



Article

Hybrid Indoor Localization Using IMU Sensors and Smartphone Camera

Alwin Poulouse  and Dong Seog Han * 

School of Electronics Engineering, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu 41566, Korea; alwinpoulosepalatty@knu.ac.kr

* Correspondence: dshan@knu.ac.kr; Tel.: +82-53-950-6609

Received: 23 September 2019; Accepted: 18 November 2019; Published: 21 November 2019



Abstract: Smartphone camera or inertial measurement unit (IMU) sensor-based systems can be independently used to provide accurate indoor positioning results. However, the accuracy of an IMU-based localization system depends on the magnitude of sensor errors that are caused by external electromagnetic noise or sensor drifts. Smartphone camera based positioning systems depend on the experimental floor map and the camera poses. The challenge in smartphone camera-based localization is that accuracy depends on the rapidness of changes in the user's direction. In order to minimize the positioning errors in both the smartphone camera and IMU-based localization systems, we propose hybrid systems that combine both the camera-based and IMU sensor-based approaches for indoor localization. In this paper, an indoor experiment scenario is designed to analyse the performance of the IMU-based localization system, smartphone camera-based localization system and the proposed hybrid indoor localization system. The experiment results demonstrate the effectiveness of the proposed hybrid system and the results show that the proposed hybrid system exhibits significant position accuracy when compared to the IMU and smartphone camera-based localization systems. The performance of the proposed hybrid system is analysed in terms of average localization error and probability distributions of localization errors. The experiment results show that the proposed oriented fast rotated binary robust independent elementary features (BRIEF)-simultaneous localization and mapping (ORB-SLAM) with the IMU sensor hybrid system shows a mean localization error of 0.1398 m and the proposed simultaneous localization and mapping by fusion of keypoints and squared planar markers (UcoSLAM) with IMU sensor-based hybrid system has a 0.0690 m mean localization error and are compared with the individual localization systems in terms of mean error, maximum error, minimum error and standard deviation of error.

Keywords: indoor positioning system (IPS); pedestrian dead reckoning (PDR); heading estimation; indoor navigation; IMU sensors; smartphone camera; Kalman filter; sensor fusion; simultaneous localization and mapping (SLAM); ArUco markers

1. Introduction

Indoor localization systems are classified as either building dependent or building independent based on the sensors used for localization [1]. The most common building independent indoor positioning technologies are pedestrian dead reckoning (PDR) systems using inertial measurement unit (IMU) sensors [2–4] and image based technologies using cameras [5–7]. The IMU sensor used in PDR systems includes the accelerometer, magnetometer and gyroscope sensors and these sensors give user position based on user heading and step length information. In image based indoor positioning, a camera is used for localization and any of monocular, stereo or RGB-D cameras can be used for localization. The camera captures the experiment area and the captured data is fed to image-based localization algorithms. The most popular image-based localization system, which consists of family

of algorithms is the simultaneous localization and mapping (SLAM) [8,9]. In SLAM based localization, the system estimates the position or orientation of the camera with respect to its surrounding and maps the environment based on the camera location. Localization using either IMU sensors or a camera system offers some level of accuracy although not the best. The limited level of accuracy in IMU sensors is due to the accumulated errors from accelerometer, drift errors from gyroscope and external magnetic fields that affect the magnetometer. These sensor errors degrade indoor position accuracy, hence the need for compensation.

In this paper, a camera-based system is introduced to the IMU localization system. However, it should be noted that in the camera based localization system, rapid user direction changes affect the effectiveness of the camera pose estimation. In the proposed hybrid systems, the results from the IMU-based system are used to compensate the heading error from the camera-based system. Similarly, the IMU sensor errors are compensated for by utilizing results from the camera-based system. Summarily, in this paper, we propose hybrid indoor localization systems that combine the results from IMU and camera based systems to improve positioning accuracy. Experimental results demonstrate the effect of the proposed hybrid fusion method and the proposed hybrid method reduces the sensor errors for IMU localization and heading error for the camera based localization system. The main contributions of this paper are as follows:

- We implemented an IMU-based indoor localization system. The proposed IMU system uses the accelerometer, gyroscope and magnetometer for position estimation. A pitch based estimator is used for step detection and step length estimation. A sensor fusion algorithm is used for heading estimation. The user position is estimated by using step length and heading information.
- We followed an oriented fast rotated binary robust independent elementary features (ORB)-SLAM algorithm proposed by Mur-Artal et al. [10] for the camera based localization system. The ORB-SLAM uses the same features for tracking, mapping, relocalization and loop closing. This makes the ORB-SLAM system more efficient, simple and reliable as compared to other SLAM techniques.
- We developed a SLAM by fusion of keypoints and squared planar markers (UcoSLAM) algorithm proposed by Munoz-Salinas et al. [11] for the camera based localization system by adding markers to the experiment area. We used Augmented Reality Uco Codes (ArUco) markers for localization and the markers improved the localization accuracy.
- We proposed hybrid indoor localization systems using an IMU sensor and a smartphone camera. The sensor fusion is achieved by a Kalman filter and the proposed systems reduced the IMU sensor errors and heading errors from camera-based localization systems.

The rest of the paper is organized as follows: in Section 2, a review of previous work is discussed. A model for indoor localization using an IMU and a camera is presented in Section 3. The experimental setup and result analysis is given in Section 4. Finally, Section 5 concludes the work and gives future directions.

2. Related Work

Indoor localization has been studied in the past and recent times based on different localization techniques and technologies used [12–15]. In this section, we discuss the existing technologies used for IMU-based localization, camera based localization and various existing hybrid approaches for indoor localization.

The IMU-based indoor localization uses an accelerometer, gyroscope and magnetometer sensors for position estimation and is also known as pedestrian dead reckoning (PDR). Different PDR approaches [16,17] have been proposed for indoor localization and these approaches have a significant role in indoor positioning. The PDR study includes step detection [18–20], step length estimation [21–25], heading estimation [26–28] and position estimation using step length and heading information. The accuracy of PDR position results depends on the accurate step length estimation and

accurate heading estimation. The basic PDR models are explained in [29–34]. In these PDR models, different algorithms for indoor localization are explained and their proposed models show significant position accuracy improvements for IMU-based indoor localization. Smartphone IMU-based indoor localization systems are discussed in [35–39]. The smartphone IMU-based PDR systems have many challenges due to the changes in smartphone coordinates. The changes in smartphone coordinates are based on the user movements, hence affecting IMU sensor readings. Several studies have been conducted on the smartphone IMU-based PDR system in indoor environments during the past years [40–42]. These proposed smartphone based IMU localization systems achieve accurate position results for indoor applications. However, the proposed PDR systems still exhibit sensor errors which affect the indoor position accuracy. More recent works on IMU-based localization are shown in [43–47]. In these works they reached significant position accuracy levels for indoor localization. From all these studies, it can be seen that IMU sensor-based localization is not free from sensor errors. It should also be noted that indoor position accuracy depends on accurate sensor reading and calibration. In this paper, we used our previous model presented in [48] for IMU sensor-based localization. In [48], we implemented a pitch based step detector, step length estimator and position estimation algorithm for indoor localization. Our proposed IMU-based localization model in [48] reduced the sensor errors and gives significant results for indoor localization.

A camera based localization system has a major role when the experiment area is independent from the building infrastructure. This method is also known as computer vision [49–52]. In this method, we use a camera and captures the environment in the form of images or video. The video or image data from the camera is used for estimating the position and orientation of an object or device. Camera based localization can be divided based on the markers used in the indoor area. If the localization system depends on the natural landmarks such as corridors, edges, doors, wall, ceiling light etc., it is referred to as markerless localization [53–57]. If we use some special type of markers such as fiducial markers or ArUco markers in the experiment area, then the localization is known as marker based localization [58,59]. The most common technique used in computer vision is SLAM based localization [60–64]. In SLAM based localization, we create a map of the experiment area and at the same time locate the camera position. The SLAM technique is classified as extended Kalman filter (EKF) SLAM [65,66], FastSLAM [67], low dimensionality (L)-SLAM [68], GraphSLAM [69], Occupancy Grid SLAM [70–72], distributed particle (DP)-SLAM [73], parallel tracking and mapping (PTAM) [74], stereo parallel tracking and mapping (S-PTAM) [75], dense tracking and mapping (DTAM) [76,77], incremental smoothing and mapping (iSAM) [78], large-scale direct (LSD)-SLAM [79], MonoSLAM [80], collaborative visual SLAM (CoSLAM) [81], SeqSLAM [82], continuous time (CT)-SLAM [83], UcoSLAM [11], RGB-D SLAM [84] and ORB SLAM [85–87]. In this paper, we used ORB SLAM and UcoSLAM for camera based localization. The ORB SLAM is a feature based localization system and it operates in real time in indoor environments. The ORB SLAM includes tracking, mapping, relocalization and loop closing. The ORB SLAM achieved significance indoor position accuracy as compared to other state-of-the-art monocular SLAM approaches. To improve the localization accuracy of camera based localization systems, a simple marker-based localization system is introduced in [88]. The localization system proposed in [88] added the markers into the map with a Tf package to the robot operating system (ROS) [89]. A 2D marker based monocular visual-inertial EKF-SLAM system is proposed in [90] and a planar marker based mapping and localization system is explained in [91]. The marker based systems explained have reduced the localization error as compared to markerless localization. However, the marker based localization systems have serious problems such as mapping distortion due to lack of correction, drift error if the markers are lost for a long time and it requires increasing computation resources to handle a multiple number of markers. In this paper, we use 4 ArUco markers for localization and the UcoSLAM based localization approach gives more accurate results as compared to ORB-SLAM localization approach.

To reduce the sensor errors from IMU-based localization and heading errors for the camera-based system, various studies have proposed a hybrid indoor localization approach. In hybrid based

localization, we use the IMU data and camera data together for position estimation. The basic models explained in [92–101] are used for IMU and camera data fusion. In their proposed methods, they combined the IMU data and vision data for better performance. Their experiment results proved that the hybrid indoor localization system has significant positional accuracy for indoor applications. A comparative study of IMU and camera based localization can be seen in [102]. Most recent work on hybrid localization is discussed in [103–109]. Among these studies, the hybrid indoor localization can be used for better performance and which reduces IMU sensor and camera position errors. In this paper, we use a linear Kalman filter (LKF) for combining IMU sensor and camera position data. The LKF is simple and easy to implement for real time applications as compared to the extended Kalman filter (EKF) [110–112], particle filter [113] and unscented Kalman filter (UKF) [114].

3. Model for Indoor Localization Using IMU Sensor and Smartphone Camera

The proposed model for indoor localization using IMU sensor and smartphone camera-based system is shown in Figure 1. In the proposed model, different localization approaches are fused with the sensor fusion frameworks. The proposed model is divided into three steps. Locating user position using IMU sensor is the first step of the proposed model. In IMU-based localization, the position is estimated by step length and heading information. The complementary features of accelerometer, gyroscope and magnetometer sensors are used for position estimation. In the second step, we use a smartphone camera for localization. The smartphone camera captures the experiment area and the captured data used in the ORB SLAM algorithm for position estimation. To improve the camera based localization, we used an UcoSLAM algorithm with ArUco markers. In the last step, we combine the position results from IMU and camera based systems with a Kalman filter. The sensor fusion framework uses a linear Kalman filter for combining position results from IMU and camera.

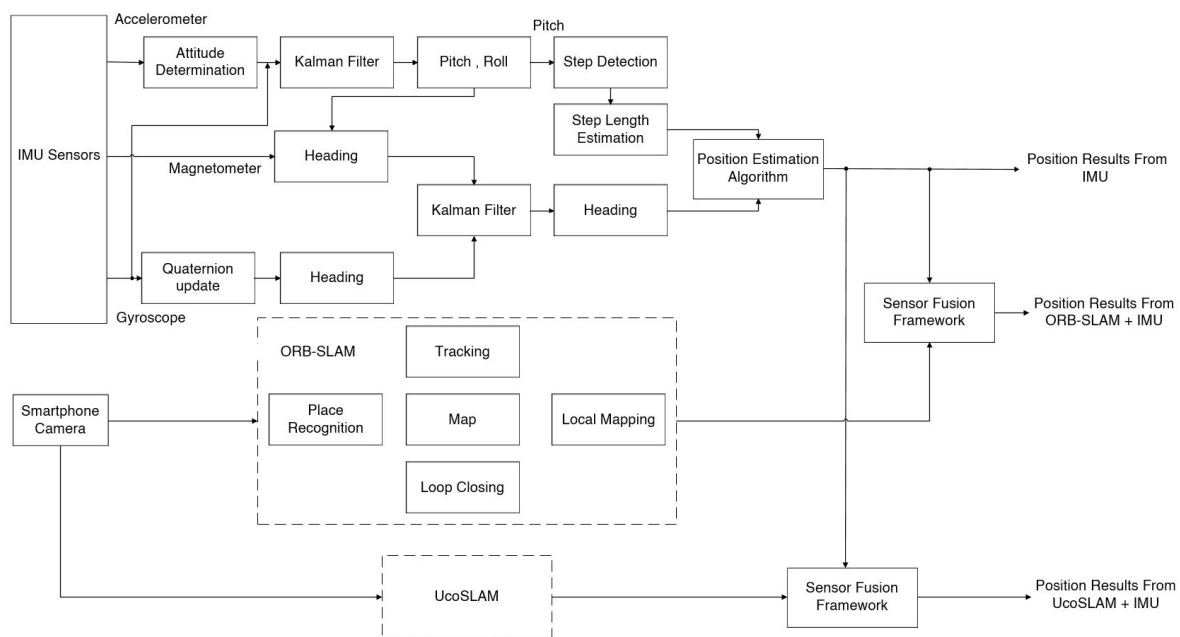


Figure 1. Proposed hybrid indoor localization model using IMU sensor and smartphone camera.

3.1. Indoor Localization Using IMU Sensor

The model presented in our previous work [48] is the same one that we used for IMU-based localization. The proposed model utilizes accelerometer, gyroscope and magnetometer for position estimation. A pitch based estimator is proposed for step detection. A sensor fusion algorithm is used for estimating pitch and roll from accelerometer and gyroscope. The sensor fusion pitch value is used for step detection. The step length is estimated from pitch amplitude. The heading is estimated from gyroscope and magnetometer fusion. Finally, the position is estimated using step length and

heading information. For more details on indoor localization using IMU sensor refer to our previous work in [48].

3.2. Indoor Localization Using Smartphone Camera

To enhance the indoor position accuracy, we used smartphone camera for localization when the privacy of the user is not a concern. The IMU sensor based localization gives accurate user position results for indoor localization. However, the user position results are not free from sensor error and it is necessary to compensate this sensor error by adding a smartphone camera to the system. In this paper, we used two algorithms for camera based localization. The most common camera based localization algorithm is the ORB-SLAM which gives the user position indoors. However, the keypoints mismatch and camera pose problems, the ORB-SLAM shows heading errors and some user position results are missing during the experiment time. To overcome these problems, we used UcoSLAM algorithm for localization. The UcoSLAM uses special markers called ArUco markers for localization. The ArUco markers solved the camera localization heading problems and improved the position accuracy for localization.

3.2.1. Localization Using ORM-SLAM

The model presented in [10] is used for camera based ORB-SLAM localization. In this localization, we use ORB features [115] instead of scale-invariant feature transform (SIFT) [116] or speeded-up robust feature (SURF) [117,118] for feature matching. The ORB feature matching allows real-time performance without GPUs and it gives best invariance to changes in viewpoint and illumination. The ORB-SLAM consists of tracking, local mapping and loop closing. In the tracking step, we estimate the camera position with every frame and control the new keyframe insertion. We used a motion only bundle adjustment (BA) for optimizing the camera pose using an initial feature matching with previous frame. If the tracking is not done, the place recognition module is used for global relocalization. When the initialization of camera poses and feature map is done, a local map is retrieved using the covisibility graph with keyframes that is used in the system. It then matches with the local map points which are searched by reprojection and finally optimize the camera pose with all matches. The last step of tracking is the new keyframe decision. If the new keyframe is inserted, it is used in the local mapping process. The local mapping uses the new keyframes and performs the local BA to get an optimal reconstruction in the surroundings of the camera pose. Loop closing is the final stage of ORB-SLAM. In loop closing stage, the system searches for loops with every new keyframe. If a loop is identified, a similarity transformation is computed which gives information about the accumulated drift in the loop. If a drift accumulation is detected, then the duplicated points are fused. Finally, to achieve global consistency, a pose graph optimization is performed. For more details on ORB-SLAM refer to [10].

3.2.2. Localization Using UcoSLAM

The ORB-SLAM approach uses natural landmarks (keypoints) for localization. However, it is unstable over time or insufficient for indoor localization. To improve the indoor localization accuracy, the UcoSLAM method proposed in [11] is used for camera based indoor localization. In UcoSLAM, we use artificial landmarks (ArCuo) in the experiment area for tracking and relocalization. Figure 2 shows the UcoSLAM architecture.

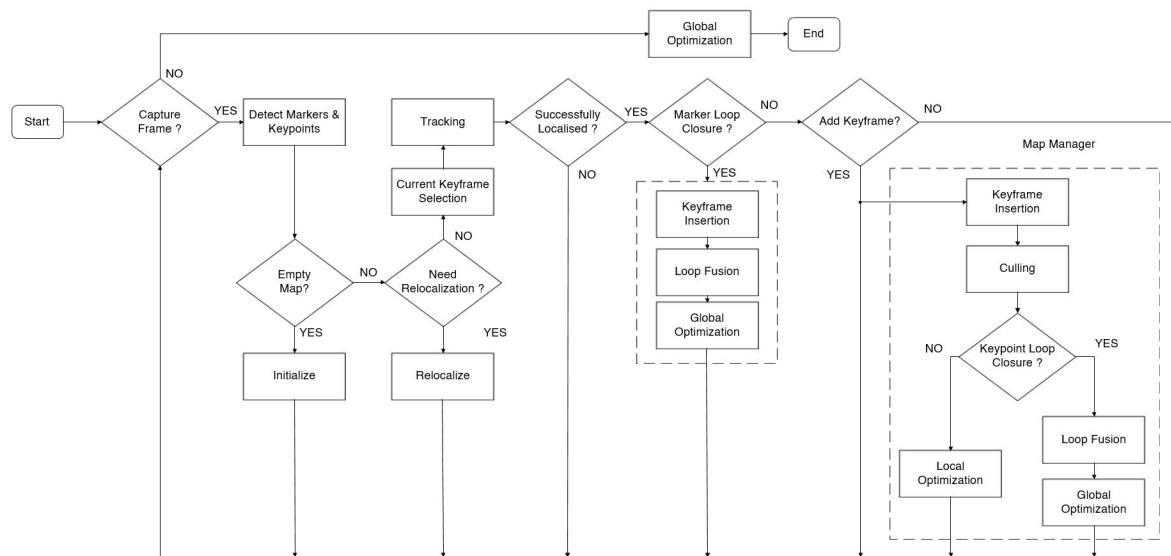


Figure 2. UcoSLAM architecture [11].

The UcoSLAM follows the same procedure used in the visual SLAM approaches for camera pose estimation. The main difference in the UcoSLAM approach is the combined use of keypoints and ArCuo markers for tracking and relocalization. The UcoSLAM system contains a map of the environment which is created and updated every time a new frame is available to the system. The UcoSLAM starts with map initialization. For map initialization, it uses homography, fundamental matrix (using keypoints) [86] and one or several markers [119]. After the map initialization in UcoSLAM, the system starts tracking or relocalization. If the system determines the camera pose in the last frame, it tries to estimate the current position using the last one as a starting point. In the UcoSLAM system, it uses the reprojection errors of keypoints and marker corners for tracking. For tracking in the UcoSLAM system, a reference frame is selected as map keyframe before tracking and map keyframe contains common matches to the frame analyzed in the previous time instant. After the tracking process, the system searches for the loop closures caused by ArCuo markers. If a loop closure is detected, then the system follows the keyframe insertion, loop fusion and global optimization steps as shown in Figure 2. If the loop closure is not detected, the system uses the map manager block, which runs the culling process when a new keyframe inserted. The culling process helps to maintain the map size manageable by removing the redundant information in the map. After the culling process, the system checks for the keypoint loop closure. If the keypoint loop closure is not detected, the system performs local optimization to integrate the new information. If the keypoint loop closure is detected, the system follows the loop fusion and global optimization steps. If the tracking process in the system failed in the last frame, then the system uses the relocalization mode. In the relocalization mode, it checks the markers already registered on the map. If the relocalization mode is unable to detect the known markers, it uses the bag-of-words (BoW) [120] process. For more details on UcoSLAM refer to [11].

3.3. Hybrid Indoor Localization Using IMU Sensor and Smartphone Camera

The objective of hybrid indoor localization is to improve the indoor position accuracy by reducing the IMU and camera sensor errors. For combining IMU localization results with camera based localization results, we used a linear Kalman filter (LKF) instead of other sensor fusion frameworks. The LKF is computationally light and we tackle the problem in a linear perspective. The model presented in [121] is used for sensor fusion frameworks. The system with controlled input and noise is given as

$$X_g(k) = AX_g(k-1) + BU(k-1) + \Gamma W(k-1) \quad (1)$$

$$A = \begin{bmatrix} F & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & F & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & F \end{bmatrix} \quad (2)$$

$$B = \begin{bmatrix} G & 0_{3 \times 3} & 0_{3 \times 3} \\ 0_{3 \times 3} & G & 0_{3 \times 3} \\ 0_{3 \times 3} & 0_{3 \times 3} & G \end{bmatrix} \quad (3)$$

$$U(k-1) = a^n(k-1) \quad (4)$$

with $F = \begin{bmatrix} 1 & t & -t/2 \\ 0 & t & -t \\ 0 & 0 & 1 \end{bmatrix}$, $G = \begin{bmatrix} t^2/2 \\ t \\ 0 \end{bmatrix}$, where k is the variable used for the recursive execution of Kalman filter, t is the sample period, A is the state-transition matrix, B is the controlled input matrix, $0_{3 \times 3}$ is the zero matrix, $U(k-1)$ is the controlled input, Γ is the noise matrix and $W(k-1)$ is the Gaussian white noise with variance $Q(k-1)$. The state vector $X_g(k)$ includes the position results from the IMU sensor. The measurement model of ORB-SLAM + IMU is the camera position $P_c(k) = (P_x^C(k), P_y^C(k))$ and the measurement function is

$$z_1(k) = \begin{bmatrix} P_x^C(k) & 0 & 0 & P_y^C(k) & 0 & 0 \end{bmatrix}^T = H_1 X_1(k) + n_1(k) \quad (5)$$

where $H_1 = \text{diag}(1, 0, 0, 1, 0, 0)$, $X_1(k)$ is the state vector of ORB-SLAM + IMU fusion, $n_1(k)$ is the measurement noise with covariance matrix $R_1(k)$. The measurement of UcoSLAM + IMU is the camera position $P_u(k) = (P_x^u(k), P_y^u(k))$ and the measurement function is

$$z_2(k) = \begin{bmatrix} P_x^u(k) & 0 & 0 & P_y^u(k) & 0 & 0 \end{bmatrix}^T = H_2 X_2(k) + n_2(k) \quad (6)$$

where $H_2 = \text{diag}(1, 0, 0, 1, 0, 0)$, $X_2(k)$ is the state vector of UcoSLAM + IMU fusion, $n_2(k)$ is the measurement noise with covariance matrix $R_2(k)$. The sensor fusion algorithm uses linear Kalman filter since the state and measurement functions are both linear. The Kalman filter consists of two processes, predicting and updating [121].

Predicting:

$$\hat{X}_i(k, k-1) = A\hat{X}_i(k-1) + BU(k-1) \quad (7)$$

$$\hat{P}_i(k, k-1) = A\hat{P}_i(k-1)A^T + Q_i(k-1) \quad (8)$$

Updating:

$$K_i(k) = \hat{P}_i(k, k-1)H_i^T \left(H_i\hat{P}_i(k, k-1)H_i^T + R_i(k) \right)^{-1} \quad (9)$$

$$\hat{X}_i(k) = \hat{X}_i(k, k-1) + K_i(k) (z_i(k) - H_i\hat{X}_i(k, k-1)) \quad (10)$$

$$P_i(k) = (I - K_i(k)H_i)\hat{P}_i(k, k-1) \quad (11)$$

The variables used in the LKF algorithm are summarized in Table 1.

Table 1. Variables used in the LKF algorithm.

\hat{X}_i	Estimated State Vector
$P_i(k)$	Estimated state covariance
$Q_i(k-1)$	Process noise
$R_i(k)$	Measurement noise covariance matrix
$K_i(k)$	Kalman gain matrix
$z_i(k)$	Measurement
A	State-transition matrix
B	Controlled input matrix
H	Observation model
$\hat{X}_i(k, k-1), \hat{P}_i(k, k-1)$	For internal computation

4. Experiment and Result Analysis

To evaluate the performance and accuracy of the proposed hybrid indoor localization systems, we considered an experiment scenario shown in Figure 3. The data from IMU and camera were collected at the fifth floor of IT building 1, Kyungpook National University, South Korea. During data collection, a user of age 27 and height 172 cm held a smartphone and the IMU sensor in his hand and walked around the table as shown in Figure 3c.

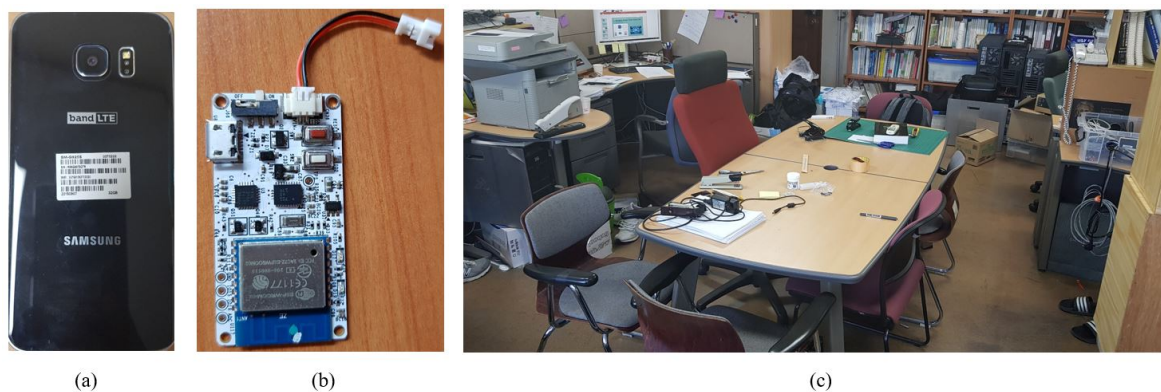


Figure 3. Experiment setup. (a) Smartphone. (b) IMU sensor. (c) Experiment area.

The experiment was conducted using an Android 5.0.2 Lollipop platform (Google, Mountain View, CA, USA) on the Samsung Galaxy S6 edge smartphone with Exynos 7420 processor and 3 GB ram. A Biscuit™ Programmable Wi-Fi 9-Axis absolute orientation sensor is used for IMU localization. To reduce the computational complexity and delay problems from the proposed algorithm, we carried out the proposed algorithm in an external server computer instead of smartphone to estimate the final user position. The length of the experiment area is 3.1 m and the width is 6 m. The experiment is carried out strictly along the reference path. The IMU sensor and smartphone camera data are collected during the user motions in the reference path. The localization algorithms use the collected data and estimate the current user position. For ground truth value estimation, the starting position of the user is assumed as zero and manually measured the coordinates of the reference path. To analyze the performance of the proposed system, we compared the estimated position results from the localization systems with ground truth values and estimated the user position error in terms of meters. Figure 4 shows the experimental results from IMU-based localization and camera based localization approaches.

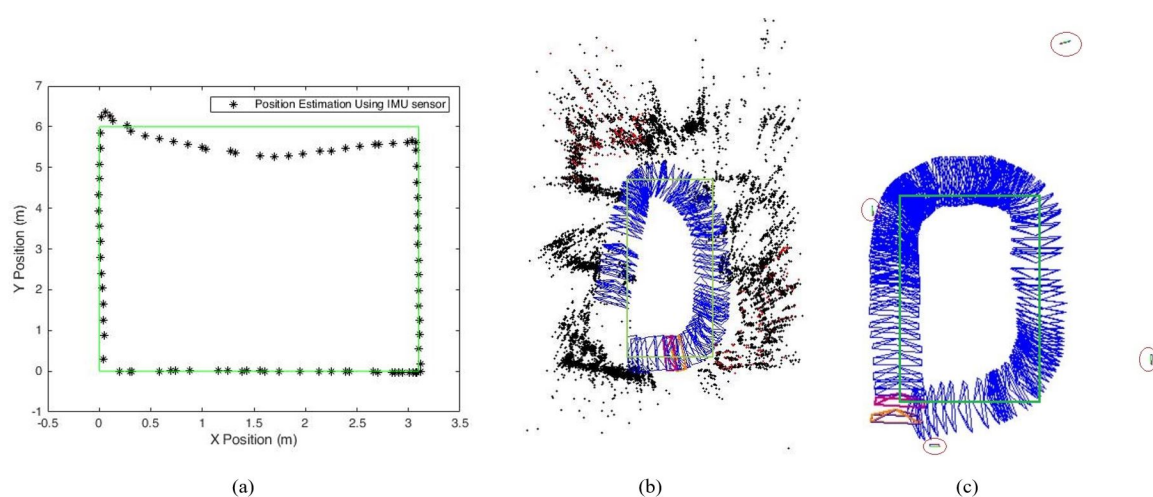


Figure 4. Indoor localization. (a) IMU-based localization. (b) ORB-SLAM localization. (c) UcoSLAM localization.

From Figure 4, the localization results show that the ArUco markers improved the localization accuracy as compared to the IMU and ORB-SLAM based localization approaches. In IMU-based localization, the accuracy of localization depends on the IMU sensor errors and the position estimation results from IMU approach is not accurate. In the case of ORB-SLAM localization approach, some user points are missing due to the absence of keypoints in the experiment area and the heading error from ORB-SLAM approach is higher than IMU-based localization. To improve the localization accuracy of ORB-SLAM approach, we used an UcoSLAM approach with ArUco markers and the ArUco markers reduced the heading error. The red circles in Figure 4c indicate the ArUco markers used for UcoSLAM localization.

To improve the indoor localization accuracy from IMU-based and camera based localization systems, we proposed a hybrid indoor localization system using IMU and camera features together and the estimated position results from proposed hybrid systems are shown in Figure 5.

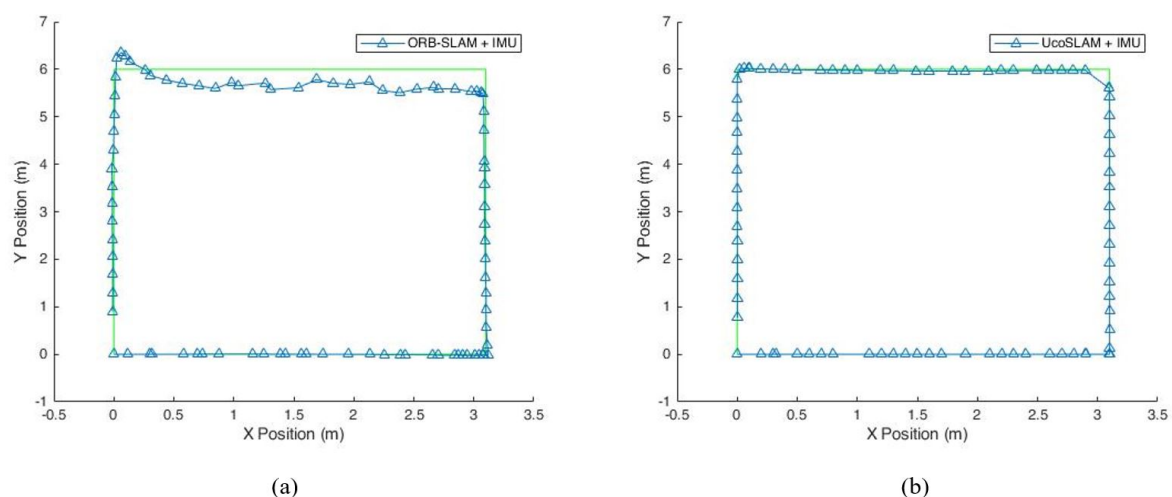


Figure 5. Proposed hybrid systems. (a) ORB-SLAM + IMU. (b) UcoSLAM + IMU.

From Figure 5, it can be seen that the proposed hybrid localization systems remove the effect of IMU sensor errors. The IMU-based localization exhibits accumulated error from the accelerometer and drift error from the gyroscope. The camera based localization cannot estimate the position accurately when user changes direction. The proposed ORB-SLAM + IMU hybrid system overcomes the challenges of IMU and camera based systems. The proposed UcoSLAM + IMU based hybrid

system is free from marker based localization problems. When the markers are not detected for a long time, the proposed UcoSLAM uses the IMU position results for localization. The mapping distortion and drift errors in the UcoSLAM are overcome by the effective utilization of the IMU position results. The IMU sensor errors are also reduced by the ArUco markers in the UcoSLAM and the proposed UcoSLAM + IMU outperforms conventional hybrid localization systems and gives high position accuracy for indoor localization.

The accuracy of the proposed hybrid localization systems is evaluated using the average localization error and probability distribution function of localization errors. The average localization error (E) is defined as

$$E = \sum_{i=1}^L \left((x_i^{true} - x_i^{est})^2 + (y_i^{true} - y_i^{est})^2 \right)^{0.5} \quad (12)$$

where (x_i^{true}, y_i^{true}) is the actual user position and (x_i^{est}, y_i^{est}) represents the estimated coordinate of unknown user position calculated by localization methods. L is the total number of data samples used for the localization. The average localization error results of IMU-based localization, camera based localization and proposed hybrid localization systems are shown in Figure 6.

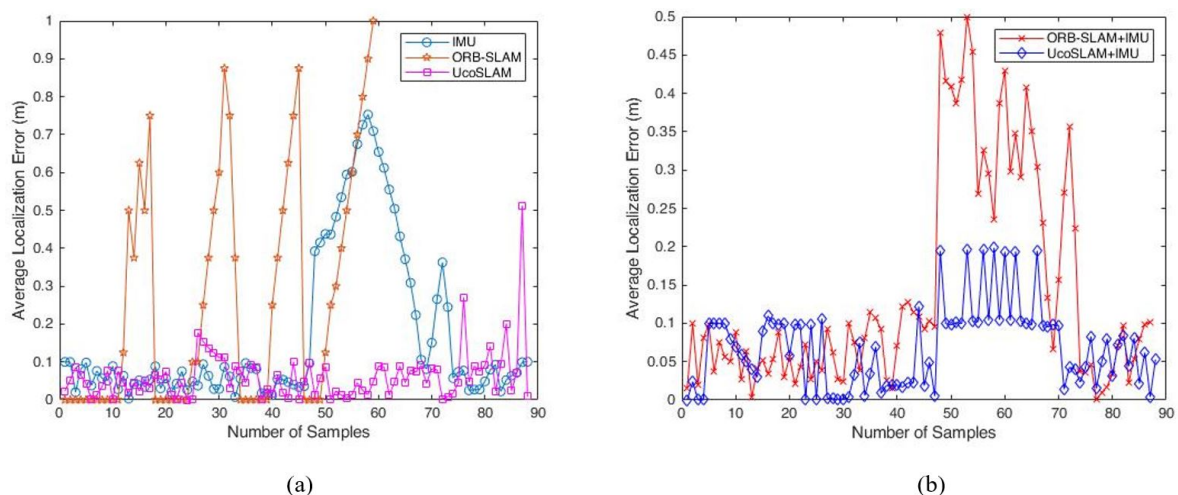


Figure 6. Average localization error. (a) IMU-based localization, camera based localization (b) Proposed hybrid localization systems.

Figure 6a shows the average localization error results from IMU, ORB-SLAM and UcoSLAM based localization approaches. The maximum error from IMU-based localization approach is 0.7528 m when compared to the ground truth values. The ORB-SLAM localization approach gives 1 m localization error when compared to the reference path. The UcoSLAM localization approach shows a maximum of 0.5120 m when compared to actual values. Figure 6b shows the average localization error results from the proposed hybrid localization systems. The proposed hybrid systems show reasonable localization accuracy when compared to the true position values. From the average localization error results, the proposed hybrid localization approaches improved the localization accuracy when compared to the independent localization approaches. Table 2 shows the performance analysis of different localization approaches.

Table 2. Performance of different localization approaches.

Localization Method	Mean Error (m)	Max. Error (m)	Min. Error (m)	Standard Deviation of Error (m)
IMU	0.1710	0.7528	0.0031	0.2088
ORB-SLAM	0.2669	1	0	0.3150
UcoSLAM	0.0617	0.5120	0	0.0685
ORB-SLAM + IMU	0.1398	0.4996	0.0011	0.1375
UcoSLAM + IMU	0.0690	0.1985	0	0.0532

From Table 2, the proposed hybrid localization systems have less localization errors as compared to the IMU and camera based localization approaches. The proposed hybrid localization gives high position accuracy as compared to other localization approaches and the accuracy of proposed hybrid approach is analyzed by probability distributions of localization errors. Figure 7 shows the probability distributions of localization errors of IMU-based localization, camera based localization and proposed hybrid localization approaches.

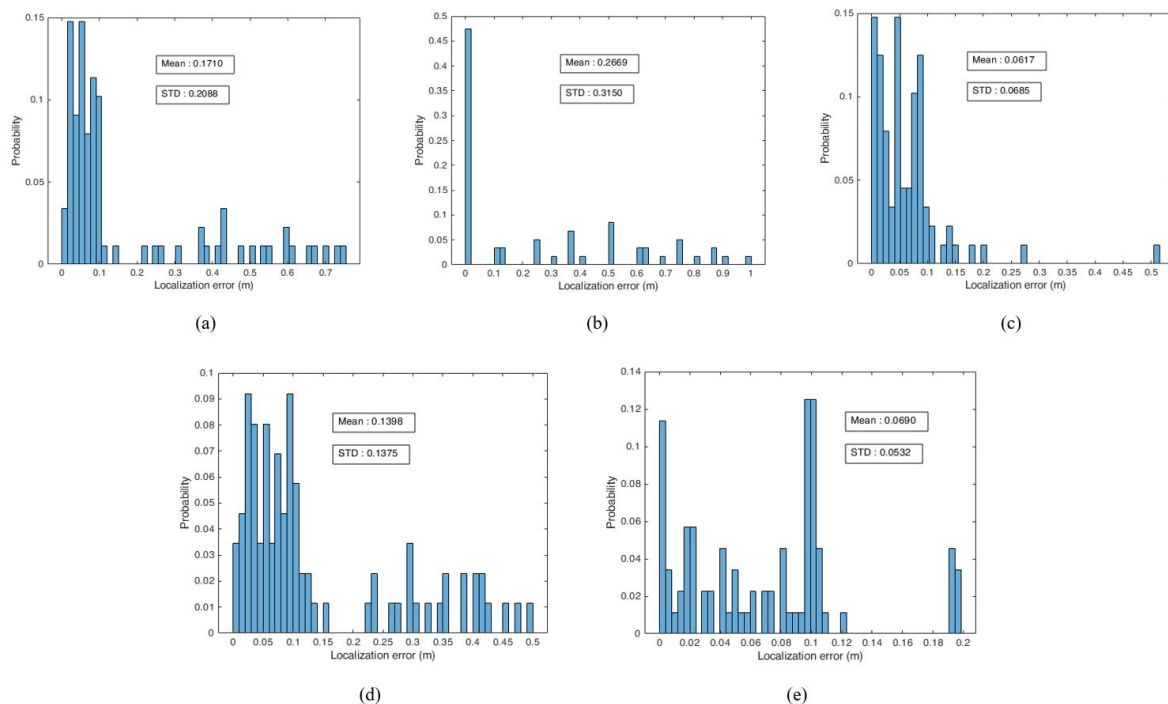


Figure 7. Probability distribution of localization errors. (a) IMU-based localization. (b) ORB-SLAM localization. (c) UcoSLAM localization. (d) ORB-SLAM + IMU. (e) UcoSLAM + IMU.

From Figure 7, it is clear that the proposed hybrid localization approach reduces the position errors and gives better performance when compared to the IMU and camera-based localization approaches. The localization system based on ORB-SLAM approach has high possibilities for zero localization error as compared to other localization systems. However, the system cannot estimate all user positions due to the lack of keypoints in the experiment area. The UcoSLAM system has less mean localization error than IMU and ORB-SLAM based systems. The IMU-based localization gives better performance compared to camera-based localization in terms of heading errors. The error analysis shows that the camera-based localization approach is affected by camera pose errors and hence degrades the indoor position accuracy. From all this experiment result analysis, the proposed hybrid localization approach gives reasonable position accuracy for indoor applications.

5. Conclusions

This paper presented hybrid indoor localization system using IMU sensor and smartphone camera. The IMU sensor errors are reduced by smartphone camera pose and the heading errors from camera-based localization is overcome by the IMU localization results. The proposed hybrid method show better positioning results compared to the individual localization method. We proposed a Kalman filter based sensor fusion framework for hybrid localization approaches which enhanced the results of the proposed methods. To improve the indoor position accuracy, the authors intend to utilize optical flow and ultrasonic sensors in the proposed system in future work.

Author Contributions: Writing-original draft, A.P.; Writing-review & editing, D.S.H.

Funding: This work was supported by Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Alarifi, A.; Al-Salman, A.; Alsaleh, M.; Alnafessah, A.; Al-Hadhrami, S.; Al-Ammar, M.; Al-Khalifa, H. Ultra wideband indoor positioning technologies: Analysis and recent advances. *Sensors* **2016**, *16*, 707. [[CrossRef](#)] [[PubMed](#)]
- Tian, Z.; Zhang, Y.; Zhou, M.; Liu, Y. Pedestrian dead reckoning for MARG navigation using a smartphone. *EURASIP J. Adv. Signal Process.* **2014**, *2014*, 65. [[CrossRef](#)]
- Zhou, R. Pedestrian dead reckoning on smartphones with varying walking speed. In Proceedings of the International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6.
- Racko, J.; Brida, P.; Perttula, A.; Parviainen, J.; Collin, J. Pedestrian dead reckoning with particle filter for handheld smartphone. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcalá de Henares, Spain, 4–7 October 2016; pp. 1–7.
- Kawaji, H.; Hatada, K.; Yamasaki, T.; Aizawa, K. An image-based indoor positioning for digital museum applications. In Proceedings of the International Conference on Virtual Systems and Multimedia, Seoul, Korea, 20–23 October 2010; pp. 105–111.
- Niu, Q.; Li, M.; He, S.; Gao, C.; Gary Chan, S.-H.; Luo, X. Resource-efficient and Automated Image-based Indoor Localization. *ACM Trans. Sensor Netw.* **2019**, *15*, 19. [[CrossRef](#)]
- Kawaji, H.; Hatada, K.; Yamasaki, T.; Aizawa, K. 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, Firenze, Italy, 29 October 2010; pp. 1–4.
- Kohlbrecher, S.; Von Stryk, O.; Meyer, J.; Klingauf, U. A flexible and scalable slam system with full 3d motion estimation. In Proceedings of the International Symposium on Safety, Security, and Rescue Robotics, Kyoto, Japan, 1–5 November 2011; pp. 155–160.
- Santos, J.M.; Portugal, D.; Rocha, R.P. An evaluation of 2D SLAM techniques available in robot operating system. In Proceedings of the International Symposium on Safety, Security, and Rescue Robotics (SSRR), Linköping, Sweden, 21–26 October 2013; pp. 1–6.
- Mur-Artal, R.; Montiel, J.M.M.; Tardos, J.D. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Trans. Robot.* **2015**, *31*, 1147–1163. [[CrossRef](#)]
- Munoz-Salinas, R.; Medina-Carnicer, R. UcoSLAM: Simultaneous Localization and Mapping by Fusion of KeyPoints and Squared Planar Markers. *arXiv* **2019**, arXiv:1902.03729.
- Zafari, F.; Gkelias, A.; Leung, K.K. A survey of indoor localization systems and technologies. *IEEE Commun. Surv. Tutor.* **2019**, *21*, 2568–2599. [[CrossRef](#)]
- Yassin, A.; Nasser, Y.; Awad, M.; Al-Dubai, A.; Liu, R.; Yuen, C.; Raulefs, R.; Aboutanios, E. Recent advances in indoor localization: A survey on theoretical approaches and applications. *IEEE Commun. Surv. Tutor.* **2016**, *19*, 1327–1346. [[CrossRef](#)]
- Brena, R.F.; García-Vázquez, J.P.; Galván-Tejada, C.E.; Muñoz-Rodríguez, D.; Vargas-Rosales, C.; Fangmeyer, J. Evolution of indoor positioning technologies: A survey. *J. Sens.* **2017**, *2017*, 2630413. [[CrossRef](#)]

15. Sakpere, W.; Adeyeye-Oshin, M.; Mlitwa, N.B. A state-of-the-art survey of indoor positioning and navigation systems and technologies. *S. Afr. Comput. J.* **2017**, *29*, 145–197. [[CrossRef](#)]
16. Harle, R.A. A survey of indoor inertial positioning systems for pedestrians. *IEEE Commun. Surv. Tutor.* **2013**, *15*, 1281–1293. [[CrossRef](#)]
17. Wu, Y.; Zhu, H.-B.; Du, Q.-X.; Tang, S.-M. A survey of the research status of pedestrian dead reckoning systems based on inertial sensors. *Int. J. Autom. Comput.* **2019**, *16*, 65–83. [[CrossRef](#)]
18. Kim, J.W.; Jang, H.J.; Hwang, D.-H.; Park, C. A step, stride and heading determination for the pedestrian navigation system. *Positioning* **2004**, *1*. [[CrossRef](#)]
19. Ho, N.-H.; Truong, P.; Jeong, G.-M. Step-detection and adaptive step-length estimation for pedestrian dead-reckoning at various walking speeds using a smartphone. *Sensors* **2016**, *16*, 423. [[CrossRef](#)] [[PubMed](#)]
20. Liu, Y.; Chen, Y.; Shi, L.; Tian, Z.; Zhou, M.; Li, L. Accelerometer based joint step detection and adaptive step length estimation algorithm using handheld devices. *J. Commun.* **2015**, *10*, 520–525. [[CrossRef](#)]
21. Vežočník, M.; Juric, M.B. Average Step Length Estimation Models' Evaluation Using Inertial Sensors: A Review. *IEEE Sens. J.* **2018**, *19*, 396–403. [[CrossRef](#)]
22. Díez, L.E.; Bahillo, A.; Otegui, J.; Otim, T. Step length estimation methods based on inertial sensors: A review. *IEEE Sens. J.* **2018**, *18*, 6908–6926. [[CrossRef](#)]
23. Xing, H.; Li, J.; Hou, B.; Zhang, Y.; Guo, M. Pedestrian stride length estimation from IMU measurements and ANN based algorithm. *J. Sens.* **2017**, *2017*, 6091261. [[CrossRef](#)]
24. Zhao, K.; Li, B.-H.; Dempster, A.G. A new approach of real time step length estimation for waist mounted PDR system. In Proceedings of the International Conference on Wireless Communication and Sensor Network, Wuhan, China, 13–14 December 2014; pp. 400–406.
25. Sun, Y.; Wu, H.; Schiller, J. A step length estimation model for position tracking. In Proceedings of the International Conference on Location and GNSS (ICL-GNSS), Gothenburg, Sweden, 22–24 June 2015; pp. 1–6.
26. Shiau, J.-K.; Wang, I.-C. Unscented kalman filtering for attitude determination using mems sensors. *J. Appl. Sci. Eng.* **2013**, *16*, 165–176.
27. Nguyen, P.; Akiyama, T.; Ohashi, H.; Nakahara, G.; Yamasaki, K.; Hikaru, S. User-friendly heading estimation for arbitrary smartphone orientations. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcalá de Henares, Spain, 4–7 October 2016; pp. 1–7.
28. Liu, D.; Pei, L.; Qian, J.; Wang, L.; Liu, P.; Dong, Z.; Xie, S.; Wei, W. A novel heading estimation algorithm for pedestrian using a smartphone without attitude constraints. In Proceedings of the International Conference on Ubiquitous Positioning, Indoor Navigation and Location Based Services (UPINLBS), Shanghai, China, 2–4 November 2016; pp. 29–37.
29. Jin, Y.; Toh, H.-S.; Soh, W.-S.; Wong, W.-C. A robust dead-reckoning pedestrian tracking system with low cost sensors. In Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom), Seattle, WA, USA, 21–25 March 2011; pp. 222–230.
30. Shin, B.; Lee, J.H.; Lee, H.; Kim, E.; Kim, J.; Lee, S.; Cho, Y.-S.; Park, S.; Lee, T. Indoor 3D pedestrian tracking algorithm based on PDR using smartphone. In Proceedings of the International Conference on Control, Automation and Systems, Jeju Island, Korea, 17–21 October 2012; pp. 1442–1445.
31. Ali, A.; El-Sheimy, N. Low-cost MEMS-based pedestrian navigation technique for GPS-denied areas. *J. Sens.* **2013**, *2013*, 197090. [[CrossRef](#)]
32. Fourati, H.; Manamanni, N.; Afilal, L.; Handrich, Y. Position estimation approach by complementary filter-aided IMU for indoor environment. In Proceedings of the European Control Conference (ECC), Zurich, Switzerland, 17–19 July 2013; pp. 4208–4213.
33. Kakiuchi, N.; Kamijo, S. Pedestrian dead reckoning for mobile phones through walking and running mode recognition. In Proceedings of the International Conference on Intelligent Transportation Systems (ITSC 2013), The Hague, The Netherlands, 6–9 October 2013; pp. 261–267.
34. Kang, W.; Han, Y. SmartPDR: Smartphone-based pedestrian dead reckoning for indoor localization. *IEEE Sens. J.* **2014**, *15*, 2906–2916. [[CrossRef](#)]
35. Tian, Q.; Salcic, Z.; Kevin, I.; Wang, K.; Pan, Y. A multi-mode dead reckoning system for pedestrian tracking using smartphones. *IEEE Sens. J.* **2015**, *16*, 2079–2093. [[CrossRef](#)]
36. Shin, B.; Kim, C.; Kim, J.; Lee, S.; Kee, C.; Kim, H.S.; Lee, T. Motion recognition-based 3D pedestrian navigation system using smartphone. *IEEE Sens. J.* **2016**, *16*, 6977–6989. [[CrossRef](#)]

37. Ilkovičová, L.; Kajánek, P.; Kopáček, A. Pedestrian indoor positioning and tracking using smartphone sensors, step detection and map matching algorithm. *Geod. List* **2016**, *1*, 1–24.
38. Zhang, P.; Chen, X.; Ma, X.; Wu, Y.; Jiang, H.; Fang, D.; Tang, Z.; Ma, Y. SmartMTra: Robust indoor trajectory tracing using smartphones. *IEEE Sens. J.* **2017**, *16*, 3613–3624. [[CrossRef](#)]
39. Kuang, J.; Niu, X.; Zhang, P.; Chen, X. Indoor positioning based on pedestrian dead reckoning and magnetic field matching for smartphones. *Sensors* **2018**, *18*, 4142. [[CrossRef](#)]
40. Zhang, R.; Bannoura, A.; Höflinger, F.; Reindl, L.M.; Schindelbauer, C. Indoor localization using a smart phone. In Proceedings of the Sensors Applications Symposium, Galveston, TX, USA, 19–21 February 2013; pp. 38–42.
41. Chen, Z.; Zhu, Q.; Jiang, H.; Soh, Y.C. Indoor localization using smartphone sensors and iBeacons. In Proceedings of the Conference on Industrial Electronics and Applications (ICIEA), Auckland, New Zealand, 15–17 June 2015; pp. 1723–1728.
42. Liu, Y.; Dashti, M.; Rahman, M.A.A.; Zhang, J. Indoor localization using smartphone inertial sensors. In Proceedings of the Workshop on Positioning, Navigation and Communication (WPNC), Dresden, Germany, 12–13 March 2014; pp. 1–6.
43. Shi, L.-F.; Zhao, Y.-L.; Liu, G.-X.; Chen, S.; Wang, Y.; Shi, Y.-F. A Robust Pedestrian Dead Reckoning System Using Low-Cost Magnetic and Inertial Sensors. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 2996–3003. [[CrossRef](#)]
44. Gu, F.; Khoshelham, K.; Yu, C.; Shang, J. Accurate Step Length Estimation for Pedestrian Dead Reckoning Localization Using Stacked Autoencoders. *IEEE Trans. Instrum. Meas.* **2018**, *68*, 2705–2713. [[CrossRef](#)]
45. Ju, H.; Lee, J.H.; Park, C.G. Pedestrian Dead Reckoning System Using Dual IMU to Consider Heel Strike Impact. In Proceedings of the International Conference on Control, Automation and Systems (ICCAS), Daegwallyeong, Korea, 17–20 October 2018; pp. 1307–1309.
46. Tong, X.; Su, Y.; Li, Z.; Si, C.; Han, G.; Ning, J.; Yang, F. A Double-step Unscented Kalman Filter and HMM-based Zero Velocity Update for Pedestrian Dead Reckoning Using MEMS Sensors. *IEEE Trans. Ind. Electron.* **2019**, *67*, 581–591. [[CrossRef](#)]
47. Cho, S.Y.; Park, C.G. Threshold-less Zero-Velocity Detection Algorithm for Pedestrian Dead Reckoning. In Proceedings of the European Navigation Conference (ENC), Warsaw, Poland, 9–12 April 2019; pp. 1–5.
48. Poulouse, A.; Eyobu, O.S.; Han, D.S. An Indoor Position-Estimation Algorithm Using Smartphone IMU Sensor Data. *IEEE Access* **2019**, *7*, 11165–11177. [[CrossRef](#)]
49. Fusco, G.; Coughlan, J.M. Indoor Localization Using Computer Vision and Visual-Inertial Odometry. In Proceedings of the International Conference on Computers Helping People with Special Needs (ICCHP), Linz, Austria, 11–3 July 2018; pp. 86–93.
50. Cooper, A.; Hegde, P. An indoor positioning system facilitated by computer vision. In Proceedings of the MIT Undergraduate Research Technology Conference (URTC), Cambridge, MA, USA, 4–6 November 2016; pp. 1–5.
51. Xiao, A.; Chen, R.; Li, D.; Chen, Y.; Wu, D. An Indoor Positioning System Based on Static Objects in Large Indoor Scenes by Using Smartphone Cameras. *Sensors* **2018**, *18*, 2229. [[CrossRef](#)]
52. Kittenberger, T.; Ferner, A.; Scheickl, R. A simple computer vision based indoor positioning system for educational micro air vehicles. *J. Autom. Mob. Robot. Intell. Syst.* **2014**, *8*, 46–52. [[CrossRef](#)]
53. Carozza, L.; Tingdahl, D.; Bosché, F.; Van Gool, L. Markerless vision-based augmented reality for urban planning. *Comput.-Aided Civ. Infrastruct. Eng.* **2014**, *29*, 2–17. [[CrossRef](#)]
54. Comport, A.I.; Marchand, E.; Pressigout, M.; Chaumette, F. Real-time markerless tracking for augmented reality: The virtual visual servoing framework. *IEEE Trans. Vis. Comput. Graph.* **2006**, *12*, 615–628. [[CrossRef](#)] [[PubMed](#)]
55. Goedemé, T.; Nuttin, M.; Tuytelaars, T.; Van Gool, L. Markerless computer vision based localization using automatically generated topological maps. In Proceedings of the European Navigation Conference GNSS, Rotterdam, The Netherlands, 16–19 May 2004; pp. 219–236.
56. Comport, A.I.; Marchand, É.; Chaumette, F. A real-time tracker for markerless augmented reality. In Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality, Tokyo, Japan, 10 October 2003; pp. 36–45.
57. Colyer, S.L.; Evans, M.; Cosker, D.P.; Salo, A.I. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports Med.-Open* **2018**, *4*, 24. [[CrossRef](#)] [[PubMed](#)]

58. Mutka, A.; Miklic, D.; Draganjac, I.; Bogdan, S. A low cost vision based localization system using fiducial markers. *IFAC Proc. Vol.* **2008**, *41*, 9528–9533. [[CrossRef](#)]
59. Li, Y.; Zhu, S.; Yu, Y.; Wang, Z. An improved graph-based visual localization system for indoor mobile robot using newly designed markers. *Int. J. Adv. Robot. Syst.* **2018**, *15*. [[CrossRef](#)]
60. Khairuddin, A.R.; Talib, M.S.; Haron, H. Review on simultaneous localization and mapping (SLAM). In Proceedings of the International Conference on Control System, Computing and Engineering (ICCSCE), George Town, Malaysia, 27–29 November 2015; pp. 85–90.
61. Kuzmin, M. Review, Classification and Comparison of the Existing SLAM Methods for Groups of Robots. In Proceedings of the 22st Conference of Open Innovations Association (FRUCT), Jyvaskyla, Finland, 15–18 May 2018; pp. 115–120.
62. Saeedi, S.; Trentini, M.; Seto, M.; Li, H. Multiple-robot simultaneous localization and mapping: A review. *J. Field Robot.* **2016**, *33*, 3–46. [[CrossRef](#)]
63. Cadena, C.; Carlone, L.; Carrillo, H.; Latif, Y.; Scaramuzza, D.; Neira, J.; Reid, I.; Leonard, J.J. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans. Robot.* **2016**, *32*, 1309–1332. [[CrossRef](#)]
64. Taketomi, T.; Uchiyama, H.; Ikeda, S. Visual SLAM algorithms: A survey from 2010 to 2016. *IPSJ Trans Comput. Vis. Appl.* **2017**, *9*, 16. [[CrossRef](#)]
65. Hwang, S.-Y.; Song, J.-B. Clustering and probabilistic matching of arbitrarily shaped ceiling features for monocular vision-based SLAM. *Adv. Robot.* **2013**, *27*, 739–747. [[CrossRef](#)]
66. Bailey, T.; Durrant-Whyte, H. Simultaneous localization and mapping (SLAM): Part II. *IEEE Robot. Autom. Mag.* **2006**, *13*, 108–117. [[CrossRef](#)]
67. Montemerlo, M.; Thrun, S. Simultaneous localization and mapping with unknown data association using FastSLAM. In Proceedings of the International Conference on Robotics and Automation (Cat. No. 03CH37422), Aipei, Taiwan, 14–19 September 2003; pp. 1985–1991.
68. Zikos, N.; Petridis, V. L-SLAM: Reduced dimensionality FastSLAM with unknown data association. In Proceedings of the International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 4074–4079.
69. Doaa, M.-L.; Mohammed, A.; Megeed Salem, H.R.; Mohamed, I.R. 3D Graph-based Vision-SLAM Registration and Optimization. *Int. J. Circuits Syst. Signal Process.* **2014**, *8*, 123–130.
70. Gil, A.; Juliá, M.; Reinoso, Ó. Occupancy grid based graph-SLAM using the distance transform, SURF features and SGD. *Eng. Appl. Artif. Intell.* **2015**, *40*, 1–10. [[CrossRef](#)]
71. Ozisik, O.; Yavuz, S. An occupancy grid based SLAM method. In Proceedings of the International Conference on Computational Intelligence for Measurement Systems and Applications, Istanbul, Turkey, 14–16 July 2008; pp. 117–119.
72. Asmar, D.; Shaker, S. 2 D occupancy-grid SLAM of structured indoor environments using a single camera. *Int. J. Mechatron. Autom.* **2012**, *2*, 112–124. [[CrossRef](#)]
73. Eliazar, A.; Parr, R. DP-SLAM: Fast, robust simultaneous localization and mapping without predetermined landmarks. In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), Acapulco, Mexico, 9–15 August 2003; pp. 1135–1142.
74. Klein, G.; Murray, D. Parallel tracking and mapping for small AR workspaces. In Proceedings of the ACM International Symposium on Mixed and Augmented Reality, Nara, Japan, 13–16 November 2007; pp. 1–10.
75. Pire, T.; Fischer, T.; Castro, G.; De Cristóforis, P.; Civera, J.; Berles, J.J. S-PTAM: Stereo parallel tracking and mapping. *Robot. Auton. Syst.* **2017**, *93*, 27–42. [[CrossRef](#)]
76. Newcombe, R.A.; Lovegrove, S.J.; Davison, A.J. DTAM: Dense tracking and mapping in real-time. In Proceedings of the International Conference on Computer Vision, Barcelona, Spain, 6–3 November 2011; pp. 2320–2327.
77. Gupta, T.; Shin, D.; Sivagnanasadan, N.; Hoiem, D. 3dfs: Deformable dense depth fusion and segmentation for object reconstruction from a handheld camera. *arXiv* **2016**, arXiv:1606.05002.
78. Kaess, M.; Ranganathan, A.; Dellaert, F. iSAM: Incremental smoothing and mapping. *IEEE Trans. Robot.* **2008**, *24*, 1365–1378. [[CrossRef](#)]
79. Engel, J.; Schöps, T.; Cremers, D. LSD-SLAM: Large-scale direct monocular SLA. In Proceedings of the European Conference on Computer Vision (ECCV), Zürich, Switzerland, 6–12 September 2014; pp. 834–849.

80. Davison, A.J.; Reid, I.D.; Molton, N.D.; Stasse, O. MonoSLAM: Real-time single camera SLAM. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 1052–1067. [[CrossRef](#)]
81. Zou, D.; Tan, P. Coslam: Collaborative visual slam in dynamic environments. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *35*, 354–366. [[CrossRef](#)]
82. Milford, M.J.; Wyeth, G.F. SeqSLAM: Visual route-based navigation for sunny summer days and stormy winter nights. In Proceedings of the International Conference on Robotics and Automation, Saint Paul, MN, USA, 14–18 May 2012; pp. 1643–1649.
83. Bosse, M.; Zlot, R. Continuous 3D scan-matching with a spinning 2D laser. In Proceedings of the International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; pp. 4312–4319.
84. Shen, X.; Min, H.; Lin, Y. Fast RGBD-ICP with bionic vision depth perception model. In Proceedings of the International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), Shenyang, China, 8–12 June 2015; pp. 1241–1246.
85. Mur-Artal, R.; Tardós, J.D. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [[CrossRef](#)]
86. Andersson, M.; Baerveldt, M. Simultaneous Localization and Mapping for Vehicles Using ORB-SLAM2. Master's Thesis, Chalmers University of Technology, Gothenburg, Sweden, May 2018.
87. Mur-Artal, R.; Tardós, J.D. Probabilistic Semi-Dense Mapping from Highly Accurate Feature-Based Monocular SLAM. In Proceedings of the Robotics: Science and Systems, Rome, Italy, 13–17 July 2015.
88. Bacik, J.; Durovsky, F.; Fedor, P.; Perdukova, D. Autonomous flying with quadcopter using fuzzy control and ArUco markers. *Intell. Serv. Robot.* **2017**, *10*, 1861–2776. [[CrossRef](#)]
89. Robot Operating System. Available online: <http://www.ros.org/> (accessed on 17 May 2019).
90. Sanchez-Lopez, J.L.; Arellano-Quintana, V.; Tognon, M.; Campoy, P.; Franchi, A. Visual Marker based Multi-Sensor Fusion State Estimation. In Proceedings of the 20th IFACWorld Congress, Toulouse, France, 13–16 July 2017; pp. 16003–16008.
91. Muoz-Salinas, R.; Medina-Carnicer, R.; Marín-Jimenez, M.J.; Yeguas-Bolivar, E. Mapping and Localization from Planar Markers. *Pattern Recognit.* **2017**, *73*, 158–171. [[CrossRef](#)]
92. You, S.; Neumann, U. Fusion of vision and gyro tracking for robust augmented reality registration. In Proceedings of the IEEE Virtual Reality 2001, Yokohama, Japan, 13–17 March 2001; pp. 71–78.
93. Tao, Y.; Hu, H.; Zhou, H. Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation. *Int. J. Robot. Res.* **2007**, *26*, 607–624. [[CrossRef](#)]
94. Hol, J.D.; Schön, T.B.; Gustafsson, F. Modeling and calibration of inertial and vision sensors. *Int. J. Robot. Res.* **2010**, *29*, 231–244. [[CrossRef](#)]
95. Park, J.; Hwang, W.; Kwon, H.-I.; Kim, J.-H.; Lee, C.-H.; Anjum, M.L.; Kim, K.-S.; Cho, D.-I. High performance vision tracking system for mobile robot using sensor data fusion with kalman filter. In Proceedings of the International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 3778–3783.
96. Aubeck, F.; Isert, C.; Gusenbauer, D. Camera based step detection on mobile phones. In Proceedings of the International Conference on Indoor Positioning and Indoor Navigation, Guimaraes, Portugal, 21–23 September 2011; pp. 1–7.
97. Kumar, K.; Varghese, A.; Reddy, P. K.; Narendra, N.; Swamy, P.; Chandra, M.G.; Balamuralidhar, P. An improved tracking using IMU and Vision fusion for Mobile Augmented Reality applications. *arXiv* **2014**, arXiv:1411.2335.
98. Erdem, A.T.; Ercan, A.Ö. Fusing inertial sensor data in an extended Kalman filter for 3D camera tracking. *IEEE Trans. Image Process.* **2014**, *24*, 538–548. [[CrossRef](#)]
99. Delaune, J.; Le Besnerais, G.; Voirin, T.; Farges, J.-L.; Bourdarias, C. Visual-inertial navigation for pinpoint planetary landing using scale-based landmark matching. *Robot. Auton. Syst.* **2016**, *78*, 63–82. [[CrossRef](#)]
100. Li, J.; Besada, J.A.; Bernardos, A.M.; Tarrío, P.; Casar, J.R. A novel system for object pose estimation using fused vision and inertial data. *Inf. Fusion* **2017**, *33*, 15–28. [[CrossRef](#)]
101. Chroust, S.G.; Vincze, M. Fusion of vision and inertial data for motion and structure estimation. *J. Robot. Syst.* **2004**, *21*, 73–83. [[CrossRef](#)]
102. Elloumi, W.; Latoui, A.; Canals, R.; Chetouani, A.; Treuillet, S. Indoor pedestrian localization with a smartphone: A comparison of inertial and vision-based methods. *IEEE Sens. J.* **2016**, *16*, 5376–5388. [[CrossRef](#)]

103. Alatise, M.; Hancke, G. Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter. *Sensors* **2017**, *17*, 2164. [[CrossRef](#)] [[PubMed](#)]
104. Liu, M.; Chen, R.; Li, D.; Chen, Y.; Guo, G.; Cao, Z.; Pan, Y. Scene recognition for indoor localization using a multi-sensor fusion approach. *Sensors* **2017**, *17*, 2847. [[CrossRef](#)] [[PubMed](#)]
105. Guo, X.; Sun, C.; Wang, P.; Huang, L. Vision sensor and dual MEMS gyroscope integrated system for attitude determination on moving base. *Rev. Sci. Instrum.* **2018**, *89*, 015002. [[CrossRef](#)]
106. Farnoosh, A.; Nabian, M.; Closas, P.; Ostadabbas, S. First-person indoor navigation via vision-inertial data fusion. In Proceedings of the IEEE/ION Position, Location and Navigation Symposium (PLANS), Monterey, CA, USA, 23–26 April 2018; pp. 1213–1222.
107. Yan, J.; He, G.; Basiri, A.; Hancock, C. 3D Passive-Vision-Aided Pedestrian Dead Reckoning for Indoor Positioning. *IEEE Trans. Instrum. Meas.* **2019**, doi:10.1109/TIM.2019.2910923. [[CrossRef](#)]
108. Huang, G. Visual-inertial navigation: A concise review. *arXiv* **2019**, arXiv:1906.02650.
109. Wu, Y.; Tang, F.; Li, H. Image-based camera localization: An overview. *Vis. Comput. Ind. Biomed. Art* **2018**, *1*, 1–13. [[CrossRef](#)]
110. Ligorio, G.; Sabatini, A. Extended Kalman filter-based methods for pose estimation using visual, inertial and magnetic sensors: Comparative analysis and performance evaluation. *Sensors* **2013**, *13*, 1919–1941. [[CrossRef](#)]
111. Tian, Y.; Zhang, J.; Tan, J. Adaptive-frame-rate monocular vision and imu fusion for robust indoor positioning. In Proceedings of the International Conference on Robotics and Automation, Karlsruhe, Germany, 6–10 May 2013; pp. 2257–2262.
112. Nilsson, J.; Fredriksson, J.; Ödöblom, A.C. Reliable vehicle pose estimation using vision and a single-track model. *IEEE Trans. Intell. Transp. Syst.* **2014**, *15*, 2630–2643. [[CrossRef](#)]
113. Fox, D.; Thrun, S.; Burgard, W.; Dellaert, F. *Particle Filters for Mobile Robot Localization*; Sequential Monte Carlo Methods in Practice; Springer Science & Business Media: New York, NY, USA, 2001; pp. 401–428.
114. Wan, E.A.; Van Der Merwe, R. The unscented Kalman filter for nonlinear estimation. In Proceedings of the Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373), Lake Louise, AB, Canada, 4 October 2000; pp. 153–158.
115. Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G.R. ORB: An efficient alternative to SIFT or SURF. In Proceedings of the International Conference on Computer Vision (ICCV), Barcelona, Spain, 6–13 November 2011; pp. 2564–2571.
116. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [[CrossRef](#)]
117. Bay, H.; Tuytelaars, T.; Van Gool, L. Surf: Speeded up robust features. In Proceedings of the European conference on computer vision, Graz, Austria, 7–13 May 2006; pp. 404–417.
118. Bay, H.; Ess, A.; Tuytelaars, T.; Van Gool, L. Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **2008**, *110*, 346–359. [[CrossRef](#)]
119. Munoz-Salinas, R.; Marin-Jimenez, M.J.; Medina-Carnicer, R. SPM-SLAM: Simultaneous localization and mapping with squared planar markers. *Pattern Recognit.* **2019**, *86*, 156–171. [[CrossRef](#)]
120. Gálvez-López, D.; Tardos, J.D. Bags of binary words for fast place recognition in image sequences. *IEEE Trans. Robot.* **2012**, *28*, 1188–1197. [[CrossRef](#)]
121. Xing, B.; Zhu, Q.; Pan, F.; Feng, X. Marker-Based Multi-Sensor Fusion Indoor Localization System for Micro Air Vehicles. *Sensors* **2018**, *18*, 1706. [[CrossRef](#)] [[PubMed](#)]

