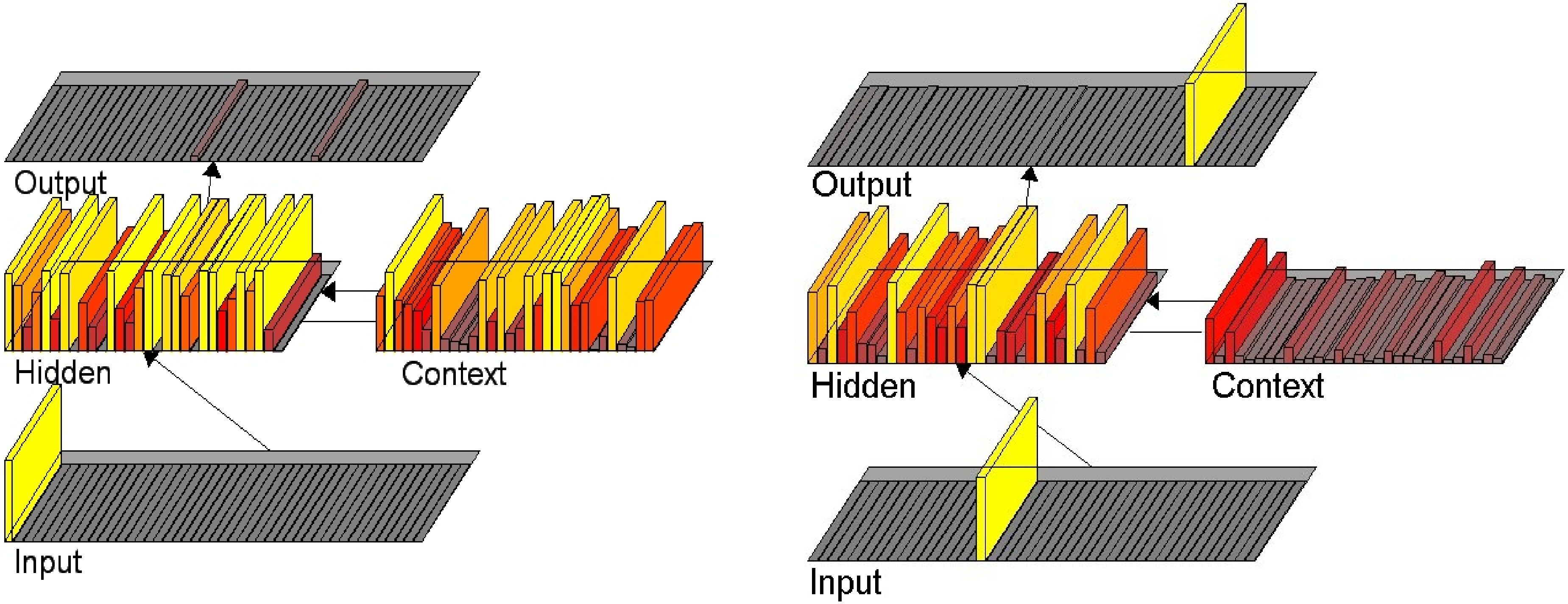
- Predictor for next angle in trajectory of source

next angle



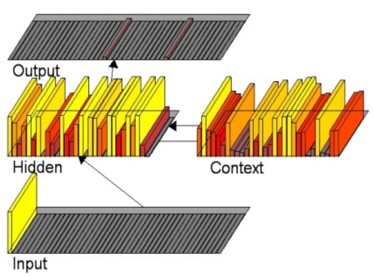
current angle

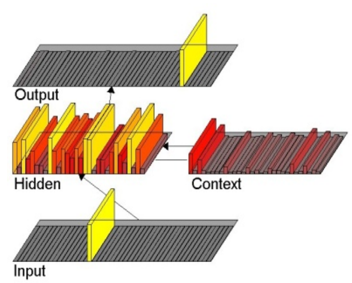
Training a network to predict over time with different (fixed) speed



Two events of first five speed sequences with expected output activations for 2° increments

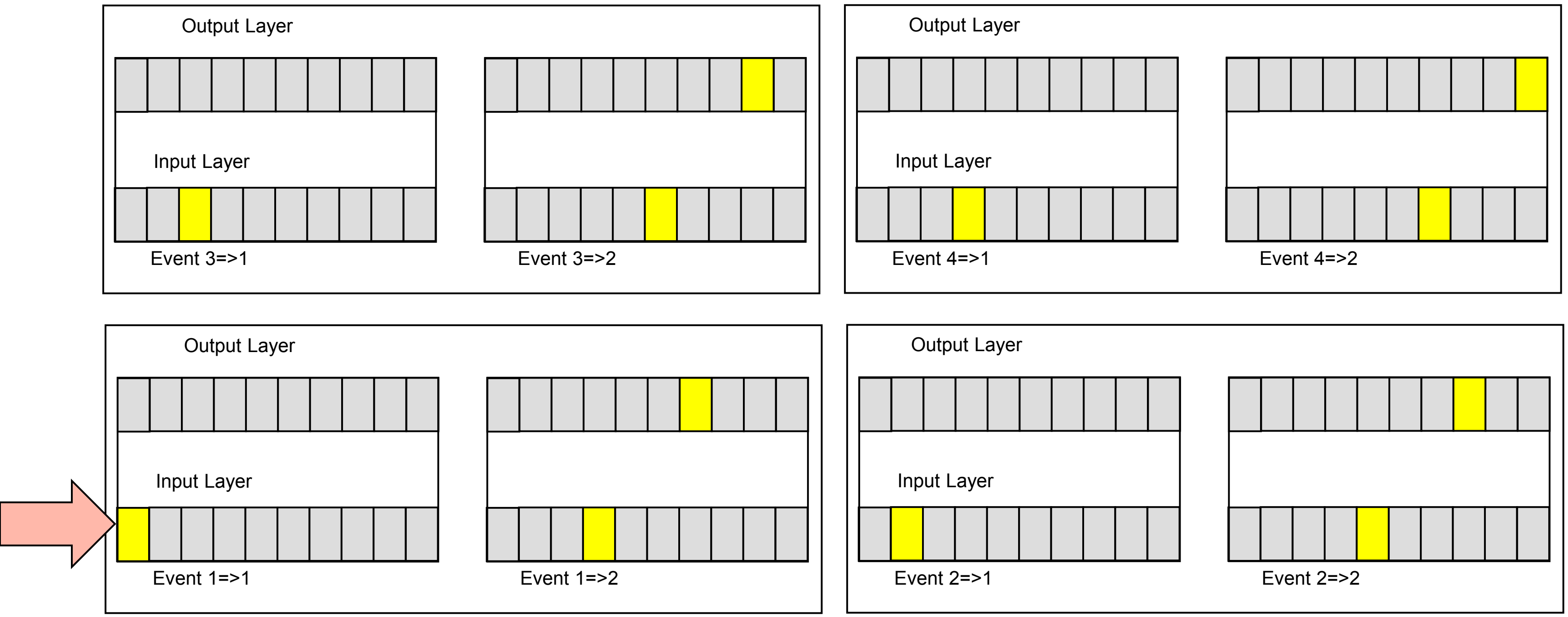
		Input Sequence										Expected Output t_{n+2}										
Speed	**Time**	0°	2°	4°	6°	8°	10°	12°	14°	16°	18°	0°	2°	4°	6°	8°	10°	12°	14°	16°	18°	20°
1	t_{n+0}																					
	t_{n+1}																					
2	t_{n+0}																					
	t_{n+1}																					
3	t_{n+0}																					
	t_{n+1}																					
4	t_{n+0}																					
	t_{n+1}																					
5	t_{n+0}																					
	t_{n+1}																					



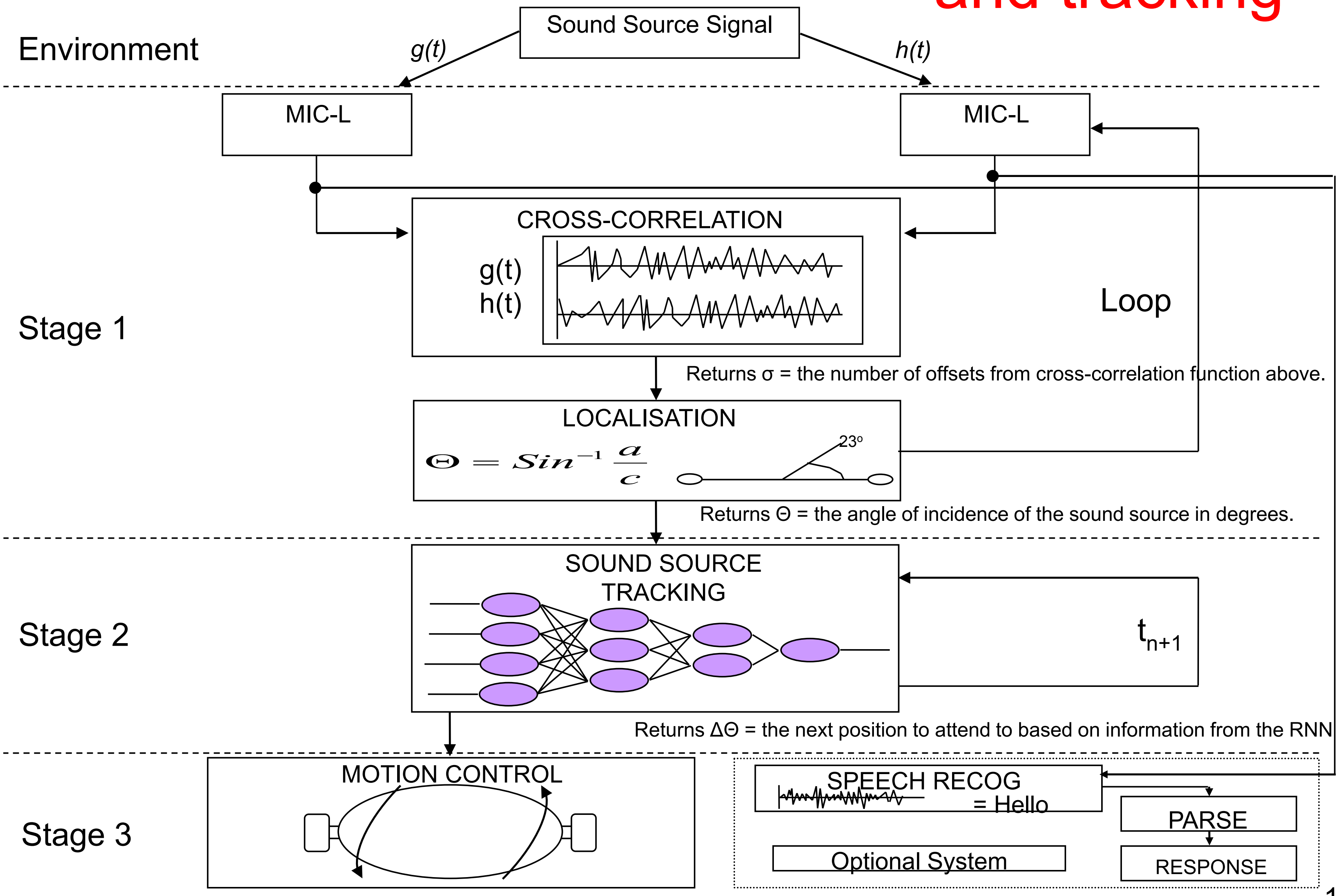


First 16 events for speed 1 and 2° increments

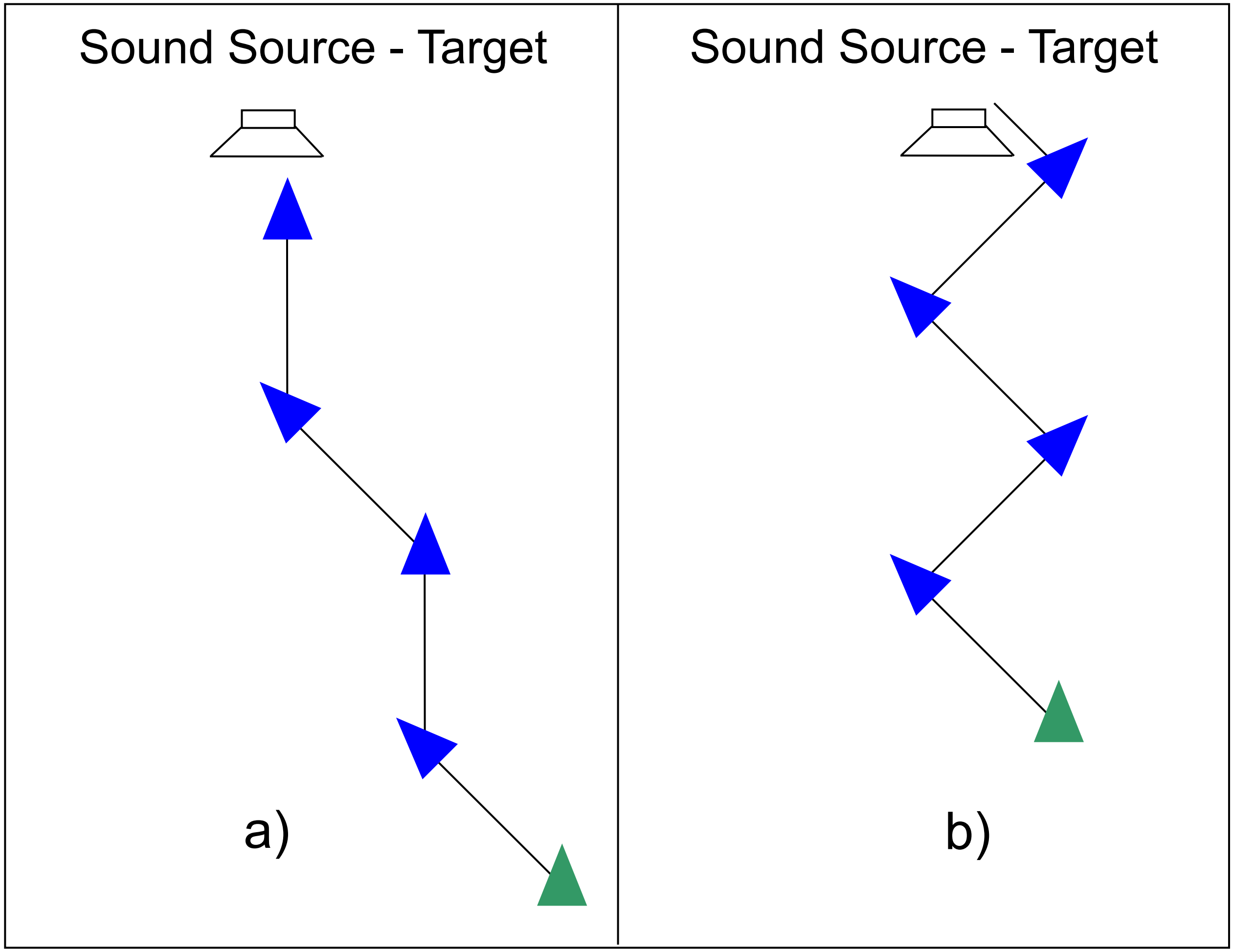
First eight event patterns for speed 3, i.e. 6° increments



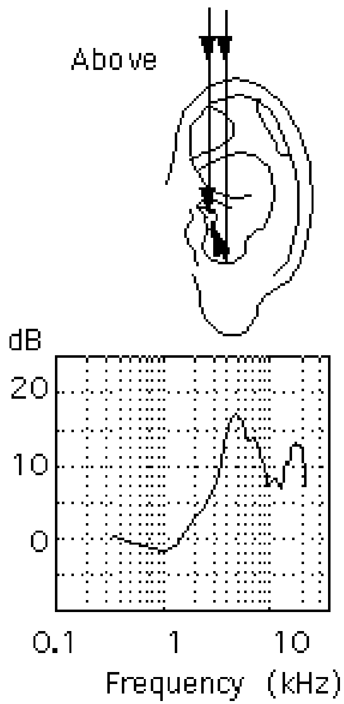
Hybrid localisation and tracking

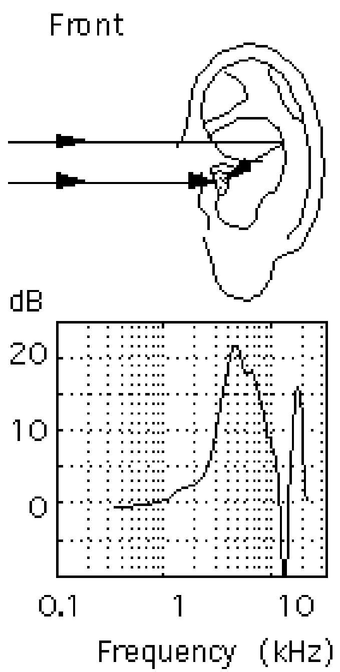


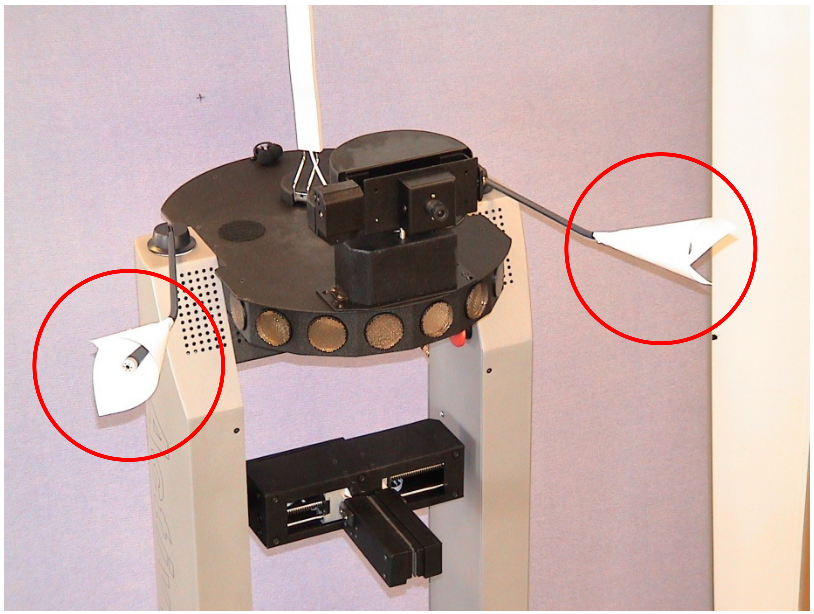
Two example solutions for localisation model



Variations: natural findings into bioinspired robotics







Why hybrid localisation and tracking?

- ***Algorithmic sound source localisation*** well understood algorithmically
- Cross-correlation does not require training to provide azimuth angle
- **Neural predicting** of source enables a quicker response and can learn temporal sequences

Acoustic Tracking

Auditory world 1: fleeing rabbit

fleeing rabbit [Ref: Marathon Intel.]

Cocktail Party Problem

- We and animals have astonishing abilities of sound localisation and sound perception in auditory cluttered environment

Biological inspiration:

How does the human auditory system do the job?

How can we use it for mobile robot sound perception?

Cocktail party effect
[Ref: Pride and prejudice (1995)]

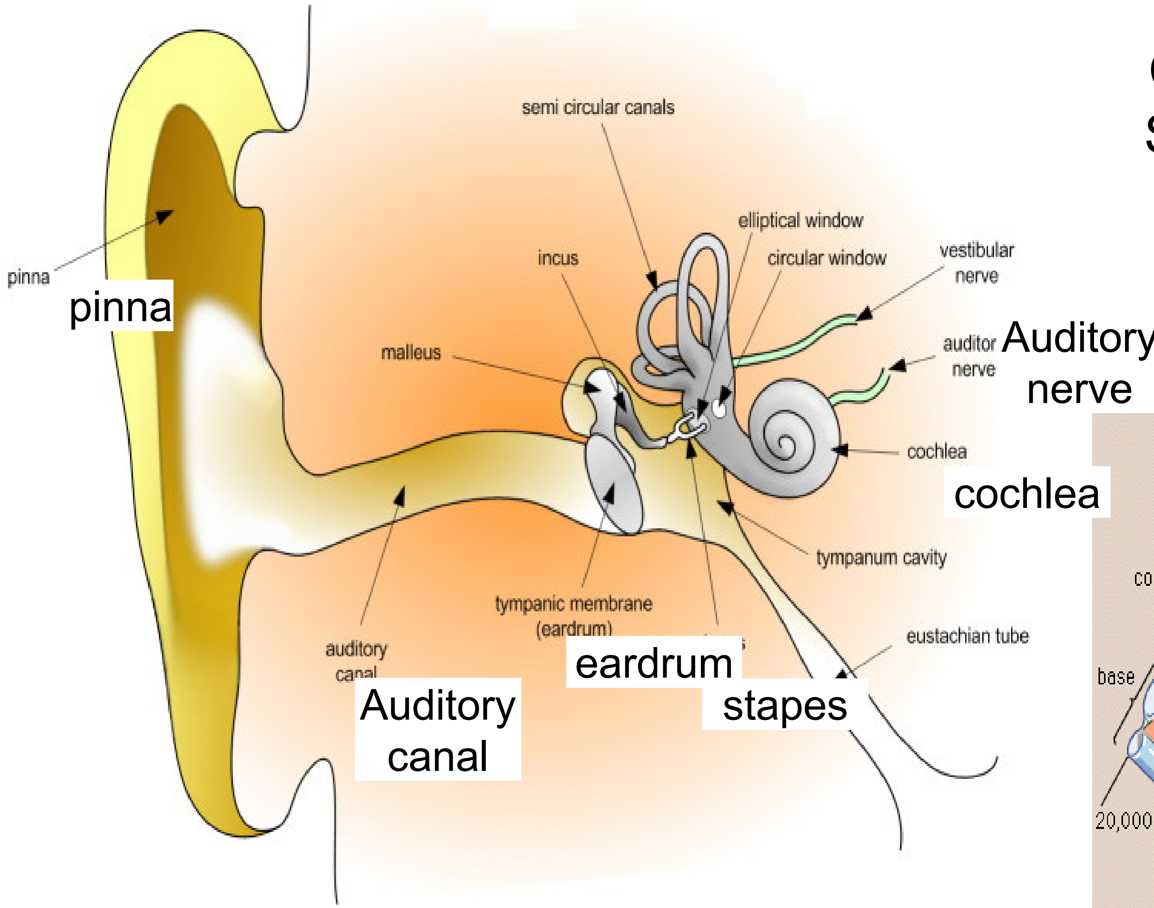
What can we gain from cognitive neuroscience for robot sound localisation?

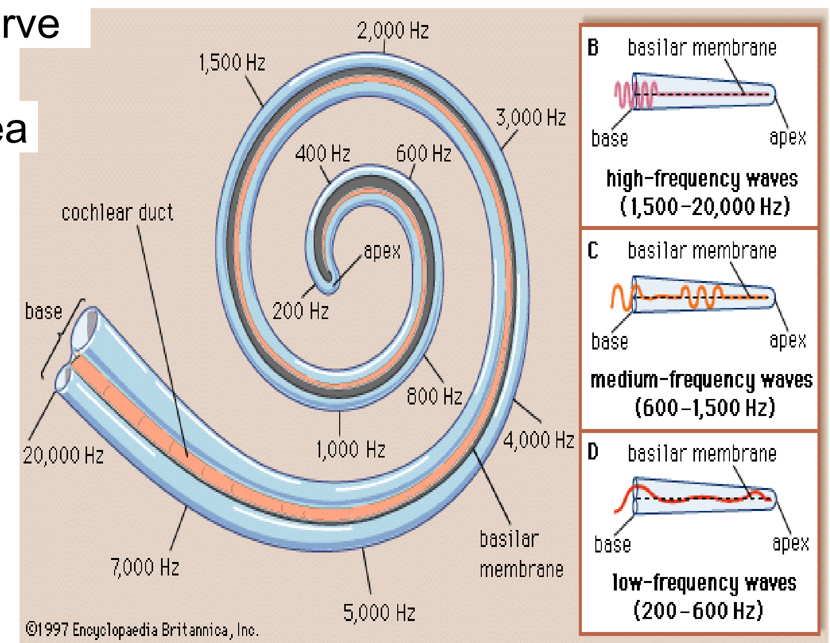
- How is sound encoded?
  - Ear pinna->middle ear->inner ear>auditory nerve
  - ***Tonotopic representation***

- How is the encoded information processed?
  - Spiking neural networks
  - ***Interaural time difference*** (ITD)
    in Medial Superior Olive (MSO)
  - ***Interaural level difference*** (ILD)
    in Lateral Superior Olive (LS0)
  - Integrated in the Inferior Colliculus (IC)

“Our microphones” for encoding sound

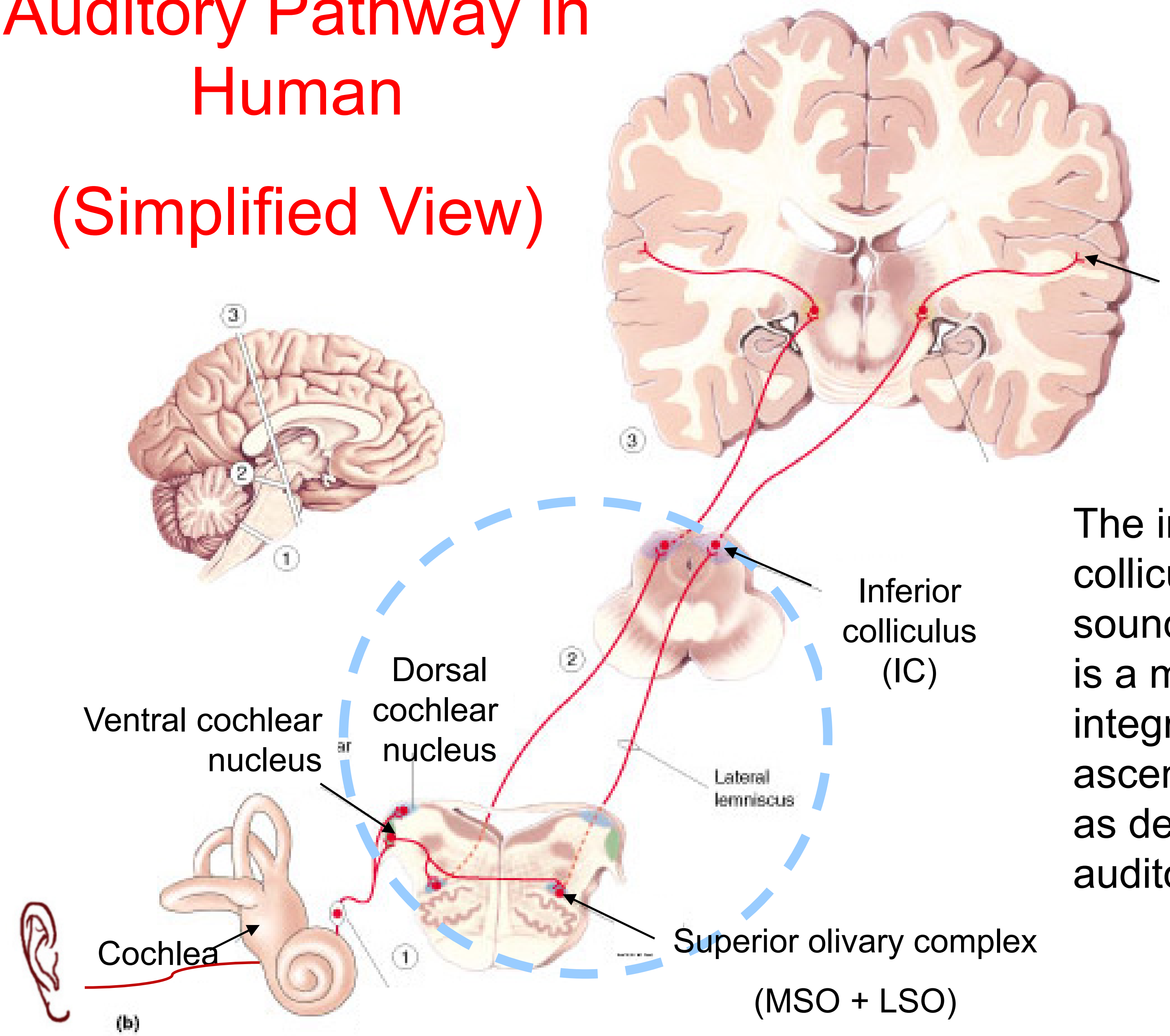
Cochlea: tonotopical
Sound repesentation





Auditory Pathway in Human

(Simplified View)

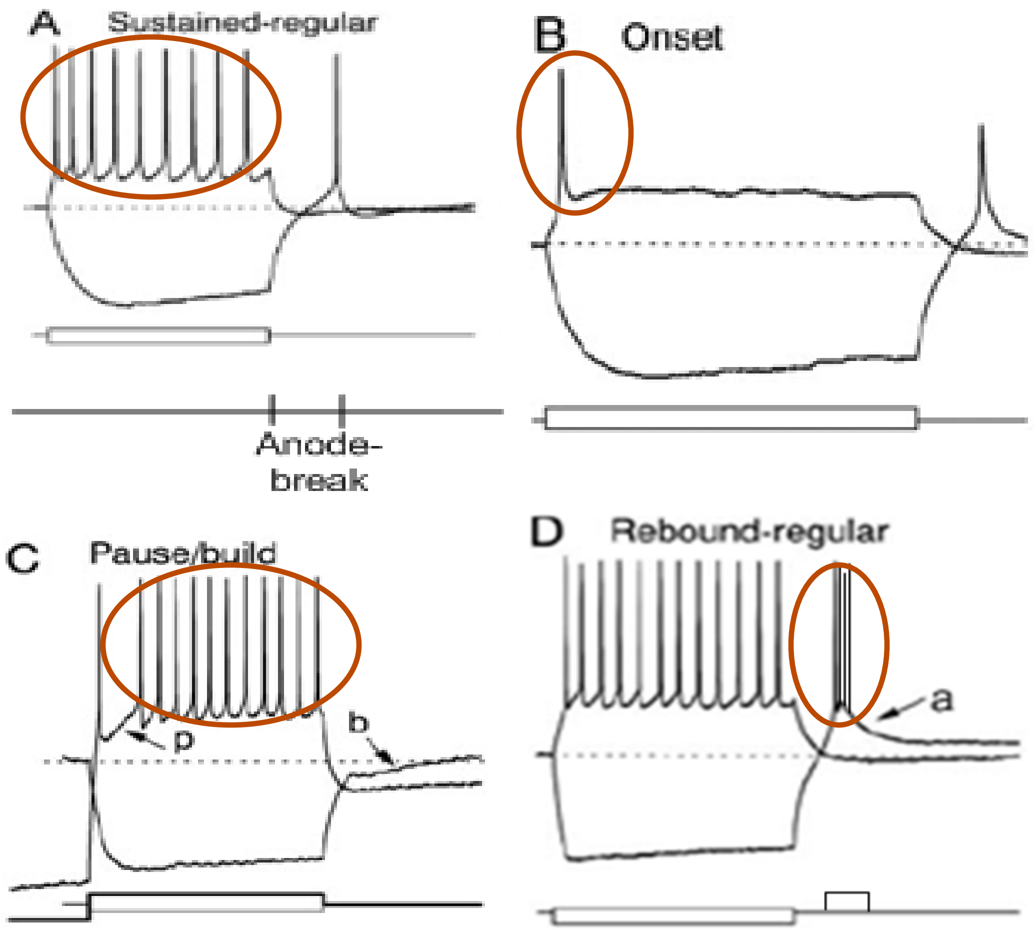


The inferior colliculus (IC) for sound localisation is a major centre of integration in the ascending as well as descending auditory pathways.

Four typical IC cells for modeling

Sustained
"I am hearing
something"

Onset
"I heard
something"



Pause-Build
"I heard
something
after a delay"

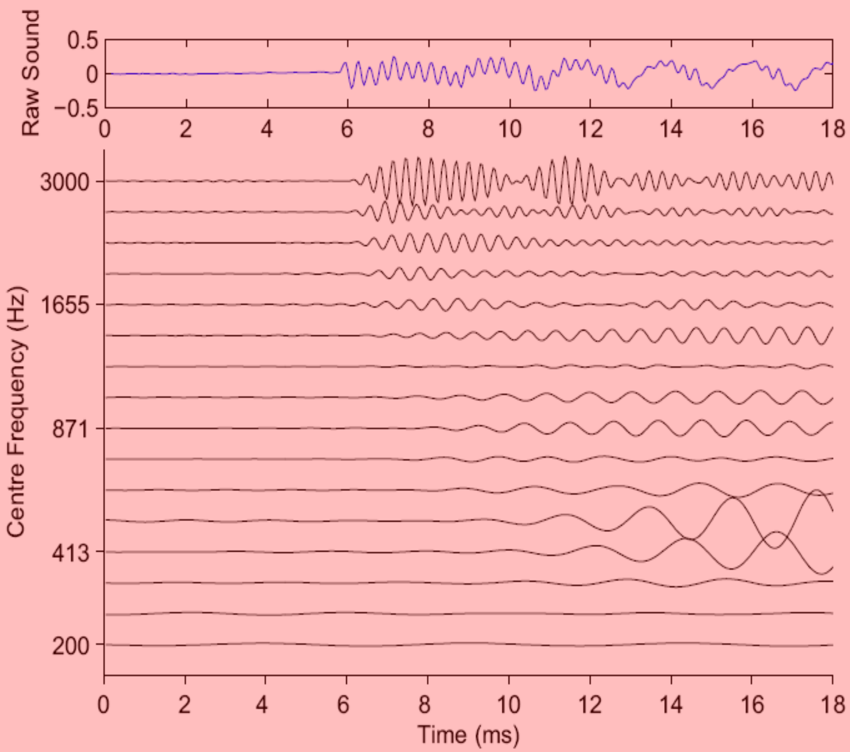
Rebound
"I expect to hear
something after
a delay"

p: pause,
b: build-up (inactive state)

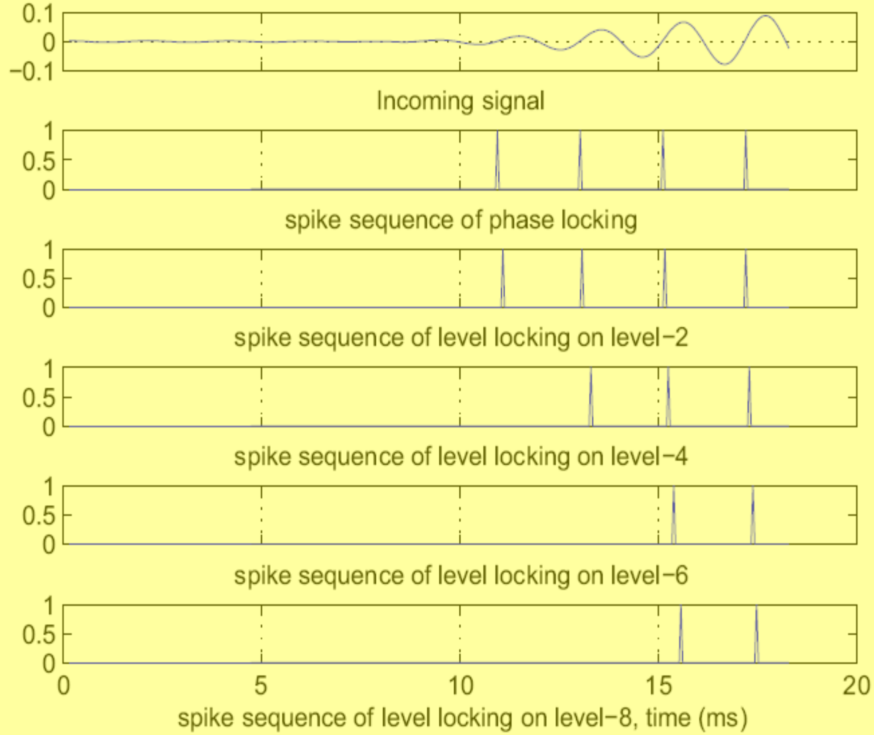
a: calcium rebound

Ref: Journal of
Neuroscience, 2001,
21(8):2861–2877

Sound encoding: From sounds to spike trains

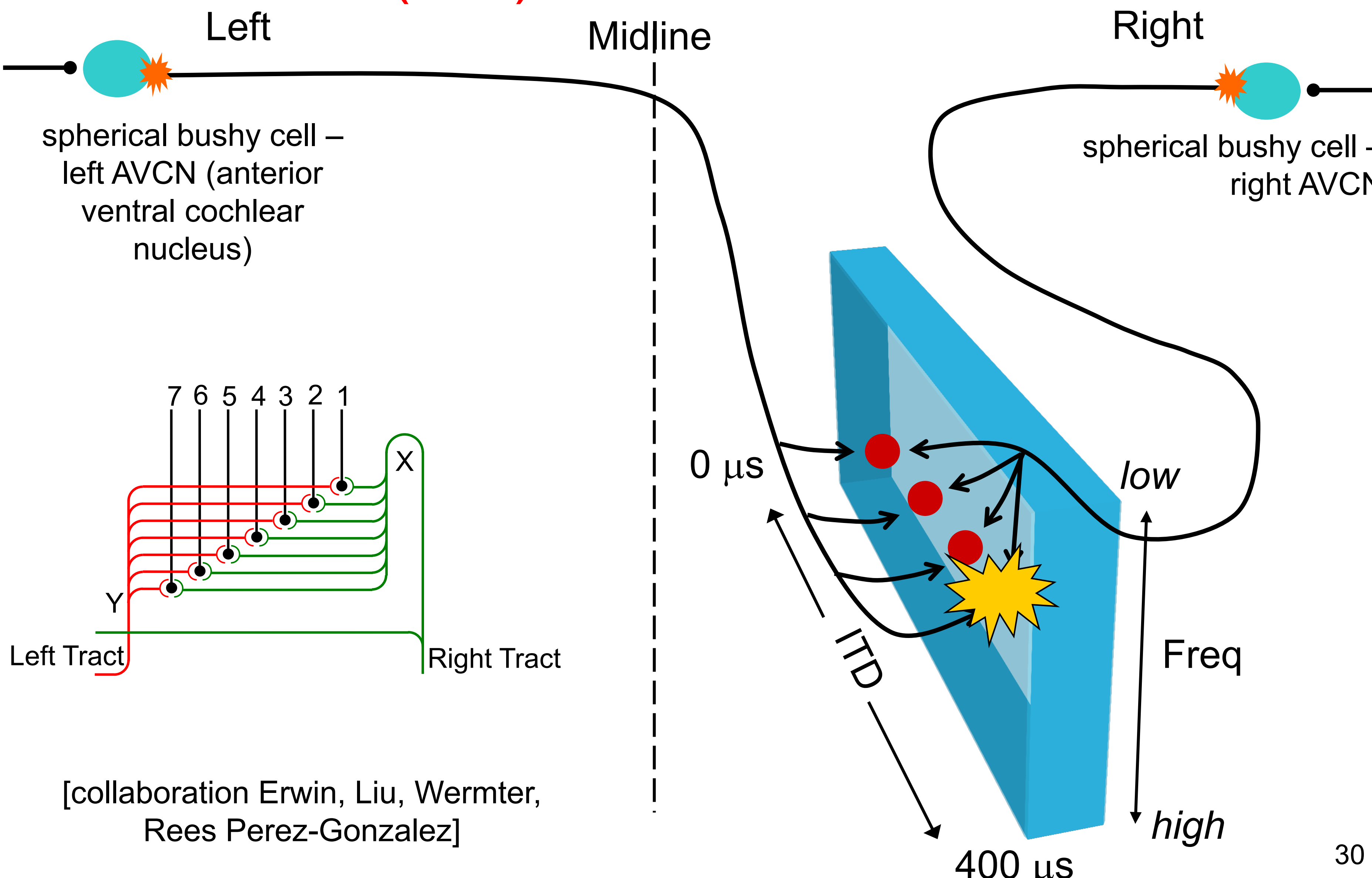


Cochlea model



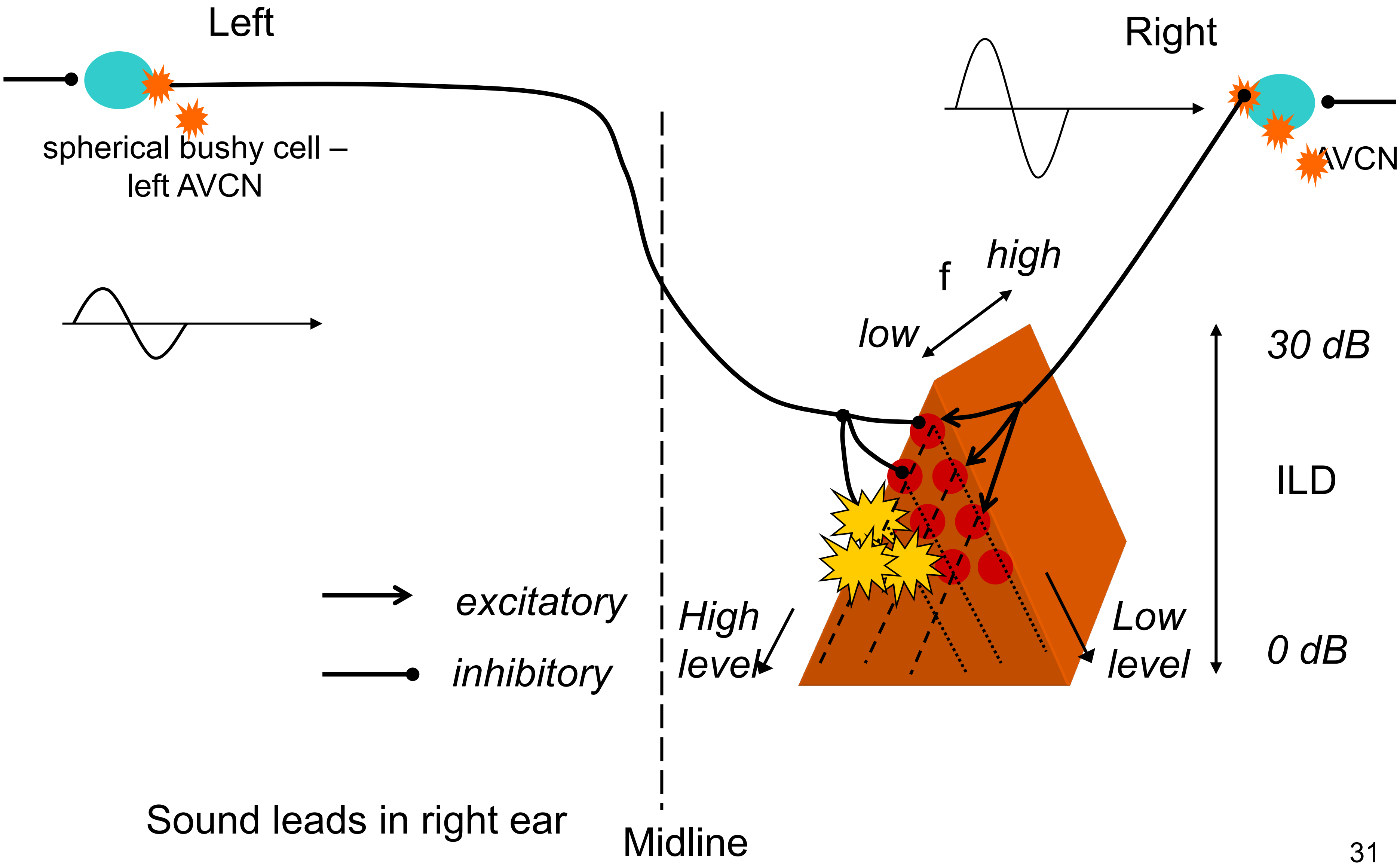
Spikes encoding time and level information

Revised Jeffress Interaural Time Difference (ITD) model in MSO

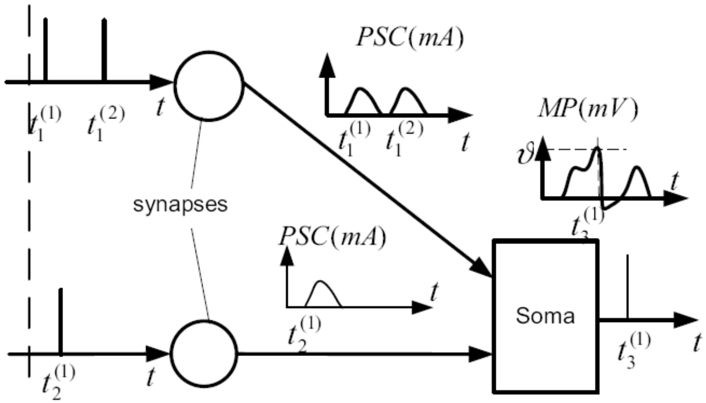


[collaboration Erwin, Liu, Wermter, Rees Perez-Gonzalez]

Interaural Level Difference (ILD) model in LSO



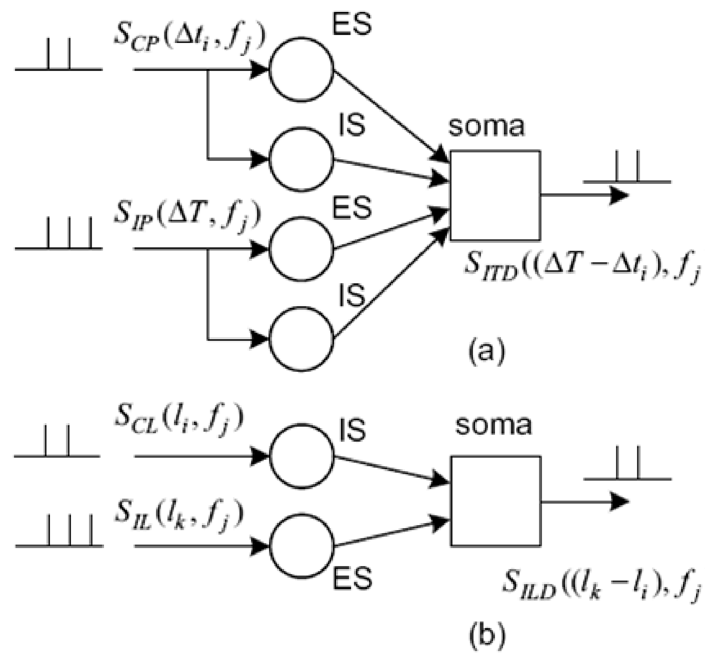
Model ITD and ILD



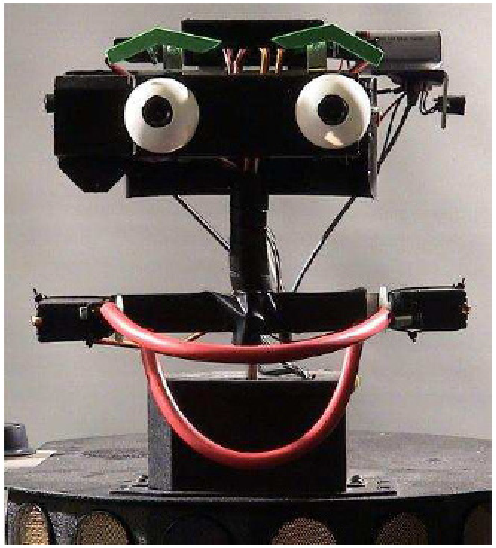
← Basic SNN structure and mechanism

(a) ITD coincidence → model

(b) ILD coincidence → model

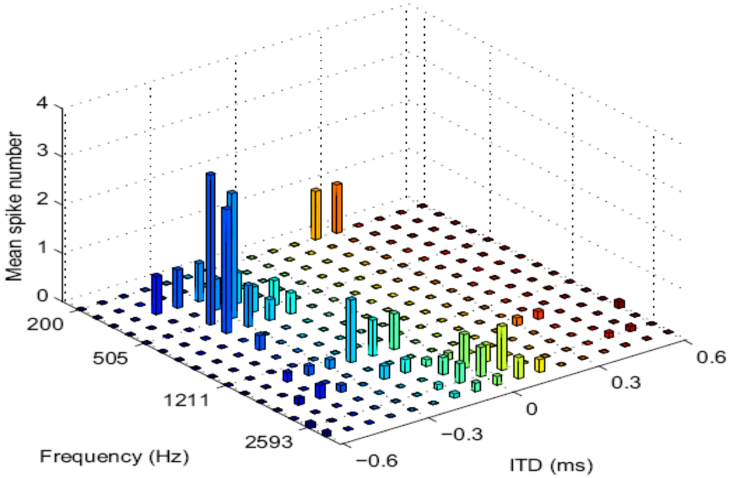


Testing the model in a realistic robot environment

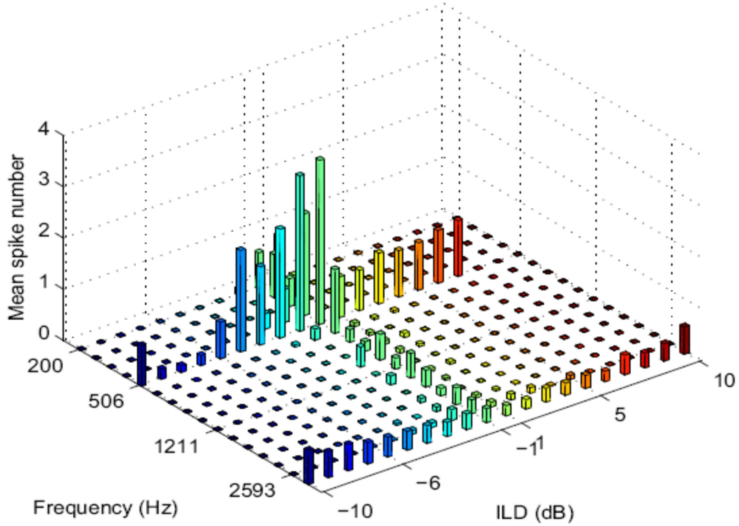


MIRA, a mobile robot for sound localisation experiments

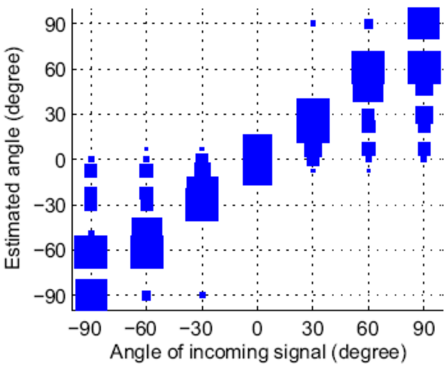
ITD results for a pure tone recorded in a noisy → environment (500 Hz at left 30 degree)



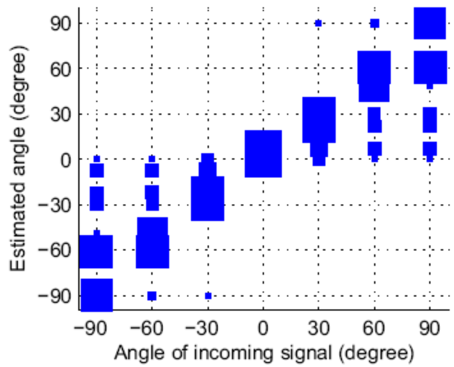
ILD results for the same recording →



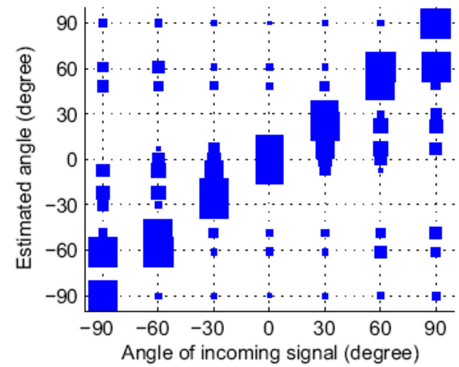
Artificial pure tone localisation



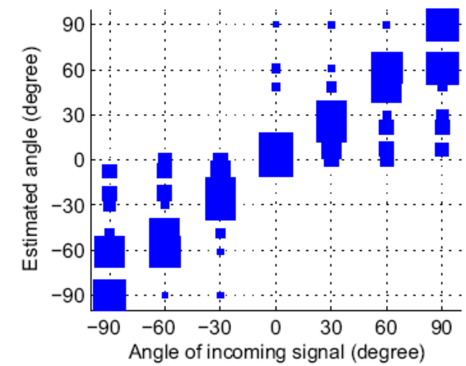
(a) 500 Hz with ITD only



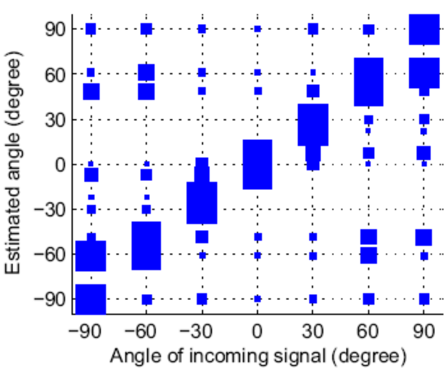
(b) 500 Hz with ITD & ILD



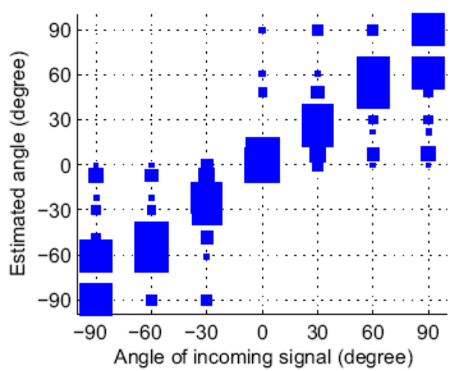
(e) 2000 Hz with ITD only



(f) 2000 Hz with ITD & ILD



(c) 1000 Hz with ITD only



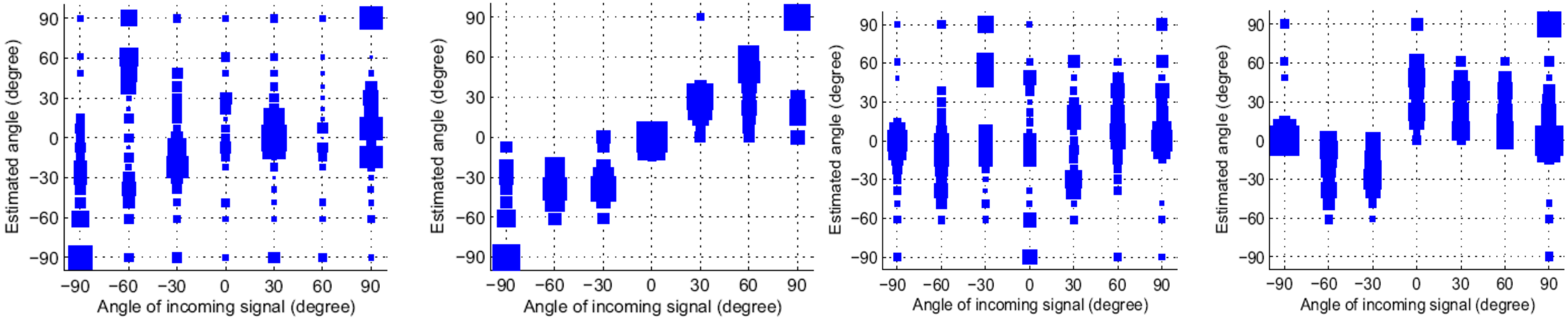
(d) 1000 Hz with ITD & ILD

1. *ITD only, localisation efficiency 70%.*
(i) highest efficiency when the sound source is in front of the observer,
(ii) efficiency fades down over 1.2 kHz

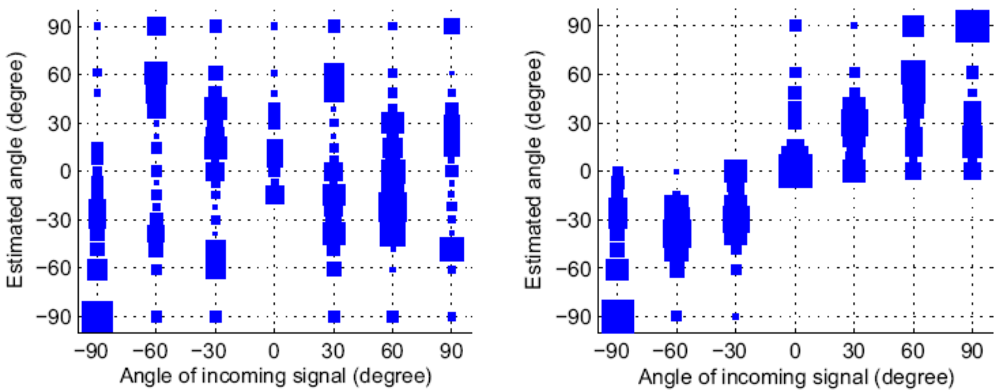
2. *ITD+ILD, localisation efficiency 80%, it is independent on the frequency.*

3. *The results match the data in human auditory system.*

Real pure tone localisation



(a) 500 Hz with ITD only (b) 500 Hz with ITD & ILD (e) 2000 Hz with ITD only (f) 2000 Hz with ITD & ILD

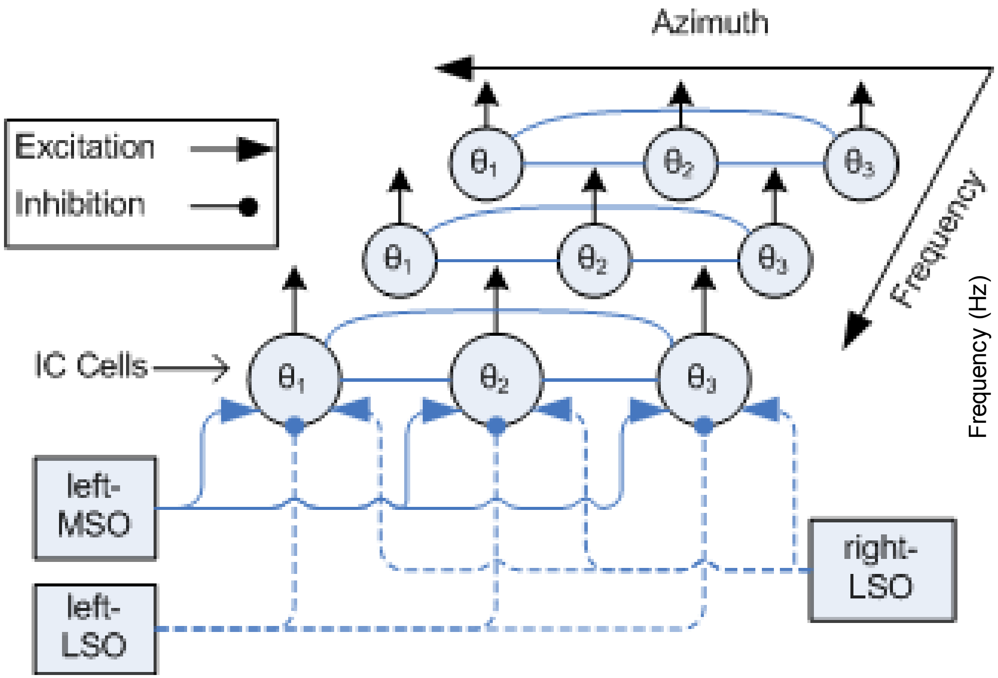


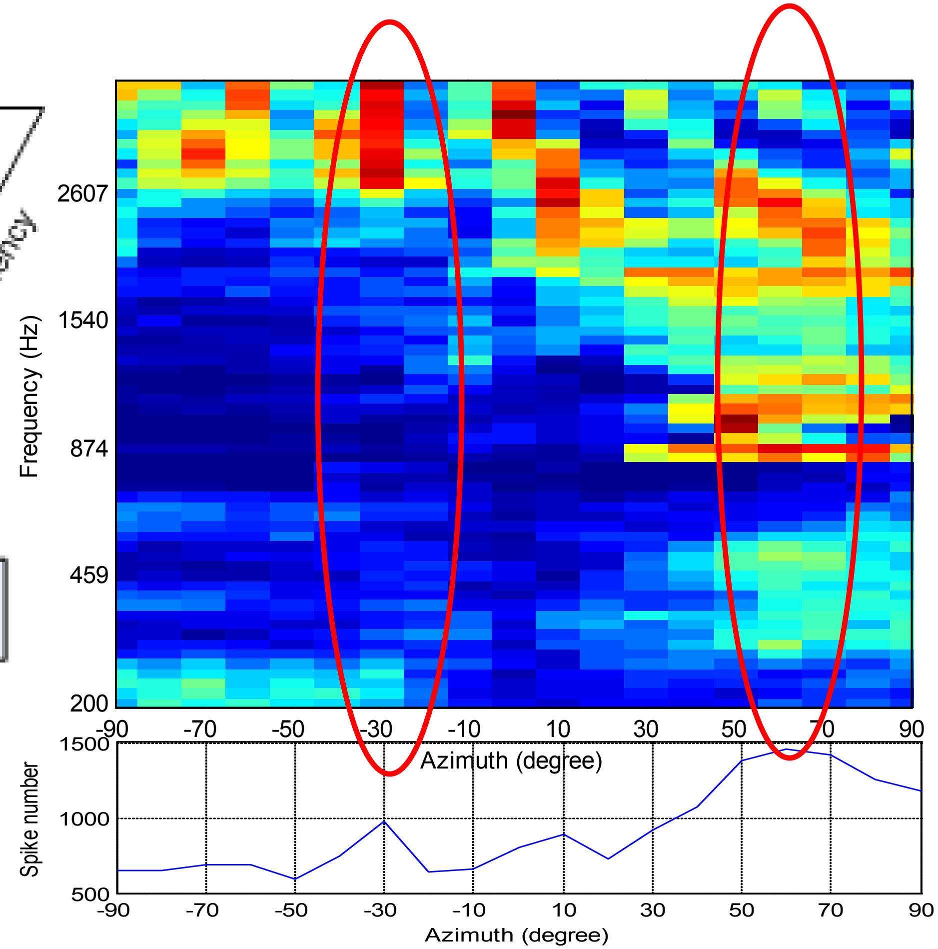
(c) 1000 Hz with ITD only (d) 1000 Hz with ITD & ILD

1. *Sound recorded on the robot, MIRA, in a general noisy background (30 dB).*
2. ITD only method only can achieve 50% localisation efficiency.
3. *ITD+ILD, increase the efficiency to 65%*

4. *The results match the data in human auditory system.*

IC Network for Sound Localisation

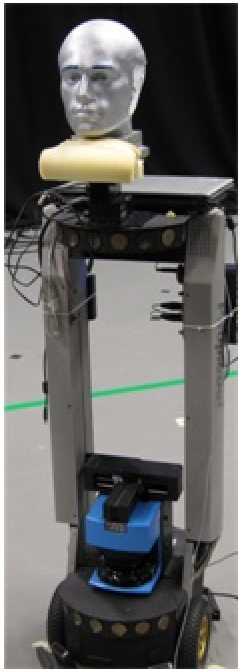


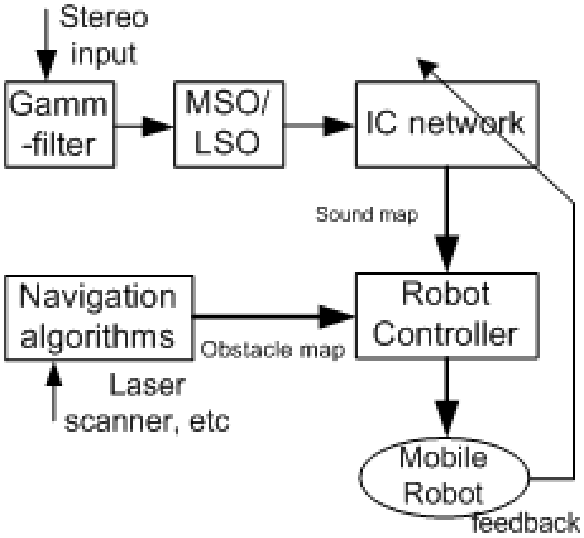


Left-IC network (there is also a mirror network of right-IC)

“hello” at 60 degree and “tea” at -30 degree

Evaluation on a Mobile Robot





Demo at NCAF conference

Summary

- ITD and ILD in one hybrid Spiking Neural Network
- Approach
  - Brings new insight into the brain mechanism of the auditory system
  - Demonstrates a practical application of sound localisation for mobile robots.
  - Improve speech recognition by user localisation

Summary and Reading

- Floreano and Mattiussi Bio-inspired Artificial Intelligence, MIT Press, 2008 (chapter 3)
- Papers Chacon, Liu, Erwin, Murray, Wermter on auditory localisation on Knowledge Technology website http://www.informatik.uni-hamburg.de/WTM/

Robust Sound Source Localization on the NAO