

3D Moving Sound Source Localization via Conventional Microphones

Mehmet Cem Catalbas¹, Simon Dobrisek²

¹*Faculty of Electrical and Electronics Engineering, Firat University,
Elazig, Turkey*

²*Faculty of Electrical Engineering, University of Ljubljana,
Ljubljana, Slovenia*

catalbas@firat.edu.tr

Abstract—In this paper, we present an approach to sound localization for moving sound sources. We use four conventional microphones and a multi-channel sound recording device for capturing sound signals. A database with different types of moving sound source scenarios is created and one of the scenarios is analysed in this paper. This scenario involves walking of a person and creating sounds simultaneously at certain positions in a room, and our aim is to determine the location of the person as three dimensions. We use time difference of arrival for 3D sound source localization. The time delays between sounds channels are obtained by excitation source information based time-delay estimation algorithm. In addition, signal processing methods are used for increasing the success of moving sound source localization. Savitzky-Golay and threshold filters are utilized to pre-filter the recorded sounds and removing the background noise signals, respectively. The non-linear equation of the sound localization is solved by using the modified Levenberg–Marquardt Algorithm (LMA) on the Time Difference of Arrival (TDOA) process. The results for the 3D localization of the moving sound source are obtained by measuring the differences between the real and estimated positions of the person.

Index Terms—Sound processing; sound source localization; generalized cross-correlation; moving sound source; time delay estimation; time difference of arrival.

I. INTRODUCTION

In this paper, the generic idea of sound localization process is examined. Sound source localization is the determination of exact locations of sound sources using some appropriate signal processing techniques [1], [2]. The location of a sound source can be determined by 2D or 3D observation space [3], [4]. Generally, 2D location information is adequate for some of the basic applications [5]. Furthermore, some applications such as robotics, drones, etc. require the information on 3D positions. Before discussing the concept of sound location, we need to explain some of the basic parameters of sound signals. The basic definition of a sound signal is an air pressure changing with time. The pressure change can be easily converted to electrical signals via microphones which are three basic types based on the application area [6]. The first one based on the generic principle of generator systems is dynamic

microphones.

The vibration of the air can be easily transformed into an electric signal via vibrating diaphragm and the coil structure. The second type is capacitor microphones affected by the changes in capacitance due to the changes of sound pressure. The last type is special purpose microphones that are specialized for different areas. In this work, the dynamic and basic microphone structure is used. As known, the sound signals vary with the time. The processing of such changes can be useful for different types of signal processing application such as sound recognition and localization, voice-based emotion detection, trajectory estimation etc. [7], [8]. The aim of this paper is to investigate the main sound localization techniques and also the whole sound process of sound source localization. Moreover, a novel sound database that can be used for developing moving sounds source localization systems is presented. A part of the presented database is used to investigate several selected sound localization techniques.

II. SOUND LOCALIZATION

In this section, some terms on sound processing are explained. The sound localization is the determination of sound sources via signal processing techniques. In literature, this concept is implemented in different ways. Generally, the sound localization techniques are inspired by the hearing system of a human. The main terms that are related to sound localization are as follows:

1. Interaural Time Differences (ITD);
2. Interaural Level Differences (ILD).

The parameter of ITD is commonly used on basic works [9]. The negative part of ITD is the sensitivity to ambient noise and it also doesn't provide the exact location of the sound source. This method gives an idea about azimuth angle of the sound source. The ILD method is related with sound pressure level [10]. This parameter also gives an idea about the location of the sound source via sound level difference. Besides, this parameter is more sensitive to environment and hardware condition. The ITD and ILD parameters can be used together. This combination is provided to determine the exact location of sound source [11]. ILD and ITD parameters are shown in Fig. 1 and Fig. 2, respectively.

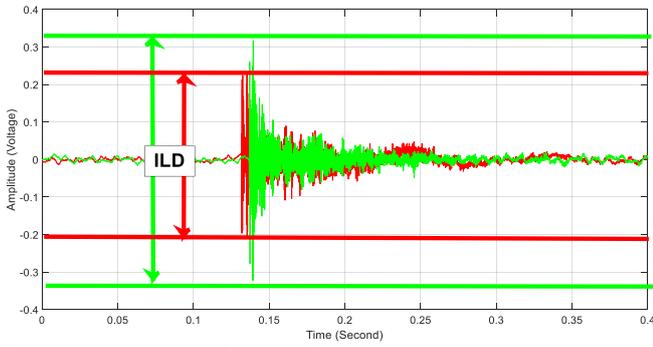


Fig. 1. Interaural level differences (ILD).

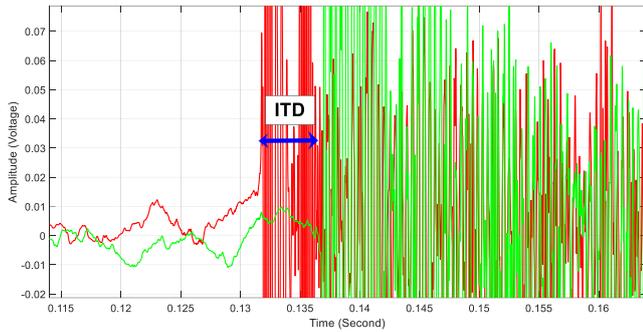


Fig. 2. Interaural time differences (ITD).

III. SOUND LOCALIZATION DATASET

In this section, we explain the basic parameters of our test database. In the presented research, one of the scenarios is tested by our databases. The scenario involves walking off one person and saying the ID numbers of location, simultaneously. These data sets are obtained from four inductive types of microphones and a sound recording system. The sampling rate of the dataset is $f_s = 44100 \text{ Hz}$ and *Scarlett18i8* desktop interface model is used for sound recording. The 2D coordinates of microphones and locations are shown in Fig. 3.

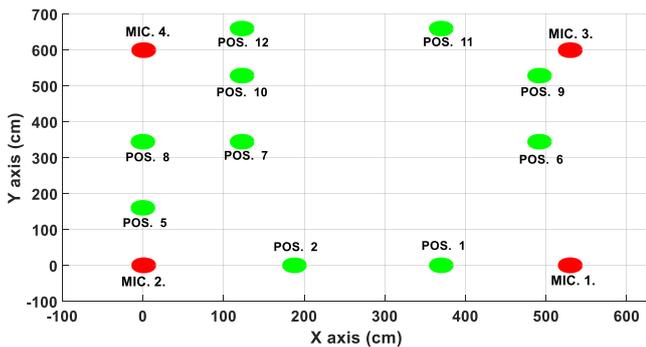


Fig. 3. Location map.

The second microphone is selected as the origin of the coordinate system to reference point. The coordinates of microphones and locations are shown in Table I and Table II, respectively.

TABLE I. COORDINATES ABOUT MICROPHONES.

Microphone	X(cm)	Y(cm)	Z(cm)
1	0	0	0
2	530	0	0
3	0	600	0
4	530	600	0

TABLE II. COORDINATES ABOUT USER LOCATIONS.

Location	X(cm)	Y(cm)	Z(cm)
1	370	0	153
2	188	0	153
3	370	161	172
4	123	161	172
5	0	161	172
6	492	345	172
7	123	345	172
8	0	345	172
9	492	529	172
10	123	529	172
11	370	660	172
12	123	660	172

The 3D modelling of the test room and the environment is shown in Fig. 4.



Fig. 4. 3D illustration of test environment.

The location of the speaker and distances are shown in Fig. 5. The distances between ceiling and floor, mount of the speaker to microphone, and ceiling and microphones are defined as follows d_{cf} , d_z , and d_{cm} , respectively.

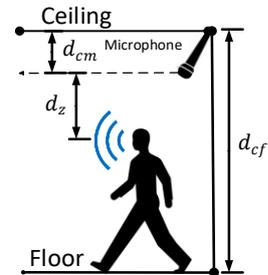


Fig. 5. Location of the speaker.

The front view of the recording environment used in sound source localization is shown in Fig. 6.

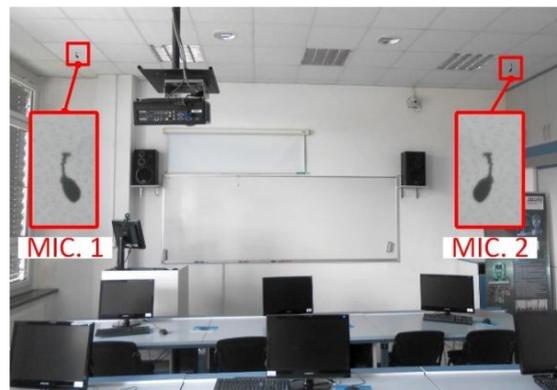


Fig. 6. Recording environment for sound source localization.

IV. SOUND LOCALIZATION PARAMETERS

In this section, we explain some methods of the sound localization process. The most common and basic one is ITD based sound localization which is used by the time lag between the receivers of sound. Several algorithms have been developed to estimate TDOA in the ideal propagation situation. The best known of these algorithms is time domain based cross-correlation [12]. The basic definition of cross correlation is shifted one waveform to other waveform and determined to provide maximum similarity point for obtained time lag between the signals. The cross-correlation formula for continuous time domain is shown in (1). The parameter of τ is the shift parameter for continuous signal

$$r_{12}(\tau) = \int_{-\infty}^{+\infty} x_1(t)x_2(t-\tau)dt. \quad (1)$$

The test sound signal is integrated with the reference sound signal and the test signal is shifted according to time to analyse the similarity. At the end of this process, the amount of delay between two signals is obtained according to time. Besides, time-based cross-correlation algorithm is not robust enough to environmental noise and echo effects. Hence, Generalized Cross-Correlation (GCC) algorithm is preferred by researchers on sound localization process [13]. The brief explanation of cross-correlation and GCC method is as follows

Assuming that there is only one (unknown) sound source in the field, the output of receiver $n(n=1,2,\dots,N)$ can be written as

$$x_n(k) = \alpha_n s(k - D_n) + b_n(k), \quad (2)$$

where α_n , which satisfies $0 \leq \alpha_n \leq 1$, is an attenuation factor due to propagation effects, D_n corresponds to the propagation time from unknown sound source to the receiver n , and $s(k)$ which is often speech from either a talker or loudspeaker is broadband in nature. The $b_n(k)$ represents Gaussian random noise and it is uncorrelated with both sound source signal and noises at the other sensors. The function of cross correlation between two signals is shown in (4):

$$p = D_2 - D_1, \quad (3)$$

$$r_{12}(p) = E[x_1(k)x_2(k+p)]. \quad (4)$$

The parameter of TDOA between received sound signals can be calculated by

$$\hat{\tau}_{21} = \arg \max_p r_{12}(p), \quad (5)$$

where $p \in [-\tau_{\max}, \tau_{\max}]$ and τ_{\max} is the maximum possible delay about signals. In digital implementation of TDOA calculation, we need some approximations. Supposing at time instant t , we have a set of observation samples of x_n , $\{x_n(t), x_n(t+1), \dots, x_n(t+k-1), \dots, x_n(t+K-1)\}$, $n=1,2$, corresponding CCF can be

estimated by (6) and (7), respectively:

$$\hat{r}_{12}(p) = \begin{cases} \frac{1}{K} \sum_{k=0}^{K-p-1} x_1(t+k)x_2(t+k+p), & p \geq 0, \\ \hat{r}_{12}(-p), & p < 0, \end{cases} \quad (6)$$

$$\hat{r}_{12}(p) = \begin{cases} \frac{1}{K-p} \sum_{k=0}^{K-p-1} x_1(t+k)x_2(t+k+p), & p \geq 0, \\ \hat{r}_{12}(-p), & p < 0. \end{cases} \quad (7)$$

Another way to estimate cross-correlation between signals is the use of Discrete Fourier Transform (DFT) and the Inverse Discrete Fourier Transform (IDFT) as shown in (8) and (9):

$$\hat{r}_{12}(p) = \frac{1}{K} \sum_{k'=0}^{K-1} X_1(w_{k'})X_2(w_{k'})e^{jw_{k'}p}, \quad (8)$$

$$w_{k'} = \frac{2\pi k'}{K}, \quad (9)$$

where $k' = 0, 1, \dots, K-1$, $w_{k'}$ is the angular frequency

$$X_n(w_{k'}) = \sum_{k=0}^{K-1} x_n(t+k)e^{-jw_{k'}p}. \quad (10)$$

In the frequency domain, DFT of $x_n(k)$ is formed as [14]:

$$X_n(w_{k'}) = \sum_{k=0}^{K-1} [a_n s_n(k-D) + b_n(k)]e^{-jw_{k'}p}, \quad (11)$$

$$X_n(w_{k'}) = a_n S(w_{k'})e^{-jw_{k'}p} + B_n(w_{k'}), \quad (12)$$

which is a Gaussian random variable. It can be illustrated by:

$$E[X_n(w_i)X_n^*(w_j)] = \begin{cases} a_n^2 P_s(w_i) + P_{b_n}(w_i), & i = j, \\ 0, & i \neq j, \end{cases} \quad (13)$$

$$E[X_1(w_i)X_2^*(w_j)] = \begin{cases} a_1 a_2 P_s(w_i) e^{-jw_i(D_1 - D_2)} + P_{b_n}(w_i), & i = j, \\ 0, & i \neq j, \end{cases} \quad (14)$$

where $n=1,2,\dots,N$ the power spectral densities of $s(k)$, $b_n(k)$, $x_1(k)$, and $x_2(k)$ are shown in (15)–(18) respectively:

$$P_s(w_i) = E[S(w_i)S^*(w_j)], \quad (15)$$

$$P_{b_n}(w_i) = E[B_n(w_i)B_n^*(w_j)], \quad (16)$$

$$P_{x_1}(w_i) = E[X_1(w_i)X_1^*(w_j)], \quad (17)$$

$$P_{x_2}(w_i) = E[X_2(w_i)X_2^*(w_j)]. \quad (18)$$

The commonly used weighting functions in GCC method are shown in Table III.

TABLE III. COMMON GCC FUNCTIONS.

Method Name	Weighting Function
Cross correlation (Unfiltered)	1
Roth Processor (ROTH)	$\frac{1}{P_{x_1}(w_k)}$
Smoothed coherence Transform (SCOT)	$\frac{1}{\sqrt{P_{x_1}(w_k)P_{x_2}(w_k)}}$
Modified SCOT ($n=scale\ factor$)	$\frac{1}{[P_{x_1}(w_k)P_{x_2}(w_k)]^n}$

In this paper, we have a moving sound source for localization. Excitation Source Information (ESI) based time-delay estimation algorithm is used for determining the time lag between sound channels [14], [15].

The optimal GCC method is determined by some experimental studies for this database. The error percentage is equal to the difference between the real number of sample delay and estimated sample delay. According to this information, we calculate the error percentage versus Signal-to-Noise Ratio (SNR) to determine the robustness of GCC methods. This comparison is implemented by our sound recording database. The result of this comparison is shown in Fig. 7. The approach of ESI has more successful results than other GCC methods on our sound dataset as shown in Fig. 7. The mean error percentage values for different GCC approach is illustrated in Table IV.

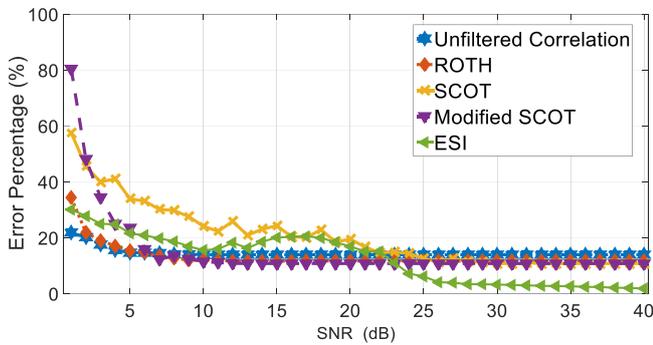


Fig. 7. Comparison of GCC methods.

TABLE IV. MEAN ERROR OF GCC METHODS.

Method Name	Mean Error (%)
Cross correlation	14.505
ROHT Filter	13.101
SCOT	20.511
SCOT-Modified	15.052
ESI	12.627

The accuracy of sample delay estimation has a vital importance for determining the exact location of sound sources. The user only wants to know the distance between microphones and time difference between received sounds [16], [17]. By means of this information, the azimuth angle of the sound source can be calculated easily. The illustration of ITD based on the determination of azimuth angle is shown in Fig. 8.

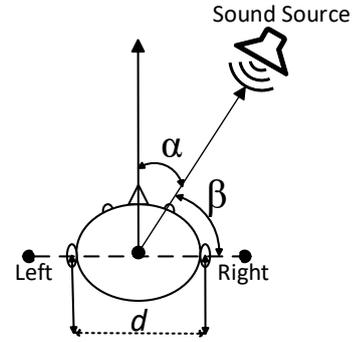


Fig. 8. ITD and azimuth angle.

One of the most important parameters for sound localization is the speed of the sound. This parameter directly affects the performance of sound source localization.

The speed of sound is highly sensitive to environment conditions. The generic application of sound localization techniques, sound speed is selected as a constant ($V_s \approx 343\ meter / second$). In addition, the speed of sound can be calculated as a function of environment temperature. The calculation of sound speed is shown in (19). The estimation of azimuth angle via ITD is also given in (20). In this equation, delay between samples, time of sample and distance between receivers are represented by s_d , T_s , and $d(meter)$, respectively [18]. The parameter of C_t represents ambient temperature in Celsius:

$$V_s = 20.05\sqrt{273.15 + C_t}, \quad (19)$$

$$\alpha = \arcsin\left(\frac{V_s s_d T_s}{d}\right). \quad (20)$$

The ITD based sound localization is illustrated in Fig. 9. The intersection points of azimuth angles provide the exact sound source location. Generally, the estimation of correct delay between sound channels is very hard due to environmental noise and echo effects. Since there will be a deviation between estimated and actual locations. It depends on the estimation success of the algorithm. In this approach, the users have to know only the delay between sound channel and locations of microphones.

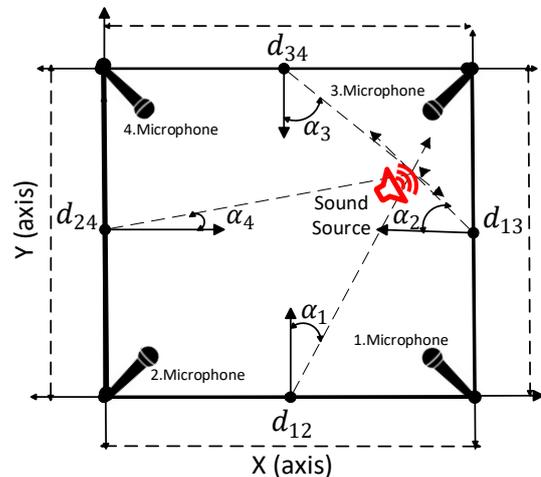


Fig. 9. ITD based 2D dimensional sound localization.

V. TIME DIFFERENCE OF ARRIVAL

Another approach for sound source localization is the time difference of arrival. The generic idea of Time Difference of Arrival (TDOA) is to determine the relative arrive time differences between receivers. This approach can be easily applied to the time difference between receiving channels and exact location of receivers [19], [20]. This approach is commonly used in the area of military, sound localization, GSM, and wireless sensor networks etc. The TDOA based 3D sound source localization can be defined as optimization problem [21].

The optimal estimated location is the value minimizing the expression in (21)

$$\min_{x_s, y_s, z_s} \sum_{i=1}^M \left(d_i^2 - (x_s - x_i)^2 - (y_s - y_i)^2 - (z_s - z_i)^2 \right)^2, \quad (21)$$

where (x_i, y_i, z_i) – location of microphones, d_i – distance between sound source and microphones, (x_s, y_s, z_s) – location of sound the source, M – number of microphones ($M \geq 3$).

This is a quadratic and unconstrained optimization problem to solve [22]. The open form of (21) is defined as in (22)–(25) for four microphones:

$$d_1 = \sqrt{(x_1 - x_{s_e})^2 + (y_1 - y_{s_e})^2 + (z_1 - z_{s_e})^2}, \quad (22)$$

$$d_2 = \sqrt{(x_2 - x_{s_e})^2 + (y_2 - y_{s_e})^2 + (z_2 - z_{s_e})^2}, \quad (23)$$

$$d_3 = \sqrt{(x_3 - x_{s_e})^2 + (y_3 - y_{s_e})^2 + (z_3 - z_{s_e})^2}, \quad (24)$$

$$d_4 = \sqrt{(x_4 - x_{s_e})^2 + (y_4 - y_{s_e})^2 + (z_4 - z_{s_e})^2}, \quad (25)$$

where $(x_{s_e}, y_{s_e}, z_{s_e})$ – estimated location of sound source.

In this work, four microphones are used for sound source localization and *six* different combinations are obtained for this relationship. This combination based distances are shown in (26)

$$d_{ij} = \sqrt{(d_i - d_j)^2}. \quad (26)$$

These distances can be easily transformed to sample delay difference s_{ij} between microphones as (27) and V_s is easily calculated depending on ambient temperature via (19). The sampling frequency (f_s) is 44100 Hz for this study

$$s_{ij} = \frac{d_{ij} f_s}{V_s}. \quad (27)$$

The TDOA approach allows determining the exact location of the sound source using the delays between the audio channels. The realization of the positioning problem with the TDOA approach is equivalent to an optimization problem. The solution of TDOA problem is performed by modified Levenberg-Marquardt algorithm [22]–[24].

VI. 3D MOVING SOUND LOCALIZATION

In this section, we mention about our approach for 3D moving sound source localization with some of the data mining and signal processing methods. The process of 3D sound source localization consists of several steps. In the first step, signals which increase SNR ratio for better signal representation and correct time delay estimation are smoothed by Savitzky-Golay filter [25]. The sample result of Savitzky-Golay filter is shown in Fig. 10. As shown in the figure, Savitzky-Golay filter is suitable for signal smoothing and the user can obtain original signal without greatly distorting the signal.

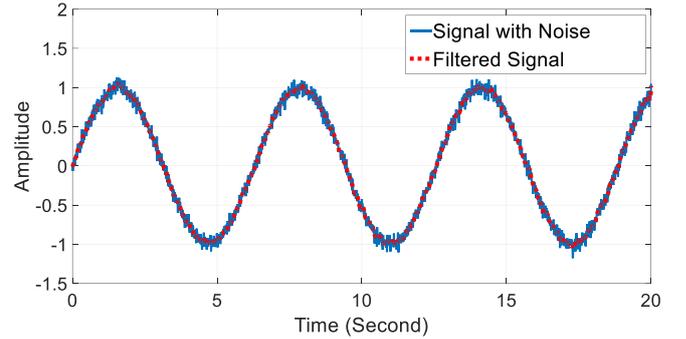


Fig. 10. Signal smoothing via Savitzky-Golay filter.

The second stage of the process is to improve difference between the sound source signal and environment noises. We use a threshold for the filtered signal in order to clarify info of locations. Local maxima points are determined by adaptive thresholding method [26]. The results of filtered signal and threshold applied signal are shown in Fig. 11.

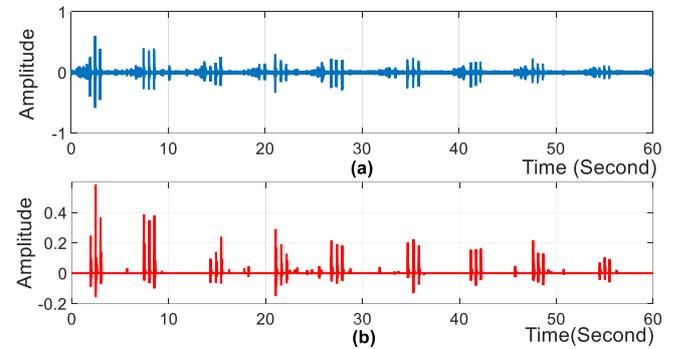


Fig. 11. Filtered signal (a); Threshold applied signal (b).

After this stage, it is required to determine a reference point for calculating correct time delay between sound channels. The k-medoids based clustering is applied to each sound signal outputs with thresholding applied [27]. Centroid points are also determined for all sound channels and calculated mean points for centroid. K-medoids algorithm is utilized for two purposes in this work. The first one is to determine the reference point for all sound channels and the other one is to obtain the exact number of sound source location.

The determination of exact location number is shown in Fig. 12. This graphic is obtained by a number of locations (cluster) versus total distances between medoids and observation points for summation of all sound channels. As shown in the figure, there is an obvious elbow point, which

gives an idea about an exact number of location adaptively. As expected, the number of sound location is $N=10$ in this study.

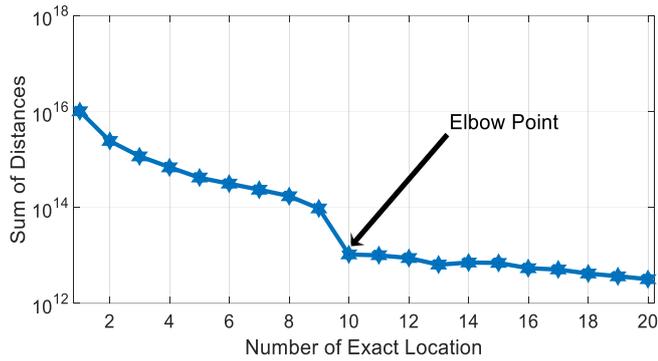


Fig. 12. Exact number of location determination.

Furthermore, the centers of clusters which are our mask to obtain a reference for all the locations for sound signals are obtained by clustering. These reference points are calculated as shown in (28). The centroid points of sound channels are defined as c_{1_i} , c_{2_i} , c_{3_i} , and c_{4_i} . The number of sound channel is followed as $c_n = 4$. The size of sample windows can be easily selected by these reference points. As known the number of location is calculated as $N = 10$:

$$Ref_i = \frac{(c_{1_i} + c_{2_i} + c_{3_i} + c_{4_i})}{c_n}, \quad (28)$$

$$W = \min(Ref_{i+1} - Ref_i), \quad (29)$$

where $i = 1, 2, 3, \dots, N-1$. The size selection of sample windows (W) is given in (29). The differences between reference points are calculated and a small safe merge is added to avoid overlap on sample window. The size of sample window and the reference line are represented in Fig. 13.

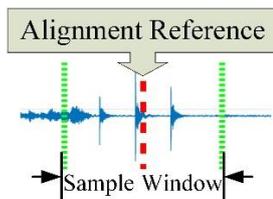


Fig. 13. Sample window size and Reference line.

The estimated reference lines and sample windows are also shown in Fig. 14.

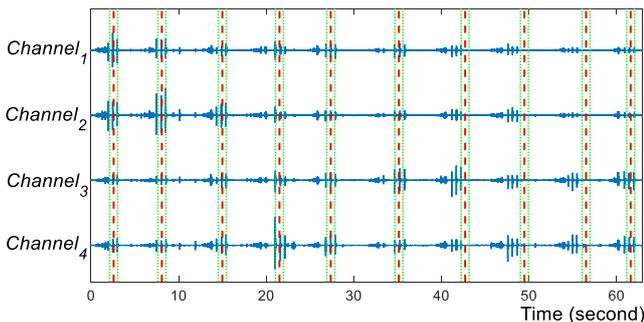


Fig. 14. Reference lines and optimal masking for sound signals.

In the proposed algorithm, 3D moving sound source sequence localization can be described as 7 steps as follows:

Step 1. Implementation of Savitzky-Golay pre-filter for input sound signal smoothing.

Step 2. Determination of local maxima points about sound signals via adaptive threshold algorithm.

Step 3. Data clustering algorithm is performed on calculated local maxima to determine optimal masking parameters.

Step 4. Determine an exact number of location for localization process.

Step 5. Obtain time delay between sound channels via excitation source information based time-delay estimation.

Step 6. Implementation of TDOA algorithm via modified Levenberg-Marquardt and determination of sound source location points in 3D space as coordinates (x_s, y_s, z_s) .

In this paper, only ITD parameter is used for determining sound source localization. The parameter of ITD is much more robust and reliable when compared to a parameter of ILD. The difference between the outputs of our algorithm and the results of real coordinates is given in Table V.

TABLE V. 2D AND 3D DISTANCE BETWEEN EXACT AND ESTIMATED LOCATIONS.

Number of Location	Exact Position	2D distance (cm)	3D distance (cm)
1	1	62.846	77.792
2	2	29.392	46.591
3	5	87.648	107.306
4	8	42.151	91.197
5	7	26.710	54.154
6	6	38.188	59.702
7	9	58.525	92.207
8	10	81.199	98.872
9	11	78.118	129.786
10	12	93.358	121.419

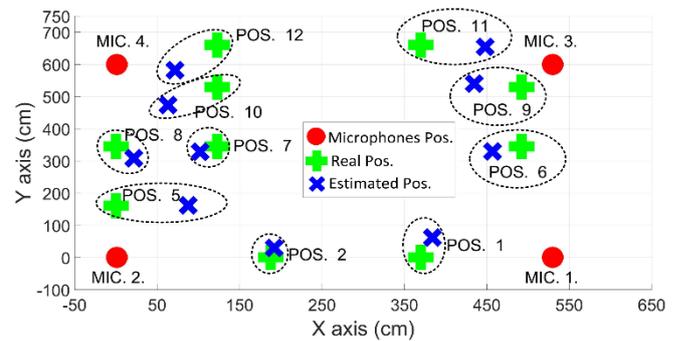


Fig. 15. 2D sound source localization results.

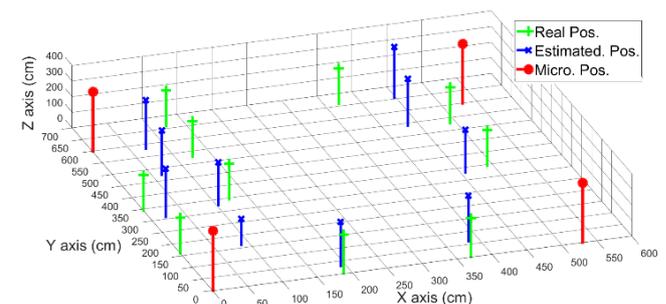


Fig. 16. 3D sound source localization results.

The mean errors in 2D and 3D distance estimation are 59.814 cm and 87.902 cm, respectively. These results are acceptable since the sound source is moved during the whole process and the reference points for the time delay are estimated adaptively.

The 2D visualization of sound source localization results is shown in Fig. 15. The 3D illustration of localization results is shown in Fig. 16.

VII. CONCLUSIONS

In this paper, we presented an algorithm to determine the moving sound sequence source localization for static systems. The generic idea of this study is to define suitable observation screen via basic data mining methods and implemented on sound sources. The excitation source information based on time-delay estimation algorithm is used for determining the time lag between sound channels. We describe some of the basic terms of sound signals and sound source localization methods. Besides, TDOA based on sound source localization is utilized for sound sequence localization. The results of axis Z are not good enough since all microphones are on the same plane. Also, 2D results are very acceptable for moving sound sources. The solution of TDOA problem is performed by modifying Levenberg-Marquardt algorithm in this work. After some processes, the clustering algorithm is used to obtain an exact number of locations of sound source adaptively. In addition, we share a new example for sound source localization, separation, and determination for different types of scenarios and explain all of the specifications about this database. It is observed that this paper is very helpful for researchers worked on sound processing and localization especially.

REFERENCES

- [1] A. N. Popper, R. R. Fay, *Sound source localization*. New York, USA: Springer, 2005. [Online]. Available: <https://dx.doi.org/10.1007/0-387-28863-5>
- [2] A. Deleforge, R. Horaud, "2D sound-source localization on the binaural manifold", *IEEE Int. Workshop on Machine Learning for Signal Processing*, 2012, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/MLSP.2012.6349784>
- [3] D. Pavlidi, S. Delikaris-Manias, V. Pulkki, A. Mouchtaris, "3D localization of multiple sound sources with intensity vector estimates in single source zones", *23rd European Signal Processing Conf. (EUSIPCO)*, Nice, 2015, pp. 1556–1560. [Online]. Available: <https://doi.org/10.1109/EUSIPCO.2015.7362645>
- [4] S. T. Birchfield, R. Gangishetty, "Acoustic localization by interaural level difference", *IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2005. [Online]. Available: <https://doi.org/10.1109/ICASSP.2005.1416207>
- [5] M. C. Catalbas, M. Yildirim, A. Gulten, H. Kurum, S. Dobrisek, "Estimation of trajectory and location for mobile sound source", *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 7, no. 9, pp. 237–241, 2016. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2016.070934>
- [6] D. A. Boyd, C. Hardy, *Understanding microphones*. In D. Boyd, S. Cohen, B. Rakerd, & D. Rehberger (Eds.), *Oral history in the digital age*. Institute of Library and Museum Services, 2012. [Online]. Available: <http://ohda.matrix.msu.edu/2012/06/understanding-microphones>
- [7] T. Lukaszewicz, Z. Kidon, D. Kania, K. Pethe-Kania, "Postural symmetry evaluation using wavelet correlation coefficients calculated for the follow-up posturographic trajectories", *Elektronika ir Elektrotechnika*, vol. 22, no. 5, pp. 84–88, 2016. [Online]. Available: <http://dx.doi.org/10.5755/j01.eie.22.5.16349>
- [8] H. Ziegelwanger, P. Majdak, W. Kreuzer, "Numerical calculation of listener-specific head-related transfer functions and sound localization: Microphone model and mesh discretization", *The Journal of the Acoustical Society of America*, vol. 138, no. 1, pp. 208–222, 2015. [Online]. Available: <http://dx.doi.org/10.1121/1.4922518>
- [9] B. Laback, "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors", pp. 488–500, 2004. [Online]. Available: <http://dx.doi.org/10.1097/01.aud.0000145124.85517.e8>
- [10] T. Hidaka, "Interaural cross-correlation, lateral fraction, and low-and high-frequency sound levels as measures of acoustical quality in concert halls", *The Journal of the Acoustical Society of America*, vol. 98, no. 2, pp. 988–1007, 1995. [Online]. Available: <http://dx.doi.org/10.1121/1.414451>
- [11] A. Pourmohammad, S. M. Ahadi, "N-dimensional N-microphone sound source localization", *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 27, no. 1, pp. 1–19, 2013. [Online]. Available: <http://dx.doi.org/10.1186/1687-4722-2013-27>
- [12] T. Padoisa, F. Sgardb, O. Doutsrea, A. Berry, "Acoustic source localization using a polyhedral microphone array and an improved generalized cross-correlation technique", *Journal of Sound and Vibration*, vol. 386, pp. 82–99, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.jsv.2016.09.006>
- [13] Z. S. Velickovic, V. D. Pavlovic, "The performance of the modified gcc technique for differential time delay estimation in the cooperative sensor network", *Elektronika ir Elektrotechnika*, vol. 19, no. 8, pp. 119–122, 2013. [Online]. Available: <http://dx.doi.org/10.5755/j01.eie.19.8.2445>
- [14] V. C. Raykar, R. Duraiswami, B. Yegnanarayana, S. R. Mahadeva Prasanna, "Tracking a moving speaker using excitation source information", *European Conf. Speech Communication and Technology*, pp. 69–72, 2003.
- [15] V. C. Raykar, B. Yegnanarayana, S. R. M. Prasanna, R. Duraiswami, "Speaker localization using excitation source information in speech", *IEEE Trans. Speech and Audio Processing*, vol. 13, no. 5, pp. 751–761, 2005. [Online]. Available: <https://doi.org/10.1109/TSA.2005.851907>
- [16] W. G. Gardner, *3-D audio using loudspeakers*. Springer Science & Business Media, 1998.
- [17] L. Calmes, "Biologically inspired binaural sound source localization and tracking for mobile robots", PhD dissertation, RWTH Aachen Univ., 2009.
- [18] C. Rascon, H. Aviles, L. A. Pineda, "Robotic orientation towards speaker for human-robot interaction", *Ibero-American Conf. Artificial Intelligence*, 2010, pp. 10–19. [Online]. Available: https://doi.org/10.1007/978-3-642-16952-6_2
- [19] F. Gustafsson, F. Gunnarsson, "Positioning using time-difference of arrival measurements", in *Proc. Acoustics, Speech, and Signal Processing (ICASSP 2003)*, 2003. [Online]. Available: <https://doi.org/10.1109/ICASSP.2003.1201741>
- [20] J. S. Hamdoun, A. Rachedi, A. Benslimane, "Comparative analysis of RSSI-based indoor localization when using multiple antennas in Wireless Sensor Networks", in *Int. Conf. Selected Topics in Mobile and Wireless Networking (MoWNeT 2013)*, 2013, pp. 146–151.
- [21] K. Shoda, M. Arakawa, M. Morikawa, T. Hisano, K. Matsumura, "A 3D location estimation method using the Levenberg-Marquardt method for real-time location system", *WCSMO-10*, 2013.
- [22] J. E. Dennis, R. B. Schnabel, *Numerical methods for unconstrained optimization and nonlinear equations*. Siam, 1996. [Online]. Available: <http://dx.doi.org/10.1137/1.9781611971200>
- [23] L. Chen, "A modified Levenberg–Marquardt method with line search for nonlinear equations", *Computational Optimization and Applications*, vol. 65, no. 3, pp. 753–779, 2016. [Online]. Available: <http://dx.doi.org/10.1007/s10589-016-9852-y>
- [24] M. Balda, "LMFsolve. m: Levenberg-Marquardt-Fletcher algorithm for nonlinear least squares problems", 2009. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/16063-lmfsolve-m--levenberg-marquardt-fletcher-algorithm-for-nl-least-squares-problems>
- [25] R. W. Schafer, "What is a Savitzky-Golay filter? [Lecture notes]", *IEEE Signal Processing Magazine*, vol. 28, no. 4, pp. 111–117, 2011. [Online]. Available: <https://doi.org/10.1109/MSP.2011.941097>
- [26] T. O'Haver, "A pragmatic introduction to signal processing with applications in scientific measurement", University of Maryland at College Park, 2015.
- [27] P. N. Tan, M. Steinbach, V. Kumar, *Introduction to data mining*. India: Pearson Education, 2006.