Fast High-Density Fingerprint Data Acquisition Based on Dense Sampling

Linfu Duan, Zepeng Hu, Feng Xu and Qun Wan
School of Electronic Engineering
University of Electronic Science and Technology of China,
Chengdu, China
*duanlinfu02@163.com

Keywords: Indoor positioning; Wi-Fi positioning; Fingerprint acquisition; Fingerprint database

Abstract. The demand of indoor positioning become more and more urgent with popularity of the smart home, smart medical, etc. In recent years, the research on indoor positioning has achieved some results, especially the method based on Wi-Fi, which is due to Wi-Fi technology is mature, Wi-Fi signal is relatively stable and has a wide range of coverage, etc. Thus, fingerprint based Wi-Fi indoor positioning is low cost, easy to acquire data and implement. However, there are still many problems in the practical application even though the theoretical research is mature. In this paper, we proposed a fast fingerprint data acquisition method based on dense sampling and interpolation aim at the problem of fingerprint database establishment, which acquire data efficiently and take into account maintenance of fingerprint database later.

1 Introduction

Indoor positioning method based on Wi-Fi received signal strength indicator (RSSI) is becoming more and more mature in the field of indoor positioning. In practical applications, fingerprint data acquisition often takes a lot of efforts. In the stage of fingerprint database establishment, the traditional fingerprint based indoor positioning should first divide area to be located into a plurality of grids with known positions. Then, the fingerprint data is sampled by the handheld device standing in the grid [1]-[3]. Finally, the sampled data and grid position are paired and stored in the fingerprint database. The density of the grid greatly affects the final positioning accuracy, the greater the density, the higher the positioning accuracy. However, the establishment of high-density grid is time-consuming and laborious. Moreover, if the indoor environment changes, such as interior decoration change the indoor layout, resulting in the fingerprint database update is also difficult, which makes the fingerprint based indoor positioning has a great limitation in practice.

For the problem of fingerprint data acquisition workload, the current solution is fitting the fingerprint data based on the signal attenuation model by interpolation and extrapolation [4]-[6]. In addition, crowdsourcing has also been used in indoor positioning in recent years, such as, in order to solve the problem of updating fingerprint database, Huang et al proposed and implemented a positioning platform based on crowdsourcing data [7, 8]. Nevertheless, it is very difficult to use attenuation model to accurately describe the influence on Wi-Fi signal by the walls, tables, chairs and people walking around in complex practical environments, thus, the interpolation does not solve the problem essentially. Furthermore, crowdsourcing get the fingerprint data without tags through the users share their own positioning data, and still need to acquire a large amount of fingerprint data in the offline phase, especially in large-scale environments. The traditional data acquisition process separate the position and fingerprint data, which is the main reason leading to a large amount of data acquisition workload. By contrast, the inertial navigation (IN) technology [9, 10] effectively realizes the dynamic fingerprint data acquisition by recording the fingerprint data and position simultaneously, which greatly improves the work efficiency.

This paper proposed a fast high-density fingerprint data acquisition method (FHFDA) aimed at solving the problem of increased time consuming of database establishment due to the high-density grid, as well as the difficulty of updating and maintaining the fingerprint database later, which both resulting in fingerprint based indoor positioning has a great limitations in practical applications.

2 Electronic Map Coordinates

In the practical indoor positioning process, the real position needs to be linked with the electronic map (E-map). In this paper, the accelerometer and the gyroscope in the IN device are used to count walking steps. The number of walking steps, the straight walking length and the E-map coordinates of each step are calculated through the condition detection method. Assume the time of the $k'th$ step or coordinate is $t_s(n)$, $n=1,2,\cdots,L$, L represents the total number of walking steps, the $k'th$ triaxial acceleration sampled by accelerometer is $a_n(1)$, $a_n(2)$ and $a_n(3)$, the $k'th$ triaxial angular velocity sampled by gyroscope is $\omega_n(1)$, $\omega_n(2)$ and $\omega_n(3)$. The condition detection method uses three conditions C_1, C_2 and C_3 to determine whether a person is still or walking.

The three conditions C_1, C_2 and C_3 defined as follows:

Condition C_1 is

$$C_1 = \begin{cases} 1 & th_{a_{\min}} < |a_n| < th_{a_{\max}} \\ 0 & \text{else} \end{cases} \tag{1}$$

where the measurement of acceleration a_n is

$$|a_n| = \sqrt{a_n^2(1) + a_n^2(2) + a_n^2(3)} \tag{2}$$

$th_{a_{\min}}$ and $th_{a_{\max}}$ is the minimum and maximum threshold of $|a_n|$.

Condition C_2 is

$$C_2 = \begin{cases} 1 & \sigma_{a_k}^2 > th_{\sigma_a}^2 \\ 0 & \text{else} \end{cases} \tag{3}$$

$$\sigma_{a_n}^2 = \frac{1}{2c+1} \sum_{j=n-c}^{n+c} (a_j - \bar{a}_n)^2 \tag{4}$$

$$\bar{a}_n = \frac{1}{2c+1} \sum_{j=n-c}^{n+c} a_j \tag{5}$$

where $\sigma_{a_n}^2$ is the local acceleration variance of the $k'th$ time, $th_{\sigma_a}^2$ is the minimum threshold of $\sigma_{a_n}^2$, $\bar{a}_n$ is the mean local acceleration, c is the window length of the mean. According to the trait of accelerometer, the local acceleration variance $\sigma_{a_n}^2$ will increase when the measurement of acceleration a_n is small.

Condition C_3 is

$$C_3 = \begin{cases} 1 & |\omega_n| < th_{\omega_{\max}} \\ 0 & \text{else} \end{cases} \tag{6}$$

where the measurement of gyroscope ω_n is

$$|\omega_n| = \sqrt{\omega_n^2(1) + \omega_n^2(2) + \omega_n^2(3)} \tag{7}$$

$th_{\omega_{\max}}$ is the maximum threshold of $|\omega_n|$.

Taking logical AND operation between the three conditions, i.e., the condition test result is $C_1 \& C_2 \& C_3$. It is easy to know that the logic output "1" denotes the acceleration and angular velocity are small, i.e., stationary state, the logic output "0" denotes walking state. Hence, change from the stationary state to the walking state is counted as one step. Assume the total number of steps walked on the current path at the $k'th$ time is $m(n)$, and the length of every step is approximated as a fixed length l, then, the straight walking length is

$$d(n) = m(n) * l \tag{8}$$

or

$$d(n) = d(n-1) + l \tag{9}$$

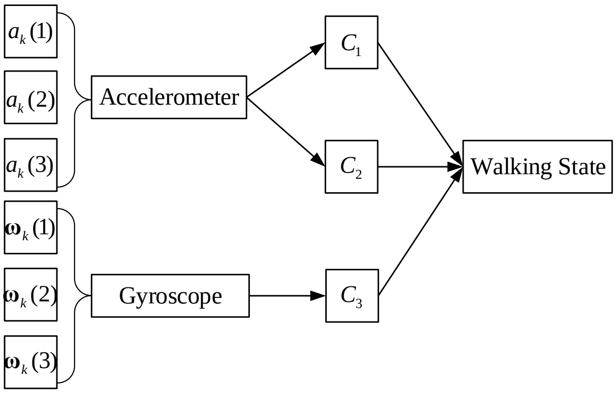


Figure 1. Determining the state of pace through the IN device.

In the offline fingerprint data acquisition, we get E-map coordinates of the unit path and save the corresponding time by determining the state of each step according to the condition detection through the IN device. Determining the state of pace through the IN device is shown in the Figure 1.

3 Fast Fingerprint Data Acquisition

This paper first elaborates the complete procedure of the fingerprint data acquisition method proposed. Secondly, compares with the traditional fingerprint data acquisition and interpolation method to analyze the performance of the proposed method. Finally, gives positioning accuracy comparison through the simulation

3.1 Procedure of Fingerprint Data Acquisition Method

The procedure of the FHFDA method is as follows:

1) Get the E-map coordinates of the unit path and save the corresponding time $t_s(n)$, $n = 1, 2, ..., L$, through the method of the previous section;

2) Acquire RSSI values from N wireless signal transmitting devices arranged in the environment in accordance with the minimum scanning time, obtain N RSSI values at the jth scanning time $t_r(j)$, $j = 1, 2, ..., T$;

3) Combine the sampled RSSI values with the E-map coordinates, calculate the absolute value of the difference between the scanning time $t_r(j)$ of N RSSI values and all E-map coordinates corresponding time $t_s(n)$, find the time $t_m(j)$ which corresponding to minimum absolute value

$$t_m(j) = \min_{t_s(n)} \left| t_r(j) - t_s(n) \right|, n = 1, 2, ..., L \tag{10}$$

The coordinates of N RSSI values scanned at time $t_r(j)$ is located at the E-map coordinates generated at time $t_m(j)$, meanwhile, they are paired and saved as a basic data unit of the fingerprint database;

4) If there are multiple RSSI values corresponding to the same E-map coordinates, then, the average of multiple RSSI values is taken as the unit data of the E-map coordinates in the fingerprint database.

The procedure of the FHFDA method is shown in Figure 2.

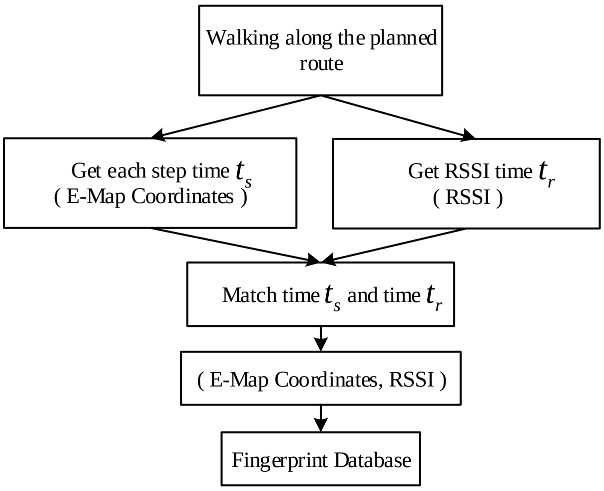


Figure 2. The procedure of the FHFDA method

3.2 IN Data Acquisition and Calibration

As shown in Figure 3, in the practical indoor positioning process, the position of people's activities is presented as a path, such as the corridor of the hospital, the walkway in the office or the flow of people in the shopping mall, etc. Therefore, we will plan the walkway form if there are open areas when establishing a fingerprint database.

First of all, we plan the route under test in the practical road, set the starting position as (x_0, y_0) and the road direction as x axis. According to the IN positioning in the previous section, the number of walking steps can be determined by the gyroscope and accelerometer, and the coordinates of each step can be obtained as (x_i^*, y_i^*), $i = 1 \cdots m$, as shown in Fig. 3. The distance between the end and initial position of the planned route is S.

We know that the biggest drawback of the IN positioning is that it will produce cumulative error after a period of positioning. Accordingly, it is necessary to perform error calibration. In this paper, the IN system is used to calibrate the position of fingerprint data sampled offline. The final abscissa of the position produced by the IN is x_m^*, which deviation is $\left|S\text{-}x_m^*\right|$. In order to calibrate the position, we could compensate for the deviation according to the acceleration of each step. Deviation calibration for each step is as follows

$$\Delta_i = \frac{\bar{a}_i \cdot t_i^2 - \bar{a}_{i-1} \cdot t_{i-1}^2}{x_m^* - x_0^*} \left|x_m^* - S\right|, i = 1, \cdots, m \tag{11}$$

Let $x_i^{'} = x_i^* \pm \Delta_i$ replace the coordinates before. This paper will not consider the error of ordinate because we establish the abscissa along the planned route in default.

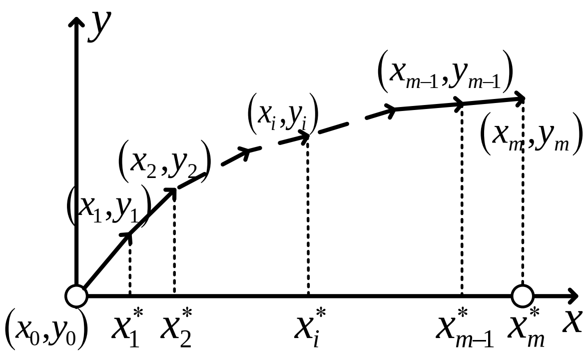


Figure 3. Indoor positioning

3.3 Data Interpolation

If acquire data by grid method, the sampled data will be interpolated to solve the problem that the grid is not dense enough. Common interpolation methods include inverse distance weighted interpolation (IDW) [4, 6], natural neighbor interpolation, Kriging interpolation, radial basis function method and linear interpolation based on Delaunay triangulation [11], etc. Kriging interpolation method is suitable for the condition that the mean of RSSI values is zero. Linear interpolation may produce points outside the planned route for the points that the sampled data is linearly distributed. All in all, we choose IDW interpolation in the experiment mainly because it is classic, simple, stable and suitable for indoor wireless signal model.

The main idea of a typical IDW interpolation method is to weight the RSSI values of the interpolation points according to the distance between the interpolation point and the known sample point. Assume the interpolation point coordinates is (x, y), its interpolation weight w_i with respect to the known sampling points (x_i, y_i), $i = 1, \cdots n$, is

$$w_i = \frac{d_i^{-z}}{\sum_{i=1}^{n} d_i^{-z}} \tag{12}$$

where d_i is the distance between the interpolation point and sampling points (x_i, y_i), z is exponential parameter. The total RSSI value of interpolation point is

$$RSS^*_{(x,y)} = \sum_{i=1}^{n} w_i RSS_{(x_i, y_i)} \tag{13}$$

3.4 Experiment

As shown in Fig. 4, the length and width of the venue are 75m and 35m, respectively. There are 115 Access Points (APs) detectable in the venue. Define the database is

$$L = \{D_1, D_2 \cdots D_n\} \tag{14}$$

$$D_i = \{RSSI_{i1}, RSSI_{i2} \cdots RSSI_{i115}\} \tag{15}$$

where n is the number of sampling points or sampled data.

Note that our device does not get RSSI values from all APs in many positions, which lead to data missing. For missing values, the uniform setting is -90dB, which represents a very weak value. Therefore, we obtain a $n \times 115$-dimensional fingerprint database L, and record the coordinates of each piece of data in the same time

$$Y = \{(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)\} \tag{16}$$

The traditional grid sampling method divide the area to be located into a plurality of grids with known positions, and acquire the fingerprint data at each grid by human-held devices to form a fingerprint database. In the experiment, we divide grid points by 1 meter intervals at open space in the venue as shown in Fig. 4. There are 408 sampled data in total, so the establishment of the database takes about 34 minutes according to each sampling point spent 5s in average. The fingerprint database sampled by the traditional grid sampling method is L_1 and the corresponding position is Y_1.

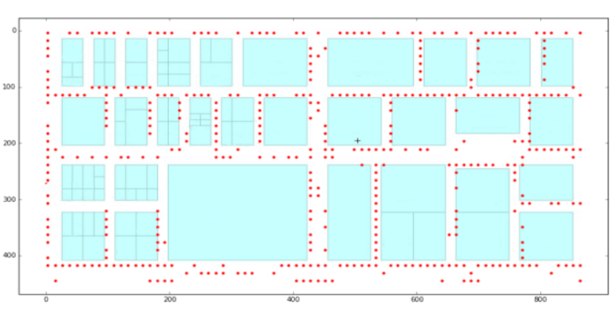
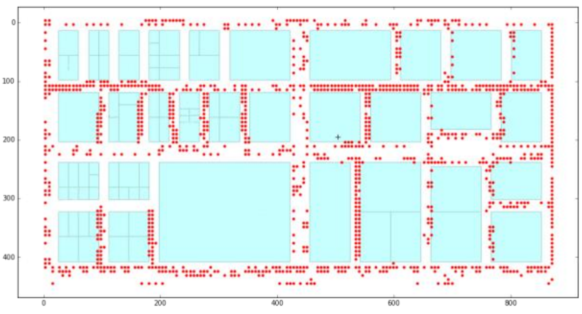


Figure 4. Distribution of points by traditional methoFigure



5. Distribution of points after interpolation

The traditional grid sampling methods cost long time and are inefficient, so the fingerprint database can be augmented with the interpolation method described in Section 3.C. The fingerprint database data points are expanded to 1183 by interpolation as shown in Fig. 5. After interpolation, the fingerprint database is L_2 and the corresponding position is Y_2.

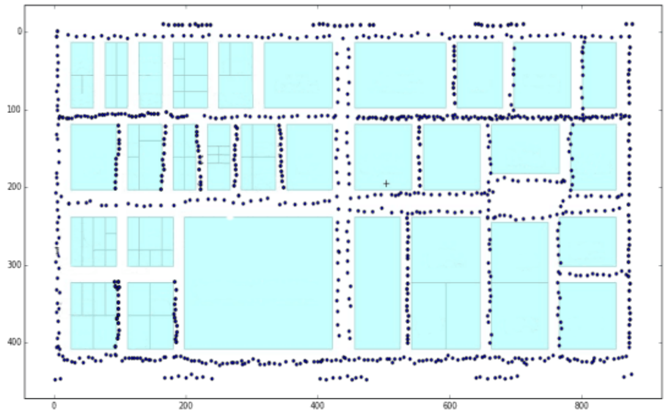


Figure 6.Distribution of sampling points of the IN device

TABLE I. COMPARISON OF SAMPLING PROCEDURE

Sampling Method	**Time (minute)**	**Sample number**
Tradition	35	408
Interpolation	More than 35	1183
Proposed	6	935

By comparison, the sampling method proposed acquire the RSSI values through the physical layer sampling. The sampling period is set to 0.5 seconds, thus, sampling density is about 0.5 meters according to walking speed is 3.5 kilometers per hour. It takes about 6 minutes to complete the whole map sampling and get 935 sampled data. The sampling points of the IN device are shown in Figure 6.

The number of sampled data and time-consuming of the traditional grid sampling method and sampling method proposed are shown in Table 1. As shown in Table 1, the sampling method proposed has obvious advantages in the time and size of the fingerprint database. While the size of the database has doubled, the database establishment time is only equivalent to one sixth of the traditional grid sampling method. The fingerprint database sampled by the IN device is L_3 and the corresponding position is Y_3.

In order to verify the positioning accuracy of the three fingerprint data acquisition methods, we randomly selected 250 grids to acquire the test data, which expressed as $T=\{T_1,T_2\cdots T_{250}\}$. Then, take use of fingerprint database L_1, L_2 and L_3 to match and analyze the positioning performance .The specific method of matching are as follows.

Take L_1 as an example, calculate the Euclidean distance between each test data T_i and each fingerprint data of L_1

$$\begin{aligned} &E_{i,j} = Euclidean\ (T_i, D_j) \\ &i=1,2,\cdots 250;\ j=1,2,\cdots 408 \end{aligned} \tag{17}$$

From(17), we can get 250×408-dimensional matrix E, where 250 indicates the number of test data, and 408 indicates the number of fingerprint data. For each test data, find out the fingerprint data in fingerprint database L_1 which closest to the test data in Euclidean distance as a result of their matching

$$y_i = Y_1[\arg\min(E[i,:])] \quad i=1,2,\cdots m \tag{18}$$

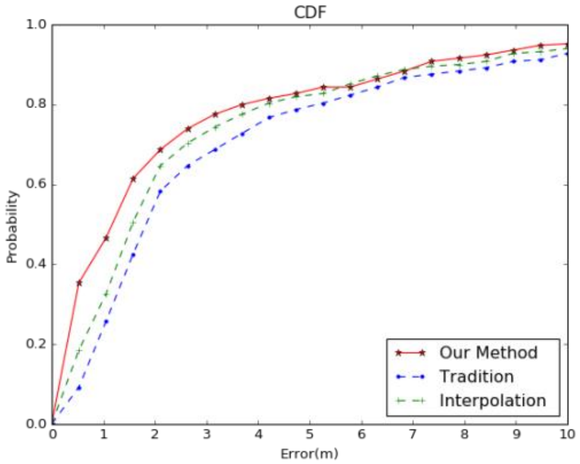


Figure7.Positioning error cumulative distribution of three methods

Using the common correlation function matching indoor positioning method for the fingerprint database L_1, L_2 and L_3 sampled by the traditional grid sampling method and the sampling method proposed. The positioning error cumulative distribution of 250 random positions is shown in Fig. 7. The positioning accuracy can be reduced to less than 2 meters with 65% probability by the method proposed, which is improved by about 30% compared with the traditional grid sampling method with 65% probability of less than 2.8 meters. It is obvious that the sampling method proposed improve the positioning accuracy significantly. The mean positioning error of three methods are shown in Table 2.

TABLE II. COMPARISON OF POSITIONING ACCURACY

Sampling Method	**Mean error (meter)**
Tradition	3.90
Interpolation	3.52
Proposed	2.81

Moreover, this method is suitable for almost all RSSI-based indoor positioning methods and is an effective supplement to the database establishment, which is of great significance for fingerprint data acquisition, and is convenient for large-scale popularization and application.

4 Conclusions

A fast high-density fingerprint data acquisition method for indoor fingerprint database is proposed, which generate a large fingerprint database by acquiring fingerprint data in the continuous movement in a short period of time, and simplify the establishment of the fingerprint database. Meanwhile, the method reduce the workload of resampling and maintenance if fingerprint database need to be reconstructed when the indoor environment changes. The sampling method proposed has the advantages of fast fingerprint data acquisition and high-density fingerprint data compared with the traditional grid sampling method. In detail, this method reduce the fingerprint data acquisition time by about five-sixths, double the fingerprint data size, reduce the cost of fingerprint based indoor positioning, and improves the positioning accuracy finally.

Acknowledgment

This work is supported in part by the National Natural Science Foundation of China (NSFC) under Grant U1533125 and 61771108, National Science and technology major project under Grant

2016ZX03001022, and the Fundamental Research Funds for the Central Universities under Grant ZYGX2015Z011.

References

[1] P. Addesso, L. Bruno, R. Restaino. 2010. Adaptive localization techniques in WiFi environments,"IEEE International Symposium on Wireless Pervasive Computing, 2010, pp.289-294.

[2] C. Feng, W.S.A. Au, S. Valaee, et al. 2012.Received-Signal-Strength-Based Indoor Positioning Using Compressive Sensing, IEE Transactions on Mobile Computing, vol. 11, no. 12, 2012, pp.1983-1993.

[3] S.H. Fang, C.H. Wang. 2015. A Novel Fused Positioning Feature for Handling Heterogeneous Hardware Problem, IEEE Transactions on Communications, vol. 63, no. 7, 2015, pp.2713-2723.

[4] Talvitie, Jukka, M. Renfors, and E. S. Lohan. 2015. Distance-Based Interpolation and Extrapolation Methods for RSS-Based Localization With Indoor Wireless Signals, IEEE Transactions on Vehicular Technology, vol. 64, no.4, 2015, pp.1340-1353.

[5] H. Zhao, B. Huang, B. Jia. 2016.Applying kriging interpolation for WiFi fingerprinting based indoor positioning systems,Wireless Communications and NETWORKING Conference. IEEE, 2016.

[6] D. Denkovski.2012. Reliability of a radio environment Map: Case of spatial interpolation techniques, International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications. IEEE, 2012.

[7] Z. Huang, Y. Chen,J. Liu, et al. 2016. Indoor Localization Method and Platform Based on Crowdsourcing Data,Journal of Geo-Information Science, 2016.

[8] S.H. Jung,S. Lee,D. Han. 2016. A crowdsourcing-based global indoor positioning and navigation system, Pervasive & Mobile Computing, vol. 31, 2016, pp.94-106.

[9] C.C. Ho. 2015. Real-Time Indoor Positioning System Based on RFID Heron-Bilateration Position Estimation and IMU Inertial-Navigation Position Estimation, IEEE, Computer Software and Applications Conference. IEEE Computer Society, 2015, pp.481-486.

[10] T.Hoshizaki,E. Tashiro. 2009. Computational scheme for MEMS inertial navigation systemes. US, US20090171585. 2009.

[11] Etal, Markde Berg. 2013. Computational geometry : algorithms and applications, 3rd ed. World Book Inc Beijing company, 2013.