An Experimental Approach to Sound Localization in 2-D and Numerical Simulation

Shakti Singh, Sai Sravanthi K, Himanshu Achra, Dheeraj Akula

Physics Department, BITS Pilani Hyderabad Campus, Hyderabad

April 27, 2019

Abstract

Sound Localization was performed using four cellphone microphones in two dimensions. Cross correlation function and Fast Fourier transform were used to find the arrival time difference of sound at different microphones. Relation between sampling rate of the sensor and position error was studied by considering alternative samples in the recorded waveform. An attempt was made to accurately measure the source position outside the perimeter of sensors.

1 INTRODUCTION

Sound localization is the process through which location of sound source is determined in terms of distance and direction. In the recent past many progressive researches have been carried out on this topic. Sound localization has found its applications in various fields such as sonar, medical, radar, defense and aerospace. Many countries are actively doing research on this topic to detect military weapons for example US boomerang system, Advanced Sound Ranging (ASR) Project by Britain and Rafael anti-sniper system of Israel. This report discusses the sound localization through time difference method. We have used Fast Fourier algorithm to calculate the time difference between different microphones. This report broadly outlines the experimental setup, real data and simulated data derived from extensive experiments conducted during the course of the study.

2 HYPOTHESIS

1. The error in the measured position of the sound source increases when the sampling rate in the experiment is decreased.
2. It is not possible to accurately determine the position of sound source outside the perimeter formed by the sensors.

3 MATERIALS USED

1. 4 similar microphones (Boat BassHeads 100)
2. 4 Android phones (With WaveEditor app)

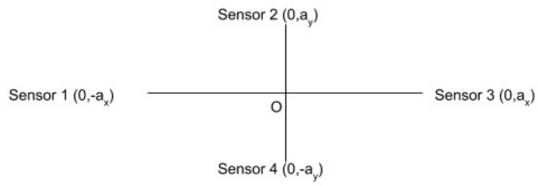

Figure 1: Experimental Setup

3. Audacity software
4. Flat Table.

4 METHODOLOGY

4.1 Experimental Setup

Setup for the experiment is shown in Figure 1.When sound waves are emitted from the source at any location (x,y) at an unknown time to, then the travel times to the four microphones is given by

$$t_1 - t_0 = \sqrt{(x-c_x)^2 + y^2}/v \tag{1}$$

$$t_2 - t_0 = \sqrt{(x)^2 + (y-c_y)^2}/v \tag{2}$$

$$t_3 - t_0 = \sqrt{(x+c_x)^2 + y^2}/v \tag{3}$$

$$t_2 - t_0 = \sqrt{(x)^2 + (y+c_y)^2}/v \tag{4}$$

where v is the speed of sound and ti is the measured time of arrival of the signal at the ith microphone. Combining Eqs. (1) and (3) to obtain

$$v(t_3 - t_1) = \sqrt{(x+c_x)^2 + y^2} - \sqrt{(x-c_x)^2 + y^2} \tag{5}$$

This equation describes a hyperbola with foci at ($+c_x$, 0) and ($-c_x$, 0). Defining the time difference as t_x = t_3 - t_1, the difference in distances from the point (x, y) to the two foci is given by $2a_x$ = vt_x. The points (a_x, 0) and ($-a_x$, 0) are the vertices of the hyperbola. After defining subvertices of the hyperbola as are the vertices of the hyperbola. After defining subvertices of the hyperbola as,

$$b_x = \sqrt{c_x^2 - a_x^2}$$

Eq. (5) can be rearranged to obtain an equation for a hyperbola in a standard form. Equation corresponds to a hyperbola with foci at (0,c_y), (0,$-c_y$) and vertices at (0,a_y) and (0,$-a_y$). Hyperbolic can be solved for the source position, yielding

$$x = a_x b_y \sqrt{(a_y^2 + b_x^2)/(b_y^2 b_x^2 - a_y^2 a_x^2)}$$

$$y = a_y b_x \sqrt{(a_x^2 + b_y^2)/(b_y^2 b_x^2 - a_y^2 a_x^2)}$$

Errors in measured times produce errors in the a and b parameters of the hyperbolas, leading to relatively small shifts in the point of intersection of the hyperbolas.

4.2 Procedure

1. Setup - Grids were marked on a flat table to keep track of the position of the sound source. Microphones,connected with mobile phones, were placed on the table as per the experimental setup shown in figure
2. WaveEditor app is used to detect and record the incident sound wave on the microphone.
3. Synchronization of clocks - To synchronize the mobile clocks, sound was produced at the origin and recorded using microphones.The arrival time difference between the microphones was used as the time lag between clocks. Sensor 1 was synchronized with sensor 3 and sensor 4 was synchronized with sensor 2.
4. Recordings were made at different source positions and analyzed using Audacity software.
5. Cross- correlation along with Fast Fourier transform was used to find the arrival time difference between the microphones as explained below.

4.3 Cross Correlation

Identification of start of the pulse didn't give consistent results, so many other methods were tried like:
1. Using the time at which maximum amplitude occurred for every pulse by exporting data from audacity.
2. Using 3 sound pulses and taking the average of onset for the time pulses.
3. Using different sound sources like sine wave , increasing amplitude sine wave, noise etc. to obtain an onset distinct from noise.

All these methods failed as the errors were very high. Correlation is nothing but sliding one signal and multiplying it by the other. The index (time) of maximum value of correlation gives us the time shift.But this implemented in discrete time requires time complexity of O(N^2) which is very large considering our N is (10^6), so FFT(Fast Fourier Transform) was used which uses time complexity of O($N * log(N)$). Signals were first transformed to fourier space using fft algorithm and then multiplied in the fourier space and then get the inverse fourier transform which gives the correlation in time domain. The code is written in python and the libraries used are numpy and scipy . Time taken or N = 264600 the time is 2.88 seconds.

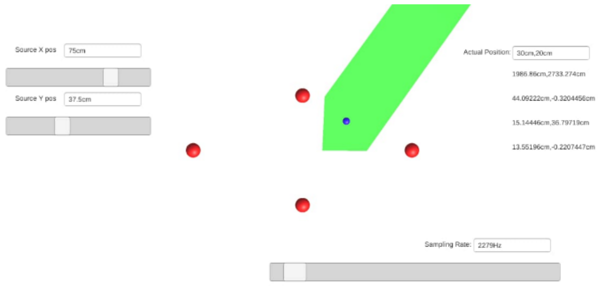

Figure 2: Simulation

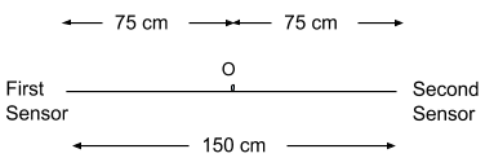

Figure 3: Experiment 1

4.4 Simulation

A simulation of above procedure is modelled using the unity software and c scripting language. Real world elements like sensors, sound source, sampling rate of the sensor, etc are modelled. Sound is simulated starting from the source and time is calculated it takes to reach the sensors as it is done in the real world. Difference of these times is calculated and after interpolating the two hyperbolas their intersection gives the exact source location. All the above steps are done to mimic the real world experimental procedure. This itself does not give us any valuable information since the actual position and the predicted position are the same because everything is ideal. Then the errors due to sampling rate are modelled which leads us to different predicted positions than the actual one. The error due to sampling rate is that because of oversampling or undersampling due to which time difference may be in the range of original time difference +- 1/sampling rate. This defines a two dimensional region in which the source can be predicted which is plotted in the simulation as shown in figure 2.

5 DATA COLLECTION AND RESULTS

5.1 Sound localization in one dimension

5.1.1 Setup

In this experiment, position of the source is constrained to be on the line segment joining the two microphones (Figure 3). Sound source was placed at 10cm, 20cm, 30cm, 40cm, 50cm and 60cm on both sides of the origin(O).

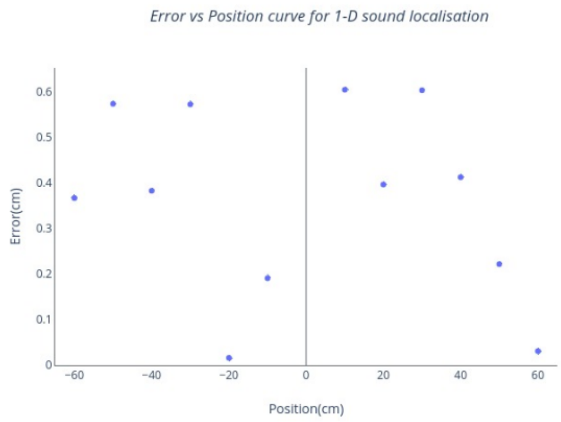

Figure 4: Error vs Position curve for 1-D sound localisation

5.1.2 Data Collected

Table 1: Experiment No.1 Data

Actual Position(cm)	Measured Position(cm)	Error(cm)
-60	-59.6331	0.3669
-50	-49.4261	0.5739
-40	-39.617	0.383
-30	-29.4273	0.5727
-20	-20.0111	0.0161
-10	-9.8091	0.1909
10	10.6049	0.6049
20	20.3967	0.3967
30	30.6037	0.6037
40	40.4128	0.4128
50	50.2219	0.2219
60	60.031	0.031

Error = mod(Actual Position - Measured Position)

RMS Position error

$$RMSError = \sqrt{(\sum(Error)^2/11)} = 0.4352cm \quad (6)$$

This value indicates the expected error if the experiment is performed at any position on the line segment.
Source of error
The only possible reason for this position error is the timing error due to a finite sampling rate of the microphone.
Approximate theoretical estimate of timing error
= Speed of sound* least count of time
= 346m/s * 1/44100 Hz
= 0.7854 cm
Figure 4 shows the change in error withe the position of the sound source.

5.2 Sound localisation in two dimension

5.2.1 Setup

In this experiment ,source is placed on different positions in a 2-D plane and its location is estimated(Figure 5).

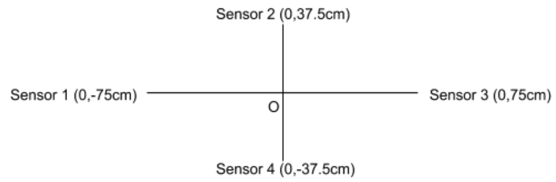

Figure 5: Experiment 2

Table 2: Experiment No.2 Data

Y-actual(cm)	X-actual(cm)	Y-measured(cm)	X-measured(cm)	Y-error(cm)	X-error(cm)
5	10	5.26	10.173	0.26	0.173
5	20	5.709	19.168	0.709	0.832
5	30	5.9	28.586	0.9	1.414
5	40	5.602	38.379	0.602	1.621
5	50	5.702	47.822	0.702	2.178
10	10	10.226	10.691	0.226	0.691
10	20	10.467	21.407	0.467	1.407
10	30	10.377	31.745	0.377	1.745
10	40	10.121	42.126	0.121	2.126
10	50	16.148	54.027	6.148	4.027
20	10	19.952	9.687	0.048	0.313
20	20	19.86	19.4	0.14	0.6
20	30	14.35	27.866	5.65	2.134
20	40	15.14	37.209	4.86	2.791
20	50	20.904	47.473	0.906	2.527
30	10	34.54	9.878	4.54	0.122
30	20	25.981	22.382	4.019	2.382
30	30	23.991	27.606	6.009	2.394
30	40	23.55	40.065	6.45	0.065

5.2.2 Data Collected

Data Collected is shown in Table 2.Figure 6 shows the actual positions and measured positions plotted in a plane.

5.3 Sound localisation on a line inclined with the axes

5.3.1 Setup

In this experiment, position of the source is constrained to be on a line inclined at 45 degree with the axes.(Figure 7) Sound source was placed at 5cm, 10cm, 15cm, 20cm, 25cm, 30cm and 35cm on the inclined line.

5.3.2 Data Collected

Data collected for the experiment is displayed in Table 3.Figure 8 shows a graphical representation of the

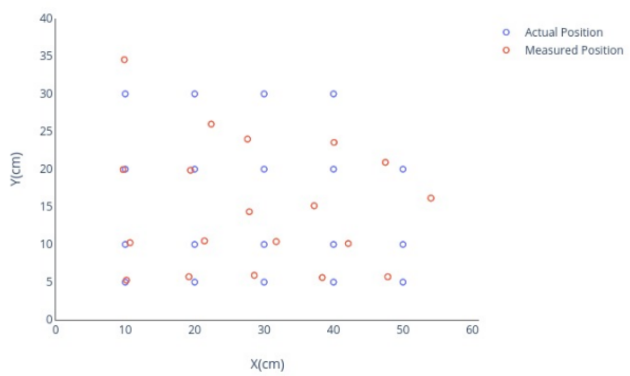

Figure 6: Actual position and measured position plotted in 2-D plane

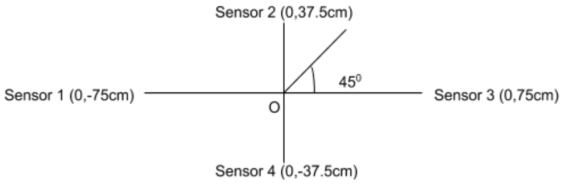

Figure 7: Experiment 3

Table 3: Error in each axis for Experiment No.3

X-actual(cm)	Y-Actual(cm)	X-Measured(cm)	Y-Measured(cm)	Error-X(cm)	Error-Y(cm)
3.53	3.53	3.5136	3.9217	0.0164	0.3917
7.07	7.07	7.4553	7.9685	0.3853	0.8985
10.6	10.6	15.0458	12.6961	4.4458	2.0961
14.14	14.14	18.4114	16.8116	4.2714	2.6716
17.67	17.67	17.823	20.3986	0.153	2.7286
21.21	21.21	24.0115	25.4855	2.8015	4.2755
24.74	24.74	27.6624	30.808	2.9224	6.068

actual and measured positions. Error in distance from the origin is shown in Table 4 and plotted in Figure 10. From the graph it can be seen that, generally, the error in the measured position increases with distance from the origin.

Table 4: Error from Origin(O)

X-Actual(cm)	Y-Actual(cm)	Distance from O(cm)	X-Measured(cm)	Y-measured(cm)	Error (cm)
3.53	3.53	5	3.5136	3.9217	0.392
7.07	7.07	10	7.4553	7.9685	0.9776
10.6	10.6	15	15.0458	12.6961	4.9151
14.14	14.14	20	18.4114	16.8116	5.038
17.67	17.67	25	17.823	20.3986	2.7328
21.21	21.21	30	24.0115	25.4855	5.1115
24.74	24.74	35	27.6624	30.808	6.735

5.4 Sound Localization outside the perimeter of sensors

5.4.1 Setup

In this experiment, source is placed outside the perimeter of the sensors(Figure 9). Total eight points were used in this experiment.
1. 4 points behind each microphone – (0,35cm), (-70cm,0), (0,-35cm), (70cm,0).
2. 4 points on the boundary of the box – (-65,30), (65,30), (-65,-30), (65,-30).

5.4.2 Data collected

Data collected is displayed in Table 5. Many source positions can not be estimated because of no intersection between the hyperbolas.

5.5 Sound Localization with different sampling rates

Aim of this experiment is to observe the effect of sampling rate on the error in measured position. The microphones used have a constant sampling rate. So, if every alternating sample is considered then the sampling rate becomes half of the original sampling rate.

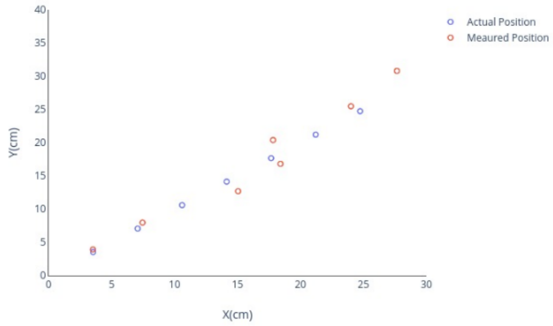

Figure 8: Actual positions and measured positions on inclined line

Table 5: Experiment No.4 Data

X-actual(cm)	Y-actal(cm)	X-measured(cm)	Y-measured(cm)	X-error(cm)	Y-error(cm)
0	35	0	25.353	0	9.647
-70	0	-65.71	1.905	4.29	1.905
0	-35	-0.403	-23.793	0.403	11.207
70	0	NA	NA	NA	NA
-65	30	NA	NA	NA	NA
-65	-30	NA	NA	NA	NA
65	-30	NA	NA	NA	NA
65	30	NA	NA	NA	NA

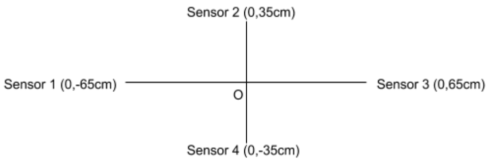

Figure 9: Experiment 4

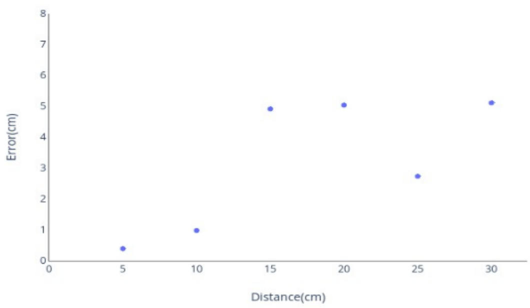

Figure 10: Error vs distance plotted for inclined line

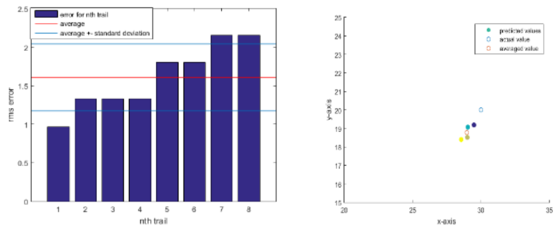

Figure 11: Error for 44100 Hz

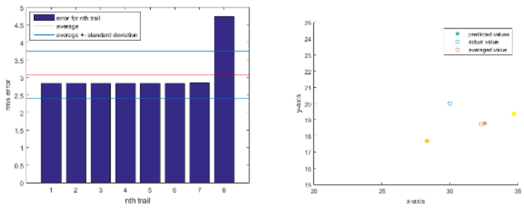

Figure 12: Error for 8820 Hz

Similar techniques are used to obtain different sampling rates. All the measurements are taken at x = 30cm and y = 20cm position.
1. Sampling Rate = 44100 Hz
Results are shown in Table 6 and Table 7 and plotted in Figure11.
2. Sampling rate = 8820 Hz
Results are shown in Table 8 and Table 9 and plotted in Figure 12.
3. Sampling Rate = 4410 Hz
Results are shown in Table 10 and Table 11 and plotted in Figure 13.
4. Sampling Rate = 2205 Hz
Results are shown in Table 12 and Table 13 and plotted in Figure 14.
5. Sampling rate = 1764 Hz
For this sampling rate,all the 8 trials do not meet. It can be seen that the average of the rms error increases as the sampling rate is decreased. As the sampling rate is decreased the predicted positions gets snapped to certain points. After a certain minimum frequency which turns out to be 1764 Hz in this case, prediction of final positions should not be possible.

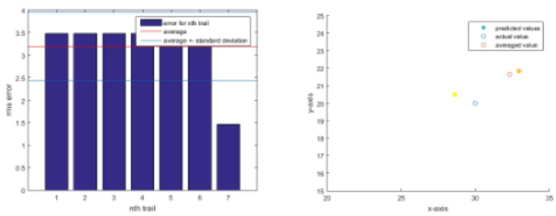

Figure 13: Error for 4410 Hz

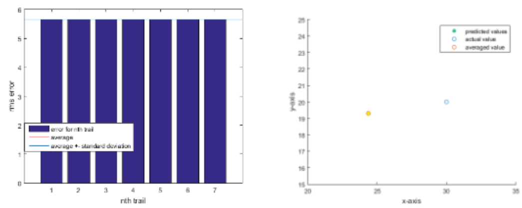

Figure 14: Error for 2205 Hz

Table 6: Sampling Rate = 44100 Hz

Trail no:	x predicted(cm)	y predicted(cm)	rms error(cm)
1	29.473	19.19	0.965
2	29.047	19.074	1.328
3	29.047	19.074	1.328
4	29.047	19.074	1.328
5	28.987	18.506	1.803
6	28.987	18.506	1.803
7	28.564	18.396	2.152
8	28.564	18.396	2.152

Table 7: Sampling Rate = 44100 Hz

	x predicted(cm)	y predicted(cm)	rms error(cm)
Average	28.964	18.777	1.607
Standard Deviation	0.293	0.353	0.433

Table 8: Sampling rate = 8820 Hz

Trail no:	x predicted(cm)	y predicted(cm)	rms error(cm)
1	32.553	18.777	2.831
2	32.553	18.777	2.831
3	32.553	18.777	2.831
4	32.553	18.777	2.831
5	32.553	18.777	2.831
6	32.553	18.777	2.831
7	28.328	17.684	2.855
8	34.702	19.363	4.745

Table 9: Sampling rate = 8820 Hz

	x predicted(cm)	y predicted(cm)	rms error(cm)
Average	32.294	18.713	3.073
Standard Deviation	1.77	0.463	0.675

Table 10: Sampling Rate = 4410 Hz

Trail no:	x predicted(cm)	y predicted(cm)	rms error(cm)
1	32.956	21.834	3.479
2	32.956	21.834	3.479
3	32.956	21.834	3.479
4	32.956	21.834	3.479
5	32.956	21.834	3.479
6	32.956	21.834	3.479
7	28.619	20.496	1.466
8	Do not intersect	Do not intersect	Do not intersect

Table 11: Sampling Rate = 4410 Hz

	x predicted(cm)	y predicted(cm)	rms error(cm)
Average	32.336	21.643	3.191
Standard Deviation	1.638	0.506	0.76

Table 12: Sampling Rate = 2205 Hz

Trail no:	x predicted(cm)	y predicted(cm)	rms error
1	24.386	19.299	5.657
2	24.386	19.299	5.657
3	24.386	19.299	5.657
4	24.386	19.299	5.657
5	24.386	19.299	5.657
6	24.386	19.299	5.657
7	24.386	19.299	5.657
8	Do not intersect	Do not intersect	Do not intersect

Table 13: Sampling Rate = 2205 Hz

	x predicted(cm)	y predicted(cm)	rms error(cm)
Average	24.386	19.299	5.657
Standard Deviation	0	0	0

6 CONCLUSION

- In all the above experiments the results obtained lie within the boundary region predicted by the simulation till a certain number of samples and after that the error is increasing rapidly. This is due to the fact that the synchronization error builds up after a certain amount of time. It can be seen that in experiment 5, the trail number 8 does not intersect after 44100 Hz sampling rate due to fact that the error has built up. If this trail number 8 is considered an outlier, then it can be seen from the tables and graphs that the average of the error is increasing and the standard deviation of error is decreasing as the sampling rate is decreased.

- From the simulation it is observed that for sampling rates below 2279 Hz, there is a good chance that the source position can not be predicted but in the experiment it is possible to predict even for 2205 Hz. This is due to the fact that 2279 is the limit only if oversampling is done both in x axis and y axis. So there is of chance of getting into trouble so this experiment survived these odds. It can be seen that the when sampling rate is 1764 Hz no output is obtained. So it is better not to conduct the experiment below 2279 Hz sampling rate.

- As observed from the table when the source is outside the sensors there is no intersection point for five out of the eight points and the mean error for other three points is very large as compared to when the source was inside the perimeter which was 2cm. This is due to the convergence of the two hyperbolas at the solution with almost same angle of incidence when the source is outside the perimeter. Hence the noise and time errors can produce large localization errors because a slight change in these asymptotic angles can shift the intersection point to a very large distance or may even completely erase the intersection point.

7 ACKNOWLEDGEMENTS

We would like to thank Dr. Aravind N. Raghavan and Dr. Meenakshi V. for their support and guidance throughout this endeavour. We would also like to thank the Physics department for letting us continually use the laboratory for weeks.

8 BIBLIOGRAPHY

1. Lamelas, F., Swaminathan, S. (2019). Laboratory demonstration of acoustic source localization in two dimensions. *American Journal of Physics*, 87(1), 24-27.

2. Fan, J., Luo, Q., Ma, D. (2010). Localization estimation of sound source by microphones array. *Procedia Engineering*, 7, 312-317. .

3. Bingqian, D., Fangmin, L. (2016, December). *An Active Sound Localization Method Based on Mobile Phone*. In 2016 5th International Conference on Computer Science and Network Technology (ICCSNT) (pp. 846-850). IEEE.

4. Adrián-Martínez, S., Bou-Cabo, M., Felis, I., Llorens, C. D., Martínez-Mora, J. A., Saldaña, M., Ardid, M. (2014, June). *Acoustic Signal Detection Through the Cross-Correlation Method in Experiments with Different Signal to Noise Ratio and Reverberation Conditions. In International Conference on Ad-Hoc Networks and Wireless* (pp. 66-79). Springer, Berlin, Heidelberg.

Figure 15: Experimental setup