Cocktail Party Hearing Aid Using a Microphone Array

Luting Yang, Edward Richter and Dr. Arye Nehorai

Department of Electrical and Systems Engineering, Washington University in St. Louis

Abstract

Beamforming is a signal processing technique used in sensor arrays for directional signal transmission or reception. This is achieved by combining elements in a phased array in such a way that signals at particular tangles experience constructive interference while others experience destructive interference. In this project, a 64-element antenna array can achieve beamforming by the algorithm which is doing Fast Fourier and Inverse Fast Fourier of signal to cancel the phase coefficient to find maximum power of 64 elements. After applying this algorithm from 1 degree to 359 degree, the algorithm can find the angle of arrival signal and track it.

Motivation

Hearing in the presence of background noise is challenging enough for people with normal hearing. The problem is much worse for the hearing impaired. It is also a situation where traditional hearing aids don't perform well. In this project, researcher will use a 64-element microphone array. We will place the array in the center of a table while several people sitting at the table carry on a normal conversation to produce the Cocktail Party Effect . Using data collected from the microphone array, the student will develop algorithms to remove the background noise and amplify the current speaker.. This filtered signal will then be transmitted to the smart phone of a hearing impaired person wearing one or two headphones.

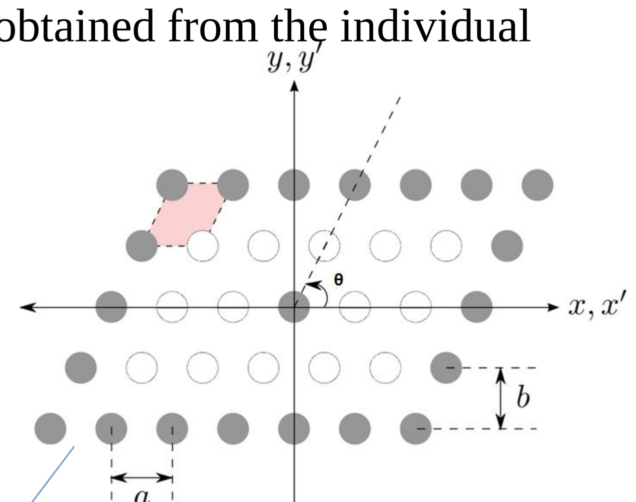
Cocktail Party Effect

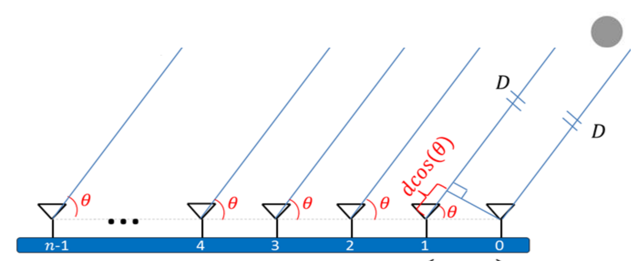
The cocktail party effect is the phenomenon of being able to focus one's auditory attention on a particular stimulus while filtering out a range of other stimuli, like people may focus on someone who speaks loudly or someone speak their mother tongue in a more international circumstance.

Antenna Array

An 64-element antenna array is a set of individual antennas used for receiving radio waves, connected together in such a way that their individual currents are in a specified amplitude and phase relationship. This allows the array to act as a single antenna, generally with improved directional characteristics than would be obtained from the individual elements.

Phase lag is caused by the difference of arrival time to each elements due to their geometric distribution on the Plane.





Algorithm

Receive signal of 64 channel

Receiving the speech signal comes from the distance d from unknown angle θ. The frequency of human speech is between 500Hz~3000Hz and the speed of speech transmitting in the air 340m/s

Frequency domain matrix

$$\begin{bmatrix} Z_{11}e^{i\phi_{11}} & Z_{12}e^{i\phi_{12}} & \dots & Z_{1N}e^{i\phi_{1N}} \\ Z_{21}e^{i\phi_{21}} & Z_{22}e^{i\phi_{22}} & \dots & Z_{2N}e^{i\phi_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{641}e^{i\phi_{641}} & Z_{642}e^{i\phi_{642}} & \dots & Z_{64N}e^{i\phi_{64N}} \end{bmatrix}$$

Matrix after elimination

$$\begin{bmatrix} Z_{11}e^{i\phi_{11}} & Z_{12}e^{i\phi_{12}} & \dots & Z_{1N}e^{i\phi_{1N}} \\ Z_{11}e^{i\phi_{11}} & Z_{12}e^{i\phi_{12}} & \dots & Z_{1N}e^{i\phi_{1N}} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{11}e^{i\phi_{11}} & Z_{12}e^{i\phi_{12}} & \dots & Z_{1N}e^{i\phi_{1N}} \end{bmatrix}$$

Sampling

Sampling the signal of each elements by Fs sample frequency, then in a certain numble of samples N, we can get a 64*N matrix of signal

Fast Fourier Transform

$$X_k = \sum_{n=0}^{N-1} X_n e^{-i2\pi k\frac{n}{N}}$$

Eliminate phase

Calculate the phase difference between every element to the reference element.(First row). Then eliminate the phase different of each element.

Beamforming

However, in the reality, the coming angle of signal is unknown, the microarray need to rotate its beam to find the signal by adding phase for each element before applying the algorithm above. Adding phase from 0 degree to 359 degree by one degree resolution.

Frequency domain matrix

$$\begin{bmatrix} Z_{11}e^{i\phi_{11}} & Z_{12}e^{i\phi_{12}} & \dots & Z_{1N}e^{i\phi_{1N}} \\ Z_{21}e^{i\phi_{21}} & Z_{22}e^{i\phi_{22}} & \dots & Z_{2N}e^{i\phi_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{641}e^{i\phi_{641}} & Z_{642}e^{i\phi_{642}} & \dots & Z_{64N}e^{i\phi_{64N}} \end{bmatrix}$$

Create 360 phase co-Efficient matrices of each one Degree by the Euler's formula and FFT.

Phase coefficient matrix

$$\begin{bmatrix} e^{-i\theta_{11}} & e^{-i\theta_{12}} & \dots & e^{-i\theta_{1N}} \\ e^{-i\theta_{21}} & e^{-i\theta_{22}} & \dots & e^{-i\theta_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-i\theta_{641}} & e^{-i\theta_{642}} & \dots & e^{-i\theta_{64N}} \end{bmatrix}$$

Multiply the those two Matrix, create 360 new Frequency matrices

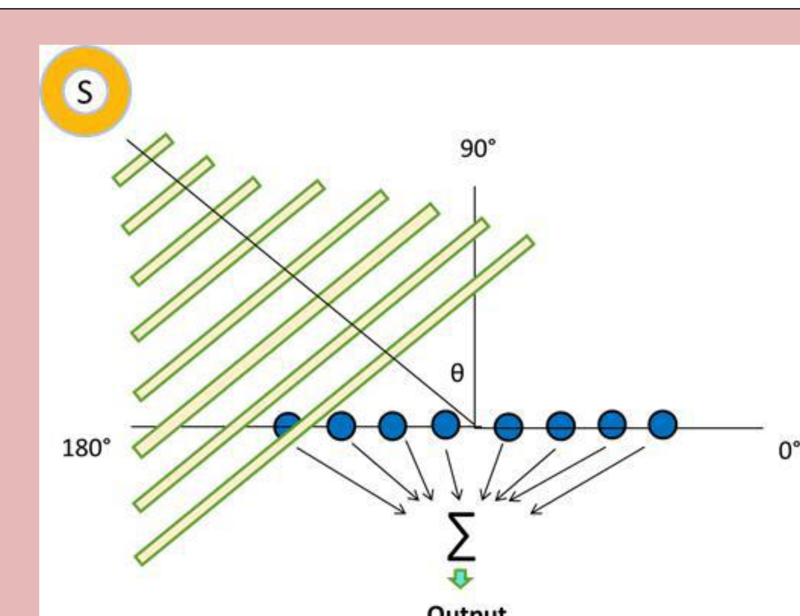
New Frequency domain matrix

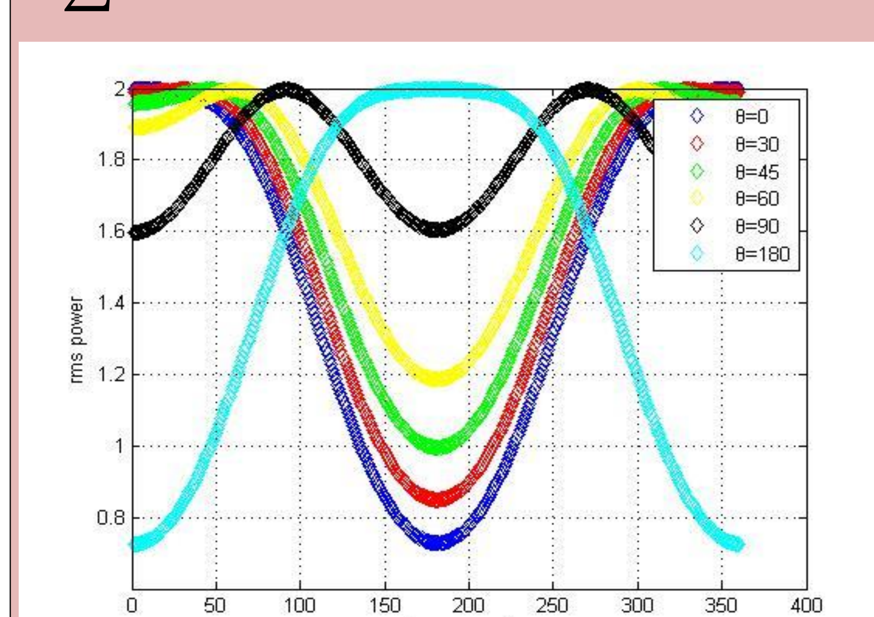
$$\begin{bmatrix} Z_{11}e^{i\phi_{11}-\theta_{1}} & Z_{12}e^{i\phi_{12}-\theta_{2}} & \dots & Z_{1N}e^{i\phi_{1N}-\theta_{N}} \\ Z_{21}e^{i\phi_{21}-\theta_{21}} & Z_{22}e^{i\phi_{22}-\theta_{22}} & \dots & Z_{2N}e^{i\phi_{2N}-\theta_{2N}} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{641}e^{i\phi_{641}-\theta_{641}} & Z_{642}e^{i\phi_{642}-\theta_{642}} & \dots & Z_{64N}e^{i\phi_{64N}-\theta_{64N}} \end{bmatrix}$$

Delay-Sum Classifiers

Since the beamforming creates 360 new frequency domain matrices, the next step is to classify which one can represent the coming angle of signal θ. So, by using delay-sum algorithm, adding all rows of one new frequency domain matrix and calculate their root mean square value. The maximum root mean square value indicate the coming angle of signal.

$$\sum output = s(f) + s(f)e^{-j2\pi\phi_1} + s(f)e^{-j2\pi\phi_2} + \cdots + s(f)e^{-j2\pi\phi_{64}}$$





An example of two elements simulation

When the right angle phase applied to create the new frequency domain matrix,

$$\sum output = s(f) + s(f)e^{-j2\pi\phi_1-\varphi_1} + s(f)e^{-j2\pi\phi_2-\varphi_2} + \cdots + s(f)e^{-j2\pi\phi_{64}-\varphi_{64}}$$

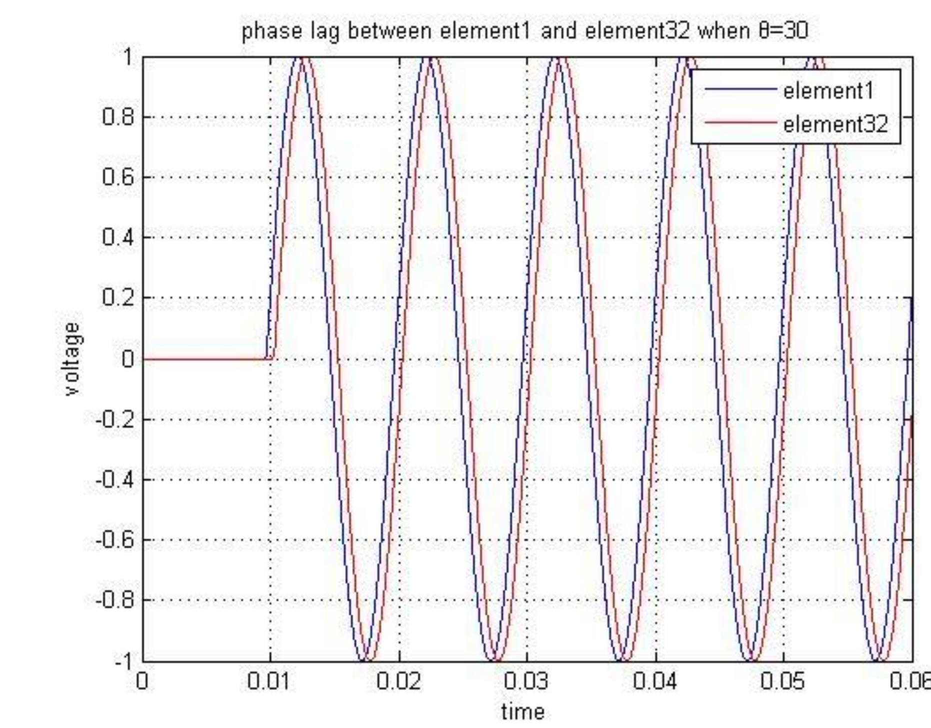
Then $\phi = \varphi$

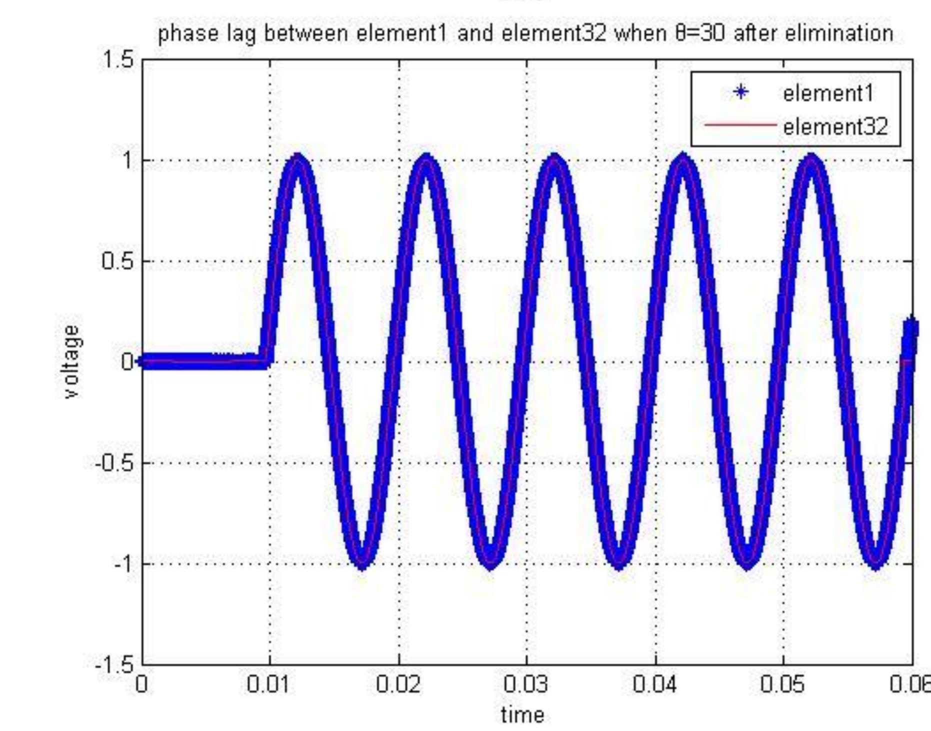
$$\sum output = 64 * s(f)$$

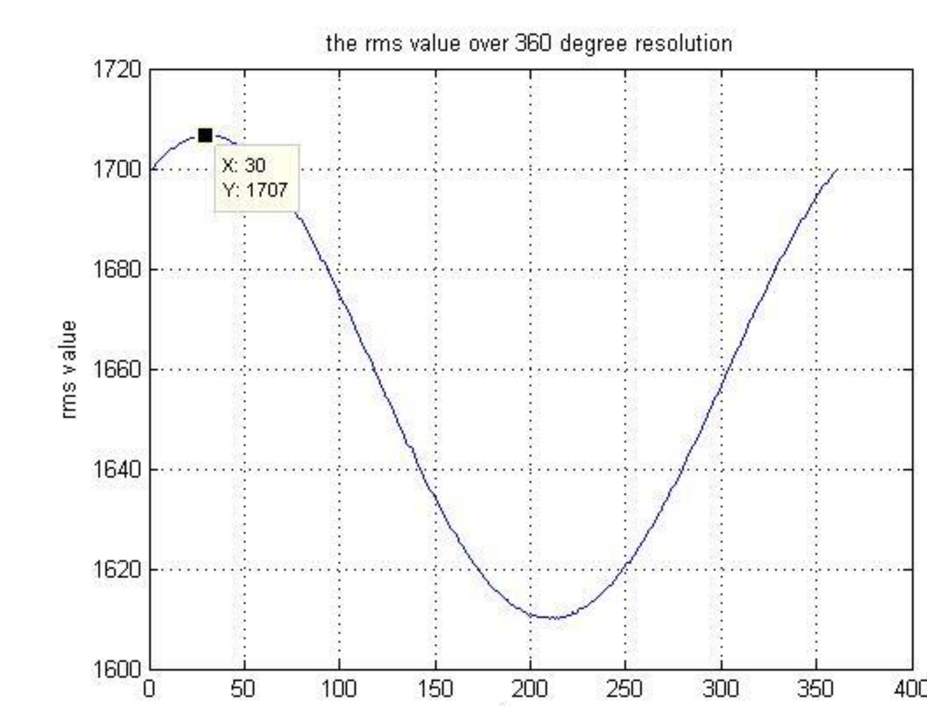
After inverse fast Fourier transform, the maximum rms in real time indicate the approximate angle θ.

Results

The signal is 50Hz sine wave with 10k sample frequency and 6k samples.







Conclusions

The simulation of 64-element works on sine wave signal relatively precise due to the large amount of microphone element. However, because of the one-degree resolution, this algorithm would not get precise angle of signal in the real time.

Also, the elapsed time of this algorithm to compute is 73.235 seconds which is not practical for tracking the real speech signal which need to be optimized.

Further Information

After this project, the research is going to more deep and practical step by setting up the real microphone array to receive the audio signal. Meanwhile, the algorithm will be optimized and become more flexible and practical.

Literature Cited

Van Veen, B.D.; Buckley, K.M. (1988). "Beamforming: A versatile approach to spatial filtering" . IEEE ASSP Magazine 5 (2)

Bronkhorst, Adelbert W. (2000). "The Cocktail Party Phenomenon: A Review on Speech Intelligibility in Multiple-Talker Conditions". Acta Acustica united with Acustica 86: 117–128.

Acknowledgements

Thank you to our research mentors Ed Richter and Arye Nehorai for their guidance and support throughout this project..