Information Fusion 89 (2023) 166–188

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Survey of Landmark-based Indoor Positioning Technologies

Beakcheol Jang [a], Hyunjung Kim [b], Jong wook Kim [2,*]

[a] Graduate School of Information, Yonsei University, Seoul 03722, South Korea
[b] Department of Computer Science, Sangmyung University, Seoul 03016, South Korea

ARTICLE INFO

Keywords:
Indoor positioning
Landmark
Wireless signal
Sensor data
Visual image
Multisource-based technologies

ABSTRACT

Owing to the increase in the time people spend indoors, coupled with the pervasiveness of high-performance smart devices, the importance of indoor positioning techniques has grown. Researchers have extensively studied indoor positioning techniques using wireless signals or mobile device built-in sensors because satellite signals are absent in indoor environments. However, the built-in sensor readings tend to be distorted by the surrounding indoor environments. Furthermore, wireless signal-based technology incurs substantial cost because of presurvey and maintenance. To address these drawbacks, landmark-based indoor positioning technologies have recently been developed. An indoor landmark is a unique point in a room that is distinguishable from other points based on its features. Because the indoor landmark itself underscores the unique features of a specific region, it acts as a reference point to help users navigate efficiently. In this paper, we review recent landmark-based indoor positioning technologies. We categorize the technologies according to whether they use wireless signals, built-in sensors, images, and multisource-based technologies. We analyze them based on six evaluation criteria, namely, accuracy, core technology, detection difficulty, cost, versatility, and robustness. Finally, we discuss future directions of research and applications on indoor landmark positioning techniques. We believe that this review provides useful viewpoints and necessary information regarding indoor positioning technology using landmarks.

1. Introduction

Today, people spend most of their time indoors. Owing to the pervasiveness of high-performance mobile devices coupled with recent advances in the development of sensor technologies, indoor positioning technologies have attracted considerable research attention. Various services and applications based on the indoor location of users have been developed and extensively used, and the importance of indoor positioning technologies is increasing [1,2].

The main infrastructure used for positioning are satellites positioned in near space. Satellite signals can be detected in almost all parts of the globe, and the Global Positioning System (GPS) technology [3,4] exploiting them achieves a maximum error rate of 3.5 m and 95% reliability, even in the worst case. However, satellite signals used by GPS are reflected or diffracted by obstacles, such as building walls, and are only available in outdoor environments; thus, using them in indoor positioning technology is difficult [5]. For this reason, indoor positioning technology must be based on alternative methods other than GPS and satellite signals [6].

Mobile devices, specifically their built-in sensors, can be efficient tools for acquiring and processing the vast amount of data needed for indoor positioning [7,8]. With the advances in technologies, recently developed mobile devices incorporate various sensors such as acceleration, gyroscope, magnetic field, air pressure, and temperature sensors. These built-in sensors make it possible to collect various data associated with the surrounding environment, which the indoor positioning system uses to track the movement of the user or the characteristics of the surrounding environment. Among these, the most frequently used are magnetic field sensors and inertial measurement unit (IMU) sensors [9,10], including acceleration and gyroscope sensors. The device continuously records the user's location by monitoring the distance covered and the direction in real time using the corresponding sensors. IMU and magnetic field sensors are simple and intuitive; therefore, it is easy to design their algorithms and software. However, these sensors are not suitable for long-term indoor positioning because they are heavily influenced by the surrounding environment. Owing to the influence of the indoor environment, the sensor data gradually changes over time or an error occurs in which the values change depending on the posture of

* Corresponding author.
E-mail addresses: bjang@yonsei.ac.kr (B. Jang), kjyy28@naver.com (H. Kim), jkim@smu.ac.kr (J. Kim).

https://doi.org/10.1016/j.inffus.2022.08.013
Received 7 September 2021; Received in revised form 5 February 2022; Accepted 11 August 2022
Available online 13 August 2022

the user [11]. The continuous accumulation of errors reduces the sensor's accuracy and effectiveness for long-term indoor positioning. [12]

Other commonly used indoor positioning technologies are wireless signal or visual image-based fingerprinting, which consists of an offline training phase and an online estimation phase [12]. In the former, the indoor space of interest is divided into cells, and the received signal strength indicator (RSSI) values of the wireless signals [13] or visual images taken with a camera [14] for each cell are collected. Subsequently, a fingerprinting map is constructed. In the online estimation phase, users detect their locations based on a map. However, whenever the interior features such as walls or even furniture are changed or access points (APs) are added, modified, or removed, the wireless signals or visual images must be changed and the fingerprint map must be completely recreated. Hence, indoor positioning technologies without such limitations are required.

To address and compensate for this drawback, the powerful concept of an indoor landmark has been adopted in indoor positioning technology [15,16]. The indoor landmark is a unique and rarely changing point in the indoor space that is distinguishable from other points by its features. The indoor positioning system stores data related to that specific point as an indoor landmark, based on which the system provides information about the location of a moving user.

The advantage of the indoor landmark technique is its robustness. For example, sensor-based positioning technology can be heavily influenced by its surroundings, so large errors can occur and those errors can accumulate. Since the landmark itself contains unique data on a specific region, it acts as a reference point and can be used to correct these errors. In the fingerprinting technique, whenever the interior features are changed, the wireless signals or visual images must be changed, and the fingerprint map must be completely recreated. The landmark technique is very robust because it uses the point with little change as the landmark, and even if some landmarks do not work, if other landmarks work correctly, positioning can be carried out using the correct landmarks. For these reasons, the landmark technique is usefully used in the environment where the indoor environment frequently changes, such as a shopping mall.

The landmark technique reduces the cost of measurement at the location of interest too. For example, the fingerprinting technique should collect data as precisely as possible at the place where a user's movement is expected to occur. In the landmark technique, only points that show a property that do not change easily in the place where positioning is to be carried out are set as landmarks. Since the amount of data to be measured and stored is relatively small, the cost of the measurement can be very reduced.

Using the basic principle of the landmark algorithm, various techniques for calculating the user's real-time movement via sensors, wireless signals, or the camera of a mobile device and correcting the error accumulation of the data with a landmark have been studied [17,18].

The purpose of this study is to present a timely, valuable, and detailed review and analysis of indoor positioning technologies that utilize indoor landmarks. We categorize them based on the data they exploit, namely, wireless signals [19,20], sensors [21,22], visual images [23,24], and multisource-based [25,26] indoor positioning technologies and described their algorithms, characteristics, advantages, and disadvantages. Fig. 1 illustrates the taxonomy of landmark-based indoor positioning techniques using landmarks.

One of the basic elements that can be selected as a landmark is the wireless signal. Wireless signals are easily detectable in indoor environments and the required infrastructure installation costs are low. In addition, there are various types of wireless signals such as Wireless-Fidelity (Wi-Fi) [27], Bluetooth [28], and millimeter wave (mmWave) [29]. Further, these individual signals are distinct; therefore, multiple signals can be combined. Recently, research on indoor location tracking technology using elements such as Bluetooth Low Energy (BLE) is being actively conducted [30–32]. Indoor positioning based on magnetic fingerprints [33], Walkie-Markie [34], and CiFi [35] use wireless signals to select landmarks.

Although sensor data are not very effective for long-term positioning, they are a good source of data for locating users [36]. The indoor positioning system can determine the structure of the room through the user's movement—calculated using the IMU sensor and the magnetic field sensor—and select it as a landmark. In addition to simple movements, like bats, the system can also exploit reflected waves using installed acoustic sensors [37]. Currently, owing to further advances in

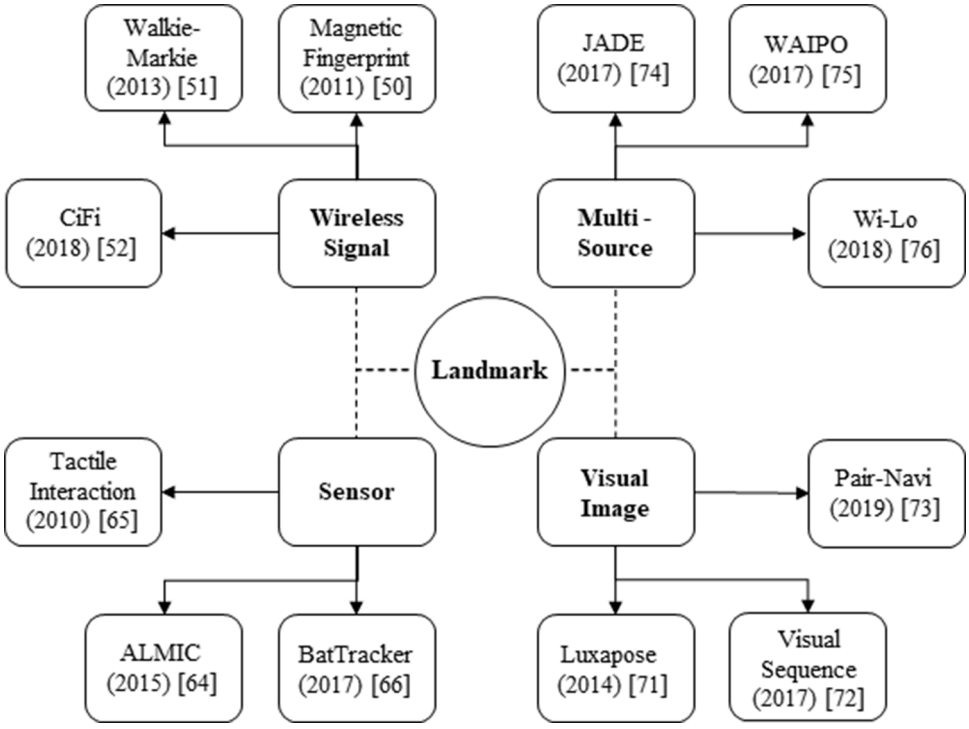

Fig. 1. Taxonomy of indoor positioning techniques using landmarks.

technology, there are systems that select landmarks using tactile data collected from visually impaired people [38]. Indoor positioning technologies such as activity landmark-based indoor mapping via crowdsourcing (ALIMC) [39], indoor positioning with tactile interaction [40], and BatTracker [41] use sensor data to select the landmarks.

One of the most natural ways for users to identify their location is through visual observation. Hence, using image matching technology [42], an algorithm that captures the surrounding images and matches them with prestored images to determine the location has been proposed. Luxapose [43] performs indoor positioning using a visual sequence [44], and Pair-Navi [45] use images to select landmarks.

The feasibility of complementing the shortcomings of each technology by combining different technologies has also been actively studied in the field of indoor positioning. The researchers did not merely integrate sources with similar properties, such as Wi-Fi and mmWave; they attempted to improve the performance of indoor positioning technology by combining various elements such as photo and built-in sensors. Joint anchor and device location estimation (JADE) [46], Wi-Fi and magnetic fingerprints, image matching and people co-occurrence (WAIPO) [47], and wireless indoor localization through multisource radio fingerprinting (Wi-Lo) [48] use complex elements.

We analyzed the performance of landmark-based indoor positioning technologies based on six criteria, namely, accuracy, core technology, detection difficulty, cost, versatility, and robustness. We suggested future research directions and then concluded the paper.

The remainder of this study is organized as follows. Section 2 presents basic algorithms and their problems of indoor positioning technologies that use wireless signals, sensor data, and visual images respectively. We also present recent surveys related to indoor positioning technologies. Section 3 covers the systems that use wireless signals as landmarks, and Section 4 presents the systems that use sensor data as landmarks. Section 5 describes the systems that use landmarks obtained using visual images, and Section 6 describes the systems that use multiple sources. Section 7 compares the technologies covered according to the selected criteria. Section 8 covers future research directions and applications, and in Section 9 we conclude.

2. Background Knowledge and Related works

2.1. Background knowledge

In this section, we explore indoor positioning algorithms that employ wireless signals, sensor data, and visual images. Specifically, we describe the basic algorithms used by these systems, as well as their problems.

Most wireless signal-based indoor positioning technologies use fingerprinting technology that divides the indoor space into cells of a certain size and stores the RSSI of the APs in the cell as a fingerprint. Subsequently, when a moving user reaches the corresponding point, the system compares the wireless signal value detected by the user's device with the previously stored fingerprint. Through the estimation process, the system determines the cell whose fingerprint value was most similar to the wireless signal value detected by the device as the cell where the current user is located.

Fingerprinting technology can locate a user through a relatively simple and easy algorithm. Because the technology uses a wireless AP signal, such as Wi-Fi, the technology is very versatile and can be easily used in an indoor environment where the AP is installed. However, fingerprinting technology incurs a higher cost from investigating the RSSI of the APs to generate a fingerprint map. This is for the preliminary generation of the map, as well as its maintenance. If the AP that is the target of the fingerprint map is modified or lost, or a new AP is added, the system's administrators must update the fingerprint map for smooth operation [49]. Therefore, the operator must go to all the cells and repeat the process of finding and saving the values as during the initial creation. Specifically, owing to the nature of the indoor environment, which is highly variable, fingerprinting technology requires frequently updating the fingerprint map for accuracy; therefore, it is costly.

The mobile device can acquire the user's moving distance and moving direction data through the built-in IMU sensors. The moving distance is the distance that the user has covered either during one step or a certain period. The moving direction is the bearing direction in which the user's front faces at the corresponding position and it is calculated by aggregating the data from various sensors. The IMU sensors that are commonly used in calculating the direction of movement are the magnetic and gyroscope sensors. The magnetic sensor has a property similar to a compass. It detects the indoor magnetic field and indicates the direction that the device is facing by making the earth's pole as a reference [57]. The gyroscope sensor detects the rotation of the device and shows the angular velocity when the device rotates or changes its posture [58]. The device calculates the change in rotation angle $\Delta\theta_t^{gyro}$ from the angular velocity s_t^{gyro} at time t using the following equation.

$$\Delta\theta_t^{gyro} = \int_{t-1}^{t} s_t^{gyro} dt \tag{1}$$

By adding the change of the rotation angle obtained by the Eq. (1) to the existing angle, the gyroscope sensor calculates the direction of movement. Then, using the algorithm, the indoor positioning system calculates the user's direction using both the value obtained by the magnetic sensor and the value obtained by the gyroscope sensor.

When both the moving distance and the moving direction are calculated, the system leverages these two values to calculate the user's location. Assuming that the user's position is located in the plane coordinate system $(x,\ y)$, the user's current position value $(x_t,\ y_t)$ is calculated by substituting the existing position value $(x_{t-1},\ y_{t-1})$ into the following equation.

$$(x_t, y_t) = \begin{cases} x_{t-1} + \Delta x \\ y_{t-1} + \Delta y \end{cases} \tag{2}$$

In Eq. (2) since the change in distance Δx and Δy can be expressed by the moving distance d and the moving direction θ_t, the equation becomes;

$$(x_t, y_t) = \begin{cases} x_{t-1} + d * sin(\theta_t) \\ y_{t-1} + d * cos(\theta_t) \end{cases} \tag{3}$$

The algorithm for calculating the user's position using IMU sensors is relatively simple, and because IMU sensors are embedded in recent smart devices, they can be used with versatility. However, the calculation of the user position obtained using IMU sensors cannot be used long term because of the continuous accumulation of errors in the sensors.

The direction of movement, which is one of the two inputs for the position calculation algorithm, is calculated using a magnetic sensor and gyroscope sensor. The magnetic sensor maintains a similar value when the sensor value is measured in the same place, but it is greatly affected by indoor elements such as electronic devices or steel in the structures of buildings [50]. The gyroscope sensor updates the direction by adding the angle variation to the value of the previous direction. However, the device is sensitive enough to measure the angular velocity, even in granular periods [51]. Owing to the error accumulation problem, the positioning technology that uses only sensors cannot be used for a long time. It uses additional landmarks to overcome the shortcomings of these sensors or utilizes only the advantages of sensor data to create a landmark to proceed with positioning.

Humans perform self-localization by referring to the visual images of their surroundings and various locations saved in their memories. Researchers have incorporated localization through image matching of people into indoor localization. Image matching is a crucial aspect of positioning technology and is based on visual images. Images of special structures or places, such as statues, monuments, and plaques, that exist indoors are stored prior to positioning. Then, when the user moves indoors, the device captures the indoor environment using a built-in

camera. Subsequently, the system receives the captured image and video and matches the image data stored in the database. If there is an image with a matching rate that is higher than a set threshold among the stored images, the system selects the image and informs the user that he/she is currently in the location where the image was taken.

The visual image-based technology uses the structures in the indoor environment, thus eliminating the need to install additional infrastructure. In terms of data collection, preliminary investigations are easy because the techniques select and store only specific indoor locations without having to search every place individually, as done in fingerprinting. Unfortunately, in an indoor environment where the visual structure tends to change frequently owing to structural alterations, the visual data do not have a long life. This implies that the replacement cycle of the data is short, necessitating frequent data updates unlike in fingerprinting techniques based on the wireless signal. Consequently, significant time and cost are incurred in further investigating the dictionary data.

2.2. Related works

In many recent studies, indoor positioning technology have been studied and numerous indoor positioning technologies have been proposed. Researchers have conducted surveys based on which they analyzed and evaluated indoor positioning techniques according to their own specific criteria. Table I summarizes the findings of key surveys on indoor positioning technologies.

Suining He and S.-H. Gary Chan reviewed indoor positioning technologies based on Wi-Fi fingerprints [52]. The study revealed that Wi-Fi is a promising replacement for GPS signals. They focus on Wi-Fi fingerprinting, which has high applicability in complex indoor environments because it does not require line-of-sight (LoS) from an AP. The researchers explained the advanced Wi-Fi positioning technology and proposed several strategies to efficiently construct these technologies. They further introduced key advanced technologies for locating a user in an indoor space by classifying them in terms of time or space signal patterns, user collaboration, and motion sensors. The proposed techniques were also evaluated to determine the most efficient system deployment approach. Their study aimed at reducing offline labor-intensive survey, adapting to fingerprint changes, calibrating heterogeneous devices for signal collection, and achieving energy efficiency for smartphones as evaluation criteria.

Xiao et al. introduced indoor positioning technologies with focus on the device perspective [53]. They classified indoor positioning systems into device-based systems and device-free systems depending on the method used by the wireless device to interact with a target. In the device-based positioning system, a wireless device, such as a smartphone, is connected to a target. The devices calculated the position of the target through cooperation with other distributed wireless devices. In the device-free positioning system, because the target does not have a wireless device, the wireless infrastructure deployed in the indoor environment determines the position of the target by analyzing the effect of the target on the wireless signal. New technologies that integrate wireless and sensor functions using smartphones were introduced. Furthermore, the social context for device-based positioning were expanded and specific wireless functions were extracted to trigger a new positioning technology without relying on user centric devices. The recent developments in technologies that drive both systems by detailing the basic radio format, basic positioning principles, and data fusion techniques were introduced. In the study, the systems were comprehensively compared based on selected evaluation criteria, namely, accuracy, cost, scalability, and energy efficiency.

Yassin et al. introduced indoor positioning techniques with a focus on indoor methodology and concepts [54]. They focused on heterogeneous approaches that combine hybrid positioning for indoor positioning with different radio access technologies. They also introduced the limitations and challenges of an indoor positioning system, as well as

Table 1
Related survey papers

Study	Year	Scope	Evaluation item
Wi-Fi fingerprint-based indoor positioning: recent advances and comparisons [52]	2015	Indoor positioning techniques based on Wi-Fi fingerprint were reviewed. Advanced Wi-Fi fingerprint localization techniques and system deployments are explained.	- Cost - Robustness - Versatility - Energy efficiency
A survey on wireless indoor localization from the perspective of the device [53]	2016	Indoor positioning techniques from the perspective of the device were considered. Advanced technologies are explained by elaborating on the basic wireless modalities, basic localization principles, and data fusion techniques.	- Accuracy - Cost - Scalability - Energy efficiency
Recent advances in indoor localization: a survey on theoretical approaches and applications [54]	2016	Indoor positioning techniques, with an emphasis on indoor methodologies and concepts, were reviewed. Different localization-based applications, specifically where the location information is critical to estimation are considered.	- Accuracy - Cost - Complexity - Security - Scalability
Indoor positioning technologies without offline fingerprinting map: a survey [55]	2018	Indoor positioning techniques that do not exploit offline fingerprinting map were reviewed. These techniques are classified according to whether they use simultaneous localization and mapping (SLAM), extrapolation and interpolation, or crowdsourcing.	- Accuracy - Calculation time - Versatility - Robustness - Security - Participation
A survey of indoor localization systems and technologies [56]	2019	Indoor positioning techniques based on the angle of arrival (AoA), time of flight (ToF), return time of flight (RTOF), Wi-Fi, ultrawideband (UWB), and Bluetooth were reviewed. Furthermore, the localization and positioning of human users and their devices are discussed.	- Energy efficiency - Availability - Cost - Reception range - Latency - Scalability - Tracking accuracy
A survey on fusion-based indoor positioning [62]	2019	Fusion-based indoor positioning technologies and systems to derive the latest technologies within the integrated fusion-based positioning framework composed of three convergence characteristics of source, algorithm, and weight space were reviewed.	- Source - Algorithm - Weight space
Self-calibration and collaborative localization for UWB positioning systems: a survey and future research directions [63]	2021	Important studies done on self-correction and collaborative localization are reviewed. Their classification and analysis provide the basis for further studies on self-calibration and cooperative positioning in the deployment of UWB	- Node type - Topology - Type of algorithm - Sequential/ Concurrent - UWB specific - Experimental evaluation

(continued on next page)

Table 1 (*continued*)

Study	Year	Scope	Evaluation item
		indoor positioning systems.	
A survey on indoor vehicle localization through RFID technology [64]	2021	The latest analysis of methods suitable for indoor positioning and utilization of vehicles using RFID technology were reviewed. The survey covers three main categories of vehicle localization systems. These include a solution using only RFID technology, a sensor fusion technology that combines data from an RFID system and proprioceptive sensors, and a sensor fusion technology that combines RFID data with other in vitro soluble sensors in addition to the RFID system itself.	- Cost - Energy Consumption - Complexity

the latest technologies and concepts used to improve positioning through an application. They reviewed the latest positioning systems for indoor and outdoor environments. Meanwhile, they discussed recent methodologies, such as data fusion and collaboration techniques, to improve the accuracy of positioning. In addition, in this paper, they present an overview of the machine learning technique recently adopted for positioning. Various positioning-based applications in the field were discussed and evaluated. In this study, accuracy, cost, complexity, security, and scalability were selected as the evaluation criteria.

Jang and Kim introduced indoor positioning technologies that did not use an offline fingerprint map [55]. They highlighted the problems of existing wireless signal fingerprinting technology. The WSN technology divides the room into cells of a predetermined size before it is used to store all the wireless cell signals. The stored radio signals are called fingerprint maps. The authors noted that the fingerprint map has a short effective period owing to the highly variable nature of the indoor environment, and rewriting the map is time consuming and labor intensive. The findings of the study proved that some indoor positioning technologies do not adopt or introduce offline fingerprint maps. These researchers analyze indoor positioning technologies that do not use offline fingerprint maps and classified them based on whether simultaneous localization and mapping (SLAM), extrapolation and interpolation, and crowdsourcing were used. The evaluation criteria in the study were accuracy, calculation time, versatility, robustness, security, and participation.

Faheem Zafari et. al introduced user and device positioning techniques [56]. The authors proposed that recent surveys introducing an accurate and reliable indoor positioning systems cannot fully realize the extensive indoor positioning services that can be provided by utilizing the increasingly available Internet of Things (IoT) [57] and ubiquitous connections [58]. In this paper, the key aspects of indoor positioning, such as the angle of arrival (AoA) [59], ToF [60], and return time of flight, are discussed based on technologies such as Wi-Fi, ultra--wideband (UWB) [61], and Bluetooth. The strengths of the proposed system are highlighted in the study. Furthermore, in these papers, several positioning systems are compared and their modes of operation were briefly described; intrinsic technologies or techniques are not compared. The evaluation criteria were energy efficiency, availability, cost, reception range, latency, scalability, and tracking accuracy.

Guo, Xiansheng, et al. presented a survey on fusion-based indoor positioning [62]. Unlike other paper, this survey paper summarizes and analyzes the existing fusion-based positioning systems and technologies with three characteristics. On the other hand, lessons, tasks, and discussions on countermeasures are also presented. This study focused on helping researchers acquire a clear concept of indoor fusion-based positioning systems and technologies and gain insights to further develop other advanced fusion-based positioning systems and technologies in the future. The evaluation criteria were source, algorithm, and weight space.

Ridolfi et al. presented a survey and future research directions of self-calibration and collaborative localization for UWB positioning systems [63]. They proposed a classification that takes these aspects into account and other aspects such as the way information is distributed in the network and the types of algorithms employed to place the nodes. The main contributions of survey were as follows. The first contribution was that they rigorously investigated solutions for the problems of self-correction and collaborative localization. They established that these are not the same by highlighting the differences and similarities. The second most focused area is UWB, which has achieved high positioning accuracy due to its essential characteristics. It also represented a general tool for designers of any indoor positioning system that may face similar problems. A third contribution was the direction of future UWB research. The evaluation criteria were node type, topology, type of algorithm, sequential/concurrent, UWB specific, and experimental evaluation.

Motroni et al. introduced a survey on indoor vehicle localization through Radio-Frequency Identification (RFID) technology [64]. The survey focused on achievable localization performance and also considered the cost and complexity of infrastructure deployment and maintenance overhead. Positioning, tracking, navigation and SLAM issues were considered. This article highlights the pros and cons of each method, along with the main challenges and perspectives of RFID-based solutions for vehicle localization. The evaluation criteria were cost, energy consumption, and complexity.

Although important aspects of the indoor positioning technology have been examined in related studies, landmark-based technologies have not been examined. Therefore, to address this research gap, we focused on landmark-based indoor positioning technologies and selected the most suitable evaluation criteria to assess their robustness.

3. Wireless Signal

In the field of indoor positioning, wireless signals are the most efficient tools for extracting indoor landmarks. The wireless signal is altered and conditioned by various factors such as the distance between the AP and the user, objects in the environment, as well as the presence of other electronic devices in the room. These factors offer diverse and unique patterns of wireless signals, which discriminating the various regions of the area. The indoor positioning technologies described hereafter differentiate landmarks based on the unique characteristics of these wireless signals.

3.1. Walkie-Markie

The absolute value of a wireless signal can be easily changed by minor factors. Even in the same place, the absolute value of a wireless signal varies depending on the people present, electronic devices deployed, and time of the day. Therefore, when an absolute value is used, the measured value of the wireless signal may not match the actual placement of the landmark. To solve this, Shen et al. proposed the Walkie-Markie system, which detects landmarks according to the increase or decrease in the Wi-Fi intensity.

As shown in Fig. 2, the architecture of the Walkie-Markie system consists of the mobile client and the backend service located in the server cloud. When the user starts moving, the mobile client collects the regular Wi-Fi scan values and the user's approximate trajectory data, such as the number of steps, step period, and movement direction. The mobile client generates Wi-Fi marks using the collected results and

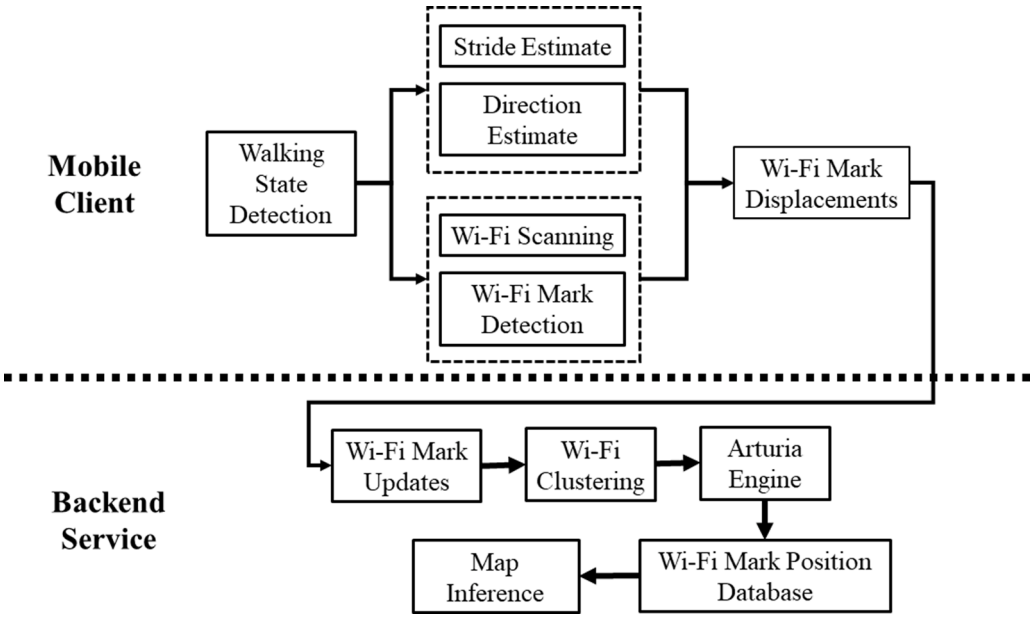

Fig. 2. System architecture of Walkie-Markie

displays them on a 2D map. The backend service builds a route map using the collected Wi-Fi marks and user trajectory data.

The Walkie-Markie uses the received signal strength trend tipping point (RTTP), the point when the received signal strength (RSS), that is, the measurement of the power present in a received radio signal, of a Wi-Fi AP switches from an increasing trend to decreasing while in an indoor environment as a landmark called Wi-Fi mark. However, many Wi-Fi marks are generated from one master AP during the collection of Wi-Fi marks using the RTTP approach. To tackle this problem, the system assigns three attributes of BSSID, (D1, D2), and N to the landmark. The BSSID is an ID of the master AP that satisfies the RTTP condition, and (D1, D2) are the approach and exit directions of the corresponding RTTP, respectively. N is the set of information regarding neighboring APs in the RTTP.

The system assigns the ID of each master AP to the detected RTTPs as a BSSID. It identifies the path and rotation of the RTTP using its (D1, D2). The system eliminates incorrect RTTPs from incorrect operation arising from situations such as the user making a u-turn. Then, for each RTTP showing similar D1 and D2, each N was collected. N includes the IDs of neighboring APs and the RSSI difference between the AP and master AP. Once all the attributes have been collected, the RTTP is stored as a Wi-Fi mark.

The system updates the Wi-Fi marks by sending the newly collected Wi-Fi mark to the backend service. As mentioned earlier, because several Wi-Fi marks can be created by multiple paths from one AP, the server storing the Wi-Fi mark uses a clustering algorithm to fuse several Wi-Fi marks related to a certain AP into a single Wi-Fi mark. The main challenge of the Walkie-Markie system is addressing the inaccuracy of IMU-based tracking measurements, such as the stride length or direction error, using the Wi-Fi mark. To correct the error in the IMU tracking, the Walkie-Markie backend service utilizes an Arturia algorithm [65] to remove noise from the measurement and assign optimal coordinates to the Wi-Fi mark. After the measurement noise problem is addressed and the coordinate is assigned using the Arturia algorithm, the system acquires and places a Wi-Fi mark in the 2D plane using a graph insertion algorithm and the path of the client device.

3.2. Geomagnetic fingerprinting

Even if there is a risk of error, the absolute value of a wireless signal is fundamental to indoor positioning. This type of signal RSSI is easy to collect and can also be useful when using wireless signals from different infrastructures. There are various types of indoor wireless signals; one of the commonest methods to utilize them was proposed by Jaewoo Chung et al. It consists of a novel positioning technology that uses the indoor geomagnetic field as a landmark.

The data generated by the geomagnetic field can be used in most buildings without the need for additional infrastructure. Owing to building steel, concrete frames, and electronic devices located indoors, the geomagnetic field inside the building is distorted, pointing to its unique strength and direction in each location. Accordingly, the system builds a landmark map according to the relative distortions of the indoor geomagnetic field.

An electronic compass (HMC6343 – 3-axis magnetic field sensor [66]) was used in the experiment with an internal tilt correction algorithm to measure the indoor geomagnetic field for map construction. With the compass positioned 1 m above the floor, the experiments confirmed that the sensor head changed every 60 cm along a 25 m corridor. While measuring the changes in headings, a step-sensing motor was used to rotate the sensor, drawing a 5 cm diameter circle per revolution every 100 steps. The measured heading direction was then compared to the actual direction.

As a result of the measurements, the deviation in the heading direction varied significantly, depending on the relative position along the corridor. The deviation errors were assumed to be related to the building structures near the location. In addition, the changes in the four directions were irregular. Therefore, deviations occurred on both large and small scales; otherwise, variable parallel lines could be observed.

To investigate the feasibility of using fingerprints to track user's location, a system was established to collect two datasets: one for creating geomagnetic field fingerprint maps and the other for testing the location influence. For the fingerprint mapping, the data were measured and collected in the experiment using the previous data collection methods. During this measurement stage, the geomagnetic field was confirmed to be a three-dimensional vector. The system identified and indexed the location and direction of the collected data. Through the above survey, a total of 6000 data features, comprising 60 location points and 100 locations, were collected for the study.

During the experiment, the location of the user was calculated based on the difference between the minimum root mean squares (RMS) [67] of the individual location d of the test dataset and the map dataset E. When the map dataset E and target point fingerprint d were given, the

d nearest neighbor was calculated as follows:

$$\forall d'' \in E,\ |d \leftrightarrow d'| \leq |d \leftrightarrow d''|, |d \leftrightarrow d'| = \sqrt{\sum_{i=1}^{k} (d_i \leftrightarrow d_i')^2} \quad (4)$$

where d_i is the i^t characteristic component of d (e.g., m). Subsequently, to measure the accuracy of this method, the system used the difference between the location index of the test dataset and that of the recently accessed neighborhood, and $d`$ was used in the map dataset E.

To ensure that the distances between the indices of the components were equal, the difference in the RMS was calculated using eight different m combinations in the subset $d_k = \{m_1 \ldots m_k\}$ of d with a common denominator k. In this analysis, all the combinations of map datasets corresponding to d_k were calculated to determine the correlation between the number of features and accuracy.

Most of the lowest RMS differences were distributed along the lines where x and y matched. However, there were some anomalies. Because the reading of geomagnetic fields depends on location and direction, the algorithm can predict direction using fingerprint maps.

3.3. CiFi

The recently developed open-source device driver for a Wi-Fi network interface card (NIC) [68] provides an interface that extracts channel state information (CSI) [69] from each packet received by the device. The CSI contains detailed channel information, including the channel measurements of the subcarrier level of the orthogonal frequency division multiplexing system. Simultaneously, the CSI can capture multipath effects. The AoA value estimated using the phase difference data of Wi-Fi is relatively more stable than the amplitude data because of the stability of the phase difference data. Based on the stability of the CSI and estimated AoA, Wang et al. proposed CiFi, an indoor location tracking technology using the CSI of commercial 5 GHz Wi-Fi and deep convolutional neural network (DCNN) for learning [70].

Fig. 3 shows the system architecture of CiFi. CiFi consists of an offline learning phase and an online location estimation phase. In the offline phase, CiFi first obtains the CSI data from three antennas for every packet received from the modified Intel 5300 firmware and extracts the phase information. Thereafter, two sets of CSI data, including a phase difference, are calculated from Antennas 1 and 2 and Antennas 2 and 3.

When the CSI phase difference data are calculated, CiFi generates a

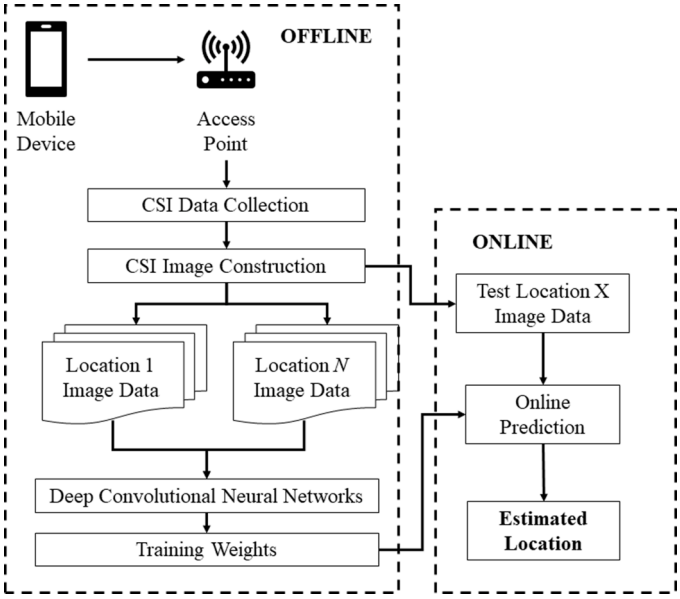

Fig. 3. System architecture of CiFi

60 × 60 AoA image for the DCNN input. For image generation, CiFi estimates the AoA value of each received packet using the following equation:

$$\theta_i = \arccos(\Delta \angle \widehat{H}_i \lambda / (2\pi d)) \quad (5)$$

where $\angle \widehat{H}_i$ is the phase difference measured in subcarrier i, d is the distance between two adjacent antennas, and λ is the wavelength. Thereafter, CiFi generates 60 × 60 AoA images. After image generation, the CSI trains a DCNN consisting of a convolutional layer, subsampling layer, and a fully connected layer using all the images generated at all training positions. The image is ideal for the CiFi DCNN to process data in the convolutional and subsampling layers. The convolution layer obtains a functional map and extracts the temporal and spatial functions of the AoA image. The subsampling layer implements an average pooling function to reduce the training time. For each input image of the first convolutional layer and subsampling layer, CiFi uses a 5 × 5 convolution filter to obtain the same number of 56 × 56 feature maps. To reduce the training data and ensure the invariability of the feature map, CiFi subsamples the image using a 2 × 2 size to obtain the same number of 28 × 28 feature maps. Thereafter, another three convolution and subsampling layers can be implemented to obtain 16 1 × 1 feature maps. Finally, CiFi combines the labels of the training data that can be used to update the training weights, such as a BP algorithm and a convolutional filter, using a loss function after obtaining the forward output result.

In the online stage, a probabilistic method for predicting the location of a mobile device is used based on the learned DCNN and the newly received CSI AoA image from the mobile device.

When M represents the number of images at one location and o_{ij} is the predicted output of the DCNN for location i using image j, CiFi is the output of the DCNN for K training locations using M images. The matrix O is obtained as follows:

$$O = \begin{bmatrix} o_{11} & o_{12} & o_{13} & \cdots & o_{1M} \\ o_{21} & o_{22} & o_{23} & \cdots & o_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ o_{K1} & o_{K2} & o_{K3} & \cdots & o_{KM} \end{bmatrix} \quad (6)$$

Using matrix O, CiFi selects R candidate positions and calculates the weighted average of the selected positions as the expected positions of the mobile device using the greedy method. First, CiFi selects the largest R position indices among the DCNN output values in all the columns of matrix O and creates a new R × M matrix S, as follows:

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1j} & \cdots & s_{1M} \\ s_{21} & s_{22} & \cdots & s_{2j} & \cdots & s_{2M} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ s_{R1} & s_{R2} & \cdots & s_{Rj} & \cdots & s_{RM} \end{bmatrix} \quad (7)$$

where s_{ij} is the position index of the i^{th} largest output in image j. All the elements of matrix S have position indexes belonging to {1,2, …, K}. The R largest position indices are obtained by calculating the frequencies of all the position indices of matrix S. In addition, the weight of the position i index can be calculated as the average of all the selected outputs for the position i index indicated by p_i. CiFi estimates the location of the mobile device as a weighted average of R selected locations using the following equation:

$$\widehat{L} = \sum_{i=1}^{R} l_i \times \frac{p_i}{\sum_{i=1}^{R} p_i} \quad (8)$$

where l_i is the i^{th} training position.

4. Sensor Data

Owing to the error accumulation problem, the positioning technology that uses only sensors cannot be used for a long time. To overcome the shortcomings of sensors, technologies use additional landmarks or utilizes only the advantages of sensor data to create unique landmarks to proceed with positioning. The sensor data can be used not only to track a user's real-time location but also to detect landmarks. As the number of sensors built into the newly developed devices increases, the indoor positioning system can leverage various sensors, such as air pressure, temperature, sound sensors, and basic IMU sensors for the detection of landmarks. Various technologies that use the data sourced from these sensors to track and map the landmarks are discussed in this section.

4.1. Activity landmark-based indoor mapping

The acceleration and gyroscope sensors are typical motion sensors built into mobile devices. Motion sensors embedded in mobile devices detect diverse pattern changes that differentiate the movements that occur on flat surfaces from those that occur on irregularly shaped structures in an indoor environment such as elevators, stairs, and corners [71]. Utilizing the characteristics captured by motion sensors, Zhou et al. proposed the activity landmark-based indoor mapping (ALIMC), a landmark-based indoor mapping algorithm that leverages the variations recorded by motion sensors based on indoor movements.

The ALIMC uses motion sensor data generated by indoor movements to convert indoor structures into links-node models, as shown in Fig. 4. The link is the user's path and the node is an indoor point of interest. During the creation of the links-node network, the ALIMC continuously collects motion sensor data from the mobile device and records the point as a link if no particular pattern is detected in the data. When the device moves into certain indoor structures, such as corners, stairs, escalators, and elevators, the ALIMC detects the patterns caused by movement and captures their difference from the patterns recorded during the movement on a flat surface. The ALIMC analyzes the detected pattern and selects a specific structure with a similar pattern, determines it as the point of interest, and extracts it as an activity landmark. The ALIMC identifies the activity landmark L using the following six attributes:

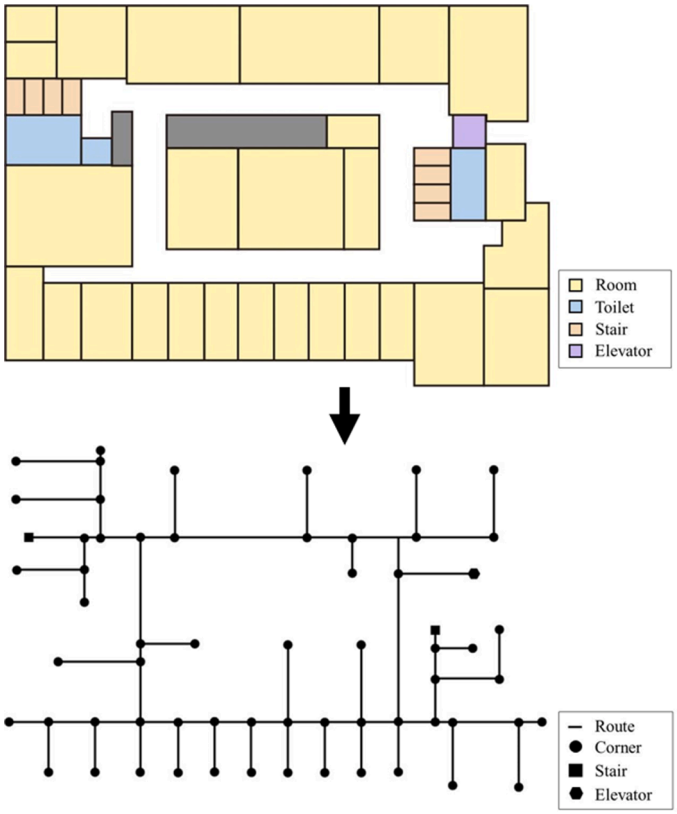

Fig. 4. Link-node map composition using indoor environment structure

$$L \underline{\underline{\Delta}} \{ID,\ type,\ F,\ \Delta H, (D_p, D_f), (ID_p, ID_f)\} \tag{9}$$

ID is the identification number of the landmark; type is the type of activity by the landmark, including rotation, elevator boarding, and movement up and down the stairs; F is the Wi-Fi fingerprint collected when activity is detected; ΔH is the heading change at the corresponding activity point; D_p is the distance between the current activity landmark and the previous landmark; D_f is the distance between the current activity landmark and the next landmark; ID_p and ID_f are the IDs of the previous and next activity landmarks, respectively.

After sufficient landmarks are extracted, the ALIMC system uses Wi-Fi to categorize the collected activity landmarks into clusters based on the distance between them. The main problem in using the RSS differences in Wi-Fi is that different types of smartphones may exhibit different RSS values owing to their internal designs. To address this, the system utilizes RSS sequence. The RSS sequence of sets of APs measured from the same location by different devices is theoretically the same. The correlation coefficients are used to evaluate similarities in this RSS sequence and the correlation coefficients between these RSSs are extended to Wi-Fi to distinguish between different activity landmarks.

A high correlation coefficient suggests that the two fingerprints are extracted from the same location. However, if the indoor APs are few, the correlation coefficient may be high, even if two fingerprints are extracted from different locations. To address this, the system uses the Jaccard similarity coefficient [72] as another metric for the activity landmark cluster. The Jaccard coefficient is a statistical metric used to compare the similarity and diversity of the sample sets.

Assume that f_i and f_j are the Wi-Fi fingerprints of activity landmarks, i and j.

First, the system calculates the intersection MAC_{int} and union MAC_{uni} of MAC_i and MAC_j which are the MAC sets of f_i and f_j.

The Jaccard factor of f_i and f_j is calculated using the following expression.

$$Jac_{ij} = num(MAC_{int}) / num(MAC_{uni}) \tag{10}$$

where $num()$ is the number of MACs in the set. To calculate f_i' and f_j', which are the correlation coefficients of f_i and f_j, the RSS of each AP existing in the intersection MAC_{int} of f_i and f_j must be extracted first.

By sorting the RSS in descending order, according to the RSS of f_i', the order index ix can be obtained. When ix is acquired, the RSSs of f_j' can be rearranged based on the sorted ix ix_{sort} and RSS_j can be obtained based on this. Then, the system calculates the correlation coefficients of f_i and f_j using the following formula:

$$Corr_{ij} = \frac{cov(RSS_i, RSS_j)}{\sigma RSS_i * \sigma RSS_j} \tag{11}$$

$cov(RSS_i, RSS_j)$ is the covariance of RSS_i and RSS_j, and σRSS_i and σRSS_j are the standard deviations of RSS_i and RSS_j.

Based on the calculated Jac_{ij} and $Corr_{ij}$, the system can verify that activity landmarks i and j are collected from the same node under the following conditions. If both $Jac_{ij} \geq jac$ and $Corr_{ij} \geq corr$ inequalities are satisfied, then i and j are located on the same node. Conversely, if these conditions are not fulfilled, the two activity landmarks are not located on the same node. Each bound cluster is treated as a single node within the graph. When clustering is completed for all activity landmarks, the ALIMC estimates the distance between all the nodes and obtains a distance matrix between them based on the estimate. Then, the ALMIC uses multidimensional scaling techniques to construct indoor maps based on the distance matrices.

4.2. Tactile interaction

It is more difficult to develop an indoor positioning system for visually impaired people owing to the difficulty of visually providing information about the indoor location to them. Therefore, the system must adopt other sources of information such as sound. Hence, it is necessary to install additional auxiliary devices or specialized sensors [73]. However, the extensive installation of these auxiliary devices or expensive sensors increases the system's cost of operation tremendously. Apostolopoulos et al. performed a pilot project to determine whether the location of visually impaired people can be tracked and guided using affordable sensors installed on mobile devices such as smartphones.

The system data d_T available at time T is the conversion $u(0:T-1)$ and observation $o(0:T)$. The transition u_T at time t corresponds to an operation in which the user acquires the heading direction u_t^θ and moves forward by u_t^f. This transition determines the behavioral transition model of the system.

$$(x_{t+1}, y_{t+1}, \theta_{t+1}) = \left(x_t + u_t^f * cos\left(u_t^\theta\right), y_t + u_t^f * \sin\left(u_t^\theta\right), u_t^\theta\right) \tag{12}$$

In the system, u_t^θ and u_t^f are measured using a compass and a pedometer, respectively. The observation o_t^j of landmark type L_j in the user's state $\varepsilon_t = (x_t, y_t, \theta_t)$ includes the following:

$$\exists\, l^i \in L_j : \| \ (x_t, y_t), \left(x^j, y^j\right) \ \| < R_{obs} \tag{13}$$

The above observation model specifies that the corresponding landmark $\varepsilon_t = (x_t,\ y_t, \theta_t)$ must exist within the predefined observation distance R_{obs} from the user's current location coordinates $(x_t,\ y_t)$ to enable the user to detect the landmark type L_j around them.

The goal of the system is to make it possible to gradually estimate the user's state ε_t at time T. To achieve this, the system uses a Bayes filter [74] that computes the reliability distribution $B_T = P(\varepsilon_t|d_T)$ for ε_t over time T with the given data d_T. Three factors are required for this calculation. The first is the initial setting B_0, the second is the conversion model $P(\varepsilon`|u, \varepsilon)$, which describes the probability that the user will be in the new position ε' when the position is switched from the previous position, ε,to u. The third factor, which is based on the map m, is an observation model $P(o|u, \varepsilon)$ that describes the possibility of observation o when the user is in ε. At this time, the map is assumed to be unchanged and accurate.

When the three elements are prepared, the system updates the reliability distribution according to the normalization factor as follows:

$$B_T = \eta * P(o_T|\varepsilon_T, m) \int P(\varepsilon_T|\mathrm{u}_{T-1}, \varepsilon_{T-1}) * B_{T-1} * d\varepsilon_{T-1} \tag{14}$$

B_T can be expressed through a set P of particle $p^i = (\varepsilon^i, w^i)(\mathrm{i} \in [1,\mathrm{N}])$. Each particle stores the state estimation ε^i and a weight w^i representing the probability that ε^i is at the actual location. The reliability distribution of the particle filter is better when the particle number is closer to infinity. To update the particle filter based on the new conversions and given observations, a similar approach was adopted to importance sampling. At each time step T, the system follows the next step based on the given list of particles $\{p_T^1, p_T^2, \ldots, p_T^p\}$, conversion u_T, and observation o_{T+1}.

In the last step, the system collects all the generated sensor readings such as the compass direction and pedometer steps. Usually, during a one-time step, the compass provides multiple direction estimates, which the system averages to obtain a new estimate for the heading direction, u_t^θ. Pedometers usually return a value of zero or one after detecting a step. The system must obtain a distance estimate by transiting the values obtained from the pedometer. To calculate the length of the steps, the system ran a short training session for each user. During this session, the user goes through two routes between two landmarks with known distances. The pedometer calculates the number of steps while moving through these paths, and the device estimates the average length of the steps. Through this estimate and the number of steps measured by the online pedometer, the system calculates the travel distance estimate u_t^f.

Based on the given u_t^θ and u_t^f, different levels of noise are added to apply the transition model to each particle. The noise parameter for the particle p^i is derived as a normal distribution, computed as $(u_t^\theta)^i = \mathcal{N}(u_t^\theta, \sigma_\theta^2)$ and $(u_t^f)^i = \mathcal{N}(u_t^f, \sigma_f^2)$, respectively. The resulting value is used to obtain a new state ε_{T+1}^i in Eq. (13). The corresponding transition from ε_T^i to ε_{T+1}^i uses the map to estimate whether the route taken is likely to cause a collision with an obstacle. If a collision occurs, the sampling of $(u_t^\theta)^i$ and $(u_t^f)^i$ is repeated up to a certain number of trials or until a collision-free transition is found.

The algorithm of the system samples higher-probability particles with higher weights. However, there might be certain cases in which all the particles weigh zero owing to particle defects. In such cases, the mixed Markov cluster method [75] continues to sample a certain number of particles from the observations, whereas the "sensor reset" approach samples from observations only when the observations deviate significantly from the previous distribution. Thus, the approach implemented follows a similar logic.

All the particles weigh zero, when, upon checking the landmark l^i, the user finds that the filter does not bring them near the landmark. The particles are sampled immediately from the observation. For each particle p_t^q, the function computes the landmark of the identified type closer to p_t^q. If the nearest landmark found is l^i, the line between p_t^q and l^i is calculated. If there is a LoS between the particle and landmark, the new particle is sampled in a region that simultaneously satisfies the top of the line segment $[p_t^q, l^i]$ and within the radius R_{obs} representing the longest distance where the landmark can be detected. The line segments are used in the calculations to ensure that new particles do not cross into a room or other corridor.

4.3. BatTracker

Recently, mobile device manufacturers have added many more features to mobile devices, including more effective sound systems [76]. Researchers have asserted that the relative distance from an object to the mobile can be inferred through an acoustic signal emitted from the device, reflected by the object, and received back by the device. Based on this, Zhou et al. proposed BatTracker, an indoor mobile device positioning system without a high-density infrastructure in a 3D space with a range similar to that of existing solutions [37].

The system uses acoustic signals that are continuously emitted by mobile devices. The emitted acoustic signals bounce off the surface of the surrounding object. The system receives the generated echo and infers the relative distance to the object. Subsequently, the system is used to correct the drift errors in the inertial data using the inferred distance estimates. BatTracker's acoustic sensing module consists of a signal emission section, recording, and a series of signal processing steps to generate the distance, amplitude, and Doppler shift [77]. Therefore, as shown in Fig. 5, the system can measure the speed of an ultrasonic candidate of a surrounding object in an indoor environment.

The first step is the generation of an acoustic signal to be emitted. The conventional acoustic distance measurement system uses a frequency-modulated pulse signal with a linear resolution frequency to improve the distance estimation. However, because it is difficult to derive the fluctuation of the frequency from such a frequency-modulated pulse signal, the system selects a signal with a regular frequency, 17 KHz, which is slightly audible to humans. A high-frequency signal reduces the robust tracking range owing to the faster signal power attenuation. In certain situations, at the normal volume, the users can barely hear the sound of the designed system owing to the background noise. Owing to this problem, the system selected a wavelength of 1 ms. A shorter pulse duration is preferred because the system requires high

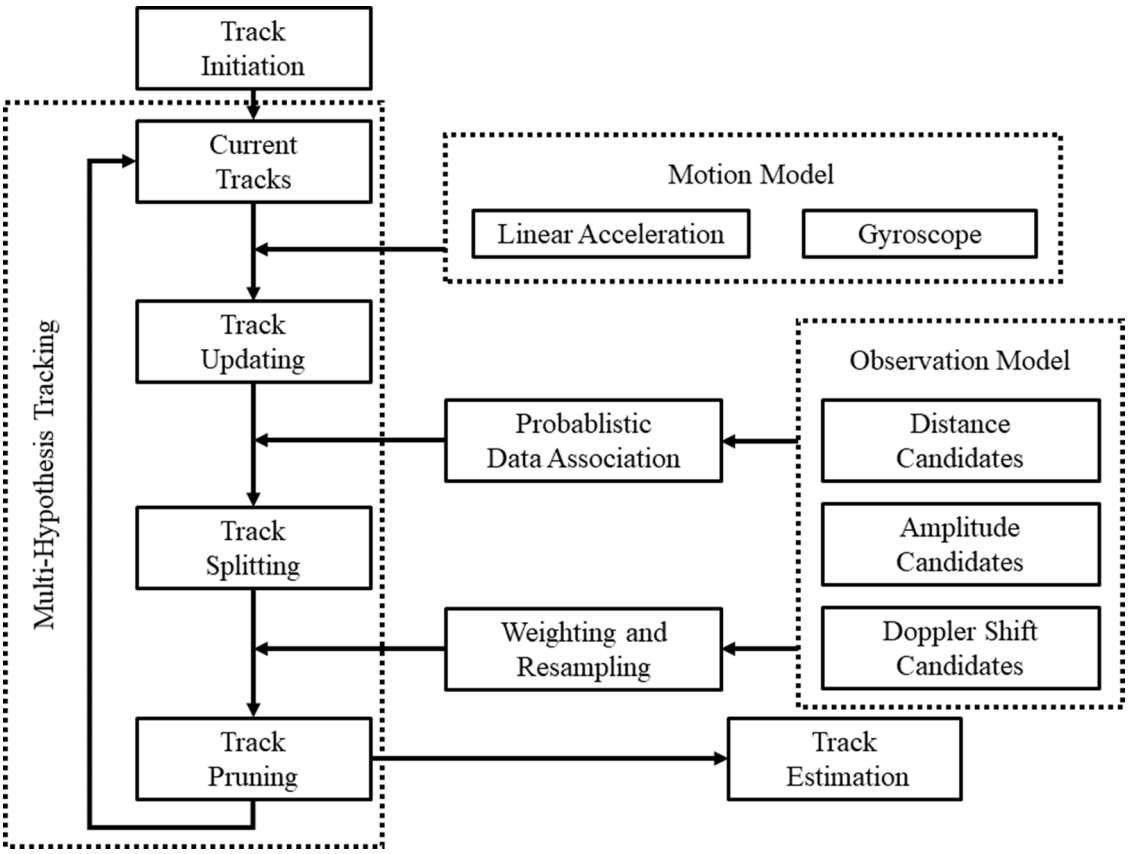

Fig. 5. System architecture of BatTracker

accuracy for distance measurement. The speed of the Doppler shift acts as a complementary input.

The system adjusts the envelope size by applying a Hanning window to the pulse, thereby enabling the system to increase the signal-to-noise ratio of the echo [78] by increasing the peak-to-side lobe ratio. To prevent the echoes of two consecutive pulses from overlapping, a considerable gap between the echoes is required. The system emits the generated sound and receives the signal reflected from the object. The system performs noise cancelation, individual wave tracking, distance estimation, and Doppler shift estimation to generate distance and speed from the received signal.

In the noise cancelation step, the system passes the received signal through a Butterworth bandpass filter with a passband of 17 K $\pm$ 200 Hz to remove background noise. Through this process, the system eliminates background noise while preserving the frequency shift caused by the Doppler effect. This step is very important for tracking in noisy environments, which can weaken noise reflection.

In the individual wave tracking step, the system correlates the signal with the wave designed through the usual general technique of generating peaks for each echo and obtains the upper envelope of the signal. The system divides the envelope into small time windows of 31 ms, each containing only echoes in one wave. To do this, the system needs to determine the starting point of the window. The first peak is always the sound that goes directly from the speaker to the microphone; it coincides with the highest amplitude. Hence, the system uses the peak as the starting point.

In the distance estimation step, the system detects several peaks corresponding to the different echoes of different objects for each emitted wave. The system uses a threshold to select only the strongest K peaks that are expected to be reflected from relatively larger and closer objects. The system then estimates the distance between the device and the surrounding objects by calculating the time delay between each echo and the starting point. It also extracts the amplitude of each echo that is used to connect the data in the tracking algorithm.

After positioning each echo in the Doppler shift estimation step, the system analyzes the frequency f_e of the received echo composed of 48 sampling points (1 ms). Because the sampling frequency is $f_s = 48$ KHz, the Fourier transform at 48 points results in a frequency resolution of $\frac{48000}{48} = 1$KHz. The system uses a method similar to the artificial padding method of the zero-value point and applies a short-term Fourier transform [79] to obtain a 1 Hz resolution. The relative velocity for each object where an echo occurs can be calculated using the following equation:

$$v = \frac{f_d}{2 * f} * c \tag{15}$$

where $f_d = f_e - f$ is the Doppler frequency shift, $f = 17$KHz is the frequency of the design signal, and c is the speed of sound propagation.

f_d represents the direction of movement, where $f_d > 0$ indicates movement toward the object, and $f_d < 0$ indicates movement away from the object. This gives us a speed resolution of $\sim$1 cm/s. Because the number of sampling points is very limited, the speed estimation may not be very accurate owing to the Doppler shift. Therefore, the system does not explicitly use the data to estimate the device movement speed but integrates it as a complementary input to enhance data connectivity.

5. Visual Image

The visual image prioritizes the visual element. For this reason, indoor positioning technology through visual image matching faces a primitive problem with images. For example, for an indoor space without special structures, image matching should be performed using the structure of the general indoor environment. However, a similar structure may exist in several buildings, which makes it difficult to sort

data and match the image captured by the user. In order to supplement the problem of the positioning algorithm through the visual image, many researchers conducted studies using new elements.

5.1. Luxapose

Artificial lighting facilitates the location of objects in a dark indoor environment, and although its purpose has always been for lighting, it has now become an indispensable part of interior decoration. Consequently, diverse technologies that leverage light for other purposes have been proposed. Kuo et al. introduced Luxapose, which uses smartphones and slightly modified commercial light-emitting diode (LED) illuminators for indoor positioning.

An illuminator adapted for a fast ON–OFF switch transmits the location encoded in a human-detectable optical pulse. Then, a smartphone equipped with a camera can detect the light emitted by the illuminator using only one image frame capture. Subsequently, the phone decodes the transmitted identifier or location, compares it with the location of the illuminator, and determines the location and orientation of the smartphone.

The algorithm enables a continuous location update by continuously capturing and processing images based on the following four factors. The first is the driver circuit of a known LED lighting system that can be easily modified to transmit data via ON–OFF keys. The second insight is to utilize the rolling shutter effect of the complementary metal-oxide-semiconductor imager [80] to receive many bits of data encoded in the optical transmission with a single frame capture. Third, the camera is essentially an AoA sensor; thus, the projection of several nearby light sources with known positions in the camera's image can be framed as an example of a sufficiently limited AoA positioning problem. The last is an optimization technique that can solve this problem. In the study, an analytical model that is utilized for the design as well as an auxiliary design option using a prototype system is used.

Luxapose uses an optical AoA positioning principle based on an ideal camera with biconvex lenses [81]. An important characteristic of biconvex lenses is that the rays passing through the center of the lens are not refracted. Therefore, the transmitter, center of the lens, and projection of the transmitter projecting on the camera imager plane form a straight line.

Suppose that the transmitter T_0 with coordinates $(x_0, y_0, z_0)_T$ in the transmitter's global reference frame has the image i_0 with coordinates $(a_0, b_0, Z_f)_R$ in the receiver's reference frame. The position of T_0 is on the line passing through $(0,0,0)_R$ and $(a_0, b_0, Z_f)_R$, where Z_f is the pixel distance from the lens to the imager. Using a similar triangle geometry, the unknown scaling factor K_0 can be defined for the transmitter T_0, and the position $(u_0, v_0, w_0)_R$ of T_0 in the reference frame of the receiver can be described as follows:

$$\begin{cases} u_0 = K_0 * a_0 \\ v_0 = K_0 * b_0 \\ w_0 = K_0 * Z_f \end{cases} \tag{16}$$

The system positioning algorithm assumes that the transmitter location is known. Hence, the bidirectional distance between the transmitter and reference frame of the transmitter can be inferred. When the equation is calculated in two different domains, only the remaining unknowns have scaling factors $\{K_0, K_1, \ldots, K_n\}$, and a set of quadratic equations is calculated.

When all the scaling factors are estimated, the position of the transmitter can be determined in the reference frame of the receiver, and the distance between the receiver and transmitter can be calculated. The relationship between the two domains can be expressed using the following equation:

$$\begin{bmatrix} x_0 & x_1 & \ldots & x_{N-1} \\ y_0 & y_1 & \ldots & y_{N-1} \\ z_0 & z_1 & \ldots & z_{N-1} \end{bmatrix} = R * \begin{bmatrix} u_0 & u_1 & \ldots & u_{N-1} \\ v_0 & v_1 & \ldots & v_{N-1} \\ w_0 & w_1 & \ldots & w_{N-1} \end{bmatrix} + T \tag{17}$$

where R is a 3 × 3 rotation matrix, and T is a 3 × 1 transformation matrix. The element T (T_x, T_y, T_z) indicates the position of the receiver in the frame of the transmitter reference. The system builds a transformation matrix according to the geometric relationship. Because the scaling factor is known, the equivalent distances of the two domains facilitate the calculation of the receiver's position in the coordinate system of the transmitter as follows:

$$(T_x - x_m)^2 + (T_y - y_m)^2 + (T_z - z_m)^2 = K_m^2\left(a_m^2 + b_m^2 + Z_f^2\right) \tag{18}$$

where (x_m, y_m, z_m) are the coordinates of the m^{th} transmitter in the reference frame of the transmitter, and (a_m, b_m) is the projection of the mth transmitter projected on the image plane. Finally, the set of receivers (T_x, T_y, T_z) to be minimized is estimated as follows:

$$\sum_{n=1}^{N} \left\{ (T_x - x_m)^2 + (T_y - y_m)^2 + (T_z - z_m)^2 \right\}^2 - K_m^2\left(a_m^2 + b_m^2 + Z_f^2\right)\} \tag{19}$$

Once the transformation matrix T is known, the rotation matrix R can be obtained by calculating its elements. The 3 × 3 rotation matrix R is represented by the three column vectors $\overrightarrow{r_1}$, $\overrightarrow{r_2}$, and $\overrightarrow{r_3}$, which are the components of the unit vectors $\hat{x}'$ $\hat{y}'$, and $\hat{z}'$ projected on the x, y, and z axes in the reference frame of the transmitter, respectively. Once the receiver orientation is confirmed, it is necessary to switch between the portrait and landscape modes to determine the orientation and compute the projection on the X–Y plane.

5.2. Visual Sequence

The camera mounted on the smartphone acts as a sensor for detecting the visual elements. It acquires the structures or features in an indoor environment as images. It then analyzes and stores the acquired images. The system can recognize the user's current location by matching the image captured when the user moves with the stored image. According to this concept, the indoor positioning of the camera-based computer vision approach can be divided into two categories: image search and motion structure search.

Qing Li et al. analyzed the key drawbacks of these two technologies and proposed a technique to extract the landmark sequences from a sequence of images and match the topological map with the landmarks generated in the indoor plan map, instead of comparing the images directly [40].

The system process is divided into two main phases, online and offline. In the offline phase, the system is trained to recognize indoor objects and scenes using a convolutional neural network (CNN) [82]. In the online phase, after an image is extracted from the video sequence, a local suggestion algorithm is used to generate an image patch containing an indoor object. Then, the image patch is fed to the trained CNN, and the system classifies the landmark into three types according to the recognition result of the indoor object. A single object landmark is defined as a single object such as a fire extinguisher or elevator. To deal with objects that are difficult to classify as a single object landmark because of their light characteristics, such as a door frame, the system classifies objects as a multiobject landmark by combining results for different objects. Finally, the scene landmarks are defined as the main locations of indoor structures such as corners, intersections, and corridors.

In the study, the system is trained to recognize indoor objects and scenes using AlexNet [83]. AlexNet has five convolutional layers with

each layer, followed by a max-pooling layer. The system uses two fully connected layers to focus on the overall features of the subsequent convolutional layers. The input layer receives the image pixels as input and the output layer uses a multiclass prediction by classifying the image using a probability distribution for each class. Therefore, the shape of the output layer is equivalent to the number of classes. Using the pretrained CNN network helps extract the image's information without retraining the system. The system deploys the transfer learning approach by using pretrained weights directly and modifying only the output layer for landmark detection.

The first step in the online phase was video frame extraction. The sampling rate for image extraction is essential for the accuracy and efficiency of landmark detection. If the sampling rate is very low, successive images will have a low overlap or no overlap at all. In this case, the system may not be able to identify the object entirely in the image, or other foreign objects may appear in the image, causing some landmarks to be missed. However, a high sampling rate leads to information duplication, resulting in a low landmark detection efficiency. Empirically, the overlap between two successive images should be greater than 90% to avoid missing landmarks. In general, the human walking speed is 1.4–2 m/s, and a 90° rotation is achieved in 0.8 Three to five frames per second empirically satisfy the requirement of over 90% overlap.

When a system captures an image in an indoor environment, with or without interest, more than one indoor object, such as a floor, chair, or table, appears in the image. The appearance of unwanted objects impedes the effectiveness of target recognition. Thus, finding an area that includes only the objects of interest can greatly increase the detection accuracy. The system chooses an optional search algorithm to generate a region of interest from the image instead of generating a patch based on the drag of the object. The algorithm consists of two steps.

First, the system selects the hyper segmentation algorithm to generate a large initial region in various color spaces with different parameter ranges. Then, a hierarchical grouping strategy based on various similarity measures, such as color, size, and texture, was applied. After the above process, candidate areas of various sizes were generated. The landmark covers a specific space in the captured image and small-sized areas are unlikely to have indoor objects of interest. Therefore, setting a threshold to filter these areas increases the detection efficiency.

In the case of the object extracted from the image, after the indoor object recognition step, the indoor object is divided into those that appear and those that do not. Images without indoor objects are useless and discarded. For images with objects, the trained CNN network recognizes the object type. If an object is used to define a single object landmark, the landmark is detected. When an object is a component of a multi-object landmark, more information is needed to determine the landmark type. To prevent the separation or omission of indoor objects, the images were sampled at a high overlap rate. The disadvantage in this case is that one object appears in more than one image. Thus, instead of determining a landmark based on one image, the system uses a continuous sequence of images to recognize the landmark.

Another reason for using a continuous image sequence is to detect multiple object landmarks that may not be detected in a single image because of its size and position in the image sequence. After landmark detection, the video sequence is divided into landmark segments. The landmark segment starts at the first frame when the landmark type is detected and ends at the last frame containing the current landmark

5.3. Pair-Navi

Dong et al. presented Pair-Navi—a peer-to-peer indoor navigation system that does not require any prebuilt infrastructure, a prebuilt localization service, nor an indoor digital map. The system investigates visual SLAM using mobile vision [84, 85]. Visual SLAM uses one or more cameras to continually locate the camera within the environment and simultaneously construct an environment map to explore unknown environments. The system is based on a single visual SLAM using a single camera mounted on a typical smartphone.

Unlike traditional navigation systems that rely on prebuilt location positioning services, Pair-Navi operates in a simple peer-to-peer (P2P) navigation mode. There are two main roles in a navigation system, leadership and followership. The basic idea underlying the P2P navigation system is to reuse the experiences of early users who have become leaders. Anyone can act as a leader for a specific route by providing positioning information along with the automatically extracted walking hint. The follower's navigation is then accomplished by synchronizing their relative position with the leader's reference path.

Pair-Navi enables this type of leader–follower navigation using the mobile vision function. The system receives all the video frames captured by the smartphone camera and sends them to the cloud server for further processing. The system inputs the video clip sent for the leader to two SLAM modules, visual odometry, and trajectory construction, to simultaneously execute route formation and mapping. When the positioning is complete, the system stores the route and map data in addition to the map-labeled origin and destination locations on the server.

When the follower arrives indoors and selects a destination, the system provides the reference path that led the leaders to the same destination. When the follower moves, the system instantly relocates based on the video frame captured by the follower to instantly perform localization based on the reference path. If the relocation is successful, the follower will also be relatively positioned according to the leader's reference path and will be instructed in real-time navigation. However, if a positioning failure occurs because a follower deviates from a given path in a subsequent move, the deviation detection module triggers the execution of an auxiliary visual odometer that tracks the follower's camera posture independently. This enables the system to receive the result back and display it along with the reference path on the device of the follower to enable the follower return to the correct path for navigation.

A key and unique component of Pair-Navi is the nonrigid context culling (NRCC), which aims to extract and subtract the dynamic content of video clips to cope with changes in the environments over time. To ensure the creation of an accurate reference path and strong follower positioning, the NRCC was applied to both the leader's and follower's videos.

The system contains two key modules: a visual odometer and a repositioning module. First, the system performs an initialization step by using epipolar geometry to position the camera on the initial map, where the 3D map point is the environmental landmark. Through consecutive video frames, the system uses a visual odometry (VO) module to continuously track the camera posture. The VO module is the core module that makes up the route map of the leader. When a new video frame arrives, the system extracts the 2D feature points from the frame and compares them with the 3D map points previously generated via a feature matching process.

Following the association of the 2D feature points with 3D map points, the system can acquire the camera pose of the frame by solving the perspective-n-point (PnP) problem. This involves determining the position and orientation of the camera by using a corresponding set between 3D points and 2D pixel points. When considering the camera pose of two frames, the system can generate new 3D map points by calculating 3D coordinates through triangulation between the two frames. As new frames emerge, the system gradually builds a path to the camera pose, a map of 3D landmarks, and corresponding keyframes.

The relocation module is useful for reusing the route maps created through the previous process, which is a key element of the follower program. The module compares the video frame to the keyframe on the map and finds the closest keyframe based on feature point matching. After the most similar keyframe is found, the feature point of the current follower frame is connected to the feature point of the selected keyframe. To obtain a camera pose based on this, the system solves the PnP

problem in the same way as the VO module and repositions the camera on the map.

The system depends heavily on the visual SLAM technology; however, the technology is not very effective in nonrigid environments where significant dynamic changes occur over time. To address this, Pair-Navi adjusts the R-CNN mask to fit the NRCC. The mask R-CNN is a framework for instance splitting. For each instance, the framework seeks to separate different instances from the image using a split mask. The system deploys a pretrained R-CNN mask using the common objects in context dataset and selects an appropriate object category for the indoor scenario.

When the system receives a video frame from the camera at the server, the system extracts the function point, detects the nonrigid context using the mask R-CNN framework, and filters the function point in the mask of the dynamic instance. After preprocessing the video frame, the user may encounter a nonrigid environment; however, the feature points in the dynamic instance's mask do not participate in the path structuring (for the leader) or repositioning (for the follower).

6. Multisources

Improving the performance of mobile devices means expanding the technologies that can be installed. Researchers have increased the accuracy of location tracking technology by using complex elements rather than one element. The use of complex elements complements the shortcomings of the individual elements and maximizes their strengths.

6.1. JADE

With the recent proliferation of mmWave [86, 87] communications in the 30–300 GHz band and the advent of commercial mmWave systems (typically 60 GHz), several researchers have asserted that mmWave can play a crucial role in indoor positioning systems. Owing to its short wavelength and large bandwidth, the mmWave signal can generate multiple reflection paths and reach different receivers according to its arrival time and angle, particularly when it is transmitted in a single direction. This makes it possible to obtain an approximate environmental map at the beginning of positioning. Palacios et al. proposed the JADE, which can track mobile users in indoor environments using only AoA information, while assuming that they do not have prior knowledge of the surroundings.

JADE uses the angle difference of arrival (ADOA) algorithm with the advantage of rotational invariance. It stores all information for both the LoS and nonLoS (NLoS) components extracted from the AoA spectrum. The LoS component is easy to measure because it is clearly in the line of sight; however, the NLoS portion is difficult to measure because it is often beyond the visual field. Therefore, the system creates a virtual AP (VA) that is defined by the symmetry between the physical AP and the NLoS elements based on the boundaries of the room such as the floor, ceiling, and wall. Therefore, the system increases the accuracy of the positioning process by using the generated VA created as an additional lever.

In this system, the interior spaces are characterized by boundaries and are assumed to contain additional obstacles. Simultaneously, the system assumes that many APs in the room are in unknown locations, and users can discover and connect to these APs via the LoS or NLoS paths. These paths are typically discovered through the beam training process, which is widely used in mmWave communication systems. The AoA of the signal from some APs changes as the user shifts positions; some APs may disappear behind obstacles and others may reappear via LoS or NLoS paths. In either case, the system assumes that the user can distinguish between the multipath components of the signal transmitted by each AP and can determine if the AP has previously appeared. Finally, the system assumes that the user has no preliminary knowledge of the environment.

The system denotes the set of indoor APs used as physical anchors, A_p. It then denotes A_v as a set of VAs created by symmetrical physical AP $\alpha \in A_p$ relative to the indoor boundaries. Consequently, it denotes a set of all the anchors present in the room as $A = A_p \cup A_v$ and designates the anchor using positions in x_i and y_t, where x_i is the position of the i^{th} anchor, and y_t is the user's position at time t.

The system then finds a pair (i,t) of anchor i whose device can measure the AoA $\phi_i^{(t)}$ of the LoS at time t and denotes a set containing all the applicable pairs as V. The system labels the ADOA generated by the signals that are obtained from the anchors $i, j \in A$ at time t as $\theta_{ij}^{(t)} = \phi_j^{(t)} - \phi_i^{(t)}$ and $\varsigma_{ij}^{(t)} = \pi/2 - \theta_{ij}^{(t)}$. During this process, the angle was measured counterclockwise.

The ADOA positioning problem involves finding a set of anchors and user positions over time so that the pairwise difference between the anchor signals and AoA is compatible with the pairwise difference between the components of the AoA spectrum detected by the users. At any given time t, it can be inferred that the user is positioned on the arc above the shortest line (segment) $\overline{X_iX_j}$ which connects the two anchors i and j when the user is at an angle $\phi_i^{(t)}$ and $\phi_j^{(t)}$ from anchors i and j, respectively, and the AoA is error free.

In this situation, the system adds a third anchor, k. When the information in anchor k is available, an estimate of its location is known via the intersection of the two anchors above chords $\overline{X_iX_j}$ and $\overline{X_jX_k}$. When the current time is t and the node position is $y^{(t)}$, the node checks the three anchors for this time and position. One is a physical AP located in x_1 and the others are VAs created by symmetrically moving the AP of x_1 relative to the lower and left walls represented by a thick black line. The two VAs are located at x_2 and x_3, respectively.

When anchors i and j at time t satisfy $(i,t),(j,t) \in V$, the system sets the center of the arc, including anchors i and j located in x_i and x_j as $c_{ij}^{(t)}$. Through spatial construction, the center of all the arcs is above the line bisecting segment $\overline{X_iX_j}^{(t)}$.

By defining $v_i^{(t)} = 2 * \| y^{(t)} - x_i \|^{-2} * (y^{(t)} - x_i)$, the equation of the bisector is expressed as follows:

$$v_i^{(t)T} * R_{\varsigma_{ij}^{(t)}}\left(x_j - x_i\right) = 2sin\theta_{ij}^{(t)} \tag{20}$$

where T denotes the intersection. Because the center of all arcs, including anchor i in time t, lies above the line, the system determines that they also intersect at one point symmetrical to anchor i in relation to the bisector in the structure. This is an estimate of the location of the current user. Thus, the system calculates the user's position in time t by solving Eq. (20).

6.2. WAIPO

Advances in crowdsourcing and sensing technologies have spurred the invention of new methods to improve the accuracy of existing location tracking systems. Crowdsourcing is a powerful method for automatically and effectively organizing an accurate fingerprint database. Ambient sensing technology differentiates locations based on their surrounding functions. The collaborative location tracking method that considers existing algorithms and human factors is a new approach in this field. F and Gu proposed WAIPO, an indoor location tracking system that combines Wi-Fi, magnetic fingerprints, image matching, and simultaneous user generation.

Fig. 6 shows the system architecture of WAIPO, which is divided into client and server applications. The client application running on the smartphone is responsible for detecting the surrounding. The server includes data processing, learning programs, and data storage. The structure of WAIPO consists of four main components: data storage, data collection, smartphone execution module, and server location tracking.

The smartphone execution module is a WAIPO kernel that runs on a smartphone. Using a smartphone execution module, WAIPO collects

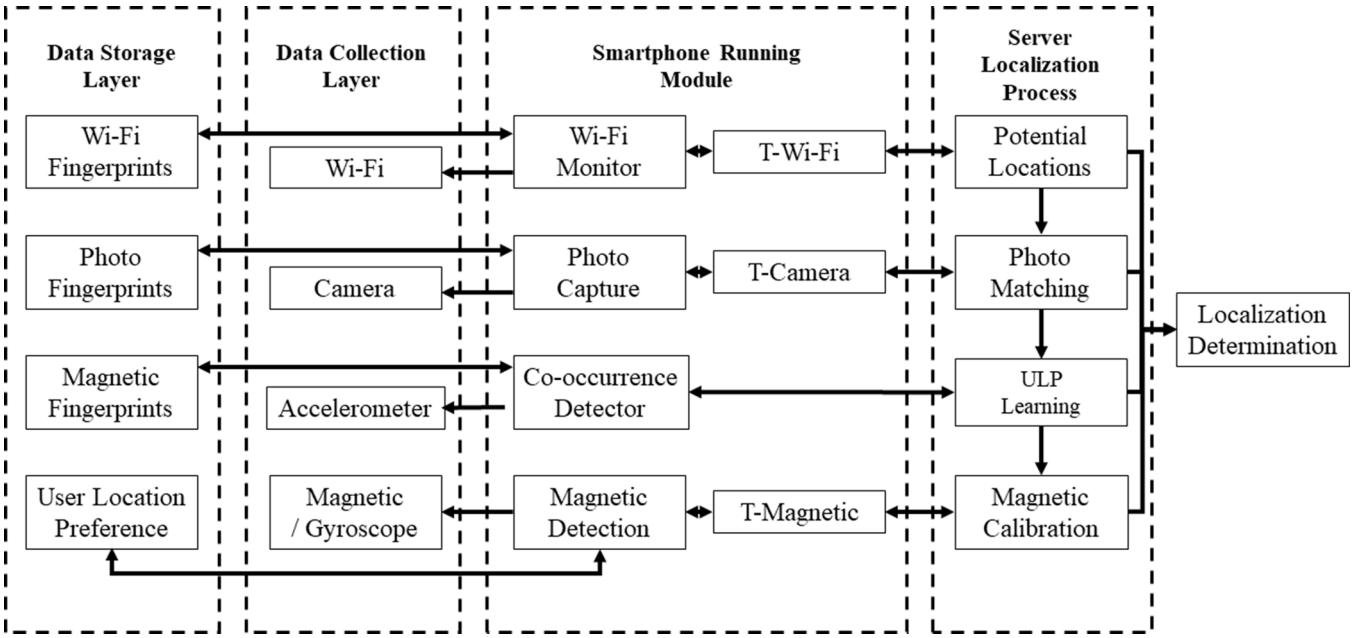

Fig. 6. System architecture of WAIPO [75]

three fingerprints, Wi-Fi, photos, and magnetic fields. The smartphone execution module consists of seven modules: Wi-Fi monitor, photo capturing module, co-occurrence detector, magnetic detection module, T-Wi-Fi, T-photo, and T-magnetic. The Wi-Fi monitor module is an observation module that detects nearby APs. The Wi-Fi monitor module sends the detected Wi-Fi RSS to the server to determine the potential location of users who demand for their location. The photo capture module was designed to identify the user's current location through photos. The smartphone randomly takes pictures and sends them to the WAIPO server. To build a photographic fingerprint database, WAIPO extracts visual features deploying a scale-invariant feature transform (SIFT) algorithm using two photos taken in each room. The server identifies the key points of the received photo, analyzes it, and compares it with the photo fingerprint in the database.

The co-occurrence detector module is a key component of the WAIPO for colocation tracking. When starting the location tracking service, the co-occurrence detector module searches for other Bluetooth devices that are open for pairing and sends the result to the server for simultaneous detection of the user. If the Bluetooth interface is not available, the module replaces it with a Wi-Fi interface. However, if both are available, the communication overhead of Wi-Fi is significantly greater than that of Bluetooth. Therefore, WAIPO first selects the Bluetooth interface to detect the user's co-occurrence. The magnetic detection module is used to detect the magnetic field value of the user's location. The readings of the magnetometer are designed to correct the position tracking results, which can improve the position tracking accuracy. The user location preference (ULP), which is the user's recording location, is stored in the data storage layer. The ULP records were synthesized from the matching results of the Wi-Fi, magnetic fields, and photographic fingerprints.

The smartphone execution module needs to scan Wi-Fi signals, take pictures, and detect magnetic field values to proceed with location tracking. To this end, the smartphone execution module includes a system timer that triggers the corresponding modules. However, if the timer is short, the client application consumes the battery quickly. Conversely, if it is too long, the location tracking system loses its basic usability. Three triggers are defined in the WAIPO to tune the timer. The T-Wi-Fi trigger drives the Wi-Fi monitor module, the T-photo trigger controls the photo capture module, and the T-magnetic trigger controls magnetic detection. All the three triggers were designed for efficient energy consumption. In addition, the T-photo has the additional task of protecting users' personal information when images are deployed.

The WAIPO server location tracking process module runs on the server. The module consists of four components: potential location calculation, photo matching, ULP learning, and magnetic field correction. The potential location calculation is a set of potential locations of the user determined by calculating the RSS of the Wi-Fi for the AP. Photo matching is designed to determine the user's location by receiving a photo from a smartphone and comparing the received photo with a photo fingerprint stored in the database. ULP learning represents a record of the locations through which the user has moved across. The magnetic field calibration is used to calibrate the location results to improve the location tracking accuracy. The predicted position, which is the output of the WAIPO system, is sent back to the smartphone to help the user determine his or her location.

6.3. Wi-Lo

Crowdsourcing is restricted to suburbs or rural areas where user density is low. Some researchers have used both wireless signals and built-in sensors to improve the accuracy of indoor positioning technology. Di Felice and Marco developed Wi-Lo, a wireless fingerprint technology that shows improved accuracy in both high-density and low-density APs owing to the simultaneous use of wireless interfaces typically embedded in smart devices, including acceleration sensors.

Fig. 7 shows the software architecture of Wi-Lo. Wi-Lo supports four wireless fingerprint input sources: Wi-Fi, LTE, BLE, and magnetic field (MAG). A barometer can optionally be used to detect user floors. The Wi-Lo user can select a wireless input source to be considered for creating a fingerprint map in the offline stage and an input source to be selected for location tracking in the online stage through the graphical user interface of the mobile app. If I_{online} and $I_{offline}$ are sets of selected input sources, then $I_{online} \subseteq I_{offline}$. The client software performs measurements on a specific RP and sends it to the server. The server aggregates the radio frequency map based on the transmitted data and stores it in the database. The server side executes the fusion method after executing the pattern-matching algorithm selected for each source selected in I_{online} and finally returns the expected RP to the client to perform the location tracking process.

The client software performs data measurements on a specific RP. The client first selectively collects Wi-Fi, LTE, BLE, and MAG samples for

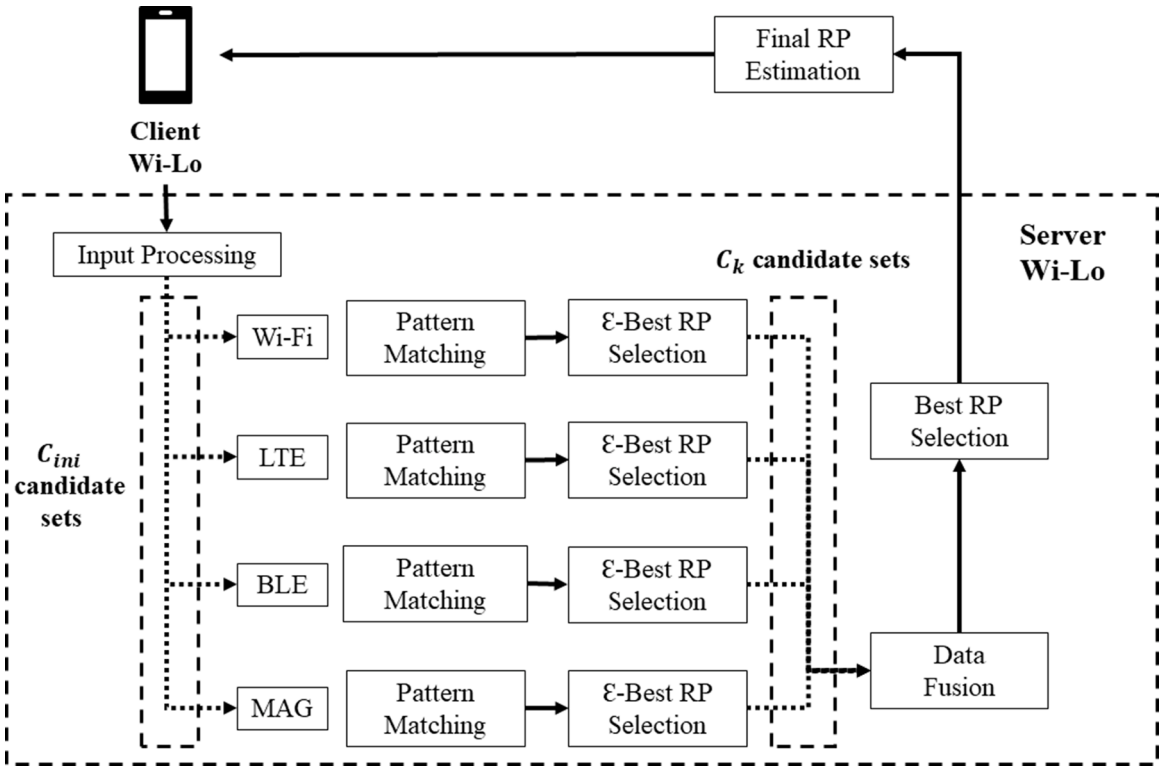

Fig. 7. System architecture of Wi-Lo [76]

T_d seconds based on the source selected by I_{online}. Optionally, the client detects the current floor using a barometer sensor. When data collection is complete, the client transmits the collected data record to the server, along with the floor location. In the case of an additional movement path search, the client transmits the previously expected RP.

The server software aggregates the RF map based on the transmitted data and stores the map in the database. If the floor location and the previously predicted RP are available at the same time, the server constructs C_{ini}, an initial set of candidate RPs, based on the two datasets. If the previously expected RP cannot be used, $C_{ini} = P_{rp}$, where P_{rp} is the set of RPs in the building. Subsequently, the server applies the selected pattern-matching algorithm to each source. Each algorithm assigns a score $S_k(i)$ to an individual RP i ∈ C and data source K. Thereafter, the server calculates the maximum score, $Max(S_k)$, and selects ϵ-Best RP to populate the candidate set C_k for the data source k ∈I_{online}, along with the RP containing the relative score of the difference less than ϵ and $Max(S_k)$. C_k is calculated using the following equation:

$$C_k = \left\{ \text{RP } j.s.t. \left| \frac{S_k(i) - Max(S_k)}{Max(S_k)} \right| \leq \varepsilon \right\} \tag{21}$$

The server also defines the reliability metric of the data source k=κ (k) as the inverse of the cardinality of the candidate set. This is denoted by κ (k) =$\frac{1}{|C_k|}$. Thereafter, when a single source k is selected, the server stops fusing the data, and the RP assigns the same score as $Max(S_k)$ is returned. If no source is selected, the server merges the results of the candidate set C_k. The identity function $U_k(i)$ returns 1, if RP i ∈ C_k, and 0 otherwise. The fusion module shown in c calculates the fusion score for RP i ∈ P_{rp}using the following equation:

$$FS(i) = \sum_{k=1}^{|I_{online}|} U_k(i) * \omega_k(i) \tag{22}$$

where $\omega_k(i)$ is a weight applied to determine source k within a range of $0 \leq \omega_k(i) \leq 1$; it varies according to the fusion method selected for location tracking. Finally, the ϵ-Best selection module performs location tracking by returning the RP that maximizes the FS function to the client as the final location. If multiple maximum values are selected, the server randomly selects and returns the corresponding value.

7. Comparison of Techniques

In this section, we analyze the landmark-based technologies we have covered thus far. To identify their individual strengths and weaknesses, we evaluated them against six criteria, namely, accuracy, core technology, difficulty of exploration, cost, diversity, and robustness. The technologies are briefly described in Table II.

- Accuracy: Accuracy is one of the most important metrics for evaluating positioning technology or map construction, mainly because indoor positioning technologies require high accuracy. Hence, various approaches aiming to minimize errors in this perspective have been attempted. In this section, we present the variation in average error rate of the various devices with location based on experiments. As a singularity, among the three categories for the technologies based on visual images, accuracy is determined based on the recognition rate rather than the distance error.
- Core Technology: Each technology uses a specific infrastructure and algorithms to explore the landmarks. We compare the infrastructure and algorithms at the core of each technology, as well as the key applications of each of these technologies.
- Detection difficulty: The difficulty of landmark detection varies significantly according to the infrastructure and algorithms at the core of each technology. Regardless of how accurate the technique is, if the landmark detection is low, significant effort is required to build it in advance. We examine the various methods used to locate landmarks and briefly describe the difficulties encountered while locating them.
- Cost: Cost is another paramount consideration when evaluating landmark detection technologies. In addition to the cost associated with the core infrastructure, each system should be evaluated in

Table 2
Analysis for the reviewed landmark based indoor positioning technologies

	Accuracy	Core Technology	DetectionDifficulty	Cost	Versatile	Robust
Walkie-Markie	1.65 m average error inside 3600 m^2	Wi-Fi AP RSSI RTTP	Easy	Low infra cost High time cost	Yes	No
Geomagnetic Fingerprint	0.16 m RMS error inside 300 m^2	Magnetic fingerprint	Easy	Low time cost	Yes	No
CiFi	1.78 m mean error inside 100.8m^2	Wi-Fi AoA CSI Image	Normal	High time cost	Yes	Yes
ALIMC	1.6 m GDM and SDM error inside 2756 m^2	Motion sensor	Easy	Low infra cost High time cost	Yes	Yes
Tactile	2.1 m average error inside 54.25 m	Tactile/ Particle filter	Normal	High infra cost Low time cost	No	Yes
BatTracker	0.015 m maximum error inside 15 m^2	Reflected echo	Hard	Low infra cost High time cost	No	Yes
Luxapose	0.07 m average error inside 5240 m^2	On-off keying light	Normal	High infra cost Low time cost	No	No
Visual	96.6% matching rate	Deep learning	Hard	Low infra cost High time cost	Yes	No
Pair-Navi	98.6% matching rate	Peer-to-Peer	Easy	Low infra cost High time cost	Yes	Yes
JADE	0.5 m average error inside 300m^2	mmWave Virtual AP	Hard	High time cost	Yes	Yes
WAIPO	87.6% matching rate inside 4500 m^2	Multidata Selection	Hard	High personal cost High time cost	Yes	Yes
Wi-Lo	+8% on urban +25% on rural	Wi-Fi, LTE, BLE, Magnetic	Normal	Low infra cost	Yes	No

terms of the additional costs incurred while searching and creating the landmark map. Hence, we evaluated various approaches based on these costs. In this study, the cost items are categorized into preliminary investigation costs and additional infrastructure installation costs.

- Versatility: Each technology is built and tested under specific locations and in the presence of various other constraints, usually to optimize and demonstrate its technological superiority. However, the technologies fail to perform under more general and nonoptimal conditions as under the optimal experimental conditions. In this section, we examined the possibility of using each technology in a general environment without core modifications.
- Robustness: The indoor environment changes over time Furthermore, there can be restrictions in the use of technology for a variety of reasons, from simple structural changes to replacement owing to outdated infrastructure. We evaluate the capability of each technology to operate in a variable indoor environment without major problems.

The Walkie-Markie shows 1.65m average error in a 3600 m^2 space, using the RTTP of Wi-Fi AP RSSI as a landmark. It also uses its own Arturia algorithm to cluster the surveyed landmarks. In terms of search difficulty, the Walkie-Markie exhibits a relatively easy difficulty because it can be used as a landmark if the RSSI value of the Wi-Fi increases and decreases. In terms of the cost of installing additional infrastructure, the Walkie-Markie utilizes Wi-Fi, a widely installed infrastructure, and does not require additional infrastructure. On the order side, in order to collect data, Walkie-Markie conducted 10 experiments in which 7 users walked a total of 30 laps around the experimental site for about 2 hours. The Walkie must travel around the room in advance to investigate the Wi-Fi signal. Although it requires less effort than the fingerprint technique, the operator must keep moving until sufficient data is accumulated. This measurement method increases the manpower and time consumption. Therefore, the Walkie-Markie approach is not cost-efficient overall. In terms of versatility, the Walkie-Markie uses Wi-Fi, and as mentioned above, Wi-Fi is one of the most versatile network infrastructures. Because locating and using landmarks is not difficult, the Walkie-Markie is superior in terms of versatility. However, in terms of robustness, it suffers from similar drawbacks as the fingerprint method. If an indoor AP changes or is removed, the system cannot utilize the landmarks obtained through the AP. Hence, if a large number of indoor APs change, the Walkie-Markie must continuously update the landmarks.

The system deploying geomagnetic fingerprints showed 0.16 m RMS error in a 300 m^2 composite atrium. This system reduces fingerprint instability arising from changes in the indoor environment by setting the indoor magnetic field as a fingerprint instead of the Wi-Fi RSSI, which is a preferred method for many existing fingerprint technologies. Because the geomagnetic search in the system is principally similar to that of a normal compass, the system is ideal in terms of detection difficulty. They collected 120 magnetic fingerprint samples at each measured position by rotating the chair 360 degrees in a 28 cm radius at approximately constant speed for 12 s to support position recognition in all directions. The specific tests they ran didn't predict the orientation, so they didn't need to accurately measure the orientation while mapping, which saved significant time. With regard to versatility, the system uses geomagnetic signals that can be obtained in all buildings. It also uses an uncomplicated algorithm; therefore, it can be easily deployed in most buildings. However, geomagnetism is affected not only by steel structures or structures in buildings but also by electronic devices. The versatility of the system greatly deteriorates with an increase in the number of electronic devices in the environment. Apart from affecting the versatility, the presence of too many electronic devices and other disturbances in the room will affect the magnetic field, as well as the robustness of the system. This is because these devices will hinder the matching of the geomagnetism obtained through the fingerprint method with the geomagnetism measured by the user.

CiFi shows 1.78 m mean error at 100.8 m^2. They used 15 training positions and 15 test positions. They set the distance between two adjacent training positions to be 1.8 m, and collected CSI data in 1000 packets for each training and test position. It deployed the CSI of commercial 5 GHz Wi-Fi and DCNN for learning. It acquired the CSI data of Wi-Fi, converted it into an image, and used a deep learning algorithm to learn the image and increase the robustness and accuracy of the user's final location selection. Although it deploys CSI, instead of RSSI, detecting CSI data is not very difficult if there is only a device driver. In terms of cost, CiFi must convert the received CSI data into an image. This process increases the location accuracy but uses additional resources. In terms of versatility, the CiFi can be used anywhere Wi-Fi is installed, as long as the user's device has an open-source driver for the Wi-FI NIC. Therefore, it exhibits high versatility. In terms of robustness, it uses CSI

images instead of volatile RSSIs and deep learning algorithms for learning. Therefore, it exhibits good robustness and the location suggestion does not change significantly even after remeasurement at the same location.

The ALIMC shows 1.6 m Graph Discrepancy Metric (GDM) and Shape Discrepancy Metric (SDM) error in a 2756 m^2 space. They repeated each follow-up 10 times to evaluate the performance of ALIMC with incremental data. A total of 300 user trajectories were collected by 3 participants using three types of smartphones, and in terms of time, these trajectories correspond to 220 minutes of data collection. The data collected included acceleration data, compass data, gyroscope data, barometer data, and WiFi fingerprints. It identifies the structure of a building using the value of a motion sensor based on human behavior and uses it as a landmark candidate. It then clusters the landmark candidates in similar locations to create a unique landmark. As for the detection difficulty, it selects landmarks by understanding the building's structure using the values of the sensors that track the continuous movement of the user. Because the patterns recorded during the movement for each indoor structure are often similar and the detection of the user's movement is not challenging, we conclude that the search difficulty of the ALIMC is low. Another advantage is in terms of installation cost; it does not require additional infrastructure for location tracking because only the sensors and indoor Wi-Fi are required for the ALIMC landmark navigation. However, with regard to the preinvestigation costs, the ALIMC's is more expensive because moving indoors is required to investigate the landmark candidate points and extract the details as frequently as possible. Although prior knowledge of the indoor structural information can amortize the cost, the ALIMC requires a thorough investigation, such as a fingerprint technique. Therefore, it is not efficient in terms of time and human resources; rather, it is prohibitive owing to the preliminary investigation cost. In terms of versatility, the ALIMC is capable of locating landmarks if the structure involved is a building. The ALIMC is only challenging in rare cases when it is operating in complex environments and with no specific structures such as elevators, intersections, corners, and stairs. Therefore, it can generally be classified as highly versatile. It exhibits significant robustness because the structure of the building remains almost unchanged, except in the case of large construction or large-scale repairs. Therefore, the ALMIC landmarks, which are based on the structure of the building, do not change easily.

The tactile technique shows 2.1 m average error in a 54.25m hallway. Unlike other technologies, the technology based on the tactile approach obtains data from people with visual impairments and then applies a particle filter to create a landmark. In terms of detection difficulty, the system examines information that can be obtained through the tactile sense conveyed by the floor material as well as door and corridor intersections using a wand, as a visually impaired person would. Because most of these elements can be observed intuitively, the detection difficulty of the system is low. With regard to installing additional infrastructure, as mentioned earlier, the system requires factors that can only be achieved with tactile sensation. Although it is easy to install a sidewalk block for the visually impaired, not many buildings have such indoor infrastructures. Therefore, the system is prohibitive in terms of the cost of installing additional infrastructure. Compared to ALIMC, the system requires detailed investigation and incurs higher preinvestigation costs. Owing to the need to investigate the floor material, the system preinvestigation costs are high because it requires more time and human resources. In terms of versatility, as mentioned earlier, few buildings have the same infrastructure as tiles for the blind. Therefore, the system is not versatile unless it is a building with a specific purpose. In terms of robustness, because the structure of the building does not change easily and the floor material is very rarely changed, the landmark of the system that is based on contact with the floor exhibits high robustness.

The error of BatTracker is maximum 0.015 m at 15 m^2. They evaluated the effect of particle number on tracking accuracy and confirmed that the 90th percentile error continued to decrease from 5 cm to 1 cm and then was relatively stable when the number of particles increased from 500 to 1500. Also, for numbers less than 500, they only displayed results of $\geq$ 500 particles, as too few particles can easily lead to tracking failure. BatTracker aims to reduce the inertial sensor error using a built-in sound device that emits a sound signal that is reflected by the indoor structures of the room. It uses sound reflected from indoor elements, which may be assumed to lower the detection difficulty. However, there are various noises in the room, making it difficult for BatTracker to easily receive the reflected sound, even with noise-canceling technology. Therefore, BatTracker suffers from high detection difficulty. In terms of the cost of installing additional infrastructure, because BatTracker does not require an AP such as Wi-Fi, the only thing that it needs to locate landmarks is a device equipped with sound equipment and an indoor object that can reflect the sound. Consequently, BatTracker has a low cost in terms of installing additional infrastructure. In terms of preinvestigation costs, BatTracker conducts continuous acoustic signal reflection surveys. However, it is not necessary to perform the survey at all locations because of the structural conditions of these locations. Therefore, BatTracker require relatively less manpower and time in terms of preliminary investigations. Combining the above two aspects, BatTracker is not costly. In terms of versatility, for BatTracker to measure the distance using the reflected sound, the building must have a specific interior structure. Hence, BatTracker is less useful in large empty spaces, such as shopping mall lobbies or gyms, making it less efficient in terms of versatility. In terms of robustness, BatTracker consistently locates landmarks, and, regardless of the user's posture, it can provide matching information. In addition, the BatTracker landmark does not change easily because of landmark discrimination based on the structure of the building. Therefore, BatTracker exhibits high robustness.

Luxapose shows 0.07 m average error in a 5240 m^2 space. They integrated five LED landmarks, a smartphone and a Cloudlet server into an indoor positioning testbed. They attached a complementary pegboard to the floor, aligned using a laser sight, and verified with a vertical bob to create a 3D grid with 2.54 cm resolution of known locations for experimental evaluation. To separate localization from communication performance, the transmitter was set to emit pure tones in the 2kHz to 4kHz range with 500Hz separation to ensure reliable communication. This is the lowest error rate among the investigated techniques. Luxapose determines the user's location by examining the extracted image using a light bulb that is capable of on–off keying. In terms of detection difficulty, Luxapose detects landmarks using images obtained through illumination. Usually, visual detection has a high level of difficulty. However, Luxapose has a relatively high detection capability; it simply uses the presence or absence of light. In terms of the cost of installing additional infrastructure, Luxapose requires on–off keying, instead of normal lighting for location tracking. However, the lights installed in most buildings are not on–off keyed; hence, to utilize Luxapose, it is paramount to replace lighting or install additional ones; thus, Luxapose incurs more cost. Luxapose does not require any particularly complicated preinvestigation compared to other technologies. Based on these factors, even though the preinvestigation cost is not high, Luxapose is expensive because of the additional cost incurred for the installation of the necessary infrastructure. In terms of versatility, Luxapose can only be operated in the presence of on–off keyed lights. However, as mentioned earlier, only a few buildings are equipped with such lights, thus reducing the versatility of Luxapose. In terms of robustness, the greatest weakness that Luxapose suffers from is that it cannot be used in situations where lighting is absent as is often the case with emergency situations. Therefore, Luxapose exhibits relative robustness.

The visual sequence showed a matching rate of 96.6%. The visual sequence uses a camera to locate the landmarks. Experimental participants walked along five paths observing different types and landmarks. Each route included 10 to 17 landmarks. An Honor Android phone was

held on the participant's arm viewed from the side while the participant was walking. A total of 5 images were obtained by collecting images for each route. Of the 5 videos, the first 2 are with the camera on the left arm and the last 3 with the camera on the right arm. Thereafter, the system uses a visual sequence to learn the image pattern using deep learning and matches the image captured by the user to determine the user's current location. In terms of detection difficulty, if the structure of the indoor structures is similar because of the characteristics of the system using the image, the system has great difficulty in landmark identification. The system attempts to overcome this problem through deep learning. However, this problem cannot be entirely solved; therefore, the detection difficulty of the system is high. The system can optionally install additional structures to improve image matching; hence, the cost of installing the additional infrastructure of the system is low. In terms of the pre-inspection cost, the system must store images in advance. However, as the system needs to take only the images of the necessary structures, the preinvestigation cost is relatively low. When both costs are aggregated, the required cost of the system decreases. In terms of versatility, the system can be used in an indoor environment with specific structures; hence, it has high versatility. In terms of robustness, the system cannot be used in environments with invisible structures owing to various factors such as power outages. In addition, even if deep learning continues, it becomes difficult to match the images in buildings where the interior design is very different. Therefore, the system exhibits poor robustness.

Pair-Navi exhibits an error rate of 98.6% in instant matching and 83.4% in the same space after two weeks. They designed 21 search paths, including 6 short paths ($\leq$ 100 m), 7 medium paths (100 m - 200 m) and 8 long paths ($\geq$ 200 m) covering all major paths in the test area. At the same time, they recruited four volunteer followers so that they could naturally follow a different path as usual. The system receives and utilizes images from multiple users through P2P navigation using the images of highly variable indoor environments. In terms of detection difficulty, Pair-Navi receives images from an unspecified number of people. Pair-Navi can find structures that can be used in various received images and use them as landmarks. For this reason, Pair-Navi exhibited low detection difficulty. In terms of cost, it has a similar tendency as technology exploiting visual sequences. Because no additional infrastructure is required and Pair-Navi allows image updates through followers, it is cost effective. In terms of versatility, Pair-Navi can be used anywhere in an indoor environment with defined structures, making it more versatile. Pair-Navi updates the images sent by the leaders using the follower's images. If the indoor environment does not change completely, Pair-Navi can maintain high robustness by continuously updating the images.

The average error of JADE was 0.5 m in a 300 m^2 space. It uses the physical AP of the mmWave, as well as the virtual AP made by mirroring the physical AP as a landmark. JADE traces the user's location by substituting the location of the input physical and virtual APs into the geometric expression. In terms of search difficulty, JADE is able to use the location of these two types of APs as landmarks. Therefore, it can be deemed effective in terms of the search difficulty metric. However, because both the physical and virtual APs require information on the indoor structure in advance (the physical AP is the location of the corresponding AP and the virtual AP is the location of the indoor boundary), JADE struggles greatly with searching for the landmarks if there is no information about the indoor structure. In terms of the cost associated with installing the additional infrastructure, JADE requires an unfamiliar AP called mmWave, which incurs extra installation cost. With respect to preinvestigation cost, JADE requires additional cost in investigating the interior structure by walking around, especially when there is no prior information available. Hence, the larger the building, the costlier the preliminary investigation. Moreover, in terms of the cost of installing additional infrastructure, the core element of the system is based on geomagnetic structures available in any building; therefore, it is not necessary to install additional infrastructure. However, in terms of preinvestigation costs, the system requires the manufacture of a special compass for investigation. In addition, because of the nature of fingerprint technology, irradiation must be performed in all the cells. Therefore, the system is labor intensive and time consuming. With regard to versatility, JADE uses landmarks to locate the user in a geometric way; hence, it can be used anywhere with sufficient mathematical knowledge. In addition, if there is no change in a plurality of APs in terms of robustness, the user's location can be continuously searched using the JADE geometry. Thus, JADE is efficient in terms of versatility and robustness.

The WAIPO showed a matching accuracy rate of 87.6% in a space of 4500 m^2. It uses Wi-Fi, photos, magnetic fingerprinting, and ULP learning to track the user's location. They conducted an experiment with the help of 51 volunteers for photo crowdsourcing. They detected 89 APs in the 11 story building they tested. They divided each room in half according to the length of the room and gathered the Wi-Fi signal strength in the center of each section to create a basic fingerprint. As a result, they obtained a total of 50 raw fingerprints from 25 rooms. They randomly collected Wi-Fi signals from room to room, repeated 20 times over 10 days. It uses the execution module of the client application to collect three types of fingerprints: Wi-Fi, photos, and magnetic fields. WAIPO uses a system timer during the collection process to adjust the energy consumption of the modules. It uses the server's location tracking process module to estimate the user's final location, which is calculated by the WAIPO's process module based on the collected data through four steps: potential location calculation, personal matching, ULP learning, and magnetic field calibration. In terms of detection difficulty, WAIPO uses different types of complex elements as landmarks, rather than as a single source. The use of complex data elements increases the difficulty of collecting data for location tracking. In terms of cost, WAIPO must process all data simultaneously while collecting various data. In the data collection process and using a client application, WAIPO aims for energy efficiency by using a system timer. However, simultaneously processing multiple data requires a higher cost than other technologies. In terms of versatility, WAIPO has a structure that can be distinguished by photographs and used anywhere an indoor wireless signal can be captured. Hence, WAIPO is versatile. In terms of robustness, WAIPO uses multiple factors to determine a user's location. This helps correct errors by using other elements, even if an error occurs in one element. Fast error correction using complex elements makes WAIPO highly robust.

Wi-Lo showed an 8% improvement in performance compared to existing technology in the established urban area and 25% in rural areas. After running more than 30 tests for each scenario and algorithm, both urban and rural, they calculated the average accuracy of detecting the RP where the user is. Experimental results, although performance decreases significantly from urban to rural scenarios, they all confirm that Wi-Fi wireless fingerprinting is more accurate on average compared to LTE and MAG, which are slightly affected. Wi-Lo tracks the user's location using a wireless signal selected by the user from Wi-Fi, LTE, BLE, and magnetic. Wi-Lo showed high accuracy in both high-density and low-density AP environments through the simultaneous use of four wireless signals. In terms of detection difficulty, the four wireless signals used by Wi-Lo have been widely used in the past. Therefore, it is not difficult to detect these signals. The cost implications of WAIPO depend on the signal selected by the user; however, Wi-Lo uses a combination of various signal data and consumes resources in collecting and analyzing each. In terms of versatility, because the signals used by Wi-Lo are not difficult to detect and have been widely used, there are many cases where the existing structures are installed and detected. Therefore, Wi-Lo can be used in various buildings. Even if various wireless signals are used simultaneously, wireless signals may be unstable, depending on the environment. Furthermore, even signals measured in the same place may vary greatly. Therefore, Wi-Lo has weak robustness.

8. Future Research

In this section, we present future research direction based on additional factors that researchers should consider when developing indoor positioning technology, and then we provide future advance direction focusing areas to which the indoor positioning technology is applied.

8.1. Future research directions

The first element is the application of technology from other fields, which is related to the accuracy of the system. Before the recent improvements in the performance of mobile devices, other fields used location estimation algorithms for indoor positioning technology. Robotics has received significant research attention [88,89]. It is necessary to determine a robot's position as it move indoors. However, the robot does not possess any previous information about the indoor space, making it less successful than humans at navigating its surroundings. Jiang et al. proposed a SLAM framework based on new graph optimization through a combination of low-cost light detection and range (LiDAR) [90,91] sensors and visual sensors [92]. The SLAM method using a LiDAR sensor is generally adopted for robot navigation. However, the cost limitations for consumer robots cause means that they use low-cost sensors. The poor performance of low-cost LiDAR can lead to errors that accumulate quickly during the SLAM process, which grow with larger maps. To resolve these concerns, the SLAM system introduces a new mechanism that considers both scan data and image data, and a bag-of-words model [93,94] with visual features was used to loop close detection. In addition, a fast relocation approach using a 2.5D map and a diagram that presented both obstacles and vision functions was proposed in the study. It was asserted that the proposed method outperformed the LiDAR or camera alone. Additionally, researchers has conducted various research in this direction. [95,96]

The second element to consider is security, which relates to the core technology of the system. An indoor positioning system utilizes various information categories. In an indoor positioning system, the position of a particular person can be tracked using personal movement information. Indoor positioning systems that use visual images may accidentally include private information or scenes. If such a situation is not properly managed, private information may be leaked, creating the risk of intrusion. Effective security technology is needed when indoor positioning systems process user information because these leakages can have serious adverse outcomes. Konstantinidis et al. suggested a temporary vector map (TVM) algorithm to constantly provide granular location updates while shielding consumers from violation of privacy by services [97]. The proposed TVM algorithm helps users to obtain accurate positioning by utilizing k-Anonymous Bloom (kAB) filters and best neighbors generators [98] in camouflaged localization requests, both of which can flexibly deflect various privacy attacks. Through empirical assessment and experiment, it was found that the TVM algorithm is historically not vulnerable to attacks that compromise k-anonymous security. It requires approximately four times less energy and message positioning than other competitive approaches. Including these studies, many researchers are conducting various studies for data privacy of users. [99,100]

The third element to consider is the filter, which is related to the detection difficulty of landmarks. The accuracy of the landmarks generated was boosted when the indoor data were reviewed for errors. However, it is very difficult for a person to review all the data during the system operation. To this end, several algorithms that automatically collect data and convert them into landmarks have been developed. However, various factors cause the data collected indoors to be prone to degradation. If these degraded data are used by the system to generate the landmarks, there will be a high risk of generating flawed landmark locations. Therefore, the system must modify or eliminate compromised data that has deteriorated owing to noise; most indoor positioning systems use filters to eliminate noise. Huu Quoc Dong Tran et al. used an extended Kalman filter (EKF) [101,102] in a mobile robot positioning system and achieved better results [103]. The EKF is a mathematical tool that can estimate variables in a wide range of processes for nonlinear systems. It functions by linearizing the nonlinear state dynamics and measurement models. When the data are measured and integrated into position estimates, the EKF recognizes and prevents situations in which the data are mixed or ignored and does not affect location estimates. The EKF amortizes the noise effects of current estimates by integrating more information from trusted data than from unreliable data. The users can observe the amount of noisy data that is collected through the extended knife-only filter and the filter estimates the location by considering the noise of the data. The study demonstrated that the EKF algorithm is the most basic and general solution for discontinuous, low-accuracy signals. Techniques using various filters for data refinement have been developed. [104,105]

The fourth element to consider is the calculation time, which is also related to the cost of the system. The nature of the indoor positioning algorithm requires real-time recurrent calculations, and the performance of indoor positioning technology varies significantly, depending on the calculation time. If the calculation time is long, the positioning system must consume additional resources, such as memory or time, to calculate the location. If a disproportionate amount of resources is consumed in calculating a single location, the system fails to search for new landmarks or update the user's location, leading to performance degradation. Kawaji et al. proposed an image-based indoor positioning system using panoramic images with additional location information [106]. They propose an indoor positioning system that includes rapid database construction, image matching, and a visualization interface using a forward camera. The system estimates the user's position within a short time and with high accuracy by combining a robust image matching process based on the integration of a principal component analysis-scale invariant feature transform (PCA-SIFT) descriptor [107] with a high-speed similarity search algorithm based on locality sensitive hashing (LSH) [108]. LSH is a probabilistic nearest-neighbor search method that uses a hash function. When using more than 20 vectors, LSH is faster than the other algorithms. The system also enhanced accuracy by introducing a "confidence" parameter in image matching. The indoor positioning system can speed up positioning by adding special functions or parameters to the algorithms. In addition to this, indoor location positioning technologies that focus on providing quick information, including shortening the calculation time, have been developed [109,110].

The fifth element to consider is device diversification, which is related to the versatility of the system. The versatility of the system is affected by other factors other than changes in the indoor environment. The diversification of mobile devices also affects the versatility and robustness of the system. There is an ever-increasing number of mobile manufacturers and mobile brands (e.g., smartphones, smartwatches, and smart glasses). Crucially, even if the parts to be mounted for each device are produced in batches, the mobile devices may exhibit various operational deviations. Therefore, even if the technology exhibits good performance with a specific experimental device, the results may change if the experiment was conducted using another device. Various indoor positioning technologies have been developed in consideration of this for use in smartphones between different models. [112,113] Laoudias et al. proposed an approach that uses crowdsourcing techniques to produce Wi-Fi fingerprint maps and solved the device diversity problems that arise in the process [111]. Crowdsourcing is a promising paradigm that can help solve difficult problems through the help of the public without relying on trained professionals.

Crowdsourcing is attractive because the burden of production is shared by the public. However, it poses new challenges such as contaminated data or incomplete filtering, uneven fingerprint distribution processing, radio map size management, and device handling. The study was aimed at solving the problem of device heterogeneity that naturally occurs in crowdsourcing fingerprint systems owing to the use

of various mobile devices such as users' smartphones, personal digital assistants, tablets, and laptops. Mobile devices do not transmit RSS values to the system in a similar manner and because measurements between devices are incompatible, it is very difficult to fuse location tag RSS values from other devices into a single radio signal map. The researchers suggest using differential fingerprints, or fingerprints that include differences in RSS, instead of absolute RSS values, as a method of combining RSS data from different models. The authors proved experimentally that the new device can be positioned using the corresponding crowdsourced radio signal map and measured RSS values at unknown locations.

The final element to consider is the introduction of artificial intelligence (AI) in relation to the robustness of the system. Continuous changes in the environment reduce the life span of the landmarks. Technologies that use visual images as landmarks are sensitive to alterations in the indoor environment and the system on which it is based can be rendered ineffective by merely removing the existing structures, changing interior fittings, or installing event decorations. To respond to changes in the indoor environment, the system can leverage AI capabilities to continuously update landmarks. The addition of an AI analytical component to an algorithm helps the system to observe the changes in the indoor environment and effectively modify the landmarks accordingly. Through this process, the system increases the life span of a landmark. Zou et al. proposed an indoor positioning algorithm based on an online sequential extreme learning machine (OS-ELM) [114] to compensate for the shortcomings of Wi-Fi fingerprinting technology [115]. The authors noted that fingerprint technology suffers from excessive mapping costs and weak environmental dynamics, which they solved using the OS-ELM. Originating from extreme learning machines (ELMs), the OS-ELM inherits the advantages of ELM—excellent generalization performance along with very fast learning speeds. The OS-ELM has online sequential learning capabilities; therefore, no retraining is required when receiving new data. Unlike other online sequential learning algorithms, such as stochastic gradient descent back-propagation [116] and growing and pruning radial basis function network [117], which require specific types of hidden nodes and process data sequentially chunkwise, the OS-ELSM can adapt to different types of hidden nodes. Therefore, Zou et al. enabled OS-ELM to collect and update Wi-Fi RSS fingerprints more flexibly. Moreover, they demonstrated that the fast learning speed of OS-ELM can reduce the time and labor cost of the offline site investigation and the online sequential learning ability of the learning system helps the proposed positioning algorithm adapt to environmental dynamics promptly. Indoor location positioning technology using artificial intelligence have been actively researched [118,119].

8.2. Applications

Indoor positioning can be used in various applications, where certain aspects of the technology should be strengthened. This section introduces the four representative applications of indoor positioning, advance directions to be strengthen, factors to be considered when the technology is used, and key elements to solve it. Table 3 shows applications of indoor positioning technologies

First applications are the epidemiological investigations of confirmed cases in the pandemic situation such as COVID19 and the real-time monitoring of criminals, where important advance direction of the technology is the real-time movement detection [120,121]. In the pandemic situation, the real-time indoor positioning system is a great help for the quarantine of confirmed cases. The quarantine manager can identify the real-time movement of the confirmed person and send a quick notification to users who overlap with the movement of the confirmed person. In the real-time monitoring of criminals, the positioning technology continuously monitors criminals wearing electronic anklets, and at the same time, it can be used to help arrest criminals when they escape. In these applications, the important factor that the technology should consider is the privacy of the user. The technology should focus on security so that users' movement data should not be leaked, and at the same time it can provide users with data of which personal information is obscured.

Table 3
Applications of indoor positioning technologies

Application	Advance direction	Factor	Key element
Disease epidemiologic investigation Criminal real-time monitoring	Real-time movement detection	Privacy of user	Security
Public places guide Landmark tour guide	Location guide in many people constantly visit	Reliability of location data	Filter
Indoor event hall guide Sports venue guide	Location guide in a short period of time	Time it takes to transmit location data to many unspecified people.	Calculation time
Shopping mall guide	Location guide in a frequently changing environment	Shortening the lifespan of landmarks	Artificial intelligence

Second application is the location guidance in public places or tourist destinations where many people constantly visit [122,123]. The important advance direction is the supply of correct information. The important factor for technology to consider is the reliability of the location data collected from many unspecified users. Some of them may be correct, but others may be unclear due to noise. If the system reflects the data as it is to the landmark, due to the degraded data with noise, there is a high risk of a defective landmark. The technology should focus on filters to remove uncertain data to maintain landmark accuracy.

Third application is the location guidance in indoor event halls or sports venues where a large number of people gather for a short period of time [124,125]. The important advance direction is the location guidance for a short period of time. In this application, the factor for technology to consider is the time it takes to transmit location data to many unspecified people. Information provided to those who enjoy the event within a limited time should be prompt. The technology should focus on calculation time to quickly deliver information to a large number of users.

Final application is the shopping mall where store locations and information change according to seasons, contracts, and events [126,127]. The important advance direction is the location guide in a frequently changing environment. In this advance direction, the factor for technology to consider is shortening the lifespan of landmarks. The technology can focus on artificial intelligence to update landmarks automatically with low time and update cost

9. Conclusion

With the performance of mobile devices and the time spent indoors increasing, many location-based services and applications have been developed. For such services and applications to become widespread, the importance of indoor positioning technology that accurately detects the user's current location is emphasized. However, indoor positioning technologies based on sensor data, wireless signals, and visual images tend to have the problems of error accumulation and maintenance difficulty. To solve these problems, researchers have studied indoor positioning technologies that use indoor spots with distinctive features as landmarks. In this study, the development of indoor positioning technology using landmarks was investigated, and the technologies were classified into four categories according to the data source utilized, namely wireless signals, sensor data, visual images, and multi-sources. We introduced landmark search methods and algorithms deployed by

the technologies. We compared the technologies based on seven evaluation criteria: accuracy, core technology, detection difficulty, cost, versatility, and robustness. We also mapped future research directions. This study provides useful viewpoints and necessary information about landmark-based indoor positioning technology.

CRediT authorship contribution statement

Beakcheol Jang: Conceptualization, Supervision, Writing – original draft, Writing – review & editing, Project administration. **Hyunjung Kim:** Visualization, Investigation, Formal analysis, Writing – original draft. **Jong wook Kim:** Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Nationam Research Foundation of Korea Fund of NRF2022R1F1A1063961.

References

[1] D. Zhang, F. Xia, Z. Yang, L. Yao, W. Zhao, Localization technologies for indoor human tracking, in: 2010 5th international conference on future information technology, 2010, pp. 1–6.
[2] H. Liu, H. Darabi, P. Banerjee, J. Liu, Survey of wireless indoor positioning techniques and systems, IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 37 (6) (2007) 1067–1080.
[3] P. Misra and P. Enge, "Global Positioning System: signals, measurements and performance second edition," Global Positioning System: Signals, Measurements And Performance Second Editions, vol. 206, 2006.
[4] S.K. Morley, et al., Energetic particle data from the global positioning system constellation, Space Weather 15 (2) (2017) 283–289.
[5] X. Wang, L. Gao, S. Mao, S. Pandey, CSI-based fingerprinting for indoor localization: A deep learning approach, IEEE Transactions on Vehicular Technology 66 (1) (2016) 763–776.
[6] K. Chintalapudi, A.Padmanabha Iyer, V.N. Padmanabhan, Indoor localization without the pain, in: Proceedings of the sixteenth annual international conference on Mobile computing and networking, 2010, pp. 173–184.
[7] I. Amundson, X.D. Koutsoukos, A survey on localization for mobile wireless sensor networks, in: International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments, 2009, pp. 235–254.
[8] N.B. Priyantha, H. Balakrishnan, E.D. Demaine, S. Teller, Mobile-assisted localization in wireless sensor networks, Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies 1 (2005) 172–183.
[9] W. Kang, Y. Han, SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization, IEEE Sensors journal 15 (5) (2014) 2906–2916.
[10] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, F. Zhao, A reliable and accurate indoor localization method using phone inertial sensors, in: Proceedings of the 2012 ACM conference on ubiquitous computing, 2012, pp. 421–430.
[11] M. Luna, G. Meifeng, Z. Xinxi, Z. Yongjian, S. Mingliang, An indoor pedestrian positioning system based on inertial measurement unit and wireless local area network, in: 2015 34th Chinese Control Conference (CCC), 2015, pp. 5419–5424.
[12] Z. Yang, C. Wu, Y. Liu, Locating in fingerprint space: wireless indoor localization with little human intervention, in: Proceedings of the 18th annual international conference on Mobile computing and networking, 2012, pp. 269–280.
[13] S. He, S.-H.G. Chan, Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons, IEEE Communications Surveys & Tutorials 18 (1) (2015) 466–490.
[14] N. Ravi, P. Shankar, A. Frankel, A. Elgammal, L. Iftode, Indoor localization using camera phones, in: Seventh IEEE Workshop on Mobile Computing Systems & Applications (WMCSA'06 Supplement), 2005, pp. 1–7.
[15] Z. Chen, H. Zou, H. Jiang, Q. Zhu, Y.C. Soh, L. Xie, Fusion of WiFi, smartphone sensors and landmarks using the Kalman filter for indoor localization, Sensors 15 (1) (2015) 715–732.
[16] Q. Wang, H. Luo, F. Zhao, W. Shao, An indoor self-localization algorithm using the calibration of the online magnetic fingerprints and indoor landmarks, in: 2016 international conference on indoor positioning and indoor navigation (IPIN), 2016, pp. 1–8.
[17] Z.-A. Deng, G. Wang, D. Qin, Z. Na, Y. Cui, J. Chen, Continuous indoor positioning fusing WiFi, smartphone sensors and landmarks, Sensors 16 (9) (2016) 1427.
[18] N. Simon, et al., Indoor localization system for emergency responders with ultra low-power radio landmarks, in: 2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings, 2015, pp. 309–314.
[19] A. Bekkali, H. Sanson, M. Matsumoto, RFID indoor positioning based on probabilistic RFID map and Kalman filtering, in: Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007), 2007, 21–21.
[20] B. Gozick, K.P. Subbu, R. Dantu, T. Maeshiro, Magnetic maps for indoor navigation, IEEE Transactions on Instrumentation and Measurement 60 (12) (2011) 3883–3891.
[21] F. Gu, K. Khoshelham, J. Shang, F. Yu, Sensory landmarks for indoor localization, in: 2016 Fourth International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services (UPINLBS), 2016, pp. 201–206.
[22] P. Nazemzadeh, D. Fontanelli, D. Macii, Optimal placement of landmarks for indoor localization using sensors with a limited range, in: 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2016, pp. 1–8.
[23] N. Rajagopal, P. Lazik, A. Rowe, Visual light landmarks for mobile devices, in: IPSN-14 proceedings of the 13th international symposium on information processing in sensor networks, 2014, pp. 249–260.
[24] A.M.G. Pinto, A.P. Moreira, P.G. Costa, Indoor localization system based on artificial landmarks and monocular vision, Telkomnika 10 (4) (2012) 609.
[25] H. Abdelnasser, et al., SemanticSLAM: Using environment landmarks for unsupervised indoor localization, IEEE Transactions on Mobile Computing 15 (7) (2015) 1770–1782.
[26] H. Tang, F. Xue, T. Liu, M. Zhao, C. Dong, Indoor Positioning Algorithm Fusing Multi-Source Information, Wireless Personal Communications 109 (4) (2019) 2541–2560.
[27] Y. Gu, M. Chen, F. Ren, J. Li, HED: Handling environmental dynamics in indoor WiFi fingerprint localization, in: 2016 IEEE wireless communications and networking conference, 2016, pp. 1–6.
[28] P. Kriz, F. Maly, T. Kozel, Improving indoor localization using bluetooth low energy beacons, Mobile Information Systems 2016 (2016).
[29] A. Yassin, Y. Nasser, A.Y. Al-Dubai, M. Awad, MOSAIC: Simultaneous localization and environment mapping using mmWave without a-priori knowledge, IEEE Access 6 (2018) 68932–68947.
[30] P. Spachos, K.N. Plataniotis, BLE beacons for indoor positioning at an interactive IoT-based smart museum, IEEE Systems Journal 14 (3) (2020) 3483–3493.
[31] M. Nagah Amr, H.M. El Attar, M.H. Abd El Azeem, H. El Badawy, An Enhanced Indoor Positioning Technique Based on a Novel Received Signal Strength Indicator Distance Prediction and Correction Model, Sensors 21 (3) (2021) 719.
[32] A.K. Taşkan, H. Alemdar, Obstruction-Aware Signal-Loss-Tolerant Indoor Positioning Using Bluetooth Low Energy, Sensors 21 (3) (2021) 971.
[33] J. Chung, M. Donahoe, C. Schmandt, I.-J. Kim, P. Razavai, M. Wiseman, Indoor location sensing using geo-magnetism, in: Proceedings of the 9th international conference on Mobile systems, applications, and services, 2011, pp. 141–154.
[34] G. Shen, Z. Chen, P. Zhang, T. Moscibroda, Y. Zhang, Walkie-Markie: Indoor pathway mapping made easy, in: Presented as part of the 10th ${$USENIX$}$ Symposium on Networked Systems Design and Implementation (${$NSDI$}$ 13), 2013, pp. 85–98.
[35] X. Wang, X. Wang, S. Mao, CiFi: Deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi, in: 2017 IEEE International Conference on Communications (ICC), 2017, pp. 1–6.
[36] L.-H. Chen, E.H.-K. Wu, M.-H. Jin, G.-H. Chen, Intelligent fusion of Wi-Fi and inertial sensor-based positioning systems for indoor pedestrian navigation, IEEE Sensors Journal 14 (11) (2014) 4034–4042.
[37] C. Medina, J.C. Segura, A. De la Torre, Ultrasound indoor positioning system based on a low-power wireless sensor network providing sub-centimeter accuracy, Sensors 13 (3) (2013) 3501–3526.
[38] T. Gallagher, E. Wise, B. Li, A.G. Dempster, C. Rizos, E. Ramsey-Stewart, Indoor positioning system based on sensor fusion for the blind and visually impaired, in: 2012 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2012, pp. 1–9.
[39] B. Zhou, Q. Li, Q. Mao, W. Tu, X. Zhang, L. Chen, ALIMC: Activity landmark-based indoor mapping via crowdsourcing, IEEE Transactions on Intelligent Transportation Systems 16 (5) (2015) 2774–2785.
[40] E. Apostolopoulos, N. Fallah, E. Folmer, K.E. Bekris, Feasibility of interactive localization and navigation of people with visual impairments, in: Proceedings of the 11th IEEE Intelligent Autonomous Systems (IAS-10), Ottawa, ON, Canada, 2010, pp. 22–32.
[41] B. Zhou, M. Elbadry, R. Gao, F. Ye, BatTracker: High precision infrastructure-free mobile device tracking in indoor environments, in: Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, 2017, pp. 1–14.
[42] G. Conte, P. Doherty, An integrated UAV navigation system based on aerial image matching, in: 2008 IEEE Aerospace Conference, 2008, pp. 1–10.
[43] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, P. Dutta, Luxapose: Indoor positioning with mobile phones and visible light, in: Proceedings of the 20th annual international conference on Mobile computing and networking, 2014, pp. 447–458.
[44] Q. Li, J. Zhu, T. Liu, J. Garibaldi, Q. Li, G. Qiu, Visual landmark sequence-based indoor localization, in: Proceedings of the 1st Workshop on Artificial Intelligence and Deep Learning for Geographic Knowledge Discovery, 2017, pp. 14–23.
[45] E. Dong, J. Xu, C. Wu, Y. Liu, Z. Yang, Pair-Navi: Peer-to-Peer Indoor Navigation with Mobile Visual SLAM, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, 2019, pp. 1189–1197.
[46] J. Palacios, P. Casari, J. Widmer, JADE: Zero-knowledge device localization and environment mapping for millimeter wave systems, in: IEEE INFOCOM 2017-IEEE Conference on Computer Communications, 2017, pp. 1–9.

[47] F. Gu, J. Niu, L. Duan, WAIPO: A fusion-based collaborative indoor localization system on smartphones, IEEE/ACM Transactions on Networking 25 (4) (2017) 2267–2280.

[48] M. Di Felice, C. Bocanegra, K.R. Chowdhury, WI-LO: Wireless indoor localization through multi-source radio fingerprinting, in: 2018 10th International Conference on Communication Systems & Networks (COMSNETS), 2018, pp. 305–311.

[49] Y. Ye, B. Wang, X. Deng, L.T. Yang, On solving device diversity problem via fingerprint calibration and transformation for RSS-based indoor localization system, in: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), 2017, pp. 1–8.

[50] S. Kaiser, C. Lang, Detecting elevators and escalators in 3D pedestrian indoor navigation, in: 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2016, pp. 1–6.

[51] M. Ma, Q. Song, Y. Gu, Z. Zhou, Use of magnetic field for mitigating gyroscope errors for indoor pedestrian positioning, Sensors 18 (8) (2018) 2592.

[52] S. He, S.-H.G. Chan, Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons, IEEE Communications Surveys & Tutorials 18 (1) (2015) 466–490.

[53] J. Xiao, Z. Zhou, Y. Yi, L.M. Ni, A survey on wireless indoor localization from the device perspective, ACM Computing Surveys (CSUR) 49 (2) (2016) 1–31.

[54] A. Yassin, et al., Recent advances in indoor localization: A survey on theoretical approaches and applications, IEEE Communications Surveys & Tutorials 19 (2) (2016) 1327–1346.

[55] B. Jang, H. Kim, Indoor positioning technologies without offline fingerprinting map: A survey, IEEE Communications Surveys & Tutorials 21 (1) (2018) 508–525.

[56] F. Zafari, A. Gkelias, K.K. Leung, A survey of indoor localization systems and technologies, IEEE Communications Surveys & Tutorials 21 (3) (2019) 2568–2599.

[57] J. Gubbi, R. Buyya, S. Marusic, M. Palaniswami, Internet of Things (IoT): A vision, architectural elements, and future directions, Future generation computer systems 29 (7) (2013) 1645–1660.

[58] J. Tomlinson, Beyond connection: Cultural cosmopolitan and ubiquitous media, International Journal of Cultural Studies 14 (4) (2011) 347–361.

[59] P. Rong, M.L. Sichitiu, Angle of arrival localization for wireless sensor networks, in: 2006 3rd annual IEEE communications society on sensor and ad hoc communications and networks 1, 2006, pp. 374–382.

[60] P.F. DeCarlo, et al., Field-deployable, high-resolution, time-of-flight aerosol mass spectrometer, Analytical chemistry 78 (24) (2006) 8281–8289.

[61] S. Gezici, et al., Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks, IEEE signal processing magazine 22 (4) (2005) 70–84.

[62] Xiansheng Guo, et al., A survey on fusion-based indoor positioning, IEEE Communications Surveys & Tutorials 22 (1) (2019) 566–594.

[63] M. Ridolfi, A. Kaya, R. Berkvens, M. Weyn, W. Joseph, E.D. Poorter, Self-calibration and Collaborative Localization for UWB Positioning Systems: A Survey and Future Research Directions, ACM Computing Surveys (CSUR) 54 (4) (2021) 1–27.

[64] A. Motroni, A. Buffi, P. Nepa, A survey on indoor vehicle localization through RFID technology, IEEE Access 9 (2021) 17921–17942.

[65] S.-J. Yu, S.-S. Jan, D.S. De Lorenzo, Indoor navigation using Wi-Fi fingerprinting combined with pedestrian dead reckoning, in: 2018 IEEE/ION Position, Location and Navigation Symposium (PLANS), 2018, pp. 246–253.

[66] R. Dutta, D. Smith, R. Rawnsley, G. Bishop-Hurley, J. Hills, Cattle behaviour classification using 3-axis collar sensor and multi-classifier pattern recognition, in: SENSORS, 2014 IEEE, 2014, pp. 1272–1275.

[67] Y. Sasaki, T. Shibuya, K. Ito, H. Arimura, Efficient Approximate 3-Dimensional Point Set Matching Using Root-Mean-Square Deviation Score, IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences 102 (9) (2019) 1159–1170.

[68] P. Liu, P. Yang, W.-Z. Song, Y. Yan, X.-Y. Li, Real-time identification of rogue WiFi connections using environment-independent physical features, in: IEEE INFOCOM 2019-IEEE Conference on Computer Communications, 2019, pp. 190–198.

[69] S. Yousefi, H. Narui, S. Dayal, S. Ermon, S. Valaee, A survey on behavior recognition using wifi channel state information, IEEE Communications Magazine 55 (10) (2017) 98–104.

[70] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, S. Liu, Towards better analysis of deep convolutional neural networks, IEEE transactions on visualization and computer graphics 23 (1) (2016) 91–100.

[71] M. Kalantari, M. Nechifor, Accuracy and utility of the Structure Sensor for collecting 3D indoor information, Geo-spatial information science 19 (3) (2016) 202–209.

[72] S. Niwattanakul, J. Singthongchai, E. Naenudorn, S. Wanapu, Using of Jaccard coefficient for keywords similarity, in, Proceedings of the international multiconference of engineers and computer scientists 1 (6) (2013) 380–384.

[73] T.V. Mataró, et al., An assistive mobile system supporting blind and visual impaired people when are outdoor, in: 2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI), 2017, pp. 1–6.

[74] M. Khalaf-Allah, K. Kyamakya, Database correlation using bayes filter for mobile terminal localization in GSM suburban environments, in: 2006 IEEE 63rd Vehicular Technology Conference 2, 2006, pp. 798–802.

[75] J.-S. Gutmann, D. Fox, An experimental comparison of localization methods continued, in: IEEE/RSJ International Conference on Intelligent Robots and Systems 1, 2002, pp. 454–459.

[76] H. Nakano, et al., Monitoring sound to quantify snoring and sleep apnea severity using a smartphone: proof of concept, Journal of Clinical Sleep Medicine 10 (1) (2014) 73–78.

[77] G.M. Gibson, et al., Reversal of orbital angular momentum arising from an extreme Doppler shift, Proceedings of the National Academy of Sciences 115 (15) (2018) 3800–3803.

[78] H. Yamaguchi, Doppler Migration Estimation for a Complex Moving Target in Low Signal to Noise Ratio Environment, in: 2018 Progress in Electromagnetics Research Symposium (PIERS-Toyama), 2018, pp. 2378–2383.

[79] J.A. Moorer, A note on the implementation of audio processing by short-term Fourier transform, in: 2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA), 2017, pp. 156–159.

[80] J. Bao, F. Xing, T. Sun, Z. You, CMOS imager non-uniformity response correction-based high-accuracy spot target localization, Applied optics 58 (16) (2019) 4560–4568.

[81] R. Sun, M. Falahati, L. Li, Numerical and experimental study on multiphase printing of polymeric biconvex micro lenses, Journal of Micromechanics and Microengineering 28 (11) (2018), 115005.

[82] J. Ding, B. Chen, H. Liu, M. Huang, Convolutional neural network with data augmentation for SAR target recognition, IEEE Geoscience and remote sensing letters 13 (3) (2016) 364–368.

[83] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," arXiv preprint arXiv:1602.07360, 2016.

[84] C. Kerl, J. Sturm, D. Cremers, Dense visual SLAM for RGB-D cameras, in: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2013, pp. 2100–2106.

[85] H. Strasdat, A.J. Davison, J.M. Montiel, K. Konolige, Double window optimisation for constant time visual SLAM, in: 2011 international conference on computer vision, 2011, pp. 2352–2359.

[86] S. Akoum, O. El Ayach, R.W. Heath, Coverage and capacity in mmWave cellular systems, in: 2012 conference record of the forty sixth Asilomar conference on signals, systems and computers (ASILOMAR), 2012, pp. 688–692.

[87] T.A. Thomas, H.C. Nguyen, G.R. MacCartney, T.S. Rappaport, 3D mmWave channel model proposal, in: 2014 IEEE 80th Vehicular Technology Conference (VTC2014-Fall), 2014, pp. 1–6.

[88] Q. Fan, B. Sun, Y. Sun, Y. Wu, X. Zhuang, Data fusion for indoor mobile robot positioning based on tightly coupled INS/UWB, The Journal of Navigation 70 (5) (2017) 1079–1097.

[89] V. Magnago, et al., Ranging-free UHF-RFID Robot Positioning through Phase Measurements of Passive Tags, IEEE Transactions on Instrumentation and Measurement (2019).

[90] W. Hess, D. Kohler, H. Rapp, D. Andor, Real-time loop closure in 2D LIDAR SLAM, in: 2016 IEEE International Conference on Robotics and Automation (ICRA), 2016, pp. 1271–1278.

[91] J. Han, D. Kim, M. Lee, M. Sunwoo, Enhanced road boundary and obstacle detection using a downward-looking LIDAR sensor, IEEE Transactions on Vehicular Technology 61 (3) (2012) 971–985.

[92] G. Jiang, L. Yin, S. Jin, C. Tian, X. Ma, Y. Ou, A simultaneous localization and mapping (SLAM) framework for 2.5 D map building based on low-cost LiDAR and vision fusion, Applied Sciences 9 (10) (2019) 2105.

[93] H.M. Wallach, Topic modeling: beyond bag-of-words, in: Proceedings of the 23rd international conference on Machine learning, 2006, pp. 977–984.

[94] L.-F. Shi, Y. Wang, G.-X. Liu, S. Chen, Y.-L. Zhao, Y.-F. Shi, A fusion algorithm of indoor positioning based on PDR and RSS fingerprint, IEEE Sensors Journal 18 (23) (2018) 9691–9698.

[95] L.-F. Shi, Y. Wang, G.-X. Liu, S. Chen, Y.-L. Zhao, Y.-F. Shi, A fusion algorithm of indoor positioning based on PDR and RSS fingerprint, IEEE Sensors Journal 18 (23) (2018) 9691–9698.

[96] L. Wu, S.C. Hoi, N. Yu, Semantics-preserving bag-of-words models and applications, IEEE Transactions on Image Processing 1869 (7) (2010) 1908–1920.

[97] A. Konstantinidis, G. Chatzimilioudis, D. Zeinalipour-Yazti, P. Mpeis, N. Pelekis, Y. Theodoridis, Privacy-preserving indoor localization on smartphones, IEEE Transactions on Knowledge and Data Engineering 27 (11) (2015) 3042–3055.

[98] S.G. Yoo, J.C. Polo, Enhancement to the Privacy-Aware Authentication for Wi-Fi Based Indoor Positioning Systems, in: International Conference on Applied Cryptography and Network Security, 2019, pp. 143–155.

[99] S. Holcer, J. Torres-Sospedra, M. Gould, I. Remolar, Privacy in indoor positioning systems: a systematic review, in: 2020 international conference on localization and GNSS (ICL-GNSS), 2020, pp. 1–6.

[100] J.W. Kim, D.-H. Kim, B. Jang, Application of local differential privacy to collection of indoor positioning data, Ieee Access 6 (2018) 4276–4286.

[101] H. Q. D. Tran et al., "Employing Extended Kalman Filter with Indoor Positioning System for Robot Localization Application," 2019.

[102] S. Bolognani, L. Tubiana, M. Zigliotto, Extended Kalman filter tuning in sensorless PMSM drives, IEEE Transactions on Industry Applications 39 (6) (2003) 1741–1747.

[103] A. Barrau, S. Bonnabel, The invariant extended Kalman filter as a stable observer, IEEE Transactions on Automatic Control 62 (4) (2016) 1797–1812.

[104] D. Feng, C. Wang, C. He, Y. Zhuang, X.-G. Xia, Kalman-filter-based integration of IMU and UWB for high-accuracy indoor positioning and navigation, IEEE Internet of Things Journal 7 (4) (2020) 3133–3146.

[105] F.S. Daníş, A.T. Cemgíl, C. Ersoy, Adaptive Sequential Monte Carlo Filter for Indoor Positioning and Tracking With Bluetooth Low Energy Beacons, IEEE Access 9 (2021) 37022–37038.
[106] H. Kawaji, K. Hatada, T. Yamasaki, K. Aizawa, Image-based indoor positioning system: fast image matching using omnidirectional panoramic images, in: Proceedings of the 1st ACM international workshop on Multimodal pervasive video analysis, 2010, pp. 1–4.
[107] Y. Ke, R. Sukthankar, PCA-SIFT: A more distinctive representation for local image descriptors, in: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004 2, 2004. II–II.
[108] B. Kulis, K. Grauman, Kernelized locality-sensitive hashing for scalable image search, in: 2009 IEEE 12th international conference on computer vision, 2009, pp. 2130–2137.
[109] Y. Wu, X. Liu, W. Guan, B. Chen, X. Chen, C. Xie, High-speed 3D indoor localization system based on visible light communication using differential evolution algorithm, Optics Communications 424 (2018) 177–189.
[110] H.-H. Liu, C. Liu, Implementation of Wi-Fi signal sampling on an android smartphone for indoor positioning systems, Sensors 18 (1) (2018) 3.
[111] C. Laoudias, D. Zeinalipour-Yazti, C.G. Panayiotou, Crowdsourced indoor localization for diverse devices through radiomap fusion, in: International Conference on Indoor Positioning and Indoor Navigation, 2013, pp. 1–7.
[112] I. Ashraf, S. Hur, Y. Park, Indoor positioning on disparate commercial smartphones using Wi-Fi access points coverage area, Sensors 19 (19) (2019) 4351.
[113] K.F. Davies, I.G. Jones, J.L. Shapiro, A Bayesian approach to dealing with device heterogeneity in an indoor positioning system, in: 2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2018, pp. 1–8.
[114] S. Scardapane, D. Comminiello, M. Scarpiniti, A. Uncini, Online sequential extreme learning machine with kernels, IEEE transactions on neural networks and learning systems 26 (9) (2014) 2214–2220.
[115] H. Zou, X. Lu, H. Jiang, L. Xie, A fast and precise indoor localization algorithm based on an online sequential extreme learning machine, Sensors 15 (1) (2015) 1804–1824.
[116] Y. Chen, T. Yin, H. Babri, A stochastic backpropagation algorithm for training neural networks, in: Proceedings of ICICS, 1997 International Conference on Information, Communications and Signal Processing. Theme: Trends in Information Systems Engineering and Wireless Multimedia Communications (Cat 2, 1997, pp. 703–707. .,.
[117] G.-B. Huang, P. Saratchandran, N. Sundararajan, A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation, IEEE transactions on neural networks 16 (1) (2005) 57–67.
[118] H.K. Yu, S.H. Oh, J.G. Kim, AI based location tracking in WiFi indoor positioning application, in: 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), 2020, pp. 199–202.
[119] E. Shahid, Q. Arain, S. Kumari, I. Farah, Images based indoor positioning using AI and crowdsourcing, in: Proceedings of the 2019 8th International Conference on Educational and Information Technology, 2019, pp. 97–101.
[120] A. Honnef, E. Sawall, M. Mohamed, A.A.S. AlQahtani, T. Alshayeb, Zero-effort indoor continuous social distancing monitoring system, in: 2021 IEEE World AI IoT Congress (AIIoT), 2021, pp. 0482–0485.
[121] M.-H. Tsai, J.-N. Luo, M.-H. Yang, N.-W. Lo, Location Tracking and Forensic Analysis of Criminal Suspects' Footprints, in: 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT), 2019, pp. 210–214.
[122] G. Lee, H. Kim, A hybrid marker-based indoor positioning system for pedestrian tracking in subway stations, Applied Sciences 10 (21) (2020) 7421.
[123] A. Handojo, R. Lim, T. Octavia, J.K. Anggita, Museum interactive information broadcasting using indoor positioning system and Bluetooth low energy: A pilot project on Trowulan museum Indonesia, in: 2018 3rd Technology Innovation Management and Engineering Science International Conference (TIMES-iCON), 2018, pp. 1–5.
[124] M. Rico-González, A. Los Arcos, F.M. Clemente, D. Rojas-Valverde, J. Pino-Ortega, Accuracy and reliability of local positioning systems for measuring sport movement patterns in stadium-scale: A systematic review, Applied Sciences 10 (17) (2020) 5994.
[125] Z. Zhao, Z. Lou, R. Wang, Q. Li, X. Xu, I-WKNN: Fast-speed and high-accuracy WIFI positioning for intelligent sports stadiums, Computers & Electrical Engineering 98 (2022), 107619.
[126] V. Renaudin, et al., Evaluating indoor positioning systems in a shopping mall: The lessons learned from the IPIN 2018 competition, IEEE Access 7 (2019) 148594–148628.
[127] Y. Kamiya, Y. Gu, S. Kamijo, Indoor positioning in large shopping mall with context based map matching, in: 2019 IEEE International Conference on Consumer Electronics *(ICCE)*, 2019, pp. 1–6.