Acoustic Source Localization

Smitha Paulose[1], Elizabeth Sebastian[2], Dr. Babu Paul[3]

Assistant Professor, Dept of Electrical and Electronics, Mar Athanasius College of Engg., Kerala, India[1]

Associate Professor, Dept of Electrical and Electronics, Mar Athanasius College of Engg., Kerala, India[2]

Assistant Professor, Dept of Electrical and Electronics, Mar Athanasius College of Engg., Kerala, India[3]

Abstract: Source Localization is a very well established technique that has a wide range of applications from remote sensing to the Global Positioning System. Sound source localization techniques are used in commercial applications like improving speech quality in hands free telephony, video conferencing to military applications like SONAR, surveillance systems and devices to locate the sources of artillery fire. A method is proposed to localize an acoustic source within a frequency band from 100Hz to 4 KHz in two dimensions using microphone array by calculating the direction of arrival (DOA) of the acoustic signals. Direction of arrival (DOA) estimation of acoustic signals using a set of spatially separated microphones uses the phase information present in signals. For this the time-delays are estimated for each pair of microphones in the array. From the known array geometry and the direction of arrival, the location of source can be obtained.

Keywords: Beam forming, Time Difference of Arrival (TDOA), Direction of Arrival (DOA), Cross Correlation.

I. INTRODUCTION

People are generally able to discern the direction that a sound is coming from using two ears. The combination of the slightly different signals that arrive at the ears enables us to deduce, intuitively, the direction of the sound. Similarly, a biologically inspired sound localization system can be built by making use of an array of microphones, which are hooked up to a computer. In addition, such a system of microphones can be made to extract any particular sound from a mixture of sounds produced simultaneously by several sources. The spatial location of a sound source can be determined based on multiple observations of the emitted sound signal. The mathematical details of these applications are basically variations of a triangulation scheme, using multiple sensors, which detect a signal emitted by the source that is to be localized, or by using multiple emitters, whose signals are sensed by the sensor (e.g. GPS receiver).The most representative methods frequently used in conjunction with sound source localization are intensity difference between microphones, Beam forming and Time Difference of Arrival (TDOA). In this current work, the phase information present in signals and the Direction of arrival (DOA) of the signals at the microphone arrays are used to find an estimate of the source location.

II. ACOUSTIC SOURCE LOCALIZATION

The process of determining the location of an acoustic source relative to some reference frame is known as acoustic source localization. Acoustic source present in the near-field can be localized with knowledge of the time difference of arrivals (TDOAs) measured with pairs of microphones. The speed of sound in the medium in which the acoustic source is present is known. With continued investigation over the last two decades, the time delay estimation (TDE) based localization has become the technique of choice.

A. Time Difference of Arrival

Time-Delay Estimation (TDE), which aims at measuring the relative time difference of arrival (TDOA) among spatially separated sensors, has played an important role in radar, sonar, and seismology for localizing radiating sources.

Suppose the source is emitting a sound signal S(t), and a two-microphone array is observing signals $X_1(t)$ and $X_2(t)$. The signal received by the microphones will be distorted due to both the acoustics of the room and the ambient noise in the environment. Furthermore, 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) , τ , between the microphones. The time τ_i of a signal S *(t)* is the amount of time required for a signal to propagate from the source to the ith microphone in an array.

$$\tau_i \;=\; \frac{\|x_s - x_i\|}{v} \qquad\qquad \text{Eqn(2.1)}$$

Where x_s and x_i are the spatial positions of the source and the ith 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.
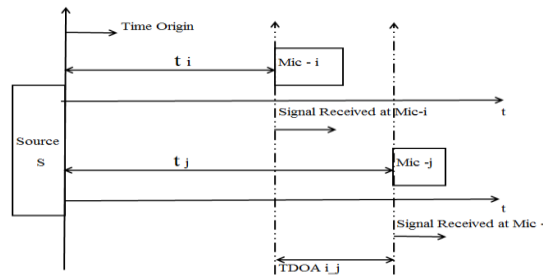

Fig .2.1 Schematic depicting the signal receiving time and TDOA

In figure 2.1, t_i is the time taken by the signal to reach microphone i while t_j is the time taken by the signal to reach microphone j. The Time Difference of Arrival between the ith and j^{th} microphone when the source 'S' is excited is given by $TDOA_{ij}$.

$$TDOA_{ij} \;=\; \frac{(\|m_i - s\| - \|m_j - s\|)}{v} \qquad\qquad \text{Eqn(2.2)}$$

A TDOA estimate can be obtained by finding the $'\tau'$ that maximizes the cross correlation between the two microphone signals. TDOA's are then converted to path differences and angle of arrivals.

B. Direction of Arrival

When the microphones are spatially separated, the acoustic signals arrive at them with differences in time of arrival. From the known array geometry, the Direction of arrival (DOA) 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.
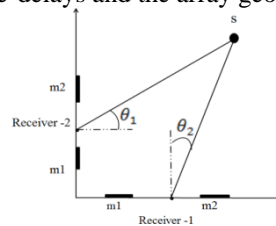

Fig.2.2. Two dimensional microphone arrays

In figure 2.2 receivers 1 and 2 consists of a pair of sensors m1 and m2 which are placed along the X and Y axes respectively. Signals coming from source S, reach the receiver 1, at an angle θ_1 with respect to line perpendicular to X- axis. In receiver 1, the extra distance travelled by source signal to reach m1 as compared to m2 will be $dcos\theta$.

C. Cross Correlations

Cross correlation is a routine signal processing technique that can be applied to find the time delays between two copies of a signal registered at a pair of microphones. 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

$$R_{xy}(\tau) = \lim_{T \to \alpha} \int_{-\frac{T}{2}}^{\frac{T}{2}} x(t) \, y^*(t - \tau) dt \qquad\qquad \text{Eqn (2.3)}$$

Where $'\tau'$ is called the delay parameter. Cross correlation represents the overlapping area between the signals and the delay parameter determines the maximum possible correlation between two signals.

III. SYSTEM DESCRIPTION

A method is proposed to localize an acoustic source within a frequency band from100Hz to 4 kHz in two dimensions using microphone array by calculating the time difference of arrival (TDOA) of the acoustic signals. The Time Difference of Arrival (TDOA) is estimated from the captured audio signals. In this paper, Time difference of arrival (TDOA) based source localization is done based on a two-step procedure. The first stage involves estimation of phase difference between the signals reaching the receivers using the time delay estimation techniques. The estimated phase differences are then transformed into range difference measurements between sensors and the direction of arrival (DOA) of the signal is computed. The range difference or the path difference can also be calculated by directly from the time delay using cross correlation methods. The second stage utilizes efficient algorithms to calculate the location of the source by using the angle of arrival of the signals.

A. Assumptions

The following conditions are assumed under which location of sound source is estimated: (1) Single sound source is present, which is omni-directional and infinitesimally small. (2) Reflections from the bottom of the plane and from the surrounding objects are negligible. (3) No disturbing noise sources are contributing to the sound field. (4) The source to be located is assumed to be stationary during the data acquisition period. (5)Microphones are assumed to be both phase and amplitude matched and without self noise. The positions of acoustic receivers are known. (6) The change in sound velocity due to change in pressure and temperature are neglected. The velocity of sound in air is taken as 330m/sec. (7) If the distance between the source and the receivers is greater than the distance between the two receivers the sound waves reaching the receivers can be considered to be planar rather than spherical.
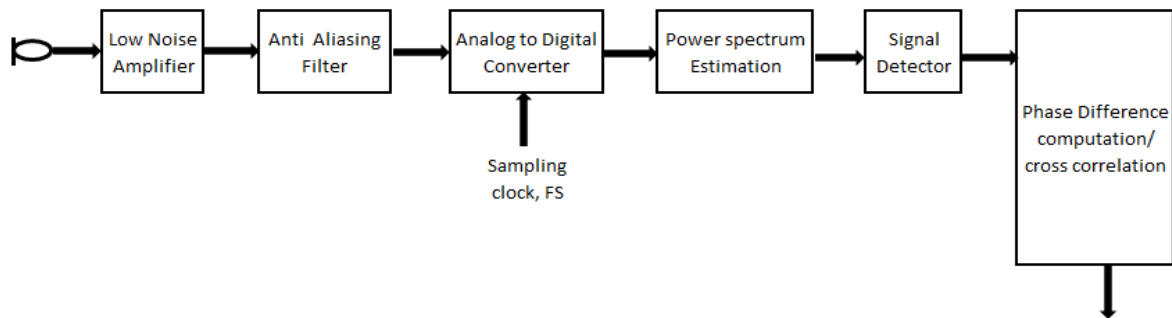

Fig 3.1 Block Diagram of System Description

Figure 3.1 shows the block diagram of the system description. The acoustic signal which is a physical signal is converted into an electrical signal using a transducer. Microphone is an acoustic to electric transducer or sensor that converts sound into an electrical signal. An amplifier is used for increasing the power of an electric signal by use of an external energy source. The dynamic range of an amplifier decides the output levels of an ADC. Filters play a key role in virtually all sampled data systems. Most analog to digital converters (ADCs) are preceded by a filter which removes frequency components that are beyond the ADC's range. In a sampled data system, frequency components greater than half the sampling rate shifts into the frequency band of interest. An anti-aliasing filter is used before a signal sampler, to restrict the bandwidth of a signal to approximately satisfy the sampling theorem. Anti aliasing Filter, is designed to minimize aliasing effects while the required signal is sampled. An analog-to-digital converter (ADC) is a device that converts a continuous signal to a discrete time representation in digital form. The frequency on which the sampling will occur is called sampling rate. The sampling rate should be at least twice the maximum frequency component of the signal of interest. The maximum frequency of the input signal should be less than or equal to half of the sampling rate. The digital output will be a two's complement binary number that is proportional to the input.

IV. SIGNAL DETECTION AND SOURCE LOCALIZATION

 From the different signals reaching the receivers, the signals of sufficient magnitudes and those coming within the specified frequency range are considered for source localization. Power Spectral estimation techniques are employed in the digitized signal to detect the source in spectral domain. The power density spectrum of a signal shows the distribution of power of a signal among the various frequency components.

To perform frequency analysis on a discrete time signal the signal is converted to its equivalent frequency domain representation, which is given by the Fourier Transform of the signal. Using FFT, the magnitude of the signal and its corresponding frequency can be found. The signals whose power spectrum crosses a certain limit and that comes in the specified frequency range are considered for source localization. Observing the power spectrum of the signals, the frequency at which the sharp peaks occur can be calculated by the equation,

$$f = n \times \frac{f_s}{N} \qquad\qquad \text{Eqn (4.1)}$$

$f_s - sampling\ frequency$
$f - frequency\ at\ which\ the\ peak\ occurs$
$N - number\ of\ FFT\ points$
$n - number\ of\ the\ particular\ frequency\ bin$

A. Phase Difference Computation

The signals reaching the microphones at different instants of time will also be having a phase difference between them. Let $\sin(wt)$ and $\sin(\omega t + \emptyset)$ are two signals reaching the receivers. Product of these signals will be,

$$\sin(wt) * \sin(\omega t + \emptyset)\ \ = \frac{1}{2}\cos\emptyset - \frac{1}{2}\ \cos(2\omega t + \emptyset) \qquad\qquad \text{Eqn (4.2)}$$

An FIR filter can be used to filter out the ac components. The remaining part provides the phase difference $'\emptyset'$ between the two signals. The phase differences between the signals can be used to obtain the time difference of arrival (TDOA). The difference in the propagation delay and the acoustic velocity in air are known. From this the path difference of the acoustic waves and the direction of arrival (DOA) with respect to the microphone pair can be found.

$$\emptyset = \frac{2\pi}{T} * \Delta t \qquad\qquad \text{Eqn(4.3)}$$

$$\Delta t = \frac{\emptyset}{2\pi f} \qquad\qquad \text{Eqn (4.4)}$$

$$\Delta d = \Delta t * v \qquad\qquad \text{Eqn(4.5)}$$

$$\therefore \quad \Delta d = \frac{\emptyset}{2\pi f} * v \qquad\qquad \text{Eqn(4.6)}$$

$$cos\theta = \frac{\Delta d}{R_{12}} \qquad\qquad \text{Eqn(4.7)}$$

$$\therefore \ \theta = \cos^{-1}\left(\frac{\emptyset}{2\pi f} * \frac{v}{R_{12}}\right) \qquad\qquad \text{Eqn(4.8)}$$

$\emptyset - phase\ difference\ between\ the\ signals$
$\Delta d - \ path\ delay,$
$\Delta t - time\ delay,$
$\theta - angle\ of\ arrival$
$v - velocity\ of\ sound,$
$f - frequency\ of\ the\ incoming\ signal,$
$R_{12} - distance\ between\ receivers\ 1\ and\ 2$

Hence by calculating the phase difference of two signals the angle of arrival or the direction of arrival (DOA) $'\theta'$ can be obtained. A microphone pair is coupled with the existing pair of microphones to obtain another value of the direction of arrival. Source location can be estimated using these angles assuming source is in two-dimensional plane.

B. *Source Location Estimation*

Microphones are placed in an L shaped array with a distance of *d* between the adjacent sensors. Receiver 1 is placed at X- axis with co-ordinates (5,0) and receiver 2 is placed at Y axis having co-ordinates(0,5). If θ_1 is the angle of arrival of the signal at the first receiver and θ_2 is the angle of arrival at the second receiver, the two dimensional location of source (x, y) is given as

$$x = \frac{5(\tan\theta_1 + 1)}{\tan\theta_1 - \tan\theta_2}$$ Eqn (4.9)

$$y = \frac{5\tan\theta_1 (1 + \tan\theta_2)}{\tan\theta_1 - \tan\theta_2}$$ Eqn (4.10)

V. EXPERIMENTAL DETAILS

Initially the scenario is created using MATLAB and the source location is computed in two dimensions from the phase difference and cross correlations of the signals arriving at the sensors. Secondly Cross-correlation methods are employed to calculate the time delay of the signals and using this value the angle of arrivals are computed.

A. *Simulation Results*

The signals reaching the two receivers are artificially stimulated in a MATLAB environment assuming a sampling frequency of 10000Hz and signal frequency varying from 100Hz to 4 kHz. By cross-correlating and finding the mean value, the phase difference between the signals are computed and the path delay is found. From the path delays direction of arrival (DOA) and location of source is estimated in two dimensions using the equation (4.9) and equation (4.10), assuming the receivers are placed in an L shaped array at known locations

The following table shows the values obtained by stimulating the environment in MATLAB. The measured phase differences of signals reaching receivers 1and 2 (ph diff1) and receivers 3 and 4 (ph diff2) are shown. From these the angle of arrivals and source location estimates are obtained. The experiment was carried out for various signal frequencies from 100Hz to 3 kHz and the results for 100Hz and 2200Hz is shown here.

TABLE I MEASURED PHASE DIFFERENCES, ANGLE OF ARRIVALS ANGLE OF ARRIVALS, LOCATION ESTIMATES FOR 100Hz ESTIMATES FOR 100Hz

TABLE II MEASURED PHASE DIFFERENCES, OF ARRIVALS, LOCATION

Sl. no	Ph diff 1, ph diff2	AOA1,AOA2	(X, Y)
1	1.63, 0.94	30.9, 60.31	(1.4 , 3.7)
2	1.2, 0.01	51.2, 90	(2.3 , -1.9)
3	1.7, 1.44	27, 40.6	(5 , 4.2)
4	0.5, 1.75	74.7, 22.5	(2.4 ,4.7)
5	0.3, 1.88	80.5, 8.1	(-7.87, 24.3)

Sl. no	Ph diff1, ph diff2	AO A1 ,AOA2	(X, Y)
1	1.87 , 1.88	87.5 , 87.4	(-4.4, -3.6)
2	1.1 , 0001	88.5 , 90	(1.4 , -1.1)
3	0.37 , 0.38	89.5 , 89.5	(α , α)
4	1.5 , 1	87.9, 88.6	(-1.07 , 0 .31)
5	0.63 , 2.5	89.1 , 86.5	(1.6 , -7.1)

Hence by computing the phase difference between the signals arriving at the sensors which are spatially separated, it is possible to find the angle of arrivals and the location of the source. But while computing the location estimates using the phase difference between the signals reaching arrays, ambiguity arises when the frequency exceeds 3.3 KHz. If $'d'$ is the distance between the two sensors of a receiver, to avoid ambiguity

$$d < \frac{v}{f}$$ Eqn (6.1)

$v - velocity\ of\ sound\ given\ as\ 330m/s$ and $f - frequency\ of\ the\ incoming\ signal$

For a distance of 1m between the sensors, the maximum frequency that provides a unique solution to direction of arrival comes around 3.3KHz. For frequencies above 3.3 KHz the measured phase differences will be ambiguous. Hence 3.3 KHz has to be kept as the upper cut off frequency while finding the source location using phase difference computation in MATLAB.

This disadvantage is overcome by cross-correlating the signals in real time. As cross correlations have no frequency constraints, by cross-correlating the signals reaching the receivers the time delay between the signals can be found. The lag at which the cross-correlation function has its maximum value is taken as the time delay between the two signals. From the time delay of the signals, angle of arrival is computed.

B. Real Signal Experiments

The real signal experiments are carried out using python programming language in real time and the signals are detected using a pair of microphones. PC sound card is used to acquire the signals from the microphones. The captured signals are processed inside the PC and the digitized signals are used for source localization.
Figure 5.2 shows the data flow diagram by which the angles of arrival of the signals are computed. The dominant frequencies present in the detected signals are selected from the maximum components of the Fast Fourier Transform (FFT) results. When the power spectrum of the signals crosses a certain limit, these dominant frequencies are detected. The envelope areas of the detected signals are cross correlated taking a portion of signals arriving at a time. From cross-correlation the time delay of arrival between the detected signals is found and substituted in the equations (5.2) and (5.3) to find the direction of arrival (DOA) of the detected signals.
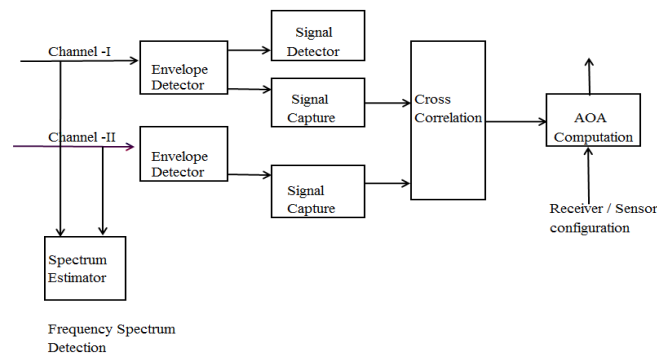

Fig 5.2 Data Flow Diagram

From Cross correlations the time delay is obtained in terms of number of samples (Δn). Time Delay is given as,

$$\therefore \Delta t = \frac{\Delta n}{f} \hspace{3cm} \text{Eqn (5.2)}$$

Using equations (2.3) and (4.5), Direction of Arrival (DOA) of the signals is given as

$$\theta = \cos^{-1}\left(\frac{\Delta n}{f} * \frac{v}{d}\right) \hspace{3cm} \text{Eqn (5.3)}$$

C. Results Based On Real Time Signals

During measurement, the source is kept stationary and at known positions with respect to the sensors. Experiments were done using a pair of microphones which are kept at a spacing of 1m. Velocity of sound is 330m/sec. Hence the worst time delay at which the signals reach the receivers can be 3milliseconds. The signals are sampled in real time at a sampling rate of 44.1 kHz. The time delay in msec and the angle of arrival of the signals are obtained from the peak value of the cross correlated signals using equations (6.2) and (6.3) keeping the source at various locations.

TABLE III ESTIMATED TDOA AND DOA FOR 45 DEGREES

Sl.No	Delay in samples (Δn)	Time delay msec (Δt)	DOA in degrees (θ)
1	39	0.406	48.295
2	18	0.188	66.452
3	23	0.239	55.465
4	20	0.208	48.057
5	22	0.229	58.662
6	24	0.25	75.267
7	21	0.2187	49.86
8	22	0.229	54.662
9	18	0.187	48.452
10	19	0.197	46.255

TABLE IV ESTIMATED TDOA AND DOA FOR 90 DEGREES

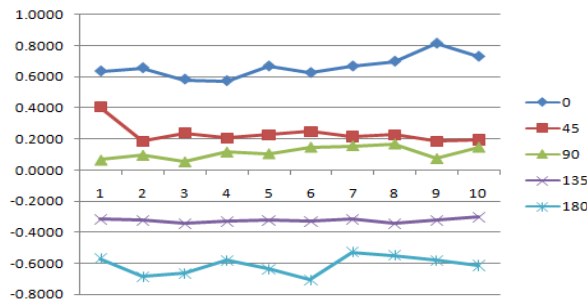| Sl.No	Delay in samples (Δn)	Time delay msec (Δt)	DOA in degrees (θ)
1	6	0.0625	88.818
2	9	9.38E-05	88.227
3	5	5.21E-05	89.010
4	11	0.1146	87.832
5	10	0.1042	88.03
6	14	0.1458	87.241
7	15	0.1563	87.044
8	16	0.1667	86.847
9	7	7.29E-05	88.621
10	14	0.1458	87.241

Fig.5.4. Estimated time delay (msec) for various angles

From the values obtained the variation in samples of delay and time delay with respect to each positions are plotted. Comparing the values obtained it is found that the time delays for positions 45 degrees and 90 degrees are minimum. At 90 degrees the signals are reaching the receivers almost at the same time and so the Time Difference of Arrival of the signals at 90 degrees is reduced to a minimum.

VI. CONCLUSION

A number of complications limit the potential accuracy of the system. Some of these are due to physical phenomena that can never be corrected, and others are due to inherent errors built into the processing, due to the design of the system. The error in DOA estimation can be reduced with increase in frame size of signal acquired. Increase in number of frequency components may also improve the DOA estimation.

REFERENCES

[1] Bob Mungamuru And Parham Aarabi, *Enhanced Sound Localization*, IEEE Transactions On Systems, Man, And Cybernetics – Part   B:Cybernetics, VOL. 34, NO. 3, pp.1526-1540, Jun 2004

[2] Ashok Kumar Tellakula, *Acoustic Source Localization Using Time Delay Estimation,* Supercomputer Education And Research Centre Indian   Institute of Science, Thesis Report, pp. 1-82, Aug 2007

[3] Ying Yu And Harvey F. Silverman, *An Improved TDOA Based Location Estimation Algorithm For Large Aperture Microphone Arrays*, IEEE Explore pp. 77-80,Sept 2009

[4] Beom-Cheol, Park Keun, Chang Kwak, Kyu-Dae Ban And Ho- Sub Yoon Computer Software And Engineering Human Robot Interaction Team, University Of Science and Technology Intelligent Robot Research Division Korea, *Sound Source Localization Based On Audiovisual   Information for Intelligent Service Robot*s, Proceeding IEEE International Workshop in Robot and Human Interactive Communication, Dec 2002

[5] Micheal Shapiro Brandstein, Ph.D, Brown University S.M.E.E Massachusetts Institute Of Technology, *A Frame Work For Speech Source Localization Using Sensor Arrays,* Thesis Report, May 1998

[6] Jingcho Chen, Jacob Benesty and Yiteng Huang, *Time Delay Estimation Using Spatial Correlation Technique,* International Workshop on Acoustic Echo Noise and Noise Control , Kyoto, Japan pp. 207-210 Sept 2003