A Hybrid Indoor Localization Framework in an IoT Ecosystem

Marc Junior Pierre Nkengue, Ivan Madjarov, Jean Luc, Rabah Iguernaissi

A Hybrid Indoor Localization Framework in an IoT Ecosystem

Marc Junior Pierre Nkengue, Ivan Madjarov, Jean Luc Damoiseaux, and Rabah Iguernaissi

Abstract The Global Position System (GPS) does not work in the indoor environment because of the satellite signal attenuation. To overcome this lack, we propose a Hybrid Indoor Positioning and Navigation System (HIPNS), based on Li-Fi (Light-Fidelity) localization and optical camera positioning analyses deployed in an indoor environment. The localization approach is based on the fuse of two positioning strategies where the camera-based part is responsible for localizing individuals and recovering their trajectories in zones with low coverage of Li-Fi LEDs. A third-party element is planned to operate in the event of loss of contact. So, the step detection technique and heading estimation are applied in a smartphone-based indoor localization context between two referenced points. The main contribution of this paper focuses on the use of techniques, algorithms, and methods from different spheres of application that generate heterogeneous data. We apply a data integration approach based on REST Web service architecture to allow localization operations in this hybrid indoor positioning system (HIPS). In this work-in-progress paper, we also present a state-of-the-art survey of techniques and algorithms for indoor positioning with the help of smartphones, as well as the main concepts and challenges related to this emergent area.

Keywords Indoor navigation · Li-Fi-based localization · Scene analysis · Smartphone-based positioning · IoT ecosystem

M. J. P. Nkengue · I. Madjarov (✉) · J. L. Damoiseaux · R. Iguernaissi
Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France
e-mail: ivan.madjarov@lis-lab.fr

M. J. P. Nkengue
e-mail: marc-junior-pierre.nkengue@etu.univ-amu.fr

J. L. Damoiseaux
e-mail: jean-luc.damoiseaux@lis-lab.fr

R. Iguernaissi
e-mail: rabah.iguernaissi@lis-lab.fr

1 Introduction

Devices providing sensing, actuation, control, and monitoring (positioning) activities are defined in [1] as the Internet of Things (IoT) ecosystem. The Indoor Positioning Systems (IPS) has been developed using a wide variety of technologies and sensors, or even combining several of them in hybrid systems. Our work is part of this approach as our indoor guidance system combines low-cost technologies that are simple to implement and operate: Li-Fi lamps and video cameras. Besides, we have chosen to process the positioning data from these sensors via a Web service platform, thus ensuring dynamic contact with the user and considering guidance constraints in real-time. Among all indoor positioning technologies, we will focus on those most often used with a mobile phone, namely Wi-Fi, low-energy Bluetooth (BLE), and inertial sensors. We will also present solutions based on the use of light and computer vision.

After a reminder of the possible technologies and the existing hybrid systems, we will then detail the architecture of our guidance system, and the tests carried out, to finally conclude with the follow-up envisaged to our work.

2 Related Work

As multiple published surveys attest [2–6], a wide variety of IPS have been proposed, for performances that are not always satisfactory in dynamic environments and often require costly investments for a significant improvement of the latter. Usually, in IPS, the position of the object or person is estimated using either the measurement of its angle of arrival (AOA), time of arrival (TOA), the difference between arrival times (TDOA), or received signal strength (RSS) [2, 4–6]. If several measurements of the same type are used to determine the position more precisely, the term lateration and angulation is used [4]. The measurement-based systems are complex to implement and expensive in terms of material.

A WLAN is a high-speed wireless network that uses high-frequency radio waves to connect and communicate between nodes and devices within the coverage area. To correctly perform indoor geolocation from a WLAN, it is necessary to densify the network infrastructure to counteract the effect of environmental and human disturbances [4, 5], and also to be able to combine several position measurements or used propagation model within the same algorithm [4, 5].

Very similar to Wi-Fi, the Bluetooth has recently seen a resurgence of interest with the development of Bluetooth Low Energy (BLE) [3, 4]. The low cost of BLE equipment and its long energy autonomy is often cited advantages as they make it easier than Wi-Fi to obtain better radio coverage also necessary for good performance [2, 4]. For geolocation systems, based on WLAN or BLE, many studies propose to improve their performance either by mapping beforehand (fingerprinting) the environment in which the object or person evolves [3–5] or by combining these technologies [2, 4, 5].

The use of the smartphone's sensors (i.e., accelerometer, gyroscope, etc.) is also a research topic tested in the context of IPS [2–4]. Most of the time, they estimate walking parameters (number of steps, length of steps, direction) or determine the nature of the movement. The performances obtained were not convincing, notably because of the difficulty of taking into account the relative position of the smartphone in motion or of integrating physiological parameters (weight, age, etc.) of the person and the nature of the surface of the movement. The current trend is, therefore, to integrate these sensors into WLAN/BLE geolocation systems [2, 3].

Other systems use LED-light for geolocation purposes [2, 4]. Because LEDs are capable of flashing very quickly without impairing human vision, they can substitute for conventional lighting while transmitting information to a smartphone. All positioning algorithms (RSS, TDOA, lateration, angulation, fingerprinting, etc.) can then be used. However, to overcome certain inherent defects of light, it's short-range or it's possible obscuring, couplings with other technologies have already been proposed (e.g., Li-Fi & Wi-Fi) [6].

Finally, there are IPS based on computer vision [2, 4]. In the simplest cases, the phone to determine its position identifies with its camera markers type QR-Codes. But there are also more complex solutions where the mobile device uses video scene analysis to estimate its location by comparing a snapshot of a scene generated by itself with several pre-observed simplified images of the scene taken from different positions and perspectives.

3 A Hybrid System Model for IPS

The localization methods in an IPS are classified into two groups as noted in [7]: (1) based on distance estimation; and (2) mapping-based localization. In the first group, the distance estimation process employs techniques based on the signal strength and/or the elapsed time between two signals. In our work, we opt for the second group where the mapping-based localization works with pre-stored signals (tags) values in a database.

We apply the mapping localization approach in a Li-Fi based positioning system that uses a signal emitted from a LED (light source) to determine the position of the user's device (receiving device). The user's device, which is equipped with a receptacle (e.g., photodiode-dongle), receives the signal from the LED i.e., its identifier. So, we use the ID as a positioning tag associated with a LED lamp installed in a known location, both data prior stored in a database.

We also use a vision-based positioning system to estimate the position and the orientation of a person indoor by identifying an image that is within a view. In [8] authors note that the commonly used methods for image-based indoor positioning are focused on calculating the Euclidean distance between the feature points of an image.

For smartphone-based indoor localization as a compliment, we opt for a Pedestrian Dead Reckoning (PDR) technique to give the position of a mobile user relative to a reference, as presented in [9]. PDR approach relies on IMU (Inertial Measurements Unit) based techniques, which typically comprise an accelerometer, gyroscope, and compass. We use the step detection technique (accelerometer) and heading estimation (gyroscope) to reassure the guided person between two identified positions in case of contact losses from other technics.

In this research and development project, we opt for a hybrid IPS system based on Li-Fi technology with path positioning from optical cameras placed in shadow zones to compensate for each other's shortcomings and take advantage of each other's strengths.

3.1 Positions Data from Camera

The camera-based positioning strategies is responsible for localizing individuals and recovering their trajectories in zones with low coverage with Li-Fi LEDs. Thus, we proposed a mono-camera tracking system that is designed in three main phases. The first ones consist of the detection of individuals and the initialization of trackers which is done in two parts the motion detection and motion segmentation. Then, the second phase consists of the tracking of detected individuals from the first phase to recover their trajectories within the camera's field of view. The last part of our strategy consists of the association of image positioning of individuals with their ground plane positioning. The system design is illustrated on Fig. 1.

The first part of our positioning system is the detection of individuals within the camera's field of view. This is done in two main parts, which are motion detection and motion segmentation. We started by using a background subtraction algorithm, which is based on the use of the Gaussian mixture model as proposed in [10], to detect the foreground of the studied scene. This model is applied to all pixels gives a binary image representing the moving objects within the current frame of the video (Fig. 2).

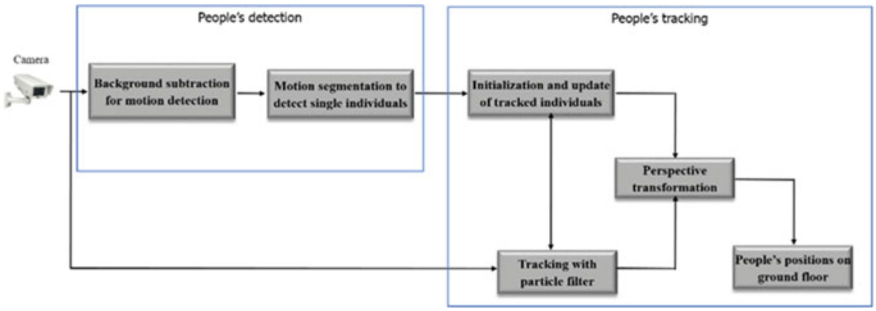

Fig. 1 Ground floor positions from a camera

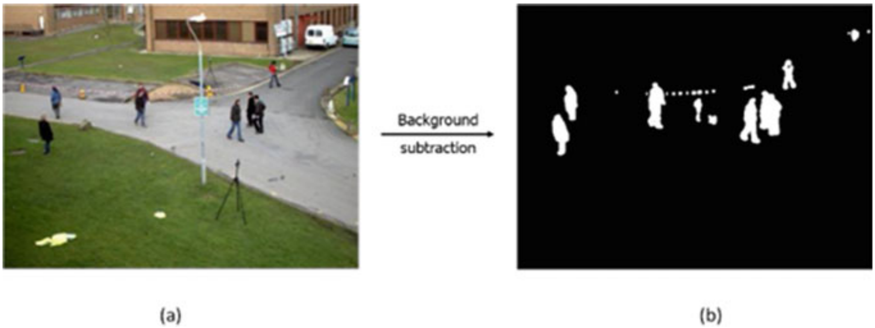

Fig. 2 Motion detection: **a** original image and **b** moving parts

The used strategy for motion detection enables the detection of blobs representing the moving objects within the studied scene at a given time *t*. The detected blobs may represent either a single individual or a group of individuals. Thus, we used a method based on connected components analysis, which is associated with some restrictions on the width and height of blobs, to separate the detected blobs into blobs each representing a single individual. We represented each blob with a rectangle of width w and height h. The properties of this rectangle are estimated based on Eq. (1).

$$\begin{cases} (x_0, y_0) = \left(\frac{x_{min}+x_{max}}{2}, \frac{y_{min}+y_{max}}{2}\right) \\ w = x_{max} - x_{min} \\ h = y_{max} - y_{min} \end{cases} \tag{1}$$

Then, we used a restriction on the ratio between the width and the height of each blob to estimate the number of individuals within the blob. This is done by the assumption of Eq. (2).

$$N_{ind} = \begin{cases} 1 & if\ Th_{min} < \frac{w}{h} < Th_{max} \\ round\left(\frac{w}{h*\left(\frac{Th_{min}+Th_{max}}{2}\right)}\right) & if\ \frac{w}{h} > Th_{max} \\ round\left(\frac{h*\left(\frac{Th_{min}+Th_{max}}{2}\right)}{w}\right) & if\ \frac{w}{h} < Th_{min} \end{cases} \tag{2}$$

The estimated number of individuals is used to perform new segmentation of blobs based on Eq. (3) for an example of a blob with a ratio $\frac{w}{h} > Th_{max}$ and an estimated number of individuals $N_{ind} = 2$ (illustration of results is shown in Fig. 3).

$$\begin{cases} (x_{01}, y_{01}) = \left(\frac{x_{min}+x_{max}}{4}, \frac{y_{min}+y_{max}}{2}\right) \\ (x_{01}, y_{01}) = \left(\frac{3*(x_{min}+x_{\max})}{4}, \frac{y_{min}+y_{max}}{2}\right) \\ w = \frac{x_{max}-x_{min}}{2} \\ h = y_{max} - y_{min} \end{cases} \tag{3}$$

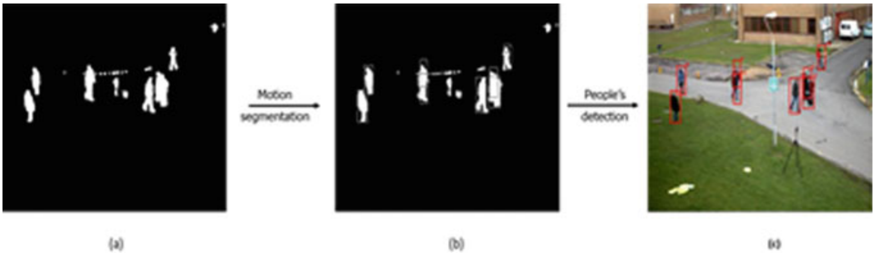

Fig. 3 Motion segmentation and people's detection: **a** moving parts, **b** segmented blobs and **c** detected individuals

The previous steps end up with the list of detected individuals at a given instant t_0. This list is used to initialize the list of tracked individuals, which are then tracked, and their trajectories recovered. For this, we used a strategy based on the use of a particle filter similar to the one proposed in [9] to estimate the position of the tracked individual at instant t based on his position at the instant $t - 1$ (Eq. 4).

$$\begin{cases} (x, y)_t = (x, y)_{t-1} + (u, v)_{t-1} * \Delta t \\ (u, v)_t = (u, v)_{t-1} \end{cases} \tag{4}$$

With $(x, y)_t$ and $(u, v)_t$ the position and velocity of the individual at instant t.

Then a set of N particles are propagated around this position and weighted based on the difference between their color histograms and the color histogram of the individual in the HSV color space.

The positions of these weighted particles are then used to refine the position of the tracked individual at the instant t. The new position of the individual within the current frame is estimated by Eq. (5).

$$\begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix} = \sum_{n=0}^{N} \boldsymbol{w}_t^{(n)} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{bmatrix}^{(n)} \tag{5}$$

The last step of our localization algorithm consists of the association of the image positioning of individuals with their ground plane positioning. In fact, the previous steps are used to recover the trajectories of the individuals on the video. These trajectories are represented by a set of detections representing the individual while moving on the camera's field of view. These detections are then used, first, to localize the individual within the image and, second, to localize the individual on the ground plane. The first part consists of the association with the bounding box of a tracked person with a single point representing his position on the ground plane on the image (u_0, v_0). This is done by considering the point of intersection between the central vertical axes of the detection with the bottom limit of the bounding box.

Then to get the positions of individuals on the ground plane, we use a perspective transformation, similar to the one used in [11], which maps the locations of individuals in the image with their corresponding positions in a plane representing the ground floor of the studied scene. This method is based on the use of four initials points, located by the user in both the image and the plane, that are used to calculate the transformation that will be used later on to map the points from the image to their corresponding positions on the plane. This perspective matrix is estimated using the Eq. (6) and the selected points. Then this perspective matrix is used to get the correspondence between any point on the image and its position on the viewing plane.

$$\begin{bmatrix} x' & y' & z' \end{bmatrix} = \begin{bmatrix} u & v & w \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{6}$$

With: x, y the coordinates of pixels on the viewing plane (ground floor), u, v the coordinates of pixels on the image. At the end of this part, we map the trajectories obtained previously to the estimated trajectories on the ground floor of the studied scene. These trajectories on the ground floor are then sent to the server as camera data that will be combined with the Li-Fi data to localize individuals.

3.2 Data from Li-Fi Lamp

The Li-Fi indoor data model is part of infrastructure-based positioning, non-GPS technologies, where fixed beacon nodes are used for location estimates. The positioning algorithms are associated with Proximity Based Localization (PBL) as classified in [7]. Proximity sensing techniques are used to determine when a user is near a known location. The provided location is the area in which the user is detected. In our case of using, a Li-Fi lamp emits tag to be detected by a mobile target when passes within the covered area. The most common manufacturers' technical parameters for a Li-Fi LED mounted in a standard ceiling height indicate a luminous flux dispersion in a range of 30°–40°. So, to calculate the detector's area, a simple cone-diameter equation can be used, as presented in (7).

$$D = 2 \times h \times tg(\alpha), S = \pi \times \left(\frac{D}{2}\right)^2 \tag{7}$$

where: D expresses the LED covered area, h indicates the ceiling height, α the angle of light dispersion, and S expressed the surface covered by the detector. In a general case, we can count on a detector area with 3 m of diameter i.e., approximatively 7 m^2. This is quite reassuring for the installation of Li-Fi lamps on points of interest in a building. So, the detection infrastructure can be developed as a mesh of Li-Fi lamps, which can be presented as nodes in a graph-path.

3.3 Data Integration and Graph Path

Hybrid System Model. Li-Fi lamps and optical cameras (OC) are two promising IPS technologies that can be implemented in all kinds of indoor environments using existing infrastructures. However, both are subjected to data heterogeneity. In this paper, we propose a hybrid IPS that integrates data from Li-Fi lamps and OC in a RESTful architecture to improve the quality-of-service (QoS) of the user's positioning and navigation in order to provide better performance in terms of accuracy, power consumption, and reduced costs of installation.

In the proposed system model, the source of data dissemination is a Li-Fi lamp and a processed image from an OC, whereas the source data collection is a user device with a photoreceptor. The collected data are analyzed and processed, and the localization is performed via a Web service. In Fig. 4, a four layers system architecture is presented: (1) data generation and image collection, (2) communication technology, (3) data management and processing, and (4) application for data interpretation.

When the user passes under a LED his smartphone can receive the tag associated with this LED lamp. Its path is followed by an optical camera to confirm the user's position. An alert message will be sent in case of remoteness from the prescribed

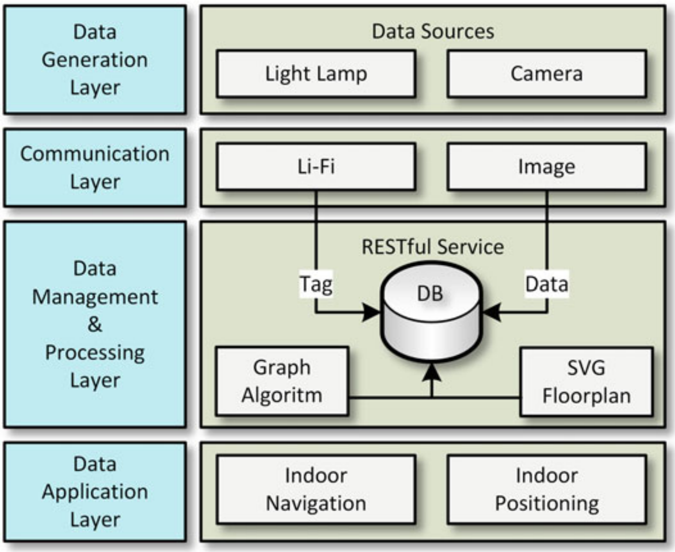

Fig. 4 System model for hybrid IPS

path or in case of unexpected barriers. A reconfigured path will be then sent to the user.

The graph-path algorithm. The BFS-based graph-path algorithm resides on the RESTful Web service side. This algorithm allows us to obtain the path to follow when a destination is defined at the beginning (e.g., the entry point of a building). With knowledge of the starting point and the endpoint, the algorithm determines all intermediary points to be followed to guide the user to the destination. These intermediary points represent the graph-vertices where the Li-Fi lamps are positioned.

The Graph algorithm is developed as a class with two methods. The first one determines the vertices in the graph corresponding to the building's plan stored in numerical format. Once the set of vertices is retrieved, the method locates for each node the set of vertices that succeeds it in a unidirectional manner (i.e., for each vertex, the edge to follow to the next vertex throughout the part of the suggested path). The second one is essential to allow us to find the available path from the starting point to the defined endpoint. Based on the graph-paths established by the previously described method, this method allows the suggested path to be highlighted on the user's screen.

The vector floorplan. The vector graphics format (SVG) used for the building's plan representation allows us to manipulate the graph directly on the plan by associating it to the user's path. Thus, the highlighting path can be displayed directly on the graph with the points of reference (i.e., graph-vertices).

4 Implementation and Evaluation

Implementation. We have focused on server-side processing as a development approach to reduce the user's client–server interactions. So, in this IoT schema, the dedicated REST service, as shown in Fig. 4, can handle multiple requests at once with correctly achieved data integration from heterogeneous sources like Li-Fi lamp, optical camera, an accelerometer as shown in Fig. 5. On the other hand, this

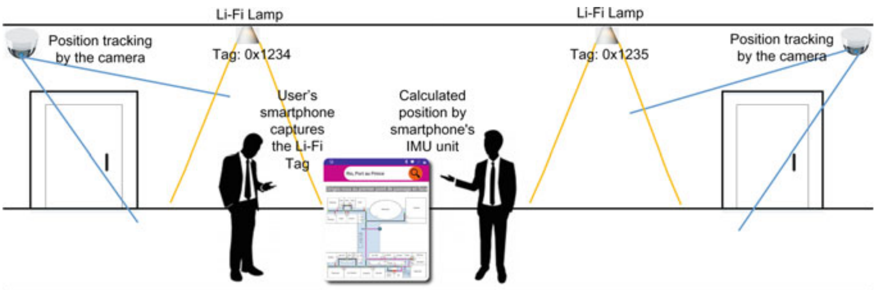

Fig. 5 Use case scenario for hybrid Li-Fi-camera-accelerometer IPS

centralization approach for indoor navigation process management allows server-side service to track simultaneously different requested paths without interferences between users.

A location-aware Android-based application for indoor navigation tracking is developed. So, when a smartphone with a light sensor, is within the range of a Li-Fi lamplight, it will compare the emitted from the lamp tag with the value recommended in the building's path-list. The graph-path is highlighted on the building's plan, already displayed on the smartphone's screen, with the highlighted intermediary point of the detected position as shown in Fig. 5. The developed Android activity is based on the Oledcomm GEOLiFi Kit [12], with GEOLiFi LED lamp, GEOLiFi Dongle to be used with a smartphone, and GEOLiFi SDK Library for Android application development.

The data integration of the camera and the Li-Fi lamps is done through the Web service installed on a Node.js server running on a Raspberry pi 4. The reference points identified by the camera for the guided person are stored in the database. When the user passes through a Li-Fi point, the retrieved coordinates are compared with those transmitted by the camera. In case of differences, the coordinates confirmed by the position of the Li-Fi lamp are considered for the user's guidance.

To improve the accuracy of the localization system, we combine different technologies. To increase the quality of the data and to reassure the user in case of failure of the main approach, an accelerometer, a gyroscope, embedded in a smartphone are employed to develop a multi-sensor fusion approach. This results in the Android application that integrates data from the IMU for the user's guidance between two reference points. However, this data is not communicated to the server and its Web service.

Evaluation. For this work-in-progress paper, the performance of each positioning approach is partially analyzed due to objective reasons. Our project started at the end of 2019. The containment imposed by the Covid-19 pandemic prevented us from deploying the entire infrastructure, namely optical cameras, and Li-Fi lamps, on a larger scale. We were planning to deploy four optical cameras and 32 Li-Fi lamps. The pretests were carried out in an enclosed space with a minimum of deployed equipment. The camera-based algorithms for localizing individuals and recovering their trajectories were tested with an extern public database. Moreover, this avoids some inconvenience in terms of image rights. The guidance activity with an accelerometer and gyroscope was tested in extern associated to the main Android application. The graph path algorithm, installed as RESTful service on Raspberry pi 4, was tested on a virtual floorplan with QR Codes in place of the Li-Fi lamps. The developed Android application for user indoor guidance gave satisfaction.

To estimate the accuracy of the IMU unit associated with the user's activity, we proceed by a test to count the number of steps over 10 m and then to compare with real values. It appears that the accuracy of the IMU unit is quite good over the tested distance. The observed error has a rate of up to 23%, which is a tolerable threshold. A real difference begins to be created between the values of the IMU unit and the real values beyond 9 m, so a distance lower than 10 m is recommended between two Li-Fi lamps.

5 Conclusion

In this article, we present a hybrid IPS system based on the integration of data from heterogeneous sources: i.e. Li-Fi tags to determine the positioning of a user on a floorplan; trajectory tracking of the user by optical cameras; step counting by a smartphone application supposed to guide the user between two reference points and in case of loss of cameras tracking due to congestion, smoke or other disruptive events.

Because it does not require any special infrastructure, the proposed solution is easy to implement and low cost, and it would be easy to install it in most indoor environments like hospitals, buildings, campuses, and malls.

References

1. ITU (2015) IoT global standards initiative. https://handle.itu.int/11.1002/1000/11559
2. Brena RF et al (2017) Evolution of indoor positioning technologies: a survey. J Sens 2017, Article ID 26304113
3. Davidson P, Piche R (2017) A survey of selected indoor positioning methods for smartphones. IEEE Commun Surv Tutor 19(2):1347–1370
4. Mendoza-Silva GM, Torres-Sospedra J, Huerta J (2019) A meta-review of indoor positioning systems. Sensor 19:4507. https://doi.org/10.3390/s19204507
5. Zafari F, Gkelias A, Leung KK (2019) A survey of indoor localization systems and technologies. arXiv:1709.01015v3 [cs.NI] 16 jan 2019
6. Wu X et al (2020) Hybrid LIFI and Wifi networks: a survey. arXiv:2001.04840v1
7. Rahman ABMM, Li T, Wang Y (2020) Recent advances in indoor localization via visible lights: a survey. Sensors 20(5):1382
8. Yang S, Ma L, Jia S, Qin D (2020) An improved vision-based indoor positioning method. IEEE Access 8:26941–26949. https://doi.org/10.1109/ACCESS.2020.2968958
9. Nummiaro K, Koller-Meier E, Van Gool L (2002) Object tracking with an adaptive color-based particle filter. Pattern Recogn, 353–360
10. Shimada A et al (2006) Dynamic control of adaptive mixture-of-Gaussians background model. In: Video and signal based surveillance, 2006. IEEE-AVSS'06, 2006, pp 5–5
11. Sun M et al (2019) See-your-room: indoor localization with camera vision. In: Proceedings of the ACM turing celebration conference-China, pp 1–5
12. Oledcomm (2020) GEOLiFi kit. https://www.oledcomm.net/lifimax-discovery-kit/