ECE ILLINOIS
Department of Electrical and Computer Engineering

ILLINOIS

Performance Tradeoffs in Overdetermined Multimicrophone Array Processing

Matthew D. Skarha, Ryan M. Corey, and Andrew C. Singer

ABSTRACT

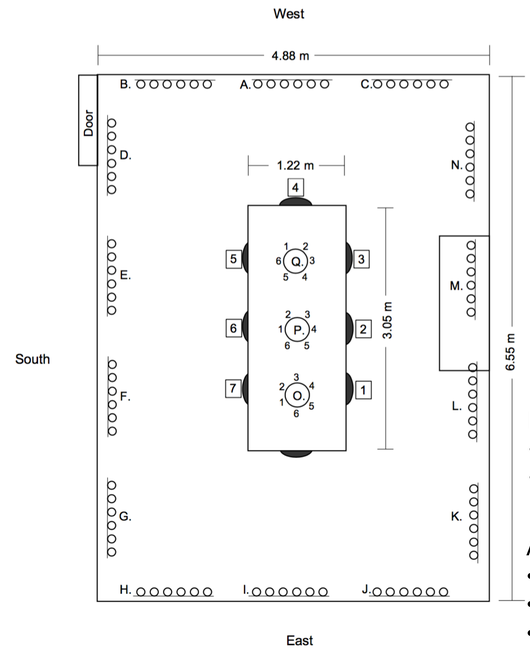
Two new multi-channel audio recording datasets are presented for overdetermined multimicrophone array processing. The datasets consist of impulse response recordings from large (100+) microphone arrays for various loudspeaker and array configurations in two conference-style meeting rooms. In each configuration, three logarithmic sine sweeps (to improve SNR) and one clip from the VCTK corpus were played and recorded. These datasets can be used to show that listening devices can achieve better performance using microphone arrays to filter acoustic signals in both space and time. The primary applications explored in this work are how well beamforming, blind source separation, and source localization scale with number of microphones.

OVERDETERMINED BLIND SOURCE SEPARATION (OBSS)

Majority of research attention has been focused on determined and underdetermined systems for BSS.

- These have limited performance in complex acoustic environments with many speakers (e.g. a cocktail party)
- Large arrays allow for subset selection and dimensionality reduction for the exploitation of redundant information

CSL 221 DATASET



Conference Room Meeting Scenario

North

Loudspeakers:
- Denoted by numbers
- Model: Presonus Eris E3.5

Arrays:
- 6cm UCA, ULA
- Denoted by letters
- Model: Audix TM1

MICROPHONE ARRAYS

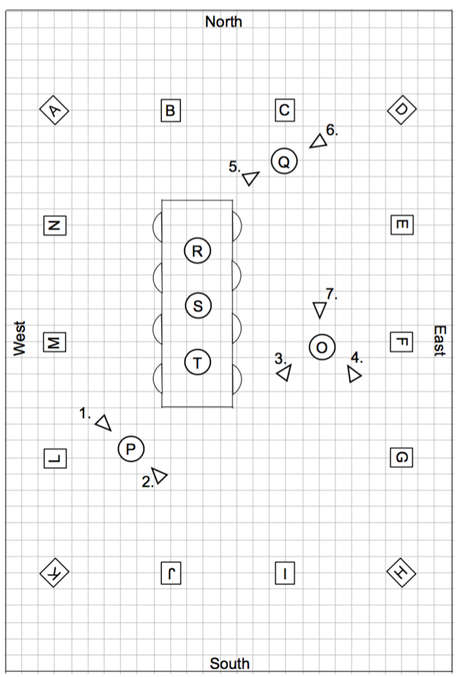
- Use multichannel processing to separate and filter sources from different directions
- Good performance in noisy environments
- Advanced algorithms are expensive, not real-time

UNIFORM LINEAR & CIRCULAR ARRAYS

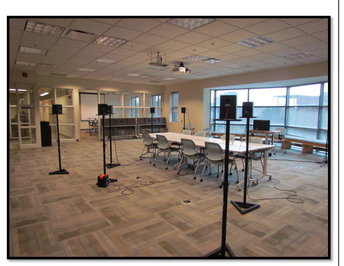
6 cm spacing ULA | 6 cm radius UCA

- Audix TM1 Measurement Microphones
- Stands modeled in OpenSCAD and 3D printed

CSL 301 DATASET



Cocktail Party Scenario



COCKTAIL PARTY PROBLEM

- Tremendous applications in speech processing and communication
- Beamforming using microphone arrays is a well-known technique for target extraction
- Can be used for localization and separation of sources

SOURCE SIGNALS

1. Logarithmic sine sweep
   - 50 Hz to 20 kHz
   - Sample rate: 48 kHz
   - Repeated three times to improve SNR
2. VCTK corpus clip
   - Speech recording

This work was supported in part by Systems on Nanoscale Information fabriCs (SONIC), one of the six SRC STARnet Centers, sponsored by MARCO and DARPA. This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant Number DGE-1144245.

CSL: Coordinated Science Lab