

REVIEW

Open Access



Indoor positioning and wayfinding systems: a survey

Jayakanth Kunhoth^{*}, AbdelGhani Karkar, Somaya Al-Maadeed and Abdulla Al-Ali

^{*}Correspondence:
jayakanth.k.chandran@gmail.com
Department of Computer
Science and Engineering,
Qatar University, Al Jamiaa
Street, Doha, Qatar

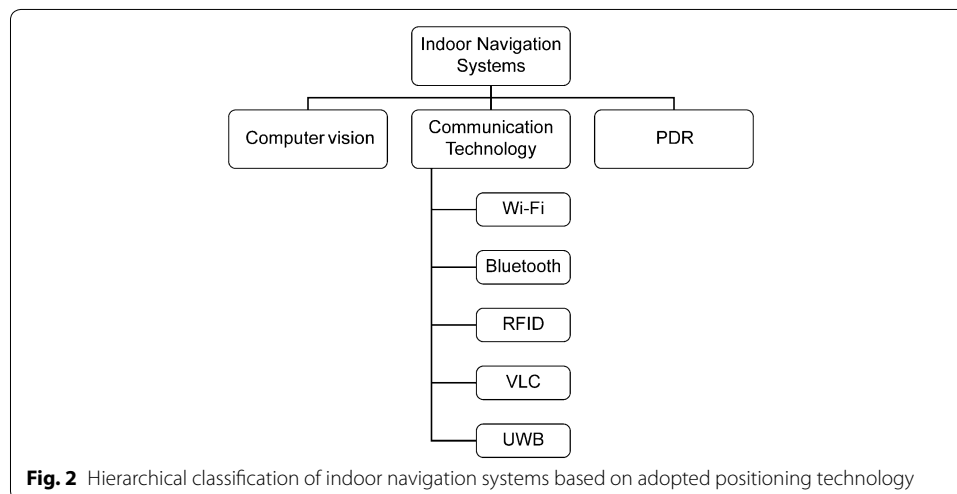
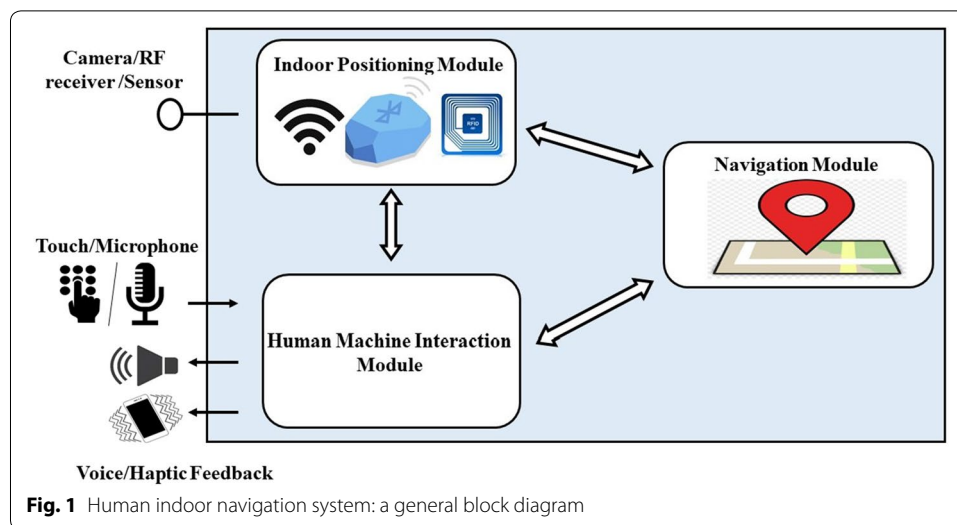
Abstract

Navigation systems help users access unfamiliar environments. Current technological advancements enable users to encapsulate these systems in handheld devices, which effectively increases the popularity of navigation systems and the number of users. In indoor environments, lack of Global Positioning System (GPS) signals and line of sight with orbiting satellites makes navigation more challenging compared to outdoor environments. Radio frequency (RF) signals, computer vision, and sensor-based solutions are more suitable for tracking the users in indoor environments. This article provides a comprehensive summary of evolution in indoor navigation and indoor positioning technologies. In particular, the paper reviews different computer vision-based indoor navigation and positioning systems along with indoor scene recognition methods that can aid the indoor navigation. Navigation and positioning systems that utilize pedestrian dead reckoning (PDR) methods and various communication technologies, such as Wi-Fi, Radio Frequency Identification (RFID) visible light, Bluetooth and ultra-wide band (UWB), are detailed as well. Moreover, this article investigates and contrasts the different navigation systems in each category. Various evaluation criteria for indoor navigation systems are proposed in this work. The article concludes with a brief insight into future directions in indoor positioning and navigation systems.

Keywords: Indoor navigation, Indoor positioning, Computer vision, PDR, RF signals, Visible lights

Introduction

The term ‘navigation’ collectively represent tasks that include tracking the user’s position, planning feasible routes and guiding the user through the routes to reach the desired destination. In the past, considerable number of navigation systems were developed for accessing outdoor and indoor environments. Most of the outdoor navigation systems adopt GPS and Global Navigation Satellite System (GLONASS) to track the user’s position. Important applications of outdoor navigation systems include wayfinding for vehicles, pedestrians, and blind people [1, 2]. In indoor environments, the GPS cannot provide fair accuracy in tracking due to nonlinear of sight issues [3]. This limitation hinders the implementation of GPS in indoor navigation systems, although it can be solved by using “high-sensitivity GPS receivers or GPS pseudolites” [4]. However, the cost of implementation can be a barrier to applying this system in real-world scenarios.



Indoor navigation systems have broad number of applications. The certain applications are wayfinding for humans in railway stations, bus stations, shopping malls, museums, airports, and libraries. Visually impaired people also benefit from indoor navigation systems. Unlike outdoor areas, navigation through indoor areas are more difficult. The indoor areas contains different types of obstacles, which increases the difficulty of implementing navigation systems. General block diagram of a human indoor navigation system is illustrated in Fig. 1.

A human indoor navigation system mainly consists of the following three modules: (1) Indoor positioning system module, (2) Navigation module, and (3) Human-machine interaction (HMI) module. The indoor positioning system estimates the user's position, the navigation module calculates routes to the destination from current location of the user, and the HMI module helps the user to interact with the system and provide instructions to the user. Since GPS-based indoor positioning is not effective, methods based on computer vision, PDR, RF signals are utilized for indoor positioning. Figure 2

illustrates the hierarchical classification of indoor navigation systems according to the positioning technologies adopted by them.

Computer vision-based systems employ omnidirectional cameras, 3D cameras or inbuilt smartphone cameras to extract information about indoor environments. Various image processing algorithms, such as Speeded Up Robust Feature (SURF) [5], Gist features [6], Scale Invariant Feature Transform (SIFT) [7], etc., have been utilized for feature extraction and matching. Along with feature extraction algorithms, clustering and matching algorithms are also adopted in conventional approaches for vision-based positioning and navigation systems. Apart from conventional approaches, computer vision based navigation systems utilized deep learning methodologies in recent years. Deep learning models contains multiple processing layers to study the features of input data without an explicit feature engineering process [8]. Thus, deep learning-based approaches have been distinguished among object detection and classification methods. Egomotion-based position estimation methods are also utilized in computer vision-based navigation systems [9]. Egomotion approach estimates the camera's position with respect to the surrounding environment.

PDR methods estimate the user's position based on past positions by utilizing data from accelerometers, gyroscopes, magnetometers, etc. The user's position is calculated by combining the step length, the number of steps and the heading angle of the user [10, 11]. Since a greater number of position errors occur in dead reckoning approaches due to drift [12], most of latest navigation systems integrate other positioning technologies with PDR or introduced some sensor data fusion methods to reduce the errors.

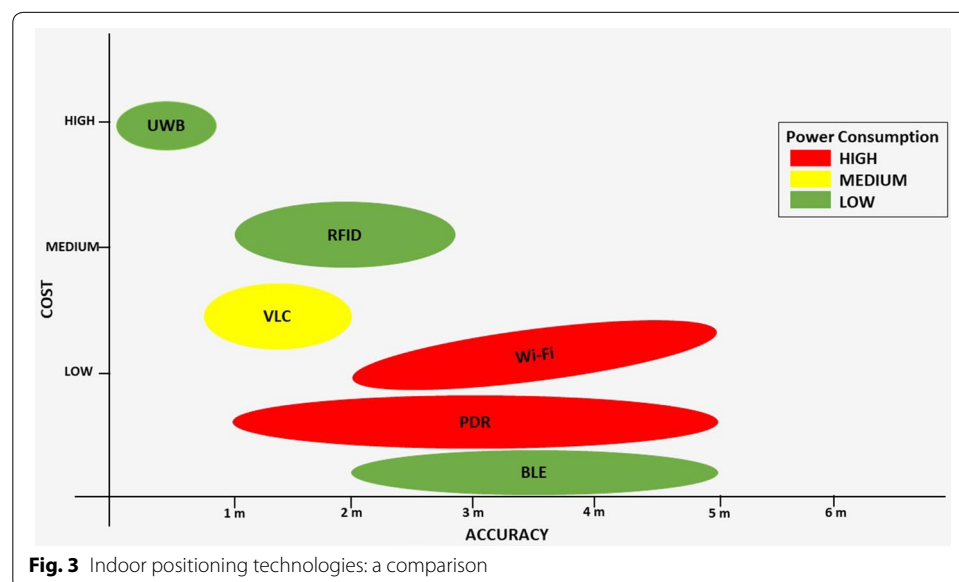
Communication-based technologies for indoor positioning includes RFID, Wi-Fi, visible light communication (VLC), UWB and Bluetooth. RFID systems consist of a RFID reader and RFID tags attached to the objects. There exist two types of RFID tags, namely, active and passive. Most of the recent RFID-based navigation systems have implemented passive tags since an external power source is not required. RFID-based systems utilize Received signal strength (RSS), Angle of arrival (AOA), Time of arrival (TOA) and Time difference of arrival (TDOA) for position estimation [13]. In indoor environments, however, all the methods except RSS may fail to estimate the user's position accurately due to nonline of sight scenarios. The popular RSS-based positioning approaches are trilateration and fingerprinting [14]. RFID technology are widely implemented in navigation systems because of their simplicity, cost efficiency, and long effective ranges. Wi-Fi-based approaches are implemented in indoor environments, where we have sufficient numbers of Wi-Fi access points, and a dedicated infrastructure is not required; instead, these approaches can utilize existing building infrastructure because most current buildings will be equipped with Wi-Fi access points. Wi-Fi-based indoor navigation systems make use of RSS fingerprinting or triangulation or trilateration methods for positioning [15]. Bluetooth-based systems have almost similar accuracy as Wi-Fi-based systems and use Bluetooth low energy (BLE) beacons as source of RF signals to track the positions of users using proximity sensing approaches or RSSI fingerprinting [16]. In recent advances, smartphones are usually used as a receiver for both Bluetooth and Wi-Fi signals. VLC-based systems utilize the existing LED or fluorescent lamps within buildings, which makes VLC-based systems low cost. These LEDs or fluorescent lamps are becoming ubiquitous in

indoor areas. The light emitted by lamps is detected using smartphone cameras or an independent photo detector. TOA, AOA, and TDOA are the most popular measuring methods used in VLC-based positioning systems [17]. UWB-based positioning systems can provide centimeter-level accuracy, which is far better than Wi-Fi-based or Bluetooth-based methods. UWB uses TOA, AOA, TDOA, and RSS-based methods for position estimation [18]. Comparison of various indoor positioning technologies in terms of accuracy, cost of implementation and power consumption are shown in Fig. 3.

The navigation module will determine the route of the user in the constructed indoor map with respect to user's current position. The navigation module mainly consists of a map which represent the areas of indoor environment and a method to plan the navigation routes. The most commonly used methods for route planning are A* algorithm [19], Dijkstra's algorithm [20], D* algorithm [21] and Floyd's algorithm [22]. In addition, there exist some systems that provide mapless navigation. All these systems are discussed in the upcoming sections.

Human-machine interaction module allows the users to communicate with the navigation system such as to set up the destination as well as change the destination. The HMI module gives proper information and guidance to the users regarding route and location by means of acoustic feedback [23] or haptic feedback [24]. In the case of visually impaired ones, audio or vibration feedback is widely implemented in the HMI module.

In the past, significant number of attractive surveys about various indoor positioning technologies and indoor navigation systems were published [16–18, 25–28]. Most of these surveys mainly concentrated on positioning systems rather than the navigation system. In addition, they considered only a single technology such as wireless-based systems or visible light-based system or vision-based systems. In this work, we provide a summary of recent advancements and developments in the field of indoor navigation and positioning systems that utilize different types of approaches, such as



computer vision, sensors, RF signals, and visible lights. The survey primarily deals with human navigation systems, including assisted systems for people with visual impairments (VI). In addition, some robotic navigation systems are also detailed in this paper.

Indoor positioning and wayfinding systems

Computer vision-based navigation and wayfinding systems

One of the main applications of indoor navigation is wayfinding for people with VI. ISANA [29] is a vision-based navigation system for visually impaired individuals in indoor environments. The proposed system prototype contains a Google tango mobile device, a smart cane with keypad and two vibration motors. The Google tango device has a RGB-D camera, a wide-angle camera, and 9 axes inertial measurement unit (IMU). The key contributions of ISANA are: (1) an “indoor map editor” to create semantic map of indoor areas, (2) “obstacle detection and avoidance method” that aids real-time path planning and (3) a Smart Cane called “CCNY Smart cane” that can alleviate issues associated with voice recognition software. The geometric entities in the floors, such as lines, text, polygons, and ellipses, were extracted by the indoor map editor from the input CAD model of the indoor areas. The indoor map editor can recognize the locations of rooms, doors, hallways, spatial and geometrical relationships between room labels, and global 2-dimensional traversal grid map layers. Prim’s minimum spanning tree algorithm is employed to draw out the above-mentioned semantic information. A novel map alignment algorithm to localize the users in the semantic map is proposed in ISANA. The proposed map alignment algorithm utilizes the 6-DOF pose estimation and area descriptive file provided by Google Tango VPS. The navigation module utilizes the global navigation graph constructed from the 2-dimensional grid map layer along with the A* algorithm for path planning. The safety of visually impaired individuals is guaranteed via obstacle detection, motion estimation and avoidance methods introduced in ISANA. To detect the obstacles, the ISANA will make use of the RGB-D camera to acquire depth data. The 3-dimensional point cloud or the depth data are rasterized and subjected to a denoising filter to remove the outliers. Three-dimensional points are aligned with the horizontal plane by utilizing the deskewing process. A 2-dimensional projection-based approach is introduced to avoid the obstacle, and it produces a time stamped horizontal map for path planning and time stamped vertical map for obstacle alerts. A connected component labeling approach-based algorithm [30] is adapted to detect the object to create horizontal and vertical maps. The Kalman filter is employed to reckon the motion of obstacles based on time-stamped maps. ISANA uses an Android text to speech library to speak out the instructions and feedback to the users and a speech to text module [31] to recognize user’s voice inputs. The CCNY smart cane provides haptic feedback to the user in noisy environments, and it also has a keypad to set the destinations and IMU to track the orientation of user and cane.

Tian et al. [32] developed a system for helping blind persons navigate indoor environments. The proposed system consist of door detector module and text recognition module. The separate module for the door detector consisted of the Canny edge detector and curvature property-based corner detector. The relative positions were detected by measuring the angle between the top edge of the door and the horizontal axis. A

mean shift-based clustering algorithm was adopted for enhancing the text extraction by grouping similar pixels. A text localization model was designed by considering that texts have shapes with closed boundaries and a maximum of two holes. Text recognition was achieved by using Tesseract and Omni Page optical character recognition (OCR) software. The demonstrated results show that the false positive rate increased for the images acquired under challenging conditions, such as low light, partial occlusion, etc.

A wearable navigation system for people with VI by utilizing a RGB-D camera was proposed in [33], and it used sparse feature and dense point clouds for estimating camera motions. The position and orientation of the objects in the indoor environment were identified using a corner-based real-time motion estimator algorithm [34] and along with that, an iterative closest point algorithm was included to prevent drift and errors in pose estimation. A simultaneous localization and mapping (SLAM) algorithm provided the mapping [35, 36]. The modified D* lite algorithm helped the user route through the shortest path. Although the D* algorithm can handle dynamic changes in the surroundings, narrow changes in the map can cause the change in the produced walking path. This issue will make the navigation of people with VI more complicate. Normal D* algorithm generates the shortest path as a set of cells on the grid map. And it connects the current location and final destination by excluding untraversable cells. In this work, instead of directly following the generated set of cells, a valid waypoint point is generated in such a way that waypoint should be traversable as well as it should be located near to the obstacles at some distance. In the valid waypoint generation method, a point that is most far, visible as well as traversable from the current location is selected from the set of cells generated by the D* algorithm. Also, another point is selected, which is far, visible as well as traversable from the first selected point. Finally, a cell near to the first selected point and with less cost function is computed. However, some of the maps were inconsistent because the map merging technique was unable to correct for deformations in the merged maps.

The indoor wayfinding system for people with VI in [37] utilized Google Glass and Android phone. The proposed object detection method used the Canny edge detector and Hough line transform. Since walls may be one of the main obstacles in indoor environments, the floor detection algorithm identified the presence of walls by finding the stature of the floor region. However, the proposed object detection method failed for bulletin boards as well as indoor low contrast wall pixels.

In [38], the Continuous adaptive mean (CAM) shift algorithm was implemented with the D* algorithm for helping blind people navigate in indoor areas. The proposed method used image subtraction for object detection and histogram backpropagation for creating a color histogram of detected objects. The CAM shift algorithm provided tracking and localization of the users, and the D* algorithm helped the user calculate the shortest route between the source and destination.

Bai et al. [39] developed a vision-based navigation system for people with VI by utilizing a cloud-computing platform. The proposed prototype is made up of a stereo camera mounted on a helmet, smartphone, web application, and cloud platform. The helmet also contains a speaker and earphones to facilitate the human-machine interaction. The stereo camera acquires all the information about the surroundings and forwards it to the smartphone using Bluetooth. The smartphone will act as a bridge

between users and cloud platform. All the core activities of the system, such as object or obstacle detection, recognition, speech processing, navigation route planning, is performed at the cloud platform. The cloud platform contains three modules, namely, speech processing, perception, and navigation. The speech processing module implemented a recurrent neural network-based natural language processing algorithm [40, 41] to analyze the user's voice commands. The perception module makes the user aware of his surroundings and aid the blind people to live as a normal person, and it fuses object detection and recognition functions [42], scene parsing functions [43, 44], OCR [45, 46], currency recognition functions [47] and traffic light recognition functions [48] to improve the blind user's awareness about the environment. All the functionalities in the perception module are based on deep learning algorithms. The navigation module implements a vision-based SLAM algorithm to construct the map. The SLAM algorithm will extract the image features of the surrounding environment and recreate the path of the camera's motion. Preassigned sighted people use web application to provide additional support for blinds in complex scenarios.

Athira et al. [49] proposed an indoor navigation system for shopping mall. The proposed vision based system used GIST feature descriptor, and it enhanced the processing of captured images and reduced memory requirements. The main functions of the proposed system are keyframe extraction, topological map creation, localization, and routing. Keyframes are the important frames extracted from walkthrough videos that are used to create a topological map. For each frame, the L2 norm between two descriptors is calculated. If the L2 norm (Euclidean distance) exceeds a specific value, the frame is considered as "keyframe". Consequently, the direction of key frames is detected by analyzing present frames with left and right parts of prior frames individually to create the map. Once the direction of keyframes is detected, 2D points are calculated in the map. For localization, images captured from user's current position is compared with existing keyframes using the L2 norm.

Bookmark [50] is an infrastructure-free indoor navigation system that utilize existing barcodes of books in a library, and it facilitates the navigation of library visitors to any book's position just by scanning the barcode of books in the library. Bookmark was developed as an application that can be used in any phone that has a camera, and it provides a detailed map of the library to the user. The detailed map contains locations of stairs, elevators, doors, exits, obstacles (pillars or interior walls) and each bookcase. Then, the map is converted to scalable vector graphics format, and the locations are represented by different color codes. To map the books in the library with the map, a book database of call numbers (a unique alphanumeric identifier associated with each book) and locations associated with call numbers is created. When a user scans the barcode of a book to know the location, the Bookmark's server-side will collect information about the book from an existing library API. This information will contain details of the book, including the call number. The system will look up the call number inside the book database to retrieve the location for the user. Bookmark implements the A* algorithm to plan the route between two points of interest. Since Bookmark does not use a positioning technique, the system will be unaware about the current position of the user until he/she completes the navigation or until

the next barcode is scanned. The major limitations of the system include the absence of barcodes on books and the misplacement of books on the wrong shelves.

Li et al. [51] proposed a wearable virtual usher to aid users in finding routes in indoor environments, and it consists of a wearable camera that captures pictures for frontal scenes, headphones to listen to verbal routing instructions to reach a specific destination, and a personal computer. The aim of the system is to aid users in wayfinding in an indoor environment using egocentric visual perception. A hierarchical contextual structure composed of interconnected nodes uses cognitive knowledge to estimate the route. The hierarchical structure can be presented in the following three levels: (1) top level, where the root node represents the building itself; (2) zones and areas inside the building; and (3) bottom level, which corresponds to the location inside each area. Generally, the structure illustrates the human mental model and the understanding of an indoor environment. SIFT has been used for scene recognition. A “self-adaptive dynamic-Bayesian network” is developed to find the best navigation route, and it is self-adaptive and can modify its parameters according to the current visual frame. Moreover, this network can address uncertainties in perception and is able to predict relevant routes. The obtained results demonstrated that the developed system is capable of assisting users to reach their destination without requiring concentration and a complex understanding about the map.

ViNav [52] is a vision based indoor navigation system developed for smartphones. The proposed system provides indoor mapping, localization and navigation solutions by utilizing the visual data as well as data from smartphone’s inbuilt IMUs. ViNav system is designed as a client–server architecture. The client is responsible for collecting the visual imageries (images and videos) and data from sensors including accelerometer, barometer, gyroscope, etc. The server will receive these data from the client and build 3-dimensional models from that. The server comprises of two modules. The first module is responsible for building 3-dimensional models of the indoor environment. Structure from motion technique is used to build 3-dimensional models from crowdsourced imageries captured by the client. The data from the accelerometer as well as gyroscope are utilized to detect trajectories of the user. Moreover, Wi-Fi fingerprints collected from the path traveled by the user are combined with the 3d model for localizing the user’s position in the indoor area. The second module facilitates the navigation of the user by calculating the navigation routes by using pathfinding algorithms. The data about the obstacles in the path are retrieved from constructed 3D models and navigation meshes are computed by adding pedestrian’s traveling path retrieved from crowdsourced user’s paths with obstacle’s data. Barometer readings are utilized to detect the stairs, elevators, and change of floors. The performance evaluation experiments demonstrate that ViNav can locate users within 2 s with an error, not more than 1 m.

Rahman et al. [53] proposed a vision-based navigation system using the smartphone. The proposed system is designed in a manner where the smartphone camera is enabled to capture the images in front of the user. The captured images will be compared with pre-stored images to check whether the captured image contains any obstacles. An algorithm is proposed for assisting people with visual impairments. The algorithm performs both obstacle detection as well as pathfinding tasks for the user. Once an image is captured by the smartphone, the obstacle detection technique will initially extract the

region of interest from the image. The extracted region of interest will be compared with images in the database. If an obstacle is detected, the pathfinding technique will suggest an alternate path for the user. It is achieved by checking the right and left of the extracted region of interest. In a test environment proposed system achieved an accuracy of 90%.

Reference [54] examined the performance of three indoor navigation systems that utilize different techniques for guiding people with visual impairments in the indoor environment. The proposed work focused on the development of three navigation systems that utilize image matching, QR code, and BLE beacons respectively for localizing the user and testing of the developed navigation system in the realtime indoor environment. Image matching based indoor navigation system included a novel CNN model that is trained with thousands of images to identify the indoor locations. QR code-based system utilized existing QR code methods such as Zxing and Zbar. BLE beacons based method adopted a commercially available indoor positioning SDK to localize the user in indoor areas. All three navigation systems are implemented in a smartphone for real-time evaluation. Evaluation results show that QR code and image matching based methods outperformed the BLE beacons based navigation system for people with visual impairments in the indoor environment.

Tyukin et al. [55] proposed an indoor navigation system for autonomous mobile robots. The proposed system utilizes an image processing-based approach to navigate the robot in indoor areas. The system consists of “a simple monocular TV camera” and “color beacons”. The color beacons are the passive device that has three areas with different colors. All these colors can be visually identified, and the surface of the beacons are matte and not glowing. The operation of the proposed system can be classified into the three steps: (1) detection of the color beacons; (2) relative map generation, which identifies the location of the detected beacons in the indoor space with respect to the TV camera; and (3) identification of robot coordinates in the absolute map. An algorithm that contains different image processing techniques was introduced for beacon detection. Initially, the image from the TV camera was subjected to noise removal and smoothing of image defects using a Gaussian filter. After preprocessing, the image is converted to HSV. Then, the algorithm will choose each color in order, and a smooth continuous function is applied for the classification of pixels. Color mask images are generated by averaging the grayscale images from each HSV channel. Finally, the algorithm will recognize the pixel with the maximum intensity in the color mask and will fill pixels around it. The algorithm will repeat this step until all colors used in the beacons are processed. Once the center of the colored areas of the beacon is identified, the magnitude and direction of the vectors connecting the center of the colored area are estimated. There will be two vectors, with one connecting the center of the first and second colored area and the other connecting the center of the second and third colored area. The differences between these two vectors are used to identify the beacons. A navigation algorithm is introduced to estimate the coordinates of the beacon’s location and absolute coordinates of the TV camera. The relative coordinates of beacons were estimated using the beacon height and the aperture angle of the lens. The created relative map is an image where the relative positions of beacons and colored beacons are represented as dots. A three-dimensional transformation is applied to the relative coordinates of beacons to create the absolute map. The demonstrated results of the experiment using the

proposed system show that the detection algorithm is able to detect the beacons only if it is within a range of 1.8 m from the TV camera. The average deviation in the calculated absolute coordinates was only 5 mm from the original value.

Bista et al. [56] proposed a vision-based navigation method for robots in indoor environments. The whole navigation process depends on 2-dimensional data of the images instead of 3-dimensional data of images utilized in existing methods. They depicted the indoor area as a collection of reference images that were obtained during the earlier learning stage. The proposed method enables the robot to navigate through the learned route with the help of a 2-dimensional line segment detector. To detect the line segment in the acquired images, a highly accurate and quick line detector called EDLine detectors [57] is employed. The indoor maps were created by utilizing key images and its line segments. During map construction, the first acquired image will be considered as a key image. The line segment of the first key image will be matched with the next image's (second image) line segment to form a set of line segments. For matching line segments, a Line Band descriptor-based matching method was adopted [58]. Matching will be mainly based on Line Band descriptor, followed by the application of geometric constraints and filters to remove false matches. Once the matched set of the key image and the next image is obtained, the method will consider the next image (third image) and perform the same line segment matching between the first acquired key image and the current image. These steps will result in two-matched set of lines. A trifocal tensor is utilized to find the two-view matches between these two sets. The trifocal tensor is a " 3×3 array that contains all the geometric relationships among three views," and it needs three-view correspondence between lines. Two-view matches (matching of the current image with the previous and next key image) utilized for initial localization and three-view correspondence generation. Three-view matches were used for mapping (matching the current image with previous, next and second next key image). The previous and next key images of a currently acquired image will share some line segment, which facilitates robot navigation and motion control. The rotational velocity of the robot is also derived from the three-view matches. The proposed navigation method was evaluated in three different indoor areas. The obtained drift in the navigation path of the robot was only 3 cm and 5 cm for the first two experiments. In the third experiment, a large drift was present in the path of the robot during the circular turn. The inclusion of obstacle avoidance module will be considered in future work to deal with the dynamic objects in the indoor environment.

Table 1 illustrates the comparison of computer vision-based indoor navigation systems.

Computer vision-based positioning and localization systems

The tasks of indoor localization, positioning, scene recognition and detection of specific objects, such as doors, were also considered in the context of indoor navigation since they can be extended for wayfinding in indoor areas.

Tian et al. [59] developed a method to detect doors for assisting people with VI to access unfamiliar indoor areas. The proposed prototype consists of a miniature camera mounted on the head for capturing the image and a computer to provide the speech output following the object detection algorithm. A "generic geometric door model"

Table 1 Computer vision-based navigation and wayfinding systems

References	Beneficiary	Computer-vision solution	Path planning solution	Remarks /findings
Lie et al. [29]	People with VI	Google Tango VPS	Prism MST algorithm, A* algorithm	(+) Haptic feedback system provided safe navigation in noisy environments
Tian et al. [32]	People with VI	Canny edge detector, Tesseract and Omni page OCRs	Not available	(−) Path planning module is absent
Lee and Medioni [33]	People with VI	Corner-based motion estimator algorithm	SLAM and D* Lite algorithm	(−) Inconsistency in constructed maps
Garcia and Nahapetian [37]	People with VI	Canny edge detector and Hough line transform	Not available	(−) Detection failed for bulletin boards as well as low contrast wall pixels
Manlises et al. [38]	People with VI	Image subtraction, Histogram back-propagation	D* algorithm	(−) Low brightness and noise in indoor areas will affect the recognition and feedback systems, respectively
Bai et al. [39]	People with VI	Deep learning-based object recognition, scene parsing, Currency recognition functions	Vision-based slam	(+) Improved location awareness for the users
Athira et al. [49]	Customers of shopping mall	Gist descriptors	Not available	(−) Does not support navigation between floors
Pearson et al. [50]	Visitors of library	Bar code recognition	A* algorithm	(−) Misplaced books and books without barcodes can limit the system functionalities
Li et al. [51]	Normal people	SIFT descriptors	Self-adaptive dynamic-Bayesian network	(+) Scalability

built on stable edge and corner features facilitates door detection. Objects with similar shapes and sizes, such as bookshelves and cabinets, were separated from the door using additional geographic information. The presented results indicate a true positive rate of 91.9%.

The Blavigator project included a computer vision module [60] for assisting blind people in both indoor and outdoor areas. The proposed object collision detection algorithm uses a “2D Ensemble Empirical Mode Decomposition image optimization algorithm” and a “two-layer disparity image segmentation algorithm” to identify adjacent objects. Two area of interests are defined near the user to guarantee their safety. Here, depth information at 1 m and 2 m are analyzed for retrieving information about the obstacles in the path from two distances.

An omnidirectional wearable system [61] for locating and guiding the individual in an indoor environment combined GIST and SURF for feature extraction. Two-levels of topological classification are defined in this system, namely, global and local. The

global classification will consider all images as references, whereas the local classification will be based on prior knowledge. A visual odometry module was developed by integrating extended Kalman filter monocular SLAM and omnidirectional sensors. The system was trained using 20,950 omnidirectional images and tested on 7027 images. Localization errors were present due to misclassified clusters.

Huang et al. [62] developed an indoor positioning system called 3DLoc, which is a 3D feature-based indoor positioning system that can operate on handheld smart devices to locate the user in real time. This system solves the limitation that exists in previous indoor navigation systems based on sensors and feature matching (e.g., SIFT and SURF), and it considers the 3D signature of pictures of places to recognize them with high accuracy. An algorithm to obtain the signatures from pictures has been proposed. The algorithm is capable of robustly decoding those signatures to identify the location. At the first stage, 3D features are extracted from the captured pictures. Therefore, a 3D model is constructed using the obtained features using the indoor geometry reasoning [63]. Pattern recognition is then performed to identify the 3D model. The authors proposed a K-locations algorithm to identify the accurate location. An augmented particle filter method is used if the captured images are insufficient for recognizing the location due to information loss. Inertial sensors of the mobile device are used to provide real-time navigation of users under motion. Based on the conducted experiments, 90% of the exposed errors are within 25 cm and 2° for location and orientation, respectively.

iNavigation [64] combines SIFT feature extraction and an approximately nearest neighbor algorithm called ‘K-d tree based on the best bin’ first for positioning from the ordinary sequential images. Inverse perspective matching was used for finding the distance when an image was queried by the user. Dijkstra’s algorithm was implemented for routing through the shortest path. In this method, locations of landmark images were manually assigned. Therefore, further expansion of landmark image datasets requires a considerable amount of manual work.

Image processing-based indoor localization method [65] for indoor navigation utilizes the principal component analysis (PCA)-SIFT [66] feature extraction mechanism to reduce the overall running time of the system compared to that of SIFT- or SURF-based methods. It also implemented a Euclidean distance-based locality sensitive hashing technique for rapid matching of the images. The precision of the system increased up to 91.1% via the introduction of a confidence measure.

The localization algorithm [67] for indoor navigation apps consists of an image edge detection module using a Canny edge detector and text recognition module using stroke width transform, Tesseract, and ABBY fine reader OCRs. Tesseract is a free OCR software that supports various operating systems, and its development has been sponsored by Google. Tesseract can support the recognition of texts in more than 100 languages, including the languages written from right to left, such as Arabic. The ABBY fine reader OCR is developed by “ABBY”, a Russia-based company, and it supports approximately 192 languages. Further, its latest version is able to convert texts in the image files to various electronic documents, such as PDF, Microsoft Word, Excel, Power Point, etc. The experimental results proved that ABBY is quick and has high recognition accuracy on a benchmark dataset used in research on OCR and information retrieval.

Xiao et al. [68] proposed a computer vision-based indoor positioning system for large indoor areas using smartphones. The system makes use of static objects in the indoor areas (doors and windows) as the reference for estimating the position of the user. The proposed system contains mainly two processes as follows: (a) static object recognition and (b) position estimation. In the static object recognition process, initially, the static object is detected and identified by implementing the Fast-RCCN algorithm [69]. The included deep learning network is similar to VGG16 network [70]. The pixel coordinate of the “control points” (physical feature points on the static object) in the image is used for position calculation of the smartphone. The pixel coordinates of “control points” were calculated by analyzing the test image and identified reference image. The SIFT feature detector is adapted for the extraction of feature points from both the test and reference images. A homographic matrix is constructed from the matching feature point pairs of test and reference images. This homographic matrix and “control point” of reference images are utilized to find the “control point” of the test image. The collinear equation model of the “control point” in the image and “control point” in the space is calculated for the position estimation of the smartphone. The results show that the system has achieved an accuracy within 1 m for position estimation.

A visual indoor positioning system that makes use of a CNN-based image retrieval method was proposed in [71]. The system database contains images for each scene, and its CNN features, absolute coordinates and quaternion are provided with respect to a given local coordinate system and scene labels. In the offline phase, the CNN features of the images related to each scene were extracted using the pretrained deep learning VGG16 network. The proposed system consists of the following two online phases: (1) image retrieval task based on CNN and (2) pose estimation task. During the image retrieval phase, the CNN will retrieve most similar images (two images) with respect to the query image. In the pose estimation phase, the “Oriented Fast and Rotated Brief (ORB)” [72] feature detector is adapted for feature extraction of three images (test image and retrieved most two identical images). The feature points of the test image are matched with each similar image using the Hamming distance. The scale of the monocular vision is calculated from the pose of the two identical images and transformation of matches between pairs of the test image and identical image. The position and orientation of the test image is calculated by utilizing the monocular vision and transformation between the test image and the identical image. Images from two benchmark datasets, the ICL-NUIM dataset [73] and the TUM RGB-D dataset [74], were used for system evaluation. The average error in pose estimation using ICL-NUIM and TUM RGB-D was 0.34 m, 3.430 and 0.32 m, 5.580, respectively. In the ICL-NUIM dataset, the proposed system exhibited less localization error compared to PoseNet [75], 4D PostNet and a RGB-D camera pose estimation method that combines a CNN and LSTM recurrent network [76].

PoseNet is a 6 DOF camera relocalization system for indoor and outdoor environments using a deep learning network. The PoseNet used a 23 convolutional layer model that is similar to GoogLeNet [77] for classification. Caffe library was utilized for implementing the PoseNet model.

A considerable number of elderly people may fall and become injured because of aging. In this scope, a smartphone-based floor detection module for structured and

unstructured environments that enables the identification of floors in front of the user is proposed in [78]. Structured environments are the areas that have a well-defined shape, and unstructured environments are the area with the unknown shape. In unstructured environments, superpixel segmentation was implemented for floor location estimation task. Superpixel segmentation will generate clusters of pixels, and they are then reshaped based on their color surroundings. For the structured environment, the Hough transform is used for line detection and the floor-wall boundary is represented by a polygon of connected lines. The results demonstrate that the system achieved an accuracy of 87.6% for unstructured environments and 93% for structured environments.

Stairs, doors, and signs are the common objects that can be used as reference points to guide people with visual impairments in indoor areas. Bashiri et al. [79] proposed an assistive system to guide people with visual impairments in indoor areas. The proposed system consists of two modules; a client mobile devices to capture the images and a processor server to detect the objects in the image. A CNN model was utilized to recognize indoor objects such as stairs, doors, and signs to assist people with visual impairments. The transfer learning technique was leveraged to build the object recognition CNN model. A popular CNN model AlexNet was utilized for the transfer learning method to create the new CNN model. The developed CNN model has evaluated in MCindoor 20000 dataset [80] and achieved recognition accuracy of more than 98%.

Jayakanth [81] examined the effectiveness of texture features and deep CNNs for indoor object recognition to assist people with visual impairments in indoor environments. The performance of three texture features LPQ, LBP, BSIF and CNN model built by the transfer learning approach using a pre-trained GoogleNet model was evaluated in this work. All of the proposed methods were evaluated in MCindoor 20000 dataset. Obtained results show that the CNN model built by the transfer learning approach using a pre-trained GoogleNet model achieved recognition accuracy of 100%. Although LPQ computation doesn't require any high-performance computing tools like what CNN computation required, the LPQ feature descriptor displayed a similar performance compared to CNN for indoor object recognition.

Afif et al. [82] extended a famous deep convolutional neural network called RetinaNet for indoor object detection to assist the navigation of people with visual impairments in indoor areas. The proposed object detection network is comprised of a backbone network and a pair of sub-networks. Among two sub-networks, the first network will perform object classification and the second network will extract the bounding box as well as the class name of objects. Feature pyramid networks are used as a backbone of the proposed detection network. Feature pyramid network-based architecture can detect objects on various scales which improves the performance of multi scales predictions. Evaluation of the proposed object detection network was carried out in a custom dataset which contains 8000 images and 16 different indoor landmark objects. During the evaluation of the proposed detector, different backbone network architectures such as ResNet, DenseNet, VGGNet have been experimented with RetinaNet. RetinaNet with ResNet network outperformed all other combinations and achieved a mean average precision of 84.16%.

An object recognition method [83] for indoor navigation of robots was developed using a SURF-based feature extractor and bag-of-words feature vectors using Support

Vector Machine (SVM) classifier. The nearest neighbor algorithm or RANSAC algorithm enabled feature vector matching. The proposed method was not able to recognize multiple objects in a single frame.

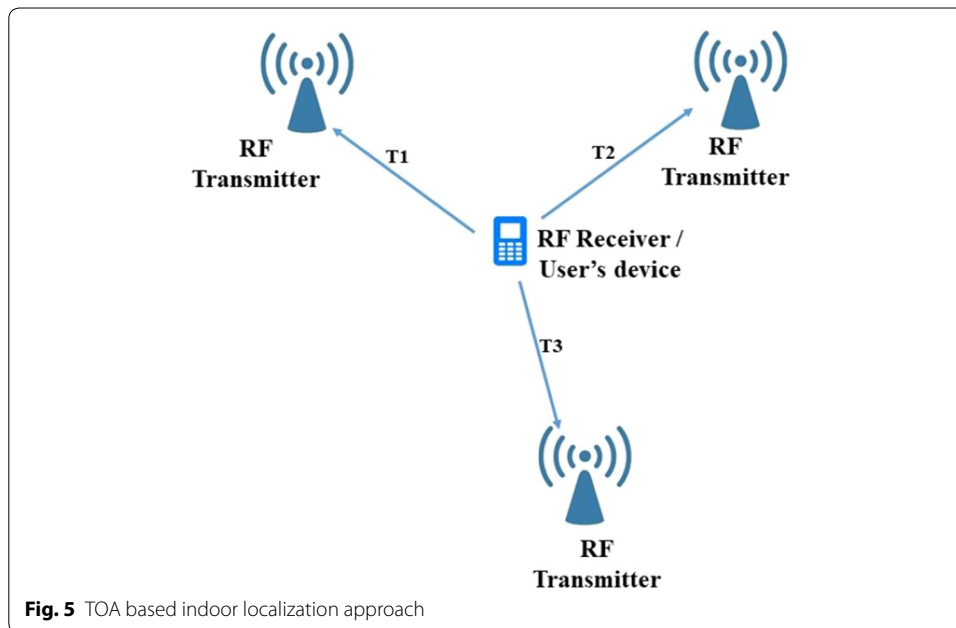
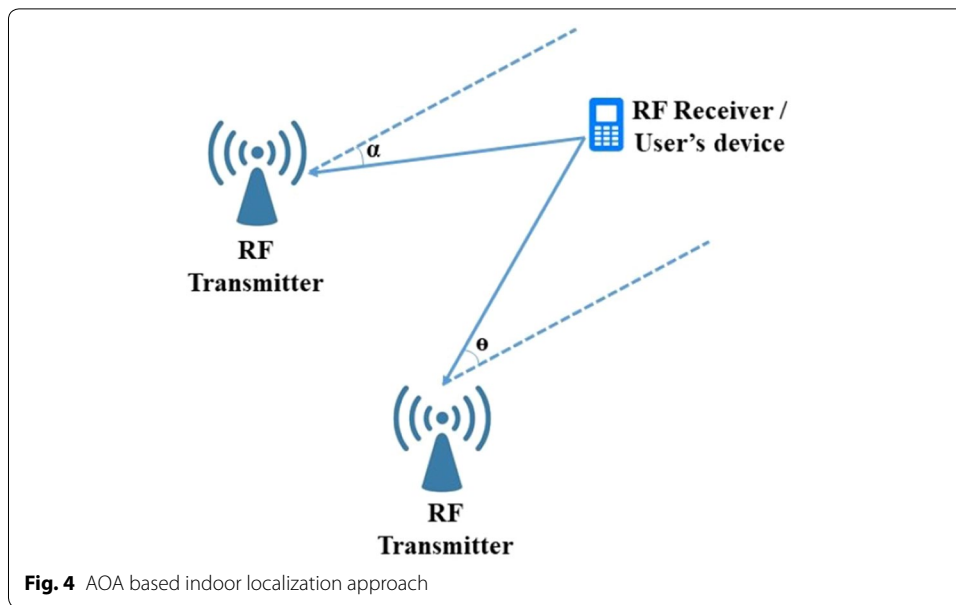
Table 2 presents a comparison of computer vision-based indoor positioning, indoor localization and indoor scene recognition systems.

Communication technology based indoor positioning and wayfinding systems

Communication technology-based positioning systems make use of various approaches to measure the signals from respective signal transmitting devices (Wi-Fi access point, BLE beacon etc.) installed in the indoor environments. The commonly used methods are time-based methods, angle-based methods, and RSS-based methods [84]. The time-based measurements include TOA and TDOA. The TOA approach utilizes the time taken for the signal propagation between the transmitter and receiver to find the range of the user, while the TDOA approach uses the difference of transmission time for two signals that have a different velocity. The angle-based method 'AOA' makes use of the angle of arrival at the target node to estimate target direction. The AOA measurement technique is rarely used in an indoor environment due to non-line of sight issues [85]. AOA and TOA based indoor localization approach are shown in Figs. 4 and 5 respectively.

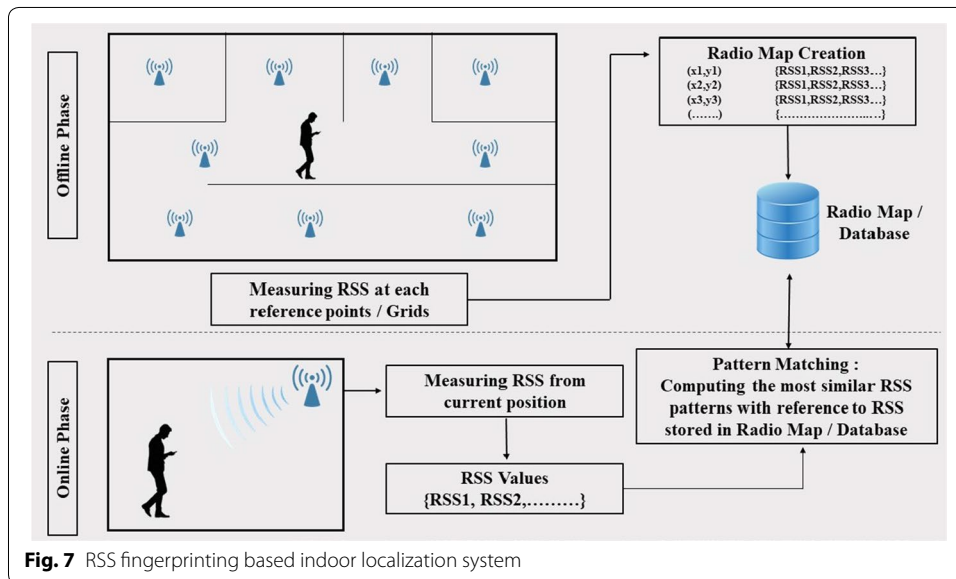
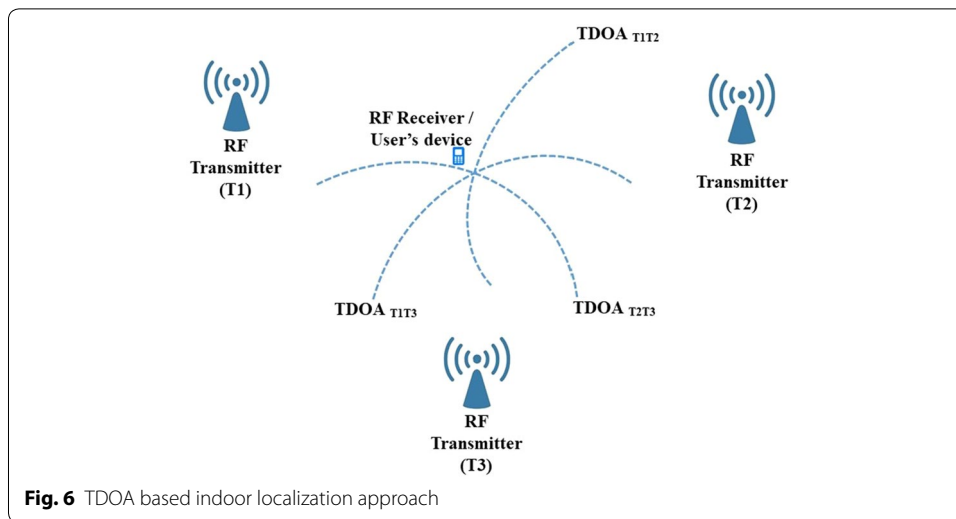
Table 2 Computer vision-based positioning, localizing and scene recognizing systems

References	Purpose	Solution	Performance /findings
Huang et al. [62]	Indoor positioning	3D signature of places for feature detection, Novel K-locations algorithm	(+) 90% of the exposed errors are within 25 cm and 2° for location and orientation respectively
Kawaji et al. [65]	Indoor positioning	PCA-SIFT features and locality sensitive hashing	(+) Running time reduced while comparing with the pure SIFT features-based system
Deniz et al. [67]	Localization using texts in boards and banners	Canny edge detector, Tesseract and ABBY fine reader OCRs	ABBY fine reader showed better recognition rate than Tesseract
Adorno et al. [78]	Floor detection method	Superpixel segmentation and Hough line transform	(+) Accuracy: 87.6% for the unstructured environment and 93% for the structured environment
Murillo et al. [61]	Personal localization in indoor areas	GIST and SURF-based feature detector, Extended Kalman filter monocular SLAM	(+) 82% correct localization
Xiao et al. [68]	Indoor positioning in large indoor areas	CNN and SIFT features	(+) Low cost, accuracy: less than 1 m
Chen et al. [71]	Indoor positioning	CNN and ORB features	(+) Average position error: less than 0.35 m
Kendall et al. [75]	Indoor and outdoor localization	CNN	(+) Robust to various lighting and motion blur scenarios
Bashiri et al. [79]	Indoor object recognition to assist people with VI	Transfer learning based on the CNN model (AlexNet)	(+) Accuracy: 98%
Jayakanth [81]	Indoor object recognition to assist people with VI	CNN and texture features	(+) Accuracy: 100%
Afif et al. [82]	Indoor object detection to assist people with VI	CNN	Mean average precision: 84.16%



TDOA method computes the difference between the TOA of the signals from two distinct RF transmitters to the mobile device. A TDOA value geometrically represents a hyperbola as shown in the figure. When there is more than one TDOA value, the intersection point of hyperbolas is estimated as the position of the mobile device. Figure 6 illustrate the TDOA based indoor localization approach.

Lateration, angulation, proximity, and radio fingerprinting are the main techniques used in communication technology-based systems for position estimation. The lateration technique calculates the distance between the receiver device and cluster of



transmitting devices (access points, tags or beacons) that are attached in predefined locations. The angulation technique is similar to the lateration technique but considers the angle or phase difference between the sender and receiver instead of distance for position estimation [86].

The proximity technique is based on the proximity of the receiver to recently known locations. Compared with lateration and angulation, the proximity technique can provide a rough location or set of possible locations. The radio fingerprinting approach is an entirely different approach compared to the other techniques and does not consider the distance, angle or nearness between sender and receiver. Instead, a pattern matching procedure is applied, where the RSS or other signal properties at a location will be compared with the RSS for different locations stored in the database [87]. General steps involved in RSS fingerprint-based localization system are explained in Fig. 7. For pattern

matching, different types of algorithms including Euclidean distance, machine learning algorithms such as KNN, SVM, etc are used in the literature.

FreeNavi [88] is a mapless indoor navigation system that relies on the Wi-Fi fingerprints of each landmark's entrance in the indoor environment. Along with Wi-Fi fingerprints, walking traces of the users between two landmarks was utilized for creating virtual maps of the indoor environment. A lowest common subsequence (LCS) algorithm [89] that finds similarities between Wi-Fi fingerprints was adopted for virtual map creation as well as indoor localization. The LCS algorithm was developed to tolerate access point (AP) changes in regions where the concentration of the APs are high. To provide reliable navigation of users, two route planning algorithms were introduced in this system. One was for finding the shortest path between two landmarks while the other was for finding the most frequently traveled route. Both of the abovementioned algorithms were implemented using Floyd's shortest path algorithms. FreeNavi was evaluated in a shopping center environment in Beijing by collecting the fingerprints of 23 landmarks and a total of 1200 m of traces. The virtual maps have a maximum accuracy of 91%, although an 11.9% error step rate was found in navigation because the user have to guess the travelling direction in junctions.

A Wi-Fi fingerprinting-based navigation system was proposed in [90]. The proposed system makes use of Wi-Fi fingerprinting combined with a radio path loss model for the estimation of locations. The position estimation algorithm was based on particle filter and K nearest neighbor (K-NN) algorithms. Dijkstra's algorithm was implemented for the shortest path calculation between the source and destination. These authors also examined the performance of three fingerprint matching algorithms, namely, Kalman filter, unscented Kalman filter and K-NN. The results showed the average error while using each algorithm and the values were similar at approximately 1.6 m. However, K-NN had the greatest maximum error.

In a Wi-Fi-based indoor navigation system, the fluctuations in RSS can result in unfair positioning accuracy. To overcome these issues, a fingerprint spatial gradient (FSG) was introduced in [91]. The proposed method makes use of the spatial relationship of RSS fingerprints between nearby numerous locations. For profiling the FSG, these authors introduced an algorithm that picks a group of nearby fingerprints that advance the spatial stability as well as fingerprint likeness. A pattern matching approach is used for comparing the stored FSG and queried FSG using similarity measures, such as the cumulative angle function, cosine similarity or discrete Fréchet distance. The average accuracy of the position estimation was between 3 and 4 m.

In Wi-Fi-based indoor positioning and navigation systems, the radio fingerprinting approach has been used widely for estimating the position of the RF signal receiver. The fingerprinting approach follows a pattern matching technique where the property of the currently received signals is compared with the properties of the signal stored during the offline or training phase. In the last 10 years, various machine learning algorithms such as SVM [92], KNN [93], neural networks [92] have been utilized for pattern matching in radio fingerprint-based indoor localization methods. Compared to traditional machine learning algorithms, deep learning algorithms such as CNN, RNN, etc have demonstrated their effectiveness in various tasks such as image classification, text recognition, intrusion detection, etc. In this context, in recent years

deep neural network-based approaches [94, 95] have been used in fingerprint-based indoor localization systems.

Jin-Woo et al. [96] proposed an indoor localization system that utilizes CNN for the Wi-Fi fingerprinting task. Since the fluctuations in RSS and multipath issues can cause errors in location estimation, training with few data can lead to the development of ineffective models. The proposed method utilized 2-D radio maps as inputs to train the CNN model. The 2-D virtual map for the input has been created from the 1-D signals. The developed deep CNN architecture consists of four convolutional layers, two max-pooling layers, and two fully connected layers. Even though It is a lite deep CNN model it has outperformed all other deep neural network-based methods proposed before that and achieved a mean accuracy of 95.41%. Since 2-D radio maps are used for training the deep CNN, it can learn signal strength and topology of radio maps. This approach makes the proposed system robust to the small RSS fluctuations.

Mittal et al. [97] have adapted CNN for Wi-Fi based indoor localization system for mobile devices. The proposed work presents a novel technique for Wi-Fi signature to image transformation and a CNN framework namely CNN-LOC. Instead of training with the available dataset, they have constructed their database by collecting RSSI data from a test environment. One of the novelties of the proposed work is the conversion of RSSI data to image data. For each location, the collected RSSI data are converted to the grayscale image using the Hadamard product method. Similar to [14], this work has used a lite deep CNN model which comprises of five CNN layers. To improve the scalability of the system, CNN-LOC is integrated with a hierarchical classifier. Hierarchical classifiers are used to scale up the lite or small CNN architecture for larger problems. The proposed hierarchical classifier consists of three layers where the first layer is used to find the floor number, the second layer for detecting the corridor and the third layer for estimating the location of the mobile device. The system has been tested in 3 indoor paths extended over 90 m. The obtained results show that the average localization error was less than 2 m.

Ibrahim et al. [98] proposed an advanced approach to improve the localization accuracy by reducing randomness and noise found in RSS values. The time series of RSS values are applied to CNN as input. The hierarchical architecture of CNN was employed for predicting the fine-grained location of the user. The first layer is responsible for detecting the building, second and third layers will predict floor number and location of the user respectively. The proposed CNN model was evaluated in the UJIIndoorLoc dataset. The dataset consists of the Wi-Fi RSS fingerprints collected from multiple-multi storied buildings. Demonstrated results show that the proposed hierarchical CNN predicts the building and floor with an accuracy of 100%. The average error in localization is 2.77 m which acceptable in the case of Wi-Fi-based systems.

Li et al. [99] proposed a multi-modal framework for indoor localization tasks in the mobile edge computing environment. Presented work focuses on the multiple models' based localization approaches, its drawback and finally proposes theoretical solutions to overcome its shortcomings. There exist many machine learning models for RSS based indoor localization tasks. Even though they displayed their effectiveness in the test environment, but failed to repeat the same performance in practical situations. There are many factors like refrigerators, temperature, doors in indoor areas which can

affect the localization performance. Theoretically, building distinct models for distinct surroundings is an effective method for indoor localization. But multiple models based approaches will also have drawbacks. Too many models have to be built, the presence of unstable factors which affect RSS are the major drawbacks. To solve these issues, two combinatorial optimization problems are formulated: external feature selection problem and model selection for location estimation problem. NP-hardness of the problems is analyzed in this work.

Wireless technology based indoor localization systems are prone to errors because of non-line of sight issues, inconsistency in received signals, fluctuation in RSS, etc. In large scale wireless-based localization systems, while comparing with the number of sensors, information is sparse. The main challenge in these systems is recovering the sparse signals for further processing to localize the user. Compressive sensing is a popular signal processing technique to efficiently acquire and reconstruct signals. This technique is used in wireless-based indoor positioning systems [100, 101] to recover sparse signals efficiently. Many of the existing compressive sensing techniques are intended to solve the issues for a single application and it lacks dynamic adaptability. Zhang et al. [102] proposed a learning-based joint compressive technique to solve the challenges in compressive sensing techniques. They introduced a learning technique that can learn the basis of sparse transformation from compressive sensing measurement output instead of historical data. Acquiring a big amount of historical data is costly and learning from specific historical data can affect the dynamic adaptability.

A hybrid navigation system that combines magnetic matching (MM), PDR and Wi-Fi fingerprinting was proposed in [103]. Since such systems combine different approaches, the user can even navigate through the regions where Wi-Fi signals are poor or environments have indistinctive magnetic feature. The location of the user was resolved by calculating the least value of the mean absolute difference between the estimated fingerprint or magnetic profile and the predetermined value of the respective candidate in the database. An attitude determination technique [104] and PDR [105] method were integrated for implementing the proposed navigation algorithm. To improve the Wi-Fi and MM results, three separate levels of quality control method using PDR-based Kalman filter were introduced. The results demonstrated that the proposed method has an accuracy of 3.2 m in regions with sufficient number of APs and 3.8 m in regions with poor numbers of APs.

iBill [106] integrates an iBeacon device, inertial sensors, and magnetometer to localize the users in large indoor areas using a smartphone. iBeacon is a variant of BLE protocol developed by Apple Inc. [107]. The proposed system contains two operational modes. If the user is within the range of the beacons, then a RSS-based trilateration algorithm is adopted for localization. Otherwise, the system will enter the particle filter localization (PFL) mode, which considers magnetic fields and data from inertial sensors for localization. Since the PFL mode itself cannot compute the initial position of the user, it will assume the last location obtained in the iBeacon localization mode as the initial position of the user. The accelerometer data and gyroscope data are used for updating the location and direction of the particles, respectively. These particles are utilized to represent the walking distance and direction of the user. To overcome the limitation of using magnetic fields only for assigning weights for a particle in the particle filter method, the

system considered the probability distribution of step length and turning angles of particles to determine the weight. The iBILL system reduced the computational overhead of PFL and solved the limitations associated with the unknown initial location and heavy shaking of smartphone. iBill achieved less error in localization compared to the dead reckoning approach and Magicol [108]. Magicol system combines magnetic fields and Wi-Fi signals using a “two-pass bidirectional particle filter” for localization. Magicol consumes less power compared to systems that rely only on Wi-Fi signals. In the Magicol and dead reckoning approaches, the error in localization increased drastically while walking for more steps. However, iBill showed consistency in localization accuracy while walking for long time (more steps case) also.

Lee et al. [109] proposed an indoor localization system that utilizes inbuilt sensors in smartphones, such as Bluetooth receivers, accelerometers, and barometers. The RSSIs of the signal received from Bluetooth beacons are used for location estimation with the help of the trilateration algorithm. PDR has been used to reduce the uncertainty in RSS identifiers, which improved location estimation by tracing the direction and steps of a normal user. Atmospheric pressure determined using a barometer was utilized for vertical location estimation. Due to the limitations of sensors in smartphones, the proposed method could not deliver satisfactory results in real-world scenarios.

A simple but efficient Bluetooth beacon-based navigation system using smartphones was proposed in [110]. The system utilizes RSSI measurements for position estimation. The positioning algorithm [111] initially measures the RSSI from each beacon and perform a noise removal operation. The “Log-Path Loss model” [112] based on the mean of the RSSI values is utilized for the estimation of distance from each beacon. The algorithm implements the proportional division method to estimate the position of the users when they are near to two or more beacons. In the proportional division method, the line representing corridor where beacons are installed is divided with respect to the distance between two nearby beacons. When only one beacon is near to the user and another one is far, the algorithm assumes the user’s position is on the other side of that nearby beacon. Dijkstra’s shortest path algorithm was adopted for finding the shortest route for navigation. The system performed well in a small indoor region.

DRBM [113] is a dead reckoning algorithm that combines a “Bluetooth propagation model” and multiple sensors for improved localization accuracy. The “Bluetooth propagation model” utilized the linear regression method for feature extraction. An individual parameter that varies with the characteristics of the users was integrated with data from accelerometers for calculating the exact steps covered by the user. Subsequently, the results from Bluetooth propagation model and sensor-based step calculation method were fused using a Kalman filter for improving the accuracy of positioning. The results demonstrated that the positioning errors were within 0.8 m.

Reference [114] examined the performance of machine learning classifiers, such as SVM, Random Forest and Bayes classifier, for the Bluetooth low-energy beacon fingerprinting method. The experimental infrastructure was created using beacons provided by Estimote and iBeeks. Both types of beacons use Eddystone profiles developed by Google. These authors evaluated the performance of algorithms for different smartphones with a preinstalled fingerprinting Android app. Eddystone packets from each beacon are scanned over a period of time to obtain the RSSI values. The MAC of the

beacon and associated RSSI values are logged for further training processes. The open source project 'Find' was adapted for the whole task. Several machine learning algorithms are already available in 'Find' servers. The results showed that Random Forest increased the accuracy of positioning by 30% compared to the Bayes classifier and a 91% correct identification of location.

In recent years BLE beacons based technology has been used for the development of assistive navigation systems for people with visual impairments. A blind or visually impaired user with a minimum knowledge of smartphones can utilize these systems to find the indoor ways in train stations, museums, university premises, etc. Basem et al. [115] proposed a BLE beacons based indoor navigation system for people with visual impairments. The proposed system utilized the fuzzy logic framework for estimating the position of the user in indoor areas. The basic methodology utilized for indoor positioning is BLE fingerprinting. Authors analyzed the performance of various versions of the fingerprinting algorithm including fuzzy logic type 1, fuzzy KNN, fuzzy logic type 2 and traditional methods such as proximity, trilateration, centroid for indoor localization. The fuzzy logic type 2 method outperformed all other methods. The average error of localization obtained in the fuzzy logic type 2 approach is just 0.43 m.

Murata et al. [116] proposed a smartphone-based indoor localization system that can be extended for blind navigation in large indoor environments that contains multistoried buildings. The proposed work addressed six key challenges for smartphone-based indoor localization in large and complex environments. The challenges are associated with the mobility of the user and the nature of large scale environments. The challenges include accurate and continuous localization, scaling the system for multiple floors, Varied RSS values from the same transmitter to different devices located in the same location, varied walking patterns of individuals, signal delay, etc. The authors improved the probabilistic localization algorithm using various techniques to address the above-mentioned challenges. RSSI from BLE beacons and data from embedded IMUs in the smartphone are utilized for location estimation. The proposed system was evaluated in a large shopping mall (21,000 m² area) with 10 individuals including blinds and people with low vision. It is observed that the proposed techniques reduced the mean localization error from 3 to 1.5 m while using the probabilistic localization algorithm.

Ahmetovic et al. [117] proposed a smartphone-based indoor navigation system for people with visual impairments. The proposed system namely, NavCog relies on RSSI from BLE beacons and inbuilt sensors of smartphones for localizing the user in indoor areas. The location of the user was estimated using a fingerprint matching algorithm. There are many fingerprint matching algorithm proposed in the literature. Here, the author chose a variant of the KNN algorithm to compute the location of the user by matching the observed RSSI value with RSSI fingerprints stored during the offline stage. Apart from basic localization and navigation service, NavCog can notify the user about their surroundings regarding the point of interest or stairs or elevators etc. NavCog was evaluated in a university campus with the help of six people with visual impairments. They recorded all the experiments using a video camera to see whether the user is missing any turn during navigation, waiting for instructions, hitting any obstacles, etc. Current version of NavCog lacks the functionality to notify the user when they are traveling in the wrong way.

Kim et al. [118] proposed a smartphone based indoor navigation assistant for people with VI impairments. The system namely, StaNavi uses the smartphone and BLE beacons attached in the indoor areas to guide the users in a large train station. Along with the RSS from BLE beacons, data from the inbuilt compass of the smartphone is utilized to estimate the position and orientation of the visually impaired users. A commonly used indoor localization method called proximity detection technique was used to compute the user's position. The StaNavi system makes use of a cloud-based server for providing navigation route information. Similar to StaNavi, GuideBeacon [119] indoor navigation system also utilizes the smartphone compass and BLE beacons to estimate the position and orientation of the visually impaired users in the indoor environment. But GuideBeacon used the low-cost BLE beacons to facilitate indoor tracking. The position estimation procedure includes identification of the nearest beacons for a user by using the proximity detection technique. GuideBeacon can provide audio, haptic and tactile feedback to the visually impaired user. Reference [120] proposed an indoor navigation system for people with visual and hearing impairments. The proposed system utilized proximity detection as well as nearness to beacons techniques in localization algorithm to track the position of the user. It is noted that in the last 5 to 6 years only the development of BLE beacons based navigation systems have become popular. It can be due to the availability of smartphones for low cost, less cost of beacons compared to other RF transmitters which have been used before. In the case of Blind navigation, only a few BLE beacon based systems have been proposed in recent years.

ISAB [121] is a wayfinding system developed for assisting people with VI in libraries, and it utilizes various technologies such as Wi-Fi, Bluetooth, and RFID. Here, each communication technology was used for different purposes. First, Wi-Fi fingerprinting was used for localization and navigating through the entrance of the building to the desired floor. Floor plans of the indoor environment were represented as graphs, and Dijkstra's algorithm was implemented for path planning. Bluetooth technology was used for navigating users to the desired shelf where the desired item is placed. Each shelf contains a shelf reader where a Bluetooth module is attached. The user can pair their smartphones with this Bluetooth module, and the shelf reader will provide instructions to the user. Finally, RFID technology was implemented to find the desired item on the shelf, where each item is embedded with a RFID tag. Additionally, an effective user interface was developed for simplifying interactions of blind people with the system. The proposed system helped the users to reach towards a target with a maximum accuracy of 10 cm.

PERCEPT [122] is a RFID technology-based navigation system developed for people with VI. PERCEPT consists of passive RFID tags pasted on the indoor areas, a "glove" that consists of a RFID reader, and kiosks placed at entrances, exits of landmarks. The kiosks contain information about key destinations and landmarks. Additionally, an Android smartphone that provides instruction to the user through a text to speech engine. An Android phone will communicate with the glove and PERCEPT server using Wi-Fi, and Bluetooth. The directions provided by the PERCEPT system lacks proximity. Moreover, the direction was not presented in terms of steps or feet.

PERCEPT II [123] includes a low-cost navigation method using smartphones alone (gloves were omitted). The cost for system deployment was decreased by creating a survey tool for orientation and mobility that aids in labeling the landmarks. NFC tags were

also deployed in specific landmarks for providing navigational instructions by means of audio. The navigation module implemented Dijkstra's algorithm for route generation.

A RFID-based indoor wayfinding system for people with VI and elderly people was proposed in [124]. The proposed system consists of a wearable device and a server. The wearable device consists of a RFID reader that can read passive tags, an ultrasonic range finder for detecting obstacles in the path and a voice controller. The server comprises of a localization module as well as navigation module. The navigation module implements Dijkstra's algorithm for path planning. For efficient localization, authors considered the normal movements of a person with vision while developing the system. The navigation module was linked to an obstacle avoidance algorithm where obstacles are categorized as expected and unexpected by assigning a predefined probability measure. Again, these obstacles were categorized as mobile and fixed and a triangle set is formulated for detecting mobile obstacles. An earphone was also embedded in the system for providing effective guidance to the user.

Roll Caller [125] introduced a method that relates the location of the user and the targeted object based on frequency shifts caused in the RFID system. The Roll Caller prototype comprises of passive RFID tags attached to items, RFID reader with multiple numbers of antennas, and a smartphone with inertial sensors, such as accelerometers and magnetometers. An anchor timestamp was used to represent the value of the frequency shift. This anchor time stamp was integrated with inertial measurements, such as acceleration and the direction from the sensors of smartphone for allocating antennas. The proposed method reduced the system overhead since the location of person and item are not calculated separately. Instead, a spatial relationship between the object and users was introduced to locate them.

DOVI [126] combined IMU and RFID technology to assist the people with VI in indoor areas. DOVI's navigation unit consists of a chip (NavChipISNC01 from InterSense Inc.) has a three-axis accelerometer, barometer, and magnetometer. An extended Kalman filter was included for compensating the sensor and gravity biases. While RFID module was implemented to reduce the drift errors in IMU. Dijkstra's algorithm was implemented to estimate the shortest navigation routes. A haptic navigation unit was present in DOVI that provides feedback/instruction to the user about navigation by means of vibration.

Traditional RFID positioning algorithms were facing fluctuations in location estimation due to multipath and environmental interference in RFID systems. To take care of this issue a new positioning algorithm called BKNN is introduced in [127]. BKNN is the combination of Bayesian probability and K-NN algorithm. In the implemented UHF-RFID system, RSS values were analyzed using Gaussian probability distribution for localization. The irregular RSS were filtered out using Gaussian filters. Integration of Bayesian estimation with K-NN improved the localization accuracy. Hence, the average error in location estimation of the proposed system was approximately 15 cm.

A VLC technology-based navigation system that utilizes existing LEDs inside an indoor environment was proposed in [128]. The proposed system comprises of four LED bulbs attached to the ceiling of the room and they were interconnected using the same circuit to operate as a single optical transmitter. Trilateration algorithm was implemented for locating the receiver/user. Target's path was tracked using Kalman filtering

and sequential important sampling particle filtering methods. They also examined the performance of the Kalman filter and particle filter for tracking the users. The demonstrated results show that particle filter is better compared to the Kalman filter for user tracking.

AVII [129] is a navigation system for visually impaired people using VLC technology. Along with VLC-based positioning, a geomagnetic sensor was introduced for providing accurate direction. A sonar sensor was also embedded in the system for detecting obstacle along the navigation path. Dijkstra's algorithm was modified and utilized in the system that enables the user to select the best and shortest navigation routes. The system give instructions to the user through the embedded earphone in the form of audio signals.

In [130], a VLC-based positioning system was integrated with magnetic sensors of the Android smartphone for assisting people with VI in indoor environments. The proposed prototype consists of an Android phone for calculating the position of the user. A speech synthesizer system inside a smartphone provides instructions to users through the earphone. The latitude and longitude of each location will be stored as visible light ID in each visible light associated with that location. Once the visible light receiver obtains information about the light ID from visible light, it will transmit the ID to a smartphone via Bluetooth. The smartphone integrates this information with the directional calculation from geomagnetic sensors and provides route instructions to the user. The smartphone was attached to a strap that hung freely around the user's neck. Due to the irregular motion of the users, the strap swung more than expected, which led to errors in reading the geomagnetic sensor and errors in position estimation.

Reference [131] proposed a method of mitigating random errors in inertial sensors and removing outliers in a UWB system. Multipath and nonline of sight conditions were the reason for outliers in the UWB systems. The proposed system consists of a UWB system and an inertial navigation system. The inertial navigation system consists of an accelerometer, a gyroscope, and a magnetometer. The UWB system makes use of the TDOA method and least square algorithm for position estimation. An "anti-magnetic ring" was introduced to remove the outliers in the UWB system under non-line of sight conditions, and it was the first method to do so. For improved positioning accuracy, the information from the accelerometer and UWB system was fused using a "double-state adaptive Kalman filter" algorithm based on the "Sage-Husa adaptive Kalman filter" and "fading adaptive Kalman filter". The results showed that the inclusion of "anti-magnetic ring" and "double state adaptive Kalman filter" algorithm reduced the positioning errors.

Table 3 provides a comparison of communication technology-based indoor navigation and positioning systems.

Pedestrian dead reckoning based indoor positioning and wayfinding systems

Hsu et al. [132] developed a system that only depends on inbuilt sensors of smartphones and is devoid of any external infrastructure. Here, the user's steps were detected by integrating values of acceleration along three axes obtained from accelerometer data. The user's step length was calculated by combining the maximum acceleration values and minimum acceleration values. Since step length varies with the person, an individual parameter that varies with the users was also fused with the maximum and minimum

Table 3 Communication technology-based systems

References	Purpose	Solution	Performance/findings
Guo et al. [88]	Indoor navigation	Wi-Fi fingerprints	(–) The user should guess his direction at junctions
Wu et al. [91]	Indoor positioning	Wi-Fi fingerprint spatial gradient.	(+) Reduced the positioning errors caused by fluctuations in RSSI
Wu et al. [106]	Indoor localization	BLE, Magnetic field, PDR, Trilateration, Particle filter	(+) Reduced the issues associated with the unwanted shaking of smartphones
Li et al. [103]	Indoor navigation	Wi-Fi, magnetic fingerprint, PDR-based Kalman filter	(+) The system is scalable
Lee et al. [109]	Indoor localization	BLE (Trilateration algorithm) and PDR	(–) Limitation of sensors
Campana et al. [114]	Evaluation of machine learning algorithms for positioning	BLE fingerprints, Random Forest and Bayes classifier	(+) Random Forest algorithm increased the accuracy by 30% with respect to the Bayesian classifier
Abu et al. [121]	Wayfinding in library for people with VI	Localization: Wi-Fi fingerprints, Book finding; Bluetooth + RFID, Dijkstra's algorithm	(+) Decimeter level accuracy
Tsirmpas et al. [124]	Indoor navigation system for people with VI	RFID for localization, ultrasonic rangefinder, and IMU for safe navigation	(+) 99% Success rate during real time testing
Lin and Guo [125]	Indoor positioning	Frequency shift caused in the RFID system	(+) System overhead is reduced
Loconsole et al. [126]	Indoor navigation system for people with VI	Kalman filter to reduce gravity and sensor biases, RFID to reduce drift errors, Dijkstra's algorithm	(+) Drift errors got reduced significantly
Xu et al. [127]	To solve the fluctuations in RFID-based localization systems	RFID and Bayesian algorithm integrated with K-NN	(+) The average location estimation error was 15 cm
Ganti et al. [128]	Indoor navigation system	VLC (Trilateration algorithm), sequential important particle filter, and Kalman filter	(+) The particle filter method is better for user tracking
Jayakody et al. [129]	Indoor navigation system for people with VI	Visible light, geomagnetic sensor, and sonar for orientation and obstacle determination	(+) Implementation cost low
Martinez-Sala et al. [23]	Indoor navigation system for people with VI	UWB (TDOA, AOA approach) and A* algorithm	(+) Centimeter level accuracy (< 20 cm)
Fan et al. [131]	To reduce the error in UWB and IMU-based navigation system	UWB (TDOA approach) and IMU, double state adaptive Kalman filter	(+) Significant reduction in positioning errors
Murata et al. [116]	Indoor localization for blind navigation	RSSI from BLE beacons, IMUs, and probabilistic localization algorithm	(+) Reduced the error in localization up to 1.5 m
Ahmetovic et al. [117]	Indoor navigation for people with VI	BLE fingerprinting and IMUs	(–) Current version lack the functionality to notify the users traveling in the wrong path

Table 3 (continued)

References	Purpose	Solution	Performance/findings
Kim et al. [118]	Indoor navigation for people with VI	BLE beacons based proximity detection and IMUs	(+) Proposed system was tested in a complex, highly crowded environment and observed results showed its potential for large scale deployment
Cheraghi et al. [119]	Indoor navigation for people with VI	BLE beacons based proximity detection and IMUs	(+) Low-cost beacons reduced the implementation cost
Jin-Woo et al. [96]	Indoor localization	CNN based Wi-Fi fingerprinting approach	(+) The proposed system is robust to small RSS fluctuations
Mittal et al. [97]	Indoor localization	CNN based Wi-Fi fingerprinting approach in smartphone	(+) Average localization error < 2 m
Ibrahim et al. [98]	Indoor localization	CNN based Wi-Fi fingerprinting using time series of RSS value as input.	(+) Building and floor prediction accuracy: 100%

acceleration values to detect steps. Direction changes of the users were decided from the data provided by the gyroscope. Since the PDR approaches may result in errors in localization, a calibration mark is provided in the map and floor. The main limitations of this work were the variations in the data provided by sensors caused by the holding position of phones (like in pocket or bag) and the absence of a path planning algorithm, which increased the difficulty of navigation.

Hasan and Mishuk [133] proposed a PDR-based navigation method for smart glasses. Since PDR methods need sensors for acquiring data, these authors introduced the smart intelligent eye ware called “JINS-MEME”, which contains a three-axis accelerometer and gyroscope. Usually, PDR methods will calculate the current position of a user from his last known position; therefore, the initial position of the user should be known for tracking his position. Calculation of the current position of a user requires the step length of the user, number of steps covered by the user from his last known position and azimuth or heading angle of the user. Data from the three-axis accelerometer is used for step detection process. If the sensors are mounted on foot, then step detection can be achieved by utilizing the zero-velocity update. However, the sensors in this system were attached on to the smart glass, and the norm of the accelerometer was utilized for detecting steps. If the norm of the accelerometer crosses a predefined value, it is then considered as one step. Since step length varies with the user, a parameter is included for step length calculation and these parameter is obtained from an experiment where 4000 steps of 23 people were analyzed. In addition, the “extended Kalman filter” was introduced to merge the values from accelerometer and gyroscope. The data merging approach was utilized for calculating the heading angle to rectify the errors of PDR and gyro sensor such as bias, noises, tilts, etc. Instead of using data from sensors individually, fused data are more accurate for further calculation.

Ju et al. [134] proposed a PDR based navigation system for smartphones that uses “multiple virtual tracking (MVR)”, The proposed system solves the limitation of existing methods in which all walls and passages are considered parallel or perpendicular. It also solves the limitation of walking indoor for a long time. Microelectromechanical system IMUs are comprised of three-axis accelerometers, magnetometers, and gyroscopes and they are employed to calculate the position. The proposed system does not rely on existing infrastructure or designed maps. The proposed MVR algorithm uses the concept of the map-matching algorithm to examine potential pedestrian’s trajectories [135], and it passes through two stages; in the first stage, “the main track utilizes the dominant direction as a matching function when it is significantly reliable”, and in the second stage, “the data obtained on an ambiguous direct straight line to utilize the dominant direction is expanded by multiple virtual tracking for diversified cases”. Generally, the PDR system passes through the following four steps: (1) step detection, in which peak detection approach is employed to detect the accurate step; (2) step length estimation, in which the relationship between the walking status and the step length is determined by the accelerometer; (3) heading estimation, in which the cumulative error over time of the “Attitude Reference System” is used to reduce the accumulated heading error and an “extended Kalman filter (EKF)” is designed to compute the heading error as well as the gyroscope biases; (4) position calculation, in which the obtained heading value through the “Attitude Reference System” and the previous step are used to compute the user’s

current location and the Mahalanobis distance is employed to obtain the similarity between the dominant direction and the estimated heading. The proposed PDR-Multiple virtual tracking system passes through (1) basic tracking steps described earlier; (2) virtual trajectory awareness, in which a dominant direction block might be added to check if the user is in the dominant direction even in unreliable situations; and (3) “virtual track extension and reduction”, in which the virtual track is extended if there is a likelihood of a dominant direction when a new straight line appears. The experimental results of the proposed PDR-Multiple virtual tracking system demonstrated its effectiveness when compared with conventional PDR systems that use the dominant direction in sophisticated trajectories.

Hsu et al. [136] proposed a navigation system for pedestrian localization, and it is composed of a triaxial accelerometer, triaxial magnetometer, triaxial gyroscope, Bluetooth transmission module, and a microcontroller. The intention of the developed system is to reduce the integral error to accurately estimate and construct the walking trajectories of the pedestrian. The system does not require external positioning techniques, and it comprises of a wearable inertial navigation device and a computer device. The navigation device can be placed on the foot of the pedestrian to construct walking trajectories. A walking trajectory algorithm that consist of trajectory height estimation function, and trajectory reconstruction function, is implemented to build user's trajectories. The signals are received by the computer through the Bluetooth wireless module. The sensor merging method, which is based on “double-stage quaternion”, uses the EKF to merge the angular velocity, acceleration, and magnetic signals. A rotation stage is implemented to provide a stable rotation rate and eliminate the interference resulted from the global earth rotation rate. During the experiment phase, users were asked to wear the device and walk for both long and short periods. The distance error and the end-to-end error were approximately 8.33 m and 4.81 m, respectively. The average height errors were 6.42% and 3.60% for walking downstairs and upstairs, respectively.

In [137], a dead reckoning approach for estimating and tracking a user using smart handled phones was proposed. The proposed approach solves the limitation of positioning mobile phones and the surrounding environment, and it depends on identifying the relative variations in the distance traveled by users' walking direction. The actual position of a user is approximated by combining distance traveled and previous position information. Calibration is required to adjust the initialization of the algorithm to recognize the movement path of the user with respect to a reference system. The navigation path is estimated by a sequence of orthogonal segments. Each orthogonal segment consists of the distance computed in steps and the movement heading determined with respect to the reference system. The movement direction, or the heading, is computed from the measurement data of the magnetic field. If the algorithm is not capable of identifying the rotation of the user or if sudden variations occurred in the data, then the algorithm will create an alarm to perform the calibration again. The current PDR approach does not consider the slight rotation of a user, and an assessment of this approach indicated that the maximum error obtained in the system is not more than four steps.

PDR based indoor positioning systems are prone to error in localization because of sensors biases, drift, etc. Recent PDR based systems have introduced the multi-sensor fusion approaches and multiple positioning technology integrations [138–140] to reduce

the integral and drift errors observed in PDR based systems. Qiu et al. [141] proposed a multi-sensor fusion approach for alleviating the error present in traditional PDR based systems. The proposed work utilized a foot-mounted magnetic/inertial sensors to estimate the location of the user. The foot mounted sensor 'Xsens' contains 3 axis accelerometer, gyroscope, and magnetometer. A walking stance phase detection algorithm which utilizes the data from the accelerometer, as well as gyroscope, was introduced in this work. Sensor fusion is achieved by implementing the extended Kalman filter method. The performance of the sensor fusion approach was evaluated in both indoor as well as outdoor environments. Demonstrated results show that the overall error obtained in the proposed system was less than 1% of the total traveled distance.

The integration of multiple positioning technologies or sensor fusion approaches can improve the accuracy of PDR based systems and vice versa. It means PDR technologies have been integrating with other technologies to alleviate the positioning errors. Kuang et al. [142] proposed a smartphone-based indoor positioning approach that combines magnetic matching and PDR technology. Magnetic matching technology has been used alone or along with Wi-Fi or BLE technology for indoor navigation in recent years. While using Magnetic matching alone, it is hard for a system to differentiate the magnetic field at a single point from other near points. In this context, the authors came with a solution to integrate the magnetic field sequence along with the traveled path contour estimated by PDR technology. Moreover, the drift errors generated in the PDR method are regulated by an extended Kalman filter with the help of the reference magnetic field sequence. Gauss–Newton iterative technique was utilized to compare the measured magnetic field sequence with reference magnetic field sequences. The demonstrated result shows that the proposed method achieved an accuracy not more than 2.5 m with less computational load compared to other existing solutions.

A real-time smartphone-based indoor localization system that combines BLE technology along with the PDR approach has been proposed in [143]. The proposed system utilizes the inertial measurement units in the smartphone and RSS from the BLE beacons to estimate the position of the user in indoor areas. A smartphone application was built to fuse the inertial data and RSS from beacons. The inertial data is used to compute the step length as well as the heading angle. Step detection was achieved by analyzing the height of each jump during every zero crossings. If the height is above the specific pre-assigned threshold value, a step is detected. Zero crossings are the instants where a signal or value changes its sign. Step length estimation technique adopted a state of the art approach which require the length of the user's leg and vertical change of body center of mass as input along with a correction factor defined in the literature. The heading estimation procedure includes the fusion of data from accelerometer and gyroscope followed by computation of device's attitude. The attitude is further converted to quaternions and Euler angle representations. The relationship between these representations is used to find the heading angle. Later the step length and heading angle are fused with RSS from BLE beacons to reset the location of the user and decrease the drift error.

Shan-Jung et al. [144] propose an indoor navigation system which integrate PDR technology along with the Wi-Fi fingerprinting method. The work mainly focuses on calibrating the fingerprinting database with the aid of inertial measurement units in the smartphone device. A quaternion based extend Kalman filter was employed for sensor

fusion to reduce the positioning error in the PDR method. In the PDR method, accelerometer data were utilized to detect valid steps. Noise obtained in the accelerometer was filtered out by using a low pass filter. The step detection algorithm can able to identify the steps from both the magnitude phase as well as the temporal phase. A pair of peaks and valleys are considered as one step. An existing method is adopted for computing the step length. The step length computation method considers the standard deviation of the acceleration data since the stride length and walking speed are related. The landmarks in indoor areas can be utilized for providing directions to the user. Here, the Wi-Fi fingerprinting method has been utilized for identifying the nearest landmark. If the user has reached any landmark, the integrated errors in the PDR method will be reseted. In this way, the drift errors in the PDR system are alleviated.

Mercury [145] is a network localization and navigation system using smartphones for indoor applications. Localization of the user is achieved by fusing the various measurements from the IMU and range measurements among the users. The range measurements between various users are acquired using acoustic signals. The Mercury system consists of a built-in IMU, speaker, earphone of a smartphone and Bluetooth transceiver. Here, the users will localize themselves by utilizing the temporal cooperation, the spatial cooperation among them along with the knowledge about the map. The acceleration and angular velocity provided by the IMU were utilized for phone orientation estimation [146]. The obtained acceleration samples based on the phone orientation were transferred from the “phone’s coordinate system” to the “Earth’s coordinate system”. The step direction and step length of users were calculated by utilizing the acquired acceleration samples. For range measurements, the acoustic signals were recorded and transmitted using earphones and speakers, and these tasks are performed when a user performs spatial cooperation. Spatial cooperation is achieved by finding the range between the user followed by exchanging the position information. A user will try to sense the acoustic signals produced by another user to measure the range. If the acoustic signals are absent, then the user will perform two-way ranging and measure the time taken for two-way propagation to estimate the range [147]. The step direction and step length calculated from the IMU measurements and range measurements were fused with map information using an algorithm called Belief Propagation . The belief propagation algorithm will find the positional belief of the user. Bluetooth is used to exchange the positional belief of the users among them. During system evaluation the map of the indoor environment was partitioned as small squares with dimensions of $0.7\text{ m} \times 0.7\text{ m}$ because the belief propagation algorithm requires this partition. In case of single user scenario, the Mercury system is compared with the two systems Mapcraft [148] and a system using a Kalman filter technique [149]. The Mercury system proved its robustness compared to the other two systems, even in the absence of the user’s initial position. In multi-user scenarios, the Mercury achieved an exceptional localization performance because of spatial cooperation. Table 4 shows a comparison of PDR-based navigation systems.

Table 4 PDR-based navigation systems

References	Solution	Performance	Findings/remarks
Hasan and Mishuk [133]	Data from the accelerometer and gyroscope fused using the Kalman filter	Average deviation: 2 m	(−) The system cannot follow the path of the user if the positioning updates from RF signal are absent for a minimum duration of 5 min
Ju et al. [134]	PDR system with multiple virtual tracking to avoid drift errors in heading angle	Position error: 0.77% of total distance (1430 m)	(+) It solves the limitation of existing methods that assume all of the walls and corridors are parallel or orthogonal
Hsu et al. [136]	Trajectory reconstruction algorithm, Double filter quaternion-based adaptive Kalman filter for sensor fusion	Distance error: 0.52% (5.28 m) of total traveled distance	(+) Reduced the integral error in trajectory reconstruction and estimation without utilizing any other external positioning technique
Liu et al. [145]	Belief propagation algorithm to localize the user by fusing sensory information, and the range information between users	80th percentile of localization error was 1.6 m and 3.5 m	(+) The system is robust towards unknown initial position scenarios and multiple user scenarios
Giorgi et al. [137]	PDR-based navigation system utilizing the embedded sensors of smartphones	Maximum error: 4 steps	(+) Low-cost system
Qiu et al. [141]	Multi-sensor fusion approach using extended Kalman filter	Distance error: 1% of traveled distance	(+) Sensor installation and path propagation errors are rectified
Kuang et al. [142]	PDR approach integrated with magnetic matching	Accuracy: less than 2.5 m	(+) Introduction of magnetic matching approach reduced the drift errors in PDR
Ciabattini et al. [143]	PDR approach integrated with BLE technology	Distance error: 0.18 m when whole beacons are functioning	(+) BLE beacon-based approach reduced the drift errors in PDR
Shan-Jung et al. [144]	PDR approach integrated with the Wi-Fi fingerprinting technique	Pathway error is almost zero in simulation experiments	Wi-Fi fingerprinting based calibration system reduced the cumulative error in PDR

Evaluation criteria

In this section, we propose criteria to be considered while evaluating indoor positioning and navigation systems. These proposed criteria will be helpful for investigations into positioning and navigation system. Moreover, considering these factors in the development stage can result in an ideal navigation system.

Accuracy and precision

Accuracy is one of the main performance metrics of a navigation system. This metric is mainly associated with the indoor positioning module of the navigation system. The error in localization is expressed in terms of accuracy. It is computed as the average Euclidean distance between the ground truth location coordinates and estimated

location coordinates. A more realistic approach compared to Euclidean distance was introduced in [150]. This method depicted floor plans, obstacles and interfloor traversing routes as polygons for error estimations.

Precision deals with the consistency of system performance, or the consistency in positioning over time and various scenarios. The precision of the system can be represented in terms of the cumulative distribution function. In normal human navigation, the fall-off in these metrics up to a limit can be acceptable, although for people with VI, these fall-offs may affect their safety.

Cost

The cost of navigation systems can be classified among the positioning, navigation and HMI modules. In particular, the cost of the positioning module includes the individual costs of infrastructure components and their maintenance and devices for position estimation. The employment of Wi-Fi or VLC-based systems can reduce the cost of the infrastructure components because most building are installed with Wi-Fi APs or LEDs. However, the initial implementation cost is high for Wi-Fi APs compared with Bluetooth beacons or RFID tags. The cost of navigation modules is associated with the adopted map construction methodologies. Google indoor maps is an open source SDK, but its service is limited to few countries. Several other paid map building SDKs are available in the market.

Usually, the HMI module does not account for a large contribution to the cost of the whole system. In smartphone-based navigation or positioning systems, users utilize speakers, microphone and earphones to interact with the system. The noises from the navigation environment create difficulties for people with VI when using audio feedback. Thus, a haptic feedback system must be implemented along with audio feedback

Scalability

The scalability of the system can be evaluated by considering two parameters, namely, geography and the number of users. Geography represents the area of the indoor environment covered. An increase in the number of users in the same region can create confusion in positioning due to interference from the signals of communication technology-based systems. Computer vision-based systems also encounter problems in indoor scene recognition due to occlusions created by other users.

Robustness

Robustness means the ability of the system to withstand adverse conditions, such as component malfunctions and losses of signal. The system should provide the navigation and tracking of the user even if one or two infrastructure components fail or malfunction. In particular, Wi-Fi based system or BLE based system should work properly even if one or two Wi-Fi APs or BLE beacons fails to work.

Usability

Navigation systems are developed for reducing manual effort and time in the wayfinding process. The system design should consider the preferences of consumers. In this

context, the size of the system, power consumption, and real-time performance have to be considered. The positioning systems should pull location information in real time, and the navigation module should provide real-time route presentation and turn by turn directions. This design suits people with no disabilities or physical impairments. However, people with VI require additional feedback, such as haptic feedback, an obstacle detection module, and a location awareness module for compensating their disability.

Future work and discussion

Computer vision-based navigation and positioning systems can provide better mindfulness about encompassing environments compared with systems that utilize communication technologies or PDR approaches. Thus, computer vision-based frameworks are more appropriate for navigation by individuals with VI. In computer vision-based systems, deep learning methodologies are observed to be more precise than pure conventional methodologies. The hybrid technique that utilizes deep learning methodologies for scene recognition or image retrieval tasks and SIFT or ORB features for position estimation achieved better accuracy compared with pure deep learning methodology-based systems. Three-dimensional feature-based localization methods solved the limitations associated with SIFT- or SURF-based matching methods. The impact of human occlusions adversely affects visual feature-based positioning [81]. The elimination of human objects from visual scene recognition process can be further explored to solve these issues for both static and dynamic camera setups.

Compared to 2-dimensional image feature based approaches 3-dimensional features and RGB-D image-based methods are more reliable for indoor navigation. Visual positioning systems which are considered as the future of indoor navigation technology utilize RGB-D images to train and learn the models for localizing the user in indoor areas. Most of the RGB-D based methods have not extended to a fully working indoor navigation system. Instead many articles have proposed the methods and done offline testing in the publicly available dataset. Only a few works had extended RGB-D indoor positioning techniques to fully working indoor navigation systems, but still, it is implemented in a client-server manner. Because every mobile device is not able to bear the high computation required for position estimation tasks. Optimization of visual indoor positioning models to deploy in mobile devices like smartphones is one of the least explored topic in this research domain. There exists some systems that utilized Google tango VPS for the development of indoor navigation. But Google tango is supported in only a few devices and currently, Google has stopped the support on Google tango. But Google's new tool ARCore provides similar features like Google tango. It can be extended for the development of visual positioning based indoor navigation systems.

Communication technology-based approaches that integrate PDR methodologies or magnetic fingerprinting methods improved the coverage and precision of the system. The drift errors and initial position estimation problems of PDR-based systems are alleviated by introducing communication technologies, such as BLE and Wi-Fi, or magnetic fingerprinting approaches along with PDR. Fingerprint spatial gradients (spatial relation between RSS fingerprints of nearby locations) reduced the issues associated with RSS fluctuations. PDR systems integrated with Bluetooth technology seem to be more precise, and such systems can be further extended to correct radio maps of Wi-Fi.

Other than seven indoor positioning technologies discussed in this article, there exist many other technologies like audio signal based localization [151], magnetic field-based localization [152], etc. Audio signal based localization or acoustic localization is more accurate and cheaper compared to other RF technologies [153]. Because acoustic localization requires microphones and speakers which are available in every smart mobile device. Moreover, RF signal speed is very much higher compared to the sound and it implies acoustic localization can provide higher accuracy. In this context, acoustic localization technology can be combined with BLE or Wi-Fi-based approaches, where BLE or Wi-Fi can be utilized for rough location estimation and acoustic signals for computing the precise location.

Step length estimation is crucial task in PDR-based systems. Step length depends on the user's movement, velocity, and physical properties, such as height. The step length estimation task is still an open research issue. Precise step length estimation can minimize the accumulated errors of PDR systems. The step length for walking, running and walking with heavy load will be different for a same person. Differentiating the different walking scenarios is still an open challenge in the research of PDR based navigation system. Another interesting research direction in the field of PDR systems is deploying deep learning algorithms to determine the type of pedestrian movement by utilizing data from the gyroscopes and accelerometers installed in smartphones.

Conclusion

This paper presents a detailed overview of the advancements in systems for indoor positioning and wayfinding. We classified the existing systems based on the adopted positioning technologies. We provided a comprehensive review of the various proposed indoor positioning and wayfinding methods in the last 6 years. Moreover, this work analyzed its advantages and limitations. This article also discussed different assessment criteria for evaluating navigation and positioning systems. We further provided potential research directions for future research in indoor positioning and wayfinding systems.

Abbreviations

GPS: Global Positioning System; RF: Radio frequency; PDR: Pedestrian dead reckoning; RFID: Radio Frequency Identification; UWB: Ultra wide band; GLONASS: Global Navigation Satellite System; HMI: Human machine interaction; SURF: Speeded Up Robust Feature; SIFT: Scale Invariant Feature Transform; VLC: Visible light communication; RSS: Received signal strength; AOA: Angle of arrival; TOA: Time of arrival; TDOA: Time difference of arrival; BLE: Bluetooth low energy; VI: Visual impairments; IMU: Inertial measurement unit; OCR: Optical character recognition; SLAM: Simultaneous localization and mapping; CAM: Continuous adaptive mean; PCA: Principle component analysis; ORB: Oriented Fast and Rotated Brief; SVM: Support Vector Machine; LCS: Lowest common subsequence; AP: Access point; K-NN: K nearest neighbor; FSG: Fingerprint spatial gradient; MM: Magnetic matching; PFL: Particle filter localization; MVR: Multiple virtual tracking; EKF: Extended Kalman filter.

Acknowledgements

Not applicable.

Authors' contributions

JK and AK carried out the work and drafted manuscript. SA supervised the work and reviewed the manuscript. AA reviewed the manuscript and provide assistance to improve it. All authors read and approved the final manuscript.

Funding

This publication was supported by a Qatar University Collaborative High Impact Grant QUHI-CENG-18/19-1. The findings achieved herein are solely the responsibility of the authors. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of Qatar University.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 27 December 2019 Accepted: 4 April 2020

Published online: 02 May 2020

References

- Godha S, Lachapelle G (2008) Foot mounted inertial system for pedestrian navigation. *Meas Sci Technol* 19(7):075202. <https://doi.org/10.1088/0957-0233/19/7/075202>
- Meers S, Ward K (2005) A substitute vision system for providing 3d perception and gps navigation via electro-tactile stimulation
- Koyuncu H, Yang SH (2010) A survey of indoor positioning and object locating systems. *IJCSNS Int J Comput Sci Netw Secur* 10(5):121–128
- Zandbergen PA, Barbeau SJ (2011) Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. *J Navig* 64(3):381–399. <https://doi.org/10.1017/S0373463311000051>
- Bay H, Ess A, Tuytelaars T, van Gool L (2008) Speeded-up robust features (surf). *Computer vision and image understanding. Similarity matching in computer vision and multimedia* 110(3), 346–359. <https://doi.org/10.1016/j.cviu.2007.09.014>
- Wang H, Zhang S (2011) Evaluation of global descriptors for large scale image retrieval. In: *International conference on image analysis and processing*, Springer, pp 626–635
- Lindeberg T (2012) Scale invariant feature transform. *Scholarpedia* 7(5):10491. <https://doi.org/10.4249/scholarpedia.10491>
- LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436. <https://doi.org/10.1038/nature14539>
- Lee YH, Medioni G (2015) Wearable RGBD indoor navigation system for the blind. In: Agapito L, Bronstein MM, Rother C (eds) *Computer vision—ECCV 2014 workshops*. Springer, Cham, pp 493–508
- Kamisaka D, Muramatsu S, Iwamoto T, Yokoyama H (2011) Design and implementation of pedestrian dead reckoning system on a mobile phone. *IEICE Trans Inf Syst* 94(6):1137–1146
- Ban R, Kaji K, Hiroi K, Kawaguchi N (2015) Indoor positioning method integrating pedestrian dead reckoning with magnetic field and wifi fingerprints. In: *2015 eighth international conference on mobile computing and ubiquitous networking (ICMU)*, pp 167–172. <https://doi.org/10.1109/ICMU.2015.7061061>
- Woodman OJ (August 2007) An introduction to inertial navigation. Technical report UCAM-CL-TR-696, University of Cambridge, Computer Laboratory. <https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-696.pdf>. Accessed 10 Nov 2019
- Bouet M, dos Santos AL (2008) Rfid tags: positioning principles and localization techniques. In: *2008 1st IFIP wireless days*, pp 1–5. <https://doi.org/10.1109/WD.2008.4812905>
- Fu Q, Retscher G (2009) Active RFID trilateration and location fingerprinting based on rssi for pedestrian navigation. *J Navig* 62(2):323–340. <https://doi.org/10.1017/S0373463308005195>
- He S, Chan S-G (2016) Wi-fi fingerprint-based indoor positioning: recent advances and comparisons. *IEEE Commun Surv Tutor* 18(1):466–490. <https://doi.org/10.1109/COMST.2015.2464084>
- Farid Z, Nordin R, Ismail M (2013) Recent advances in wireless indoor localization techniques and system. *J Comput Netw Commun* 2013:12. <https://doi.org/10.1155/2013/185138>
- Do T-H, Yoo M (2016) An in-depth survey of visible light communication based positioning systems. *Sensors*. <https://doi.org/10.3390/s16050678>
- Alarifi A, Al-Salman A, Alsaleh M, Alnafessah A, Al-Hadhrani S, Al-Ammar MA, Al-Khalifa HS (2016) Ultra wideband indoor positioning technologies: analysis and recent advances. *Sensors*. <https://doi.org/10.3390/s16050707>
- Hart PE, Nilsson NJ, Raphael B (1968) A formal basis for the heuristic determination of minimum cost paths. *IEEE Trans Syst Sci Cybern* 4(2):100–107. <https://doi.org/10.1109/TSSC.1968.300136>
- Johnson DB (1973) A note on dijkstra's shortest path algorithm. *J ACM* 20(3):385–388. <https://doi.org/10.1145/321765.321768>
- Stentz A et al (1995) The focussed d* algorithm for real-time replanning. *IJCAI* 95:1652–1659
- Floyd RW (1962) Algorithm 97: shortest path. *Commun ACM* 5(6):345. <https://doi.org/10.1145/367766.368168>
- Martinez-Sala AS, Losilla F, Sánchez-Aarnoutse JC, García-Haro J (2015) Design, implementation and evaluation of an indoor navigation system for visually impaired people. *Sensors* 15(12):32168–32187. <https://doi.org/10.3390/s151229912>
- Katzschmann RK, Araki B, Rus D (2018) Safe local navigation for visually impaired users with a time-of-flight and haptic feedback device. *IEEE Trans Neural Syst Rehabil Eng* 26(3):583–593. <https://doi.org/10.1109/TNSRE.2018.2800665>
- Fallah N, Apostolopoulos I, Bekris K, Folmer E (2013) Indoor human navigation systems: a survey. *Interact Comput* 25(1):21–33. <https://doi.org/10.1093/iwc/iws010>
- Harle R (2013) A survey of indoor inertial positioning systems for pedestrians. *IEEE Commun Surv Tutor* 15(3):1281–1293. <https://doi.org/10.1109/SURV.2012.121912.00075>
- Davidson P, Piché R (2017) A survey of selected indoor positioning methods for smartphones. *IEEE Commun Surv Tutor* 19(2):1347–1370. <https://doi.org/10.1109/COMST.2016.2637663>
- Hassan NU, Naeem A, Pasha MA, Jadon T, Yuen C (2015) Indoor positioning using visible led lights: a survey. *ACM Comput Surv* 48(2):20–12032. <https://doi.org/10.1145/2835376>
- Li B, Muñoz JP, Rong X, Chen Q, Xiao J, Tian Y, Arditi A, Yousuf M (2019) Vision-based mobile indoor assistive navigation aid for blind people. *IEEE Trans Mob Comput* 18(3):702–714. <https://doi.org/10.1109/TMC.2018.2842751>

30. Cabaret L, Lacassagne L (2014) What is the world's fastest connected component labeling algorithm? In: 2014 IEEE workshop on signal processing systems (SiPS), pp 1–6. <https://doi.org/10.1109/SiPS.2014.6986069>
31. Rong X, Li B, Munoz JP, Xiao J, Arditi A, Tian Y (2016) Guided text spotting for assistive blind navigation in unfamiliar indoor environments. In: International symposium on visual computing, Springer, pp 11–22
32. Tian Y, Yang X, Yi C, Arditi A (2013) Toward a computer vision-based wayfinding aid for blind persons to access unfamiliar indoor environments. *Mach Vis Appl* 24(3):521–535. <https://doi.org/10.1007/s00138-012-0431-7>
33. Lee YH, Medioni G (2016) RGB-D camera based wearable navigation system for the visually impaired. *Comput Vis Image Underst* 149, 3–20, Special issue on Assistive Computer Vision and Robotics–Assistive Solutions for Mobility. Communication and HMI. <https://doi.org/10.1016/j.cviu.2016.03.019>
34. Huang AS, Bachrach A, Henry P, Krainin M, Maturana D, Fox D, Roy N (2017) Visual odometry and mapping for autonomous flight using an RGB-D camera. In: Robotics research, Springer, pp 235–252
35. Labbé M, Michaud F (2014) Online global loop closure detection for large-scale multi-session graph-based slam. In: 2014 IEEE/RSJ international conference on intelligent robots and systems, pp 2661–2666. <https://doi.org/10.1109/IROS.2014.6942926>
36. McDonald J, Kaess M, Cadena C, Neira J, Leonard JJ (2011) 6-dof multi-session visual slam using anchor nodes. In: European conference on mobile robots (ECMR), pp 69–76. <http://mural.maynoothuniversity.ie/6497/>
37. Garcia G, Nahapetian A (2015) Wearable computing for image-based indoor navigation of the visually impaired. In: Proceedings of the conference on wireless health. WH '15, ACM, New York, NY, USA, pp 17–1176. <https://doi.org/10.1145/2811780.2811959>
38. Manlises C, Yumang A, Marcelo M, Adriano A, Reyes J (2016) Indoor navigation system based on computer vision using camshift and d* algorithm for visually impaired. In: 2016 6th IEEE international conference on control system, computing and engineering (ICCSCCE), pp 481–484. <https://doi.org/10.1109/ICCSCCE.2016.7893623>
39. Bai J, Liu D, Su G, Fu Z (2017) A cloud and vision-based navigation system used for blind people. In: Proceedings of the 2017 international conference on artificial intelligence, automation and control technologies. AIACT '17, ACM, New York, NY, USA, pp 22–1226. <https://doi.org/10.1145/3080845.3080867>
40. Chen X, Liu X, Wang Y, Gales MJF, Woodland PC (2016) Efficient training and evaluation of recurrent neural network language models for automatic speech recognition. *IEEE/ACM Trans Audio Speech Lang Process* 24(11):2146–2157. <https://doi.org/10.1109/TASLP.2016.2598304>
41. Cho K, Van Merriënboer B, Gulcehre C, Bahdanau D, Bougares F, Schwenk H, Bengio Y (2014) Learning phrase representations using RNN encoder–decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*
42. Liang M, Hu X (2015) Recurrent convolutional neural network for object recognition. In: The IEEE conference on computer vision and pattern recognition (CVPR)
43. Farabet C, Couprie C, Najman L, LeCun Y (2013) Learning hierarchical features for scene labeling. *IEEE Trans Pattern Anal Mach Intell* 35(8):1915–1929. <https://doi.org/10.1109/TPAMI.2012.231>
44. Noh H, Hong S, Han B (2015) Learning deconvolution network for semantic segmentation. In: The IEEE international conference on computer vision (ICCV)
45. Akbani O, Gokrani A, Quresh M, Khan FM, Behlil SI, Syed TQ (2015) Character recognition in natural scene images. In: 2015 international conference on information and communication technologies (ICICT), pp 1–6. <https://doi.org/10.1109/ICICT.2015.7469575>
46. Neumann L, Matas J (2016) Real-time lexicon-free scene text localization and recognition. *IEEE Trans Pattern Anal Mach Intell* 38(9):1872–1885. <https://doi.org/10.1109/TPAMI.2015.2496234>
47. Shao K-Y, Gao Y, Wang N, Zhang H-Y, Li F, Li W-C (2010) Paper money number recognition based on intersection change. In: Third international workshop on advanced computational intelligence, pp 533–536. <https://doi.org/10.1109/IWACI.2010.5585167>
48. Weber M, Wolf P, Zöllner JM (2016) Deeptlr: a single deep convolutional network for detection and classification of traffic lights. In: 2016 IEEE intelligent vehicles symposium (IV), pp 342–348. <https://doi.org/10.1109/IVS.2016.7535408>
49. Athira SV, George M, Jose BR, Mathew J (2017) A global image descriptor based navigation system for indoor environment. *Procedia Comput Sci* 115:466–473. <https://doi.org/10.1016/j.procs.2017.09.086>
50. Pearson J, Robinson S, Jones M (2017) Bookmark: appropriating existing infrastructure to facilitate scalable indoor navigation. *Int J Hum Comput Stud* 103:22–34. <https://doi.org/10.1016/j.jhcs.2017.02.001>
51. Li L, Xu Q, Chandrasekhar V, Lim J, Tan C, Mukawa MA (2017) A wearable virtual usher for vision-based cognitive indoor navigation. *IEEE Trans Cybern* 47(4):841–854. <https://doi.org/10.1109/TCYB.2016.2530407>
52. Dong J, Noreikis M, Xiao Y, Ylä-Jääski A (2018) Vinav: a vision-based indoor navigation system for smartphones. *IEEE Trans Mob Comput* 18(6):1461–1475
53. Rahman Su, Ullah S, Ullah S (2019) A mobile camera based navigation system for visually impaired people. In: Proceedings of the 7th international conference on communications and broadband networking, pp 63–66
54. Kunhoth J, Karkar A, Al-Maadeed S, Al-Attayah A (2019) Comparative analysis of computer-vision and ble technology based indoor navigation systems for people with visual impairments. *Int J Health Geogr* 18(1):29
55. Tyukin A, Priorov A, Lebedev I (2016) Research and development of an indoor navigation system based on the digital processing of video images. *Pattern Recogn Image Anal* 26(1):221–230. <https://doi.org/10.1134/S1054661816010260>
56. Bista SR, Giordano PR, Chaumette F (2016) Appearance-based indoor navigation by IBVS using line segments. *IEEE Robot Autom Lett* 1(1):423–430. <https://doi.org/10.1109/LRA.2016.2521907>
57. Akinlar C, Topal C (2011) Edlines: a real-time line segment detector with a false detection control. *Pattern Recogn Lett* 32(13):1633–1642. <https://doi.org/10.1016/j.patrec.2011.06.001>
58. Zhang L, Koch R (2013) An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. *J Vis Commun Image Represent* 24(7):794–805. <https://doi.org/10.1016/j.jvcir.2013.05.006>
59. Tian Y, Yang X, Arditi A (2010) Computer vision-based door detection for accessibility of unfamiliar environments to blind persons. In: International conference on computers for handicapped persons, Springer, pp 263–270

60. Costa P, Fernandes H, Martins P, Barroso J, Hadjileontiadis LJ (2012) Obstacle detection using stereo imaging to assist the navigation of visually impaired people. *Procedia Comput Sci* 14, 83–93. In: Proceedings of the 4th international conference on software development for enhancing accessibility and fighting info-exclusion (DSAI 2012). <https://doi.org/10.1016/j.procs.2012.10.010>
61. Murillo AC, Gutiérrez-Gómez D, Rituerto A, Puig L, Guerrero JJ (2012) Wearable omnidirectional vision system for personal localization and guidance. In: 2012 IEEE computer society conference on computer vision and pattern recognition workshops, pp 8–14. <https://doi.org/10.1109/CVPRW.2012.6239189>
62. Huang Z, Gu N, Hao J, Shen J (2018) 3DLoC: 3D features for accurate indoor positioning. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 1(4):141–114126. <https://doi.org/10.1145/3161409>
63. Lee DC, Hebert M, Kanade T (2009) Geometric reasoning for single image structure recovery. In: 2009 IEEE conference on computer vision and pattern recognition, pp 2136–2143. <https://doi.org/10.1109/CVPR.2009.5206872>
64. Wang E, Yan W (2014) iNavigation: an image based indoor navigation system. *Multimed Tools Appl* 73(3):1597–1615. <https://doi.org/10.1007/s11042-013-1656-9>
65. Kawaji H, Hatada K, Yamasaki T, Aizawa K (2010) 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. MPVA '10, ACM, New York, NY, USA, pp 1–4. <https://doi.org/10.1145/1878039.1878041>
66. Ke Y, Sukthankar R et al (2004) PCA-sift: a more distinctive representation for local image descriptors. *CVPR* 2(4):506–513
67. Deniz O, Paton J, Salido J, Bueno G, Ramanan J (2014) A vision-based localization algorithm for an indoor navigation app. In: 2014 eighth international conference on next generation mobile apps, services and technologies, pp 7–12. <https://doi.org/10.1109/NGMAST.2014.18>
68. Xiao A, Chen R, Li D, Chen Y, Wu D (2018) An indoor positioning system based on static objects in large indoor scenes by using smartphone cameras. *Sensors*. <https://doi.org/10.3390/s18072229>
69. Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: Towards real-time object detection with region proposal networks. In: Cortes, C, Lawrence ND, Lee DD, Sugiyama M, Garnett R (eds) *Advances in neural information processing systems*, Curran Associates, Inc. 28, pp 91–99. <http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf>. Accessed 8 Nov 2019
70. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv pre-print [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
71. Chen Y, Chen R, Liu M, Xiao A, Wu D, Zhao S (2018) Indoor visual positioning aided by cnn-based image retrieval: training-free, 3D modeling-free. *Sensors*. <https://doi.org/10.3390/s18082692>
72. Rublee E, Rabaud V, Konolige K, Bradski GR (2011) Orb: an efficient alternative to sift or surf. In: ICCV, Citeseer, vol 11, p 2
73. Handa A, Whelan T, McDonald J, Davison AJ (2014) A benchmark for RGB-D visual odometry, 3d reconstruction and slam. In: 2014 IEEE international conference on robotics and automation (ICRA), pp 1524–1531. <https://doi.org/10.1109/ICRA.2014.6907054>
74. Sturm J, Magnenat S, Engelhard N, Pomerleau F, Colas F, Cremers D, Siegwart R, Burgard W (2011) Towards a benchmark for RGB-D SLAM evaluation. In: RGB-D workshop on advanced reasoning with depth cameras at robotics: science and systems conf. (RSS), Los Angeles, United States. <https://hal.archives-ouvertes.fr/hal-01142608>
75. Kendall A, Grimes M, Cipolla R (2015) PoseNet: a convolutional network for real-time 6-DOF camera relocalization. In: The IEEE international conference on computer vision (ICCV), pp 2938–2946
76. Guo F, He Y, Guan L (2017) RGB-D camera pose estimation using deep neural network. In: 2017 IEEE global conference on signal and information processing (GlobalSIP), pp 408–412. <https://doi.org/10.1109/GlobalSIP.2017.8308674>
77. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
78. Adorno J, DeLaHoz Y, Labrador MA (2016) Smartphone-based floor detection in unstructured and structured environments. In: 2016 IEEE international conference on pervasive computing and communication workshops (PerCom workshops), pp 1–6. <https://doi.org/10.1109/PERCOMW.2016.7457136>
79. Bashiri FS, LaRose E, Badger JC, D'Souza RM, Yu Z, Peissig P (2018) Object detection to assist visually impaired people: a deep neural network adventure. In: International symposium on visual computing, Springer, pp 500–510
80. Bashiri FS, LaRose E, Peissig P, Tafti AP (2018) Mcindoor20000: a fully-labeled image dataset to advance indoor objects detection. *Data Brief* 17:71–75
81. Jayakanth K (2019) Comparative analysis of texture features and deep learning method for real-time indoor object recognition. In: 2019 international conference on communication and electronics systems (ICCES), IEEE, pp 1676–1682
82. Afif M, Ayachi R, Said Y, Pissaloux E, Atri M (2020) An evaluation of retinanet on indoor object detection for blind and visually impaired persons assistance navigation. *Neural Process Lett*. <https://doi.org/10.1007/s11063-020-10197-9>
83. Takács M, Bencze T, Szabó-Resch MZ, Vámosy Z (2015) Object recognition to support indoor robot navigation. In: 2015 16th IEEE international symposium on computational intelligence and informatics (CINTI), pp 239–242. <https://doi.org/10.1109/CINTI.2015.7382930>
84. Wang Y, Ma X, Leus G (2011) Robust time-based localization for asynchronous networks. *IEEE Trans Signal Process* 59(9):4397–4410. <https://doi.org/10.1109/TSP.2011.2159215>
85. Zhang D, Xia F, Yang Z, Yao L, Zhao W (2010) Localization technologies for indoor human tracking. In: 2010 5th international conference on future information technology, pp 1–6. <https://doi.org/10.1109/FUTURETECH.2010.5482731>
86. Maccabe AB, Mielke AM, Brennan SM, Torney DC (2004) Radiation detection with distributed sensor networks. *Computer* 34(08):57–59. <https://doi.org/10.1109/MC.2004.103>
87. Werner M (2014) Indoor location-based services: prerequisites and foundations. Springer, Berlin

88. Guo Y, Wang W, Chen X (2017) FreeNavi: Landmark-based mapless indoor navigation based on wifi fingerprints. In: 2017 IEEE 85th vehicular technology conference (VTC Spring), pp 1–5. <https://doi.org/10.1109/VTCspring.2017.8108350>
89. Chen X, Kong J, Guo Y, Chen X (2014) An empirical study of indoor localization algorithms with densely deployed aps. In: 2014 IEEE global communications conference, pp 517–522. <https://doi.org/10.1109/GLOCOM.2014.7036860>
90. Han B, Zhao L (2017) An indoor positioning and navigation technique based on wi-fi fingerprint and environment information. In: China satellite navigation conference, Springer, pp 381–393
91. Wu C, Xu J, Yang Z, Lane ND, Yin Z (2017) Gain without pain: accurate wifi-based localization using fingerprint spatial gradient. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 1(2):29–12919. <https://doi.org/10.1145/3090094>
92. Liu H, Darabi H, Banerjee P, Liu J (2007) Survey of wireless indoor positioning techniques and systems. *IEEE Trans Syst Man Cybern Part C* 37(6):1067–1080
93. Dayekh S, Affes S, Kandil N, Nerguizian C (2010) Cooperative localization in mines using fingerprinting and neural networks. In: 2010 IEEE wireless communication and networking conference, IEEE, pp 1–6
94. Zhang W, Liu K, Zhang W, Zhang Y, Gu J (2016) Deep neural networks for wireless localization in indoor and outdoor environments. *Neurocomputing* 194:279–287
95. Félix G, Siller M, Alvarez EN (2016) A fingerprinting indoor localization algorithm based deep learning. In: 2016 eighth international conference on ubiquitous and future networks (ICUFN), IEEE, pp 1006–1011
96. Jang J-W, Hong S-N (2018) Indoor localization with wifi fingerprinting using convolutional neural network. In: 2018 tenth international conference on ubiquitous and future networks (ICUFN), IEEE, pp 753–758
97. Mittal A, Tiku S, Pasricha S (2018) Adapting convolutional neural networks for indoor localization with smart mobile devices. In: Proceedings of the 2018 on great lakes symposium on VLSI, pp 117–122
98. Ibrahim M, Torki M, ElNainay M (2018) CNN based indoor localization using RSS time-series. In: 2018 IEEE symposium on computers and communications (ISCC), IEEE, pp 01044–01049
99. Li W, Chen Z, Gao X, Liu W, Wang J (2019) Multimodel framework for indoor localization under mobile edge computing environment. *IEEE Internet Things J* 6(3):4844–4853
100. Wei Y, Li W, Chen T (2016) Node localization algorithm for wireless sensor networks using compressive sensing theory. *Pers Ubiquitous Comput* 20(5):809–819
101. Liu C, Yao X, Luo J (2019) Multiregional secure localization using compressive sensing in wireless sensor networks. *ETRI J* 41(6):739–749
102. Zhang P, Wang J, Li W (2020) A learning based joint compressive sensing for wireless sensing networks. *Comput Netw* 168:107030
103. Li Y, Zhuang Y, Lan H, Zhou Q, Niu X, El-Sheimy N (2016) A hybrid wifi/magnetic matching/pdr approach for indoor navigation with smartphone sensors. *IEEE Commun Lett* 20(1):169–172. <https://doi.org/10.1109/LCOMM.2015.2496940>
104. Ren H, Kazanzides P (2012) Investigation of attitude tracking using an integrated inertial and magnetic navigation system for hand-held surgical instruments. *IEEE/ASME Trans Mechatron* 17(2):210–217. <https://doi.org/10.1109/TMECH.2010.2095504>
105. Huang C, Liao Z, Zhao L (2010) Synergism of INS and PDR in self-contained pedestrian tracking with a miniature sensor module. *IEEE Sens J* 10(8):1349–1359. <https://doi.org/10.1109/JSEN.2010.2044238>
106. Wu X, Shen R, Fu L, Tian X, Liu P, Wang X (2017) iBill: using ibeacon and inertial sensors for accurate indoor localization in large open areas. *IEEE Access* 5:14589–14599. <https://doi.org/10.1109/ACCESS.2017.2726088>
107. Betters E (2013) Apple's ibeacons explained: What it is and why it matters. online publication dated Sep 18, 1–14
108. Shu Y, Bo C, Shen G, Zhao C, Li L, Zhao F (2015) Magicol: indoor localization using pervasive magnetic field and opportunistic wifi sensing. *IEEE J Sel Areas Commun* 33(7):1443–1457. <https://doi.org/10.1109/JSAC.2015.2430274>
109. Lee K, Nam Y, Min SD (2018) An indoor localization solution using bluetooth rssi and multiple sensors on a smartphone. *Multimed Tools Appl* 77(10):12635–12654. <https://doi.org/10.1007/s11042-017-4908-2>
110. Satan A (2018) Bluetooth-based indoor navigation mobile system. In: 2018 19th international carpathian control conference (ICCC), pp 332–337. <https://doi.org/10.1109/CarpathianCC.2018.8399651>
111. Satan A, Toth Z (2018) Development of bluetooth based indoor positioning application. In: 2018 IEEE international conference on future IoT technologies (Future IoT), pp 1–6. <https://doi.org/10.1109/FIOT.2018.8325586>
112. Davis J (2015) Indoor wireless RF channels. <http://wireless.per.nyu.edu/reference/chapter03/indoor.html>. Accessed 10 May 2019
113. Yu N, Zhan X, Zhao S, Wu Y, Feng R (2018) A precise dead reckoning algorithm based on bluetooth and multiple sensors. *IEEE Internet Things J* 5(1):336–351. <https://doi.org/10.1109/JIOT.2017.2784386>
114. Campana F, Pinargote A, Domínguez F, Peláez E (2017) Towards an indoor navigation system using bluetooth low energy beacons. In: 2017 IEEE second ecuador technical chapters meeting (ETCM), pp 1–6. <https://doi.org/10.1109/ETCM.2017.8247464>
115. Al-Madani B, Orujov F, R Maskeliūnas, Damaševičius R, Venčkauskas A (2019) Fuzzy logic type-2 based wireless indoor localization system for navigation of visually impaired people in buildings. *Sensors* 19(9):2114
116. Murata M, Ahmetovic D, Sato D, Takagi H, Kitani KM, Asakawa C (2019) Smartphone-based localization for blind navigation in building-scale indoor environments. *Pervasive Mob Comput* 57:14–32
117. Ahmetovic D, Gleason C, Ruan C, Kitani K, Takagi H, Asakawa C (2016) Navcog: a navigational cognitive assistant for the blind. In: Proceedings of the 18th international conference on human-computer interaction with mobile devices and services, ACM, pp 90–99
118. Kim J-E, Bessho M, Kobayashi S, Koshizuka N, Sakamura K (2016) Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In: Proceedings of the 31st annual ACM symposium on applied computing, ACM, pp 604–611
119. Cheraghi SA, Nambodiri V, Walker L (2017) Guidebeacon: beacon-based indoor wayfinding for the blind, visually impaired, and disoriented. In: 2017 IEEE international conference on pervasive computing and communications (PerCom), IEEE, pp 121–130

120. Bilgi S, Ozturk O, Gulnerman AG (2017) Navigation system for blind, hearing and visually impaired people in ITU ayazaga campus. In: 2017 international conference on computing networking and informatics (ICCNII), pp 1–5
121. Abu Doush I, Alshatnawi S, Al-Tamimi A-K, Alhasan B, Hamasha S (2016) ISAB: integrated indoor navigation system for the blind. *Interact Comput* 29(2):181–202. <https://doi.org/10.1093/iwc/iww016>
122. Ganz A, Schafer J, Gandhi S, Puleo E, Wilson C, Robertson M (2012) Percept indoor navigation system for the blind and visually impaired: architecture and experimentation. *Int J Telemed Appl* 2012:19–191919. <https://doi.org/10.1155/2012/894869>
123. Ganz A, Schafer JM, Tao Y, Wilson C, Robertson M (2014) Percept-II: Smartphone based indoor navigation system for the blind. In: 2014 36th annual international conference of the IEEE engineering in medicine and biology society, pp 3662–3665. <https://doi.org/10.1109/EMBC.2014.6944417>
124. Tsirmpas C, Rompas A, Fokou O, Koutsouris D (2015) An indoor navigation system for visually impaired and elderly people based on radio frequency identification (RFID). *Inf Sci* 320:288–305. <https://doi.org/10.1016/j.ins.2014.08.011>
125. Lin Qiongzhen, Guo Y (2016) Accurate indoor navigation system using human-item spatial relation. *Tsinghua Sci Technol* 21(5):521–537. <https://doi.org/10.1109/TST.2016.7590321>
126. Loconsole C, Dehkordi MB, Sotgiu E, Fontana M, Bergamasco M, Frisoli A (2016) An IMU and RFID-based navigation system providing vibrotactile feedback for visually impaired people. In: International conference on human haptic sensing and touch enabled computer applications, Springer, pp 360–370
127. Xu H, Ding Y, Li P, Wang R, Li Y (2017) An RFID indoor positioning algorithm based on bayesian probability and k-nearest neighbor. *Sensors*. <https://doi.org/10.3390/s17081806>
128. Ganti D, Zhang W, Kavehrad M (2014) VLC-based indoor positioning system with tracking capability using Kalman and particle filters. In: 2014 IEEE international conference on consumer electronics (ICCE), pp 476–477. <https://doi.org/10.1109/ICCE.2014.6776093>
129. Jayakody A, Meegama CI, Pinnawalage HU, Muwenwella RMHN, Dalpathado SC (2016) AVII [assist vision impaired individual]: an intelligent indoor navigation system for the vision impaired individuals with vlc. In: 2016 IEEE international conference on information and automation for sustainability (ICIAFS), pp 1–6. <https://doi.org/10.1109/ICIAF.2016.7946526>
130. Nakajima M (2013) New indoor navigation system for visually impaired people using visible light communication. *EURASIP J Wirel Commun Netw* 1:37. <https://doi.org/10.1186/1687-1499-2013-37>
131. Fan Q, Sun B, Sun Y, Zhuang X (2017) Performance enhancement of MEMS-based INS/UWB integration for indoor navigation applications. *IEEE Sens J* 17(10):3116–3130. <https://doi.org/10.1109/JSEN.2017.2689802>
132. Hsu H-H, Chang J-K, Peng W-J, Shih TK, Pai T-W, Man KL (2018) Indoor localization and navigation using smartphone sensory data. *Ann Oper Res* 265(2):187–204. <https://doi.org/10.1007/s10479-017-2398-2>
133. Hasan MA, Mishuk MN (2018) Mems IMU based pedestrian indoor navigation for smart glass. *Wirel Pers Commun* 101(1):287–303. <https://doi.org/10.1007/s11277-018-5688-3>
134. Ju H, Park SY, Park CG (2018) A smartphone-based pedestrian dead reckoning system with multiple virtual tracking for indoor navigation. *IEEE Sens J* 18(16):6756–6764. <https://doi.org/10.1109/JSEN.2018.2847356>
135. Shin SH, Park CG, Choi S (2010) New map-matching algorithm using virtual track for pedestrian dead reckoning. *ETRI J* 32(6):891–900
136. Hsu Y, Wang J, Chang C (2017) A wearable inertial pedestrian navigation system with quaternion-based extended kalman filter for pedestrian localization. *IEEE Sens J* 17(10):3193–3206. <https://doi.org/10.1109/JSEN.2017.2679138>
137. Giorgi G, Frigo G, Narduzzi C (2017) Dead reckoning in structured environments for human indoor navigation. *IEEE Sens J* 17(23):7794–7802. <https://doi.org/10.1109/JSEN.2017.2725446>
138. Huang H-Y, Hsieh C-Y, Liu K-C, Cheng H-C, Hsu SJ, Chan C-T (2019) Multi-sensor fusion approach for improving map-based indoor pedestrian localization. *Sensors* 19(17):3786
139. Luo J, Zhang C, Wang C (2020) Indoor multi-floor 3D target tracking based on the multi-sensor fusion. *IEEE Access* 8:36836–36846
140. Poulouse A, Eyobu OS, Han DS (2019) A combined PDR and wi-fi trilateration algorithm for indoor localization. In: 2019 international conference on artificial intelligence in information and communication (ICAIIIC), IEEE, pp 072–077
141. Qiu S, Wang Z, Zhao H, Qin K, Li Z, Hu H (2018) Inertial/magnetic sensors based pedestrian dead reckoning by means of multi-sensor fusion. *Inf Fus* 39:108–119
142. Kuang J, Niu X, Zhang P, Chen X (2018) Indoor positioning based on pedestrian dead reckoning and magnetic field matching for smartphones. *Sensors* 18(12):4142
143. Ciabattini L, Foresi G, Monteriù A, Pepa L, Pagnotta DP, Spalazzi L, Verdini F (2019) Real time indoor localization integrating a model based pedestrian dead reckoning on smartphone and BLE beacons. *J Ambient Intell Humaniz Comput* 10(1):1–12
144. Yu S-J, Jan S-S, De Lorenzo DS (2018) Indoor navigation using wi-fi fingerprinting combined with pedestrian dead reckoning. In: 2018 IEEE/ION position, location and navigation symposium (PLANS), IEEE, pp 246–253
145. Liu Z, Dai W, Win MZ (2018) Mercury: an infrastructure-free system for network localization and navigation. *IEEE Trans Mob Comput* 17(5):1119–1133. <https://doi.org/10.1109/TMC.2017.2725265>
146. Madgwick SOH, Harrison AJL, Vaidyanathan R (2011) Estimation of IMU and marg orientation using a gradient descent algorithm. In: 2011 IEEE international conference on rehabilitation robotics, pp 1–7. <https://doi.org/10.1109/ICORR.2011.5975346>
147. Peng C, Shen G, Zhang Y, Li Y, Tan K (2007) Beepbeep: a high accuracy acoustic ranging system using cots mobile devices. In: Proceedings of the 5th international conference on embedded networked sensor systems. *SenSys '07*, ACM, New York, NY, USA, pp 1–14. <https://doi.org/10.1145/1322263.1322265>
148. Xiao Z, Wen H, Markham A, Trigoni N (2015) Indoor tracking using undirected graphical models. *IEEE Trans Mob Comput* 14(11):2286–2301. <https://doi.org/10.1109/TMC.2015.2398431>
149. Hilsenbeck S, Bobkov D, Schroth G, Huitl R, Steinbach E (2014) Graph-based data fusion of pedometer and wifi measurements for mobile indoor positioning. In: Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing. *UbiComp '14*, ACM, New York, NY, USA, pp 147–158. <https://doi.org/10.1145/2632048.2636079>

150. Mendoza-Silva GM, Torres-Sospedra J, Huerta J (2017) A more realistic error distance calculation for indoor positioning systems accuracy evaluation. In: 2017 international conference on indoor positioning and indoor navigation (IPIN), pp 1–8. <https://doi.org/10.1109/IPIN.2017.8115950>
151. Cai C, Zheng R, Li J, Zhu L, Pu H, Hu M (2019) Asynchronous acoustic localization and tracking for mobile targets. *IEEE Internet Things J*
152. Wu H, Mo Z, Tan J, He S, Chan S-HG (2019) Efficient indoor localization based on geomagnetism. *ACM Trans Sens Netw* 15(4):1–25
153. Liu M, Cheng L, Qian K, Wang J, Wang J, Liu Y (2020) Indoor acoustic localization: a survey. *Hum-Centric Comput Inf Sci* 10(1):2

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
