Desgin of a Microphone Array for Rollin‘ Justin

German Aerospace Center (DLR)[1], Technical University Munich (TUM)[2]

Marco Sewtz[1], Tim Bodenmüller[1], Rudolph Triebel[1,2]

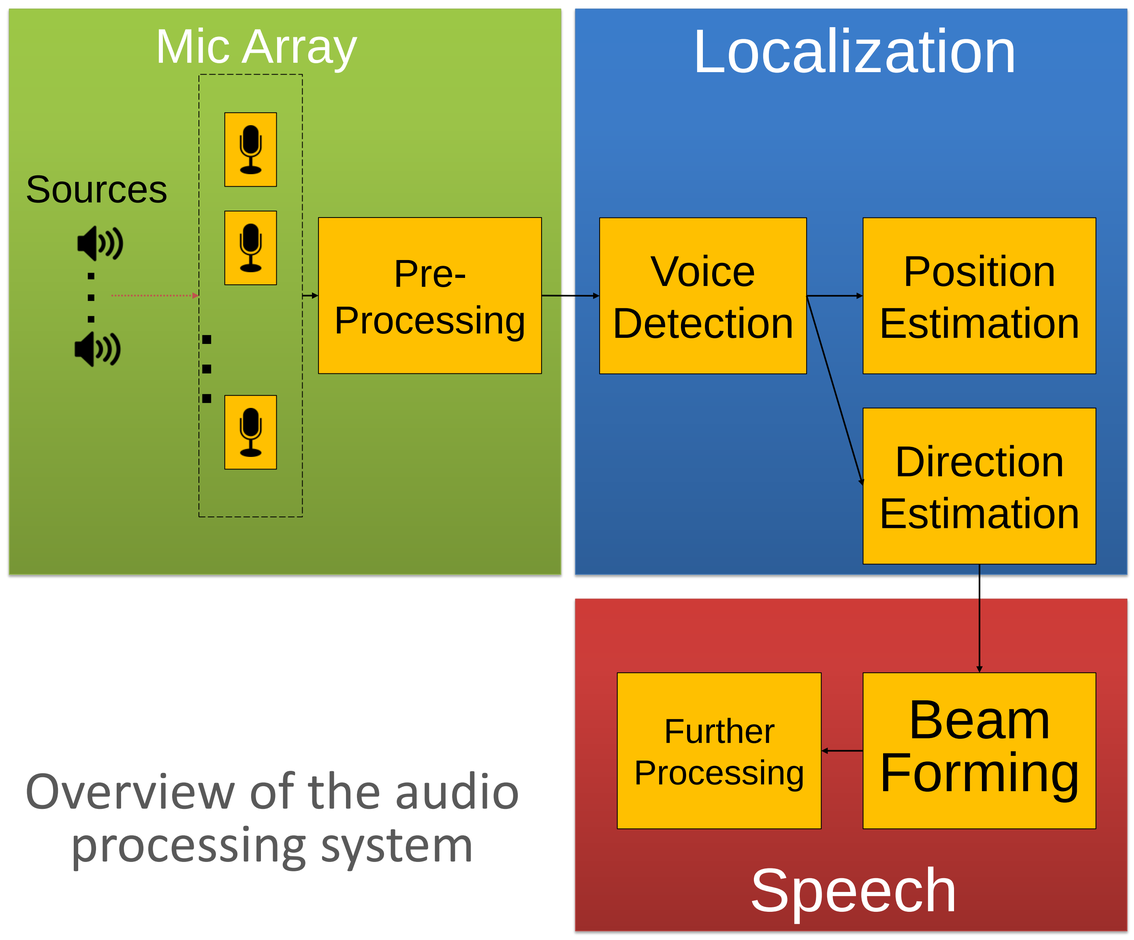
Motivation

For a humanoid assistant-robot operating in populated environments it is a major key competence to **interact with humans naturally and intuitively**.

Robot audition is a suitable modality, as it allows for **detection and tracking of speakers** from arbitrary positions around the robot and also from distant places. In addition, it allows the robot to be instructed by **voice commands**.

Robot Rollin‘ Justin in a lab environment

Processing Pipeline

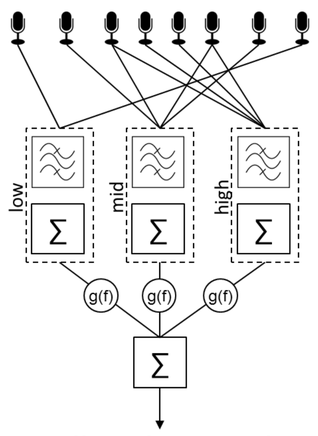


Overview of the audio processing system

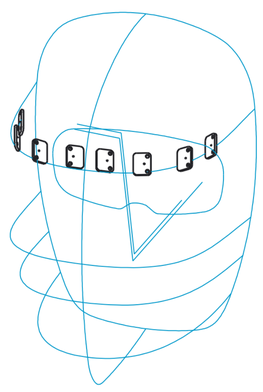
The processing is organized in three steps. Audio acquisition, localization (1D/2D) and subsequent speech processing.

Microphone Array

- Broadband Microphone Sub-Array[1]
- Target spectrum is divided intro sub-bands
- Positions of microphones optimized accordingly



Broadband Microphone Sub-Array concept



CAD drawing of the proposed microphone positions

Localization

- Estimation based on MUSIC[2] algorithm
- Extended to real time evaluation
- Main principle the delay between received signals
- For our array

$$\Delta t_i = (p_i{}^T e_\theta)/c_0$$

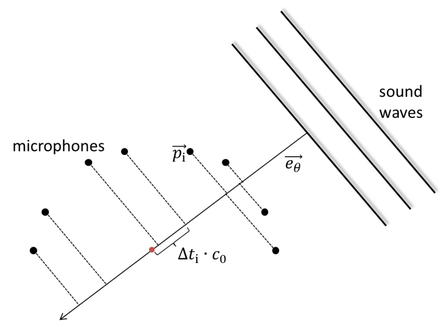
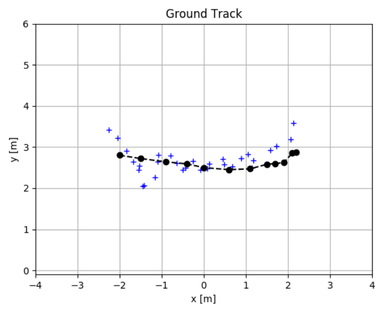


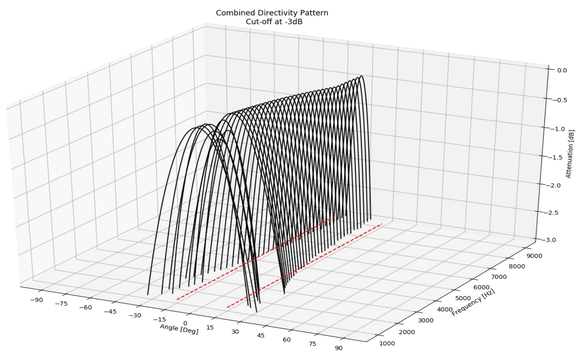
Illustration of the delay calculation for an arriving sound wave



Estimated positions of a tracked person using microphone arrays

Beam Forming

- Calculation of delay for focused direction
- Alignment of signal vector
- Frequency weighted sum of each sub-array



Simulated directivity pattern for the proposed beam former

Conclusion & Further Processing

We presented the **design of a microphone array** for sound source localization, beam forming and speech processing.

The latter one will be using the steered input from the beam former to **enhance quality and robustness**. We will use this to gain **voice commanding** ability.

All steps will be **processed online** on the robotic platform by its processing capabilities.

[1]I.A. McCowan, „Robust speech recognition using microphone arrays“ Ph.D. dissertation, Queensland University of Technology, 2001

[2]R. Schmidt, „Multiple emitter location and signal parameter estimation“ IEEE Trans. Antennas Propag., 1986