Where is the Sound?

By Callaghan Berry, ECE '23

Introduction

Humans have a natural acoustic awareness that allows us to detect the locations of sound sources relative to our location. Whether we realize it or not, our brain can perceive differences in sound magnitude and delay between our two ears and uses this information to understand their origin. People in hard of hearing communities often struggle with this task as the use of hearing aids and cochlear implants can lead to an imbalance between sides of the brain. To aid in this challenge, our project team is designing a device that captures audio from the surrounding area and computes its direction of arrival relative to the user. Audio direction tracking has been an area of research for the better half of a century now, and the fundamental strategy is not too dissimilar from our body. Just like our brain uses the different audio characteristics between ears, audio systems can take advantage of the difference in information between microphones to compute the angle of arrival of incoming sounds. In this paper, we will analyze how microphone arrays can be used to compute the time delay between microphones and how that relates to angle of arrival.

Microphone Array Setup

In order to compute the time delay between microphones, we must first understand how the microphone array is set up. For this paper, we will use a simplified example with two microphones in a line, shown in Figure 1. Utilizing two microphones separated in space gives us the ability to measure the time delay of arrival between microphones, or TDoA.

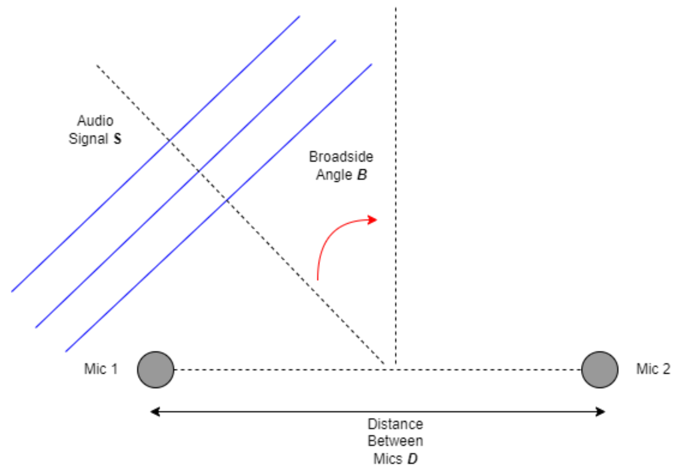


Figure 1: Microphone Array Layout

The two microphone array as shown above consists of two sensors, **Mic 1** and **Mic 2**, separated by a distance **D**. Sound is a wave traveling at 343m/s, and waves take time to propagate between different points in space. Sound waves approaching from the left side will first pass over **Mic 1**, and then over **Mic 2** after some time delay. This delay is directly proportional to the angle the signal arrives on, as well as the distance between traveled between microphones. In this project, we will take advantage of the physical setup of our array to compute the time delay between microphones, and ultimately arrive upon the angle of arrival using constants and trigonometry.

Time Delay Estimation

For this project, our team elected to use a Time Delay Estimation method to compute the angle of arrival. The angle of arrival is directly proportional to the time delay measured between microphones as we will explore, so this method can provide an accurate way to compute angle of arrival. There are many ways to calculate time delay, and in this project, we will explore a method called cross-correlation.

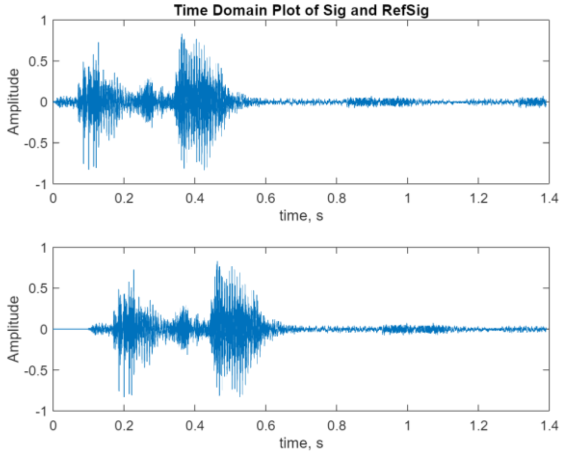


Figure 2: Recordings of Two Microphones Separated in Space

Imagine a single audio source being recorded by two separate microphones, like in Figure 1. Each microphone samples the same audio signal, however since the signal will contact one microphone before the other, a delay is introduced into the second microphone channel, as seen in Figure 2. We would like to recover this delay in order to estimate the angle of arrival, and we can do this by correlating the signals together.

Correlation is a mathmatical measure of likeness between two signals or series of numbers. A high correlation value means the two input signals are highly similar. In the audio example above, we are recording the same exact audio signal in both microphones, and we would expect a high correlation value between the inputs. However, since there is a time delay applied in the second channel, the correlation value is quite low. The cross-correlation method essentially shifts the delayed signal forward in time in small increments until the two signals match, thus being highly correlated. The amount of shift can then be recorded and used as the time delay estimate. If instead we reverse the order of the signals, applying the delay to the opposite channel, the resulting value would have the opposite sign.

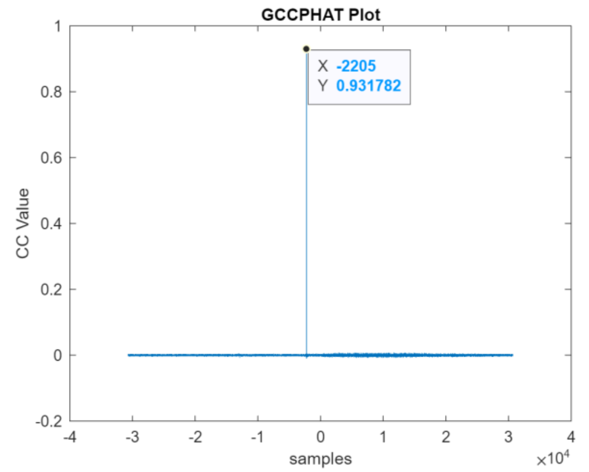
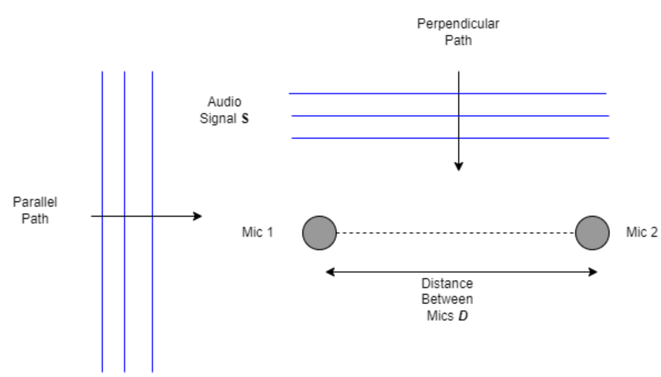


Figure 3: Cross-Correlation values Plotted vs. Sample Delay

In the example audio recordings shown in Figure 2, the second signal was delayed by 100ms, which works out to 2205 samples given the 22050Hz sample rate of the microphones used. Performing the cross-correlation yields a large correlation value after the signal was advanced in time 2205 samples, the same amount as our delay. The conclusion here is that maximum of our cross-correlation function occurs at the time delay difference between recorded inputs. Taking the sample delay and converting it to a time delay gives us the Time Delay of Arrival.

Angle Calculation

Having the time delay is great, but users really care about the angle of arrival. The good news is that we can use trigonometry and the physical characteristics of the microphone array to help us convert the delay into an angle.



$$\theta = arcsin\left(\frac{c\tau}{d}\right) - 90^{\circ}$$

Figure 4: Signal Path Length Differences

The computed time delay is a function of the angle of arrival of the audio signal. If the sound crosses parallel to the microphone array, the time delay will be at its maximum as the parallel path distance is the longest. On the other hand, if the signal arrives perpendicular to the line of microphones, it will cross each sensor with no delay, as the perpendicular path is shortest. The ratio between the actual distance traveled **cτ** and the distance between microphones **d** can ultimately be used to glean the exact angle of arrival by using the trigonometric equations from Figure 4.

Conclusion

Time Delay Estimation via cross-correlation provides a relatively simple way to calculate the angle of arrival. In its simplest form, you only need two microphones to compute the angle, and the underlying code is not computationally complex. It is not a perfect method and has its failures but works well in the right situations.

There are two main shortcomings with the proposed approach. The first issue is that with two microphones set up as in Figure 1, there is no way to detect whether the signal arrived above or below the sensors. For example, if you calculate an angle of arrival of 45 degrees, that angle could be referencing a sound source at positive or negative 45 degrees. This is the biggest limitation with a linear array but can easily be solved with other microphone geometries. For our project, we are using a circular array with 6 microphones to combat this issue. We can look at the angle of arrival relative to all microphones, giving us a much better spatial picture of the audio source, allowing us to report the angle of arrival over the entire 360 degree space.

The second issue is much harder to solve. Cross-correlation searches for the same audio source in each input channel. If multiple sources are present, the input channels are uncorrelated, giving an inaccurate and sometimes random time delay estimate. TDoA approaches are best left for single source cases, or when combined with other methods that can separate out different sources before delay estimation. Other more advanced methods such as beamforming offer much better multi-source detection and could be explored more in the future.

References

1) B. Van Den Broeck, A. Bertrand, P. Karsmakers, B. Vanrumste, H. Van hamme and M. Moonen, "Time-domain generalized cross correlation phase transform sound source localization for small microphone arrays," 2012 5th European DSP Education and Research Conference (EDERC), Amsterdam, Netherlands, 2012, pp. 76-80, doi: 10.1109/EDERC.2012.6532229.
2) C. Junu Jahana, M. Sinith and P. Lalu, "Direction Of Arrival Estimation using Microphone Array," 2021 Fourth International Conference on Microelectronics, Signals & Systems (ICMSS), Kollam, India, 2021, pp. 1-6, doi: 10.1109/ICMSS53060.2021.9673617.