Dataset of real recordings in an outdoor wireless acoustic sensor network

Anastasios Alexandridis[1], Anthony Griffin[2], and Athanasios Mouchtaris[1]

[1]Institute of Computer Science,
Foundation for Research and Technology-Hellas, Heraklion 70013, Greece
Department of Computer Science, University of Crete, Heraklion 70013, Greece

[2]High Performance Computing Research Laboratory,
Electrical and Electronic Engineering Department,
School of Engineering, Computer and Mathematical Sciences,
Auckland University of Technology, Auckland 1024, New Zealand

1 Introduction

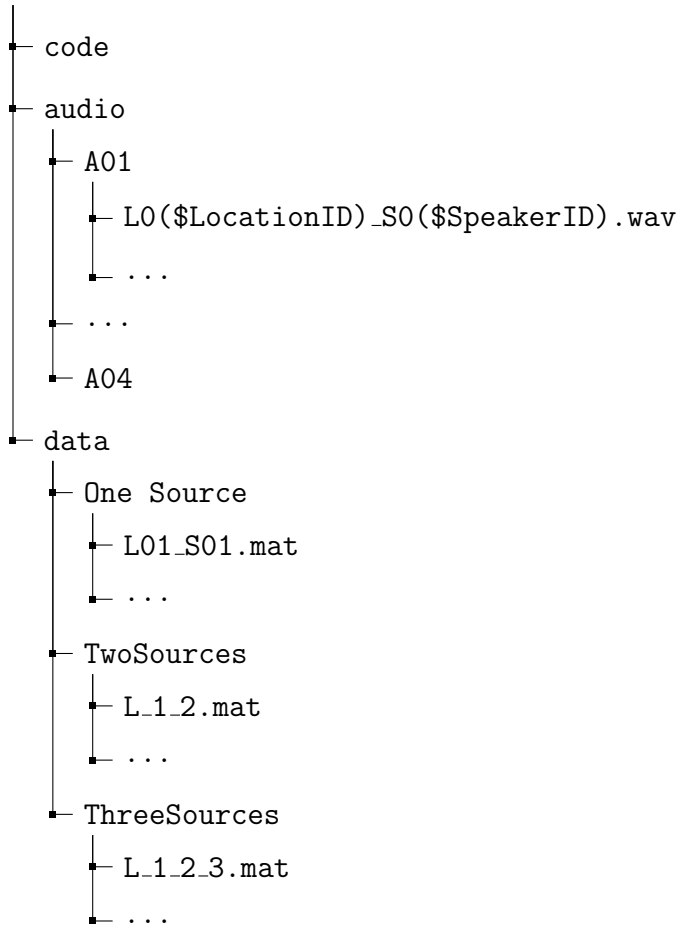
This documentation describes the structure of the dataset of real recordings in a wireless acoustic sensor network (WASN). For a detailed description of the dataset and the recording procedure, please refer to [1]. For any comments or questions please contact `analexan@ics.forth.gr`.

2 Dataset structure

The dataset contains:

- the audio recordings
- helper MATLAB functions
- direction of arrival (DOA) and location estimates from our experiments in [1]

The dataset has the following structure:



2.1 `audio` folder

The audio folder contains the recordings of each speaker (S01 - S05) at each location (L_1 - L_7). Each recording is an 8-channel wav file sampled at 44.1 kHz. The file name of each recording has the following format:

`L0($LocationID)_S0($SpeakerID).wav`

where `($LocationID)` and `($SpeakerID)` represent the location and speaker IDs. For example, `L03_S01.wav` is the recording of speaker S01 at location L_3.

2.2 `data` folder

The `data` folder contains the DOA and location estimation results, which are presented in [1]. DOA estimation (after the calibration procedure [1]) and location estimation results are given for the scenarios of one, two, and three simultaneously active sound sources.

For the single source case we tested all speakers at all locations. For the two sources case we tested speakers S03 and S04 at all possible location pairs (21 combinations), while for the three sources case we tested speakers S03, S04, and S01 at all possible location combinations (35 combinations). Thus, the naming of the .mat files has the following format:

- One Source: `L0($LocationID)_S0($SpeakerID).mat`
- Two Sources: `L_($LocationID of S03)_($LocationID of S04).mat`
- Three Sources: `L_($LocationID of S03)_($LocationID of S04)_($LocationID of S01).mat`

For example, `L_1_2.mat` represents an experiment with two sources where speaker S03 is at location L_1 and speaker S04 is at location L_2 and `L_1_2_3.mat` represents and experiment with three sources where speaker S03 is at location L_1, speaker S04 is at location L_2, and speaker S01 is at location L_3.

Each .mat file contains the following variables:

- `DOAEstimates`: 4×1 cell array that contains the DOA estimates from each microphone array. Each element of the cell array is another cell array that contains the DOA estimates for each time frame.
- `NumSourcesEstimates`: 4×1 cell array that contains the estimated number of sources for each microphone array. Each element of the cell array is a vector that contains the estimated number of sources for each time frame.
- `EstLocations`: Cell array containing the estimated locations of the sources at each time frame.
- `DOAs`: Matrix with the true DOAs of the sources. Each column denotes the true DOAs of a source with respect to the four microphone arrays (A_1 - A_4).

- `sensorXYs`: Matrix with the microphone array locations. Each row contains the x and y coordinates of a microphone array.
- `sourceXYs`: Matrix with the true sources locations. Each row contains the x and y coordinates of an active sound source.

Note that in `DOAEstimates`, `NumSourcesEstimates`, and `EstLocations` the first 42 entries are empty. This is due to the DOA estimation method which requires 1 second initialization time. Hence, DOA estimates (and thus location estimates) start from frame 43.

2.3 `code` folder

Helper MATLAB functions are provided in order to load the audio signals, plot intermediate DOA estimation results and calculate the localization error. The following functions are provided:

- `load_signal.m`: loads the signals of speakers at specific locations.
- `GenTrueDOAsForSourcePosition.m`: calculates the DOAs of a source with respect to the microphone arrays.
- `mic_array_coordinates.m`: Generates the coordinates of the microphones of a uniform circular microphone array given the array center, the array radius and the number of microphones.
- `PlotFrameDOAEstimatesAllArrays.m`: Plots the DOA estimates for each source and each array along with the ground truth DOA estimates. For example, the following code:

  ```
  load('../data/TwoSources/L_1_5.mat')
  PlotFrameDOAEstimatesAllArrays(DOAEstimates, NumSourcesEstimates, DOAs)
  ```

  produces Figure 1.
- `SingleSourceErrors.m, TwoSourcesErrors.m, ThreeSourcesErrors.m`: calculates the root-mean squared error (RMSE) for each location or location combination of one, two, and three active sources and the overall RMSE. For example:

  ```
  [RMSE_per_location,OverallRMSE] = TwoSourcesErrors('../data/TwoSources')
  ```

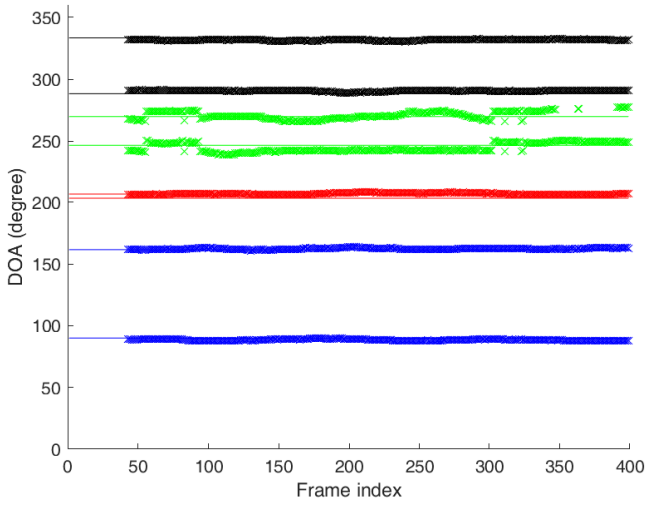


Figure 1: Estimated DOAs for each microphone array at each time frame for a scenario with two active sound sources. Different colors are used to denote the different microphone arrays. The solid lines correspond to the true DOAs of the sources and the x's correspond to the estimated DOAs at each time frame.

References

[1] A. Alexandridis, A. Griffin, and A. Mouchtaris, "Multiple source location estimation on a dataset of real recordings in a wireless acoustic sensor network," in *International Workshop on Multimedia Signal Processing (submitted)*, 2018.