Acoustic Source Localization by using Time Difference of Arrival Estimation

Ajay Verma[1] Ritu[2]
[1]Student [2]Assistant Professor
[1,2]Department of Electronics & Communication Engineering
[1,2]Guru Jambheshwar University of Science and Technology, Hisar

Abstract— Acoustic Source localization is the technique which is used to locate the sound source in a given environment. The sound source can be located with the help of sound signal recorded by the microphone. There is one original signal that comes straight to the microphone and other signals are reflections of the original signal reflected from the walls of the room. In this paper, we proposed a method of locating the sound source using the microphone array. We locate the sound source by calculating the time difference using time difference of arrival method. We process this time difference by using subband and broadband approach. The results of this paper show that we can find the position of sound source by calculating the angle towards the sound source with an accuracy of about 60-70%.

Key words: Acoustic Source Localization, Time Difference of Arrival Estimation

I. INTRODUCTION

Today, the microphones which can adjust electronically have some new variety of applications, such as the interaction of human being with the computer and the conference in intelligent room. The microphone array can be aimed electronically on a sound source and the noise from the surrounding environment can be attenuated.

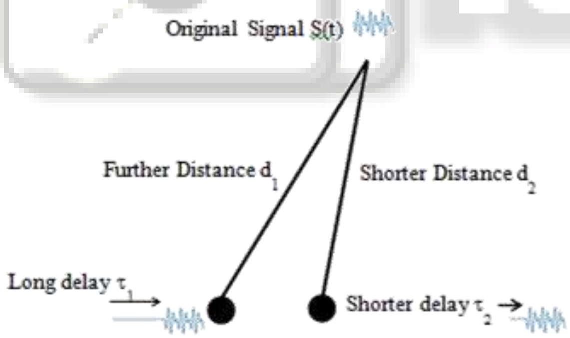

Fig. 1: Example of acoustic source localization

It can be used to control the computer with the help of sound recognition system by capturing the sound signal using the microphone array. It can be used in video conferences.

A large number of sound localization techniques and algorithms are present with different accuracy and features. In this paper, we limit our attention on the algorithms of sound source localization by time difference of arrival estimation [3] using subband and broadband approach. These microphone arrays generally have the following set of characteristics.

- The number of microphones is limited in a linear array. For better result we use the circular microphone array.
- Both directional and unidirectional microphones are popular. Here, we use the directional microphone array for the better results and accuracy.
- Microphones are placed at an acceptable distance to capture the signal efficiently and to find the direction of source more precise.
- In the conference room, there are reflections from the walls and other objects. We have to find the location of the source with the help of these reflected delayed signals.

Based upon the above characteristics the choices for the localization of sound source are limited. Due to this reason the algorithm based upon the correlation, is useful. Correlation can be done between the two signals reached to the side by side microphones and the location where the correlation is high the angle of arrival of signal can be determined. The algorithm also has to be robust to remove the noises.

II. SOURCE LOCALIZATION METHODS

The procedure of finding the location of an acoustic source relative to some reference is known as acoustic source localization. Acoustical sound source present can be located with the knowledge of time difference of arrival of signals at two side by side microphones [1]. The speed of the sound varies from medium to medium. Over the last some years the localization techniques based upon the time difference of arrival are very popular and mostly used.

A. *Time Difference of Arrival*

It aims at the relative time difference of arrival among spatially separated microphone sensors. This technique played an important role in locating the sound source in RADAR, SONAR etc.

Suppose the source is emitting a sound signal and a two-microphone array is observing signals and the signal received by the microphones will be distorted due to both the characteristics of the room and the environmental noise. Also, due to the distance between the two microphones in the array, there will be a measurable time difference between observations of the signal at each microphone. This is referred to as the time difference of arrival (TDOA) of signals between the microphones. The time of a signal is the amount of time required for a signal to propagate from the source to the microphone in an array.

$$\tau_i = \frac{|x_s - x_i|}{v}$$

Where x_s & x_i are the spatial positions of the source and the microphone, and v is the speed of sound in air (m/s). As the sound wave travels through different paths to reach spatially separated microphones, the signal from the source reaches the microphones at different instants of times. The TDOA for a given pair of microphones and the source is defined as the time difference between the signals received by the two microphones. It is computed using the spatial positions of the source and microphone.

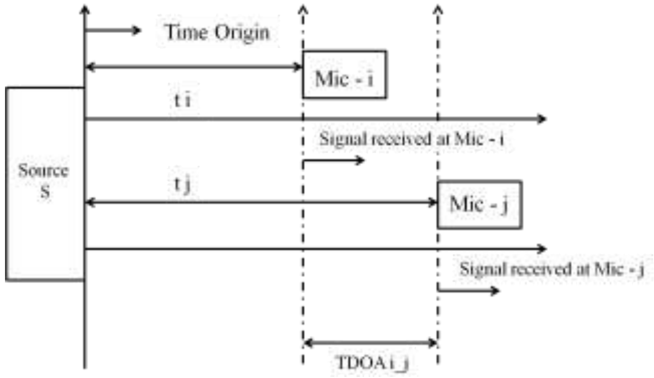

Fig. 2 : Schematic showing the time delay of the arrival of signal

In Fig. 2, ti is the time taken by the signal to reach microphone i while tj is the time taken by the signal to reach microphone j. The Time Difference of Arrival between the i^{th} and j^{th} microphone when the source 'S' is excited, is given by $TDOA_{ij}$.

$$TDOA_{ij} = \frac{\|m_i - s\| - \|m_j - s\|}{v}$$

A TDOA estimate can be obtained by finding the 'distance between microphones' that maximizes the cross correlation between the two microphone signals. TDOA's are then converted to angle of arrivals.

B. *Cross Correlation*

Cross correlation [6] is a general signal processing technique that can be applied to determine the time delay between two signals, e.g. for determining time delays for the propagation of acoustic signals across a microphone array. The cross correlation between two signals is the measure of similarity between one signal and time delayed version of another signal. Cross correlation between two signals explains how much one signal is related to the time delayed version of another signal. The cross correlation between two signals x(t) and y(t) is defined as

$$R_{xy}(\tau) = \lim_{T\to\infty} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t)y(t-\tau)dt$$

After calculating the cross-correlation between the two signals, the maximum of the cross-correlation function indicates the point in time where the signals are best aligned, i.e. the time delay between the two signals is determined by the argument of the maximum, or arg max of the cross-correlation, as in

$$\tau_{delay} = \arg\max_{\tau}(R_{xy}(\tau))$$

Where τ_{delay} is called the delay parameter at which the value of cross correlation is maximum. Cross correlation represents the overlapping area between the signals and the delay parameter determines the maximum possible correlation between two signals.

III. DIRECTION OF ARRIVAL

From the known microphone positions and the geometry, the Direction Of Arrival (DOA) [4, 5] of the signal can be obtained from the measured time-delays. The time-delays are estimated for each pair of microphones in the array. Then the best estimate of the DOA is obtained from time-delays and the array geometry.

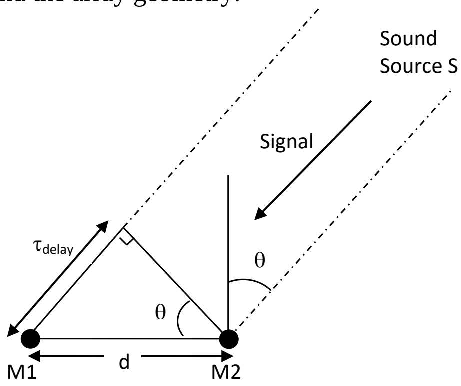

Fig. 3: Two dimensional microphone array

The angle of arrival can be found with the help of Triangulation Method which gives the following formula: -

$$\theta = \sin^{-1}\left[\frac{\tau_{delay} * C}{f_s * d}\right]$$

Where

τ_{delay} = Time Difference between the Microphones at which the cross correlation is maximum.

C = Speed of the Sound (it can vary from medium to medium)

f_s = Signal Sampling Frequency

d = Inter Microphone Distance

IV. PROBLEM STATEMENT

Given a set of two microphones in known locations, the goal is to estimate the direction of the sound source. It is to be assumed that the sound source is present in a defined coordinate system.

We know the number of microphones present and that single sound source is present in the system. Given that the sound source is stationary. The sound source is excited using a broad band signal with defined bandwidth and the signal is captured by each microphone. The Time Difference of Arrival (TDOA) is estimated from the captured audio signals.

The TDOA for a given pair of microphones and sound source is defined as the difference in the time taken by the acoustic signal to travel from the sound source to the microphones. We assume that the signal emitted from the speaker does not interfere with the noise sources. The direction measurement algorithm finds the direction of the sound source according to the estimated value of TDOA. Based on the measurements we estimate the source direction knowing the geometry of the microphone array.

V. RESULTS AND ANALYSIS

The results of source localization algorithm are presented as shown below with the given simulation parameters.

A. *Simulation Parameters:*

The root mean square error(RMSE) can be found out by the following formula: -

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n}(\hat{\theta}_t - \theta)^2}{n}}$$

Where

$\hat{\theta}_t$ = the angle calculated by the algorithm
θ = the original angle
n = total no of iteration

The parameters that are used for the calculation of sound source direction is as follows: -

PARAMETER	VALUE
Sampling Frequency(Hz)	48000
Window Size (seconds)	0.03
Lowest Center frequency(Hz)	1000
Highest Center frequency(Hz)	12000
No. of competing signals	2
Minimum distance between competing signals(degree)	5
No. of rooms	5
No. of acoustic mixtures	5
Absolute error boundary (degree)	10

The following diagrams show the cross correlation pattern of broad band and subband approach. According to the correlation pattern, where the correlation is maximum the direction of source can be estimated. The concept of subband approach and broadband approach is defined as under:-

In signal processing, Sub band coding (SBC) is any form of transform coding that breaks a signal into a number of different frequency bands and encodes each one independently. This decomposition is often the first step in data compression for audio and video signals.

Broadband Approach: - A broadband signal is the signal which does not occupy the minimum range, but instead a higher range, 1MHz to 3MHz, for example. A wire may have only one baseband signal, but it may hold any number of broadband signals, because they can occur anywhere in the spectrum.

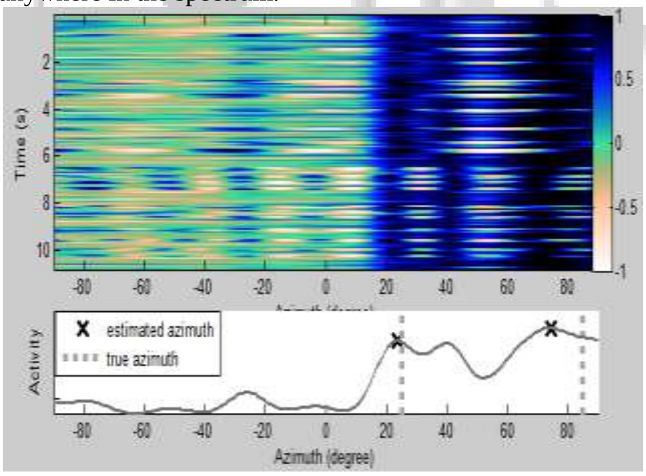

Fig. 4: Broadband correlation pattern

B. Subband Approach:

A subband signal is the signal that occupies the frequency range from 0Hz up to a certain cut-off. It is called the subband because it occupies some part of the spectrum. Here the signal part which is considered is from 100Hz to 12000Hz

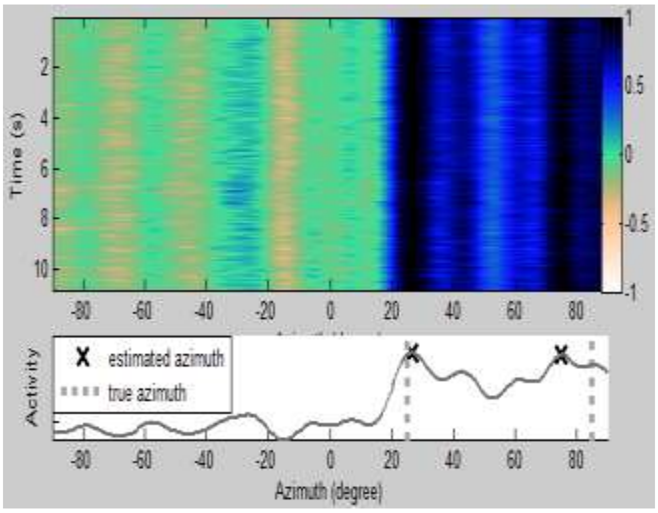

Fig. 5 : Subband correlation pattern

From the correlation pattern one can see that the results in the broadband approach are not so clear as in the subband approach. Also, the value of root mean square error is more in the broadband approach as compared to the subband approach.

The given below diagram shows the estimation of azimuth angle that shows the direction of the sound source.

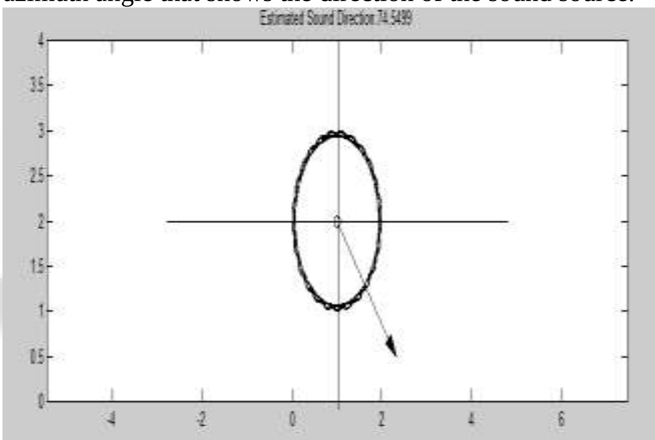

Fig. 6 : Estimated sound direction

This experiment is performed using many rooms. Every room gives different results. The accuracy varies with the rooms. Also, the root mean square error varies with different room. The following diagrams show the results for various rooms for both subband and broadband approach. The level of accuracy is better for the broadband approach.

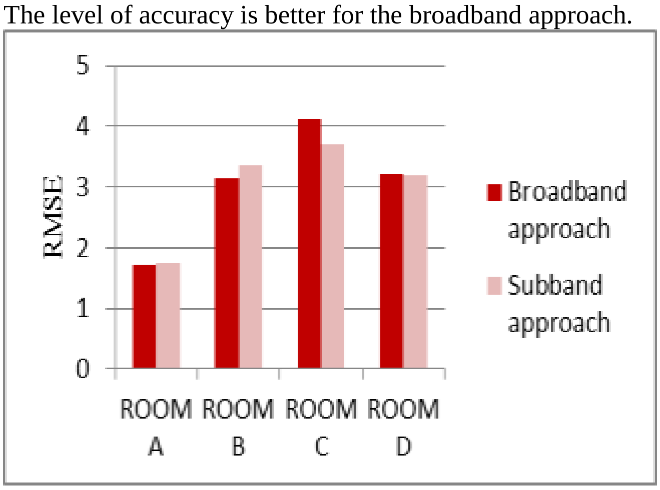

Fig. 7 : Root mean square error for different rooms

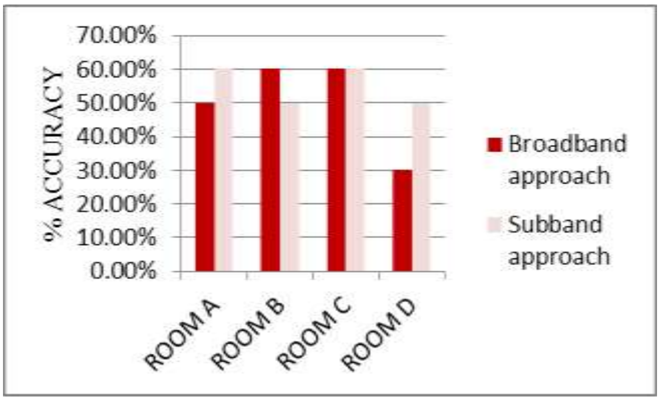

Fig. 8 : Percentage accuracy of results for different rooms

VI. CONCLUSION

This paper presented a different outline for sound source localization and beam forming with microphone arrays. The main contribution of this paper is the novel adoption of an efficient algorithm which works very well in practice. Our results show that our microphone array system is able to estimate the position of an acoustic sound source by estimating the angle to the source with reasonable accuracy about 70-80%. However, many limitations occurs which prevent us from estimating more accurately the location of the sound source. One such limitation is the reflection of the signal from the walls and the other limitation is the surrounding environmental noise. These parameters limit the accuracy.

REFERENCES

[1] Paulose, Smitha, Elizabeth Sebastian, and Babu Paul. "Acoustic source localization." International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 2, no. 2 (2013): 933-939.
[2] Liu, Kaikai, Xinxin Liu, Lulu Xie, and Xiaolin Li. "Towards accurate acoustic localization on a smartphone." In INFOCOM, 2013 Proceedings IEEE, pp. 495-499. IEEE, 2013.
[3] Yost, William A., Louise Loiselle, Michael Dorman, Jason Burns, and Christopher A. Brown. "Sound source localization of filtered noises by listeners with normal hearing: A statistical analysis." The Journal of the Acoustical Society of America 133, no. 5 (2013): 2876-2882.
[4] Wenyi Zhang and Bhaskar D. Rao, “Two microphone based direction of arrival estimation for multiple speech sources using spectral properties of speech”, IEEE International Conference on, Speech and Signal Processing pp.2193-2197,2009.
[5] Bilik, I. "Spatial compressive sensing for direction-of-arrival estimation of multiple sources using dynamic sensor arrays." Aerospace and Electronic Systems, IEEE Transactions on 47, no. 3 (2011): 1754-1769.
[6] Van Den Broeck, Bert, Alexander Bertrand, Peter Karsmakers, Bart Vanrumste, and Marc Moonen. "Time-domain generalized cross correlation phase transform sound source localization for small microphone arrays." In Education and Research Conference (EDERC), 2012 5th European DSP, pp. 76-80. IEEE, 2012.

