IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 9, Issue 5, Oct-Nov 2021
**ISSN: 2320 – 8791 (Impact Factor: 2.317)**
**www.ijreat.org**

# 3D Localization of Sound Source Using Microphone Array

**Hanumanta Rao. P[1], Poorna Chandra Prasad[2]**
[1,2]Assistant Professor, KIET Group of Institutions, Kakinada, AP.

**ABSTRACT**: Speaker identification is found to be an important feature of modern-day contactless communication establishment. This presents a unique and urgent need to create an accurate and efficient localizer. This work aims to demonstrate a 3D localization procedure with the help of a GPS/GSM module enabled with a set of mics. A mathematical analysis is performed to find elevation and azimuth angle for the sound source. The integrated GPS system computes the latitude and longitude of the mics and GPS system receive the coordinates of the mic and transmits this value with azimuth and elevation angle to the smartphone interface.

**KEYWORDS:** Interaural level difference (ILD), Interaural time difference (ITD), Head related transfer function (HRTF), Generalized cross-correlation (GCC), and Time differences of arrival (TDoA).

## 1. INTRODUCTION

Sonar [1], radar [2], teleconferencing [3], mobile phone location [4], navigation and global positioning [5], earthquake location [6], microphone arrays [7], robots [8], micro seismic events in mines [9], sensor networks [10, 11], human-computer interface [12], speaker monitoring [13], surveillance [14], and sound source tracking [15] are variety of applications, where sound source tracking [15] is used for different applications.

The ongoing pandemic situation demands contactless connection for different essential services such as banking. The voice signal can be used as a very simple base for source localization in contactless identification. Studying energy, temporal, and/or directional features from incoming sound at different microphones and using a suitable model that relates those features with the spatial location of the source of interest can be used to localize acoustic sources.

The following is the flow of this paper: section 2 discusses the relevant literature, section 3 describes the proposed TDE-ILD-HRTF method of source localization, section 4 summarizes our experimental results, section 5 discusses the integration with the Geo-location module, and section 6 concludes the paper.

## 2. REVIEW OF LITERATURE

Estimating the arrival time differences (TDoA) for the acquired signal and mapping the signals to the source position is the most common localization method. Since the distance between the source and the array is more than ten times the aperture's area, the far-field hypothesis is confirmed. Only if the cross-correlation of all the acquired signals is known then only TDoA can be calculated for two or more than two microphones. The maximum cross-correlations of the signals acquired include the lag that leads directly to the TDoA estimate. The cross-correlation approach provides greater robustness by applying the weighting function in the cross-spectrum. The generalized cross-correlation (GCC) mechanism is this robust technique. Changing the weighting function will result in a different algorithm. A GCC weighting function based on phase transforms (PHAT) has been introduced. Sound source localization using the steered response power (SRP) method is an extension of the GCC technique, and this method will be dominated by classical SRP (C-SRP).

Arrival route (DoA) [16], TDoA or interaural time difference (ITD)-[17], [18], [19], and [20], methods based on ILD-[21], [22], [23], and [24], and Head Dependent Transfer Function are the most popular passive sound source localization methods (HRTF). DoA-based beamforming and subspace approaches require a large number of microphones to approximate narrowband source locations in far-field and wideband source locations in near-field instances.

A variety of localization methods have been proposed for multiple signal estimation, covariance approximation, and signal parameter estimation using rotational invariance techniques. The ILD-based approach, on the other hand, is commonly applicable to [high signal-to-noise ratio (SNR)] [21], [22], [23], and [24] single dominant sound sources in the case of a single dominant sound source. For ILD- or TDE-based methods, three microphones are required for 2D positioning and four for 3D [17], [25, 26], [27], [28], [29]. The arrival angle can only be calculated using HRTF-based methods [30].

However, GCC with an ML estimator is the most widely used tool for TDE. Later, a slew of methods for improving GCC in the presence of noise were proposed. In the last decade, several papers have attempted to use HRTF in conjunction with TDE to estimate elevation and azimuth angle of arrival [31, 32]. In this work we try to explore TDE-ILD-HRTF model for source localization.

## 3. Proposed TDE-ILD-HRTF Method

### 3.1 Source Counting Method

One sound signal (source signal) is considered as:

$$s(t) = s(t + T) \qquad (1)$$

and two $s\,(t)$ and $s^1\,(t)$ signal in high SNR localization area. Therefore equation (1) can be rewrite as:

$$s_1(t) = s^1(t - T_1') + s(t - T_1) \qquad (2)$$

$$s_2(t) = s^1(t - T_2{}') + s(t - T_2) \qquad (3)$$

And correlation function is given as

$$Rs_1s_2(\tau) = \int_{-\infty}^{\infty}\left(s^1(t - T_1{}') + s(t - T_1)\right)\left(s^1(t - T_2{}' + \tau) + s(t - T_2 + \tau)\right)dt \qquad (4)$$

$$Rs_1s_2(\tau) = R_{1s_1s_2}(\tau) + R_{2s_1s_2}(\tau) + R_{3s_1s_2}(\tau) + R_{4s_1s_2}(\tau) \qquad (5)$$

Using above equations, maximum value for $R_{1v_1v_2}(\tau)$ is given by $\tau_1 = T_2 - T_1$ and similar expressions follow for $R_{2v_1v_2}(\tau)$, $R_{2v_1v_2}(\tau)$, $R_{3v_1v_2}(\tau)$ and $R_{4v_1v_2}(\tau)$.Therefore, four peak values are there in cross-correlation vector and by counting the values of dominant cross-correlation vector, number of dominant and active sources in localization area can be easily found.

### 3.2 Extension of Dimensions to Three

Half-cylinder reflector is only applicable in 2D cases in the proposed localization system. Instead of half-cylinder, a half-sphere reflector is more effective for 3D sound source localization Half-sphere reflector allows localization structure to localize sources of sound in 3D case as shown in Figure 1.

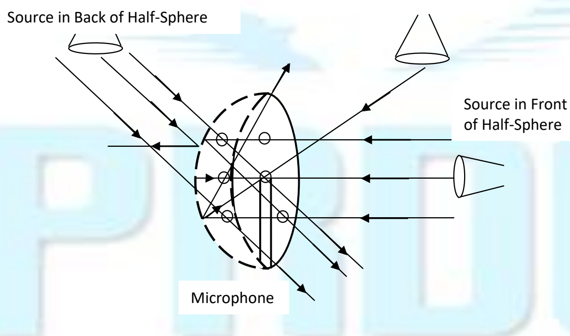


**Figure 1** Half sphere as a coarse approximation

## 4. EXPERIMENTAL OBSERVATIONS

### 4.1 Experimental Setup

The proposed hardware is implemented using a ‘MB800' motherboard and DELTA IOIO LT sound card from M-Audio, with up to 16-bit resolution and 96-kHz sampling frequency, and a B-29L preamplifier with three ‘C-417' style microphones, as shown in Figure 2.

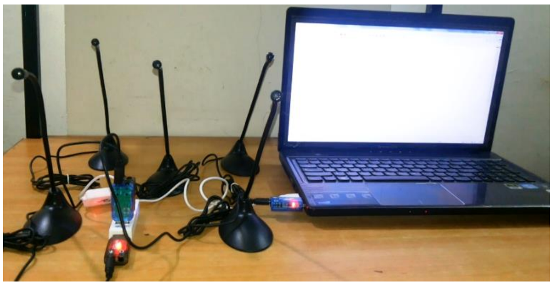

**Figure 2** Experimental mic setup

A 2-cm-diameter custom reflector is mounted behind the third microphone, as shown in Figure 3.

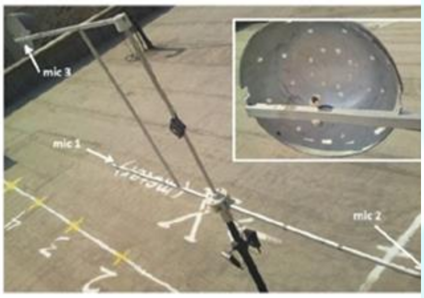


**Figure 3** The proposed 3D sound source localization setup

## 5. INTEGRATION WITH GEO-LOCATIONING

A set of mics integrated with GPS/GSM module is placed in a small geographical area. The received signal strength of all the mics is monitored using simulation tool and two of these mics are identified which are providing the highest value of sound magnitude.

Figure 4 shows the integration of acoustic-based directionality with GPS device.


**Figure 4** Integration of acoustic-based directionality with GPS devices

Table 1 and Table 2 indicate the angle of arrival measurements.

**Table 1** Hardware implementations based measured values of the azimuth angle (∅) of Arrival

| $Real\ angle\ \emptyset\ in\ (°)$ | $\emptyset_{measured}\ in\ (°)$ | $Error\ in\ (°)$ |
|---|---|---|
| 0 | 0.18 | 0.18 |
| 15 | 15.13 | 0.13 |
| 30 | 29.85 | 0.15 |
| 45 | 45.18 | 0.18 |
| 60 | 60.14 | 0.14 |
| 75 | 74.91 | 0.09 |
| 90 | 89.83 | 0.17 |
| 105 | 104.82 | 0.18 |
| 120 | 120.14 | 0.14 |
| 135 | 135.11 | 0.11 |
| 150 | 150.14 | 0.14 |
| 165 | 165.09 | 0.09 |
| 180 | 179.93 | 0.07 |

**Table 2** Hardware implementations based measured values of the elevation angle (θ) of arrival

| $Real\ angle\ \theta\ in$ (°) | $\theta_{measured}\ in$ (°) | $Error\ in$ (°) |
|---|---|---|
| -10 | -10.18 | 0.18 |
| -5 | -5.16 | 0.16 |
| 0 | 0.05 | 0.05 |
| 15 | 14.15 | 0.15 |
| 30 | 29.91 | 0.09 |
| 45 | 45.19 | 0.19 |
| 60 | 60.13 | 0.13 |
| 75 | 75.11 | 0.11 |
| 90 | 89.82 | 0.18 |

The integrated GPS system computes the latitude and longitude of the mics and GPS system receives the coordinates of the mic and transmits this value with azimuth and elevation angle to the smart phone interface. Figure 5 shows/locates the user in larger geography.

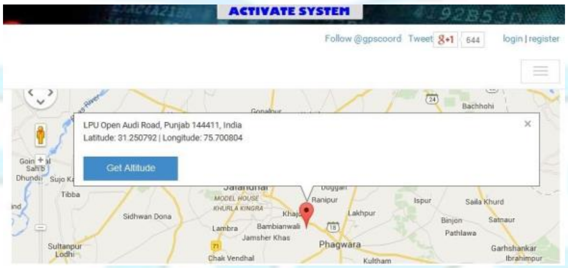


**Figure 5** Integration with Google maps to increase usability

## 6. CONCLUSION

In this paper, the designed project aims to localize the 3D sound source (direction only) without recognize the speaker, i.e., who is speaking and the content of the speech. To demonstrate this 3D localization procedure, a set of mics integrated with GPS/GSM module are placed in a small geographical area. The received signal strength of all the mics is monitored using simulation tool and two of these mics are identified which are providing the highest value of sound magnitude. A mathematical analysis is done after this to calculate the approximate both elevation and azimuth angle for the sound source with respect to the two identical mics (with highest output). The integrated GPS system computes the latitude and longitude of the mics and GPS system receives the coordinates of the mic and transmits this value with azimuth and elevation angle to the smart phone interface.

## REFERENCES

1. Carter, G.C. (1981) 'Time delay estimation for passive sonar signal processing', *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 29, No. 3, pp.463–470.

2. Weinstein, E. (1982) 'Optimal source localization and tracking from passive array measurements', *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol. 30, No. 1, pp.69–76.

3. Wang, H. and Chu, P. (1997) 'Voice source localization for automatic camera pointing system in videoconferencing', *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997. ICASSP-97.*, Vol. 1, IEEE, ICASSP '97, Munich, Germany, 21–24 April 1997, pp.187–190.

4. Caffery Jr, J. and Stuber, G.L. (1998) 'Subscriber location in cdma cellular networks', *IEEE Transactions on Vehicular Technology*, Vol. 47, No. 2, pp.406–416.

5. Tsui, J.B-Y. (2000) *Fundamentals of Global Positioning System Receivers*, New York: A Wiley-Inter-science.

6. Klein, F.W., Frankel, A.D., Mueller, C.S., Wesson, R.L. and Okubo, P.G. (2001) 'Seismic hazard in Hawaii: high rate of large earthquakes and probabilistic ground-motion maps', *Bulletin of the Seismological Society of America*, Vol. 91, No. 3, pp.479–498.

7. Huang, Y., Benesty, J., Elko, G.W. and Mersereati, R.M. (2001) 'Real-time passive source localization: a practical linear-correction least-squares approach', *IEEE Transactions on Speech and Audio Processing*, Vol. 9, No. 8, pp.943–956.

8. Valin, J-M., Michaud, F., Rouat, J. and Létourneau, D. (2003) 'Robust sound source localization using a microphone array on a mobile robot', *IEEE/RSJInternational Conference on Intelligent Robots and Systems, 2003. (IROS 2003). Proceedings*, Vol. 2, IEEE, IROS 2003, Las Vegas, USA, 27–31 October 2003, pp.1228–1233.

9. Daku, B.L., Salt, J.E. and Sha, L. (2004) 'An algorithm for locating microseismic events', *Canadian Conference on Electrical and Computer Engineering, 2004,* Vol. 4, IEEE, pp.2311–2314.

10. Gezici, S., Tian, Z., Giannakis, G.B., Kobayashi, H.,Molisch, A.F., Poor, H.V. and Sahinoglu, Z. (2005) 'Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks', *IEEE Signal Processing Magazine*, Vol. 22, No. 4, pp.70–84.

11. Patwari, N., Ash, J.N., Kyperountas, S., Hero III, A.O., Moses, R.L. and Correal, N.S. (2005) 'Locating the nodes: cooperative localization in wireless sensor networks', *IEEE Signal Processing Magazine*, Vol. 22, No. 4, pp.54–69.

12. Lee, J-Y., Hahn, M. and Cho, Y-J. (2005) 'Real-time sound localization using time difference for human robot interaction', *World Congress*, Vol. 16, pp.1410–1410, Citeseer.

13. Ma, W-K.,Vo, B-N., Singh, S.S. and Baddeley, A. (2006) 'Tracking an unknown time-varying number of speakers using TDoA measurements: a random finite set approach', *IEEE Transactions on Signal Processing*, Vol. 54, No. 9, pp.3291–3304.

14. Ho, K., Lu, X. and Kovavisaruch, L-O. (2007) 'Source localization using TDoA and FDoA measurements in the presence of receiver location errors: analysis and solution', *IEEE Transactions on Signal Processing*, Vol. 55, No. 2, pp.684–696.

15. Cevher, V., Sankaranarayanan, A.C., McClellan, J.H. and Chellappa, R. (2007) 'Target tracking using a joint acoustic video system', *IEEE Transactions on Multimedia*, Vol. 9, No. 4, pp.715–727.
16. Richardson, M.H. and Formenti, D.L., 1985, January. Global curve fitting of frequency response measurements using the rational fraction polynomial method. In *Proceedings of the Third International Modal Analysis Conference* (pp. 390-397).

17. Hahn, W.R. (1975) 'Optimum signal processing for passive sonar range and bearing estimation', *The Journal of the Acoustical Society of America*, Vol. 58, No. 1, pp.201–207.

18. Lee, H.B. (1975) 'A novel procedure for assessing the accuracy of hyperbolic multi-lateration systems', *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES- 11, No. 1, pp.2–15.

19. Schmidt, R. (1972) 'A new approach to geometry of range difference location', *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 6, No. AES-8, pp.821–835.

20. Van Etten, J. (1970) 'Navigation systems. Fundamentals of low-and very-low-frequency hyperbolic techniques', *Electrical Communication*, Vol. 45, No. 3.

21. Birchfield, S.T. and Gangishetty, R. (2005) 'Acoustic localization by interaural level difference', *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP'05)*, Vol. 4, IEEE, Philadelphia Convention Center Philadelphia, PA, USA, 19–23 March 2005, p.iv-1109.

22. Ho, K. and Sun, M. (2007) 'An accurate algebraic closed form solution for energy-based source localization', *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 8, pp.2542–2550.

23. Ho, K. and Sun, M. (2008) 'Passive source localization using time differences of arrival and gain ratios of arrival', *IEEE Transactions on Signal Processing*, Vol. 56, No. 2, pp.464–477.

24. Julián, P., Andreou, A.G., Riddle, L., Shamma, S., Goldberg, D.H. and Cauwenberghs, G. (2004) 'A comparative study of sound localization algorithms for energy aware sensor.

25. Brandstein, M.S. and Silverman, H.F. (1997) 'A practical methodology for speech source localization with microphone arrays', *Computer Speech & Language*, Vol. 11, No. 2, pp.91–126.

26. Brandstein, M.S., Adcock, J.E. and Silverman, H.F. (1997) 'A closed-form location estimator for use with room environment microphone arrays', *IEEE Transactions on Speech and Audio Processing*, Vol. 5, No. 1, pp.45–50.

27. Koyano, Y., Yatabe, K., Ikeda, Y. and Oikawa, Y., 2016, March. Physical-model based efficient data representation for many-channel microphone array. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (pp. 370-374), IEEE.

28. Lleida, E., Fernández, J. and Masgrau, E. (1998) 'Robust continuous speech recognition system based on a microphone array', *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, 1998*, Vol. 1, IEEE, Washington StateConvention and Trade Center Seattle, WA (USA), 12–15 May 1998, pp.241–244.

29. Omologo, M. and Svaizer, P. (1994) 'Acoustic event localization using a cross power-spectrum phase-based technique', *IEEE International Conference on Acoustics, Speech, and Signal Processing, 1994. ICASSP-94*, Vol. 2, IEEE, ICASSP 1994, Adelaide, South Australia, Australia, 19–22 April 1994, pp. II–273.

30. Cui, W., Cao, Z. and Wei, J. (2006) 'Dual-microphone source location method in 2-d space', *IEEE International Conference on Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings*, Vol. 4, IEEE, Toulouse Pierre Baudis Congress Center Toulouse, France, 14–19 May 2006, pp.IV–IV.

31. Kulaib, A., Al-Mualla, M. and Vernon, D. (2009) '2D binaural sound localization for Urban search and rescue robotic', *The 12th International Conference on Climbing and Walking Robots*, Istanbul, Turkey, pp.9–11.

32. Sommerville, D.M.Y. (2012) *The Elements of Non-Euclidean Geometry*, Dover Books on Mathematics Courier Corporation. 15 Wellman Avenue North Chelmsford, MA, USA, p.288.