

## Article

# Continuous Indoor Positioning Fusing WiFi, Smartphone Sensors and Landmarks

Zhi-An Deng <sup>1,\*</sup>, Guofeng Wang <sup>1</sup>, Danyang Qin <sup>2</sup>, Zhenyu Na <sup>1</sup>, Yang Cui <sup>3</sup> and Juan Chen <sup>1</sup>

<sup>1</sup> School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China; gfwangsh@163.com (G.W.); nazhenyu@dlmu.edu.cn (Z.N.); chenjuan@dlmu.edu.cn (J.C.)

<sup>2</sup> School of Electronics Engineering, Heilongjiang University, Harbin 150080, China; qindanyang@hlju.edu.cn

<sup>3</sup> School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China; ycui@hit.edu.cn

\* Correspondence: dengzhian@dlmu.edu.cn; Tel.: +86-411-8472-9223

Academic Editor: Xue-Bo Jin

Received: 14 May 2016; Accepted: 18 August 2016; Published: 5 September 2016

**Abstract:** To exploit the complementary strengths of WiFi positioning, pedestrian dead reckoning (PDR), and landmarks, we propose a novel fusion approach based on an extended Kalman filter (EKF). For WiFi positioning, unlike previous fusion approaches setting measurement noise parameters empirically, we deploy a kernel density estimation-based model to adaptively measure the related measurement noise statistics. Furthermore, a trusted area of WiFi positioning defined by fusion results of previous step and WiFi signal outlier detection are exploited to reduce computational cost and improve WiFi positioning accuracy. For PDR, we integrate a gyroscope, an accelerometer, and a magnetometer to determine the user heading based on another EKF model. To reduce accumulation error of PDR and enable continuous indoor positioning, not only the positioning results but also the heading estimations are recalibrated by indoor landmarks. Experimental results in a realistic indoor environment show that the proposed fusion approach achieves substantial positioning accuracy improvement than individual positioning approaches including PDR and WiFi positioning.

**Keywords:** indoor positioning; WiFi; inertial sensor; landmark; extended Kalman filter

## 1. Introduction

Indoor location-based service [1] has attracted much attention due to its great social and commercial values, which enable many applications including object finding, human tracking, personalized advertisement, and assisted living, etc. As Global Navigation Satellite Systems (GNSS) signal cannot penetrate well in indoor environment and provide reasonable accuracy performance, various other signals, such as WiFi [2], ultra-wideband [3], Bluetooth [4], radiofrequency identification (RFID) [5], ZigBee [6], and inertial sensors [7] have been developed for indoor positioning. Among them, WiFi positioning using received signal strength (RSS) fingerprinting [8] has been considered as one of the most popular indoor positioning solutions due to its low cost. RSS values from several access points (APs) can be easily gathered by common smartphones under existing WiFi infrastructure. However, severe RSS fluctuations [9] always render inaccurate positioning results.

Another popular indoor positioning solution is pedestrian dead reckoning (PDR) [10] system deploying inertial sensors embedded in a smartphone. PDR determines the user's location by adding the currently estimated displacement to previously estimated location. The displacement is achieved by combining step detection, step length estimation with user heading estimation using accelerometers, gyroscopes, and magnetometers. PDR using a smartphone is self-contained and requires no additional infrastructures. However, since PDR relies on low cost but noisy inertial sensors, the positioning error may accumulate along walking distances [11].

Both WiFi positioning and PDR approaches have disadvantages and complementary advantages. WiFi positioning approach can provide absolute positioning results to bind the cumulative error drifts of PDR. On the other hand, PDR achieves relatively more accurate positioning over short walking distance, which enables continuous tracking and smoothens the jumping effects of WiFi positioning caused by RSS fluctuations. Additionally, human motion recognition [12] using smartphone sensors may be used to improve indoor positioning. User motion states, like walking stairs, entering and walking out of elevators, and taking escalators, can be treated as indoor landmarks [13], to reset location estimation by PDR. Therefore, recent works tend to fuse WiFi positioning, PDR, and landmarks to enhance the indoor positioning accuracy.

Existing data fusion framework mainly includes particle filter [14–16], and Kalman filter [12,17]. The particle filter may achieve reasonable accuracy by deploying a large number of particles [14], but a large amount of computational cost is required and is not suitable for resource limited smartphones. The Kalman filter-based approaches are computational lightweight. However, an explicit measurement equation connecting user's position with RSS measurements is unavailable due to complex indoor radio propagation, thus rendering the measurement noise statistics unavailable. Previous Kalman filter-based fusion approaches manually and empirically set the related measurement noise covariance matrix. As a result, the fusion process cannot adapt the uncertainty of WiFi positioning results and, thus, rendering a degraded positioning accuracy.

In this paper, a novel extended Kalman filter (EKF)-based approach is deployed to fuse WiFi positioning, PDR, and landmarks. Instead of setting the measurement noise covariance matrix manually, we develop a kernel density estimation (KDE) based measurement model to accurately describe the RSS-position relationship and adaptively measure the related noise statistics. Furthermore, to reduce computational cost and avoid large WiFi positioning errors, we define a trusted area to limit the search space and propose a strategy for detection of outlier RSS measurements, respectively. For PDR, another EKF is designed for user heading estimation by fusing inertial sensors, magnetometers, and landmarks. Unlike previous approaches just deploying landmarks to reset positioning results of PDR, we also exploit the heading estimation information contained in landmarks to recalibrate user heading of PDR.

## 2. Related Works

For WiFi indoor positioning, the most widely used architecture is fingerprinting approach [18]. This approach consists of two phases: offline site survey and online pattern matching. During offline phase, RSS values of several APs are collected at different predefined calibration points to create a fingerprint database named radio map. Then, during online phase, RSS samples are collected in real-time and matched with RSS patterns pre-stored in radio map to derive the user's location. Due to the complementary advantages of WiFi positioning and PDR, fusing both approaches has been paid increasing attention.

Reference [19] firstly proposes a particle filter fusing inertial signals and WiFi positioning. New particles of user's location are generated according to the PDR model, once the current user step is identified, the step length and user heading are estimated. Then, the weight of a particle is updated by computing a Gaussian kernel function with variables of particle locations and related WiFi positioning results. Building map information is also exploited to remove the unexpected particles with unreasonable movements. Therefore, the user position is ultimately determined by averaging all weighted particles.

Similarly, [14] also proposes a particle filter fusing low-cost accelerometers, compasses, and WiFi RSS signals. The main difference is the particle weights update scheme, which is achieved by computing the inverse Euclidean distance between real-time observed RSS and mean RSS signal of a particle's nearest calibration point in the radio map. The particle filter may integrate WiFi positioning, human motion and building map information effectively by deploying a large number of particles. However, it is power consuming and requires extensive computational cost, because location estimation of every walking step requires certain operations on each particle, whose quantity may

reach more than several hundreds. Overall, the particle filter-based fusion approach is unsuitable for running on resource limited devices.

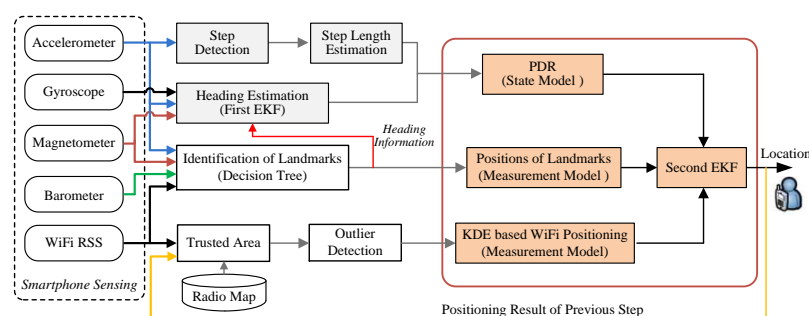
Recently, [12] proposes a light computation approach by deploying Kalman filter to fuse WiFi positioning, PDR, and landmarks. This approach formulates PDR as a linear function with a known user heading and step length. Consequently, inaccurate estimation of the error covariance of the state vector may be achieved due to the nonlinear nature of PDR versus the variables, including user heading and step length, thus rendering degraded positioning accuracy performance of the Kalman filter. For WiFi positioning, the user position is estimated by averaging the weighted coordinates of APs. The related weights are inversely proportional to the physical distances between smartphones and APs, which are computed by constructing and exploiting the RSS signal propagation model. Though a radio map is not required by WiFi positioning, due to the complexity and uncertainty of the indoor wireless propagation environment [18], the indoor signal propagation model may be so inaccurate that positioning results with large errors will be likely introduced to the Kalman filter. Furthermore, the measurement noise variance is set empirically, which cannot adapt the uncertainty of WiFi positioning results. More recently, Xin Li et al. [17] develop an EKF to fuse WiFi fingerprinting positioning with PDR. Similarly to [12], the measurement noise variance matrix is empirically set as a constant value, thus rendering degraded positioning accuracy of EKF.

We propose an EKF-based fusion approach with lightweight computation. For WiFi positioning, due to complex wireless propagation factors, such as non-line-of-sight propagation, it is difficult to develop an explicit measurement equation for WiFi RSS measurements [20]. Therefore, we develop a KDE-based [21,22] measurement model, which may adaptively measure uncertainty of WiFi positioning. To reduce the computational cost, we restrict the WiFi positioning search space into a trusted area defined by fusion position results of the previous step. Additionally, to avoid introducing large positioning errors into the EKF fusion process, we propose a detection algorithm to filter out WiFi signal outliers.

For heading estimation of PDR, rather than deploy Euler angles [23] and a direction cosine matrix (DCM) [24], we deploy a quaternion [7,25] because it can avoid the singularity problems and consume much less computational resources than DCM. Furthermore, to compensate the heading estimation drift, we design another EKF to fuse the three dimensional gyroscopes, accelerometers, and magnetometers. To reduce accumulation error of PDR, not only positioning results are reset by landmarks, but also the user heading estimation may be recalibrated, since the landmarks always indicate a specific region of user heading.

### 3. Methodology

Figure 1 overviews the proposed fusion approach for indoor positioning using smartphones. The proposed approach consists of two EKFs. The first one is used for heading estimation of PDR to integrate gyroscopes, accelerometers, magnetometers, and landmarks. The second one estimates the user's location by fusing WiFi positioning, PDR, and landmarks.



**Figure 1.** Overview of the proposed extended Kalman filter (EKF)-based fusion approach.

For the second EKF, the measurement model includes the KDE based WiFi positioning algorithm and landmarks for position resets. For WiFi positioning, we exploit the positioning results of previous step to adaptively restrict search space of WiFi positioning into a trusted area, and develop a strategy for detection of outlier RSS measurements. The state model is PDR system, which consists of step detection, step length estimation, and quaternion-based heading estimation modules. For heading estimation of PDR, it is also recalibrated by exploiting user heading information contained in landmarks. The landmarks can be identified by a decision tree using accelerometers, barometers, magnetometers, and WiFi RSS signals. For convenience, we provide a list of key symbols used in the proposed fusion approach in Table 1.

**Table 1.** Key symbols used in the proposed fusion approach.

Notations	Definitions
$\mathbf{l}_i$	Coordinate vector of the $i$ -th calibration point
$\mathbf{L}_i$	Coordinate vector of the $i$ -th walking step
$\hat{\mathbf{l}}_i$	Ultimate location estimation result at the $i$ -th step
$\mathbf{L}_i^r$	WiFi positioning result using RSS vector collected at the $i$ -th step
$\bar{\mathbf{r}}_i$	Mean RSS vector at the $i$ -th calibration point
$\wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}})$	Gaussian kernel function with mean vector $\bar{\mathbf{r}}_i$ and covariance matrix $\Sigma_{\mathbf{r}}$
$d$	Number of APs used in positioning
DCS/UCS/GCS	Device/User/Global Coordinate System
$SL_i$	Step length of the $i$ -th walking step
$\psi_{i\_step}$	Yaw angle (user heading) estimated at the $i$ -th step at GCS
$\mathbf{C}_{GCS}^{DCS}$	Direction cosine matrix between DCS and GCS
$\mathbf{q}_k$	Rotation quaternion vector at time instants $kT_s$

### 3.1. WiFi Positioning

#### 3.1.1. Kernel Density Estimation Based Measurement Update

This section describes the KDE-based WiFi positioning for constructing measurement model of the second EKF. Due to the complex propagation environment [26], such as non-line-of-sight propagation and human body absorptions, an explicit measurement equation relating user position to RSS measurements and related measurement noise statistics are unknown. Instead, fingerprinting approach collects RSS samples at a predefined set of calibration points to implicitly construct the RSS position relationship. Previous Kalman filter-based approaches always set the measurement noise covariance matrix empirically through a rough analysis of the positioning error. Consequently, an inaccurate estimation of measurement noise statistics will degrade the performance of Kalman filter.

Though measurement noise statistics of WiFi positioning is unavailable, we can approximately obtain it if we know the posterior density of the user's location with related RSS vector samples. Assume that  $\mathbf{r}$  is the RSS vector sample collected in real-time,  $\mathbf{l}$  is the user's two-dimensional coordinate vector, and the posterior density of the user's location is  $f(\mathbf{l}|\mathbf{r})$ . We give the location estimation and its covariance matrix based on Minimum Mean Squared Error (MMSE) [27] criteria:

$$\hat{\mathbf{l}} = \int \mathbf{l} f(\mathbf{l}|\mathbf{r}) d\mathbf{r} \quad (1)$$

$$\mathbf{P}_{\mathbf{r}} = \int (\mathbf{l} - \hat{\mathbf{l}}) (\mathbf{l} - \hat{\mathbf{l}})^T f(\mathbf{l}|\mathbf{r}) d\mathbf{r} \quad (2)$$

To obtain the posterior density, we deploy kernel density estimation (KDE), a nonparametric density estimator, due to its successful applications in probability density distributions estimation [21]. Given a set of RSS training pairs  $\{(\bar{\mathbf{r}}_i, \mathbf{l}_i) | i = 1, \dots, N\}$  stored in the radio map with mean RSS vector  $\bar{\mathbf{r}}_i$  at the calibration point  $\mathbf{l}_i$ , and deploying the widely used Gaussian kernel function  $\wp$ , the posterior density estimation can be rewritten as follows:

$$\hat{f}(\mathbf{l}|\mathbf{r}) = \frac{\hat{f}(\mathbf{l}, \mathbf{r})}{\hat{f}(\mathbf{r})} = \frac{\sum_{i=1}^N \wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}}) \wp(\mathbf{l}; \mathbf{l}_i, \Sigma_{\mathbf{l}})}{\sum_{i=1}^N \wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}})} = \sum_{i=1}^N \tilde{w}_i(\mathbf{r}) \wp(\mathbf{l}; \mathbf{l}_i, \Sigma_{\mathbf{l}}) \quad (3)$$

$$\tilde{w}_i(\mathbf{r}) = \frac{\wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}})}{\sum_{i=1}^N \wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}})} \quad (4)$$

$$\wp(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}}) = (2\pi)^{-d/2} |\Sigma_{\mathbf{r}}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{r} - \bar{\mathbf{r}}_i)^T \Sigma_{\mathbf{r}} (\mathbf{r} - \bar{\mathbf{r}}_i) \right\} \quad (5)$$

$$\wp(\mathbf{l}; \mathbf{l}_i, \Sigma_{\mathbf{l}}) = (2\pi)^{-1} |\Sigma_{\mathbf{l}}|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{l} - \mathbf{l}_i)^T \Sigma_{\mathbf{l}} (\mathbf{l} - \mathbf{l}_i) \right\} \quad (6)$$

where  $N$  is the number of calibration points at the target positioning environment,  $d$  is the number of APs used in RSS vector,  $\bar{\mathbf{r}}_i$  and  $\mathbf{l}_i$  are mean vector of the Gaussian kernel functions, and diagonal matrices  $\Sigma_{\mathbf{r}}$  and  $\Sigma_{\mathbf{l}}$  are the corresponding covariance matrices of the Gaussian kernel functions, respectively. To obtain the parameter  $\Sigma_{\mathbf{r}}$ , we use the estimation method given by Silverman [22],

$$\Sigma_{\mathbf{r}} = \sigma^* \mathbf{I}_d \quad (7)$$

$$\sigma^* = ((2d+1)n/4)^{-1/(d+4)} \tilde{\sigma} \quad (8)$$

where  $\mathbf{I}_d$  is  $d \times d$  identity matrix,  $\tilde{\sigma} = 1/d \sum_{i=1}^d \sigma_i^2$  is the averaged value of marginal variances for RSS values of each AP. Such an estimation of the kernel parameter has been shown to be effective in WiFi positioning [26] and can be also verified in our experimental results. For the parameter  $\Sigma_{\mathbf{l}}$  of Gaussian kernel, it represents the walking velocity of pedestrians and can be set as the classical pedestrian velocity values.

As seen in Equation (3), the posterior distribution can be considered as a Gaussian mixture of  $\wp(\mathbf{l}; \mathbf{l}_i, \Sigma_{\mathbf{l}})$ ,  $i = 1, \dots, N$  with the weighting value  $\tilde{w}_i(\mathbf{r})$  for each calibration point. Therefore, the MMSE location estimation and its error covariance using RSS measurements only may be determined by the first two moments [27] as follows:

$$\hat{\mathbf{l}} = \sum_{i=1}^N \tilde{w}_i(\mathbf{r}) \mathbf{l}_i \quad (9)$$

$$\mathbf{P}_{\mathbf{r}} = \sum_{i=1}^N \tilde{w}_i(\mathbf{r}) \left( \Sigma_{\mathbf{l}} + (\mathbf{l}_i - \hat{\mathbf{l}}) (\mathbf{l}_i - \hat{\mathbf{l}})^T \right) \quad (10)$$

Equations (9) and (10) can be used for WiFi positioning and the corresponding covariance matrix used in the measurement update step of the second EKF, respectively.

### 3.1.2. Trusted Area for Calibration Points Selection

For KDE-based WiFi positioning, the location estimation in Equation (9) is a weighted average of calibration points, where the weight is inversely proportional to the distance between the RSS vectors collected in real-time and the fingerprints pre-stored in the radio map. It may incur extensive computational cost if the real-time collected RSS vector is matched with the whole radio map consisting of a large number of calibration points. Furthermore, due to RSS fluctuations, some calibration points far away from the user's true position may receive relatively high weights and degrades the WiFi positioning accuracy.

To address these problems, we define a trusted area to restrict the WiFi positioning into a small set of the calibration points instead of the whole radio map. The basic idea for defining a trusted area is that the physical distance between adjacent walking steps is rather limited during pedestrian walking. Thus, we define the center of the trusted area by the positioning output of previous step and its size by location estimation covariance of the fusion approach. Define that  $\hat{\mathbf{L}}_{\text{Previous}}$  is the positioning result of the proposed approach of previous walking step,  $D_{\text{Th}}$  is the radius of the trusted area. The calibration points involved in KDE-based WiFi positioning for the current step can be selected:

$$S_{\text{Current}} = \{(\bar{\mathbf{r}}_i, \mathbf{l}_i) \mid \|\mathbf{l}_i - \hat{\mathbf{L}}_{\text{Previous}}\|_2 \leq D_{\text{Th}}; i = 1, \dots, N\} \quad (11)$$

where the radius of the trusted area  $D_{\text{Th}}$  is determined by location estimation covariance of the previous walking step, which may be set appropriately by controlling the probability that the error vector lies in the trusted area reaching 98.89% [28]. This 98.89% is just the probability that the variable lies within the three-sigma point if the variable is assumed to be Gaussian distributed. The cardinality of the set  $S_{\text{Current}}$  is denoted as  $N_{\text{Current}}$ , which is always much smaller than the total number of calibration points  $N$ . Therefore, the defined trusted area greatly reduces the number of calibration points involved and, thus, reduces the computational cost of WiFi positioning. This strategy also avoids deploying a clustering analysis technique [29,30], which is used in many previous works to reduce computational cost.

### 3.1.3. Outlier Detection of RSS Measurements

For EKF-based fusion approach, an erroneous location estimation caused by outlier of real-time RSS measurements can propagate to future location estimations, degrading positioning accuracy. To avoid introducing large positioning errors caused by RSS measurements outliers, we propose an outlier detection algorithm. We use the averaged weight of the calibration points in the trusted area as an indication of RSS reliability:

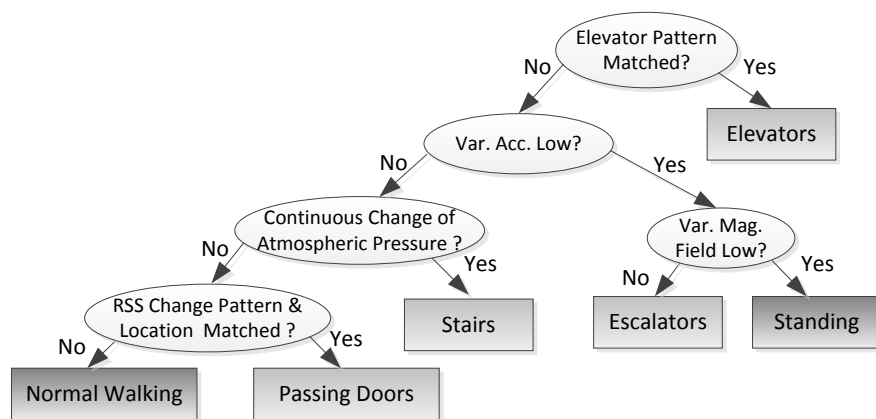
$$\eta(\mathbf{r}) = \frac{\sum_{i=1}^{N_{\text{Current}}} w_i(\mathbf{r})}{N_{\text{Current}}} = \frac{\sum_{i=1}^{N_{\text{Current}}} \phi(\mathbf{r}; \bar{\mathbf{r}}_i, \Sigma_{\mathbf{r}})}{N_{\text{Current}}} \quad (12)$$

where the averaged weight is bounded and  $0 < \eta(\mathbf{r}) < (2\pi)^{-d/2} |\Sigma_{\mathbf{r}}|^{-1/2}$  [21],  $d$  is the dimensionality of related RSS vectors. A small value of the averaged weight indicates that the real-time RSS measurement does not match most of the fingerprints in the trusted area and may, therefore, be considered as an outlier. The threshold value of the averaged weight may be set by statistical analysis of the averaged weights of normal RSS measurements during offline phase. Once an outlier is detected, the measurement update from WiFi positioning is omitted and a large positioning error introduced into the fusion process is avoided.

### 3.2. Identification of Landmarks

Previous works have shown that the landmarks may be used to reset location estimation of PDR. Particularly, landmarks may accurately restart location estimation of PDR once pedestrians reach these landmarks. The landmarks explored in this paper consist of elevators, escalators, stairs, and doors. When users walk into or out of elevators, take escalators, walk upstairs/downstairs, or pass through a door when entering or leaving a room, the user headings are usually restricted into a small specific range. Therefore, we further explore landmarks to recalibrate the heading estimation. We deploy a decision tree including four levels to detect these landmarks and separate them from normal walking and standing motion states, as shown in Figure 2. To exploit the landmarks, accurate mapping of the whole indoor environment, such as the size of the rooms and their relative locations, is not required, while knowing the locations of these landmarks is enough.





**Figure 2.** Identification of landmarks using a decision tree.

In the first level, we firstly identify the elevator landmarks based on their unique pattern of acceleration signals [13]. A typical process of taking an elevator consists of normal walking, followed by standing sometime to wait for the elevator, entering it, standing inside for a short period, a pair of hyper-gravity/hypo-gravity effect with positive/negative impulses of vertical acceleration, a stationary period between two impulses depending on the elevation variations of the elevator, and finally a walk-out. In order to exploit the user heading information, we also deploy the magnitude of the magnetic field to accurately capture the epoch of walking into or out of an elevator. A significant decrease or increase of the total magnitude of magnetic field occurs when a pedestrian walks into or out of an elevator.

In the second level, we exploit the acceleration variances to distinguish between taking escalators/standing and walking stairs/walking, since the walking stairs/walking states involve the higher locomotion intensity and, consequently, the larger acceleration variances than those of taking escalators/standing. In the third level, we exploit the variances of the magnetic field values to separate escalators from standing, since the magnetic field values vary rapidly with the motion of escalators.

To further distinguish between walking and walking stairs, we deploy a barometer, whose atmospheric pressure value decreases with the increasing elevation [12]. We compute the absolute difference values of neighboring samples, and deploy a sum of the values within a sliding window as a flag. If the flag exceeds a threshold, walking upstairs/downstairs are identified. To reduce the influence of the sampling noise, we generate one sample by averaging all raw samples collected within one second. Though the barometer may indicate the variations in altitude and help three-dimensional positioning, we just focus two-dimensional positioning in this paper, and only exploit the barometer for landmark identification. We will expand to three-dimensional positioning in our future work.

In the fourth level, we deploy the RSS change pattern and location of the pedestrian to identify the landmarks of doors. For RSS values, it is well known that passing through doors may render large changes of RSS values from multiple APs due to the present or absent of signal attenuations of walls. Furthermore, to avoid confusion between different doors and false positive events, we also compare the current pedestrian location with the true location of doors to determine whether there is a door nearby.

Note that uncertainties of heading estimation information obtained from different landmarks are different. For example, user heading of taking an escalator has a smaller variation than that of walking upstairs/downstairs. Practically, measurement noises of user headings and user positions from different landmarks may be set according to realistic environments. Through identification of landmarks based on the decision tree approach, the average classification accuracy of landmarks can reach as high as 99%. We collected 200 test samples for each motion state (landmark) and a total number of 1200 test samples in our indoor office environment, except for escalator samples collected

in a shopping mall environment. The confusion matrix for the identification of different landmarks is shown in Table 2. Table 2 shows that nearly all landmarks can be identified correctly, except for the negligible confusion between walking and stairs, walking and passing doors, due to the variations of atmospheric pressure and magnetic field values, respectively. Thus, we assume that all of the landmarks can be correctly identified in the proposed fusion approach.

**Table 2.** Confusion matrix for landmarks identification.

Motion States	Stairs	Elevators	Escalators	Walking	Standing	Doors
Stairs	0.995	0	0	0.005	0	0
Elevators	0	1.0	0	0	0	0
Escalators	0	0	1.0	0	0	0
Walking	0	0	0	1.0	0	0
Standing	0	0	0	0	1.0	0
Doors	0	0	0	0.02	0	0.98

### 3.3. Pedestrian Dead Reckoning (PDR)

PDR determines the position of current step by adding displacement to that of a previous step,

$$\mathbf{L}_i = \mathbf{L}_{i-1} + SL_i \begin{bmatrix} \cos(\tilde{\psi}_i) \\ \sin(\tilde{\psi}_i) \end{bmatrix} \quad (13)$$

where  $\mathbf{L}_i = (x_i, y_i)$  and  $\mathbf{L}_{i-1} = (x_{i-1}, y_{i-1})$  are the two-dimensional coordinate vectors of current (the  $i$ -th) and the related previous steps,  $SL_i$  is the current step length, and  $\tilde{\psi}_i$  is the user heading. In this paper, the initial position and user heading of PDR are assumed to be known, which can be obtained by various methods [31,32], such as the user's input and the landmarks. For example, when a pedestrian walks out of an elevator, we can reset the position and user heading of PDR.

#### 3.3.1. Step Detection and Step Length Estimation

We deploy the widely used windowed peak detection algorithm [7] to recognize each walking step. The step detection relies on periodic acceleration fluctuation patterns, whose peak points associate with the heel strike during walking. The total magnitude of acceleration vector is used as an input of the peak detection algorithm, which is insensitive to device orientations and more robust than the widely used vertical acceleration values. Furthermore, coordinate transformation process for the acceleration signal is also not required during the pedestrian step detection. To reduce the signal noise, we average the raw acceleration vector sample by a centered sliding window. Then, the peak detection algorithm is applied to detect every distinct walking step. To avoid the false peaks caused by acceleration jitters, we use two thresholds to limit the magnitude of peaks and the time interval between neighboring peaks:

$$\begin{cases} |AccNorm - g| \geq A_{Th} \\ \Delta T \geq T_{Th} \end{cases} \quad (14)$$

where  $AccNorm$  is the total magnitude of acceleration vector,  $g$  is the local gravity value,  $A_{Th}$  represents the magnitude threshold of peaks,  $\Delta T$  is the time interval between neighboring peaks,  $T_{Th}$  represents the shortest time interval allowed between neighboring steps. In fact, we set the time threshold around two-thirds of the duration of one step. For the false peak of a small magnitude, it will be filtered out by the magnitude threshold. For the false peak with magnitude exceeding the magnitude threshold, it can also be effectively filtered out, since the time interval between the false peak and the true one within the same step will be much shorter than the time threshold.

Various linear or nonlinear step length estimation models [33] have been developed. These models relate estimated step length to variables such as step frequency and pedestrian height. In fact,



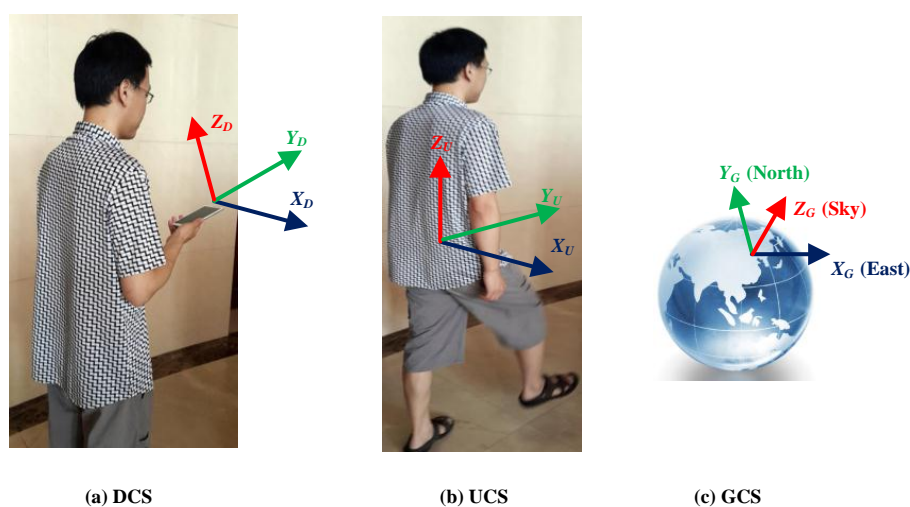
it is not an easy task to construct a generic step length estimation model to adapt different users with different environments. We deploy the following linear estimation model:

$$\text{StepLength } SL = \alpha \cdot fre + \beta \cdot var + \gamma \quad (15)$$

where  $fre$  is step frequency;  $var$  is variance of accelerations within one step;  $\alpha$  and  $\beta$  are the related coefficients;  $\gamma$  is the step constant. These parameters can be accurately estimated by offline training.

### 3.3.2. First EKF: User Heading Estimation Fusing Inertial Sensors, Magnetometers, and Landmarks

The state model of the first EKF deploys quaternion to describe device orientation time-evolutions, since the singularity problems can be avoided [25]. The measurement model of the first EKF includes an accelerometer for calibrating pitch and roll angles of the device, a magnetometer and landmarks for calibrating the yaw angle of the device. For user heading estimation, we define three coordinate systems, including device coordinate system (DCS), user coordinate system (UCS), and global coordinate system (GCS), as seen in Figure 3. As seen in Figure 3a, DCS is defined by axes  $X_{DCS}$ ,  $Y_{DCS}$ , and  $Z_{DCS}$ .  $X_{DCS}$  and  $Y_{DCS}$  axes point to right and forward, respectively, and both of them are parallel to the front face of the device screen; the  $Z_{DCS}$  axis is perpendicular to  $X_{DCS}$  and  $Y_{DCS}$  axes and points outside of screen. All sensor signals consisting of three-dimensional angular velocities, accelerations, and magnetic field are collected at DCS. As seen in Figure 3b, The UCS is defined by axes  $X_U$ ,  $Y_U$ , and  $Z_U$ , where  $Y_U$  is tangential to the walking trajectory,  $Z_U$  is the opposite direction of the gravity vector, and  $X_U$  is the cross product of  $Y_U$  and  $Z_U$ . As seen in Figure 3c, The GCS is the Earth coordinate system defined by axes  $X_G$ ,  $Y_G$ , and  $Z_G$ , which point east, north, and coincide with  $Z_U$ , respectively. User heading estimation can be considered as finding the relative orientation of UCS corresponding to the GCS. The user heading angle is defined as the anticlockwise rotation angle from the real-north direction  $Y_G$  to walking direction  $Y_U$ . In this paper, as in many other works [12,17], we only consider one of the most common device carrying positions: hand-held position, as seen in Figure 3a. The horizontal component of device forward axis ( $Y_D$ ) coincides with walking direction  $Y_U$ . Therefore, user heading is determined if DCS is specified corresponding to GCS by deploying rotation quaternion. Besides, for smartphone placed in the other positions, such as in the trouser pocket, principal component analysis can be used to extract the walking direction. Reference [34], which is our previous preliminary work in that research direction, is recommended.



**Figure 3.** Definitions of three coordinate systems: (a) device coordinate system (DCS); (b) user coordinate system (UCS); and (c) global coordinate system (GCS).

Firstly, we connect user heading with a rotation quaternion vector by transforming a  $3 \times 1$  column-vector from GCS to DCS as follows:

$$\mathbf{h}^{DCS}(t) = \mathbf{C}_{GCS}^{DCS}(\mathbf{q}(t)) \mathbf{h}^{GCS}(t) \quad (16)$$

where  $\mathbf{C}_{GCS}^{DCS}(\mathbf{q}(t))$  is the direction cosine matrix (DCM);  $\mathbf{h}^{DCS}(t)$  and  $\mathbf{h}^{GCS}(t)$  are the column-vector at DCS and GCS;  $\mathbf{q}(t) = [q_0 \ q_1 \ q_2 \ q_3]^T$  is the normalized orientation quaternion with the scalar part  $q_0$  and the vector part  $q_1, q_2, q_3$ . For simplicity, we omit the argument of time  $t$ . DCM can be described by the quaternion:

$$\mathbf{C}_{GCS}^{DCS}(\mathbf{q}) = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 2(q_1q_3 - q_0q_2) \\ 2(q_1q_2 - q_0q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_0q_1 + q_2q_3) \\ 2(q_1q_3 + q_0q_2) & 2(q_2q_3 - q_0q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \quad (17)$$

The yaw angle of a smartphone at GCS is given as a function of quaternion vector  $f(\mathbf{q})$ :

$$yaw = f(\mathbf{q}) = \begin{cases} \arcsin\left(2(q_0q_3 - q_1q_2)/\sqrt{1 - 4(q_0q_1 + q_2q_3)^2}\right), & q_0^2 - q_1^2 + q_2^2 - q_3^2 > 0 \\ \pi - \arcsin\left(2(q_0q_3 - q_1q_2)/\sqrt{1 - 4(q_0q_1 + q_2q_3)^2}\right), & \text{otherwise} \end{cases} \quad (18)$$

Secondly, according to the kinematic law of rigid body [25], the discrete time evolution model of orientation quaternions is:

$$\mathbf{q}_{k+1} = \exp(0.5 \times \Omega(\mathbf{w}_k T_s)) \mathbf{q}_k = (\mathbf{I}_4 \cos(0.5 \times \Delta\theta_k) + \Omega(\mathbf{w}_k T_s) \sin(0.5 \times \Delta\theta_k) / \Delta\theta_k) \mathbf{q}_k \quad (19)$$

where  $\mathbf{q}_k$  and  $\mathbf{q}_{k+1}$  are the quaternions at time instants  $kT_s$  and  $(k+1)T_s$ ,  $T_s$  is the system interval,  $\mathbf{w}_k = [w_k^x \ w_k^y \ w_k^z]^T$  is the angular velocity vector measured at DCS with time instants  $kT_s$ ,  $\mathbf{I}_4$  is  $4 \times 4$  identity matrix,  $\Delta\theta_k = T_s \sqrt{(w_k^x)^2 + (w_k^y)^2 + (w_k^z)^2}$ ,  $\Omega(\mathbf{w}_k T_s)$  is given as follows:

$$\Omega(\mathbf{w}_k T_s) = T_s \begin{bmatrix} 0 & -w_k^x & -w_k^y & -w_k^z \\ w_k^x & 0 & w_k^z & -w_k^y \\ w_k^y & -w_k^z & 0 & w_k^x \\ w_k^z & w_k^y & -w_k^x & 0 \end{bmatrix} \quad (20)$$

Therefore, the quaternion  $\mathbf{q}_{k+1}$  is determined, starting from initial condition  $\mathbf{q}_0$ , which can be easily calculated by initial known device orientation at GCS.

Finally, we give the state and measurement model of the first EKF for user heading estimation. We deploy the related device quaternions as the state vector and, thus, the state model can be given as:

$$\mathbf{q}_{k+1} = F_k \mathbf{q}_k + \mathbf{w}_k^q \quad (21)$$

where state transition matrix  $F_k = \exp(\Omega(\mathbf{w}_k T_s))$ , and:

$$\mathbf{w}_k^q = \Xi_k \mathbf{w}_k^{gyro} = -\frac{T_s}{2} \begin{bmatrix} [\mathbf{e}_k \times] + q_0^k \mathbf{I}_3 \\ -\mathbf{e}_k^T \end{bmatrix} \mathbf{w}_k^{gyro} \quad (22)$$

where  $\mathbf{q}_k = [q_0^k \ q_1^k \ q_2^k \ q_3^k]^T$  includes the scalar part  $q_0^k$  and the vector part  $\mathbf{e}_k = [q_1^k \ q_2^k \ q_3^k]^T$ ,  $\mathbf{I}_3$  is  $3 \times 3$  identity matrix,  $\mathbf{w}_k^{gyro}$  is the white Gaussian noise [25] of gyroscope outputs at time instants  $kT_s$ , and  $[\mathbf{e}_k \times]$  is a standard vector cross-product operator. To obtain the zero-bias offset and scaling factor of gyroscope, we deploy the widely used calibration method called zero attitude

update [35], which is implemented at the initial time of each experimental run. Equations (21) and (22) describing the time evolution of quaternion can be considered as a first order approximation of the exact model Equation (19). The measurement noise vector of gyroscope  $\mathbf{w}_k^{gyro}$  is assumed small enough after calibration process, thus a first order approximation always performs well. The process noise covariance matrix  $\mathbf{Q}_k$  can be given as:

$$\mathbf{Q}_k = \mathbf{\Xi}_k \mathbf{Q}_k^{gyro} \mathbf{\Xi}_k^T \quad (23)$$

The measurement model of the first EKF is generated by relating real-time measured acceleration vectors at DCS, real-time magnetic field vectors at DCS to the local gravity and magnetic field vectors at GCS using the DCM as seen in Equation (16). The user heading derived from landmarks are also added in the measurement model.

$$\mathbf{z}_{k+1} = \begin{bmatrix} \mathbf{a}_{k+1} \\ \mathbf{m}_{k+1} \\ yaw_{k+1} \end{bmatrix} = \phi(\mathbf{q}_{k+1}) + \mathbf{v}_{k+1} = \begin{bmatrix} \mathbf{C}_{GCS}^{DCS}(\mathbf{q}_{k+1}) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{GCS}^{DCS}(\mathbf{q}_{k+1}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{g} \\ \mathbf{h} \\ f(\mathbf{q}_{k+1}) \end{bmatrix} + \begin{bmatrix} \mathbf{v}_{k+1}^a \\ \mathbf{v}_{k+1}^m \\ v_{k+1}^{yaw} \end{bmatrix} \quad (24)$$

where  $\mathbf{a}_{k+1}$  and  $\mathbf{m}_{k+1}$  are the acceleration and magnetic field vectors of the smartphone measured at DCS,  $\mathbf{g}$  and  $\mathbf{h}$  are the local gravity and magnetic field vectors at GCS,  $\mathbf{v}_{k+1}^a$  and  $\mathbf{v}_{k+1}^m$  are the white Gaussian measurement noise vectors of the accelerometer and gyroscope;  $f(\mathbf{q}_{k+1})$  is the yaw angle at GCS derived from the quaternion vector as seen in Equation (18);  $yaw_{k+1}$  is the yaw angle derived from a landmark and  $v_{k+1}^{yaw}$  is the related measurement noise. The covariance matrix of the measurement model  $\mathbf{R}_{k+1}$  can be given as follows:

$$\mathbf{R}_{k+1} = \begin{bmatrix} \mathbf{R}_{k+1}^a & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{k+1}^m & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{R}_{k+1}^{yaw} \end{bmatrix} = \begin{bmatrix} {}^R\sigma_a^2 \mathbf{I}_3 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & {}^R\sigma_m^2 \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & {}^R\sigma_{yaw}^2 \end{bmatrix} \quad (25)$$

To filter out the acceleration perturbations and severely perturbed magnetic field values, we adaptively tune the parameters of measurement covariance matrix for observed signals:

$${}^R\sigma_a^2 = \begin{cases} \sigma_a^2, & \|\mathbf{a}_{k+1} - \mathbf{g}\|_2 < \varepsilon_{a1} \cap \text{var}(\|\mathbf{a}_{k+1-Na/2}\|_2 : \|\mathbf{a}_{k+1+Na/2}\|_2) < \varepsilon_{a2} \\ \infty, & \text{otherwise} \end{cases} \quad (26)$$

$${}^R\sigma_m^2 = \begin{cases} \sigma_m^2, & \|\mathbf{m}_{k+1} - \mathbf{h}\|_2 < \varepsilon_{m1} \cap \text{var}(\|\mathbf{m}_{k+1-Nm/2}\|_2 : \|\mathbf{m}_{k+1+Nm/2}\|_2) < \varepsilon_{m2} \\ \infty, & \text{otherwise} \end{cases} \quad (27)$$

where  $\varepsilon_{a1}$  and  $\varepsilon_{m1}$  represent the allowed maximum deviations of acceleration vector and magnetic field vector from the local gravity vector and the local magnetic field vector, respectively;  $\varepsilon_{a2}$  and  $\varepsilon_{m2}$  represent the corresponding allowed maximum variances, respectively;  $N_a$  and  $N_m$  define the number of acceleration and magnetic field samples involved in computing the variances of acceleration and magnetic field signals, respectively;  $\text{var}(\cdot)$  is the variance of windowed signals. For observed yaw angle derived from a landmark, the measurement noise variance  ${}^R\sigma_{yaw}^2$  is set during offline phase according to the style of a landmark and related realistic environments, as described in Section 3.2.

To obtain the linearized version of (24), the EKF approach linearizes the nonlinear function  $\phi(\cdot)$  of Equation (24) by computing the related Jacobian matrix:

$$\mathbf{H}_{k+1} = \left. \frac{\partial}{\partial \mathbf{q}_{k+1}} \phi(\mathbf{q}_{k+1}) \right|_{\mathbf{q}_{k+1}=\mathbf{q}_{k+1}^-} \quad (28)$$

where the true values of state  $\mathbf{q}_{k+1}$  are unavailable. We replace  $\mathbf{q}_{k+1}$  by the best estimation available, called the a priori state estimation  $\mathbf{q}_{k+1}^-$ .

Based on the state model in Equation (21), the linearized version of measurement model in Equation (24), and related process and measurement noise covariance matrices, the a *posteriori* estimation of the quaternion  $\hat{\mathbf{q}}_{k+1}$  and the related covariance matrix  $\mathbf{P}_{\mathbf{q}}^{k+1}$  can be obtained by typical procedure of the EKF. We briefly summarize the detailed procedures as follows:

**Predicting:**

$$\mathbf{q}_{k+1}^- = F_k \hat{\mathbf{q}}_k \quad (29)$$

$$\mathbf{P}_{\mathbf{q}}^{k+1} = F_k \mathbf{P}_{\mathbf{q}}^k F_k^T + Q_k \quad (30)$$

**Updating:**

$$\mathbf{K}_{k+1} = \mathbf{P}_{\mathbf{q}}^{k+1} H_{k+1}^T \left( H_{k+1} \mathbf{P}_{\mathbf{q}}^{k+1} H_{k+1}^T + \mathbf{R}_{k+1} \right)^{-1} \quad (31)$$

$$\hat{\mathbf{q}}_{k+1} = \mathbf{q}_{k+1}^- + \mathbf{K}_{k+1} \left( \mathbf{z}_{k+1} - \phi \left( \mathbf{q}_{k+1}^- \right) \right) \quad (32)$$

$$\mathbf{P}_{\mathbf{q}}^{k+1} = \mathbf{P}_{\mathbf{q}}^{k+1} - \mathbf{K}_{k+1} H_{k+1} \mathbf{P}_{\mathbf{q}}^{k+1} \quad (33)$$

After obtaining  $\hat{\mathbf{q}}_{k+1}$ , the user heading can be ultimately determined as in Equation (18).

### 3.4. Second EKF: Fusing WiFi Positioning, PDR, and Landmarks

This section describes the second EKF based fusion approach, which exploits complementary strengths of WiFi positioning, PDR, and landmarks. We deploy KDE-based WiFi positioning and landmarks for measurement updates, and deploy PDR for the state model.

Firstly, we construct the state model, whose state vector is the user's two-dimensional coordinate vector. For the target indoor environment, a specific predefined environment coordinate system may be given according to building structures or personalized requirements. Assume  $\psi_{i\_step}$  is the yaw angle estimated at the  $i$ th step from Equation (18) of the first EKF at GCS, and  $\psi_{x\_axis}$  is the yaw angle of the positive  $x$ -axis of the environment coordinate system. Therefore, the anticlockwise rotation angle from the positive  $x$ -axis of a predefined environment coordinate system to the walking direction at the  $i$ th step, which is denoted as  $\tilde{\psi}_i$ , can be given as  $\tilde{\psi}_i = \psi_{i\_step} - \psi_{x\_axis}$ . Then, PDR model can be rewritten by substituting  $\tilde{\psi}_i = \psi_{i\_step} - \psi_{x\_axis}$  into Equation (13):

$$\mathbf{L}_i = \mathbf{L}_{i-1} + \tilde{f}(\mathbf{q}\mathbf{d}_i) = \mathbf{L}_{i-1} + SL_i \begin{bmatrix} \cos(\psi_{i\_step})\cos(\psi_{x\_axis}) + \sin(\psi_{i\_step})\sin(\psi_{x\_axis}) \\ \sin(\psi_{i\_step})\cos(\psi_{x\_axis}) - \sin(\psi_{x\_axis})\cos(\psi_{i\_step}) \end{bmatrix} \quad (34)$$

$$\begin{cases} \cos(\psi_{i\_step}) = \left( 1 - 2(q_1^{i\_step})^2 - 2(q_3^{i\_step})^2 \right) \left( 1 - 4(q_0^{i\_step}q_1^{i\_step} + q_2^{i\_step}q_3^{i\_step})^2 \right)^{-0.5} \\ \sin(\psi_{i\_step}) = \left( 2q_0^{i\_step}q_3^{i\_step} - 2q_1^{i\_step}q_2^{i\_step} \right) \left( 1 - 4(q_0^{i\_step}q_1^{i\_step} + q_2^{i\_step}q_3^{i\_step})^2 \right)^{-0.5} \end{cases} \quad (35)$$

where  $\tilde{f}(\cdot)$  is the relative displacement function of variable  $\mathbf{q}\mathbf{d}_i = [\mathbf{q}_{i\_step}^T SL_i]^T$ ,  $\cos(\psi_{i\_step})$  and  $\sin(\psi_{i\_step})$  are derived from Equation (18),  $\mathbf{q}_{i\_step} = [q_0^{i\_step} q_1^{i\_step} q_2^{i\_step} q_3^{i\_step}]^T$  is the rotation quaternion estimated by the first EKF at the acceleration peak point of the  $i$ th step.

Secondly, compute the a *priori* state estimation  $\mathbf{L}_i^-$  and related error covariance matrix  $\mathbf{P}_i^-$ :

$$\mathbf{L}_i^- = \hat{\mathbf{L}}_{i-1} + \tilde{f}(\mathbf{q}\mathbf{d}_i) \quad (36)$$

$$\mathbf{P}_i^- = \mathbf{P}_{i-1} + \tilde{F}_i \mathbf{P}_{\mathbf{q}\mathbf{d}_i} \tilde{F}_i^T \quad (37)$$

where  $\hat{\mathbf{L}}_{i-1}$  is the a *posteriori* location estimation at the  $(i-1)$  th step,  $\tilde{F}_i = \partial \tilde{f} / \partial \mathbf{q}\mathbf{d}$  is the Jacobian matrix computed at  $\mathbf{q}\mathbf{d} = \mathbf{q}\mathbf{d}_i$ ,  $\mathbf{P}_{i-1}$  is the a *posteriori* error covariance matrix at the  $(i-1)$  th step:

$$\mathbf{P}_{\mathbf{q}\mathbf{d}_i} = \begin{bmatrix} \mathbf{P}_{\mathbf{q}}^{i\_step} & \mathbf{0} \\ \mathbf{0} & \sigma_{SL}^2 \end{bmatrix} \quad (38)$$

where  $\mathbf{P}_{\mathbf{q}}^{i\_step}$  is the a *posteriori* error covariance matrix of  $\mathbf{q}_{i\_step}$ , which can be obtained from Equation (33) of the first EKF; and  $\sigma_{SL}^2$  is the error covariance of the step length estimation.

Thirdly, we combine KDE-based WiFi positioning results with observed landmarks to construct the measurement model:

$$\mathbf{L}_i^{observed} = \begin{bmatrix} \mathbf{L}_i^r \\ \mathbf{L}_{landmark} \end{bmatrix} = \mathbf{L}_i + \tilde{\mathbf{v}}_i = \mathbf{L}_i + \begin{bmatrix} \tilde{\mathbf{v}}_i^r \\ \tilde{\mathbf{v}}_i^{landmark} \end{bmatrix} \quad (39)$$

where  $\mathbf{L}_i^r$  is the KDE-based WiFi positioning result using RSS vector collected at the  $i$ th step  $\mathbf{r}_{i\_step}$ . Similar to Equations (9) and (10):

$$\mathbf{L}_i^r = \sum_{i=1}^{N_{i\_step}} \tilde{w}_i (\mathbf{r}_{i\_step}) \mathbf{l}_i \quad (40)$$

$$\tilde{\mathbf{R}}_i^r = \tilde{\mathbf{v}}_i^r (\tilde{\mathbf{v}}_i^r)^T = \sum_{i=1}^{N_{i\_step}} \tilde{w}_i (\mathbf{r}_{i\_step}) \left( \Sigma_1 + (\mathbf{l}_i - \mathbf{L}_i^r) (\mathbf{l}_i - \mathbf{L}_i^r)^T \right) \quad (41)$$

where  $\tilde{\mathbf{R}}_i^r$  is the measurement noise covariance matrix of WiFi positioning,  $(\tilde{\mathbf{r}}_i, \mathbf{l}_i) \in S_{i\_step}$  is the training pairs included in the trusted area of the  $i$ th step derived from Equation (11),  $N_{i\_step}$  is the number of calibration points defined by the related trusted area:

$$\tilde{\mathbf{R}}_i^{landmark} = \tilde{\mathbf{v}}_i^{landmark} (\tilde{\mathbf{v}}_i^{landmark})^T = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad (42)$$

$$\tilde{\mathbf{R}}_i = \tilde{\mathbf{v}}_i (\tilde{\mathbf{v}}_i)^T = \begin{bmatrix} \tilde{\mathbf{R}}_i^r & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{R}}_i^{landmark} \end{bmatrix} \quad (43)$$

where  $\tilde{\mathbf{R}}_i^{landmark}$  is the measurement noise covariance matrix of landmarks, which can be set according to the style of a landmark and related realistic environments.

Similarly to [28], the measurement noise of each landmark is set during the offline phase. Firstly, define an area covering all possible locations for a specific landmark. For example, when the pedestrian takes an elevator, the area inside the elevator includes all possible locations. Secondly, for simplicity, the measurement noise is assumed to follow two-dimensional Gaussian distribution. We choose a three-sigma ellipsoid to describe the covered area, where the pedestrian lying in the three-sigma ellipsoid can reach up as high as 98.89%. Thirdly, the center point of the defined area is set as the location of the landmark and the radius of the area is set to three-sigma. For example, the radius of an elevator is set as the half of the larger diagonal line of the rectangle. Finally, the standard deviation of the measurement noise is set as a third of the radius. The parameter setting may change according to different shapes of the landmarks and related environments. The influence of different settings will be studied in the future.

Finally, the a *posteriori* location estimation at the  $i$ th step  $\hat{\mathbf{L}}_i$  and covariance matrix  $\mathbf{P}_i$  are given:

$$\mathbf{K}_i = \mathbf{P}_i^- \left( \mathbf{P}_i^- + \tilde{\mathbf{R}}_i \right)^{-1} \quad (44)$$

$$\hat{\mathbf{L}}_i = \mathbf{L}_i^- + \mathbf{K}_i (\mathbf{L}_i^r - \mathbf{L}_i^-) \quad (45)$$

$$\mathbf{P}_i = \mathbf{P}_i^- - \mathbf{K}_i \mathbf{P}_i^- \quad (46)$$

The computational cost of the proposed EKF-based fusion approach is comparable with that of the Kalman filter-based approach [12], while it is much less than that of the particle filter-based [14] approach. In our experimental environments as seen in Section 4.1, we obtain the averaged calculation time per walking step of twenty traces by running the approaches in Matlab. The averaged calculation time per walking step of the proposed fusion approach takes 0.166 s, including 0.115 s for the first EKF and 0.051 s for the second EKF. For the particle filter based approach with 1000 particles, it takes 10.4 s and, thus, unsuitable for running on a resource-limited smartphone.

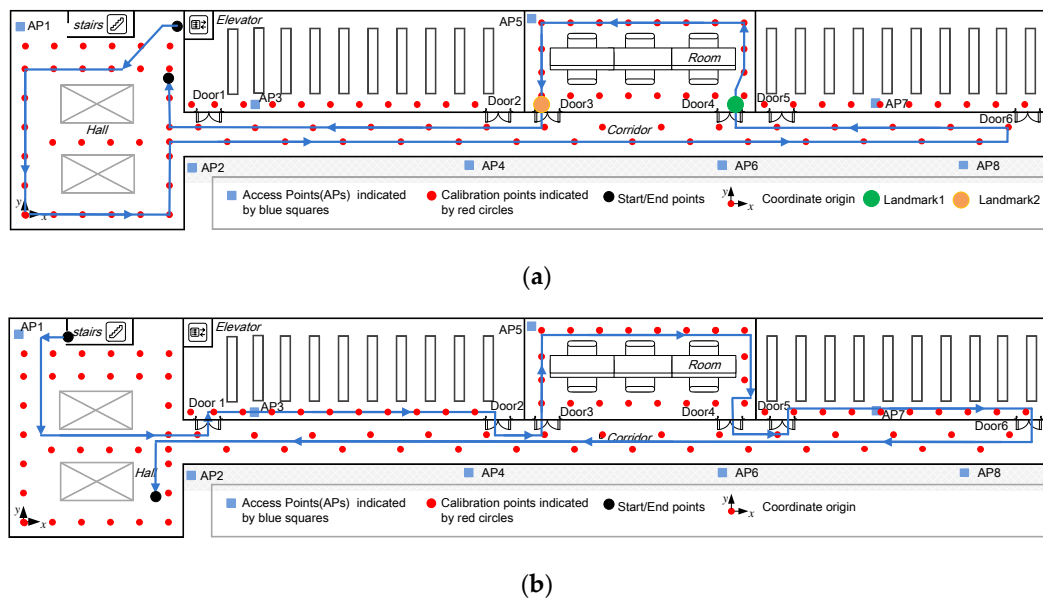
## 4. Evaluation

### 4.1. Experimental Setup

Experiments were performed in a typical indoor office area at the building of our department, whose size are  $43.5\text{ m} \times 11.2\text{ m}$ , as seen in Figure 4. We deploy a Samsung Galaxy S4 smartphone with the Android 4.2 operating system as the device involved in the experiments. The environment coordinate system is defined as seen in the lower-left of the hall, where the coordinate origin,  $x$ -axis, and  $y$ -axis are specified. For WiFi positioning, we construct the radio map by collecting RSS values at calibration points from eight APs distributed in the office environment during offline phase. All eight APs are hung in a relatively high place to reduce the signal attenuations caused by obstacles. The calibration points in the corridor are set two rows, while the distance between adjacent points along  $x$ -axis is the same as that in the hall. This setting is beneficial for improving the WiFi positioning accuracy in the corridor, especially along the  $y$ -axis, while requiring the same number of calibration points. Handheld device collected 100 RSS sample vectors per calibration point along four directions:  $x$ -axis,  $y$ -axis, and their opposite directions, each direction corresponding a quarter of the total number of samples. It should be noted that, the number of calibration points and their deployment may affect the WiFi positioning accuracy greatly. If the number of calibration points reduces or the walking path is far away from the calibration points, the WiFi positioning accuracy may deteriorate greatly [36] and even degrade the accuracy performance of the proposed fusion approach. During the online phase, real-time RSS vectors, gyroscope, accelerometer, and magnetometer data at a sampling rate of 2 Hz, 100 Hz, 50 Hz, and 10 Hz were collected, respectively.

The participant walks along the true path one and path two (indicated by the blue solid line) at least ten times, as seen in Figure 4a,b, respectively. The initial location and orientation of each experimental run are assumed to be known as *a priori* by exploiting the elevator and stair landmarks. At the beginning of each experimental run, the gyroscope is already well calibrated by the zero attitude update [35] method. We compare the proposed fusion approach with PDR using gyroscopes only, PDR combining gyroscopes with accelerometers, magnetometers, and landmarks, and WiFi positioning approaches. To validate the effectiveness of the KDE-based measurement model, we also compare the fusion approach whose measurement noise covariance matrix of WiFi positioning is set empirically [12], while the other modules of the fusion approach are the same as the proposed fusion approach. In our experiments, the measurement noise covariance matrix of WiFi positioning for the empirical fusion approach is set as a constant one, which makes the empirical fusion approach performs best for the testing path. To label the ground truth of the tracking trajectory and each walking step accurately, we exploit the square floor tiles and use a video to record the entire process of each walking path. Experiments for each path are carried out individually, and the positioning performance are the statistical results of the ten repeated tests for both two paths. The positioning error is defined as the Euclidean distance between estimated coordinate vectors and the true ones along the walking trajectory.

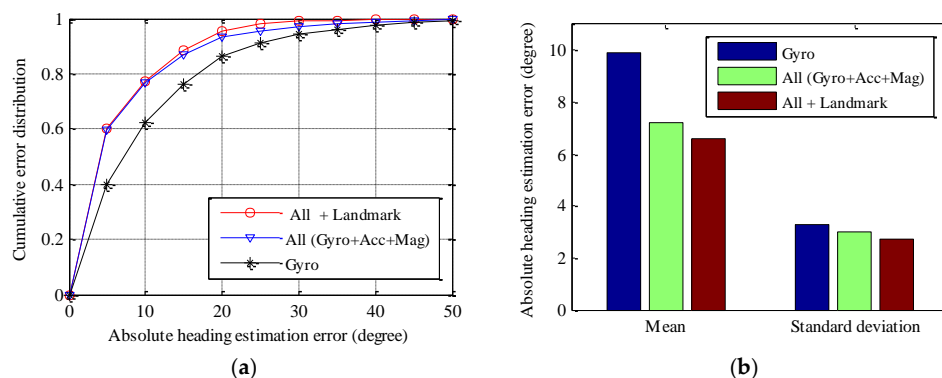




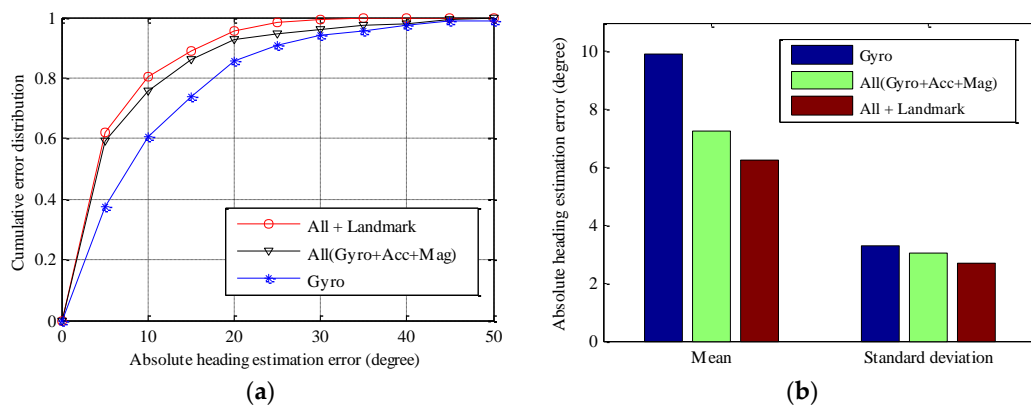
**Figure 4.** Experimental environment and the true walking path: (a) path one; (b) path two.

#### 4.2. Performance Analysis

Firstly, we evaluate the user heading estimation accuracy of various PDR-based approaches, since heading estimation bias is a main source of tracking error. Figure 5 compares user heading estimation accuracy of PDR using only a gyroscope (Gyro), using all sensors including gyroscopes, accelerometers, and magnetometers (All (Gyro + Acc + Mag)), and combining all sensors with landmarks (All + Landmark). As seen in Figure 5a, the All (Gyro + Acc + Mag) approach performs significantly better than that of the Gyro approach. This is because the former approach may fuse both acceleration and magnetic field signals, which contain tilt angle and yaw angle information of the device, respectively. For the All + Landmark approach, since the landmarks including passing through doors 2 and 3 are exploited to recalibrate the heading estimation, it may further improve heading estimation performance. As a result, the probability of large heading estimation errors reduces significantly. As seen in Figure 5b, mean absolute heading estimation error of the All + Landmark and the All (Gyro + Acc + Mag) approach reduce by 27.1% and 33.2% than that of the Gyro approach, respectively. Similarly, heading estimation accuracy improvement may also be obtained by combining all sensors with landmarks for the path two, as seen in Figure 6. The only difference is that, compared with All (Gyro + Acc + Mag) approach, the All + Landmark approach may obtain more significant heading estimation accuracy improvement, since more landmarks will be exploited to recalibrate the heading estimation.



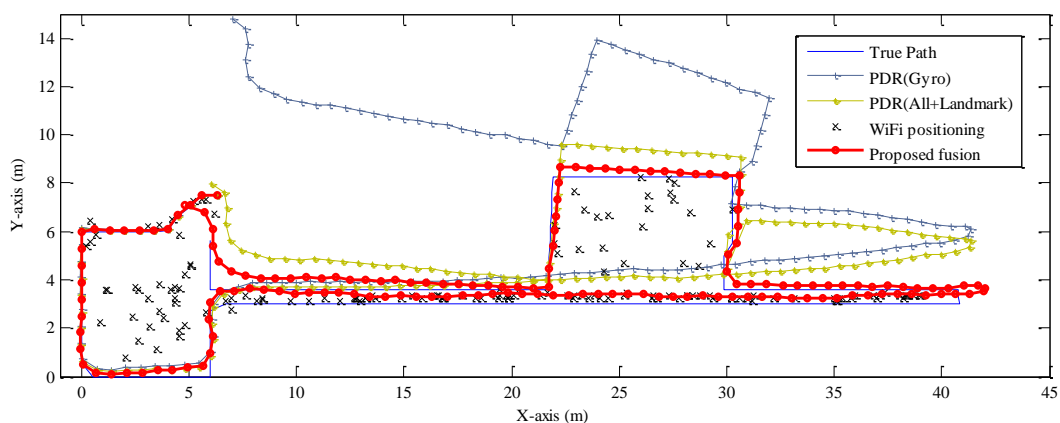
**Figure 5.** Performance comparisons of absolute heading estimation error of path one: (a) cumulative error distribution; and (b) mean and standard deviation.



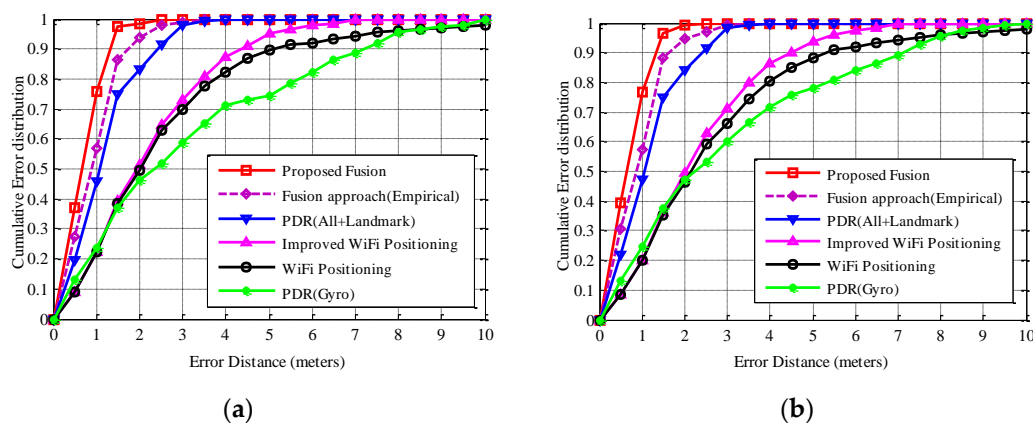
**Figure 6.** Performance comparisons of absolute heading estimation error of path two: (a) cumulative error distribution; and (b) mean and standard deviation.

Secondly, we compare pedestrian navigation performance along path one of the proposed fusion approach, PDR (Gyro), PDR (All + Landmark), and WiFi positioning. As seen in Figure 7, PDR-based approaches have comparable tracking accuracy with the proposed fusion approach along the initial walking path in the hall, while performing progressively worse with the heading estimation error accumulated. Tracking performance of PDR (All + Landmark) is significantly better than that of PDR (Gyro), since the heading estimation error of the former approach is smaller than that of the latter one. Particularly, when a pedestrian passes through Doors 4 and 3, the related landmark 1 and landmark 2 are identified and exploited to recalibrate the positioning and heading estimation results. Therefore, tracking errors of PDR (All + Landmark) reduce significantly immediately after passing through Doors 4 and 3.

Though individual WiFi positioning obtains relatively inaccurate tracking results along the whole walking path, the tracking error does not accumulate as PDR approach. By deploying the trusted area of WiFi positioning and detection of outlier RSS measurements, probability of large positioning error reduces significantly, as seen in Figure 8. WiFi positioning also provides useful localization information. For example, when the pedestrian walks along the corridor, the  $y$ -axis coordinates of the WiFi positioning results are always accurate and restricted into a small region, since the matched calibration points are almost all in the corridor. Moreover, by deploying the KDE-based model, uncertainty of WiFi positioning results is measured accurately, thus rendering WiFi positioning results with a larger error gain a smaller weight during the fusion process of EKF. Therefore, the proposed fusion approach performs best, due to an effective correction of PDR by fusing WiFi positioning and landmarks.



**Figure 7.** Comparisons between the true path one, trajectories of the proposed fusion approach, PDR (Gyro), PDR (All + Landmark), and WiFi positioning.



**Figure 8.** Accuracy comparisons between the proposed fusion approach, fusion approach (Empirical), PDR (All + Landmark), PDR (Gyro), WiFi positioning, and the improved WiFi positioning: (a) path one; (b) path two.

Finally, we compare positioning accuracy performance between the proposed fusion approach and the other approaches. Figure 8 shows the corresponding cumulative error distributions of compared approaches of both paths. Clearly, the proposed fusion approach obtains substantial accuracy improvement than compared individual approaches. As seen in Figure 8a for path one, particularly, accuracy within 1 m of the proposed fusion approach is 75.8%, while those of the fusion approach (empirical), PDR (All + Landmark), PDR (Gyro), and WiFi positioning are 56.9%, 45.8%, 23.9%, and 22.2%, respectively. Accuracy within 2 m of the proposed fusion approach is 98.2%, while those of the fusion approach (empirical), PDR (All + Landmark), PDR (Gyro), and WiFi positioning are 93.7%, 83.0%, 46.4%, and 49.7%, respectively. Accuracy within 6 m of the proposed improved WiFi positioning approach is 97.8%, while that of the original KDE-based WiFi positioning approach is only 92.2%. This is because most of large WiFi positioning errors are excluded by restricting the matched calibration points within a trusted area and identifying outlier RSS measurements.

Compared with the proposed fusion approach, the fusion approach (empirical) obtains significant accuracy performance degradation, because the measurement noise of WiFi positioning is set empirically and the uncertainty of WiFi positioning results cannot be handled. For example, when real-time RSS samples are corrupted in the complex propagation environment, the fusion approach (empirical) empirically setting the measurement noise may even bias the fusion results by giving WiFi positioning results of a large error a large weight during the fusion process. This case can be avoided effectively by the proposed fusion approach, whose measurement noise is adaptive tuned by the KDE-based method. PDR (Gyro) approach performs the worst, since the tracking error progressively increases with the accumulated heading estimation errors and positioning errors. Similar conclusions can also be derived from Figure 8b for path two. Additionally, all approaches involving PDR along path two perform slightly better than those of corresponding approaches along path one, since there are more landmarks exploited to calibrate the location estimation and heading estimation.

Table 3 shows the performance comparisons of various positioning error (m) measures. For simplicity, we give statistical results of two paths, since the performance of two paths is very close. Compared with the Fusion Approach (Empirical), PDR (All + Landmark), PDR (Gyro), WiFi positioning, and our proposed improved WiFi positioning, combining a trusted area with outlier detection, the proposed fusion approach reduces mean positioning errors by 25.0% (0.24 m), 41.0% (0.50 m), 76.9% (2.40 m), 76.1% (2.29 m), and 70.4% (1.71 m), while it reduces standard deviations by 24.1% (0.14 m), 40.5% (0.30 m), 83.3% (2.19 m), 83.6% (2.25 m), and 77.4% (1.51 m), respectively.

In summary, the proposed fusion approach obtains significant positioning accuracy improvement than various individual approaches and the fusion approach (empirical) for the following three reasons. Firstly, heading estimation performance of PDR is improved by fusing gyroscopes, accelerometers, and

magnetometers, and recalibrated by exploiting landmarks. Secondly, measurement model parameters of WiFi positioning are adaptively tuned by the KDE-based method and, thus, WiFi positioning results are fused to improve the positioning accuracy more effectively. Thirdly, various landmarks are identified by a decision tree approach and exploited to recalibrate the location estimation.

**Table 3.** Performance comparisons of various positioning error (m) measures.

Compared Approach	Proposed Fusion	Fusion Approach (Empirical)	PDR (All + Landmark)	PDR (Gyro)	WiFi Positioning	Improved WiFi Positioning
Mean error	<b>0.72</b>	0.96	1.22	3.12	3.01	2.43
Standard deviation	<b>0.44</b>	0.58	0.74	2.63	2.69	1.95
Medium error	<b>0.68</b>	0.89	1.08	2.30	2.12	2.03
75 percentile	<b>0.92</b>	1.32	1.55	4.66	3.44	3.21
90 percentile	<b>1.35</b>	1.73	2.10	7.05	5.26	4.42

## 5. Conclusions

In this paper, we propose an extended Kalman filter-based positioning approach by fusing WiFi, PDR, and landmarks. For WiFi positioning, instead of setting measurement noise parameters manually and empirically, we deploy a kernel density estimation model to measure it accurately and adaptively. Furthermore, by deploying a trusted area of WiFi positioning to restrict matched calibration points and detecting outlier RSS measurements, probability of WiFi positioning results with a large error distance reduces significantly. For user heading estimation of PDR, we develop another EKF model to integrate gyroscopes, accelerometers, magnetometers, and landmarks. Landmarks including elevators, escalators, stairs, and doors are identified by a decision tree approach. Then, they are exploited to recalibrate both user heading estimations and positioning results. Experimental results show that, compared with the fusion approach empirically setting measurement noise parameters of WiFi positioning, PDR using gyroscopes, accelerometers, magnetometers, and landmarks, WiFi positioning and PDR using a gyroscope alone, the proposed fusion approach reduces mean positioning errors by 25.0% (0.24 m), 41.0% (0.50 m), 76.1% (2.29 m), and 76.9% (2.40 m), respectively.

In future works, we will further consider the situation when orientation and carrying position of a smartphone are both free, such as placed in a pocket, which has been paid few attentions. Additionally, the proposed fusion approach will be tested in a larger and more complicated indoor environment including multi-floors and escalators.

**Acknowledgments:** This research is supported by National Natural Science Foundation of China (Granted No. 61301132, 61300188, and 61301131), China Postdoctoral Science Foundation (Granted No. 2014M561347) and the Fundamental Research Funds for the Central Universities No. 3132016024, 3132016311, and 3132016312.

**Author Contributions:** Zhi-An Deng proposed the original idea and wrote this paper; Guofeng Wang gave some valuable suggestions and improved the paper; Danyang Qin developed the decision tree for identification of landmarks; Yang Cui analyzed the data; Zhenyu Na and Juan Chen designed the experiments and revised the paper.

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

## References

1. Greßmann, B.; Klimek, H.; Turau, V. Towards Ubiquitous Indoor Location Based Services and Indoor Navigation. In Proceedings of the 7th Workshop on Positioning Navigation and Communication (WPNC), Dresden, Germany, 11–12 March 2010; pp. 107–112.
2. Zhou, M.; Zhang, Q.; Xu, K.; Tian, Z.; Wang, Y.; He, W. PRIMAL: Page Rank-Based Indoor Mapping and Localization Using Gene-Sequenced Unlabeled WLAN Received Signal Strength. *Sensors* **2015**, *15*, 24791–24817. [[CrossRef](#)] [[PubMed](#)]

3. Gezici, S.; Tian, Z.; Giannakis, G.B.; Kobayashi, H.; Molisch, A.F.; Poor, H.V.; Sahinoglu, Z. Localization via Ultra-Wideband Radios: A Look at Positioning Aspects for Future Sensor Networks. *IEEE Signal Proc. Mag.* **2005**, *22*, 70–84. [[CrossRef](#)]
4. Zhuang, Y.; Yang, J.; Li, Y.; Qi, L.; El-Sheimy, N. Smartphone-Based Indoor Localization with Bluetooth Low Energy Beacons. *Sensors* **2016**, *16*, 596. [[CrossRef](#)] [[PubMed](#)]
5. Ni, L.M.; Liu, Y.; Lau, Y.C.; Patil, A.P. LANDMARC: Indoor Location Sensing Using Active RFID. *Wirel. Netw.* **2004**, *10*, 701–710. [[CrossRef](#)]
6. Zhao, Y.; Dong, L.; Wang, J.; Hu, B.; Fu, Y. Implementing Indoor Positioning System via ZigBee Devices. In Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 26–29 October 2008; pp. 1867–1871.
7. Deng, Z.-A.; Hu, Y.; Yu, J.; Na, Z. Extended Kalman Filter for Real Time Indoor Localization by Fusing WiFi and Smartphone Inertial Sensors. *Micromachines* **2015**, *6*, 523–543. [[CrossRef](#)]
8. Ma, L.; Xu, Y.B. Received Signal Strength Recovery in Green WLAN Indoor Positioning System Using Singular Value Thresholding. *Sensors* **2015**, *15*, 1292–1311. [[CrossRef](#)] [[PubMed](#)]
9. Deng, Z.A.; Xu, Y.B.; Ma, L. Indoor Positioning via Nonlinear Discriminative Feature Extraction in Wireless Local Area Network. *Comput. Commun.* **2012**, *35*, 738–747. [[CrossRef](#)]
10. Xiao, Z.; Wen, H.; Markham, A.; Trigons, N. Robust Indoor Positioning With Lifelong Learning. *IEEE J. Sel. Areas Commun.* **2015**, *33*, 2287–2301. [[CrossRef](#)]
11. Renaudin, V.; Combettes, C. Magnetic, Acceleration Fields and Gyroscope Quaternion (MAGYQ)-Based Attitude Estimation with Smartphone Sensors for Indoor Pedestrian Navigation. *Sensors* **2014**, *14*, 22864–22890. [[CrossRef](#)] [[PubMed](#)]
12. Chen, Z.; Zou, H.; Jiang, H.; Zhu, Q.; Soh, Y.C.; Xie, L. Fusion of WiFi, Smartphone Sensors and Landmarks Using the Kalman Filter for Indoor Localization. *Sensors* **2015**, *15*, 715–732. [[CrossRef](#)] [[PubMed](#)]
13. Abdelnasser, H.; Mohamed, R.; Elgohary, A.; Farid, M.; Wang, H.; Sen, S.; Choudhury, R.; Youssef, M. SemanticSLAM: Using Environment Landmarks for Unsupervised Indoor Localization. *IEEE Trans. Mob. Comput.* **2015**, *1*, 1–14. [[CrossRef](#)]
14. Wang, H.; Lenz, H.; Szabo, A.; Bamberger, J.; Hanebeck, U.D. WLAN-Based Pedestrian Tracking Using Particle Filters and Low-Cost MEMS Sensors. In Proceedings of the 2007 4th Workshop on Positioning, Navigation and Communication (WPNC'07), Hannover, Germany, 22 March 2007; pp. 1–7.
15. Rai, A.; Chintalapudi, K.K.; Padmanabhan, V.N.; Sen, R. Zee: Zero-Effort Crowdsourcing for Indoor Localization. In Proceedings of the 18th Annual International Conference on Mobile Computing and Networking, Istanbul, Turkey, 22–26 August 2012; pp. 293–304.
16. Jan, S.-S.; Hsu, L.-T.; Tsai, W.-M. Development of an Indoor Location Based Service Test Bed and Geographic Information System with a Wireless Sensor Network. *Sensors* **2010**, *10*, 2957–2974. [[CrossRef](#)] [[PubMed](#)]
17. Li, X.; Wang, J.; Liu, C.; Zhang, L.; Li, Z. Integrated WiFi/PDR/Smartphone Using an Adaptive System Noise Extended Kalman Filter Algorithm for Indoor Localization. *ISPRS Int. J. Geo-Inf.* **2016**, *5*, 8. [[CrossRef](#)]
18. Zhou, M.; Qiu, F.; Tian, Z.; Wu, H.; Zhang, Q.; He, W. An Information-Based Approach to Precision Analysis of Indoor WLAN Localization Using Location Fingerprint. *Entropy* **2015**, *17*, 8031–8055. [[CrossRef](#)]
19. Evennou, F.; Marx, F. Advanced Integration of WiFi and Inertial Navigation Systems for Indoor Mobile Positioning. *EURASIP J. Appl. Signal Proc.* **2006**. [[CrossRef](#)]
20. Deng, Z.A.; Xu, Y.B.; Ma, L. Energy Efficient Access Point Selection and Signal Projection for Accurate Indoor Positioning. *China Commun.* **2012**, *9*, 52–65.
21. Scott, D.W. *Multivariate Density Estimation*; Wiley: Hoboken, NJ, USA, 1992.
22. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; Chapman and Hall: London, UK, 1986.
23. Colombo, A.; Fontanelli, D.; Macii, D.; Palopoli, L. Flexible Indoor Localization and Tracking Based on a Wearable Platform and Sensor Data Fusion. *IEEE Trans. Instrum. Meas.* **2014**, *63*, 864–876. [[CrossRef](#)]
24. Edwan, E.; Zhou, J.C.; Zhang, J.Y.; Loffeld, O. A New Loosely Coupled DCM Based GPS/INS Integration Method. *Navig.-J. Inst. Navig.* **2012**, *59*, 93–106. [[CrossRef](#)]
25. Sabatini, A.M. Quaternion-Based Extended Kalman Filter for Determining Orientation by Inertial and Magnetic Sensing. *IEEE Trans. Biomed. Eng.* **2006**, *53*, 1346–1356. [[CrossRef](#)] [[PubMed](#)]
26. Kushki, A.; Plataniotis, K.; Venetsanopoulos, A. Kernel-Based Positioning in Wireless Local Area Networks. *IEEE Trans. Mob. Comput.* **2007**, *6*, 689–705. [[CrossRef](#)]

27. Bar-Shalom, Y.; Li, X.-R.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation*; Wiley: New York, NY, USA, 2001.
28. Gu, Y.; Song, Q.; Li, Y.; Ma, M.; Zhou, Z. An Anchor-Based Pedestrian Navigation Approach Using Only Inertial Sensors. *Sensors* **2016**, *16*, 334. [[CrossRef](#)] [[PubMed](#)]
29. Deng, Z.A.; Xu, Y.B.; Ma, L. Joint Access Point Selection and Local Discriminant Embedding for Energy Efficient and Accurate Wi-Fi Positioning. *KSII Trans. Internet Inf. Syst.* **2012**, *6*, 794–814.
30. Mo, Y.; Zhang, Z.Z.; Lu, Y.; Meng, W.X.; Agha, G. Random Forest Based Coarse Locating and KPCA Feature Extraction for Indoor Positioning System. *Math. Probl. Eng.* **2014**, 850926. [[CrossRef](#)]
31. Zou, D.Y.; Meng, W.X.; Han, S. Euclidean Distance Based Handoff Algorithm for Fingerprint Positioning of WLAN System. In Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC) 2013, Shanghai, China, 7–10 April 2013; pp. 1564–1568.
32. Li, F.; Zhao, C.; Ding, G.; Gong, J.; Liu, C.; Zhao, F. A Reliable and Accurate Indoor Localization Method Using Phone Inertial Sensors. In Proceedings of the 14th ACM International Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012; pp. 1–10.
33. Jahn, J.; Batzer, U.; Seitz, J.; Patino-Studencka, L.; Boronat, J.G. Comparison and Evaluation of Acceleration Based Step Length Estimators for Handheld Devices. In Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Zurich, Switzerland, 15–17 September 2010; pp. 1–6.
34. Deng, Z.-A.; Wang, G.; Hu, Y.; Wu, D. Heading Estimation for Indoor Pedestrian Navigation Using a Smartphone in the Pocket. *Sensors* **2015**, *15*, 21518–21536. [[CrossRef](#)] [[PubMed](#)]
35. Sabatini, A.M.; Martelloni, C.; Scapellato, S.; Cavallo, F. Assessment of Walking Features from Foot Inertial Sensing. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 486–494. [[CrossRef](#)] [[PubMed](#)]
36. Zhou, M.; Xu, Y.B.; Ma, L.; Tian, S. On the Statistical Errors of RADAR Location Sensor Networks with Built-In Wi-Fi Gaussian Linear Fingerprints. *Sensors* **2012**, *12*, 3605–3626. [[CrossRef](#)] [[PubMed](#)]



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).