THE ANDREW & ERNA VITERBI FACULTY OF ELECTRICAL ENGINEERING

SIPL
Signal and Image Processing Lab

TECHNION Israel Institute of Technology

Binaural Graphic Analyzer

Dvir Marsh and Yossi Marciano, Supervised by Hadas Offir

In collaboration with WAVES

Introduction

- Binaural recording is a method of recording sound that uses two microphones, arranged with the intentto create a 3-D stereo sound sensation for the listener of actually being in the room with the performers or instruments
- Such recording can be obtaind phisically, by mounitng the microphones on a dummy head simulating the head and torso response; Or by a simulated dummy head recording, where the signal is sent through a complex mathematical algorithm that tries to simulate the binaural effect.

Dummy Head

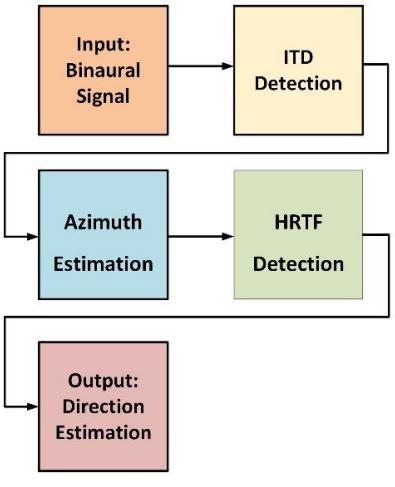
Goals

- Given a binaural recording consisted of multiple sources, we would like to determine (for each source)
  - Direction: azimuth and elevation
  - Main spectral content
  - Power
- Present and visualize that information on a GUI

Challenges

- Generation method is unknown: Signal could be generated by phisical recording or by simulation
- Spectral ques vary from different generation methods

Direction Estimation



ITD Detection

- A possible option for direction estimation is based on internal time difference - ITD
- The ITD was calculated by calculating the maximal cross correlation between the two channels

Azimuth Estimation

- A possible option for direction estimation is based on internal time difference - ITD
- We used an estimated formula by Brown & Duda to calculate the azimuth from the ITD
- Pros:
  - Easy and fast to calculate
  - Differ slightly between different techniques of binaural audio synthesis
- Cons:
  - Can't solve front back ambiguity
  - Can't indicate elevation
- We used this block to lower degrees of freedom in the next levels

What is HRTF

- The change inflicted on the signal as it propagates from the source to each ear is represented as a Head related transfer function - HRTF
- It's possible to record HRTFs in a studio
- We used the MIT database for our project

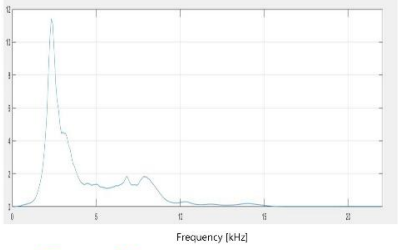
HRTF Detection

- It is common to use HRTF bank based algorithms to solve sound localization problems
- A solution reported to be succesful is the Cross Convolution (CC) Algorithm
- Direction Estimation according to CC is the solution to the following optimazation problem:

$$\min_{\hat{\theta},\hat{\varphi}} \sum \left\| R_{Left} * H_{Right}^{(\hat{\theta},\hat{\varphi})} - R_{Right} * H_{Left}^{(\hat{\theta},\hat{\varphi})} \right\|^2$$

- Azimuth estimation was used to lower degrees of freedom (otherwise have to check all directions)
- HRTF heuristic was combined in the algorithm to give extra weight to informative frequency bands
- Those bands are bands in which the variance of the HRTFs is highest:

Variance of HRTFs Along frquency bins



Results

- Successful Azimuth detection (not considering front back ambiguity)
- Azimuth RMSE: 64.5 degrees
- Elevation RMSE: 53.8 degrees