

Bluetooth, Floor-plan, and MEMS Assisted Wide-area Audio Indoor Localization System: Apply to Smartphones

Shihao Xu, Ruizhi Chen, Guangyi Guo, Zheng Li, Long Qian, Feng Ye, Zuoya Liu, and Lixiong Huang

Abstract—During the past decade, there has been an increasing interest in the audio-based indoor localization approach, which has benefited from the high precision, high privacy, and high security of the passive time difference of arrival (TDoA) positioning. However, the limited bandwidth, and complexity in real-time ID encoding and decoding make it difficult to achieve wide-area ($>1500\text{m}^2$) localization in off-the-shelf smartphones. Plenty of previous work applied to a single region localization only with a small range. In this paper, a hybrid acoustic signal transmission architecture based on frequency division multiple access (FDMA), time division multiple access (TDMA), and space division multiple access (SDMA) is perfected to cover the acoustic signal to a wider range. With the prior information of the Bluetooth and floor-plan, smartphones estimate which cell themselves are located in accurately. An enhanced particle filter (PF) algorithm, along with the micro-electro-mechanical systems (MEMS)-based pedestrian dead reckoning (PDR) and the floor-plan restraint, was proposed to make the smartphones tracking smarter and more robust. The proposed prototype system provided sub-meter-level localization accuracy and stable 1Hz update rate in several typical wide indoor environments with less infrastructure, which demonstrates some values of reference and application owing to the work of this paper.

Index Terms—Indoor localization, wide-area positioning, acoustic signal, PDR, floor-plan map, particle filter.

I. INTRODUCTION

LOCATION-based services (LBS) have generated hundreds of millions of economic values in the past few decades. The global navigation satellite system (GNSS) already has the capacities in high precision positioning and timing full-weather and around the clock outdoors. However, the unreachable or low unavailable satellite signal indoors makes people need a reliable indoor positioning technique to achieve seamless

Manuscript received March 02, 2021; revised April 20, 2021; accepted August 26, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2016YFB0502200, and 2016YFB0502201, in part by the Postdoctoral Research Foundation of China under 2020M682480.

The authors are with the Department of Localization and Navigation, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, 430079, China (e-mail: shihao.xu@whu.edu.cn; ruizhi.chen@whu.edu.cn; guangyi.guo@whu.edu.cn; lizhenglwf@whu.edu.cn; yefeng92@whu.edu.cn; lqian1995@whu.edu.cn; zyliu0017@whu.edu.cn; lx.huang@whu.edu.cn) (Corresponding author: Guangyi Guo, phone: 157-7242-9786; e-mail: guangyi.guo@whu.edu.cn).

navigation. Considering the complex indoor topology and channel environment, and the arbitrary and changeable attitude of pedestrians, researchers concern with the solutions of accuracy better than 1m, robust, and adaptable to the public. Over the past few decades, various technologies around sound, light, electricity, and field have been proposed and pushed. Among which the radio frequency (RF) and visual signals have been utilized the most widely in mobile computing. The IEEE 802.11-2016 [1] introduced the Wi-Fi fine timing measurement (FTM) protocol, which achieves geometric positioning by high precision ($<3\text{m}$) round trip time (RTT) ranging [2, 3]. But it only supports the Pixel series smartphones and has a limitation on the number of users in the same scenario. Ultra-wide-band (UWB) has become popular in the industry. The Apple [4], Xiaomi [5], and Samsung [6] have implanted UWB chips to their smartphones. However, they only consider directivity and near detection and do not develop the capability of positioning at present. Bluetooth beacon [7] relies on the statistical characteristics of the received signal strength (RSS) to form fingerprints or modeling to ranging. Nevertheless, a single beacon has a small effective coverage ($<10\text{m}$) and the positioning accuracy is not high (1-5m) in general. Bluetooth 5.1 [8] uses signal phase and amplitude information to perform high-precision positioning ($<1\text{m}$) with low cost and power consumption. However, it is restrained with an antenna array and needs complex error compensation in advance. Simultaneous localization and mapping (SLAM) [9, 10] has been well applied on robots and drones, but smartphones are not yet able to efficiently calculate locations in real-time.

Be different from the RF and visual signals, the acoustic signal has advantages in low propagation speed, wide-coverage, and high availability, etc. In general, the audio-based localization can be divided into acoustic source localization and acoustic receiver localization. The former plays an important role in many practical speech acquisitions and recognition systems [11], and the latter focuses on real-time location service providing [12-16, 18-20, 32, 33]. For the latter, researchers adopted methods widely based on the time delay estimation (TDE), or more accurately, the methods of generalized cross-correlation (GCC) [12-14, 16, 18] to acquire the TDoA pairs from acoustic signals. The linear frequency modulation (LFM) [12-14] and the binary pulse amplitude modulation (2-PAM) [20] are commonly applied to generate the acoustic signal. The LFM signal, also known as the Chirp signal, is more widely used due to its ability to better against the

Doppler effect [17, 18, 32]. When a Chirp signal is transmitted to the microphone, it will be compared with the pre-stored digital template in a sliding time window and stop until matched at timestamp i . The timestamp i is then recorded as the first relative time of arrival (ToA) of the signal. After a while, when the next signal is matched, the second relative ToA₂ is acquired. Since all the anchors are synchronized in the same network, these two relative ToAs are subtracted to get a TDoA pair. A 2-D location then can be estimated with at least 2 TDoA pairs using hyperbolic positioning methods, such as Fang [21], Chan [22], and Taylor series expansion [23]. However, these methods suffer from problems in limited station number requirements, or poor accuracy under none-line-of-sight (NLOS) or multipath circumstances.

Comparing with single data source positioning, the multi-source fusion positioning methods typically achieve higher positioning accuracy and stability. Different from the Kalman filter and its extend filters, which rely on the Gaussian distribution of errors, the particle filter (PF) [24, 25] based on Bayesian statistical analysis and sequential Monte Carlo makes better use of the statistical characteristics of the observation information. It generates a set of random particles (or samples) with associated weights in state space to represent the required posterior density function (PDF). By resampling and recursion, PF gets the sub-optimal state estimation even in the processing of nonlinear and non-Gaussian systems. Although errors can be introduced when the data is greatly contaminated by noise (such as NLOS), and make the filter overestimated, there are many adaptive and robust methods [28] available that can correct them well, fortunately. Besides, PF seems to be innately compatible with the map information by the simulation through a large number of particles [25-27]. In the topological complex indoor environment, the algorithm can be told at quite a low computation and memory costs that this is an obstacle or that is forbidden to pass. The PF, which combines observations and floor-plan information, smartly reduces logic errors and improves the reliability of positioning.

To sum up the most of existing studies in audio localization [12-16, 18-20, 32, 33], the main limitation is that they all have only been carried out in a single and small region. The difficulty in encoding and decoding the identification bits becomes the fundamental reason for the inability to easily achieve wide-area audio positioning. According to the Nyquist-Shannon sampling theorem, the upper limit frequency of the received acoustic signal in most off-the-shelf smartphones is 22.05kHz which with respect to the highest 44.1kHz sampling rate. Hence, it could be quite a narrow bandwidth for these mobile devices to utilize to achieve senseless positioning (which means people can hardly hear the acoustic signal). Even if after which the encoding of identification bits to the bandwidth, it takes extra time to decode in smartphones, which increasing the latency and workload, and decreasing the fluency in positioning. In [16], an Orthogonal Chirp Spread Spectrum (OCSS) modulation technique was used to double the data rate in signal and performed well in a laboratory environment. However, the research does not take into account an actual reverberate

environment such as detecting the identification bits accurately in a crowded airport. Relying on the existing acoustic signal templates alone, a smartphone could not speculate where it is -- it can be anywhere since the conditions are not enough to meet the requirement of unique recognition of the signal in a wide range. Given the above reasons, there is no mature wide-area audio-based positioning solution applied to smartphones yet.

This paper proposes an integration algorithm of acoustic ranging, floor-plan, and MEMS sensors for locating commercial smartphones indoors. A hybrid strategy is established to identify the acoustic signal by multiple access technologies and external Bluetooth. An enhanced PF algorithm is proposed to improve the accuracy and stability of the real-time target position estimation task. In summary, the major contributions of this paper are listed as follows:

- 1) A hybrid architecture based on FDMA, TDMA, and SDMA is perfected to cover available acoustic signal to a wider area;
- 2) The prior information of Bluetooth and floor-plan are integrated to generate a coordinate-unique and accurate TDoA pairs for the region waiting to be located;
- 3) The naïve particle filter is enhanced with data bias integrity monitoring, dynamic variance adjustment, and floor-plan restraint to make the localization more in line with the behavior logic of indoor pedestrians;
- 4) A prototype system is developed to track smartphones indoors and verified that the ideas achieve decimeter-level localization accuracy and stable 1Hz update rate in several typical and wide indoor environments such as underground parking and airport check-in hall.

The remainder of this paper is organized as follows: the next Section presents the system design, including the anchor net design, signal design, and how the system works on smartphones. Section III describes the Bluetooth and floor-plan assisted wide-area TDoA acquirement in detail, and Section IV shows the particle filter algorithm. In the fifth Section, the test setup, data, and results are presented. This is followed by concluding remarks and plans for future research in Section VI.

II. SYSTEM STRUCTURE

The overall system structure is shown in Fig.1 that two main parts are included: the anchor network and the smartphone.

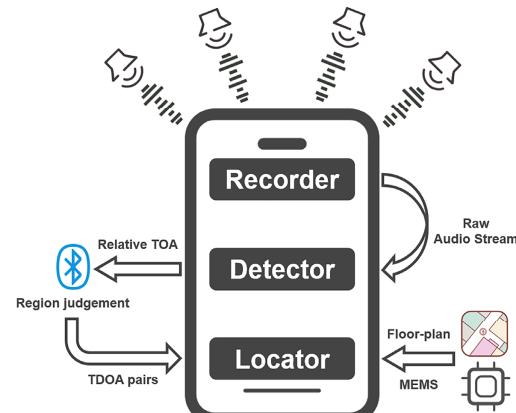


Fig. 1. Structure of wide-area audio-based indoor localization system.

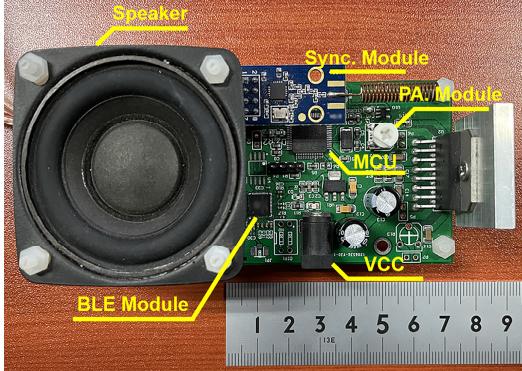


Fig. 2. Custom-designed audio anchor.

A. Anchor Design

In our system, the smallest positioning unit called sub-cell that consists of four anchors. Each anchor, as shown in Fig. 2, is equipped with a micro controller unit (MCU), a speaker, a wireless synchronous module, a Bluetooth module, a power amplification (PA) module, and a VCC interface. Once supplied with a 5V power input through the VCC, the MCU will deliver the pre-stored Chirp signals to the digital-to-analog converter (DAC). By adjusting the output power of the PA, the speaker can transmit the Chirp signal in distances from 0m to the maximum of 50m. To maintain the time consistency of the periodic acoustic signal through the synchronization module, we set one master synchronous node in the whole anchor network and at least one forwarding node in each sub-cell. Each MCU handles interrupts from the transmission of the 433MHz synchronous messages on the SPI pin and controls the utterance order of every anchor in the network. The Bluetooth module would then broadcast data packets with the designated power and period according to the utterance order.

B. Signal Design

In order to distinguish from the environmental noise, as well as be less sensitive to people's ears, we choose the lower limit of the Chirp signal as 15kHz. On the other hand, in fact, insufficient sampling often occurs when the received signal approaches the upper limit. Therefore, the maximum frequency of 21kHz is adopted in our system even the highest audio frequency that smartphones support is 22.05kHz. The designed ideal Chirp signal can be expressed in complex form as:

$$s(t) = A(t)e^{j2\pi(f_0t + \frac{1}{2}\frac{f_e - f_0}{T}t^2)}, t \in [0, T] \quad (1)$$

where $A(t)$ is the signal amplitude, T is the signal time duration, f_0 and f_e are the initial and cut-off frequency, respectively. Then the received signal considering with the indoor fading channel is formulated as:

$$x(t) = s(t) * h(t) = \sum_{i=0}^{L-1} \alpha_i s(t - \tau_i) + N_i(t) \quad (2)$$

where $h(t)$ is the continuous expression of the indoor audio channel impulse response (CIR); α_i , τ_i , and $N_i(t)$ are the channel fading coefficient, propagation delay, and additive noises of the i th propagation path, respectively. Generally, τ_0 is supposed to be the ToA of the first line-of-sight (LOS) component.

C. Application Implement

In our system, three components are coupled tightly in the smartphone application to execute accurate location estimation, as shown in Fig. 3.

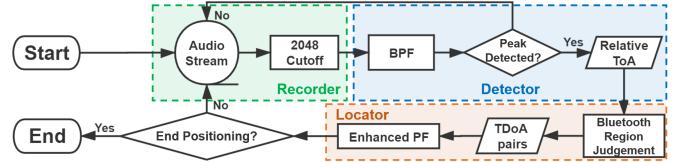


Fig. 3. Flowchart of the efficient component collaboration.

The recorder captures the raw audio stream from the hardware and maintains a cache pool of 4096 bytes. To carry out the GCC, the data in a unit of length 2048 bytes will be cut off in each round of calculation. In the detector, band pass filter (BPF) is used to filter the environment noise and improve the signal-noise ratio (SNR). The ToA peaks detection was performed with a two-step method, which could be found in detail in [32, 33]. Since smartphones cannot be told when the signals were sent from the anchors, we record the detected results as ToAs relative to the start timestamp of the recorder. With the assistance of the Bluetooth and floor-plan information in the locator, the relative ToAs are paired into TD_oA₂₁, TD_oA₃₂, and TD_oA₄₃ according to the anchor order. Finally, the location is estimated by the enhanced PF. Throughout the collaboration, peak detection takes the most resource consumption of 20ms and the PF with 200 particles takes 7ms. As a whole, the application features a low load to the CPU (15%~20% or so) extra compared with some business navigation softwares like 'Gaode Map'.

III. ACCURATE TDOA PAIRS GENERATION IN WIDE AREA

A. Acoustic Signal Hybrid Transmission Architecture

FDMA and TDMA are the two basic strategies to expand audio-based indoor positioning systems [12] and we implement a similar mode demonstrated in Fig. 4. The designed bandwidth of the Chirp signal is divided into two channels: 15-18kHz and 18-21kHz, respectively, and each of them has two opposite patterns. In this way, 4 separate signal templates are available for use. These templates of downward and upward within each channel are transmitted at 0.2s intervals in sequence inside the TDMA period of 1s. Each template is broadcast only for 45ms to protect every peak detection from being interfered. Notice the tailing on both sides of the signals in Fig. 4 (in the red dotted box), it is usually called the spectrum leakage, which is caused by the low-cost audio-related commercial off-the-shelf (COTS) infrastructure [14]. Such a phenomenon, as well as the Doppler shift [17, 18], brings a considerable offset in frequency and makes the TDE unreliable when making GCC using the biased data with the ideal signal. Because the tails can be considered as a systematic error, we reduce it through calibration and constant compensation.

Performing the FDMA and the TDMA technologies, eight anchors can make up 2 independent areas and form regional complexes, which we call it the prima-cell. To further expand the localizable range, it is almost inevitable to multiplex these acoustic signal templates in spatial, which called SDMA as

shown in Fig. 5. Each sub-cell is designed as a rectangle with a side range from 10m to 30m to guarantee that the same template in the same sub-cell is able to maintain a logical order all the time. In addition, the acoustic signal is limited inside the sub-cell as much as possible by adjusting the PA. As a result, 4 relative ToAs from a single sub-cell can be obtained with the detector. However, we are still puzzled about which prima-cell these ToAs belong to since the ambiguity may exist in every prima-cell.

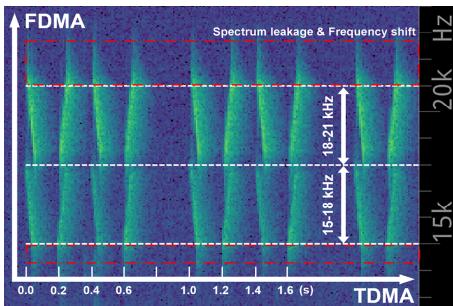


Fig. 4. FDMA and TDMA demonstrated in the real-time time-frequency diagram in the application of smartphone.

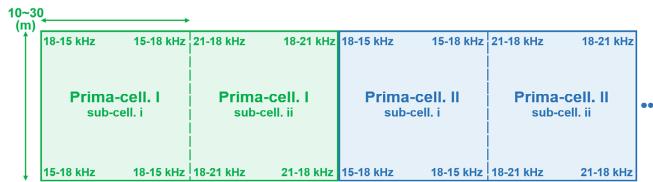


Fig. 5. SDMA design.

B. Bluetooth and Floor-plan Assisted Recognition

Bluetooth usually has a significant indication over short ranges and is suitable for existential detection. In general, 4 relative ToAs $\{ToA_i\}_{i=1}^4$ within one period can only be determined after the first signal being recognized in the next period. The sub-cell channel number S can be extracted from the identified $\{ToA_i\}_{i=1}^4$. We make 4 anchors in the same cell broadcast the same Bluetooth packets containing the floor-plan information, which in the format as "A_Pc_Sc", Pc is the prima-cell number and Sc is the sub-cell number. In addition, all Bluetooth inside the anchors are configured in non-connection and broadcast-only mode, and the power level is set to -12dBm to minimize the RSS in adjacent sub-cells employing the same templates.

Given multiple Bluetooth measurements $\{BT_i\}_{i=1}^n$ during the detecting period and in total M prima-cells in the floor-plan, it is available to uniquely identify current prima-cell code \mathbf{P} denoted by the probability $P(Pc_m | \{BT_i\}_{i=1}^n)$. However, this conditional probability is intractable to solve while the likelihood function $P(\{BT_i\}_{i=1}^n | P_{cm})$ can be derived from mass tests. Therefore, the posterior probability of P_{cm} given the joint probability $P(\{BT_i\}_{i=1}^n | P_{cm})$ can be calculated as:

$$P(P_{cm} | \{BT_i\}_{i=1}^n) = \frac{P(\{BT_i\}_{i=1}^n | P_{cm})P(P_{cm})}{\sum_{m=1}^M P(P_{cm}) \prod_{i=1}^n P(BT_i | P_{cm})} \quad (3)$$

where $P(P_{cm})$ is the prior probability of P_{cm} and the denominator is the total probability $P(\{BT_i\}_{i=1}^n)$. Then, the region decision where the signal most likely to belong can be made by:

$$\mathbf{P} = D(P_{cm}) = \arg \max_m (P(P_{cm} | \{BT_i\}_{i=1}^n)) \quad (4)$$

In our system, the maximum function is performed by counting and comparing all the entries in $\{BT_i\}_{i=1}^n$ that contains sub-cell code S . Finally, $\{ToA_i\}_{i=1}^4$ are combined into $\{TDoA_k\}_{k=1}^3$ with the unique coordinates of the anchors in the judged cell. The TDoA pairs generation is described in Algorithm 1.

Algorithm 1 TDoA pairs generation

Input: $S, \{BT_i\}_{i=1}^n, \{ToA_i\}_{i=1}^4$

Output: $\{TDoA_k\}_{k=1}^3$

```

1: Generate prima-cell candidate list  $\{Pc_j\}_{j=1}^M$  with elements 0
   according to M prima-cells in the f loor-plan
2: for  $i = 1 : n$  do
3:   if  $BT_i$  contains  $S$  then
4:     Find prima-cell code  $p$  in  $BT_i$ 
5:      $Pc_p \leftarrow Pc_p + 1$ 
6:   end if
7: end for
8: Find subscript  $p_{max}$  of the maximum in  $\{Pc_j\}_{j=1}^M$ 
9:  $\mathbf{P} \leftarrow p_{max}$ 
10: for  $k = 1 : 3$  do
11:    $TDoA_k \leftarrow ToA_{k+1} - ToA_k$ 
12:   Give coordinates of  $Anchor_k$  and  $Anchor_{k+1}$  in prima-
      cell  $\mathbf{P}$  to  $TDoA_k$ 
13: end for

```

IV. SMARTPHONE LOCALIZATION

Data in different dimensions and properties is complemented by tightly coupled integrated systems, and the fusion technique achieves finer localization results. The PDR algorithm based on MEMS sensors provides relative dynamic information with high precision and fine granularity in a short time while rapid divergence can occur over time. The acquired TDoA pairs in Section III.B maintain a relatively stable accuracy and give absolute guidance from a global level. Therefore, the MEMS and TDoA are combined in this paper to estimate locations smoothly.

A. Pedestrian Motion Judgement

Ignoring the differences of the device grip angle, different motions of the pedestrian can be converted into the attitudes of the smartphone. The built-in MEMS sensors accelerometer and gyroscope record the variations around tri-axis acceleration and angular velocity. Footstep-related parts are deduced from the accelerometer readings such as the step frequency (SF) detection [29] and step length (SL) estimation [30]. On the other hand, the heading update is performed according to the orientation quaternion as shown below:

$$\dot{\mathbf{q}} = \begin{bmatrix} \dot{q}_0 \\ \dot{q}_1 \\ \dot{q}_2 \\ \dot{q}_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix}$$

$$R(\mathbf{q}_{t-1}') = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 - q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (5)$$

$$\begin{bmatrix} \theta \\ \phi \\ \psi \end{bmatrix} = \begin{bmatrix} \arcsin(2(q_2 q_3 - q_0 q_1)) \\ \arctan(-\frac{2(q_1 q_3 - q_0 q_2)}{q_0^2 - q_1^2 - q_2^2 + q_3^2}) \\ \arctan(\frac{2(q_1 q_2 - q_0 q_3)}{q_0^2 - q_1^2 + q_2^2 - q_3^2}) \end{bmatrix}$$

where the four elements $[q_0 \ q_1 \ q_2 \ q_3]$ represent the rotation in the 3D space of the frame at timestamp t with respect to the frame at timestamp $t-1$, $[\omega_x \ \omega_y \ \omega_z]$ are the angle velocity readings, $[\theta \ \phi \ \psi]$ are pitch, roll, and heading, respectively. Furthermore, the complementary on gyroscope and accelerometer could improve the stability of the motion judgement [31].

B. Data Bias Integrity Monitoring

Autonomous integrity monitoring (AIM) is one of the essential components of conventional navigation systems. AIM detects signal changes or anomalies to determine whether the positioning is contaminated. The confidence level or alerting should be pushed to the users, indicating how much the data or results are reliable to the whole system. In this paper, we keep monitoring to the data bias between two consecutive epochs by the variation of each relative ToA and make appropriate adjustment decisions.

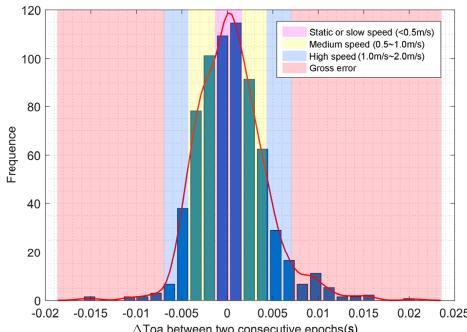


Fig. 6. Accuracy verification of the detected relative ToAs.

As shown in Fig. 6, 919 test samples of the ToA changes of two anchors between two consecutive epochs were collected in real-time, which the test smartphone kept static or moved back and forth between these two anchors at different speeds (0~2m/s). Neglect of the Doppler effect, these ToA variations are related to the moving speed and orientation of smartphones. For example, theoretically, walking 1s with speed of 1m/s brings single ToA offset of 3ms when away from the anchor and -3ms when approaching the anchor. Fig. 6 indicates that the detector gives most of the correct ToA variation intervals, and only 51 gross errors were estimated. The results also showed

Algorithm 2 Data bias integrity monitoring

Input: $\{ToA_i^t\}_{i=1}^4, \{ToA_i^{t-1}\}_{i=1}^4, \mu = 0, \sigma = 0.0038$

Output: $\{VIF_k^t\}_{k=1}^3$

```

1: for  $i = 1 : 4$  do
2:    $\Delta ToA_i^{t/t-1} \Leftarrow ToA_i^t - ToA_i^{t-1} - 1$ 
3:   Perform hypothesis testing:
4:    $H_0^{t,i}: \Delta ToA_i^{t/t-1}$  has no gross error
5:    $|u_i^t| = |\Delta ToA_i^{t/t-1} - \mu|/\sigma$ 
6:   if  $|u_i^t| > Z_{\alpha/2}$  then
7:     Reject  $H_0^{t,i}, vi f_i^t \Leftarrow 1.2 * |u_i^t|/Z_{\alpha/2}$ 
8:   else
9:     Accept  $H_0^{t,i}, vi f_i^t \Leftarrow 1$ 
9:   end if
10: end for
11: for  $k = 1 : 3$  do
12:    $VIF_k^t \Leftarrow vi f_{k+1}^t * vi f_k^t$ 
13: end for

```

Algorithm 3.1 Enhanced PF initialization

Input: P, S, N , floor-plan

Output: $\{\hat{x}_i^0\}_{i=1}^N$

```

1: Find the boundary  $[X_{max}, X_{min}, Y_{max}, Y_{min}]$  according to the
   floor-plan cell $_P S$ 
2: for  $i = 1 : N$  do
3:    $\hat{x}_i^0 \Leftarrow (\mathbb{U}[X_{min}, X_{max}], \mathbb{U}[Y_{min}, Y_{max}], \mathbb{U}[0, 2\pi], 1/N)$ 
4: end for

```

that the ToAs detected below 2m/s fit to $N(0, 0.0038^2)$ and a level of significance of $\alpha: 0.05$ approximately. Therefore, hypothesis testing can be performed under $H_0^{t,i}$: There is no gross error in the t th ToA of i th anchor. Given a new observation $\overline{ToA_i^{t+1}}$, we would reject $H_0^{t+1,i}$ and consider the observation as a gross error, if

$$|u_i^{t+1}| = \frac{|\overline{ToA_i^{t+1}} - ToA_i^t - 1|}{0.0038} > Z_{\alpha/2} \quad (6)$$

where $|u_i^{t+1}|$ is the statistic of test, and $Z_{\alpha/2}$ is the upper $\alpha/2$ point of the standard normal distribution. According to each test result, the variance inflation factor (VIF) adjusting the weight of each TDoA pair is decided as described in Algorithm 2. Furthermore, the gross errors are more likely to occur on the side facing back to the anchors because of the NLOS [14, 15], which needs pay more attention to.

C. Motion Judgement and Floor-plan Restraint Enhanced Particle Filter

Naïve particle filter is contained of sampling and resampling, and is described by state model and observation model. The smartphone tracking follows the state vector \mathbf{X} that consists of position (x, y) and heading θ (in degree). In the initialization

phase, N particles $\{\hat{x}_i^0\}_{i=1}^N$ are generated from the uniform distribution in the state space randomly according to the judged area code P and S and the floor-plan, and each particle weight is set to $1/N$, as shown in Algorithm 3.1.

1) State transition based on motion judgement.

The PF obeys the Markov Chain Monte Carlo. During each iteration, the state transition follows the Markov process of order one, which is $p(\mathbf{X}' | \mathbf{X}^{0:t-1}) = p(\mathbf{X}' | \mathbf{X}^{t-1})$. The transition probability distribution $p(\mathbf{X}' | \mathbf{X}^{t-1})$ is usually executed by the PDR, which is described as

$$\begin{bmatrix} \tilde{\theta}_i^t \\ \tilde{x}_i^t \\ \tilde{y}_i^t \end{bmatrix} = \begin{bmatrix} \hat{\theta}_i^{t-1} + \tilde{\Delta}_{\theta}^t + \delta_{\theta} \\ \hat{x}_i^{t-1} + (\tilde{\Delta}_s^t + \delta_s) \sin(\tilde{\theta}_i^t) \\ \hat{y}_i^{t-1} + (\tilde{\Delta}_s^t + \delta_s) \cos(\tilde{\theta}_i^t) \end{bmatrix} \quad (7)$$

where the $\tilde{\Delta}_{\theta}^t$ and $\tilde{\Delta}_s^t$ are the heading increment and position variation estimated given by the PDR. The δ_{θ} and δ_s indicate the state transition noise, and both are Gaussian white-noise with null means and variances σ_{θ}^2 and σ_s^2 , respectively. The noise describes the range of the interesting area (IA) covered by the particles in state space. To generate the IA more accurately, the variances should be adjusted dynamically according to different pedestrian motions.

In static mode, the TDoA pairs contribute most of the valuable information input, while the PDR only gives effective feedback of the velocity of 0. Thus, the IA needs to be intelligent enough to contain the location indicated by the TDoA pairs. In this paper, the basic σ_{θ} and σ_s are set to 10 and 0.7, respectively. However, the detector could make mistakes and then the TDoA pairs become unreliable. In order to obtain reasonable parameters, different noise levels should be chosen inversely proportional to the VIF vector decided in Algorithm 2.

In dynamic mode, the filter should be more confident in previous state estimation and current PDR output, therefore the range of IA is shrunken moderately. For example, $\sigma_{\theta} = 3$ and $\sigma_s = 0.4$ are suitable to restrain the IA and against some outliers in observations to a certain extent.

2) Floor-plan restraint assisted importance sampling.

A set of weighted particles is the key to approximate the true posterior density in PF. Each weight w_i^t is generated from a probability density function (PDF) called q that $q \propto p(\mathbf{Z}' | \mathbf{X}')$, where \mathbf{Z}' is the current observations that $\mathbf{Z}' = \{TDoA_k^t\}_{k=1}^3$.

Since \mathbf{Z}' obeys the normal distribution according to the data quality monitor, w_i^t can be calculated as

$$\tilde{w}_i = (2\pi\Sigma)^{-1/2} \exp\left\{-\frac{1}{2}[\mathbf{Z}' - h']^T \Sigma^{-1} [\mathbf{Z}' - h']\right\} \quad (8)$$

where h is the observation model related to \mathbf{X} and a sequence of zero mean white-noise with variance σ_R^2 , and Σ is the covariance matrix of the noise that $\Sigma = \text{diag}[VIF_1, VIF_2, VIF_3]$. Inevitably, these weighted particles cover inaccessible areas or cross walls due to the complex topology and narrow space indoors, which will result in frequent path replanning during navigation phase and bad user experience. However, from another perspective, the indoor layout provides guides for filtering out particles that do not meet the semantic floor-plan which is consisted of the logical matrix of values 0 and 1. The

Algorithm 3.2 Enhanced PF iteration

Input: $\{\hat{\mathbf{x}}_i^{t-1}\}_{i=1}^N$, \mathbf{Z}' , $\{VIF_k^t\}_{k=1}^3$, $\text{PDR}^t\{\tilde{\Delta}_{\theta}^t, \tilde{\Delta}_s^t\}$ and floor-plan
Output: $\hat{\mathbf{X}}^t$, $\{\tilde{w}_i^t\}_{i=1}^N$

- 1: Set $t := 1$
- 2: SAMPLING:
- 3: **for** $i = 1 : N$ **do**
- 4: Draw $\tilde{\mathbf{x}}_i^t \sim p(\mathbf{x}' | \hat{\mathbf{x}}_i^{t-1}, \mathbf{Z}')$ according to PDR t
- 5: **if** $\tilde{\mathbf{x}}_i^t$ in approachable region in floor-plan **then**
- 6: Assign \tilde{w}_i^t according to (8)
- 7: **else**
- 8: Assign $\tilde{w}_i^t \Leftarrow 0$
- 9: **end if**
- 10: **end for**
- 11: $\tilde{w}_i^t = \tilde{w}_i^t / \text{SUM}[\{\tilde{w}_i^t\}_{i=1}^N]$
- 12: **end for**
- 13: RESAMPLING:
- 14: **for** $i = 1 : N$ **do**
- 15: Draw $\hat{\mathbf{x}}_i^t, \tilde{w}_i^t$ from the $\mathbb{U}[0,1]$ -resampling
- 16: **end for**
- 17: ITERATION:
- 18: $\hat{\mathbf{X}}_k^t = \sum_{i=1}^N \hat{\mathbf{x}}_i^t * \tilde{w}_i^t$, set $t := t+1$, and go to step 2

matrix value will be indexed according to the updated state of each particle in the floor-plan: if the verification is passed, the weight is obtained by (8); otherwise, the weight is set to 0. As a result, most of the particles stay in the reachable area and given valuable weights while others are regarded as dead. All weights are then normalized to improve the degree of distinction among particles.

Finally, the random resampling is adopted referred to [24] to mitigate the sample impoverishment after multiple iterations. Current state estimation could be derived from the approximate posterior probability distribution as

$$\hat{\mathbf{X}}^t = \sum_{i=1}^N \tilde{\mathbf{x}}_i^t \tilde{w}_i^t \quad (9)$$

The proposed PF iteration is presented in Algorithm 3.2.

V. SYSTEM EVALUATION

A. System Ranging Performance

In this part, performances of multiple typical ranging technologies using in indoor localization systems were compared, that are UWB, audio, and Wi-Fi RTT. We adopted Decawave chips in UWB and Intel Wireless-AC8260 in RTT. As shown in Fig. 7(a), all stations or anchors were fixed at the end of a long corridor with a height of 1.5m, and reference distances of the receivers and smartphones were varied from 1m to 40m with step of 1m, from the corridor end. The ranging frequency and test time on each point were set to 1Hz and 30s, respectively. The ranging results of UWB with official two-way-ranging firmware were sent to the Vivo X50 Pro⁺ through Bluetooth. Estimated TDoAs can be directly regarded as distances since the test phone Huawei Honor8 was close enough to an audio anchor. The official ranging demo was used in Google Pixel 1 to figure out the distances of Wi-Fi RTT. All collected ranging results were compared to each ground truth

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

measured by the SNDWAY laser distance meter SW-80G to calculate the ranging errors.

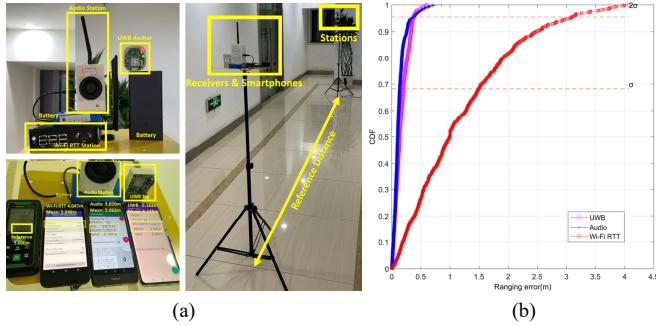


Fig. 7 Ranging comparison during multiple indoor localization systems.

The cumulative distribution function (CDF) of those errors was shown in Fig. 7(b). The proposed audio-based prototype system achieved the ranging precision of mean 0.155m and 0.392m in 2σ level, which indicates the system has the same ranging accuracy and stability as the UWB that had ranging errors of mean 0.121m and 0.384m in 2σ . Meanwhile, the precision of Wi-Fi RTT reached 3.133m in 2σ level and more big errors over 2m appeared when increasing the ground reference. Therefore, the proposed system was verified satisfying the premise of high-precision positioning.

B. Localization Comparison among Different Algorithms

We evaluated the Chan [22] and Taylor [23] algorithms along with the proposed PF by some numerical simulations that are assumed to carry out in a 4-stations-laid and 30m*30m area shown in Fig. 8(a). The observations of TD₂₁, TD₃₂, and TD₄₃ were calculated inversely according to the 317 pre-designed points shown in the area. The Cramér-Rao Lower Bound (CRLB), which describes the best-estimated localization accuracy in theory that only relies on the measurements, was compared to confirm the simulation performances. Fig. 8(b) and 8(c) demonstrate the positioning accuracy and robustness of the algorithms under $\sigma_R = 0.2$ and $\sigma_R = 2$. The Chan algorithm is an analytical solution in non-recursive, and it is based on the two-step least square (LS). Therefore, Chan had bad results more easily when variance growing, and it becomes unreliable even in low noise because of the terrible geometry of stations. Local LS and iteration are followed in the Taylor method. Such a method is able to approach the theoretical optimal accuracy under ideal conditions. However, more gross errors make it difficult to converge to the correct position rapidly. The proposed PF with 200 particles was less affected by the fluctuation of observation quality and even has better performance than CRLB with the assistance of additional information.

C. Localization in Single and Small Space Indoors

The indoor environment usually contains rooms and passageways. We first evaluated the system performance in a 13.8m*18.4m center-hollowed corridor. 4 stations were laid out in diamond shape at a height of 2.2m and the estimated positions by Xiaomi 10 Ultra were compared to 12 check points. Fig. 9 shows the localization results in real-time with and

without floor-plan assistance. Compared with the naïve PF that achieved mean accuracy of 0.72m, the proposed PF with the same 200 particles improved 33% with a mean error of 0.44m. The floor-plan restraint avoided particles entering the hollow region well and prevented the results from deviating too much from the trajectory.

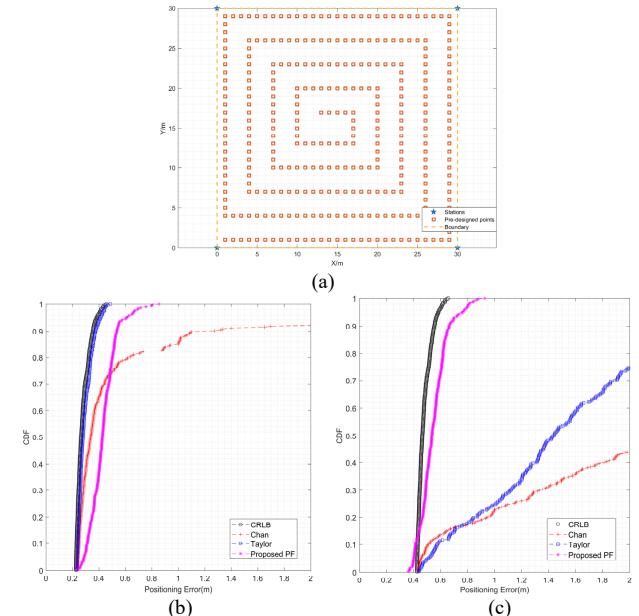


Fig. 8 Simulations of localization with variances in 0.2 (b) and 2(c).

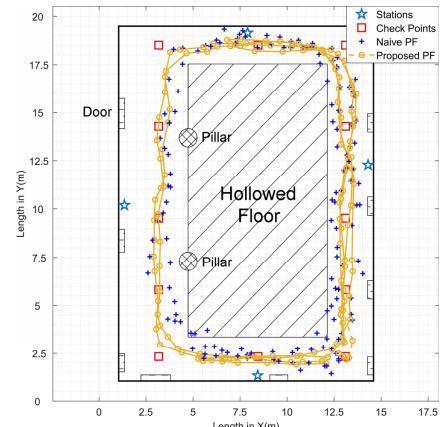


Fig. 9 Localization results of the naïve PF and the proposed PF.

D. Localization Comparison with Other System

In this section, we compared the proposed system with the outstanding system studied in [12], which is called RAIPS (robust acoustic indoor positioning system). According to the descriptions in [12], the RAIPS has similar modulation, broadcasting, and detection strategies on the acoustic signal with ours, and that's the main reason why we choose the target system. These two systems were deployed in a 7.5*12.8m office, and when one broadcast the signal, another kept silent, and vice versa. The Huawei Honor8 was used to carry out all tests with two different apps and record the data in real-time. The tester handheld Honor8 and collected 30s data at each of the 12 uniformly distributed static test points, as well as the dynamic data on the reference route around these test points. Two trajectories of these two systems were shown in Fig. 10,

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

and the detailed positioning error statistics compared to the ground truth were listed in Table II. From the results, we held the opinion that both systems have similar capabilities in indoor localization. Furthermore, RAIPS has higher precision because of their Doppler compensation, however, its non-filtering positioning method makes the system relatively unstable sometimes.

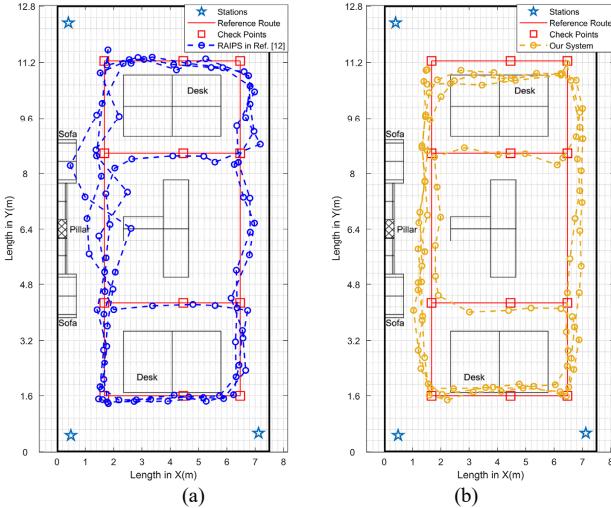


Fig. 10 Dynamic localization comparison in small area, (a) RAIPS in [12], (b) our proposed system.

TABLE I

STATISTICS OF SYSTEMS PERFORMANCE COMPARISON (HONOR8).

	Static / m				Dynamic / m			
	Max	Min	Mean	Std	Max	Min	Mean	Std
RAIPS	0.79	0.03	0.21	0.09	1.12	0.02	0.31	0.41
Our	0.51	0.02	0.17	0.07	1.01	0.02	0.51	0.30

E. Localization in Wide Area Indoors

To evaluate the localization ability of our system in the wide area, we implemented static and dynamic experiments in the underground parking and airport check-in hall with different smartphones of Huawei Honor 8, Huawei P9, Xiaomi 10 Ultra, and Google Pixel 3, as illustrated in Fig. 10. The shading colors represent different acoustic signal templates that blue means prima-cell. I, sub-cell. I; yellow means prima-cell.I, sub-cell. ii; and red means prima-cell. II, sub-cell. i, respectively. The reference coordinates of all check points and audio stations were measured by the Leica total station TS60. In the underground parking, each sub-cell is 30m in both length and width. Taking the Huawei Honor 8 as a study case, Fig. 10(a) and 10(b) shows the static results of 30 selected points of the total 83 check points, and Fig. 10(c) and 10(d) show two different estimated trajectories in dynamic. In the 58m*27m*6m airport check-in hall, we held the luggage trolley and Xiaomi 10 Ultra and walked continuously at normal speed for 6 minutes, as shown in Fig. 10(e).

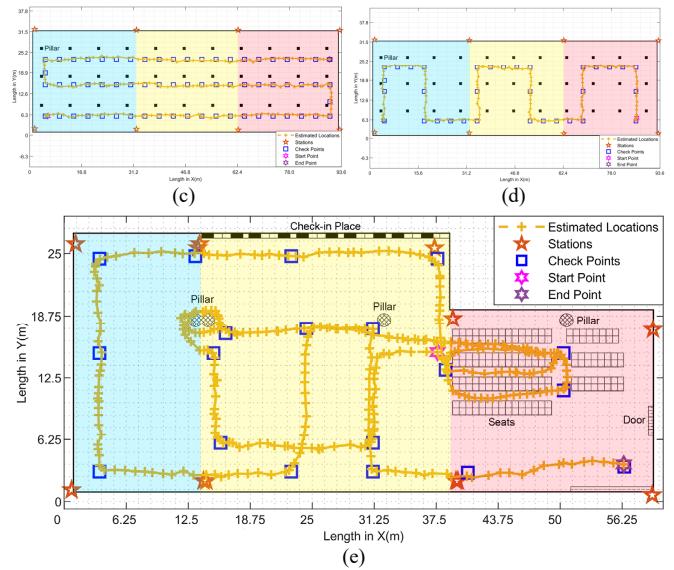
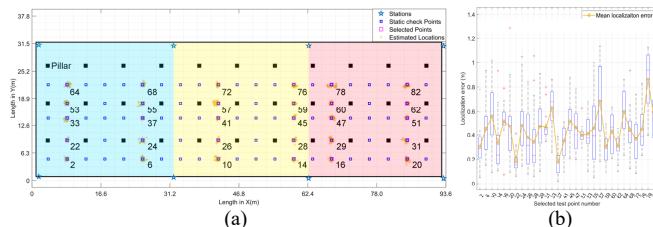


Fig. 11 Localization in underground parking and airport check-in hall.

TABLE II
STATISTICS OF THE SYSTEM PERFORMANCE ON DIFFERENT PHONES.

	Static / m				Dynamic / m			
	Max	Min	Mean	Rmse	Max	Min	Mean	Rmse
Honor8	1.29	0.02	0.44	0.50	1.52	0.05	0.59	0.67
P9	1.53	0.04	0.50	0.57	1.49	0.06	0.64	0.70
Pixel3	1.48	0.09	0.66	0.69	1.82	0.11	0.79	0.85
Mi10	--	--	--	--	1.55	0.08	0.64	0.75

The statistics of the root mean square error (RMSE), maximum and minimum error, and mean error of all the system outputs compared to the reference values are listed in Table II, which indicates that the proposed system achieved localization accuracy of 0.66m RMSE under both static and dynamic conditions and had little difference among the test smartphones. In addition, compared the performances of the PF with 200 particles and 500 particles, the latter reduced the RMSE to 0.56m but increased 114% computing overhead of 8ms.

VI. DISCUSSION

Any indoor positioning technology can only be really widely used if it has the ability to be low-cost, highly accurate, deployable over a wide area, and easily accessible to the public, etc. Here we make a comparison of the existing audio-based indoor localization systems in Table III. It is clear that regardless of changes in the signal modulation patterns and positioning algorithms, audio as a localization source always has sub-meter-level positioning capability. However, the size of test scenes involved in these outstanding works is also worth noting. Admittedly, we have no idea whether these systems will expand their effective coverage to a larger range. We hold the opinion that these systems still can be improved according to the existing data they provided.

Therefore, in the present study, we provide a novel idea to organically combine the Bluetooth, MEMS, and acoustic signal to achieve the high accuracy wide-area indoor localization. In Section V, we have verified the feasibility of the idea at the levels of localization source, localization algorithm, and practical tests respectively. From Table III, it shows that the prototype system achieved sub-meter-level positioning

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

accuracy no matter in small-area or wide-area. Meanwhile, it supports multiple regions over 30*60m without significant compromise in localization performance. While we think this is not a qualified result for artificial intelligence applications such as robotics, it is more than adequate for pedestrian localization and navigation.

TABLE III

COMPARISON OF SOME EXISTING ACOUSTIC-BASED LOCALIZATION WORKS.

System	Signal	Synchroni-zation	Scenes size	Positioning algorithm	Typical accuracy
ASSIST [19], 2012	Chirp	WLAN	12×13m ²	Iterative LS	18~34cm
GuoGuo [20], 2015	2-PAM	Zigbee	3.0×8m ²	KF	6~15cm
AALTS [16], 2019	OCSS	Async.	8×12m ²	PF	20~49cm
RAIPS [12], 2019	Chirp	WLAN	6×10m ²	COM-W	20~50cm
Our small-area	Chirp	433MHz	Max 30×30m ²	Enhanced PF	15~60cm
Our wide-area			>30×60m ²		20~80cm

Although there are some interesting discoveries revealed by these tests, the prototype system still suffers from limitations, which are vertical constraint and the lack of finer NLOS detection and Doppler effect compensation. As we described in Section III, the prima-cell could theoretically be extended infinitely in the lateral direction like a ribbon. When there is an audio signal in the vertical direction of this ribbon, the ambiguity will be even worse to determine the source of the estimated TDoA pairs. As for NLOS and Doppler, they both interfere with the determination of the normal wave peak in the GCC results, thus generating errors in the estimation of ToA. Further studies, informed by the work reported here, are needed to complete the problems of how to fix the region ambiguity, and how to detect the signal more rapidly and accurately.

VII. CONCLUSION

This study provides new insights into audio-based localization system that applies to the smartphones and supports indoor LBS in wide areas. To address the challenges of the limited positioning capability in a large area, which is caused by the difficulty in audio multi-ID encoding and decoding, we perfected the TDMA, FDMA, and SDMA schemes and acquired the unique and precise TDoA pairs by Bluetooth and floor-plan judgement. The enhanced PF constrains the estimated locations and maintains the localization logic. The prototype system was proven to achieve sub-meter-level positioning accuracy based on multiple actual tests in some typical scenes.

From this study, we could foresee that as a relatively novel positioning source, the audio-based localization technology will make important contributions to the consumer-grade indoor positioning and internet of things field in the future. Meanwhile, it should also be noted that the NLOS and the Doppler effect will reduce the stability of positioning, which is worth discussing. In addition, the idea of combining the acoustic signal with the environment as the fingerprint information may trade some accuracy for calculation convenience and be possible to avoid the complex ToA detection.

REFERENCES

- [1] *IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems Local and Metropolitan Area Networks-Specific Requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE Standard 802.11-2016 and 802.11-2012, Dec. 2016.
- [2] M. Ibrahim, H. Liu, M. Jawahar, V. Nguyen, M. Gruteser, R. Howard, B. Yu, and F. Bai, "Verification Accuracy evaluation of wifi fine time measurements on an open platform," in *Proceedings of the 24th Annual International Conference on Mobile Computing and Networking*, 2018, pp. 417-427.
- [3] L. Banin, O. Bar-Shalom, N. Dvorecki, and Y. Amizur, "Scalable Wi-Fi Client Self-Positioning Using Cooperative FTM-Sensors," in *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3686-3698, 2019.
- [4] J. S. Hammerschmidt, and E. Sasoglu, "Time instant reference for ultra wideband systems," U.S. Patent 10 567 034, Feb. 2019.
- [5] L. Zheng, *Faming Zhanli Shenqing Gongkai Shuomingshu*, CN Patent 201911129633.1, 2020.
- [6] S. M. Han, Y. H. Kim, S. S. Lee, H. K. Yoon, Y. J. Yoon, Y. H. Lim, and H. R. Kim, "Ultra wideband antenna for filtering predetermined frequency band signal and system for receiving ultra wideband signal using the same," U.S. Patent 7 557 755, July, 2009.
- [7] N. Yu, X. Zhan, S. Zhao, Y. Wu, and R. Feng, "A Precise Dead Reckoning Algorithm Based on Bluetooth and Multiple Sensors," in *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 336-351, 2018.
- [8] M. Cominelli, P. Patras, and F. Gringoli, "Dead on Arrival An Empirical Study of The Bluetooth 5.1 Positioning System," in *Proceedings of the 13th International Workshop on Wireless Network Testbeds, Experimental Evaluation & Characterization (WiNTECH '19)*, New York, USA, 2019, pp. 13-20.
- [9] C. Forster, Matia Pizzoli, and Davide Scaramuzza, "SVO: Fast semi-direct monocular visual odometry," in *IEEE International Conference on Robotics & Automation (ICRA)*, Hong Kong, China, 2014, pp. 15-22.
- [10] S. Yang and S. Scherer, "CubeSLAM: Monocular 3-D Object SLAM," in *IEEE Transactions on Robotics*, vol. 35, no. 4, pp. 925-938, 2019.
- [11] W. Robert and E. Lehmannand, "Importance sampling particle filter for robust acoustic source localisation and tracking in reverberant environments," 2005.
- [12] S. Cao, X. Chen, X. Zhang, and X. Chen, "Effective Audio Signal Arrival Time Detection Algorithm for Realization of Robust Acoustic Indoor Positioning," in *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 7341-7352, 2020.
- [13] X. Chen, Y. Chen, S. Cao, L. Zhang, X. Zhang, and X. Chen, "Acoustic Indoor Localization System Integrating TDMA+FDMA Transmission Scheme and Positioning Correction Technique" in *Sensors*, 2019, no. 10: 2353, doi: 10.3390/s19102353.
- [14] L. Zhang, M. Chen, X. Wang, and Z. Wang, "TOA Estimation of Chirp Signal in Dense Multipath Environment for Low-Cost Acoustic Ranging," in *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 2, pp. 355-367, 2019.
- [15] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, E. S. Chng, and H. Li, "A learning-based approach to direction of arrival estimation in noisy and reverberant environments," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, South Brisbane, QLD, Australia, 2015, pp. 2814-2818, doi: 10.1109/ICASSP.2015.7178484.
- [16] C. Cai, R. Zheng, J. Li, L. Zhu, H. Pu, and M. Hu, "Asynchronous Acoustic Localization and Tracking for Mobile Targets," in *IEEE Internet of Things Journal*, vol. 7, no. 2, pp. 830-845, 2020.
- [17] I. Shames, A. N. Bishop, M. Smith, and B. D. O. Anderson, "Doppler Shift Target Localization," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 49, no. 1, pp. 266-276, Jan. 2013, doi: 10.1109/TAES.2013.6404102.
- [18] T. Liu, X. Niu, J. Kuang, S. Cao, L. Zhang, and X. Chen, "Doppler Shift Mitigation in Acoustic Positioning Based on Pedestrian Dead Reckoning for Smartphone," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-11, 2021.
- [19] F. Höflinger, R. Zhang, J. Hoppe, A. Bannoura, L. M. Reindl, J. Wendeberg, M. Bührer, and C. Schindelhauer, "Acoustic Self-calibrating System for Indoor Smartphone Tracking (ASSIST)," in *2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1-9, 2012, doi: 10.1109/IPIN.2012.6418877.

IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS

- [20] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1144-1156, 2016.
- [21] B. T. Fang, "Simple solutions for hyperbolic and related position fixes," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 26, no. 5, pp. 748-753, Sept. 1990, doi: 10.1109/7.102710.
- [22] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," in *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905-1915, Aug. 1994, doi: 10.1109/78.301830.
- [23] W. H. FOY, "Position-Location Solutions by Taylor-Series Estimation," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-12, no. 2, pp. 187-194, March 1976, doi: 10.1109/TAES.1976.308294.
- [24] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation." *IEE proceedings F (radar and signal processing)*, vol. 140, no. 2, pp. 107-113, 1993.
- [25] H. Nurminen, M. Raitoharju and R. Piché, "An efficient indoor positioning particle filter using a floor-plan based proposal distribution," *2016 19th International Conference on Information Fusion (FUSION)*, Heidelberg, 2016, pp. 541-548.
- [26] H. Nurminen, A. Ristimäki, S. Ali-Löytty, and R. Piché, "Particle filter and smoother for indoor localization," in *International Conference on Indoor Positioning and Indoor Navigation*, 2013, pp. 1-10, doi: 10.1109/IPIN.2013.6817903.
- [27] P. Davidson, J. Collin, and J. Takala, "Application of particle filters for indoor positioning using floor plans," in *2010 Ubiquitous Positioning Indoor Navigation and Location Based Service*, 2010, pp. 1-4, doi: 10.1109/UPINLBS.2010.5653830.
- [28] R. Mah, and A. Tamhane, "Detection of gross errors in process data," *AIChe Journal*, vol. 28, no. 5, pp. 828-830, 1982.
- [29] W. Kang, and Y. Han, "SmartPDR: Smartphone-Based Pedestrian Dead Reckoning for Indoor Localization," *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2906-2916, 2015.
- [30] R. Chen, L. Pei, and Y. Chen, "A smart phone based PDR solution for indoor navigation," in *Proceedings of the 24th international technical meeting of the satellite division of the institute of navigation (ION GNSS 2011)*, pp. 1404-1408.
- [31] R. G. Valenti, I. Dryanovski, and J. Xiao, "Keeping a Good Attitude: A Quaternion-Based Orientation Filter for IMUs and MARGs," in *Sensors*, vol. 15, no. 8, pp. 19302-30, Aug 6, 2015.
- [32] R. Chen, G. Guo, F. Ye, L. Qian, S. Xu, and Z. Li, "Tightly-coupled integration of acoustic signal and MEMS sensors on smartphones for indoor positioning," in *Acta Geodaetica et Cartographica Sinica*, vol. 50, no. 2, pp. 143-152, 2021, doi: 10.11947/j.AGCS.2021.20200551.
- [33] R. Chen, Z. Li, F. Ye, G. Guo, S. Xu, L. Qian, Z. Liu, L. Huang, "Precise Indoor Positioning Based on Acoustic Ranging in Smartphone," in *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2021, Art no. 9509512, doi: 10.1109/TIM.2021.3082269.



SHIHAO XU received the B.S. degree in Surveying and Mapping Engineering from China University of Mining and Technology, Jiangsu, China, in 2019.

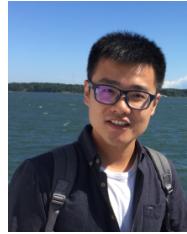
He is currently pursuing the Ph.D. degree in Geodesy and Survey Engineering from Wuhan University, Wuhan, China. His research interests include the development of location-based services, indoor positioning and navigation technology, and information fusion.



RUIZHI CHEN received the Ph.D. degree in Geophysics from University of Helsinki, Finland, in 1991.

He is currently a Professor and Director of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing at Wuhan University. He used to work as an Endowed Chair Professor at Texas A&M University Corpus Christi, US; Head and Professor of the Department of Navigation and Positioning at Finnish Geodetic Institute, Finland, and Engineering Manager in Nokia, Finland. He has published two books and more than 260 scientific papers. His current research interests include indoor positioning, satellite navigation and location-based services.

Positioning at Finnish Geodetic Institute, Finland, and Engineering Manager in Nokia, Finland. He has published two books and more than 260 scientific papers. His current research interests include indoor positioning, satellite navigation and location-based services.



Guangyi Guo received the Ph.D. degree in the Cartography and Geographical Information Science, Wuhan University, Wuhan, Hubei, China, in 2020.

He is currently a Postdoctor of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing at Wuhan University. His research interests include indoor positioning, machine learning and Internet of Things.



Zheng Li received the M.S. degree in Geodesy and Survey Engineering from Wuhan University, Hubei, China in 2019 and the B.S. degree in Surveying and Mapping Engineering from Chang'an University, Shanxi, China in 2016.

Since 2019, he has been pursuing the Ph.D. degree in the Geodesy and Survey Engineering, Wuhan University, Wuhan, Hubei, China. His research interests include indoor positioning and the Internet of Things.



Long Qian received the M.S. degree in Cartography and Geographical Information Engineering from Wuhan University, Hubei, China in 2019.

Since 2019, he has been pursuing the Ph.D. degree in the Geodesy and Survey Engineering, Wuhan University, Wuhan, Hubei, China. His research interests include indoor positioning and the Internet of Things.



Feng Ye received the M.S. degree in Optical Engineering from Huazhong University of Science and Technology, Hubei, China in 2017 and the Ph.D. degree in the Geodesy and Survey Engineering, Wuhan University, Wuhan, Hubei, China.

He is currently a Postdoctor of the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing at Wuhan University. His research interests include indoor positioning, ubiquitous computing, machine learning and Internet of Things.



Zuoya Liu received the M.S. degree in Electronic Circuit and System from Central China Normal University, Hubei, China in 2015 and the B.S. degree in Electronic Information Science and Technology from Central China Normal University, Hubei, China in 2012.

Since 2018, he has been pursuing the Ph.D. degree in the Cartography and Geographical Information Science, Wuhan University, Wuhan, Hubei, China. His research interests include indoor positioning, navigation, machine learning and Internet of Things.



Lixiong Huang received the B.S. degree in Geodesy and Geomatics Engineering from Wuhan University, Hubei, China in 2019.

He is currently pursuing the Ph.D. degree in Geodesy and Survey Engineering from Wuhan University (WHU), Wuhan, China. His research interests include indoor positioning, Internet of Things, and information fusion.