Indoor Pedestrian Dead Reckoning Calibration by Visual Tracking and Map Information

Jingjing YAN
International Doctoral Innovation Centre
University of Nottingham
Ningbo 315100, China
Jingjing.YAN@nottingham.edu.cn

Gengen HE
Department of Geographical Science
University of Nottingham
Ningbo 315100, China
Gengen.He@nottingham.edu.cn

Anahid BASIRI
Centre for Advanced Analysis
University College London
London, UK
a.basiri@ucl.ac.uk

Craig HANCOCK
Department of Civil Engineering
University of Nottingham
Ningbo 315100, China
Craig.Hancock@nottingham.edu.cn

***Abstract*—Currently, Pedestrian Dead Reckoning (PDR) systems are becoming more attractive in market of indoor positioning. This is mainly due to the development of cheap and light Micro Electro-Mechanical Systems (MEMS) on smartphones and less requirement of additional infrastructures in indoor areas. However, it still faces the problem of drift accumulation and needs the support from external positioning systems. Vision-aided inertial navigation, as one possible solution to that problem, has become very popular in indoor localization with satisfied performance than individual PDR system. In the literature however, previous studies use fixed platform and the visual tracking uses feature-extraction-based methods. This paper instead contributes a distributed implementation of positioning system and uses deep learning for visual tracking. Meanwhile, as both inertial navigation and optical system can only provide relative positioning information, this paper contributes a method to integrate digital map with real geographical coordinates to supply absolute location. This hybrid system has been tested on two common operation systems of smartphones as iOS and Android, based on corresponded data collection apps respectively, in order to test the robustness of method. It also uses two different ways for calibration, by time synchronization of positions and heading calibration based on time steps. According to the results, localization information collected from both operation systems has been significantly improved after integrating with visual tracking data.**

Keywords—pedestrian dead reckoning; visual tracking; smartphone positioning; sensor fusion

I. INTRODUCTION

Indoor positioning techniques and related navigation technologies are becoming more important as there is a high demand with urbanization [1]. One reason is due to the widely used outdoor positioning system, Global Positioning System (GPS), is problematic for indoor applications, caused by unavailability of Global Navigation Satellite System (GNSS) signals [1-4]. The other reason is Location Based Services (LBS) has been widely used by people around the world [5, 6] and are substantially affected by the popularization of smart devices [7]. Common indoor positioning techniques are can be divided into three categories as signal-based methods, dead reckoning, and device-free methods [8]. Although there is still no optimal solution which satisfies the requirements of accuracy, availability, continuity, and reliability when comparing to GPS for outdoor positioning [9, 10], the latter two classes have the advantages of higher flexibility in operation and lower cost in infrastructure installation [11]. One the other head, the signal-based approaches need the pre-installed beacons, which are more expensive, less flexible and easily affected by physical variations from environment and multipath effects [12-14]. Common sensors for dead reckoning and device-free approaches are MEMS-based Inertial Measurement Units (IMU) and Charged-Couple-Device (CCD) cameras, respectively. These two kinds of sensors have experienced great advancements in manufacturing with products in more compact, cheap, low energy consumption and precise format [8, 15-19]. In addition, with the ubiquity of IMU sensors in smartphones and surveillance cameras in public building areas, the possibility of applying these sensors in daily life is growing as well [11]. With a wide range of applications in indoor scenarios, it infers a promising future market for the technologies based on these two kind of sensors.

PDR systems or Inertial Navigation Systems (INSs), are defined as dead reckoning based systems for pedestrians [16]. They can provide relative user positions, orientation and velocity in indoor area by using triad accelerometers and gyroscopes for step detection and heading estimation [20-25]. PDR-based methods can be divided into several categories depended on the implementation of IMU: foot-mounted [26, 27], hand-held [17, 28], backpack [29-31], in-pocket [32] or head-mounted [27, 33, 34], and this study mainly focuses on hand-held ones on smartphones. This is because that smartphones have been integrated into ordinary and spaces of daily life [6], and Location-Based-Services (LBS) such as navigation and tracking, has been widely used by people around the world [5, 6]. Tiny sensors which are necessary for INSs, such as accelerometer, gyroscope, magnetometer, and gyroscope, have already been implemented in these smart devices. Moreover, they are affordable to the public as well as for infrastructure-less navigation [12]. Commonly used smartphones for PDR can be divided into two types as iPhone [13, 14, 35] and Android systems [12, 17, 36-40], and the latter one seems to be more popular in current researches. However, as the accuracy of IMU sensors can be compromised by bias drift with the accumulation of time period, especially for the inertial sensors on smartphones as they are less precise. This can cause problems on the long-term use of PDR for independent positioning, and thus external positioning information are needed for calibration and absolute localization [3, 16, 22, 41, 42]. There are two ways to solve this issue: configuration of system dynamics [43, 44] and compensation from other sensing system [3, 41].

Meanwhile, the high accuracy of Optical Positioning Systems (OPS) has encouraged many pedestrian based applications, including indoor positioning [12, 45]. In addition, the introduction of optical systems can also enrich information during positioning process by object detection from visual data such as video, with highly accurate localization results [19, 45]. Conventional methods are depended on either optical flow or feature detection [41]. The former, although with higher accuracy, can be more computationally expensive and may require precise conditions of lighting and precise cameras. In addition, it assumes that between-frame motions are small and limited enough to be ignored [24], which might not be true in real-world applications and scenarios. Feature-based methods are more popular and simple approaches. They extract features from landmarks in images for positioning, and thanks to the ubiquity of indoor landmarks, they can provide solution in many indoor applications with relatively less computation power required [41]. However, the performance of OTS can be easily affected by occlusion in surrounding environment as the Line-of-Sight (LOS) between camera and targets are compulsory for OTS [27].

In order to compensate the drawbacks from the above mentioned two positioning systems, the integration of them can be a good answer. The OTS can be used to calibrate the drift accumulation with its higher accuracy, while the inertial systems can solve the incontinuity problem of OTS caused by occlusion in LOS as it can provide relatively accurate results in a short time. The fusion of these two systems is expected to keep the merits both two positioning systems, providing localization service with higher overall accuracy, continuity, accessibility and reliability [3, 24]. This kind of vision-aided inertial positioning systems has been widely used in many applications and researches, as it can provide three dimensional location information and orientation estimation for motion tracking. The typical applications are in robotics – such as Simultaneous Localization and Mapping (SLAM) – and for unmanned vehicle system [23, 41]. The usual implementation of this system is to attach a monocular/stereo camera and IMU sensor on a fixed platform and the fusion of inertial and visual data is based on egomotion heading estimation, by slowing the sample rate of IMU data [44], or using Particle Filter (PF) [e.g. 11, 43] or applying variants of Kalman Filter, such as Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF) [e.g. 11, 18, 24, 27, 41]. Some of studies have tried to utilize the built-in cameras and IMU sensors in smartphones for indoor localization. They use built-in camera to film the ground for camera's relative position and orientation estimation based on ground-plane feature matching, and IMU sensors for step detection and heading estimation [29, 30]. However, this kind of approach is not practical for commercial application as the video recording by camera is quite energy consuming and cannot support for a long time for indoor localization. Previous studies have also proved that the combination of floor plan as environmental constraints, supported by the application of PF, can help to improve the accuracy of indoor positioning [22, 29, 30, 32, 42, 46-51]. This paper also takes this idea to constrain the positioning results but without using of PF. Instead, it is processed by georeferencing in order to provide absolute position information to the results.

This paper proposes a hybrid system with a different set-up, which attaches inertial sensors and camera on independent platforms other than the same platform. The video data is taken from static and fixed camera and inertial data is taken from devices carried by users. Instead of installing additional sensors, it tries to utilize the current indoor infrastructures of surveillance systems in indoor area and inertial sensors on smartphones to provide user location, with the support of digital floor plan with georeference. In comparison with previous studies using landmark-based image matching for localization and camera orientation estimation, this paper uses deep-learning-based object detection for pedestrian localization, with the prior information of camera location inside the building and digital floor plan. Then, the estimated 2D trajectories calculated from inertial and visual data are both pre-processed by coordinate transformation based on real geographical information provided by digitized floor plan. The visual data is then used to calibrate inertial positioning in visible area based on similar time steps. There are two ways of calibration being introduced in this paper as position replacement and heading correction. The final results acquired from these two methods are compared for a better solution based on the accuracy of tracking. In addition, this paper, other than previous studies only focus on one type of smartphone, the system is tested on both types of common smartphone systems for the robustness of the approach application and it also gives an answer about which kind of smartphone may have more precise IMU sensors based on the experiments in this study.

II. System Architecture

This paper uses a hybrid system which is consisted of one main positioning system as smartphone-based PDR and an aided system as camera-based visual tracking (Fig.1). During the operation, the smartphone-based PDR works continuously while the visual tracking system only provides positioning service in visible area. For data collection, an Android phone and an iPhone are simultaneous held by the user to collect accelerations and angular velocity during walking, which can later be used for relative positions and poses estimation (Section II A). The video recording is triggered at the same time. Once user enters the LOS area of camera, the positions can also be calculated by pedestrian detection with the support of depth information of each frame (Section II B) and the heading of user is determined by user positions in two consecutive frames. In order to acquire absolute locations, the corresponded digital floor plan with georeference in WGS 1984 is used as the reference map for the results of PDR and visual tracking before calibration (Section II C). This is beneficial to further development of seamless indoor-outdoor positioning transition by sharing same positioning coordinate. The visual positioning results are then used to calibrate the inertial positioning, based on similar time steps of PDR. The calibration methods can be divided into two ways which are described in Section II D and their inertial positioning results are compared on accuracy in order to provide an optimal solution in Section IV. This paper also compares the calibrated PDR results between two different types of smartphones to see whether this method is a potentially ubiquitous solution for both two types of mobile phones.

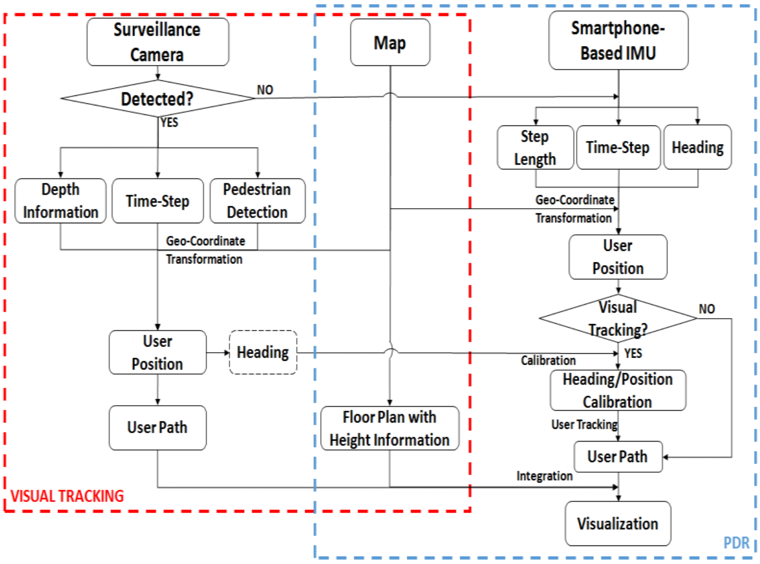


Fig. 1. The architecture of vision-aided PDR system (The left part in red box is visual tracking and the right part in blue box is PDR system).

A. *Smartphone-Based PDR System*

The inertial positioning algorithm in this paper is developed based on previous studies using Step-and-Heading System (SHS), which uses 2D description of pedestrian strides as length and heading. Meanwhile, it also transform its coordinate from body frame to global frame [16, 36, 37, 52, 53].

1) Step Detection

The algorithm of step detection in this paper is depended on gait cycle detection, which identifies gait cycles by searching for repetitive data pattern. In this paper, it applies a threshold on pre-processed accelerations, based on the principle of stationary inertial sensor during stance and the corresponded threshold is then settled, in order to be recognized that activity and reported to the system [16, 54, 55]. The measured acceleration is first filtered by a low-pass filter [37] with frequency condition depended on accelerometer sampling rate [17]. Then, a synthetic acceleration respect to time is taken in three axes ($a_x(t), a_y(t)$, and $a_y(t)$) as described in formula (1). This is due to distribution of vertical signals, which mainly contribute to step peaks, may appear in all axes based on the current device's altitude and orientation [35].

$$a(t) = \sqrt{\left(a_x(t)\right)^2 + \left(a_y(t)\right)^2 + \left(a_z(t)\right)^2 - g} \qquad (1)$$

where g is the earth's gravity, which needs to be excluded from synthetic results. The synthetic accelerations are then processed by applying pre-settled threshold to classify heel-off, toe-off, heel-strike and stance phases of gait cycle. The next step is to apply zero-crossing approach for cyclic property detection [56]. The detected steps after acceleration processing by Android and iPhone are represented in Fig.2.

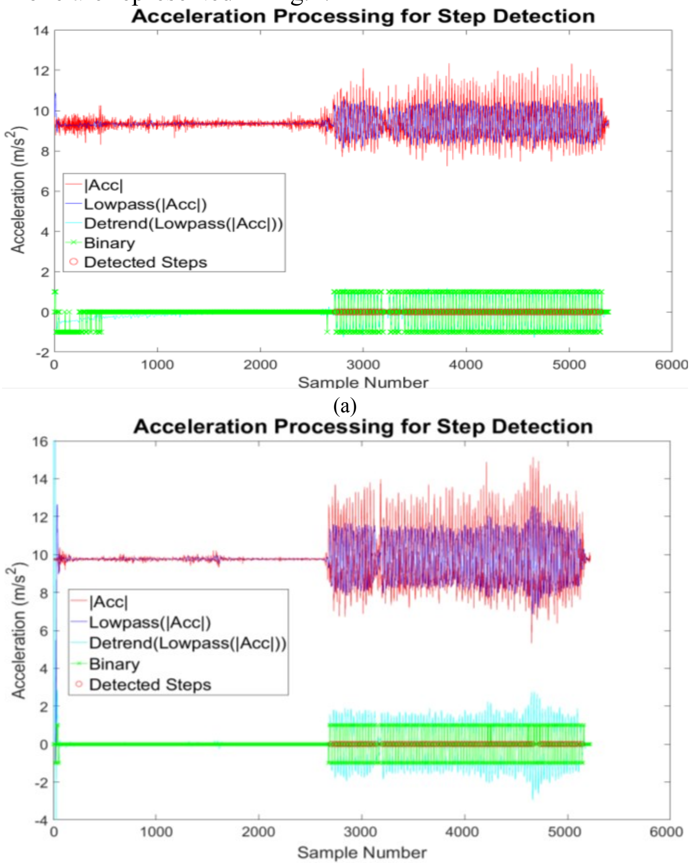


Fig. 2. The processed synthetic accelerations and detected steps by Android (a) and iPhone (b).

2) Step Length Estimation

This paper estimates step length based on Weinberg's algorithm as in (2), which uses a non-linear model with the value of maximum ($a_{max}(i)$) and minimum ($a_{min}(i)$) of synthetic accelerations of each step event [57].

$$SL_i = \sqrt[4]{a_{max}(i) - a_{min}(i)} * k \ (i = 1,2, \dots, n) \qquad (2)$$

where SL_i is the step length of the i^{th} step and k is an empirical value of penalty for estimation [37].

3) Heading Estimation

The orientation of each step is determined by its corresponded angular velocity changes of body frame at time t and the heading of previous state. Before this step, it first needs to modify the angular velocity changes from body frame into global frame by applying a rotation matrix $R(t)$. The transformation process from body frame to global frame is described in (3)-(6):

$$R_x(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi(t)) & -\sin(\phi(t)) \\ 0 & \sin(\phi(t)) & \cos(\phi(t)) \end{pmatrix} \qquad (3)$$

$$R_y(t) = \begin{pmatrix} \cos(\theta(t)) & 0 & \sin(\theta(t)) \\ 0 & 1 & 0 \\ -\sin(\theta(t)) & 0 & \cos(\theta(t)) \end{pmatrix} \quad (4)$$

$$R_z(t) = \begin{pmatrix} \cos(\psi(t)) & -\sin(\psi(t)) & 0 \\ \sin(\psi(t)) & \cos(\psi(t)) & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (5)$$

$$R(t) = R_z(t) R_y(t) R_x(t) \quad (6)$$

where $R_x(t), R_y(t)$, and $R_z(t)$ represents the sub rotation matrix in roll $\phi(t)$, pitch $\theta(t)$ and yaw $\psi(t)$ directions of body frame respectively, as a function of time t. The overall rotation matrix $R(t)$ is determined by the integration of these three components. The initial states of roll and pitch angle are determined by average changes of initial accelerations in same directions and the initial yaw will be zero as the starting point of heading. The related updating of $R(t)$ to $R(t + \Delta t)$ is described in (7)-(8):

$$X = \begin{pmatrix} 0 & -\omega_z^t \Delta t & \omega_y^t \\ \omega_z^t \Delta t & 0 & -\omega_x^t \Delta t \\ -\omega_y^t \Delta t & \omega_x^t \Delta t & 0 \end{pmatrix} \quad (7)$$

$$R(t + \Delta t) = R(t) * \exp(X) \quad (8)$$

where X is the updating variable, based on angular velocity changes in three axis as ω_x^t, ω_y^t and ω_z^t with sampling interval Δt. In this paper, as the smartphone is held stably in hand pointing to the walking direction, so the heading $\Psi(t)$ of each step is only related to the changes in yaw direction [17, 37] and can be calculated as:

$$\Psi(t) = \arctan_2(R_{2,1}(t), R_{1,1}(t)) \quad (9)$$

4) Position Estimation and Positioning Error

The user position P_i is then calculated by adding corresponded estimated step length SL_i to the previous location with estimated heading $\Psi(t)$:

$$P_i = \begin{bmatrix} P_{E_i} \\ P_{N_i} \end{bmatrix} = \begin{bmatrix} P_{E_{i-1}} + SL_i * \sin(\Psi(t)) \\ P_{N_{i-1}} + SL_i * \cos(\Psi(t)) \end{bmatrix} \quad (10)$$

where P_{E_i} and P_{N_i} represent the position in east and north direction [17, 37]. Before the calculation of position error, the estimated positions need to be transformed into real geographic system as the reference positions are measured in that way. The positioning error D_i is then defined as distance between estimated position P_i and reference position R_i:

$$D_i = \|P_i - R_i\| \quad (11)$$

B. Pedestrian Detection Based Visual Tracking

1) Deep-Learning Based Pedestrian Detection

The conventional feature-extraction based pedestrian detection, the common methods are based on figure-ground segmentation of video data [58]. Many previous studies have utilized background subtraction for foreground detection to identify people in the images (e.g. [59, 60]). After detecting human in each image, the next step is to transform the human position in image space to other coordinate systems. For example, the research by National Central University of Taiwan applied a 2D Direct Linear Transform (DLT) to processed results to get the relationship between two coordinate system for later conversion from image coordinate to bird's-eye view [60]. The study by Wuhan University integrated the results with depth image extracted from rendered background to get 3D coordinates and adjusted them into pre-established 3D networks [59].

However, feature based methods require many manual selections of best features and they are limited to the extent of the applications and environments as their parameters need to be modified regularly, which is affected by the ambient. To overcome these problems, deep learning methods can reduce the manual work to improve the flexibility and ubiquity of solution. This paper uses Faster R-CNN for pedestrian detection (Fig.3), which is based on Regional Proposal Network (RPN) and Region-Based Convolutional Neural Networks (R-CNNs), and is one of the state-of-art methods for deep learning with higher accuracy. The RPN is used for predicting Bounding Box (BB) and classifying objectness, and a Fast R-CNN method is applied for object detection using the predicted BBs with a detector based on VGG-16 model [61].

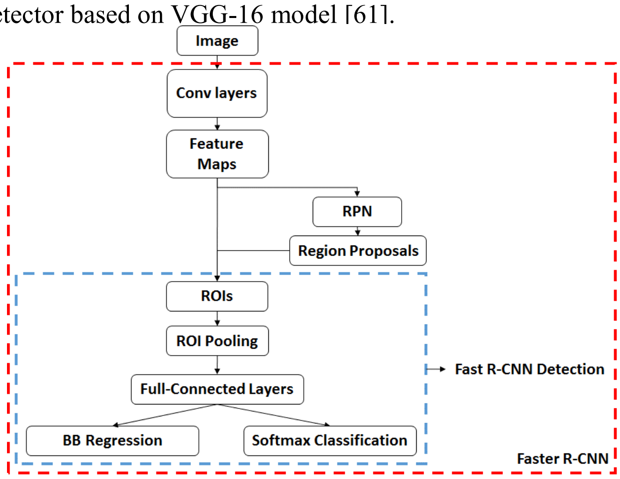


Fig. 3. The architecture of Faster R-CNN.

Fig. 4. An example of extracted BB in frame (left) and entire user path (right).

In this study, it uses pre-trained human detector of deep network based on data from MS COCO and PASCAL VOC 2007 + 2012. After being processed by Faster R-CNN, the BBs are extracted from frames and corresponded frame numbers are also recorded for later acquisition of related time stamps. The middle points of bottom boundaries of BBs are then regarded as the lowest points of user or user's mobility aid, and these points can be constructed into the entire user path (Fig.4). Although this study only uses one user, Faster R-CNN has the potential to handle multiple users for pedestrian detection. However, the overlapping of people in camera will be a big challenge at that time. Meanwhile, the relatively long filming distance in the beginning between targeted user and corresponded camera will also cause the problem of missing detection of user.

2) Depth Information

The distance D_i between the user and the camera for i^{th} frame which is regarded as depth information, could be determined by a pinhole camera model [62] as in (12) with pixel height h_i, focal pixel length f and real height H_p of human. f is determined by camera resolution and field of view (FOV).

$$\frac{h_i}{f} = \frac{H_p}{D_i} \; (i = 1,2, \dots, n) \tag{12}$$

C. Map Information Integration

Before calibration, both results from smartphone-based PDR and camera-based visual tracking need to be transformed into same coordinate system provided by map information. In this paper, the georeferenced digital map is used as the reference for absolute positions with simplified semantic representations of indoor building information. It is digitized from floor plan with georeferencing processing, and geographic system used in this paper is WGS 1984 with prior building height information (9.5 m) for 3D positioning. The application of WGS 1984 makes the future development of transition between indoor and outdoor environment possible, as it is also the common Spatial Reference System (SRS) for GPS.

D. Calibration of Smartphone-Based PDR

According previous studies [12, 19, 45], visual positioning is relatively accurate than PDR in LOS area. Thus, its results can be used to calibrate the drift during using PDR-based positioning. This paper introduces two approaches for PDR calibration supported by visual tracking with map information. In Section IV, their results are compared based on the positioning error and a relative optimal solution is then selected.

1) Position Replacement

The first idea is to directly replace PDR positioning results by visual tracking results based on time synchronization. As both PDR and visual tracking results have recorded time stamps, their results with similar time stamps can then be matched together by replacing results from PDR positioning with vision-based tracking. The time stamps of PDR are deduced from detected step events and related time stamps from accelerometer readings, while that of videos are inferred from frame number and filming frequency. This method has the advantage of simple implementation and less computation cost.

2) Heading Calibration

In reality, however, the time stamps of two positioning systems cannot be perfectly matched, and a more realistic situation is that the time stamp of current detected step from PDR is between two successive detected positions from frames with similar time stamps. This leads to the development of second method for heading calibration. This method is more close to real time simulation as it replaces PDR's heading of each step by the direction determined by two consecutive frames based on similar time steps. Then the calibrated headings are used with the previous estimated step lengths to re-calculate user positions. This method needs slightly difficult implementation and a little more computation power when dealing with large amount of data.

III. Experimental Setup

In this study, the proposed system tested in an experimental site located on the 4th floor of Sir Peter Mansfield Building at University of Nottingham, Ningbo, China. The whole walking trajectory is 51.84m, where only a short part (8.82m) which is not visible by camera. This short part is tracked by only using smartphone-based PDR. The rest, which is a long corridor, could be used for testing the calibrated PDR (43.02m) by visual tracking results. The reference map is digitized from floor plan of experimental site by using ArcGIS 10.3, with simple semantic representations of indoor structures (Fig.5). The reference path is represented in a red line with red crosses, starting from the door of Room 443 and ends at the end of corridor beside Room 413. It will be later used for accuracy tests by comparing with the estimated position points.

There are two types of smartphones are used for PDR measurements. The smartphone model selected for Android system is HUAWEI MT7-TL00, and that selected for iOS system is iPhone 7 Plus. The data collection app for the former is GetSensorData [37] and for the latter is MATLAB Mobile. The sampling frequency for two smartphones are all settled to be 100 Hz. During the experiment, both smartphones are held horizontally, pointing to heading direction. The smartphone-based positioning method is already described in Section II A.

For visual tracking system, the camera used for experiments is located on the ceiling in front of Room 416 and it starts filming simultaneously with the initialization of smartphone-based PDR. The resolution of camera is 960×544, vertical FOV is 27°, and thus the pixel length for the camera is about 1.05×10^3 per inch. The frame frequency is 16 frames per second.

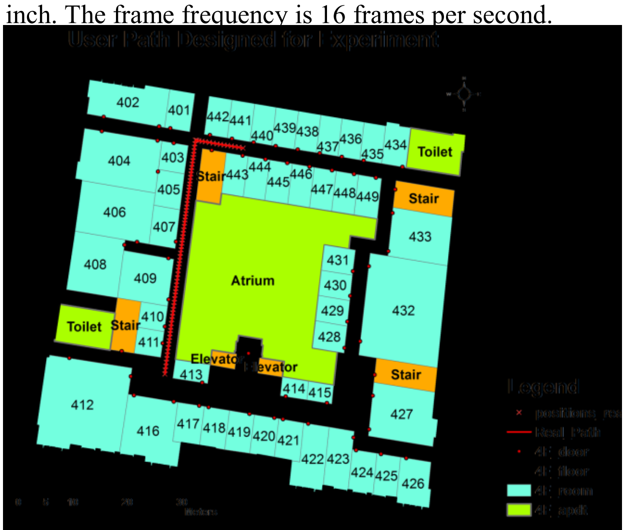


Fig. 5. The reference path (with 4F_floor, 4F_door, 4F_room, 4F_adpt and 4F_con represent floor shape, doors, rooms, non-functional places and connections such as staircases and elevators).

IV. RESULTS AND ANAYSIS

The results can be divided into two parts as pre-calibration and post-calibration. In pre-calibration, it includes the results from visual tracking and smartphone-based PDR of two types of mobile phones. In post-calibration, it shows the results of two different calibration methods. In addition, it also compares the results between two types of phones to see whether this method can improve their accuracy to a similar level.

A. Pre-Calibration

1) Visual Tracking

After extracting the visual positioning points, the overall path in visible area is shown in Fig.6. In all, the overall visual path matches the reference path in visible area, however, it can be found that the positioning points are not evenly distributed. In the beginning, the positioning points are quite dense while in the ending stage, the positioning points start to be sparser. There are two reasons for this phenomenon. First, as mentioned previously in Section II B, the target is too far away to be detected by the camera, leading to mistakes in pedestrian detection. Second, as the depth information is calculated based on a pinhole model which mainly relies on the pixel height h_i changes in frames, it also affects the results when calculating the distance. In the initial stage, the changes of h_i are trivial, this leads to the dense distribution of positioning points, while in the ending part, the changes of h_i are becoming more significant and thus the distribution of positioning are more scattered.

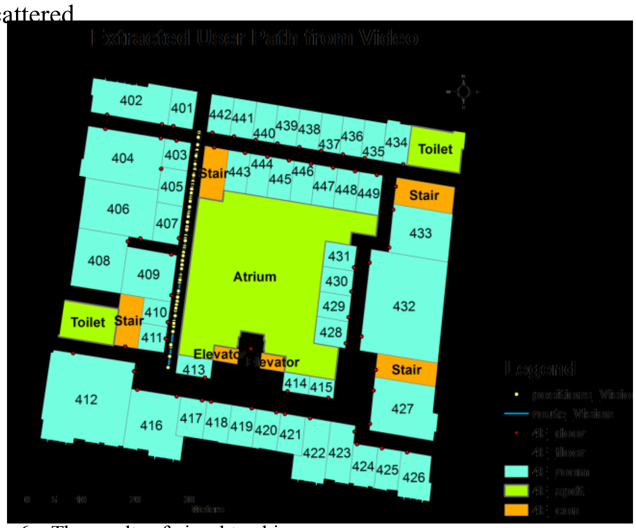
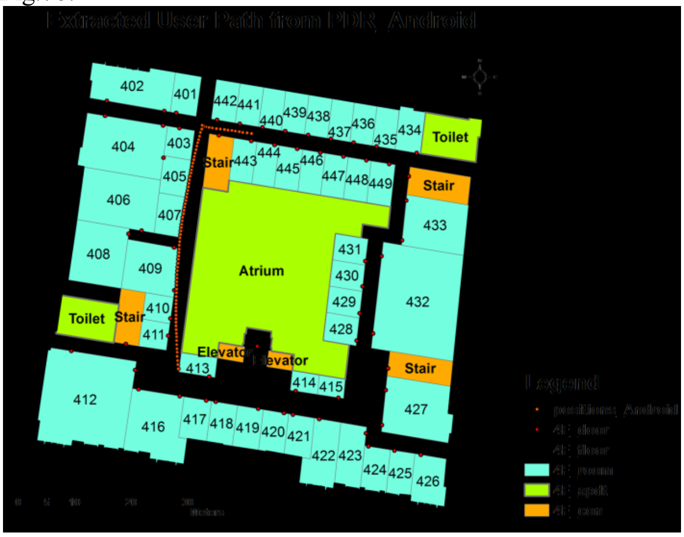


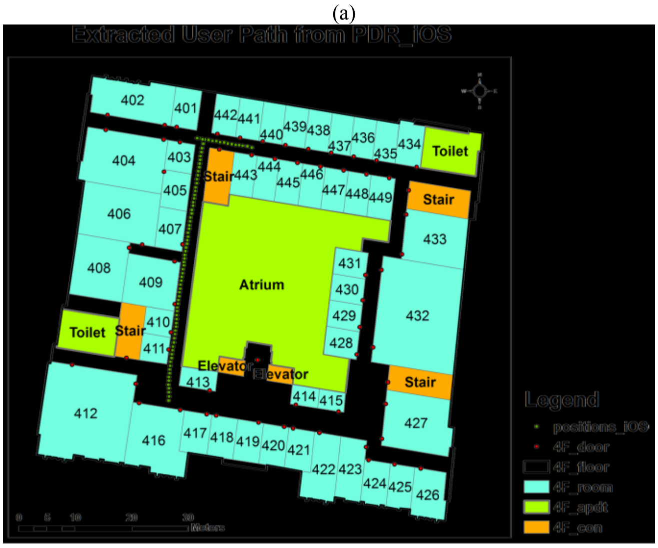
Fig. 6. The results of visual tracking.

2) Smartphone-Based PDR

The results of smartphone-based positioning are shown in Fig.7 and the accuracy is summarized in Table I. During the experiment, the user walked 83 steps. The Android phone detected 83 steps while the iPhone detected 84 steps. This may due to iPhone has more sensitive accelerometer and detects one more step in the end, which can also be viewed from Fig.2 as its acceleration changes are more significant than Android phone. This leads to the drift accumulation in step length estimation by iPhone. This unexpected However, iPhone seems to have a better gyroscope, which can be viewed after passing the corner as path estimated by iPhone has less tilt compared to Android system. In all, Android-based PDR seems to have better performance than iPhone-based PDR as Root Mean Square Error (RMSE) of Android-based positioning is 0.82m, and that for iPhone is 1.85m. Before calibration, the last step by iPhone detection needs to be removed as the reference path only has 83 steps. It will not affect the final calibration result as the whole length of the path estimated by iPhone is also much longer than the reference path when comparing Fig.5 and Fig.7b.



(a)



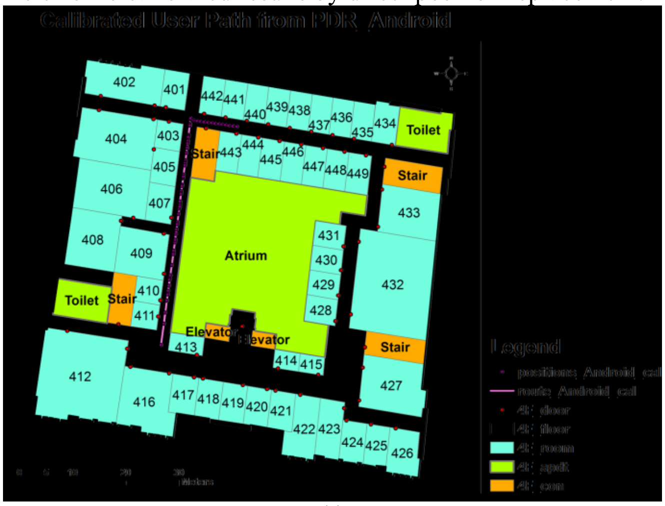
(b)

Fig. 7. The results of smartphone-based PDR by Android (a) and iOS (b).

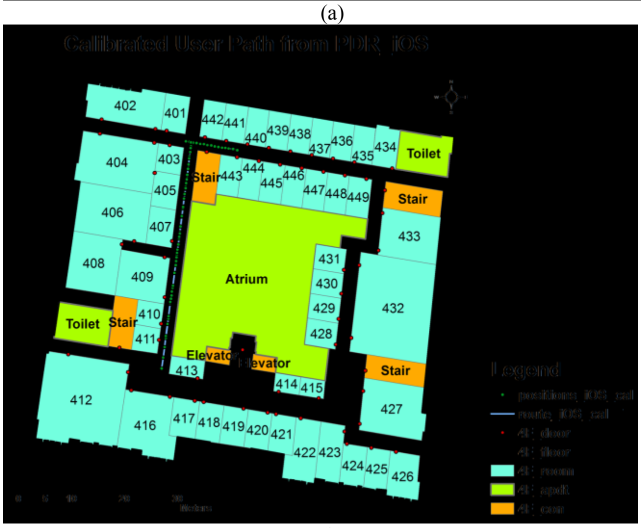
B. Post-Calibration

1) Position Replacement

The position-replacement method directly replaces the PDR results by vision-based tracking positions based on their similar time stamps as described in Section II D(1). The results for two smartphones are shown in Fig.8. Comparing to the results from pre-calibration, it provides a better solution for positioning than using single PDR tracking system as it takes the advantage of accurate positioning by visual tracking in LOS area. Moreover, as the ending positioning point of calibrated PDR is matched to that of reference path, it can provide a correct starting point for the following tracking if a second camera is introduced to the current system. The RMSEs for two smartphones reach a similar level, which are 0.73m (Android) and 0.75m (iOS) respectively, meaning this method is able to handle the positioning calibration regardless of the smartphone models. In addition, it is more effective on iPhone-based calibration as it improves its performance about 59% than 11% by Android system. However, as previously mentioned in Section IV A(1), visual tracking is affected by pinhole effect which has the problem of uneven distribution of step points, it also introduces the error to calibrated results by direct position replacement.



(a)

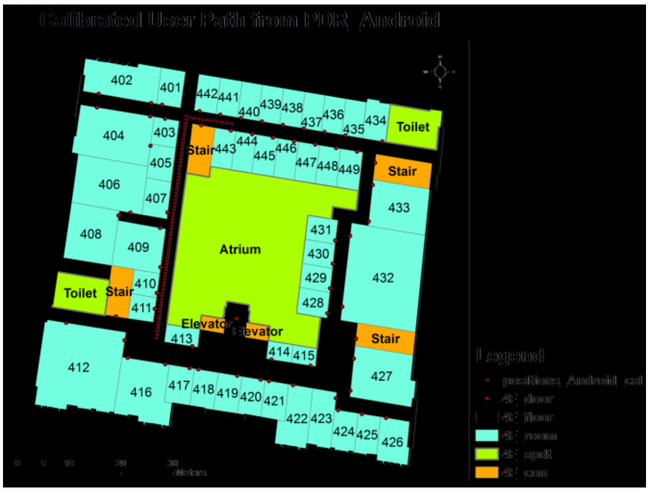


(b)

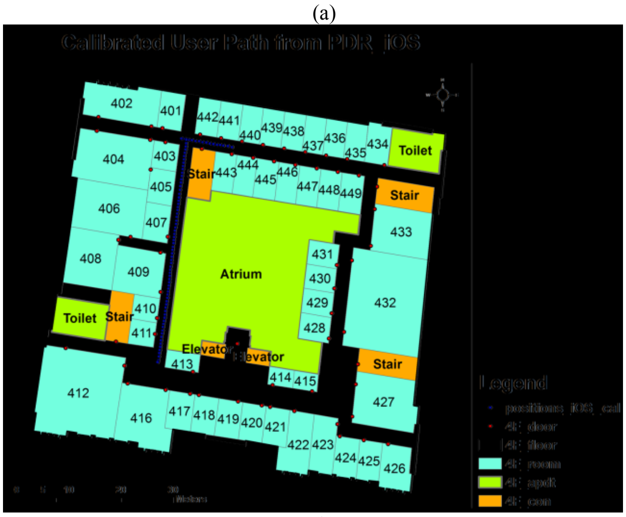
Fig. 8. Calibrated results of smartphone-based PDR by position replacement by Android (a) and iOS (b).

2) *Heading Calibration*

As mentioned in Section II D(2), the time stamps for two positioning systems cannot be exactly same, this method is more close to reality as it corrects the heading information for position re-estimation. The calibrated positioning results are shown in Fig.9. Comparing to previous results by using the first method, this method is more accurate as the uneven distribution effect caused by pinhole model is removed by only calibrating the orientations but keeping original step lengths for position estimation. Meanwhile, it still takes the merits of previous position-replacing-based hybrid system. This leads to lower RMSEs as 0.51m (Android) and 0.56m (iOS), with 37.8% and 69.7% improvement respectively than pre-calibrated results (Table I). Moreover, it also has the potential to be adjusted in online calibrations as the heading information acquired between two consecutive frames can be directly used in real time PDR heading calibrations for position estimation other than post processing. However, it has the problem that the ending point may not be perfectly matched with the ending point of reference path as there are still some errors in step-length estimation process.



(a)



(b)

Fig. 9. Calibrated results of smartphone-based PDR by heading calibration by Android (a) and iOS (b).

TABLE I. RMSE OF TWO SMARTPHONE-BASED PDR BEFORE AND AFTER CALIBRATION

		Device	
Mean Location Accuracy (RMSE)		***HUAWEI MT7-TL00***	***iPhone 7 Plus***
Pre-Calibration		0.82 m	1.85 m
Post-Calibration	Position Replacement	0.73 m	0.75 m
	Improvement	10.9%	59.5%
	Heading Calibration	0.51 m	0.56 m
	Improvement	37.8%	69.7%

V. DISCUSSION AND FUTURE WORK

In all, the hybrid system has higher accuracy than single PDR system, regardless of using either position-replacing based or heading-correction based calibration. The heading calibration based approach is more accurate in this study as it keeps estimated step length information. As this study uses a fixed k for coefficient of step length estimation, it can be modified into a real-time value which is determined by the ratio between estimated distance and real distance. This may help to improve the accuracy of final positioning results as well. Meanwhile, the whole system can also be adjusted to an online mode in order to do calibration in real time instead of post processing in this paper. In addition, the results indicate that this design of hybrid system can handle both smartphone models by achieving similar level of accuracy after calibration. As both types of phones are common models in the market, it suggests that this system has the potential to become a ubiquitous solution for indoor positioning. In addition, as this system utilizes the existed indoor infrastructures and user devices, it can be a low cost solution as well as it needs no additional installations of sensors. However, a limitation of this system is that the surveillance cameras may not existed in indoor environment of some residential areas, the current solution is more suitable in public space with completed surveillance system.

This paper also compares the positioning performances of two different smartphones as well. According RMSEs of experimental results in this paper, the Android system seems to provide slightly better positioning service before and after calibration. However, this does not mean that iPhone is not good for indoor positioning, as the final RMSEs after calibration between two phones are not significantly different. In addition, the gyroscope of iPhone seems to have better performance than that in Android, as the headings after turning detected by iPhone have a better match with the directions in reference path. As the generations of smartphones are updating in a fast speed, the quality of built-in sensors will be more precise in future, providing better services of indoor positioning.

In the next step, this system will be developed into a more comprehensive one with the ability to track the entire movement of a single user in the building with multiple floors as this study only tests on a single floor. The smartphone-based PDR system will be further tested on staircase-walking with the support of a barometer for height detection, in order to automatically change to related floor plan. The visual tracking can also be updated into a multi-camera system, as this paper only uses one camera. As the whole building already has a complete surveillance system, the next step is to utilize these cameras to work cooperatively to calibrate the entire movement estimated by smartphone-based PDR. This needs to solve the problem of linking the ending point detected by the previous camera to the starting point detected by the following camera by calibrated PDR. With the smartphone-based PDR as the main component of system, this hybrid system can still work even under the situations when some of the cameras is down. The georeferenced floor plans of different floors in WGS 1984 need to be prepared as well as a support of system for absolute positioning information. . After achieving tracking user's entire movement in indoor environment, a future test is needed to check whether the transition between indoor and outdoor spaces is smooth and accurate as well. This will compare the accuracy of last point at the exit before entering the outdoor space estimated by this system and measured by GNSS signals.

The tracking of multiple users is also an objective in future work. However, there are some difficulties required to be solved. First, the re-identification of a specific person between two successive cameras will be a great challenge and requires more computation power. The current idea is to see whether calibrated smartphone-based PDR can help to identify the same person by using the ending point of last camera and to estimate the starting point for the next camera. Second, even in the same camera, the distinguishing of different users may also be a problem if there is some overlapping parts between users during movement. This may cause some errors in calibrating smartphone-based PDR. To overcome this problem, the standardization of walking paths into a uniform network may be a possible solution.

VI. CONCLUSION

This paper contributes a design of a hybrid system for smartphone-based PDR aided by OPS with the support of digital map information in WGS 1984 on distributed platforms, in order to improve its indoor positioning performance. Both of the sub-systems can work independently with the support of digital map information and the accuracy of inertial system in visible area will be improved by additional visual tracking information. It is tested on two common smartphones to see whether it is a potential ubiquitous solution for different smartphones in market. This hypothesis has been proved in this experiment as the accuracy of calibrated results from both smartphones has been significantly improved and achieved similar levels. It also contributes two kinds of calibration algorithms as direct position replacement and heading calibration. The latter one is more accurate than the former one as it follows the mechanism of PDR with the maintenance of estimated step information. This system also has the potential to be a low-cost solution as it does not need additional installation of sensors but only utilizing available sensors from user and indoor environment. In future, it also can be developed into a more comprehensive system in order to track the entire indoor movements of user.

ACKNOWLEDGMENT

The author acknowledges the financial support from the International Doctoral Innovation Centre, Ningbo Education Bureau, Ningbo Science and Technology Bureau, and the University of Nottingham. This work was also supported by the UK Engineering and Physical Sciences Re-search Council [grant number EP/L015463/1].

REFERENCES

[1] L. Jiang, L. N. Hoe, and L. L. Loon, "Integrated UWB and GPS location sensing system in hospital environment," in 2010 5th IEEE Conference on Industrial Electronics and Applications, 2010, pp. 286-289.

[2] M. B. Kjærgaard, H. Blunck, T. Godsk, T. Toftkjær, D. L. Christensen, and K. Grønbæk, "Indoor positioning using GPS revisited " in International conference on pervasive computing, 2010, pp. 38-56.

[3] A. Vu, A. Ramanandan, A. Chen, J. A. Farrell, and M. Barth, "Real-Time Computer Vision/DGPS-Aided Inertial Navigation System for Lane-Level Vehicle Navigation," IEEE Transactions on Intelligent Transportation Systems, vol. 13, pp. 899-913, 2012.

[4] L. Niu, "A survey of wireless indoor positioning technology for fire emergency routing," in IOP Conference Series: Earth and Environmental Science, 2014, p. 012127.

[5] J. Bao, Y. Zheng, D. Wilkie, and M. Mokbel, "Recommendations in location-based social networks: a survey," GeoInformatica, vol. 19, pp. 525-565, 2015.

[6] F. Bentley, H. Cramer, and J. Müller, "Beyond the bar: the places where location-based services are used in the city," Personal and Ubiquitous Computing, vol. 19, pp. 217-223, 2015.

[7] H. Yun, D. Han, and C. C. Lee, "Understanding the use of location-based service applications: do privacy concerns matter?," Journal of Electronic Commerce Research, vol. 14, p. 215, 2013.

[8] A. Basiri, E. S. Lohan, T. Moore, A. Winstanley, P. Peltola, C. Hill, et al., "Indoor location based services challenges, requirements and usability of current solutions," Computer Science Review, 2017.

[9] V. Filonenko, C. Cullen, and J. Carswell, "Investigating ultrasonic positioning on mobile phones," in Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on, 2010, pp. 1-8.

[10] H. S. Maghdid, I. A. Lami, K. Z. Ghafoor, and J. Lloret, "Seamless outdoors-indoors localization solutions on smartphones: implementation and challenges," ACM Computing Surveys (CSUR), vol. 48, p. 53, 2016.

[11] T. C. Dong-Si and A. I. Mourikis, "Estimator initialization in vision-aided inertial navigation with unknown camera-IMU calibration," in Ieee/rsj International Conference on Intelligent Robots and Systems, 2012, pp. 1064-1071.

[12] W. Elloumi, A. Latoui, R. Canals, A. Chetouani, and S. Treuillet, "Indoor Pedestrian Localization With a Smartphone: A Comparison of Inertial and Vision-Based Methods," IEEE Sensors Journal, vol. 16, pp. 5376-5388, 2016.

[13] Y. Sun, Y. Zhao, and J. Schiller, "An autonomic indoor positioning application based on smartphone," in Wireless Communications and Networking Conference (WCNC), 2014 IEEE, 2014, pp. 3326-3331.

[14] Q. Tian, Z. Salcic, I. Kevin, K. Wang, and Y. Pan, "An enhanced pedestrian dead reckoning approach for pedestrian tracking using smartphones," in Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on, 2015, pp. 1-6.

[15] C. Fuchs, N. Aschenbruck, P. Martini, and M. Wieneke, "Indoor tracking for mission critical scenarios: A survey," Pervasive and Mobile Computing, vol. 7, pp. 1-15, 2011.

[16] R. Harle, "A Survey of Indoor Inertial Positioning Systems for Pedestrians," IEEE Communications Surveys and Tutorials, vol. 15, pp. 1281-1293, 2013.

[17] J. Racko, P. Brida, A. Perttula, J. Parviainen, and J. Collin, "Pedestrian Dead Reckoning with Particle Filter for handheld smartphone," in Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on, 2016, pp. 1-7.

[18] J. P. Tardif, M. George, M. Laverne, and A. Kelly, "A new approach to vision-aided inertial navigation," in Ieee/rsj International Conference on Intelligent Robots and Systems, 2010, pp. 4161-4168.

[19] A. I. Mourikis and S. I. Roumeliotis, "A Multi-State Constraint Kalman Filter for Vision-Aided Inertial Navigation," vol. 22, pp. 3565-3572, 2007.

[20] O. J. Woodman, "An introduction to inertial navigation," University of Cambridge, Computer Laboratory2007.

[21] S. Rajagopal, "Personal dead reckoning system with shoe mounted inertial sensors," Master's Degree Project, Stockholm, Sweden, 2008.

[22] K. Abdulrahim, C. Hide, T. Moore, and C. Hill, "Aiding low cost inertial navigation with building heading for pedestrian navigation," Journal of navigation, vol. 64, pp. 219-233, 2011.

[23] P. C. Lin, J. C. Lu, C. H. Tsai, and C. W. Ho, "Design and Implementation of a Nine-Axis Inertial Measurement Unit," IEEE/ASME Transactions on Mechatronics, vol. 17, pp. 657-668, 2012.

[24] D. Griesbach, D. Baumbach, and S. Zuev, "Stereo-vision-aided inertial navigation for unknown indoor and outdoor environments," International Journal of Cultural Policy, vol. 20, pp. 281-295, 2014.

[25] J. A. B. Link, P. Smith, N. Viol, and K. Wehrle, "FootPath: Accurate map-based indoor navigation using smartphones," in IPIN, 2011, pp. 1-8.

[26] T. N. Hung and Y. S. Suh, "Inertial sensor-based two feet motion tracking for gait analysis," Sensors, vol. 13, pp. 5614-5629, 2013.

[27] C. He, P. Kazanzides, H. T. Sen, S. Kim, and Y. Liu, "An Inertial and Optical Sensor Fusion Approach for Six Degree-of-Freedom Pose Estimation," Sensors, vol. 15, pp. 16448-16465, 2015.

[28] N. Parnian and F. Golnaraghi, "Integration of a multi-camera vision system and strapdown inertial navigation system (SDINS) with a modified Kalman filter," Sensors, vol. 10, pp. 5378-5394, 2010.

[29] C. Hide, T. Botterill, and M. Andreotti, "Low cost vision-aided IMU for pedestrian navigation," in Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS), 2010, 2010, pp. 1-7.

[30] C. Hide, T. Botterill, and M. Andreotti, "Vision-aided IMU for handheld pedestrian navigation," in Proceedings of the institute of navigation GNSS 2010 Conference, Portland, Oregon, 2010.

[31] T. Liu, M. Carlberg, G. Chen, J. Chen, J. Kua, and A. Zakhor, "Indoor localization and visualization using a human-operated backpack system," in Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on, 2010, pp. 1-10.

[32] U. Steinhoff and B. Schiele, "Dead reckoning from the pocket-an experimental study," in Pervasive Computing and Communications (PerCom), 2010 IEEE International Conference on, 2010, pp. 162-170.

[33] E. Azimi, J. Doswell, and P. Kazanzides, "Augmented reality goggles with an integrated tracking system for navigation in neurosurgery," in Virtual Reality Short Papers and Posters (VRW), 2012 IEEE, 2012, pp. 123-124.

[34] P. Sadda, E. Azimi, G. Jallo, J. Doswell, and P. Kazanzides, "Surgical navigation with a head-mounted tracking system and display," in Studies in Health Technology and Informatics, ed: IOS Press, 2013.

[35] W. Kang, S. Nam, Y. Han, and S. Lee, "Improved heading estimation for smartphone-based indoor positioning systems," in Personal Indoor and Mobile Radio Communications (PIMRC), 2012 IEEE 23rd International Symposium on, 2012, pp. 2449-2453.

[36] J. Torres-Sospedra, A. R. Jiménez, S. Knauth, A. Moreira, Y. Beer, T. Fetzer, et al., "The Smartphone-Based Offline Indoor Location Competition at IPIN 2016: Analysis and Future Work," Sensors, vol. 17, p. 557, 2017.

[37] F. Zampella, A. R. Jiménez, and F. Seco, "LOPSI repository: GetSensorData application," 2017.

[38] R. Faragher, C. Sarno, and M. Newman, "Opportunistic radio SLAM for indoor navigation using smartphone sensors," in Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION, 2012, pp. 120-128.

[39] J. Liu, R. Chen, L. Pei, R. Guinness, and H. Kuusniemi, "A hybrid smartphone indoor positioning solution for mobile LBS," Sensors, vol. 12, pp. 17208-17233, 2012.

[40] J. Liu, R. Chen, Y. Chen, L. Pei, and L. Chen, "iParking: An intelligent indoor location-based smartphone parking service," Sensors, vol. 12, pp. 14612-14629, 2012.

[41] [41] G. Panahandeh and M. Jansson, "Vision-Aided Inertial Navigation Based on Ground Plane Feature Detection," IEEE/ASME Transactions on Mechatronics, vol. 19, pp. 1206-1215, 2014.

[42] [42] J. Pinchin, C. Hide, and T. Moore, "The use of high sensitivity gps for initialisation of a foot mounted inertial navigation system," in Position Location and Navigation Symposium (PLANS), 2012 IEEE/ION, 2012, pp. 998-1007.

[43] [43] A. Ramanandan, A. Chen, and J. A. Farrell, "Inertial Navigation Aiding by Stationary Updates," IEEE Transactions on Intelligent Transportation Systems, vol. 13, pp. 235-248, 2012.

[44] [44] I. Skog, P. Handel, J. O. Nilsson, and J. Rantakokko, "Zero-Velocity Detection—An Algorithm Evaluation," IEEE transactions on bio-medical engineering, vol. 57, pp. 2657-2666, 2010.

[45] [45] R. Mautz and S. Tilch, "Survey of optical indoor positioning systems," in Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on, 2011, pp. 1-7.

[46] [46] S. Beauregard and M. Klepal, "Indoor PDR performance enhancement using minimal map information and particle filters," in Position, Location and Navigation Symposium, 2008 IEEE/ION, 2008, pp. 141-147.

[47] [47] B. Krach and P. Robertson, "Cascaded estimation architecture for integration of foot-mounted inertial sensors," in Position, Location and Navigation Symposium, 2008 IEEE/ION, 2008, pp. 112-119.

[48] [48] B. Krach and P. Robertson, "Integration of foot-mounted inertial sensors into a Bayesian location estimation framework," in Positioning, Navigation and Communication, 2008. WPNC 2008. 5th Workshop on, 2008, pp. 55-61.

[49] [49] O. Woodman and R. Harle, "Pedestrian localisation for indoor environments," in Proceedings of the 10th international conference on Ubiquitous computing, 2008, pp. 114-123.

[50] [50] P. Robertson, M. Angermann, and B. Krach, "Simultaneous localization and mapping for pedestrians using only foot-mounted inertial sensors," in Proceedings of the 11th international conference on Ubiquitous computing, 2009, pp. 93-96.

[51] [51] J. Pinchin, C. Hide, and T. Moore, "A particle filter approach to indoor navigation using a foot mounted inertial navigation system and heuristic heading information," in Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on, 2012, pp. 1-10.

[52] [52] N. Castaneda and S. Lamy-Perbal, "An improved shoe-mounted inertial navigation system," in Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on, 2010, pp. 1-6.

[53] [53] A. R. Jimenez, F. Zampella, and F. Seco, "Light-matching: A new signal of opportunity for pedestrian indoor navigation," in Indoor Positioning and Indoor Navigation (IPIN), 2013 International Conference on, 2013, pp. 1-10.

[54] [54] S. Godha and G. Lachapelle, "Foot mounted inertial system for pedestrian navigation," Measurement Science and Technology, vol. 19, p. 075202, 2008.

[55] [55] A. R. Jimenez, F. Seco, C. Prieto, and J. Guevara, "A comparison of pedestrian dead-reckoning algorithms using a low-cost MEMS IMU," in Intelligent Signal Processing, 2009. WISP 2009. IEEE International Symposium on, 2009, pp. 37-42.

[56] [56] P. Goyal, V. J. Ribeiro, H. Saran, and A. Kumar, "Strap-down pedestrian dead-reckoning system," in Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on, 2011, pp. 1-7.

[57] [57] H. Weinberg, "Using the ADXL202 in pedometer and personal navigation applications," Analog Devices AN-602 application note, vol. 2, pp. 1-6, 2002.

[58] [58] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," Computer vision and image understanding, vol. 104, pp. 90-126, 2006.

[59] [59] Y. Zhou, S. Zlatanova, Z. Wang, Y. Zhang, and L. Liu, "Moving Human Path Tracking Based on Video Surveillance in 3D Indoor Scenarios," ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. Volume III-4, pp. 97-101, 2016.

[60] [60] T.-H. Tsai, C.-H. Chang, and S.-W. Chen, "Vision based indoor positioning for intelligent buildings," in Intelligent Green Building and Smart Grid (IGBSG), 2016 2nd International Conference on, 2016, pp. 1-4.

[61] [61] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91-99.

[62] [62] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," IEEE transactions on pattern analysis and machine intelligence, vol. 34, pp. 743-761, 2012.