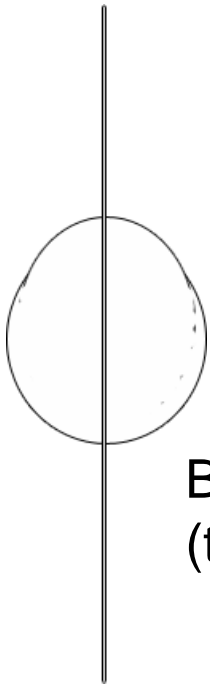

Sound Source Localization for Robot Auditory Systems

Zehaie Hailu



INTRODUCTION

- “You cannot hear blindfold! Can you?”



Blindfold head
(top view diagram)



Blindfold

Source:
http://teamworkscom.com/wp-content/uploads/2009/11/Blindfold_woman.png

TOPICS

BACKGROUND

SOUND SOURCE LOCALIZATION

SRP-PHAT + TDOA

TDOA + SRP-PHAT W/ SSC

EXPERIMENTAL SETTING

RESULTS

CONCLUSION



BACKGROUND

- **Communication: human-to-human**
 - Language: verbal or textual conversation
 - Signaling: body language, sign language, or posting sign
- **Communication: machine-to-machine**
 - Electromagnetism
 - Protocols
- **Communication: machine-to-human**
 - Sound: beeps , buzzes, and recorded prompts
 - Visual: text, color, LED's, and signs

BACKGROUND

- Human to machine: Passing instructions
 - Switches
 - Buttons
 - Graphic User Interface
- Domestic robots
 - Easy to use and communicate
 - Natural interaction by interpreting visual cue and deciphering certain sounds
- Hearing in robots
 - How loud (power): message or noise
 - Where it comes from (localization)
- Sound source localization
 - Robotic dog illustration

SOUND SOURCE LOCALIZATION

Signal Based:

- Time difference of arrival (TDOA)
 - Cross correlation
- Spectral analysis
 - Spatial spectra
 - Doppler Effect

Power Based:

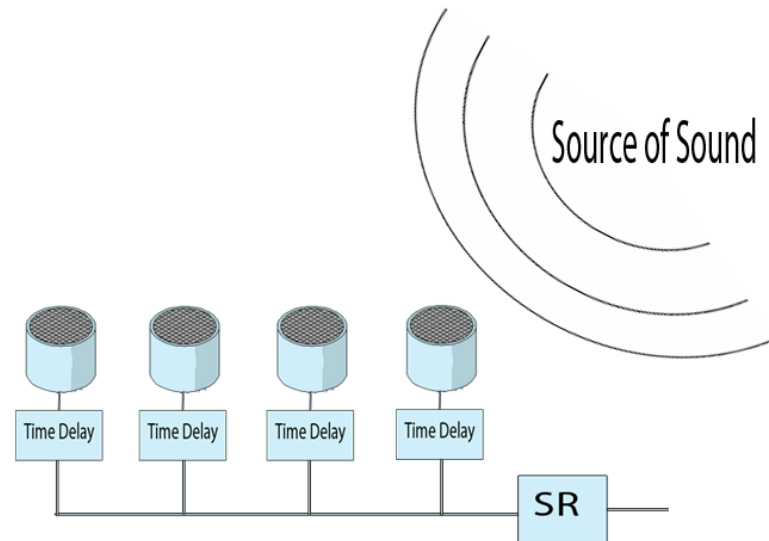
- Steered response power
- Steered response power with the phase transform (SRP-PHAT)
 - Grid search

SRP-PHAT + TDOA

- Hybrid: signal based + power based
 - SRP-PHAT + TDOA
- SRP-PHAT + TDOA
 - TDOA: Candidate locations
 - SRP: Beamforming
- Computational difficulties make SRP-PHAT hard to implement in real-time.

Beamforming

Photoshop



TDOA + SRP-PHAT W/ SSC

Search Space Clustering (SSC) Algorithm

Given the initial block b representing the entire search space:

$B \leftarrow \{b\}$

$C \leftarrow \varnothing$

while B is not empty

Calculate the TDOAs of each microphone pair at every vertex of b .

if the TDOAs at all vertices of b are the same

$B \leftarrow B - \{b\}$

$C \leftarrow C \cup \{b\}$

else

The block b is divided into a set of smaller size blocks,

b_1, b_2, \dots, b_n .

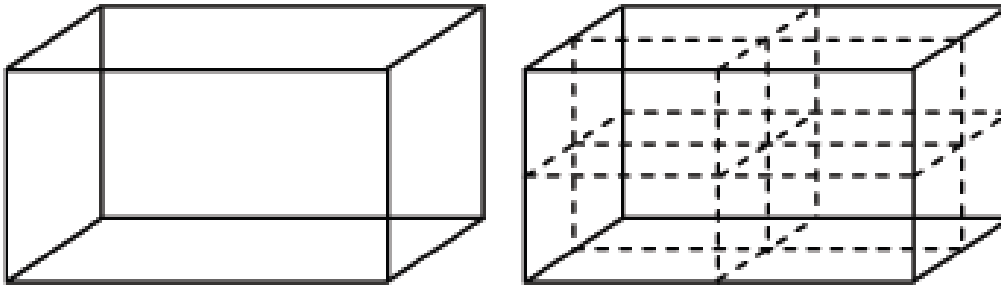
end

Any two blocks with the same TDOAs in C are merged.

The centroids of the blocks in C are stored in a look-up table.

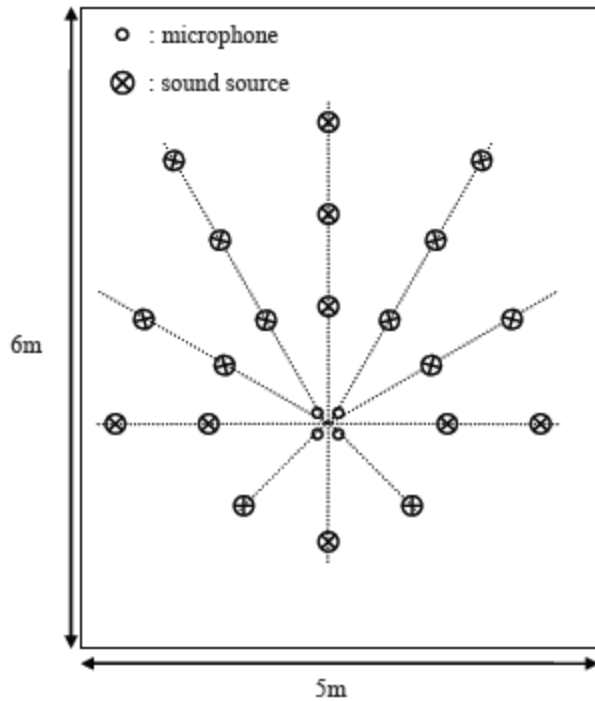


TDOA + SRP-PHAT W/ SSC



Search Space Clustering

EXPERIMENTAL SETTING



The room

Measurements

- Direction of arrival (DOA): azimuthal
- Elevation

RESULTS

LOCALIZATION ACCURACY, NUMBER OF SEARCH POINTS AND REALTIME FACTOR

	Grid Search				SSC
	20cm	15cm	10cm	1cm	
Azimuth (%)	89.5	89.9	93.6	93.9	93.9
Elevation (%)	85.4	88.8	89.2	92.8	92.8
Search Points	7500	19000	60000	6×10^7	5203

CONCLUSION

- Sound: important medium for human-machine interaction
- Microphone: hearing device
- Localization techniques use more than one ear – array of microphones