BatMapper-Plus: Smartphone-Based Multi-level Indoor Floor Plan Construction via Acoustic Ranging and Inertial Sensing

Chuize Meng, Shan Jiang, Mengning Wu, Xuan Xiao, Dan Tao, and Ruipeng Gao(✉)

Beijing Jiaotong University, Beijing 100044, China
{mengchuize,hilljiang,meningwu,xiaoxuan,dtao,rpgao}@bjtu.edu.cn

Abstract. The lack of floor plans is one of the major obstacles to ubiquitous location-based services indoors. Dedicated mobile robots with high-precision sensors can scan and produce indoor maps, but the deployment remains low. Existing smartphone-based approaches usually adopt computer vision techniques to build the 3D point cloud, at the cost of extensive image collection efforts and the risk of privacy issues. In this paper, we propose BatMapper-Plus which constructs accurate and complete indoor floor plans by acoustic ranging and inertial sensing on smartphones. It employs acoustic signals to measure the distance to a nearby wall segment, and produces the accessible area by surrounding the building during walking. It also refines the constructed floor plan to eliminate scattered segments, and identifies connection areas including stairs and elevators among different floors. Extensive experiments in a teaching building and a residential building have shown our effectiveness compared with the state-of-the-art, without any privacy concerns and environmental limitations.

Keywords: Floor plan construction · Acoustic ranging · Inertial sensing · Smartphone

1 Introduction

Indoor location-based services (LBS) brings great convenience to our modern life, especially at large-scale hospitals, multi-level shopping malls, and underground parking lots. However, its deployment is still not yet pervasive, and one of its major obstacles is the lack of floor plans for indoor localization [1] and navigation [2].

At present, existing dedicated mapping systems [3] rely on mobile robots with cameras and other high-precision sensors to construct accurate indoor maps. However, such systems always cost expansively and are not wide-spread at large scale. With commodity smartphones, some AR/VR applications adopt computer

L. Wang et al. (Eds.): WASA 2022, LNCS 13472, pp. 155–167, 2022.
https://doi.org/10.1007/978-3-031-19214-2_13

vision techniques [4] to build 3D point clouds for indoor objects, but images are affected by ambient light condition and risk privacy disclosure. Therefore, it is necessary to construct indoor floor plans without environmental supports and privacy concerns.

In a recent work [5], we have proposed BatMapper which uses acoustic signals to measure the distance for indoor mapping. Specially, it adopts a bilateral acoustic ranging mechanism that is effective to construct narrow corridors. SAMS [6] follows this work with higher accuracy by FMCW-based distance measurement. However, indoor environments are not limited to corridors, but also include wide areas such as rooms and lobbies. In such places, BatMapper can not produce satisfactory floor plans due to data association mistakes between wall segments and distance measurements. In addition, the constructed floor plan should be further adjusted to eliminate scattered points/segments, and augmented with connection areas to other floors.

In this paper, we propose BatMapper-Plus which is a smartphone-based indoor floor plan construction system by acoustic ranging and inertial sensing. It employs an unilateral ranging mechanism which measures the precise distance to a nearby wall segment, and produces the accessible area by walking around the room. It also refines scattered wall segments, and identifies connection areas to produce a multi-level floor plan. The system can build the indoor floor plan through a smart phone, which makes it have a low cost. Furthermore, the system does not need indoor images, so it is not affected by lighting conditions, and has high privacy.

Specially, we make the following contributions in this work:

– We explore a novel unilateral acoustic ranging method on smartphones. It emits acoustic signals for distance measurement to a side wall, thus users can simply surround the building to construct its floor plan with lightweight human efforts.
– We propose a map refinement algorithm to produce accurate and complete multi-level floor plans. It automatically adjusts and merges the scattered wall segments. It also detects and marks connection areas (e.g., stairs and elevators) on the map.
– We build a prototype and conduct extensive experiments in a teaching building and a residential building. Results have shown our improvements with around 2.8% in three experimental scenarios on F-score compared to BatMapper.

2 Overview

In this section, we present how the BatMapper works, explain its limitations during deployment in reality, and depict the overview of our BatMapper-Plus.

Background on BatMapper. BatMapper designs a two-pulse signals with linear frequency increasing sine waves and Hanning window reshaping for bilateral ranging on smartphones. It further explores a probabilistic evidence accumulation (PEA) method to associate the distance measurements to corresponding wall segments along the long corridor.

Limitation in BatMapper. 1) BatMapper measures the distance from echo signals which are reflected by two-side walls. This bilateral ranging mechanism is suitable in narrow areas such as a long corridor, but not designed for spacious spaces such as large rooms and lobbies. 2) BatMapper produces coarse floor plans which are composed of scattered points, and they are not consistent with the actual maps made of line segments. 3) BatMapper only generates the floor plan for one level, while modern buildings are always comprised of multiple levels with various connection areas.

BatMapper-Plus Overview. As Fig. 1 shows, BatMapper-Plus employs acoustic signals and inertial data as inputs. We adopt the Frequency Modulated Continuous Wave (FMCW) as the speaker's output signal, use two microphones (top/bottom) to receive the echo signal reflected from the side-wall, and calculate the distance between the wall and the smartphone. In addition, we collect the inertial data from smartphone to track the user and identify connection areas (e.g., stairs and elevators). Finally, we fuse distance measurements, walking trajectories and connection areas to construct and refine the multi-level indoor floor plan.

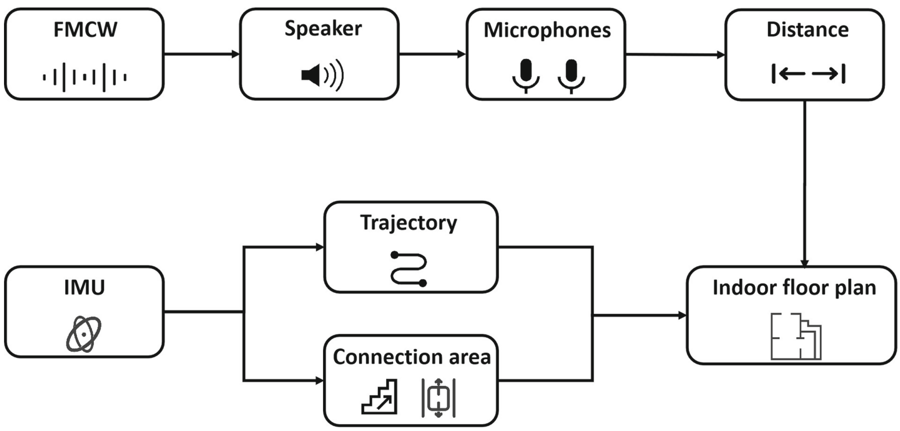


Fig. 1. Overview of BatMapper-Plus, which constructs multi-level indoor floor plan by acoustic ranging and inertial tracking.

3 Methods

In this section, we present the detailed design of BatMapper-Plus with three modules: unilateral acoustic ranging, map construction and refinement, and connection area detection.

3.1 Unilateral Acoustic Ranging

Acoustic Signal Design. The frequency of acoustic signal should balance both physical ability of smartphone and background noises. The sound frequency of smartphone is usually between $110\,Hz \sim 20\,KHz$, while the frequency of human voice is always less than $1\,KHz$. Through experiments, we found that a wider frequency range will make the echo peak more obvious. Furthermore, with the same emission energy, low-frequency sound spreads farther than high-frequency sound. Therefore, we generate the duration of acoustic signal as 3 ms and its frequency range is $8\,KHz \sim 16\,KHz$. Specially, We employ the Frequency Modulated Continuous Wave (FMCW) to produce the signal (the blue line in Fig. 2).

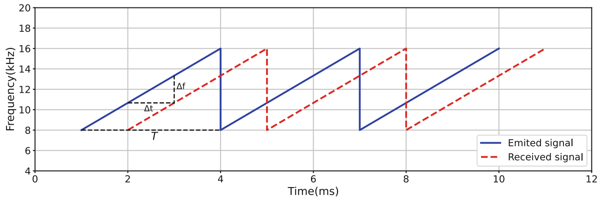


Fig. 2. Calculating distance by emited and received signals. (Color figure online)

Delay of Echo Signal. When the emitted acoustic signal meets the wall, its echo signal is reflected back and received by smartphone's microphone. At this time, the waveform of the received signal and the emitted signal are consistent, with a time delay $\varDelta t$ which is shown in Fig. 2. Specially, the time delay $\varDelta t$ is computed as:

$$\varDelta t = \frac{\varDelta f \cdot T}{f_{max} - f_{min}} \tag{1}$$

where f_{min} is the minimum frequency of the emitted signal, f_{max} is the maximum frequency of the emitted signal, and T is the duration of the emitted signal.

Distance to Wall. As shown in Fig. 3(a), when a user holds the smartphone horizontally during walking, Path 1 indicates the propagation path of the signal received by the top microphone, and Path 2 indicates the propagation path of the signal received by its bottom microphone. The propagation distance difference of received signals between by two microphones is the length l of the smartphone. Therefore, when the distance difference to the side-wall measured by two microphones is close to $\frac{1}{2}l$, the received signal is likely to be the echo reflected from the wall.

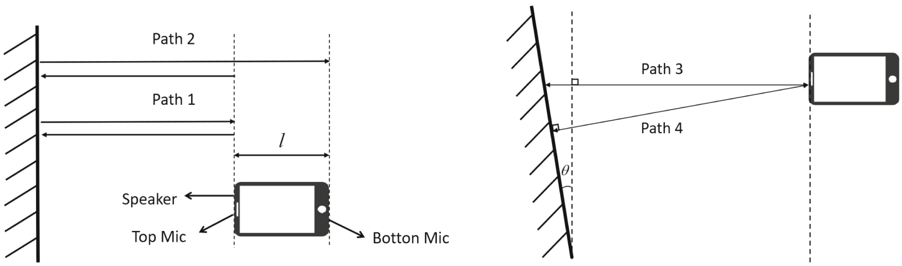
In this case, the distance d between the wall and the smartphone (by its top microphone in Fig. 3(a)) is expressed as:

$$d = \frac{1}{2} \Delta t \cdot v_{sound} \tag{2}$$

where v_{sound} is the sound propagation speed.

Through experiments, we found that when there is an angle θ between the mobile phone and the wall, the detected distance is the vertical distance from the top microphone to the wall (Path 4 in Fig. 3(b)), rather than the distance towards the top of the mobile phone (Path 3 in Fig. 3(b)). Therefore, when the mobile phone is not completely parallel to the wall, the accurate distance can still be obtained.

In addition, we use one-dimensional sound signal, which can only represent distance information and cannot distinguish indoor conditions. So this method has high privacy.



(a) Parallel to the wall (b) Not parallel to the wall

Fig. 3. The position of the speaker and microphone, and the sound signal propagation path received by the microphone.

3.2 Map Construction and Refinement

Inertial Tracking. We use the dead-reckoning to track the walking user by his/her stride length, step count and heading orientation. 1) The normal stride length for an adult is about 60 cm, we use such value as default, and it can be customized by the outdoor trajectory with GPS [7]. 2) The step count is calculated by detecting the peak and valley values of vertical accelerations. Based on our experiments, we set the threshold of their difference as $3\,\mathrm{m/s^2}$, and set the duration threshold between two steps as 400 ms in order to avoid errors caused by hand shaking. 3) In order to eliminate drifts from gyroscope and noises from magnetometer, we calculate the orientation by *gamerv* API of smartphone, which fuses accelerometer, gyroscope and magnetometer for robustness.

Door and Window Detection. We judge the existence of doors and windows by detecting the distance variation between the smartphone and the side-wall. When the change exceeds the threshold (20 cm in our system), we regard the point as a door/window. When the length of door/window is too short, we remove these points as outliers.

Floor Plan Refinement. The wall segment are positioned as scattered points on the map by distance measurements oriented to user's trajectory. Not only it contains outliers with extreme errors, but also such floor plan is not consistent with the actual map.

Intuitively, most walls are made of line segments, and their intersections could be identified by turning events of the walking user. After removing the detected doors and windows, we fit the rest points as line segments, i.e.,

$$f(x) = kx + b \tag{3}$$

where k is the gradient and b is the offset. In order to minimize the errors between a wall segment and the scattered points, we construct the objective function as:

$$e = \sum_{i}^{n}(kx_i + b - y_i) \tag{4}$$

where e represents the sum of errors, n represents the number of points on the wall, (x_i, y_i) is the ordinate of the i_{th} point by distance measurement. In order to minimize e, we calculate the partial derivatives of k and b respectively:

$$\frac{\partial e}{\partial k} = 2(\sum_{i}^{n}(kx_i + b - y_i)x_i) = 0 \tag{5}$$

$$\frac{\partial e}{\partial b} = 2(\sum_{i}^{n}(kx_i + b - y_i)) = 0 \tag{6}$$

thus

$$k = \frac{\sum_i^n(x_i y_i) - n\overline{xy}}{\sum_i^n(x_i)^2 - n\overline{x}^2} \tag{7}$$

$$b = \overline{y} - k\overline{x} \tag{8}$$

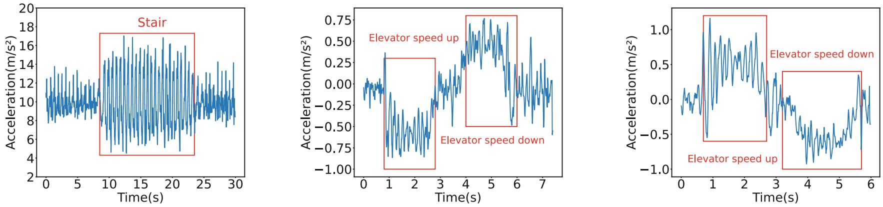
where $\overline{x}$ is the average value. Thus, we refine the reconstructed floor plan with wall segments, doors/windows, and corners.

3.3 Connection Area Detection

Since modern buildings are always multi-levels with different types of connection areas, we automatically identify stairs and elevators to associate each level of floor plan.

Stair Detection. Intuitively, there are always large variations in our accelerations when climbing stairs. In order to eliminate the influence of smartphone's attitude, we calculate the amplitude value of three-axis accelerations on smartphone (Fig. 4(a)). Next, we use a sliding window to dynamically detect the peaks and valleys along acceleration sequence. In order to avoid errors caused by hand shaking, the minimum time gap between peaks and valleys is set as 400 ms. Because when holding a mobile phone, it takes one second to take one step.

Elevator Detection. When an elevator starts or stops, the smartphone's acceleration along gravity direction varies accordingly, and it remains stable when the elevator moves at an uniform speed. As shown in Fig. 4(b) and Fig. 4(c), the vertical acceleration first decreases than increases when the elevator goes down, and verse versa. We adopt a sliding window of 3 s to detect its rising/sinking interval, and verify if such two periods are within a reasonable period (30 s in our system).



(a) Stair (b) An elevator sinks (c) An elevator sinks

Fig. 4. Three-axis acceleration variations at stairs and acceleration variations along gravity direction when an elevator rises and sinks.

4 Evaluation

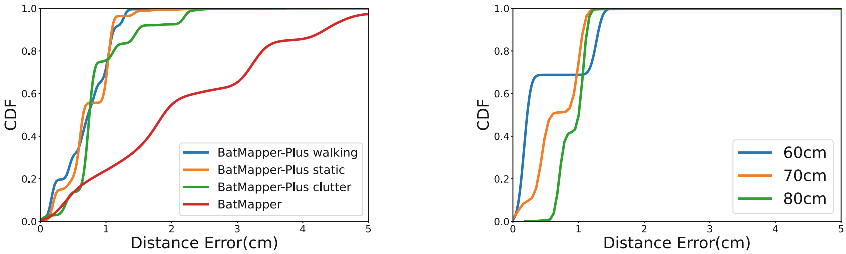
We have developed the prototype of BatMapper-Plus on Android Studio and installed it on MI 10S smartphone for data collection. Experiments are carried out in a residential building and a teaching building, both with multiple levels. The ground true distance and floor plan are measured by a laser rangefinder. Our evaluation includes three aspects, i.e., unilateral acoustic ranging, floor plan construction, and connection area detection.

4.1 Unilateral Acoustic Ranging

We carried out the acoustic ranging measurements indoors, with different user states and distances to wall.

User States. We test the accuracy with static users, walking users and a cluttered environment with many obstacles, and compare with the previous BatMapper. As shown in Fig. 5(a), our distance measurement error in the static state is close to that in walking state, with the median value around 0.7 cm and the maximum value less than 1.5 cm, both are obviously lower than the BatMapper. In a cluttered environment, the accuracy decreases slightly. This demonstrates our effectiveness of one-side ranging.

Distance to Wall. Three distances are tested at 60 cm, 70 cm and 80 cm to the same wall. As shown in Fig. 5(b), all distance measurement errors are less than 2 cm. In addition, such error increases with the farther distance.



(a) Different user states

(b) Different distances to wall

Fig. 5. Unilateral acoustic ranging with different user states and distances to wall.

4.2 Floor Plan Construction

Construction Effect. We illustrate the construction effect in an 8.7 m $\times$ 6.3 m classroom in teaching building, a 7 m $\times$ 4.5 m living room and a 6 m $\times$ 11 m corridor in residential building. The three experimental scenarios contain several doors, windows, and blackboards/closets. In order to construct the floor plan, a user hold the smartphone horizontally and walk along the experimental scenarios border, and we produce the position of wall segments by acoustic ranging. Such scattered points are drawn on the map based on the walking trajectory (Fig. 6(b), Fig. 7(b) and Fig. 8(b)). Next, our refinement algorithm improve the constructed floor plan with corners and line segments (Fig. 6(c), Fig. 7(c) and Fig. 8(c)).

Quantitative Results. In order to evaluate the reconstructed floor plan precisely, we overlay it onto the ground truth to achieve the maximum overlap (Fig. 9), and observe that the location errors of wall segments are all within 0.3 m. Next, we define the precision, recall and F-score as:

$$P = \frac{S_{re} \bigcap S_{gt}}{S_{re}}, R = \frac{S_{re} \bigcap S_{gt}}{S_{gt}}, F = \frac{2P \cdot R}{P + R} \tag{9}$$

where S_{re} denotes the shape of reconstructed map, S_{gt} denotes the shape of the ground truth, and $S_{re} \bigcap S_{gt}$ represents their overlap area.

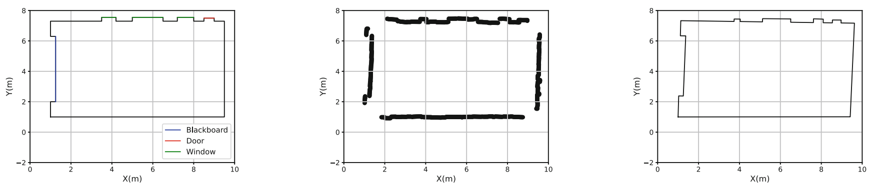
Table 1 shows the quantitative results for floor plan construction in classroom, living room and corridor. Compared with BatMapper and CrowdInside [8], the recall and F-score of our BatMapper-Plus are significantly higher than the other methods, which indicates that we produce more precise indoor maps. In addition, since the other methods generate floor plans with larger areas than the ground truth, their recall values are higher.

Table 1. Shape evaluation of floor plans.

Sense	Classroom			Living room			Corridor		
Criterion	R(%)	P(%)	F(%)	R(%)	P(%)	F(%)	R(%)	P(%)	F(%)
CrowInside	77.28	100	87.18	74.80	100	85.58	58.82	100	74.07
BatMapper	96.36	99.61	97.96	97.46	96.21	96.58	84.71	94.12	89.17
BatMapper-Plus	97.89	99.48	98.68	99.05	98.46	98.75	97.65	91.76	94.61

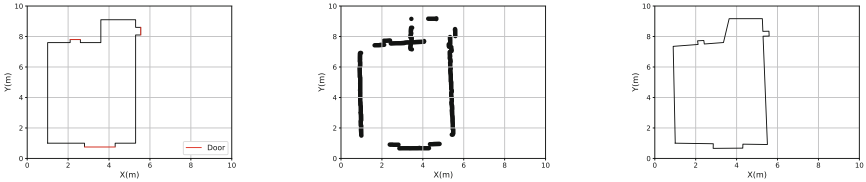
4.3 Connection Area Detection

Stairs. In order to evaluate the stair detection accuracy, users conduct a 2-minute walk either on the ground or climbing stairs, and we predict the location type for each walking step. In addition, we collect the inertial data with different postures during walking, i.e., holding the smartphone horizontally or with an arbitrary posture. As shown in Table 2, there are at most six incorrect steps on each walk for all postures, with an approximate accuracy of 97%. We look into such incorrect steps and find them are mainly located at the junction area on each floor, with slight impacts on stair detection.



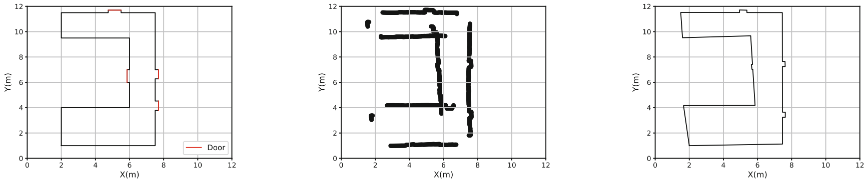
(a) Ground truth (b) Scattered points (c) Constructed map

Fig. 6. Construction process of a classroom in teaching building.



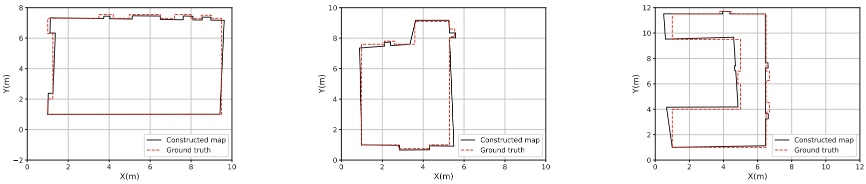
(a) Ground truth (b) Raw points (c) Constructed map

Fig. 7. Construction process of a living room in residential building.



(a) Ground truth (b) Raw points (c) Constructed map

Fig. 8. Construction process of a corridor in residential building.



(a) Classroom (b) Living room (c) Corridor

Fig. 9. Comparison between reconstructed floor plans and the ground truth.

Table 2. Accuracy of stair and elevator detection.

Area	Stair		Elevator		
State	Arbitrary posture	Horizontal posture	Rising	Stable	Sinking
Accuracy	96.84%	97.06%	100%	100%	100%

Elevator. As shown in Table 2, for a statically standing user, all test data are correctly detected (either on rising/sinking elevators or on the ground), thus the detection accuracy of elevators reaches 100%.

5 Related Work

Indoor Floor Plan Construction. At present, the construction of indoor floor plan is mainly realized by the combination of image and inertial sensor. Jigsaw [4] obtains the spatial relationship between adjacent landmark objects from the image taken by the user and the inertial sensor data, and combines the user's trajectory and the position of the captured image to generate a complete plan. Plansketcher [9] uses deep learning technology to extract new comprehensive features to identify different landmarks. Then the indoor floor plan is constructed based on sensing data, depth data and images. [10] combines the magnetic fingerprint map and user trajectory to build the indoor floor plan. IndoorCrowd2D [11] uses the image information and sensory data in crowdsourcing data to restore the building's structure. MapGENIE [12] uses syntax to represent the structural information of buildings, which has a better effect than simple trajectory mapping.

Acoustic Ranging. Acoustic ranging is a relatively new content in the field of mobile computing. DeepRange [13] uses a depth neural network to estimate the distance. [14] proposes an improved TOA estimation method to maintain high ranging accuracy and robustness in the closed environment of reverberation room. [15] and DopEnc [16] calculate the distance by measuring the time between the initial pulse of the smartphone and its reflection.

Inertial Tracking. Indoor tracking has a lot of related researches, mainly through image and inertial sensor. Walkie-Markie [17] uses WiFi tags and inertial sensor data to build user's trajectories. Zee [18] uses inertial sensors and WiFi signals for tracking. VeTrack [19] uses the inertial sensor of the mobile phone to track the position of the vehicle in real time. [20] tracks the user by gradually integrating WiFi interface and inertial sensors of smartphones. EasiTrack [21] uses the RF signal to accurately infer the moving distance of the target to achieve tracking. DeepIT [22] achieves higher precision tracking by evaluating the reliability of inertial data and synthesizing its data opportunitically. Imulet [23] adopts machine learning method to reduce the error of inertial data and improve the accuracy of tracking.

6 Conclusion

In this paper, we propose BatMapper-Plus to construct multi-level indoor floor plans without heavy human efforts and privacy/copyright concerns. Our unilateral ranging technique eliminates the limitation of existing bilateral ranging and achieves accurate distance measurements with less than 2 cm errors. We also refine the reconstructed map with line segments to replace the scattered points, and identify connection areas among different floors. We build a prototype and conduct experiments in two buildings, and the results have shown our effectiveness.

Acknowledgments. This work was supported in part by the Fundamental Research Funds for the Central Universities 2021JBM029, NSFC under Grant 62072029 and Grant 61872027, Beijing NSF Grant L192004, DiDi Research Collaboration Plan, and OPPO Research Fund.

References

1. Wang, X., Marcotte, R.J., Olson, E.: Glfp: Global localization from a floor plan. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1627–1632 (2019)
2. Anderson, R., Curro, J.: Indoor navigation using convolutional neural networks and floor plans. In: Proceedings of ION GNSS+, pp. 2133–2150 (2021)
3. Harithas, S.S., Pardia, B.: Gennav: a generic indoor navigation system for mobile robots. In: Proceedings of IEEE I-SMAC, pp. 182–187 (2020)
4. Gao, R., et al.: Multi-story indoor floor plan reconstruction via mobile crowdsensing. IEEE Trans. Mob. Comput. **15**(6), 1427–1442 (2016)
5. Zhou, B., Elbadry, M., Gao, R., Ye, F.: Batmapper: acoustic sensing based indoor floor plan construction using smartphones. In: Proceedings of ACM MobiSys, pp. 42–55 (2017)
6. Pradhan, S., Baig, G., Mao, W., Qiu, L., Chen, G., Yang, B.: Smartphone-based acoustic indoor space mapping. Proc. ACM on Interact. Mobile Wearable Ubiquit. Technol. **2**(2), 1–26 (2018)
7. Gao, R., et al.: Glow in the dark: smartphone inertial odometry for vehicle tracking in gps blocked environments. IEEE Internet Things J. **8**(16), 12955–12967 (2021)
8. Alzantot, M., Youssef, M.: Crowdinside: automatic construction of indoor floorplans. In: Proceedings of GIS, pp. 99–108 (2012)
9. Peng, Z., Gao, S., Xiao, B., Wei, G., Guo, S., Yang, Y.: Indoor floor plan construction through sensing data collected from smartphones. IEEE Internet Things J. **5**(6), 4351–4364 (2018)
10. Luo, H., Zhao, F., Jiang, M., Ma, H., Zhang, Y.: Constructing an indoor floor plan using crowdsourcing based on magnetic fingerprinting. Sensors **17**(11), 2678 (2017)
11. Chen, S., Li, M., Ren, K., Fu, X., Qiao, C.: Rise of the indoor crowd: reconstruction of building interior view via mobile crowdsourcing. In: Proceedings of ACM SenSys, pp. 59–71 (2015)
12. Philipp, D., et al.: Mapgenie: grammar-enhanced indoor map construction from crowd-sourced data. In: Proceedings of IEEE PerCom, pp. 139–147. IEEE (2014)
13. Mao, W., Sun, W., Wang, M., Qiu, L.: Deeprange: acoustic ranging via deep learning. Proceed. ACM Interact. Mobile Wearable Ubiquit. Technol. **4**(4), 1–23 (2020)
14. Liu, Z., Chen, R., Feng Ye, F., Guo, G., Li, Z., Qian, L.: Improved toa estimation method for acoustic ranging in a reverberant environment. IEEE Sens. J. **2**, 4844–4852 (2020)
15. Graham, D., Simmons, G., Nguyen, D.T., Zhou. G.: A software-based sonar ranging sensor for smart phones. IEEE Internet Things J. **2**(6), 479–489, 2015
16. Zhang, H., Du, W., Zhou, P., Li, M., Mohapatra. P.: Dopenc: acoustic-based encounter profiling using smartphones. In: Proceedings of ACM MobiCom, pp. 294–307 (2016)
17. Shen, G., Chen, Z., Zhang, P., Moscibroda, T., Zhang, Y.: {Walkie-Markie}: indoor pathway mapping made easy. In: Proceedings of USENIX NSDI, pp. 85–98 (2013)

18. Rai, A., Chintalapudi, K.K., Padmanabhan, V.N., Sen, R.: Zee: zero-effort crowdsourcing for indoor localization. In: Proceedings of ACM MobiCom, pp. 293–304 (2012)
19. Zhao, M., Ye, T., Gao, R., Ye, F., Wang, Y., Luo, G.: Vetrack: real time vehicle tracking in uninstrumented indoor environments. In: Proceedings of ACM SenSys, pp. 99–112 (2015)
20. Martínez del Horno, M., Orozco-Barbosa, L., García-Varea, I.: A smartphone-based multimodal indoor tracking system. Inf. Fusion **76**, 36–45 (2021)
21. Wu, C., Zhang, F., Wang, B., Liu, K.J.R.: Easitrack: decimeter-level indoor tracking with graph-based particle filtering. IEEE Internet Things J. **7**(3), 2397–2411, 2019
22. Gong, J., Zhang, X., Yuanjun Huang, J., Zhang, R.Y.: Robust inertial motion tracking through deep sensor fusion across smart earbuds and smartphone. Proc. ACM Interact. Mobile Wearable Ubiquit. Technol. **5**(2), 1–26 (2021)
23. Alloulah M., Tuominen, L.: Imulet: a cloudlet for inertial tracking. In: Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications, pp. 50–56 (2021)