AN INDOOR POSITIONING SYSTEM BASED ON COMBINED AUDIO CHIRP/MEMS/FLOOR MAP: PERFORMANCE ANALYSIS OF KEPLER A100

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

The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, Wuhan, 430079.

Commission IV, WG IV/5

KEY WORDS: Indoor Positioning, Audio Chirp, MEMS, Information Fusion, Pedestrian Tracking

ABSTRACT:

In recent years, audio positioning technology has been more favoured because it is not only highly accurate but also universally applicable to the mass users. In line with this trend, a chip dedicated to indoor positioning, the Kepler A100, has been released lately, which ranges by estimating the arrival time of the audio chirp. As with radio ranging techniques, sound-based ranging methods is subject to significant observation errors due to Non-Line-Of-Sight (NLOS) or multi-path effects, while Pedestrian Dead Reckoning (PDR) based on inertial data have the disadvantage of error accumulation. Both make localization challenging. In this paper, the ranging performance of the Kepler A100 chip is evaluated and experiments are conducted in two typical indoor scenarios to verify the stability and sub-metre accuracy of the ranging results. Meanwhile, this paper implements an indoor positioning system combining audio chirp/Micro-Electro-Mechanical Systems (MEMS)/floor map based on particle filtering algorithm. A new mapping constraint method is applied in the system, which can effectively constrain the particles in a reasonable space. Experiments are conducted in two complex office environments and the results show that the error of the proposed algorithm is about 0.65m in 95% of cases, a reduction of more than 40% compared to only-PDR or only-audio. The effectiveness and stability of the method is demonstrated.

1. INTRODUCTION

The architecture of the Internet of Things (IoT) can be divided into three layers: the sensing layer, the network layer, and the application layer (Puthal, 2016). For the perception layer, location awareness is indispensable and it also plays an important role in location-based services (LBS). With the advent of Global Navigation Satellite Systems (GNSS), location is readily available outdoors and performs excellently. However, due to the obstruction of buildings, it suffers from severe signal attenuation and degraded positioning performance. As a result, there is a huge market demand for providing reliable indoor positioning services, which has led to great interest in related theoretical research in both academic and industrial circles.

Currently, available indoor location technologies e.g., Wi-Fi (Yan, 2021), Bluetooth (Zuo, 2018), Ultra-Wide Band (UWB) (Dabove, 2018), Pedestrian Dead Reckoning (PDR) (Chen, 2021), Radio Frequency Identification (RFID) (Wu, 2019), Visible Light (Wang, 2018) etc. have been successfully applied in indoor scenarios. Because these technologies have their pros and cons, there is no universal positioning scheme that can be used in mostly indoor environments. Although Wi-Fi access points (APs) are almost ubiquitous in indoor scenes, the methods of inferring distance based on signal strength fading are always less accurate; besides, the most common positioning method of Wi-Fi and Bluetooth is fingerprinting, which is time-consuming and laborious for the construction and maintenance of fingerprints; and, the UWB system is currently only supported by iPhone 11, MIX 4 and Galaxy Note20 Ulta, and the system deployment cost is relatively high. A RFID system is composed of a scanner with an antenna, which queries the nearby RFID tags and transmits the data from the tag to the scanner by radio waves. It is generally used for short-range identification and communication because of its small scope. PDR is the common positioning approach used for sensor-embedded in smartphone. It's low-cost, low-power-consuming, and interference-free, can achieve high accuracy in a short term; whereas, the precision degenerates severely for long-term navigation or positioning for the accumulation of error with time. Compared with Wi-Fi, Bluetooth or UWB, acoustic (Hoflinger, 2014) positioning is more universal due to the utilization of the built-in microphones in smartphones, and all smartphones have microphones. The sound signal not only covers a wide range, but also has high ranging accuracy. Besides, its lower propagation speed than the electromagnetic wave also makes time synchronization easier. And it's worth mentioning that the application of acoustic ranging to indoor positioning has been favoured by more and more scholars.

To continuously improve the accuracy and pervasiveness of Indoor Positioning Systems (IPS), techniques on audio positioning and ranging have been extensively researched. Farrokhi et al. (2005) present an audible spread spectrum acoustic ranging system, where the signal is a linear frequency modulated (LFM) pulse (chirp). And experimental results demonstrate that an LFM pulse with 3 kHz bandwidth, centre frequency of 2 kHz, and chirp rate of 60 Hz/ms is better. Traditional signal processing methods usually rely on correct assumptions about the signal model and accurate estimation of time difference of arrival (TDoA), suffering from noise and reverberation distortions. Xiao et al. (2015) propose a learning-based approach to estimate direction of arrival (DoA) from microphone input. And, more the training data available, more accurate the estimation is. To overcome the difficult of extracting the first path signal in complex indoor environment, Zhang et al. (2018) propose a novel time of arrival (ToA) estimation method based on an iterative cleaning process, which eliminates the strongest multipath component at each iterative procedure with a band-elimination filter in fractional Fourier Domain (FrFD). Cao et al. (2020) present a novel algorithm detecting audio signal arrival time with coarse and fine searches. For the coarse search, a cosine theorem is used to extract audio data segments from the received signal, and for the fine search, the audio ToA is estimated with the

https://doi.org/10.5194/isprs-archives-XLVI-3-W1-2022-53-2022

autocorrelation function of the waveform characteristic source signal. TDoA is one of the common methods for indoor positioning. However, when the device is in motion, the Doppler shift produces an additional error in TDoA, inducing degradation in the localization performance. Liu et al. (2020) propose an indoor localization system which overcomes the effects of the Doppler shift by the fusion of PDR and acoustic.

Many acoustics-based solutions have emerged, indicating acoustics is one of the promising technologies for indoor precise positioning. However, in a complex indoor environment, the ranging results are still likely to be affected by gross errors under the non-line-of-sight propagation of audio signals, inevitably (Yu, 2018). Based on the data from Micro-electromechanical Systems (MEMS), PDR tends to provide reliable displacement and heading information within a short time, but the error accumulation along with time needs to be prevented.

To address these challenges, a combined audio chirp/ MEMS/ floor map system for indoor localization is proposed, in which the recently released Kepler A100 is used. The ranging performance of this chip is evaluated comprehensively. This paper also proposes a fusion method based on particle filtering algorithm to eliminate the severe ranging errors caused by NLOS and the cumulative errors caused by PDR. Also, to avoid unreasonable localisation results, the weight of the particles is adjusted by using a floor plan during the filtering process. The proposed positioning system is shown in Figure 1.

Figure 1. The proposed positioning system

The rest of this paper is organized as follows. Section II briefly introduces Kepler A100 and gives an overview of the selected hardware equipment in our indoor positioning system, including the designs of tag and anchors. Section III introduces the pedestrian dead reckoning algorithm, the particle filter-based fusion algorithm and the efficient map constraint method proposed in this paper. Section IV conducts extensive experiments to evaluate the ranging performance of Kepler A100, and validates the proposed system. Finally, the conclusions are summarized in Section V.

2. SYSTEM OVERVIEW

2.1 Kepler A100

A few months ago, a chip specially used for indoor positioning was released, which is called Kepler A100. It is referred to as A100 in the following description. A screen printing of the chip is shown in Figure 2.

Figure 2. Kepler A100

A100 can provide both Time Difference Of Arrival (TDOA) and Time of Flight (TOF) measurements by accurately estimating the time of arrival of audio chirps. That means distance information is estimated with knowing the speed of sound. The manufacturer estimates a real-time accuracy of ±50 cm in an ideal environment using a prescribed linear frequency modulation signal.

A100 provides an interface for connecting a digital microphone, using the I2S standard protocol with a sampling rate of 48K. The external communication protocol, as with most sensor chips, is the standard I2C protocol. The external host microprocessor can configure the chip, e.g., the sound track, range mode and synchronization function. In the short-range mode, the distance between speakers should not exceed 17m, and the distance update rate is 2 Hz; while in the long-range mode, up to 50m can be supported, but the observation interval is extended to 1.1s. In addition, after the synchronization function of A100 is enabled, the host microprocessor needs to pull up the sync-pin of A100 periodically to notify the sounding time of the horn, and then distance observation would be output. If it is disenabled, only distance differences can be obtained. Besides, three pins in A100 are used to detect the running status of the chip. The toggles of these pins indicate that state of the chip is working properly, the audio data flow is normal, and sync-signal is successfully captured respectively, which are convenient for users to troubleshoot problems.

An AM100 development kit was also provided with the release of the chip by the manufacturer. The kit contains an AM100 development board and four matching speakers. In addition, the official also provides many basic projects to help users to quickly drive the A100 to obtain ranging information to meet their business needs.

2.2 Design of Anchor

The speaker in the development board is hemispherical in overall shape, as shown in Figure 3. The front is designed as a mesh to ensure that the sound can come out. There is a knob on the side that can adjust the loudness of the speaker. When the knob is turned to the maximum, the ear can clearly hear the sound when approaching to the speaker, and there is almost no sound at a distance of about 2m. There is a rod antenna behind the speaker, which is used to send and receive RF signals to ensure the time synchronization between the speakers. In this way, the sound can be broadcast in an orderly manner between the speakers. It can also be used to synchronize A100. The power supply interface is on the side of the speaker and supports 12V DC.

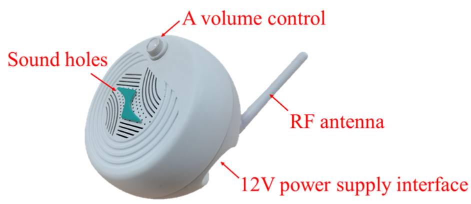


Figure 3. Hemispherical speaker

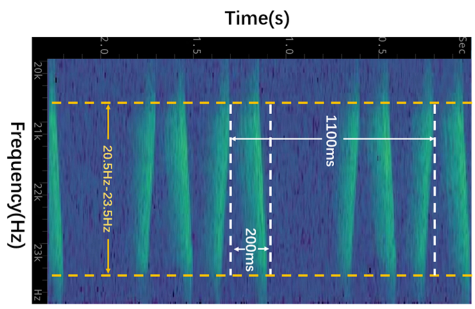


Figure 4. Time-frequency diagram processed in the smartphone

After the four speakers are powered on, a faint sound will soon be emitted. Figure 4 illustrates the time-frequency diagram processed in the smartphone. The four speakers send audio chirp signals in sequence. The first and third speaker send 23.5KHz-20.5KHz chirps, and the 20.5KHz-23.5KHz chirps are sent by the 2nd and 4th speaker. The interval between speakers is 200ms in turn, and the transmission period is 1.1s.

2.3 Design of Tag

As shown in Figure 5, the whole device consists of AM100 development board and WT931 module connected by wires and fixed by soluble glue, where AM100 module occupies a relatively large area. AM100 module mainly includes the ranging chip, the processing unit, and the RF transceiver under the shield. Besides, the digital microphone and the ceramic antenna for Bluetooth is exposed on the outside of the module. On the bottom plate, there is an antenna for RF signals. For the inertial measurement unit (IMU), we choose the WT931 module from Wit-Motion, which integrates a high-precision gyroscope, an accelerometer, and a magnetometer. It adopts advanced digital filtering technology, which can effectively reduce measurement noise and improve measurement accuracy.

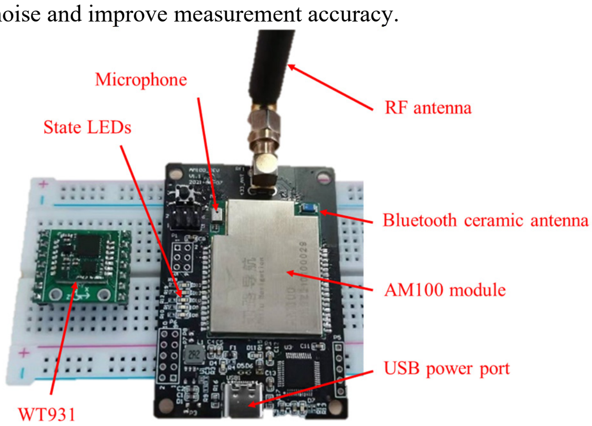


Figure 5. The front appearance of the tag

The overall structure of the tag is shown as Figure 6. Power is supplied to the development board through the type-c interface on the development board, and the host microprocessor. starts to work. First, the RF transceiver is configured through SPI, so that it accepts and receives radio frequency signals, and waits for the sync-packet from anchors. Next, WT931 module is initialized through UART, and the baud rate is 961200. The update rate of inertial data is set as 100Hz.Then through the I2C protocol, the initialization of A100 is completed setting the long-range mode and enabling the synchronization function. Finally, the host microprocessor calculates the positions of the tag according to all the data, and sends the results to the smartphone via Bluetooth.

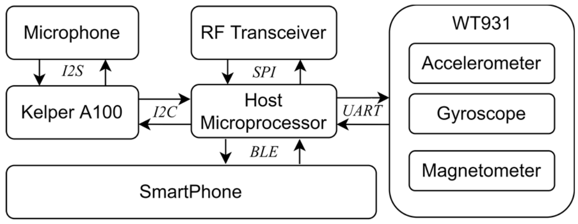


Figure 6. Overall structure of the tag

3. METHOD

The architecture of our proposed fusion algorithm combining audio chirp/MEMS/floor map based on particle filtering is shown in Figure 7. The inertial data is involved in the particle state update, while the range results from A100 are used to update the particle weights, which are also constrained by the map information. The relevant methods are described in detail in the following parts.

3.1 Pedestrian Dead Reckoning

PDR was first proposed by Levi and Judd in 1996, and it iteratively estimates step length and heading angle according to the IMU to compute a location. The PDR method mainly includes three aspects: step detection, step length estimation, and heading estimation. To detect gaits, common methods are peak detection (Jin, 2011), zero detection (Ayub, 2012), and autocorrelation detection (Rai, 2012). Considering both accuracy and complexity, peak detection method is adopted in this paper. And first of all, to obtain the acceleration of pedestrian movement without high-frequency noise, the gravity component is removed, and a first-order Butterworth low-pass filter with a 3 Hz cut-off frequency was used.

$$a = \|\boldsymbol{a}\| - \|\boldsymbol{g}\| \tag{1}$$

where $\boldsymbol{a}$ is the output of accelerometer, and $\boldsymbol{g}$ is the gravity vector. $\|\bullet\|$ is the modulo operation. The peak detection process in performed with (2) and (3).

$$\begin{aligned} peak(i) = \{ n \mid a(n) > a(n+1), a(n) > a(n-1), \\ a(n) > \gamma_{peak}, n = 2,3,4,\ldots,N \} \end{aligned} \tag{2}$$

$$\begin{aligned} peak'(k) = \{ j \mid t_{peak}(j) - t_{peak}(j-1) > \gamma_t, \\ j = 2,3,4,\ldots,i \} \end{aligned} \tag{3}$$

where N is the size of acceleration series, γ_{peak} and γ_t are the threshold of peak value and time difference between adjacent peaks, respectively. The smoothing results and peak detection results are shown in Figure 8, peaks of each step are marked as red circles.

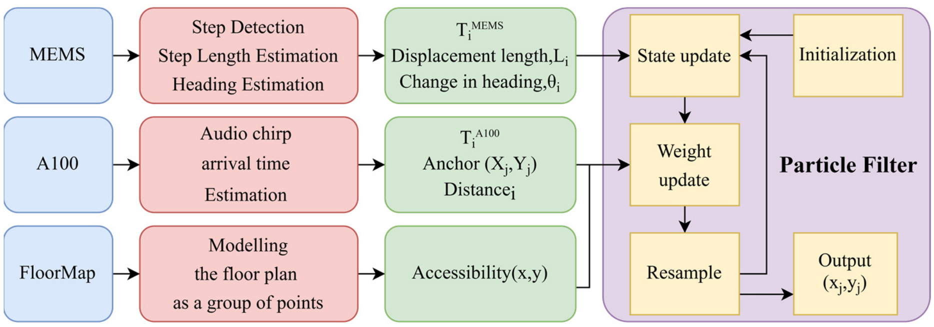


Figure 7. Architecture of our proposed fusion algorithm combining audio chirp/MEMS/floor map based on particle filtering

https://doi.org/10.5194/isprs-archives-XLVI-3-W1-2022-53-2022

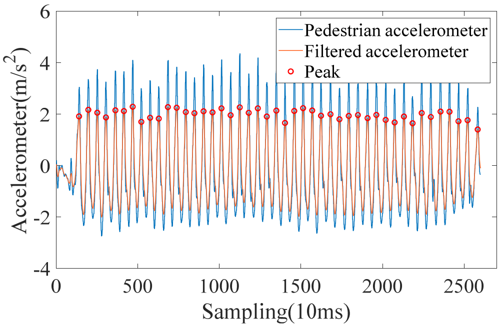


Figure 8. Acceleration smoothing and peak detection.

Typical methods for step length estimation include static models (Pratama, 2012) and dynamic models (Chen, 2015). The Weinberg model (Weinberg, 2002)is used in this paper, and the step length is estimated as

$$l_k = K \bullet \sqrt[4]{a_{\max} - a_{\min}} \tag{4}$$

where $a_{\max}$ is the maximum acceleration during the k_{th} step, $a_{\min}$ is the minimum, and K is a constant parameter equal to 0.41 in this paper

For the heading of a gait, two methods can be used: gyroscope based and magnetometer based (Lan, 2015). The gyroscope-based method is reliable only in a short time due to the accumulated measurement error of the gyroscope. The magnetometer-based method is subject to strong external magnetic interference. To improve heading accuracy more effectively, a heading estimation method with quasi-static and magnetic interference detection is derived from the Complementary Filter (CF).

To avoid the singularity problem of Euler angle, the three-dimensional attitude is expressed by the quaternion q.

$$q = q_0 + q_1 \cdot i + q_2 \cdot j + q_3 \cdot k \tag{5}$$

where $q_i\,(i = 0,1,2,3)$ are real numbers, i, j, k are imaginary units. Therefore, the attitude change rate is

$$\dot{\boldsymbol{q}} = \frac{1}{2}\hat{\boldsymbol{q}} \otimes \boldsymbol{\omega} = \frac{1}{2}\begin{bmatrix} -\boldsymbol{q}_v{}^T \\ q_0 I_{3\times3} + \left[\boldsymbol{q}_v^{\times}\right] \end{bmatrix}\boldsymbol{\omega} \tag{6}$$

where $\otimes$ is the quaternion product, $\hat{\boldsymbol{q}}$ is the normalization of $\boldsymbol{q}$, $\boldsymbol{q}_v = [q_1, q_2, q_3]^T$. $\boldsymbol{\omega} = \begin{bmatrix}0 & \omega_x & \omega_y & \omega_z\end{bmatrix}^T$, ω_x, ω_y, ω_z are outputs of the gyroscope. $\left[\boldsymbol{q}_v^{\times}\right]$ is the skew symmetric matrix

$$\left[\boldsymbol{q}_v^{\times}\right] = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix}^{\times} = \begin{bmatrix} 0 & -q_3 & q_2 \\ q_3 & 0 & -q_1 \\ -q_2 & q_1 & 0 \end{bmatrix} \tag{7}$$

To suppress drift error in (6), the Complementary Filter with quasi-static (8) and magnetic interference detection (9) is implemented.

$$\big|\|\boldsymbol{a}\| - \|\boldsymbol{g}\|\big| \le \gamma_a \tag{8}$$

$$\big|\|\boldsymbol{h}\| - \|\boldsymbol{m}\|\big| \le \gamma_m \tag{9}$$

where $\boldsymbol{a}$ and $\boldsymbol{g}$ are defined the same as in (1), $\boldsymbol{h}$ is the magnetometer outputs, $\boldsymbol{m}$ is the earth magnetic field, γ_a and γ_m the thresholds.

The kth corrected $\dot{\boldsymbol{q}}$

$$\dot{\hat{\boldsymbol{q}}}_k = \left(\frac{1}{2}\begin{bmatrix} -\hat{\boldsymbol{q}}_{k-1,v}{}^T \\ \hat{q}_{k,0} I_{3\times3} + \left[\hat{\boldsymbol{q}}_{k-1,v}^{\times}\right] \end{bmatrix}{}^S\boldsymbol{\omega}_t\right) \otimes C(\hat{\boldsymbol{y}}) \tag{10}$$

where $\boldsymbol{y} = \begin{bmatrix}\boldsymbol{a}^T & \boldsymbol{h}^T\end{bmatrix}^T$, $C(\hat{\boldsymbol{y}})$ is the correction for the fusion of acceleration and magnetic, which is obtained through Levenberg-Marquardt.

Hence, the attitude quaternion at k_{th} instance

$$\boldsymbol{q}_k = \hat{\boldsymbol{q}}_{k-1} + \dot{\hat{\boldsymbol{q}}}_k \Delta t \tag{11}$$

Δt is the sampling interval.

3.2 Particle Filter

3.2.1 Particle filter based on the Bayesian Inference and the Monte Carlo method

Particle filter (PF) is based on Bayesian inference, which provides mathematical machinery for a nonlinear system. Considering the uncertainty, it makes decisions according to the rational principle. In this positioning system, the fusion of MEMS /sound is described as a state model and a measurement model as follows:

$$\begin{cases} X_k = f\left(X_{k-1}, W_k\right) \\ Z_k = h\left(X_k, V_k\right) \end{cases} \tag{12}$$

where X_k and Z_k are the system state vector and the measurement vector at kth instance. W_k and V_k are the process noise and measurement noise, which obeying the Gaussian distribution with covariance matrix Q_k and R_k, i.e., $W_k \sim N\left(0, Q_k\right)$ and $V_k \sim N\left(0, R_k\right)$. $f(\bullet)$ is the nonlinear function reflects the relationship between the current and the previous state, while $h(\bullet)$ is between the state and the measurement.

To obtain an optimal estimation, the state estimation problem is converted to calculate the expectation of the posterior distribution, with an arbitrary function $g(\bullet)$, the expectation is

$$E\left[g\left(X_k\right) | Z_{1:k}\right] = \int g\left(X_k\right) p\left(X_k | Z_{1:k}\right) dX_k \tag{13}$$

that is to compute the posterior probability density function (pdf) $p\left(X_k | Z_{1:k}\right)$. The posterior PDF is derived from prior PDF $p\left(X_k | Z_{1:k-1}\right)$ with the conditional probability formula and the total probability formula

$$\begin{aligned} p\left(X_k | Z_{1:k}\right) &= \frac{p\left(Z_k | X_k, Z_{1:k-1}\right) p\left(X_k | Z_{1:k-1}\right)}{p\left(Z_k | Z_{1:k-1}\right)} \\ &= \frac{p\left(Z_k | X_k\right) p\left(X_k | Z_{1:k-1}\right)}{p\left(Z_k | Z_{1:k-1}\right)} \end{aligned} \tag{14}$$

where $p\left(Z_k | Z_{1:k-1}\right)$ is the normalization constant,

$$p\left(Z_k | Z_{1:k-1}\right) = \int p\left(Z_k \middle| X_k\right) p\left(X_k \middle| Z_{1:k-1}\right) dX_k \tag{15}$$

and $p\left(Z_k \middle| X_k\right)$ is the observation likelihood model. Assuming that the state transition obeys the first-order Markov process, the state X_k is determined only by X_{k-1}. The prior PDF $p\left(X_k | Z_{1:k-1}\right)$ is updated as follow:

$$
\begin{aligned}
p\left(X_k \mid Z_{1:k-1}\right) &= \int p\left(X_k, X_{k-1} \mid Z_{1:k-1}\right) dX_{k-1} \\
&= \int p\left(X_k \mid X_{k-1}, Z_{1:k-1}\right) p\left(X_{k-1} \mid Z_{1:k-1}\right) dX_{k-1} \\
&= \int p\left(X_k \mid X_{k-1}\right) p\left(X_{k-1} \mid Z_{1:k-1}\right) dX_{k-1}
\end{aligned} \tag{16}
$$

Since the system model is generally nonlinear and affected by non-Gaussian noise, there is of great difficulties on the solution of the prior PDF and the posterior PDF. Practically, in the Monte Carlo method, an importance distribution $x_k^{(i)} \sim q\left(X_k \mid Z_{1:k}\right)$ is used to update the weights by (17) and normalize by (18)

$$
\mathrm{w}_k^{(i)} \propto \frac{p\left(Z_k \mid X_k^{(i)}\right) p\left(X_k^{(i)} \mid X_{k-1}^{(i)}\right)}{q\left(X_k^{(i)} \mid X_{0:k-1}^{(i)}, Z_{1:k}\right)} \mathrm{w}_{k-1}^{(i)} \tag{17}
$$

$$
\overline{w}_k^{(i)} = \frac{w_k^{(i)}}{\sum_{i=1}^{N} w_k^{(i)}} \tag{18}
$$

And form an approximation as follows:

$$
p\left(X_k \mid Z_{1:k}\right) \approx \sum_{i=1}^{N} \overline{w}_k^{(i)} \delta\left(X_k - X_k^{(i)}\right) \tag{19}
$$

where $\delta(\bullet)$ is the Dirac function. As the number of particles N tends to infinity, the approximation will approach the true posteriori density. Generally, to avoid the weight of particles degenerating, the resample method is adopted to optimize. Finally, the state estimation of time k is performed as the average of particles:

$$
X_k = \left(\sum_{i=1}^{N} X_k^{(i)}\right) / N \tag{20}
$$

3.2.2 System Description

The indoor positioning system proposed in this paper is formulated as a dynamic state estimation problem in the discrete-time state-space domain. First, since we only consider ground targets, we choose to estimate the plane coordinates and heading of the target. Thus, the state vector can be represented as

$$
X_k = \begin{bmatrix} x_k & y_k & \theta_k \end{bmatrix}^T \tag{21}
$$

where (x_k, y_k) is the coordinate of the target at k_{th} instance in a relative coordinate frame, and θ_k is the direction.

The dynamic model describe the motion of the target from time $(k\text{-}1)$ to k is given as

$$
X_k = \begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} x_{k-1} + (l_k + \delta l_k) \cdot \sin(\theta_k) \\ y_{k-1} + (l_k + \delta l_k) \cdot \cos(\theta_k) \\ \theta_{k-1} + d\theta_k + \delta\theta_k \end{bmatrix} \tag{22}
$$

where l_k is the length of the displacement, $d\theta_k$ is the change of heading, δl_k and $\delta\theta_k$ are the corresponding measurement errors of l_k and $d\theta_k$, respectively, which obeying the Gaussian distribution with the standard deviation σ_l and σ_θ, i.e., $\delta l_k \sim N\left(0, \sigma_l^2\right)$ and $\delta\theta_k \sim N\left(0, \sigma_\theta^2\right)$.

In every cycle, the distance observations between the tag and every anchor are obtained by the audio chirp ranging technology of A100. The measurement vector is given as

$$
Z_k = [\; Dis_k^1 \; Dis_k^2 \; \cdots Dis_k^n \;]^T \tag{23}
$$

where $Dis_k^i \left(i = 1, 2, \cdots, n\right)$ is the distance between the mobile target (x_k, y_k) and i_{th} base station $\left(x_{BS}^i, y_{BS}^i\right)$ is as follows:

$$
Dis_k^i = \sqrt{(x_k - x_{BS}^i)^2 + (y_k - y_{BS}^i)^2} + \delta Dis_k^i \tag{24}
$$

where δDis_k^i is the corresponding measurement noise following a Gaussian distribution $\delta Dis_k^i \sim N\left(0, \sigma_{r,i}^2\right)$, and $\sigma_{r,i}$ is the distance standard deviation.

The corresponding prior value of Z_k is estimated from j_{th} particle

$$
Z_k^{pj} = \begin{bmatrix} Dis_{k,1}^{pj} & Dis_{k,2}^{pj} & \cdots & Dis_{k,n}^{pj} \end{bmatrix}^T \tag{25}
$$

$Dis_{k,i}^{pj}$ is an estimated distance vector calculated by i_{th} base station and j_{th} particle, it is formulated as:

$$
Dis_{k,i}^{pj} = \sqrt{(x_k^j - x_{BS}^i)^2 + (y_k^j - y_{BS}^i)^2} \tag{26}
$$

With n base stations, the observation noise covariance matrix R_k is expressed as

$$
R_k = diag\left(\begin{bmatrix} \sigma_{r,1}^2 & \sigma_{r,2}^2 & \cdots & \sigma_{r,n}^2 \end{bmatrix}\right) \tag{27}
$$

The weights are updated with the Gaussian distribution

$$
\mathrm{w}_k^{(j)} = \frac{1}{2\pi R_k} \bullet \exp\left\{-\frac{\left(Z_k - Z_k^{pj}\right)^2}{2R_k}\right\} \tag{28}
$$

3.3 Modelling the floor plan as a group of points

Map information is also very important for positioning. There are many areas of indoor space that cannot be crossed or accessed, such as walls, pillars, tables and other obstacles. It can be used to correct unreasonable positioning results. For particle filtering, it is also useful to reduce the number of particles for which weights need to be calculated, thereby reducing the amount of computation. Especially when approaching an obstacle, the observation signal is more likely to be affected by NLOS. Dividing the positioning area into two parts, the accessible and the inaccessible, is significant for practical applications.

Many scholars have already proposed using map information to determine whether a particle crosses an obstacle during its movement, which requires modelling the map to obtain a surface or line describing the obstacle, and then determining whether it intersects or is tangent to the obstacle based on the line connecting the positions of the particle before and after its movement, which is very computationally intensive. In this paper, the map is divided into small areas, considered as one point after another, and the two-dimensional plane is transformed into a two-dimensional matrix that represents accessibility by a simple binary value, such as 0 for unreachable and 1 for reachable.

$$
Floor\ Map \approx Accessibility(x, y) \tag{29}
$$

During the filtering, it can quickly find the accessibility value of the location based on the predicted point coordinates of the particle, determine whether the particle has stepped into an obstacle, and then and then implement the constraint of the map to the location as follows:

$$
\mathrm{w}_k^{(j)} = \begin{cases} \mathrm{w}_k^{(j)} & Accessibility\left(x_k^j, y_k^j\right) = 1 \\ 0 & Accessibility\left(x_k^j, y_k^j\right) = 0 \end{cases} \tag{30}
$$

4. EXPERIMENTAL VERIFICATION AND DISCUSSION

A. RANGING PERFORMANCE

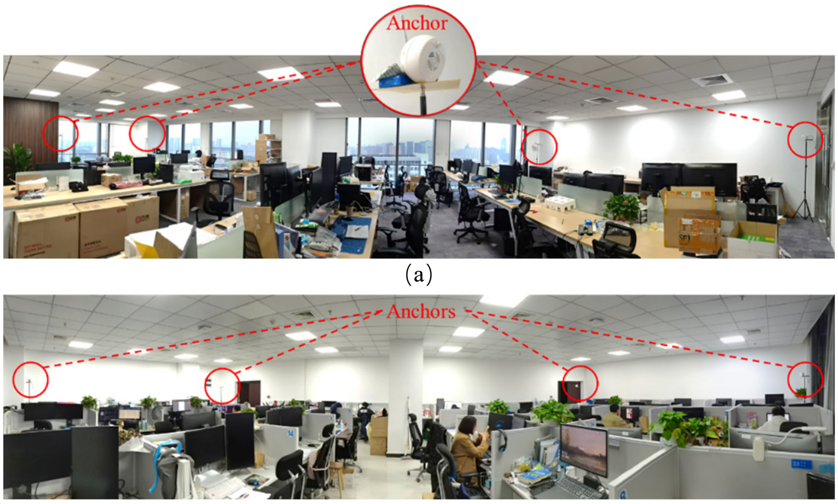
In this part, the ranging performance of A100 over a long distance is evaluated. The experiments are conducted in two scenarios:

(a) **(b)**

Figure 9. Ranging experimental scenarios: (a) The corridor; (b) The underpass.

Scenario		Facing-to-speakers					Back-to-speakers				
		Ground Truth(m)	Average Error(m)	Max Error(m)	RMSE (m)	Detection Rate	Ground Truth(m)	Average Error(m)	Max Error(m)	RMSE (m)	Detection Rate
Corridor		8.00	0.26	1.81	0.45	99.68%	8.00	0.22	0.92	0.32	96.51%
		16.00	0.19	1.45	0.40	100.00%	16.00	0.22	1.48	0.38	100.00%
		24.00	0.28	1.08	0.41	100.00%	24.00	0.77	2.07	0.97	93.02%
		32.00	0.36	1.21	0.50	100.00%	32.00	1.14	2.20	1.22	93.01%
Underpass		8.00	0.46	1.40	0.55	86.62%	8.00	0.57	1.88	0.62	95.72%
		16.00	0.06	1.59	0.40	99.84%	16.00	0.33	2.41	0.56	91.50%
		24.00	0.01	1.03	0.37	99.19%	24.00	0.58	2.47	0.69	90.90%
		32.00	0.16	0.39	0.25	99.36%	32.00	0.70	2.96	0.94	83.45%

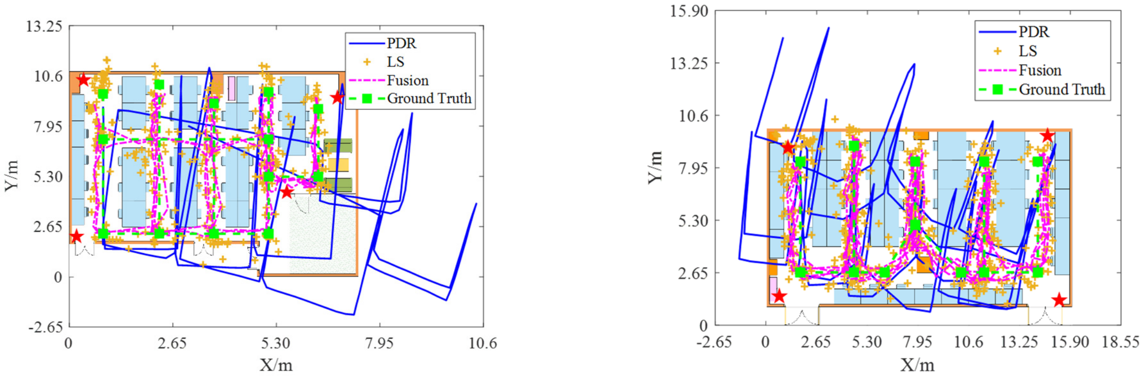
Table 1. Evaluation of A100 ranging error.



(a)

(b)

Figure 10. Positioning experimental scenarios: (a) The office; (b) The laboratory.



(a) (b)

Figure 11. Trajectories estimated by different algorithms: (a) Test 1; (b) Test 2

one is an ideal quiet corridor about 33m without any interference shown in Figure 9a, the enclosed narrow structure prevents rapid attenuation of sound signals; and the other is a noisy underpass (8m*70m) with pedestrian activity shown in Figure 9b. In each set of experiments, speakers were placed at the end of the scene at a height of 1.5 meters. The development board with the same height measured distances at 8m, 16m, 24m and 32m, respectively. Connect the development board to the computer, collect and save about a minute of data through the serial port. In addition, to assess the impact of human occlusion, we collected two sets of data at each point, one is speaker-oriented, the other is back-to-speaker.

As summarized in Table 1, in the corridor scene, the signal detection rate of A100 is close to 100 % at 8m, 16m, m, 24m, 32m, during the test of speaker-oriented; the average error is about 0.3m, and the root mean square error (RMSE) is less than 0.5m, which indicates that in the ideal environment, the attenuation of sound signal with distance is inconspicuous, and A100 obtain ranging information stably; while in the back-to-speaker test, the direct path of the signal is blocked by the body, resulting in the detection rate decreases significantly. When A100 is close to the speaker, the strong signal is received, and the ranging accuracy is almost the same as the forward. While the distance increases to 32 m, the signal becomes weaker, coupled with occlusion, the average error is more than 1 m, and changes fiercely with a RMSE exceeding 1.2 m. But overall, the ranging performance is still very considerable.

However, in the complex environment of underpass, where social activities are frequent, and environmental noise is larger, the overall signal detection rate and ranging accuracy are significantly decreased, and due to the interference of human flow, the average error has no obvious rule with the change of distance (there is no control of human flow in the experimental process, and more people flow, more serious the occlusion of the signal is, resulting in the loss of ranging data or the degeneration of accuracy). However, the ranging quality in the backward situation is significantly lower than that in the positive situation, and the detection rate is only 83.45 % at 32 m. The results indicate that the signal occlusion has a greater impact on the A100 audio chip ranging.

B. SYSTEM LOCATION PERFORMANCE

Through static experiments, we have glimpsed the ranging performance of A100. But in the actual positioning scene, most of the time the target is in motion, which will introduce the Doppler effect, making the chirp signal collected by the microphone offset, reducing the ranging accuracy. In addition, there are not only social activities in the environment, but also many furniture and facilities, making the propagation of sound more complex, which also cause deviation or even error in ranging. In order to achieve a more stable positioning system, this paper proposes a multi-source data fusion based on particle filter. The Kepler A100 is used to receive the chirp signal and estimate the distance between the target and the base station, while PDR based MEMS is used to calculate the moving step length and course of the target, and the map information of the positioning environment is used to adjust the weight of the particles.

To verify the feasibility and effectiveness of the proposed solution, field experiments are carried out in two scenarios shown in Figure 10, one is the student laboratory; the other is the office of the enterprise. On one hand, the main feature of such scenarios is that facilities account for a large proportion of the area, while the passable area is narrow, and walking inside requires more turning, i.e., frequent changes in the state of motion, which aggravate the positioning scheme using PDR-only and the error accumulates faster. On the other hand, for the positioning scheme using signal ranging, there are desks and chairs close to the height of the hand-held tag, resulting in the multi-path of the signal more obvious and complex, and more prone to NLOS.

For demonstrating the performance of the proposed positioning system more intuitively, the trajectories obtained by various schemes are visualized, as shown in Figure 11. The ground truth is marked with the green dotted line, and the turning points in the movement are marked with the green solid block represents. In the experiment, we carry a development board that integrates an IMU and a A100, and walk along the reference trajectory from one point. Every time we reach at a green block, the course is changed once. After all the paths are completed, we backtrack to the starting point along the path, and finally repeat the process, i.e., the whole path took two rounds. The blue trajectory is the result of the PDR trajectory that recursively from a starting point. The trajectory is smooth, and similar to the real at the beginning, but drifts severely; and as time goes by, the positioning result is gradually far away from the actual. And the yellow mark indicates the results of the least squares (LS) method. It is shown that most points are distributed around the real path, while a few points are obviously unreasonable, going into the table or flying out of the wall. Finally, the results of the proposed algorithm are represented by the dotted line of magenta. It can be clearly seen that there are very few positioning points into the unreachable area, which benefits from map information constraint. Moreover, compared with the results of LS method, the fusion result is more suitable for the real path, profits from the short-term relative displacement provided by PDR, weakens the gross error caused by NLOS in some degree and makes the trajectory smoother.

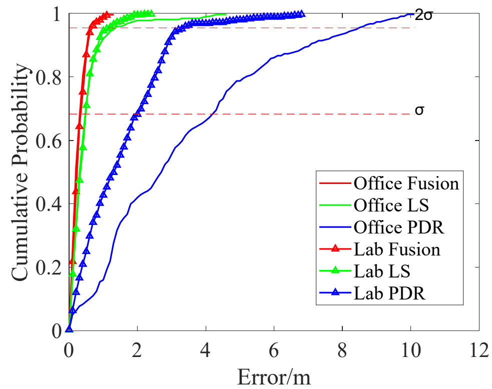


Figure 12. Positioning error CDF of different algorithms.

Indicator	Office			Laboratory		
	PDR	LS	Fusion	PDR	LS	Fusion
Max(m)	10.11	4.69	1.33	6.89	2.47	1.12
Average(m)	3.20	0.46	0.25	1.49	0.41	0.27
Median(m)	2.70	0.30	0.20	1.30	0.33	0.23
STD(m)	2.47	0.61	0.20	1.22	0.36	0.21
CDF-68%(m)	4.20	0.50	0.30	1.90	0.50	0.40
CDF-95%(m)	8.50	1.40	0.60	3.30	1.20	0.70

Table 2. Error statics of different positioning algorithms.

Figure 12 shows the cumulative distribution function (CDF) of localization errors of these three methods. The smooth curves and the curves with triangles represent the results of the laboratory and the office, respectively. It indicates that the accuracy of PF or LS does not change much in different environments; however, the PDR result in the office is significantly larger than that in the laboratory. We speculate that the PDR drifts more severely in the office due to the larger scope and the longer test time.

As summarized in Table 2, the 95% positioning error of PF in both tests are less than 1 m (0.60m for the office, and 0.7m for the laboratory), which is a reduction of more than 78.7% over PDR (5.03 m), 41.6% over LS. Additionally, the error of PF changes most stably in any scenario a standard deviation of about 0.2m, and the maximum is better than 1.5 m. The results illustrate that the proposed PF based on sound/PDR/floor map achieves approximately 0.65 m during 95% of the time. This outperforms solutions PDR-only-based and sound-only-based.

5. CONCLUSION

This paper presents an integrated audio chirp/MEMS/map system for indoor localisation. The audio chirp is received by the Kepler A100 chip and we evaluate the ranging performance of the A100 in two typical indoor environments. The results show a signal detection rate of over 80% and a ranging accuracy of sub-metre. The sensor fusion algorithm is derived from a particle filter with a map constant approach. The filter predicts the state vector from the PDR and updates it with distance observations from the A100. During important sampling, particles are restricted to accessible regions with map information. Experiments were conducted in two complex scenarios and the results show that the proposed system has a 95% error of 0.65 m, which is at least 40% lower than the error of a single PDR and a single sound, demonstrating the efficiency and stability of the system.

REFERENCES

Ayub, S., Zhou, X., Honary, S., Bahraminasab, A., Honary, B., 2012. Indoor Pedestrian Displacement Estimation Using Smartphone Inertial Sensors. *Int. J. Innovative Computing and Applications*, 4, 35-42.

Chen, R.Z., Chu, T., Liu, K., Liu, J., Chen, Y., 2015. Inferring Human Activity in Mobile Devices by Computing Multiple Contexts. *Sensors*, 15(9), 21219–21238.

Chen, R.Z., Chen, L., 2021. Smartphone-Based Indoor Positioni ng Technologies. *in Urban Informatics*, 467-490. [Online]. Avai lable: https://link.springer.com/book/10.1007/978-981-15-8983-6.

Cao, S., Chen, X., Zhang, X., Chen, X., 2020. Effective Audio Signal Arrival Time Detection Algorithm for Realization of Robust Acoustic Indoor Positioning. *IEEE Transactions on Instrumentation and Measurement*, 69(10), 7341 - 7352.

Dabove, P., Di Pietra, V., Piras, M., Jabbar, A. A., Kazim, S. A., 2018. Indoor positioning using Ultra-wide band (UWB) technologies: Positioning accuracies and sensors' performances. *2018 IEEE/ION Position, Location and Navigation Symposium (PLANS)*. doi.10.1109/plans.2018.8373379.

Farrokhi, H., Palmer, R. J., 2005. The designing of an indoor acoustic ranging system using the audible spread spectrum LFM (CHIRP) signal. *Canadian Conference on Electrical and Computer Engineering*. doi.10.1109/ccece.2005.1557409.

Hoflinger, F., Hoppe, J., Zhang, R., Ens, A., Reindl, L., Wendeberg, J., Schindelhauer, C. 2014. Acoustic indoor-localization system for smart phones. *2014 IEEE 11th International Multi-Conference on Systems, Signals & Devices (SSD14)*. doi.10.1109/ssd.2014.6808774.

Jin, Y., Toh, H.S., W.S. Soh, Wong, W.C., 2011. A Robust Dead Reckoning Pedestrian Tracking System with Low-Cost Sensors. *IEEE International Conference on Pervasive Computing and Communications*. doi.10.1109/PERCOM.2011.5767590.

Lan, H., Yu, C., El-Sheimy, N, 2015. An Integrated PDR/GNSS Pedestrian Navigation System. *China Satellite Navigation Conference (CSNC) 2015 Proceedings*, 3, 677-690.

Liu, T., Niu, X., Kuang, J., Cao, S., Zhang, L., Chen, X., 2020. Doppler shift mitigation in acoustic positioning based on pedestrian dead reckoning for smartphone. *IEEE Transactions on Instrumentation and Measurement*, 70, 1–1.

Pratama, A.R., Hidayat, R., 2012. Smartphone-based Pedestrian Dead Reckoning as an Indoor Positioning System. *International Conference on System Engineering and Technology (ICSET)*.

Puthal, D., Nepal, S., Ranjan, R., Chen, J., 2016. Threats to Networking Cloud and Edge Datacenters in the Internet of Things. *IEEE Cloud Computing*, 3(3), 64–71.

Rai, A., Chintalapudi, K. K., Padmanabhan, V. N., Sen, R., 2012. Zee: Zero-Effort Crowdsourcing for Indoor Localization. *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, 293–304.

Tian, Z.S., Zhang, Y., Zhou, M., Liu, Y., 2014. Pedestrian dead reckoning for MARG navigation using a smartphone. *EURASIP Journal on Advances in Signal Processing*.

Wu, C., Wang, X., Chen, M., Kim, M. J., 2019. Differential received signal strength-based RFID positioning for construction equipment tracking. *Advanced Engineering Informatics*, 42, 100960. doi.10.1016/j.aei.2019.100960.

Weinberg, H., 2002. Using the ADXL202 in Pedometer and Personal Navigation Applications. Analog Devices AN-602 Application Note.

Wang, Q., Luo, H., Men, A., Zhao, F., Gao, X., Wei, J., Huang, Y., 2018. Light positioning: A high-accuracy visible light indoor positioning system based on attitude identification and propagation model. *International Journal of Distributed Sensor Networks*, 14(2). doi:10.1177/1550147718758263.

Xiao, X., Zhao, S., Zhong, X., Jones, D. L., Chng, E. S., Li, H., 2015. A learning-based approach to direction of arrival estimation in noisy and reverberant environments. *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. doi.10.1109/icassp.2015.7178484.

Yan, J., Cao, Y. M., Kang, B., Wu, X.H., Chen, L., 2021. An ELM-Based Semi-Supervised Indoor Localization Technique with Clustering Analysis and Feature Extraction. *IEEE Sensors Journal*,,21(3), 3635-3644.

Yu, K., Wen, K., Li, Y., Zhang, S., Zhang, K., 2019. A Novel NLOS Mitigation Algorithm for UWB Localization in Harsh Indoor Environment. *IEEE Transactions on Vehicular Technology*, 68(1), 686 - 699. doi.10.1109/tvt.2018.2883810.

Zuo, J., Liu, S., Xia, H., 2018. Multi-phase fingerprint map based on interpolation for indoor localization using iBeacons. *IEEE Sensors J.*, 18(8), 3351-3359.

Zhang, L., Chen, M., Wang, X., Wang, Z., 2018. TOA Estimation of Chirp Signal in Dense Multipath Environment for Low-Cost Acoustic Ranging. *IEEE Transactions on Instrumentation and Measurement*, 68(2), 355–367. doi:10.1109/tim.2018.2844942.