Self-Synchronizing Acoustic Positioning System Based on TDoA

Shuai Cao [1], Ya P. Liu [1], Jian Xu [1], and Sheng B. Pei [2]

[1] *Institutes of Physical Science and Information Technology, Anhui University, Hefei 230039, China*
[2] *School of Computer Science and Technology, Anhui University, Hefei 230039, China*

Abstract
A self-synchronizing system for acoustic positioning using low-cost acoustic devices is proposed in this paper. It includes a master base station and multiple slave base stations. During a positioning cycle, the master base station first transmits an audio signal with synchronization and positioning functions. Then, after detecting the audio signal transmitted by the master base station, each slave base station transmits the positioning audio signal after delaying the set time. Finally, the target realizes its own position estimation by detecting the arrival time of the audio signal transmitted by the master base station and each slave base station. The proposed system realizes self-synchronization by using known information such as the distance between base stations and the delay time of each slave base station. After synchronization, the audio arrival time of each base station received by the target is used for position estimation. Compared with the previous synchronization method, the adopted self-synchronization method avoids complicated wiring and radio interference and greatly reduces the cost of synchronization. The simulation results show that under the test conditions when the detection noise level of the target is less than 2.5 ms, the localization accuracy of the proposed system is better than 1 m.

Keywords 1
acoustic positioning, self-synchronizing, master base station, slave base station

1. Introduction

Location-aware technology has important applications in smart cities, the Internet of Things, medical monitoring, etc. Based on the type of signals used for positioning, location-aware technology is mainly divided into radio, motion signal, geomagnetic, image, audio, etc. In radio positioning technology, Bluetooth and WIFI based on received signal strength indication (RSSI) usually use signal attenuation model ranging or fingerprint method to achieve positioning, which has low positioning accuracy (meter level) [1]-[2], the large workload of offline fingerprint collection and update, and signal attenuation model is highly susceptible to environmental interference. The angle of arrival (AoA) based on antenna array[3], WIFI-based round-trip time (RTT)[4]-[6] and channel status information (CSI)[7], and ranging-based UWB technology[8] can provide centimeter-level to sub-meter-level high-precision positioning, but wireless base stations and smart terminals are required to support related protocols, the deployment complexity is high, and the wide-area coverage requires extremely high costs. Pedestrian dead reckoning (PDR) based on motion signal [9] has the advantages of not requiring additional infrastructure and being compatible with mobile phones, but there is an accumulation of errors, and accurate positioning cannot be achieved for a long time. Geomagnetic-based positioning [10] can use the mobile phone's magnetic sensor to achieve positioning without additional infrastructure but requires the offline collection of the indoor geomagnetic field distribution. Image-based localization method [11] can achieve high-precision localization of the target, but it needs to establish an image feature library in advance. Besides, the hardware cost is high, the

calculation is complex, and the performance is easily affected by environmental textures and shooting conditions.

The audio-based positioning technology [12] uses acoustic devices such as microphones and speakers to achieve positioning. To support high concurrency, acoustic positioning systems (APS) usually adopt a passive architecture, in which the base stations only transmit audio and the targets receive and process the audio. The characteristics of this architecture are 1) Low system building complexity. Since the propagation speed of audio is much lower than that of radio, the requirement of clock synchronization accuracy to achieve high-precision acoustic positioning is not very high, 2) Low system cost. On the one hand, due to the low cost of commercial audio components, although the realization of acoustic positioning requires the deployment of acoustic base stations, location service providers are expected to provide users with high-precision location services through low-cost infrastructure investment. Furthermore, if the positioning of smart terminals is realized with the help of broadcast speakers in large shopping malls, stations, and airports, there is no need to deploy additional base stations. On the other hand, for the users, the microphones and the speakers are the standard configurations of the hand-held mobile terminals, and there is no need to provide additional overhead for users, 3) Good privacy. The targets can realize their own positioning only by receiving signals without information exchanges with the acoustic base stations. Without the user's permission, the user's location cannot be obtained by others, which is suitable for occasions with high privacy requirements, and 4) Suitable for positioning occasions with severe electromagnetic interference and metal substances. Compared with radio signals, the transmission of audio signals will not be disturbed by electromagnetic waves, nor is it not affected by metal substances.

Because synchronization between base stations and targets is not required, time difference of arrival (TDoA) is a common measurement type used in high-precision acoustic positioning systems. Accurate acquisition of TDoA is the key to ensuring the performance of the acoustic positioning system, which not only requires the target to accurately detect the audio arrival time but also requires accurate synchronization between acoustic base stations. The acoustic positioning systems involved in previous studies mainly use wireless or wired methods to achieve synchronization [13]-[16]. The wireless method uses radio to achieve synchronization between the base stations, which may cause radio interference. The wired method uses connection lines to achieve synchronization between the base stations, which is difficult to deploy. Both above synchronization methods are costly.

In this paper, a novel self-synchronizing acoustic positioning system is proposed, which consists of a master base station and multiple slave base stations. In a positioning cycle, the master base station first transmits an audio signal with synchronization and positioning functions. Then, each slave base station delays the set time after detecting the audio signal transmitted by the master base station and starts to transmit the positioning audio signal. Finally, the target realizes its real-time position estimation by detecting the arrival time of the audio signals transmitted by the master base station and each slave base station. The proposed acoustic positioning system utilizes the acoustic devices themselves to realize self-synchronization. Compared with the previous synchronization methods, complicated wiring and radio interference are avoided, and the synchronization cost is greatly reduced.

The rest of the paper is organized as follows. The basic principle of proposed self-synchronizing acoustic positioning systems is presented in Section II. The numerical simulations are described in Section III, and a conclusion is drawn in Section IV.

2. Self-synchronizing Acoustic Positioning

The basic principle described below takes two-dimensional positioning as an example and is also applicable to three-dimensional situations. As shown in Figure 1, it is assumed that there are four acoustic base stations in the positioning area, which are represented by A, B, C, and D respectively. The coordinates of the acoustic base stations are known, and each acoustic base station contains a speaker and a microphone. To ensure high concurrency, the target only uses its microphone to passively receive audio signals and realizes position calculation by detecting the arrival time of the audio transmitted by each acoustic base station.

The proposed self-synchronizing acoustic positioning system consists of a master base station and multiple slave base stations. In a positioning cycle, the master base station first transmits an audio

signal that will be received by the target and each slave base station. Then, after detecting the arrival time of the audio signal transmitted by the master base station, each slave base station delays the set time and immediately emits the positioning audio signal. Finally, the target detects the time of arrival of audio transmitted by all base stations, calculates the TDoA, and estimates the position using a TDoA-based localization algorithm. Taking the timing diagram as an example, as shown in Figure 2, the positioning processes of the self-synchronizing acoustic positioning system are described as follows:
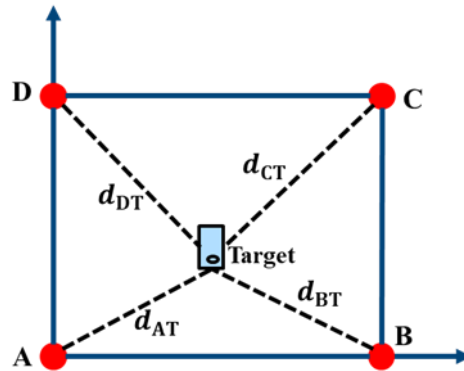

Figure 1: Schematic diagram of the self-synchronizing acoustic positioning system
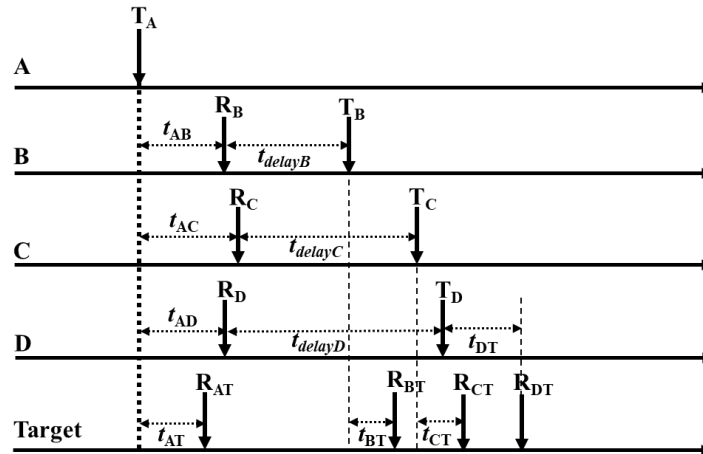

Figure 2: Timing diagram of the self-synchronizing acoustic positioning system

1. In a positioning period, such as 1 second, the master base station A transmits an audio signal at time TA. To suppress background noise, the audio signal above is modulated, such as chirp signal, orthogonal code modulation signal, etc.

2. After receiving the audio signal transmitted by base station A, base stations **B**, **C**, and **D** detect the arrival times of the audio, which are represented by R_B, R_C, and R_D respectively. The target also detects the arrival time of the audio signal transmitted by base station **A**, which is represented by R_{AT}. In Figure 2, t_{AB}, t_{AC}, and t_{AD} are used to represent the time taken for the audio transmitted by the base station **A** to reach the base stations **B**, **C**, and **D**, respectively, we have:

$$\begin{cases} t_{AB} = R_B - T_A \\ t_{AC} = R_C - T_A \\ t_{AD} = R_D - T_A \end{cases} \tag{1}$$

3. Delayed by t_{delayB} from moment R_B, namely at moment T_B, base station **B** transmits positioning audio signal, which is then received by the target at moment R_{BT}. t_{BT} is the flight time for the positioning audio signal transmitted by base station **B** to reach the target. Similarly, Delayed by t_{delayC} from moment R_C, namely at moment T_C, base station **C** transmits

positioning audio signal, which is then received by the target at moment $\mathbf{R_{CT}}$. $t_{\mathbf{CT}}$ is the flight time for the positioning audio signal transmitted by base station **C** to reach the target. For base station **D**, the processing method is the same as above, and the corresponding time parameters are t_{delayD}, $\mathbf{T_D}$, $\mathbf{R_{DT}}$ and $t_{\mathbf{DT}}$.

4. According to Figure 2, formula (2) can be obtained:

$$\begin{cases} \mathbf{R_{BT}} - \mathbf{R_{AT}} = t_{\mathbf{AB}} + t_{delayB} + t_{\mathbf{BT}} - t_{\mathbf{AT}} \\ \mathbf{R_{CT}} - \mathbf{R_{AT}} = t_{\mathbf{AC}} + t_{delayC} + t_{\mathbf{CT}} - t_{\mathbf{AT}} \\ \mathbf{R_{DT}} - \mathbf{R_{AT}} = t_{\mathbf{AD}} + t_{delayD} + t_{\mathbf{DT}} - t_{\mathbf{AT}} \end{cases} \quad (2)$$

Formula (3) is derived from formula (2):

$$\begin{cases} t_{\mathbf{BT}} - t_{\mathbf{AT}} = (\mathbf{R_{BT}} - \mathbf{R_{AT}}) - (t_{\mathbf{AB}} + t_{delayB}) \\ t_{\mathbf{CT}} - t_{\mathbf{AT}} = (\mathbf{R_{CT}} - \mathbf{R_{AT}}) - (t_{\mathbf{AC}} + t_{delayC}) \\ t_{\mathbf{DT}} - t_{\mathbf{AT}} = (\mathbf{R_{DT}} - \mathbf{R_{AT}}) - (t_{\mathbf{AD}} + t_{delayD}) \end{cases} \quad (3)$$

Multiply both sides of the equal sign of formula (3) by the speed of sound v (unit: m/s) to obtain formula (4).

$$\begin{cases} d_{\mathbf{BT}} - d_{\mathbf{AT}} = v \cdot (\mathbf{R_{BT}} - \mathbf{R_{AT}}) - (d_{\mathbf{AB}} + v \cdot t_{delayB}) \\ d_{\mathbf{CT}} - d_{\mathbf{AT}} = v \cdot (\mathbf{R_{CT}} - \mathbf{R_{AT}}) - (d_{\mathbf{AC}} + v \cdot t_{delayC}) \\ d_{\mathbf{DT}} - d_{\mathbf{AT}} = v \cdot (\mathbf{R_{DT}} - \mathbf{R_{AT}}) - (d_{\mathbf{AD}} + v \cdot t_{delayD}) \end{cases} \quad (4)$$

where $d_{\mathbf{AT}}$, $d_{\mathbf{BT}}$, $d_{\mathbf{CT}}$ and $d_{\mathbf{DT}}$ are the distances between the target and base stations **A**, **B**, **C** and **D**, respectively. $d_{\mathbf{AB}}$, $d_{\mathbf{AC}}$, and $d_{\mathbf{AD}}$ are the distances between base station **A** and base stations **B**, **C**, and **D**, respectively. $\mathbf{R_{AT}}$, $\mathbf{R_{BT}}$, $\mathbf{R_{CT}}$ and $\mathbf{R_{DT}}$ are the arrival times extracted by the target from the received audio signal using the audio detection algorithm [17]. On the right side of formula (4), $d_{\mathbf{AB}}$, $d_{\mathbf{AC}}$, $d_{\mathbf{AD}}$, v, t_{delayB}, t_{delayC} and t_{delayD} are known, $(\mathbf{R_{BT}} - \mathbf{R_{AT}})$, $(\mathbf{R_{CT}} - \mathbf{R_{AT}})$ and $(\mathbf{R_{DT}} - \mathbf{R_{AT}})$ are the time interval of target detection, which are measured values and can be obtained according to the number of sampling points and sampling period. Therefore, the distance differences on the left side of the formula (4) are determined, i.e., with base station **A** as the reference point, the distance difference$(d_{\mathbf{BT}} - d_{\mathbf{AT}})$ between the target and base stations **B** and **A**, the distance difference$(d_{\mathbf{CT}} - d_{\mathbf{AT}})$ between the target and base stations **C** and **A**, and the distance difference$(d_{\mathbf{DT}} - d_{\mathbf{AT}})$ between the target and base stations **D** and **A** are obtained.

5. Substitute the distance difference information in formula (4) into the TDoA-based positioning algorithm to estimate the target's position.

For each positioning cycle, the acoustic position system repeats the above steps to achieve target positioning. In the whole process, the target only receives the audio signal transmitted by the base stations and does not interact with the base stations, which makes the system supports high concurrency, i.e., the number of users is not limited. This paper realizes a self-synchronization acoustic positioning system through the transmitted audio signal. There are no connection lines between the base stations, which is convenient for installation and layout. It also avoids the radio interference problem existing in traditional radio synchronization. Furthermore, the proposed synchronization method is low cost considering the acoustic components such as commercial loudspeakers, microphones, etc., are cheap.

3. Numerical Simulations

To analyze the positioning performance of the proposed system, simulations were performed to investigate the influence of the detection accuracy of the target on the position estimation accuracy.

The factors affecting the positioning accuracy of the proposed system include: (1) The errors of the arrival time of the audio signal transmitted by the master base station detected by the slave base stations, that is, the errors of $\mathbf{R_B}$, $\mathbf{R_C}$ and $\mathbf{R_D}$ in Figure 2, are represented by $e_{\mathbf{B}}$, $e_{\mathbf{C}}$ and $e_{\mathbf{D}}$ respectively. Since the coordinates of the base stations are fixed, the paths between the master and slave base stations are in the line-of-sight (LOS) state, the acoustic channel states are stable, and if the system is calibrated, the noise will be at a low level. (2) The errors of the delay t_{delayB}, t_{delayC}, and

t_{delayD} caused by the system clock deviations of the slave base stations are relatively small, so this simulation ignored the influence of this factor. (3) The detection errors of the audio detection algorithm used by the target, corresponding to the four base stations A, B, C, and D, are represented by e_{AT}, e_{BT}, e_{CT}, and e_{DT}, respectively. Due to the influence of strong multipath and non-line-of-sight (NLOS), these errors are a major factor in degrading system performance. This simulation mainly examines this factor. According to Figure 2, the noisy audio arrival times $\mathbf{R_{AT}}$, $\mathbf{R_{BT}}$, $\mathbf{R_{CT}}$, and $\mathbf{R_{DT}}$ are obtained by formula (5).

$$\begin{cases} \mathbf{R_{AT}} = \mathbf{T_A} + \dfrac{d_{AT}}{v} + e_{AT} \\[2mm] \mathbf{R_{BT}} = \mathbf{T_A} + \dfrac{d_{AB}}{v} + e_B + t_{delayB} + \dfrac{d_{BT}}{v} + e_{BT} \\[2mm] \mathbf{R_{CT}} = \mathbf{T_A} + \dfrac{d_{AC}}{v} + e_C + t_{delayC} + \dfrac{d_{CT}}{v} + e_{CT} \\[2mm] \mathbf{R_{DT}} = \mathbf{T_A} + \dfrac{d_{AD}}{v} + e_D + t_{delayD} + \dfrac{d_{DT}}{v} + e_{DT} \end{cases} \tag{5}$$

where d_{AT}, d_{BT}, d_{CT}, and d_{DT} are the distances between the base stations and the target, and v is the speed of sound and is set to 340 m/s.

In this simulation, formula (5) is substituted into formula (4) to calculate the noisy distance differences, which are then substituted into the TDoA-based combined weighted (COM-W) positioning algorithm[18] to obtain the estimated position. By comparing the localization error (LE) between the true position $P_r(x_r, y_r)$ and the estimated position $P_e(x_e, y_e)$, the influence of the detection accuracy of target on the system positioning accuracy is evaluated. LE is the Euclidean distance, which is calculated using (6).

$$LE = \sqrt{(x_e - x_r)^2 + (y_e - y_r)^2} \tag{6}$$

To avoid the influence of a single abnormal noise, it is usually necessary to add multiple noises to measurement at a specific test point. Each time the noise is added, the COM-W algorithm estimates a position and can use (6) to obtain its LE, and multiple noises can be added to obtain multiple LEs. The mean positioning error (MPE) is calculated using (7).

$$MPE = \sqrt{\frac{\sum_{i=1}^{L} LE_i^2}{L}} \tag{7}$$

where L represents the total times of adding noise, and LE_i is the LE of the position estimated by the positioning algorithm under the i-th noise addition. In this simulation, L is set to 1000.
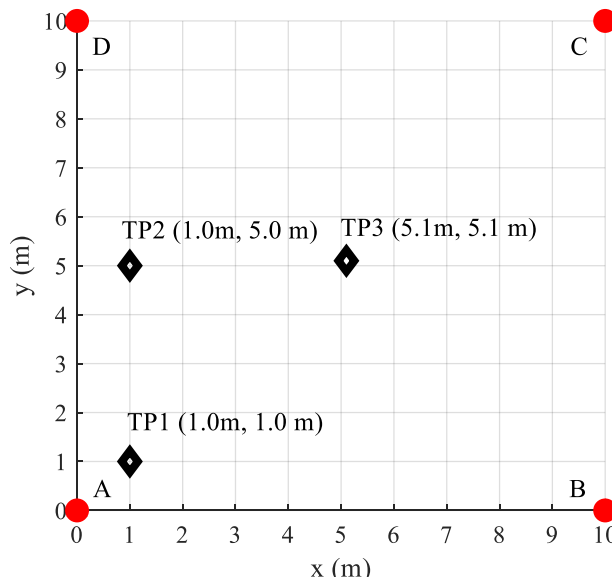

Figure 3: Base stations and test points

As shown in Fig 3, four base stations enclose a positioning area, and their coordinates are A(0 m, 0 m), B(10 m, 10 m), C(10 m, 10 m), and D(0, 10 m), respectively. Three representative test points were selected, namely TP1 (1.0 m, 1.0 m) near a base station, TP2 (1.0 m, 5.0 m) near the edge of the positioning area, and TP3 (5.1 m, 5.1 m) near the central area. In formula (5), it is assumed that e_B, e_C, and e_D conform to Gaussian white noise with a mean of zero and a standard deviation of 0.1 millisecond (ms), and e_{AT}, e_{BT}, e_{CT}, and e_{DT} conform to Gaussian white noise with a mean of zero and a standard deviation of σ (ms).
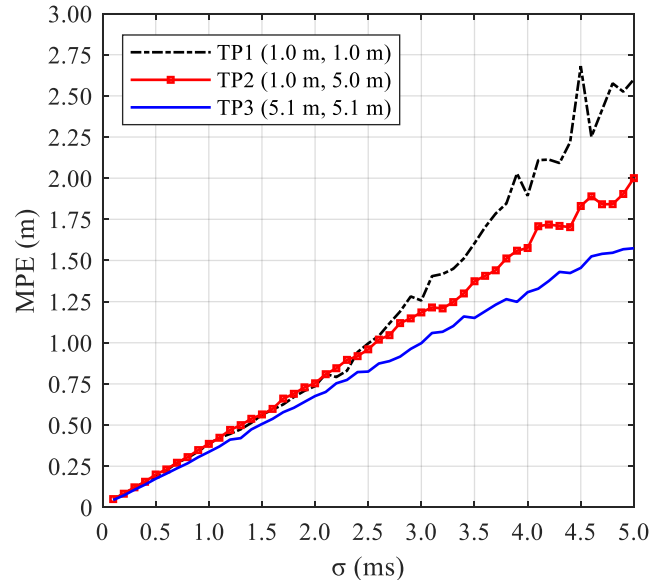

Figure 4: MPE versus standard deviation of target detection noise

Figure 4 shows the relationship between the MPE at the 3 test points and the standard deviation of the detection noise of the target. It can be seen that as the detection noise of the target increases, the MPE also increases, i.e., the accuracy of the position estimation of the target decreases. At low noise levels, the MPEs at the three test points are relatively close, but at high noise levels, the MPEs at the three test points are quite different, which is caused by the difference in geometric conditions[18]. From Figure 4, it can also be found that under the current simulation conditions, if the positioning accuracy of the self-synchronized acoustic positioning system is to be better than 1 m, the detection error of the target should be less than 2.5 ms.

4. Conclusion

This paper proposes a self-synchronizing acoustic positioning system. During a positioning period, a master base station transmits an audio signal for synchronization and positioning, multiple slave base stations receive the audio signal and then send positioning audio signals after delay setting times. Finally, the target receives the audio signals transmitted by the master base station and all slave base stations, detects their arrival times, and estimates the positions. The proposed system does not need to be synchronized by wire or radio, which reduces the cost of system layout and avoids radio interference. The simulation results verify that the proposed system can achieve positioning better than 1 m, if the detection error of the target is less than 2.5 ms.

5. References

[1] C. Wu, Z. Yang, Z. Zhou, Y. Liu, and M. Liu, "Mitigating Large Errors in WiFi-Based Indoor Localization for Smartphones," IEEE Trans. Veh. Technol., vol. 66, no. 7, pp. 6246-6257, Jul. 2017.

[2] J. Luo, Z. Zhang, C. Wang, C. Liu, and D. Xiao, "Indoor Multifloor Localization Method Based on WiFi Fingerprints and LDA," IEEE Trans. Ind. Infomat., vol. 15, no. 9, pp. 5225-5234, Sept. 2019.

[3] X. Qiu, B. Wang, J. Wang, and Y. Shen, "AOA-Based BLE Localization with Carrier Frequency Offset Mitigation," in Proc. IEEE Int. Conf. Commun. Workshops, (ICC Workshops), 2020, pp. 1-5.

[4] C. Gentner, M. Ulmschneider, I. Kuehner, and A. Dammann, "WiFi-RTT Indoor Positioning," in Proc. IEEE/ION Position, Locat. Navig. Symp., (PLANS), 2020, pp. 1029-1035.

[5] O. Hashem, M. Youssef and K. A. Harras, "WiNar: RTT-based Sub-meter Indoor Localization using Commercial Devices," in Proc. IEEE Int. Conf. Pervasive Comput. Commun., (PerCom), 2020, pp. 1-10.

[6] H. Cao, Y. Wang, and J. Bi, "Smartphones: 3D Indoor Localization Using Wi-Fi RTT," IEEE Commun. Lett., vol. 25, no. 4, pp. 1201-1205, Apr. 2021.

[7] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-Based Fingerprinting for Indoor Localization: A Deep Learning Approach," IEEE Trans. Veh. Technol., vol. 66, no. 1, pp. 763-776, Jan. 2017.

[8] F. Mazhar, M. G. Khan, and B. Sallberg, "Precise Indoor Positioning Using UWB: A Review of Methods, Algorithms and Implementations," Wireless Pers. Commun., vol. 97, no. 3, pp. 4467-4491, Dec. 2017.

[9] L. Ciabattoni, G. Foresi, A. Monteriu, L. Pepa, D. P. Pagnotta, L. Spalazzi, and F. Verdini, "Real time indoor localization integrating a model based pedestrian dead reckoning on smartphone and BLE beacons," J. Ambient Intell. Humanized Comput., vol. 10, no. 1, pp. 1-12, Jan. 2019.

[10] S. -C. Yeh, W. -H. Hsu, W. -Y. Lin, and Y. -F. Wu, "Study on an Indoor Positioning System Using Earth's Magnetic Field," IEEE Trans. Instrum. Meas., vol. 69, no. 3, pp. 865-872, Mar. 2020.

[11] M. Zhao, M. Yan, and T. Li, "Vision-Based Positioning: Related Technologies, Applications, and Research Challenges," in Proc. IEEE 9th Int. Conf. Software Eng. Serv. Sci., (ICSESS), 2018, pp. 531-535.

[12] M. N. Liu, L. S. Cheng, K. Qian, J. L. Wang, J. Wang, and Y. H. Liu, "Indoor acoustic localization: a survey," Hum.-Centric Comput. Inf. Sci., vol. 10, no. 1, Jan 6, 2020.

[13] S. I. Lopes, J. M. N. Vieira, J. Reis, D. Albuquerque, and N. B. Carvalho, "Accurate smartphone indoor positioning using a WSN infrastructure and non-invasive audio for TDoA estimation," Pervas. Mobile Comput., vol. 20, pp. 29-46, Jul. 2015.

[14] J. Urena, A. Hernandez, J. J. Garcia, J. M. Villadangos, M. C. Perez, D. Gualda, F. J. Alvarez, and T. Aguilera, "Acoustic Local Positioning with Encoded Emission Beacons," Proc. IEEE, vol. 106, no. 6, pp. 1042-1062, Jun. 2018.

[15] P. Pajuelo, M. C. Perez, J. M. Villadangos, E. Garcia, D. Gualda, J. Urena, and A. Hernandez, "Implementation of indoor positioning algorithms using Android smartphones," in Proc. IEEE 20th Conf. Emerging Technol. Factory Autom., (ETFA), 2015, pp. 1-4.

[16] L. Zhang, M. L. Chen, X. H. Wang, and Z. Wang, "TOA Estimation of Chirp Signal in Dense Multipath Environment for Low-Cost Acoustic Ranging," IEEE Trans. Instrum. Meas., vol. 68, no. 2, pp. 355-367, Feb. 2019.

[17] S. Cao, X. Chen, X. Zhang, and X. Chen, "Effective Audio Signal Arrival Time Detection Algorithm for Realization of Robust Acoustic Indoor Positioning," IEEE Trans. Instrum. Meas., vol. 69, no. 10, pp. 7341-7352, Oct. 2020.

[18] S. Cao, X. Chen, X. Zhang, and X. Chen, "Combined Weighted Method for TDOA-Based Localization," IEEE Trans. Instrum. Meas., vol. 69, no. 5, pp. 1962-1971, May. 2020.