# EE513 Project 3: Beamforming and Cocktail Party Noise

Matt Ruffner
matthew.ruffner@uky.edu

## ABSTRACT

In this project, the cocktail party problem (listening in on a specific voice in a room full of chatter) was analyzed. SNR levels are determined at which 90% of spoken subject verbiage is intelligible. Preforming this analysis with both an inverse distance beamforming algorithm (IDBF) and a masked version of the IDBF, over two two microphone arrangements, the effectiveness of both methods will be quantized.

## 1. APPROACH

Two scripts were written, `room_sim.m` and `analysis.m`. These scripts simulated the experimental parameters and saved the results as individual files to be listened to all at once. For the experiment, three noise voices were considered, that of `woman1.wav`, `woman2.wav`, `woman3.wav`. The SOI was kept as `man2.wav`.

### 1.1 Description

The first script is responsible for reading in the room noise and target speaker sound sources and recording them on the 8 microphones in the room. The Speaker of Interest (SOI) is recorded separately, the three noise speakers are recorded in one noise file. These noise and speaker files are generated for each microphone layout, for a total of 4 generated `.wav` files for each trial of the experiment.

In order for the second script to correctly beamform on the SOI, speaker and microphone location are recorded in `.dat` files for each trial as well.

The second script is in charge of creating a set of sound clips over varying SNR levels, to be listened to in order to determine the lowest SNR value at which 90% of the SOI's words can be understood. Eight SNR values were iterated over, from .1 to .8 in .1 step increments. Audio files were output for each SNR value, with each microphone layout, as well as with each type of beamformer (IDBF vs. Masked IDBF). This resulted in a total of 32 `.wav` files to be listened to for each experimental trial.

ACM ISBN N.A..
DOI: N.A.

### 1.2 Microphone Positions

| | Adjacent Clusters | | Equidistant Perimeter | |
|---|---|---|---|---|
| Mic # | X | Y | X | Y |
| 1 | 1.5677 | 0 | 0 | 1.0667 |
| 2 | 1.5892 | 0 | 0 | 2.1333 |
| 3 | 1.6108 | 0 | 1.0667 | 3.2 |
| 4 | 1.6323 | 0 | 2.1333 | 3.2 |
| 5 | 0 | 1.5677 | 3.2 | 2.1333 |
| 6 | 0 | 1.5892 | 3.2 | 1.0667 |
| 7 | 0 | 1.6108 | 2.1333 | 0 |
| 8 | 0 | 1.6323 | 1.0667 | 0 |

**Table 1: Microphone positions, all microphones at a height of 1.5m.**

### 1.3 Speaker Positions

Table 2 shows the speaker positions over all three trials.

| | Speaker Positions (X,Y,Z) | | | |
|---|---|---|---|---|
| Trial | SOI | Noise Speaker 1 | Noise Speaker 2 | Noise Speaker 3 |
| 1 | 2.4,1.8,1.0 | 1.0,2.4,1.9 | 2.3,1.4,2.0 | 1.0,1.7,2.0 |
| 2 | 1.0,1.0,1.9 | 0.4,1.1,1.3 | 0.8,2.0,1.7 | 1.7,2.1,1.7 |
| 3 | 0.9,1.9,0.5 | 1.7,2.6,1.9 | 1.6,1.0,1.4 | 2.8,1.7,2.0 |

**Table 2: Speaker locations for each trial using IDBF**

## 2. RESULTS

Tables 3 and 4 show the results for determining minimum subject (Speaker of Interest) intelligibility SNR values when using the Inverse Distance BeamFormer (IDBF) and Masked IDBF. Table 3 shows these SNR values when subjects were recorded using two adjacent clusters of microphones. Table 4 shows these SNR values when subjects were recorded using a perimeter array with equidistantly spaced microphones.

### 2.1 Adjacent Cluster Mics

The adjacent cluster microphone array was designed so that the spacing between microphones does not allow aliasing of frequencies over 8kHz. These linear arrays of microphones were positioned on adjacent walls of the room at a height of 1.5 meters. This would help when distinguishing female subject voices, although in this case I don't think it was that useful since my target voice was male and all of the noise sources were selected to be female voices.

The performance of the adjacent linear arrays of microphones was consistently less than that of the equidistantly spaced perimeter microphones.

The performance of the Masked IDBF was consistently better than that of the single iteration IDBF.

| Trial | IDBF SNR | Masked IDBF SNR |
|---|---|---|
| 1 | -4dB | -5.2dB |
| 2 | -5.2dB | -5.2dB |
| 3 | -4dB | -5.2dB |

**Table 3: SNR for intelligibility results for an adjacent cluster microphone spacing.**

## 2.2 Equidistant Perimeter Mics

The equidistant array of perimeter microphones performed better in this experiment than the adjacent cluster microphones. The minimum level of target speaker signal to noise speaker ratio was lower for each trial arrangement of speaker positions.

Similar to the results of the adjacent arrays, the minimum SNR for intelligibility was lower in each trial when using the Masked IDBF.

Equidistant Perimeter

| Trial | IDBF SNR | Masked IDBF SNR |
|---|---|---|
| 1 | -7dB | -10dB |
| 2 | -5.2dB | -7dB |
| 3 | -5.2dB | -7dB |

**Table 4: SNR for intelligibility results for equidistantly spaced microphones for trials with varying speaker locations.**

# 3. CONCLUSIONS

From this experiment we can conclude that the Masked Inverse Distance BeamForming algorithm is more effective than a single iteration of the IDBF.

If more microphone spacings were tried, it might become more evident that the Masked IDBF is more effective, since the delays between incoming speaker signals would be greater and possibly more noticeable with greater microphone spacings.

The results also suggest that the equidistant perimeter microphone arrangement is more effective than the adjacent cluster arrangement. This might be a product of the spacing of the linear arrays in the adjacent layout, however it also makes sense that having at least one microphone on each wall would help the Masked IDBF greatly.

From the results we can conclude that preforming Masked Inverse Distance Beam Forming on a SOI recorded with an equidistantly spaced perimeter array of microphones is the most effective means of eliminating cocktail party noise.