Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Entwurf vom January 6, 2023

Exposé for Master's Thesis

Sound Source Localization (and more TBD) using the Azure Kinect's Microphonearray on a Robot

Department of Informatics
MIN Faculty
Universität Hamburg

Hamburg, January 6, 2023

Roland Fredenhagen

dev@modprog.de
M.Sc. Informatics
Matriculation number: 7031533

Not sure if this should be the university e-mail as I wouldn't keep access to it

First Reviewer:
Second Reviewer:
Supervisor:

1 Research Question

TAMS has a robot equipped with an Azure Kinect providing not only a visual data, but also multichannel audio from a 7 microphone array [Mic23]. This spacial audio data should be used to perform sound source localization (SSL) and separation as well as reduction of ego and environmental noise.

2 Existing Work

2.1 Sound Source Localization

A very extensive collection of algorithms and their performance in different scenarios is provided by Evers et al. [Eve+20] with the LOCATA challenge comparing different algorithms in situations very similar to the proposed usage in the research question (chapter 1). Containing scenarios with moving sensors and sources on top of background noise and distortions such as reverberations. Their work also contains tests with multiple audio sources.

The microphone setups matching ours most closely is the Robot head though theirs consists of a total of 12 microphones that are placed spherical opposed to the planar circle of the Kinect and the DICIT which has a two-dimensional configuration, but whose microphones are placed a lot further apart with a total range of almost 2 meters.

The best performing algorithm was by Madmoni et al. [Mad+18] using the Robot head microphone array was a combination of the Direct-Path Dominance Test [NR14] and the MUSIC-Algorithm [Sch86], though it was only tested with the static sound sources, and they assumed a priori knowledge of the number of sound sources.

The only submitted algorithm applied to all scenarios using the Robot Head was by Li et al. [Li+18]. They used three modules to provide both localization and tracking.

- A recursive direct-path relative transfer function (DP-RTF) estimation module based on the estimation of the convolutive transfer function proposed by Li et al. [Li+16].
- An online multiple-speaker localization module assigning DP-RTF features to sources adopting a complex Gaussian mixture model.
- A multiple-speaker tracking module using a variational expectation maximization algorithm.

This allows their algorithm to not only locate static sound sources but also allows scenarios with either the sound sources, the microphone array or both moving.

2.2 Noise Reduction

3 Planned

As three-dimensional localization is difficult to achieve using a static microphone array [Eve+20] the localization efforts will focus on the direction of arrival (DoA).

4 Time Plan

Bibliography

[Eve+20] Christine Evers et al. The LOCATA challenge: Acoustic source localization and tracking. In: IEEE/ACM Transactions on Audio, Speech, and Language Processing 28 (2020), pp. 1620–1643.

[Li+16] Xiaofei Li et al. Estimation of the direct-path relative transfer function for supervised sound-source localization. In: IEEE/ACM Transactions on Audio, Speech, and Language Processing 24.11 (2016), pp. 2171–2186.

[Li+18] Xiaofei Li et al. A cascaded multiple-speaker localization and tracking system. In: arXiv preprint arXiv:1812.04417 (2018).

[Mad+18] Lior Madmoni et al. Description of algorithms for Ben-Gurion University submission to the LOCATA challenge. In: arXiv preprint arXiv:1812.04942 (2018).

[Mic23] Microsoft. Azure Kinect DK. 2023. url: https://azure.microsoft.com/en-us/products/kinect-dk/ (visited on 01/05/2023).

[NR14] Or Nadiri and Boaz Rafaely. Localization of multiple speakers under high reverberation using a spherical microphone array and the direct-path dominance test. In: IEEE/ACM Transactions on Audio, Speech, and Language Processing 22.10 (2014), pp. 1494–1505.

[Sch86] Ralph Schmidt. Multiple emitter location and signal parameter estimation. In: IEEE transactions on antennas and propagation 34.3 (1986), pp. 276–280.

Todo list

Not sure if this should be the university e-mail as I wouldn't keep access to it . 1