

Recommended Paper

Method to Improve Accuracy of Indoor PDR Trajectories Using a Large Number of Trajectories

KOSUKE YOTSUYA¹ KATSUHIRO NAITO¹ NAOYA CHUJO¹ TADANORI MIZUNO¹ KATSUHIKO KAJI^{1,a)}

Received: March 29, 2019, Accepted: October 3, 2019

Abstract: Building structure information is essential for achieving various indoor location-based services (ILBSs). Our approach integrates a large amount of pedestrian trajectories acquired by pedestrian dead reckoning (PDR) for generating a pedestrian network structure. To generate highly accurate pedestrian network structures, the accuracy of each trajectory must be improved. In this paper, we propose a method to improve the accuracy of indoor PDR trajectories by using many such trajectories. First, we select reliable trajectories based on the stability of the sensing data. Next by analyzing the trend of the step lengths, we correct the length of the trajectories. Finally, with same-route trajectories, we generate average trajectories for each route. We experimentally used HASC-IPSC and found that our proposed method improved the accuracy of the trajectories. The cumulative error rate of the original pedestrian trajectories was 0.1111 m/s. After adapting our proposed method, the rate improved to 0.0622 m/s.

Keywords: pedestrian dead reckoning (PDR), data integration, crowd-sensing, building structure information

1. Introduction

As location estimation techniques continue to evolve, various indoor location-based services (ILBSs) are being deployed. GPS and wireless LAN methods are generally used to acquire position information outdoors. Unfortunately, an indoor GPS cannot be used in many cases. Therefore, a method is required that stably acquires the position information of various buildings.

One indoor location estimation technique that has been shown to be effective is pedestrian dead reckoning (PDR). In this technique, we use the values of various sensors incorporated in a terminal such as a smartphone to estimate the walking and traveling directions of the person holding the smartphone and their movements with respect to a starting point [1], [2], [3]. One merit of PDR is that new infrastructure equipment does not have to be installed, and it is not limited to available places.

Building structure information is essential for various ILBSs. However, since building structure information is not necessarily available for every building at present, research is estimating the information of building structure using the sensing data from walking in the buildings themselves [4], [5], [6], [7], [8]. To generate highly accurate building structure information, we must accurately estimate each pedestrian trajectory.

Based on the above background, we are integrating a large amount of pedestrian trajectories obtained by PDR to generate a pedestrian network structure that expresses a route where pedestrians pass and the basis of route calculation. We assume a crowd-sensing approach for collecting pedestrian trajectories. We acquired daily activity sensing data from the smartphones of volun-

tary users in a specific building. After collecting them, the pedestrian trajectories are estimated based on PDR technology. Then the pedestrian network structure is automatically integrated using a large number of trajectories (**Fig. 1**).

We previously proposed a 3D pedestrian trajectory estimation method that uses walking sensing data [9]. In this study, we accumulated error of about 1 m in ten-second measurements. Accuracy can be estimated in room units if the sensing is within about 40 seconds. However, the accuracy of the estimated pedestrian trajectory must be improved for generating building structure information where ILBS is used.

In this paper, we propose a method that improves the accuracy of indoor PDR trajectories using a large number of them. We assume that we know whether the sensing data from each pedestrian reflects walking on the same route. This correspondence can be obtained by manually labeling the crowd-sensing or route-matching methods using the similarity of the trajectory shapes and the observed Wi-Fi signals while walking [10]. First, we selected reliable trajectories based on the stability of the sensing data. The reliability of the data is high if the sensor's value is relatively stable and the pedestrian's movement is reasonable. Therefore, we selected the most reliable pedestrian trajectory for each identical route and averaged the pedestrian trajectories based on this most reliable trajectory. The next stage is correcting the step lengths. Since accurately estimating them is difficult in PDR trajectories, the estimated walking distance tends to be biased. We examined the deviation of the walking distance from many pedestrian trajectories to improve the accuracy of the PDR trajec-

¹ Graduate School of Business Administration and Computer Science, Aichi Institute of Technology, Toyota, Aichi 470-0392, Japan

^{a)} kaji@aitech.ac.jp

The preliminary version of this paper was published at The 11th International Conference on Mobile Computing and Ubiquitous Networking, October 2018. The paper was recommended to be submitted to the Journal of Information Processing (JIP) by the chief examiner of SIGMBL.

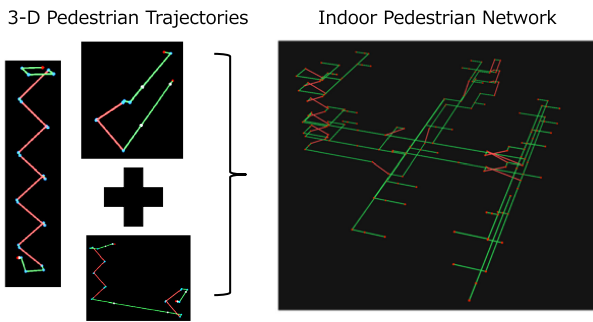


Fig. 1 Generation of pedestrian network structure by integration of multiple pedestrian trajectories.

tory. Finally with same-route trajectories, we generated average trajectories for each route and assume that the living spaces inside offices are our data collection environment. These processes obtained accurate pedestrian trajectories in our experiments.

2. Related Research

Several methods have generated the structural information of buildings. Current methods are already using existing floor maps, CAD data, or newly collected building data. However, costs are accrued, including the preparation of sensing equipment, sensing requests, and so on. Therefore, several research studies are estimating the information of building structures using walking sensing data from the buildings themselves [4], [5], [6], [7], [8]. We also adopted such an approach because crowd-sensing technology can collect such data from buildings.

To generate accurate building structure information, we must accurately estimate each pedestrian trajectory. Pedestrian trajectories by general PDR are estimated based on the information of step counts, step lengths, and walking direction. This information can be detected using various sensors, such as accelerometers, angular velocity sensors, and barometers. We previously proposed a 3D pedestrian trajectory estimation method that uses walking sensing data [9] (**Fig. 2**). In that study, we accumulated error of about 1 m in 10-second measurements. An accurate estimation in room units can be achieved if the sensing is within about 40 seconds. However, for achieving ILBS, we must improve the accuracy of the estimated pedestrian trajectories.

The estimation accuracy of the pedestrian trajectories is influenced by the sensor's stability while collecting walking sensing data with PDR. We introduced the concept of a stable-walking zone and performed reliable data detection in our previous research [9]. A stable-walking zone is a section where walking without turning or meandering can be stably sensed for longer than a certain period of time. Much of the movement is detected as a stable-walking zone when the movement is rational and the sensor's motion is stable because most buildings consist of straight corridors. Therefore, data are detected with high reliability based on the ratio of the stable-walking zone. We experimentally evaluated the accuracy of estimated pedestrian trajectories and found they were higher when using data with high reliability. For this reason, we believe that the selection of a pedestrian trajectory based on the stability of the sensing data effectively improves the estimation accuracy of pedestrian trajectories.

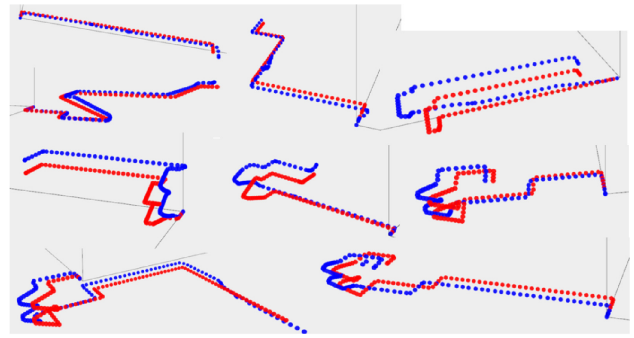


Fig. 2 Estimated pedestrian trajectory [9].

One problem is that the walking distance tends to be biased in its estimation in PDR because the bias estimation of the step length depends greatly on sensors and user profiles. A general calculation method of the step length is gauged using fixed values such as the height of the pedestrian or the amplitude of acceleration sensors [11], [12]. The walking distance in PDR can be obtained by multiplying the step length by the step count. Therefore, the total walking distance tends to be greatly biased because of the differences in each step length. This problem can be solved if the pedestrian step length is obtained by careful calibration beforehand. However, this task is very laborious due to its crowd-sensing. Since we thought that investigating the step lengths with a great quantity of walking data of pedestrians might solve these problems, we investigated the bias of the walking distances from a large amount of walking data by multiple pedestrians. We believe that the correction of the step lengths based on that bias will improve the accuracy of the PDR trajectories.

One approach improves trajectory accuracy by averaging multiple identical characters and figures. Niino et al. improved the accuracy by averaging multiple trajectories [13]. In that study, multiple users handwrote the same Japanese characters and figures. Then multiple identical trajectories were balanced to create average characters. Their evaluation experiments determined which trajectory was more beautiful: the handwritten character or the averaged character. The trajectory of the average characters was deemed more beautiful than the handwritten characters. We assume that the living spaces inside offices are the data collection environment. Therefore, it is possible to collect walking data of people's identical route to the same departments. Based on this reason, we expected to improve the accuracy by averaging the trajectories using many people who walk on the same route.

3. Method to Improve Accuracy of Indoor PDR Trajectories Using Many Trajectories

In this section, we propose a method that improves the accuracy of indoor PDR trajectories using three approaches. First, we select reliable trajectories based on the stability of the sensing data. Next, we analyze the trends of the step lengths and correct them. Finally with the same-route trajectories, we generate an average trajectory for each route. In the proposal method, we only utilize general building structure knowledge (e.g.: corridor is straight, corridors are always parallel or have vertical relation-

ships with each other) and building information that is easy to acquire and commonly applied to buildings (e.g.: step length of stair section). Although correspondence labeling is currently a manual task, a partial matching estimation method of the walking trajectories [10] should alternate the tasks in the future.

Here, the term “route” means a path in terms of the building structure that consists of stairs and corridors on the map. The term “trajectory” denotes an estimated path by PDR.

3.1 Selection of Reliable Trajectories Based on Stability of Sensing Data

In this section, we describe the selection process before the correction of the walking trajectories. Here the processing only selects a trajectory that seems highly reliable from among a large number of estimated walking trajectories without correcting each one. Therefore the section process is just pre-processing.

Since we believe that more accurate pedestrian trajectories can be obtained by correction based on accurate trajectories, we determined which trajectory to correct from multiple pedestrian trajectories. The accuracy of the estimated pedestrian trajectories may be different even in those pedestrian trajectories taking the same route because the accuracy of the estimated trajectory depends on the sensor operation’s stability. For example, **Fig. 3** shows the PDR trajectories in which two pedestrians walk straight down the center of the same route. However, since the trajectories of the section where walking straight ahead are crooked due to sensor failure, the pedestrian trajectories can be estimated with high accuracy when the sensor operation is stable and the data reliability is high.

We focus on a straight-line route in this study. Our main target is normal buildings that consist of straight corridors. When the pedestrian orientation changes slowly, it is hard to distinguish between the drift of angular velocity sensors and walking down a wavy corridor. Drift is an essential problem of PDR. The slight offset of the angular velocity causes drift because the value continues to accumulate. The slight offset varies with the terminal’s heat. Therefore, since the orientation may be crooked even if a pedestrian walks directly ahead, distinguishing between a change in the orientation due to drift or walking on a wavy route is difficult. Consequently, we focus on straight-line routes.

We assume that highly accurate estimation is possible for pedestrian trajectories with high data reliability. In a previous study [9], we set the thresholds for data reliability. The accuracy of the pedestrian trajectories only improved when we used reliable route data, suggesting the effectiveness of the selection of reliable trajectories based on the data’s reliability.

The reliability of the estimated pedestrian trajectory for each identical route can be calculated by the ratio of the stable-walking zone [9], which is a section of stable linear walking for a certain period of time.

Here, we regard that reasonable/natural pedestrian movement is walking linearly stably along corridor. Also, general building consists of straight corridors. Therefore, when a person walks reasonably in general building, most of the movement should be linear walking. Then, we regard the ratio of stable-walking zone as the reliability of the trajectory. The reliability is combination

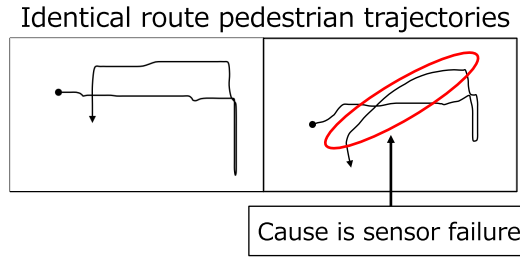


Fig. 3 Trajectory’s bend when walking straight on same route.

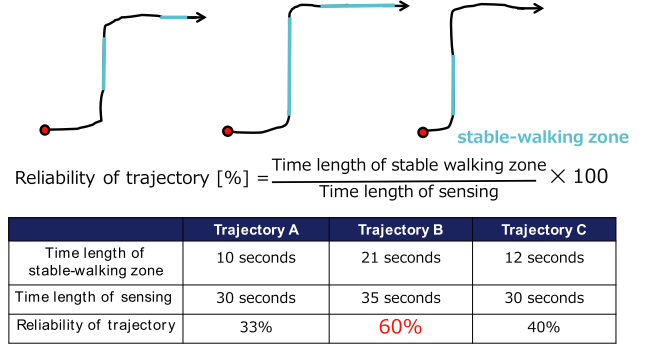


Fig. 4 Selection of pedestrian trajectory based on data reliability.

of the factor of reasonability of pedestrian movement and stability of the sensors. The highest reliable trajectories for each route are corrected (**Fig. 4**).

Specific processing is shown below. Corresponding pseudo-code is shown in Algorithm 1. First, let the set of all the walking trajectories be $T = [t_1^1, t_2^1, \dots, t_{m_i}^1, \dots, t_{m_n}^n]$. Here a route is assumed to exist from 1 to n . Also, we set the number of walking trajectories obtained by PDR by walking route i to m_i . t_j^i represents the j -th walking trajectory of route i .

Assuming that the total length of walking trajectory t_j^i is $l_{whole}(t_j^i)$ and the length of the stable-sensing zone is $l_{stable}(t_j^i)$, reliability of trajectory t_j^i can be calculated as $l_{stable}(t_j^i)/l_{whole}(t_j^i)$. For each route i , we selected path t^i with the highest reliability of stable-sensing zones among the trajectories from t_1^i to $t_{m_i}^i$. Hereafter, a walking trajectory set, which only consists of the most reliable estimated trajectories for each route, is denoted as $T_{reliable} = [t^1, \dots, t^n]$. We gradually refine walking trajectory set $T_{reliable}$ in the following sections.

3.2 Correction of Trajectory Length by Analyzing Trends of Step Length

Since estimating the exact step length on PDR trajectories is difficult, walking distance tends to be easily biased. Therefore, we improved the accuracy of the PDR trajectories correcting the step lengths based on the walking distance’s deviation. In a 3D pedestrian trajectory estimation method [9], the walking distance of the estimated pedestrian trajectory was deemed to be larger than the route’s correct answer data. In a previous method [1] that calculated the walking distance, the step length was set to 0.46 times the height. The movement distances were obtained based on the number of steps and their lengths. If we apply these methods, the estimated distance is not necessarily distributed around the correct value. Therefore, we must correct the trajectory length. Here we assumed that crowd-sensing can collect a

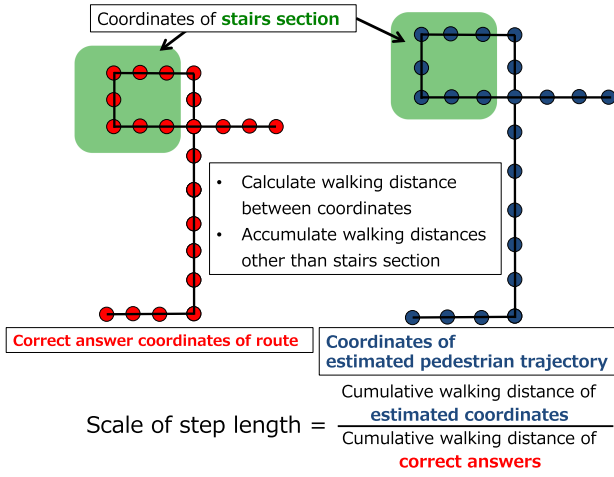


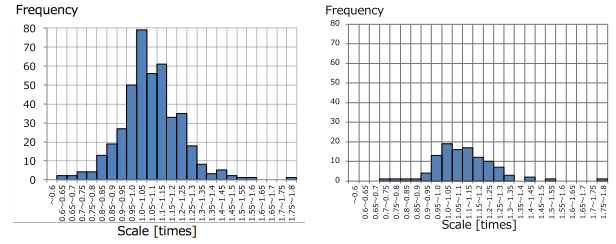
Fig. 5 Calculation of scale of step lengths.

large amount of walking sensing data whose correct distance is known. Since the HASC-IPSC data contains the correct distance value about the length of the walking trajectory, they are used to correct the distance estimation.

Furthermore, since the step length (due to ascending and descending stairs) is often shorter than the step length during normal times, building knowledge was introduced and the step length when ascending and descending stairs was fixed to 0.3 m [14]. In our previous method [9], the stairs section was detected based on barometer values and the step length in the stairs section was set to 0.3 m. In many cases, since the height and length of each stair is constant within a building, collecting such building knowledge may prove inadequate for each building. Therefore, we introduce stair length and height as building knowledge that falls within a reasonable range. We investigated the bias of the walking distance (other than the stairs section) using a large amount of walking data from multiple pedestrians and corrected the step lengths from the investigated values.

Next we compared the correct answer data of the route with the cumulative walking distance of the estimated pedestrian trajectories other than the stairs section and calculated the ratio as a scale of the step length (Fig. 5). A change in height is seen in the Z-axis coordinate of the coordinate data of the 3D pedestrian trajectories. We extracted the part that is outside of the stairs section of the estimated trajectory and the corresponding correct route. We then calculated the length of the trajectory and the correct route other than the stairs section. We compared the length of the estimated trajectory and the correct route, and calculated the step length's scale from the result.

Next we calculated the scale of the step lengths with respect to all the trajectories and checked the scale's frequency distribution. The results are shown in Fig. 6 and Table 1. We used 424 bits of walking data of 3D pedestrian trajectories estimated by our method [9] with the HASC-IPSC [16] indoor pedestrian sensing corpus. Fig. 6(a) shows the distribution of the scales of all the estimated pedestrian trajectories T . Fig. 6(b) shows the distribution of the scales that were selected as the most reliable trajectories among all identical routes T_{reliable} . Since this distribution's shape has a normal distribution, the step length is corrected using the scale's average value. Based on the data's reliability, it is



(a) All estimated pedestrian trajectories (b) Selected estimated pedestrian trajectories

Fig. 6 Frequency distribution of scales of step length.

Table 1 Result of frequency distribution of scale.

	All estimated pedestrian trajectories	Selected estimated pedestrian trajectories
Total amount of data	424	109
Maximum value	1.756	1.756
Minimum value	0.635	0.704
Standard deviation	0.145	0.143
Average value	1.070	1.108

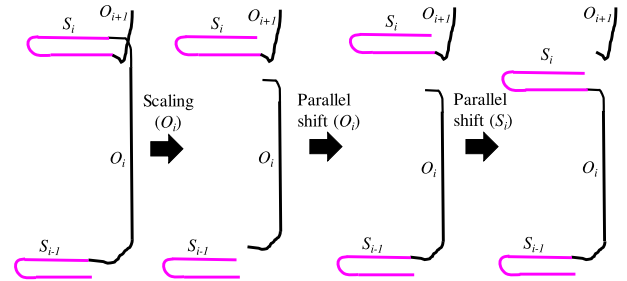


Fig. 7 Correction of step lengths.

corrected using the average value of the scale of the most reliable trajectories among each identical route.

Next we corrected the step length using the calculated average of the scale. We corrected the X and Y coordinates of the coordinate data (other than the stairs section) and enlarged or reduced the pedestrian trajectories using the step length's scale. We shifted the coordinates of the stairs section and followed up on it after it was enlarged or reduced. These processes were repeated until the end coordinates of the trajectory.

The following is the procedure for correcting the distance estimation error of the estimated walking trajectory (Fig. 7). A corresponding pseudo-code is shown in Algorithm 2. For walking trajectory t , partial trajectory o (meaning other) other than the stairs, and partial trajectory s (meaning stair) of the stair section appear alternately. If there are k stair sections, the sequence of walking trajectory t can be expressed as $(o_1 \rightarrow s_1 \rightarrow o_2 \rightarrow s_2 \rightarrow \dots s_k \rightarrow o_{k+1})$. Here we assume that the trajectory does not start or end at the stairs.

The series of walking trajectories is followed sequentially, and the partial trajectories other than the stairs are scaled based on the values obtained in Table 1. Here let c_i be the partial trajectory currently being focused on. If c_i is a partial trajectory other than the stairs, the following scaling process is applied. Here we set the scaling factor to a . Since Table 1 shows that the estimated trajectory is estimated to be 1.108 times longer than the correct trajectory, the scale of each partial trajectory is $a = 1/1.108$, and the walking trajectory is scaled down using scaling factor a .

Plane components of all coordinates (x, y, z) are converted to $(a \cdot x, a \cdot y, z)$ by scale a . The trajectory after scaling processing is replaced by c_i . The scaling process is only for plane components of the trajectory. When partial trajectories other than the stairs are scaled, the sequences of the partial trajectories become discontinuous. Therefore, c_i is shifted in parallel so that the end coordinates of previous partial trajectory c_{i-1} and the start coordinate of partial trajectory c_i just overlap. The trajectory after parallel translation is replaced by c_i . The parallel shift process is applied to both of partial trajectories of the stair section and other than stairs.

The trajectory obtained by this processing for original walking trajectory t is expressed as t' . This process is performed for each element of all estimated walking trajectories T and $T_{reliable}$. Each walking trajectory set after the process is expressed as T' and $T'_{reliable}$.

3.3 Generation of Average Trajectories

One problem is that the drift error gradually accumulates from the start of sensing in the PDR. Such drift reflects the value of the angular velocity sensor self-fluctuates over time even when the terminal is stationary. In addition, this value changes randomly depending on the type of angular velocity sensor and the terminal's heat. Therefore, we obtained an average value, as is done with the central limit theorem [17], by a comparison with a large amount of walking data. The required number of samples is generally 30 or more when using the central limit theorem. This study had about 3 to 7 samples per route. Therefore, the number of samples is inadequate for the central limit theorem. However, we expect that averaging the contribution will improve the accuracy even though the number of samples is insufficient.

We automatically detected a stable-walking zone from the estimated pedestrian trajectories. Here, stable-walking zone means an interval in which the angular velocity is within 20 degrees of change lasts for 3 seconds or more [9]. Also, we manually labeled the correspondence of each stable-walking zone on the same straight-line route in the multiple pedestrian trajectories of identical routes. Although correspondence labeling is currently a manual task, a partial matching estimation method of the walking trajectories using the similarity of the trajectory shapes and the observed Wi-Fi signals while walking [10] could alternate the tasks in the future. Our comparison focused on one linear route from all of the pedestrian trajectories using the labeling result of the correspondence relation.

The correction of a pedestrian trajectory using multiple pedestrian trajectories of identical routes is defined as the generation of average trajectories. We corrected the estimated pedestrian trajectories using the correspondence relation of stable-walking zones.

Next we explain the procedure of the averaging process. A corresponding pseudo-code is shown in Algorithm 3. A set of walking trajectories with route ID of i in T' is expressed as $T'^i = [t'_1, t'_2, \dots, t'_{m_i}]$. The trajectories contained in T'^i are averaged by the procedure described below. We assume that each walking trajectory of T'^i includes multiple stable-sensing zones, and stable-sensing zones walking on the same linear route are

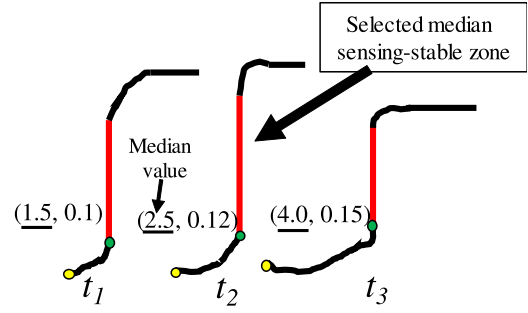


Fig. 8 Selection of median stable-walking zone.

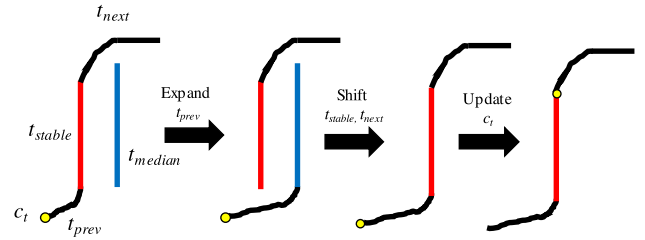


Fig. 9 Expansion/Contraction process.

associated with each other. Associated stable-sensing zones are parallel relationship each other. Also, stable-sensing zones are always along the X or Y axis due to applying general building knowledge [9].

As a preparatory step before processing, for each element of walking trajectory $T'^i = [t'_1, t'_2, \dots, t'_{m_i}]$, a reference point is identified to memorize the completion point of the averaging process. Let c_t be a reference point in trajectory t . The initial value of the reference point is the starting point of trajectory t .

We also define function *get-partial-trajectories*(t, j), which requires walking trajectory t and the number of the stable-sensing zones j as arguments. The following are the return values: t_{stable} , which is a partial trajectory of the j -th stable-sensing zone of walking trajectory t , t_{prev} , which is a partial trajectory from reference point c_t to the start point of the j -th stable-sensing zone, and t_{next} , which is a partial trajectory from the end point of the j -th stable-sensing zone.

The process of averaging based on the first stable-sensing zone is explained using Fig. 8 and Fig. 9. In Fig. 8, for each walking trajectory of t_1, t_2, t_3 , the reference point is expressed as a yellow circle, and the first stable-sensing zone is expressed as a red line.

Here, these stable-sensing zones are along Y axis. Therefore, we find by median value of X axis of each stable-sensing zones. In the example, the median value is 2.5 and the trajectory is t_2 . When stable-sensing zones are along X axis, we find by median value of Y axis of each stable-sensing zones. Consequently, stable-sensing zone of t_2 is regarded as median stable-sensing zone t_{median} .

Next, we calculated vertical distance d between reference point c_t and line segment t_{stable} of the stable-sensing zone. Also, we calculated vertical distance d_{median} between reference point c_t and line segment t_{median} of the stable-sensing zone. Then, expansion/contraction factor a is calculated as d_{median}/d .

Based on proper expansion/contraction factor a , the t_{prev} of each trajectory is expanded/contracted in the perpendicular direction of c_t and t_{stable} . Let t'_{prev} be the partial trajectory after

expansion and contraction.

Next, a parallel shift is performed so that the partial trajectory after the expansion/contraction and the subsequent partial trajectory are connected continuously. t_{stable} is shifted so that the start point of t_{stable} is the same as the end point of t'_{prev} . Let t'_{stable} be the trajectory after the parallel shift. Also, t_{next} is shifted so that the start point of t_{next} is the same as the end point of t'_{stable} . Let t'_{next} be the trajectory after the parallel shift. Trajectory t is replaced with the averaged trajectory represented by t'_{prev} , t'_{stable} , t'_{next} obtained by the above process.

Finally, the reference points are updated. In the processing up to this point, averaging was done for the interval from t'_{prev} to t'_{stable} . Therefore, reference point c_t is updated to the coordinates of the end point of t'_{stable} . This completes the process that focused on the first stable-sensing zone. Then next stable-sensing zone is focused on for the averaging process.

Assuming that the number of associated stable-sensing zones of route i is o_i , the above averaging processing is sequentially processed by tracing the IDs of the stable-sensing zones from 1 to o_i . Let t'' be a trajectory obtained by averaging t' , where the set of the averages of the walking trajectories with route ID of i is $T''^i = [t''_1, t''_2, \dots, t''_{m_i}]$. Among these elements, walking trajectory t''^i with the largest ratio of stable-sensing zones is added as an element of final estimated walking trajectory $T''_{reliable}$. The set of walking trajectories that are finally obtained is expressed as $T''_{reliable} = [t''^1, \dots, t''^n]$.

We adopted a median for correcting trajectories. If we use mean value, then one outlier has a large effect on the mean value. Median is robust for outliers rather than mean value. The formula for calculating the median changes was based on the number of correspondence relations. For odd numbers, we used the median. With even numbers, we used the average values of the two middle numbers.

Processing separated the coordinate data up to the start coordinate of the stable-walking zone and the subsequent coordinate data because the influence of correcting the large error of part of the trajectory should not affect the entire trajectory. Pedestrian activities in the building consist of such actions as walking straight, climbing up and down stairs, and turning left and right. Although PDR captures these actions to estimate a trajectory, the estimation accuracy may vary greatly depending on the type of action. For example, our PDR algorithm [9] has a characteristic where the estimation error of the moving distance increases when the direction largely changes in a short time like half-turn stairs. If the scaling for correcting the error of the trajectory of the half-turn stairs is applied to the entire trajectory, a part of the trajectory other than the half-turn stairs is also deformed. Therefore, the trajectory correction is limited to a narrow range.

4. Evaluation Experiments

4.1 Configuration

We experimentally evaluated the improved accuracy of our proposed method with coordinate data from the 3D pedestrian trajectories estimated by our previously proposed method [9] with the HASC-IPSC [16] indoor pedestrian sensing corpus. The data used this time include the walking data of 109 routes. In addition,

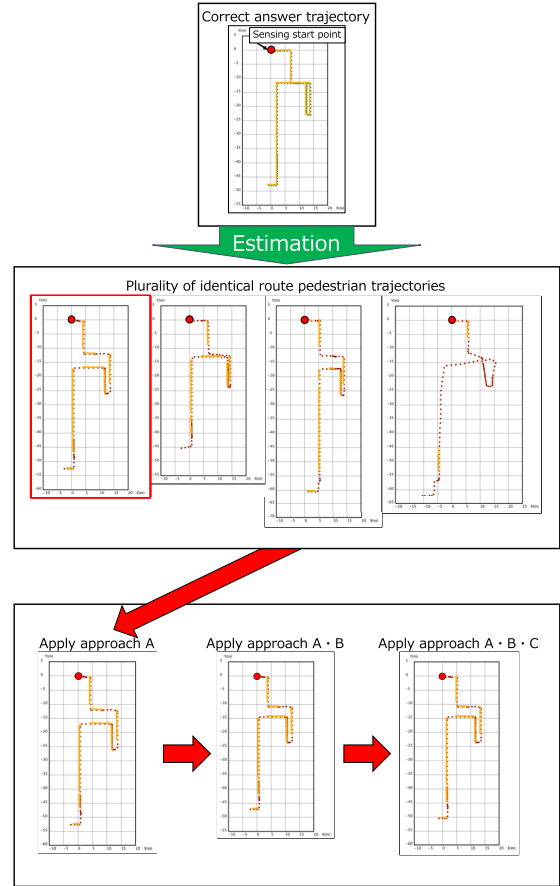


Fig. 10 Correction of pedestrian trajectory.

it also included 424 bits of walking data from multiple pedestrians who walked on each route. The baseline trajectory set is T (Section 3.1). We applied the proposed method to these data. The selection of the reliable trajectories based on the stability of the sensing data is defined as Approach A. Trajectory set $T_{reliable}$ generated in Section 3.1 is the dataset of Approach A. The correction of the trajectory length by analyzing the trends of the step length is defined as Approach B. Trajectory set $T'_{reliable}$ generated in Section 3.2 is the dataset of Approach B. The generation of the average trajectories is defined as Approach C. Trajectory set $T''_{reliable}$ generated in Section 3.3 is the dataset of Approach C. Fig. 10 shows part of the pedestrian trajectories, the estimated pedestrian trajectories, and the highly accurate pedestrian trajectories for this evaluation experiment.

We investigated the relationship between the elapsed time and the estimated position error to evaluate the accuracy of the estimated pedestrian trajectories [18]. The amount of change in the relative position in PDR was obtained by the change in the movement distance, the orientation, and the height. Investigation by distance can evaluate the moving distance and height. However, since fairly evaluating the orientation is impossible, a previous study used time for its evaluations. So we adopted this evaluation method. The results are shown in Figs. 11, 12, 13, and 14. These scatter plots are the Euclidean distance plots of the estimated errors of the positions based on comparing the coordinates of each pedestrian trajectory estimated every second with the correct coordinates (Fig. 15). If the value of Y is lower, it shows bet-

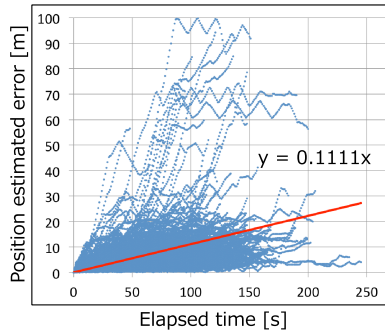


Fig. 11 Accuracy of estimated pedestrian trajectory (Baseline).

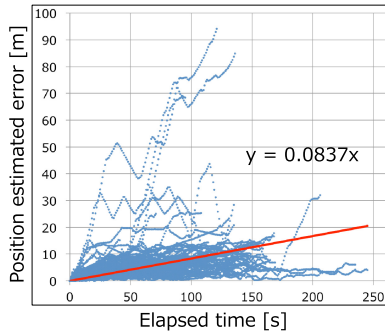


Fig. 12 Accuracy of estimated pedestrian trajectory (Approach A).

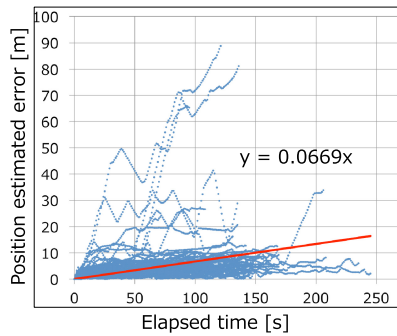


Fig. 13 Accuracy of estimated pedestrian trajectory (Approaches A and B).

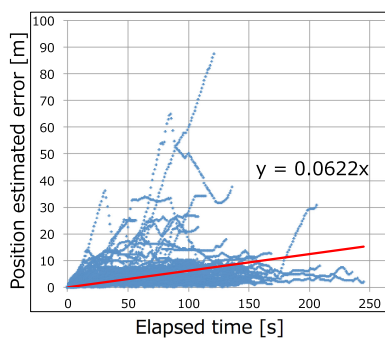


Fig. 14 Accuracy of estimated pedestrian trajectory (Approaches A, B, and C).

ter performance. We also identified a regression line that passes through the point of the origin and obtained the straight line's slope to calculate the cumulative error rate. The accuracy of the estimated pedestrian trajectories has improved if the slope of the straight line is small.

The target accuracy was set to 0.0278 m/s as the accuracy used for indoor navigation that is one of the most promising ILBSs. This target accuracy is the accuracy when the allowable error is

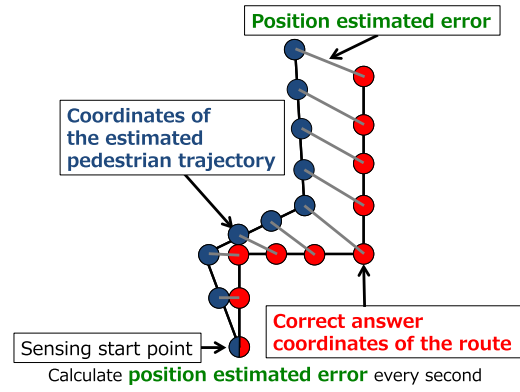


Fig. 15 Position estimated error.

considered to be 5 m assuming the travel time in the building to be about 180 seconds. The reason for setting the allowable error to be 5 m this time is because the accuracy of “available for navigation and so on as location information” in the reference document is within 10 m [19]. Also, we thought that the allowable error during navigation needs to be further improved when considering the accuracy used for indoor navigation. Therefore, the allowable error was set to about 5 m. In guidance display targeting pedestrians, the movement speed of pedestrians is generally set at 80 m/min [20]. Furthermore, Lamberg's research [21] clarified that the movement speed decreased by about 30% when pedestrians were using smartphones. From these points, it is considered that the movement speed during indoor navigation is about 56 m/min. Therefore, the travel time is about 5 seconds even if there is a position error of about 5 m. So we regarded it as allowable as an error.

4.2 Evaluation Results

We evaluated four types of estimated data and compared their accuracy with our target accuracy. For the first kind of data, we used the data of all the routes provided by our previous research [9] and defined the estimated accuracy when using all the route data as the baseline. The cumulative error rate of the baseline was 0.1111 m/s (Fig. 11).

As the second kind of data, we used the data to which Approach A was applied to the baseline. The estimation accuracy of the pedestrian trajectories is influenced by the sensor's stability during the data collection with PDR. We calculated the reliability of the estimated trajectory by Approach A and obtained the most accurate pedestrian trajectory among the multiple pedestrian trajectories of identical routes. The cumulative error rate at this time was 0.0837 m/s (Fig. 12).

As the third kind of data, we used the data to which Approaches A and B were applied to the baseline. One problem is that the walking distance tends to be biased in the walking distance estimation in PDR. To solve for it, we investigated the step length tendency using a large amount of indoor pedestrian trajectories and modified the walking distance based on Approach B to correct the walking distance. The cumulative error rate at this time was 0.0669 m/s (Fig. 13).

As the fourth kind of data, we used the data to which Approaches A, B and C are applied to the baseline. A problem is

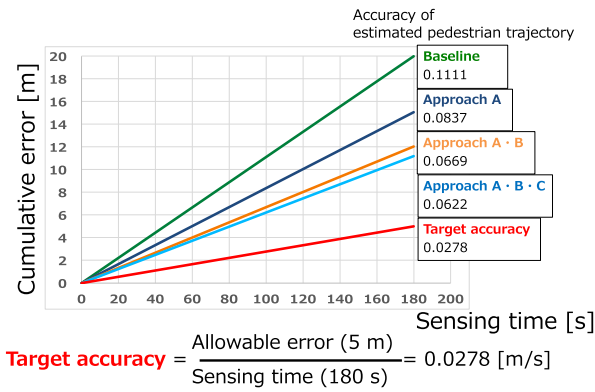


Fig. 16 List of accuracy of estimated pedestrian trajectory.

that the drift error gradually accumulates from the start of sensing in the PDR. This value also changes randomly depending on the type of angular velocity sensor and the device's heat. We thought that the drift would be averaged by Approach C like in the central limit theorem. In Approach C, the trajectory was corrected to align the positional relationship between the stable-walking zone of walking on the same straight-line route. We corrected the positional relationship of the straight-line route using multiple pedestrian trajectories of identical routes. The cumulative error rate was 0.0622 m/s (Fig. 14).

4.3 Discussion

Figure 16 shows the results of comparing these accuracies with the target accuracy. We found that the accuracy of the estimated pedestrian trajectories was improved by selecting reliable trajectories based on the stability of the sensing data, the correction of the trajectory length by analyzing the step length trends, and the generation of average trajectories rather than using the estimated pedestrian trajectories.

The cumulative error rate was about 2.2 times when we compared the final and target accuracies because we included those trajectories with low reliability when we averaged them. This time, we improved the accuracy using the multiple pedestrian trajectories of identical routes in Approach C for the most reliable trajectory selected by Approach A. The averaging included the low reliability trajectories. Therefore, the accuracy of the pedestrian trajectories became one reason why we failed to reach the target accuracy. Another cause is that we failed to eliminate those pedestrian trajectories with large error by our three proposed approaches. The data used for the evaluation experiment include pedestrian trajectories with large error due to a deviation in the traveling direction at the start of sensing. We used the value of the angular velocity sensor to estimate the traveling direction by PDR. The data used this time where the sensing started from an upright state with the travel direction were specified at the start of the sensing. Since the direction of the body was slightly inclined, the direction of the travel deviated due to the body's twist at the sensing's start. A trajectory with large error was included even if a trace with high reliability was selected. For these reasons, the accuracy of the estimated pedestrian trajectory must be improved to achieve the target accuracy. We think that Approach C will improve the accuracy quite well if there is sufficient data to generate

the average trajectory.

5. Conclusion

We propose a method to improve the accuracy of indoor PDR trajectories using many trajectories by selecting reliable trajectories based on the stability of the sensing data. For the selected trajectories, we corrected the step lengths using the actual route lengths and the walking distances of the estimated pedestrian trajectories and the generated average trajectories for each route.

From an evaluation experiment using an indoor pedestrian sensing corpus called HASC-IPSC, our proposed method improves the trajectory accuracy. After adapting it, the cumulative error rate improved to 0.0622 [m/s]. However, since achieving our target accuracy was impossible, a method that improves accuracy even further is needed.

Future work will improve the accuracy of our estimated pedestrian trajectory. In this paper, we select the most reliable trajectory in each identical route and correct it without investigating the reliability of the same-route walking trajectories. We averaged the trajectories even when their reliability was as low as 20%. If such distorted trajectories were included, they would greatly affect the pedestrian trajectory estimation and correction. The trajectory's accuracy will be further improved if we only correct the pedestrian trajectories using those whose reliability exceeds a particular threshold.

References

- [1] Kamisaka, D., Muramatsu, S., Iwamoto, T. and Yokoyama, H.: Design and Implementation of Pedestrian Dead Reckoning System on a Mobile Phone, *IEICE Trans. Information and Systems*, Vol.94, No.6, pp.1137–1146 (2011).
- [2] Kourogi, M., Sakata, N., Okuma, T. and Kurata, T.: Indoor/Outdoor Pedestrian Navigation with an Embedded GPS/Rfid/Self-contained Sensor System, *The 16th International Conference on Advances in Artificial Reality and Tele-Existence*, pp.1310–1321 (2006).
- [3] Ban, R., Kaji, K., Hiroi, K. and Kawaguchi, N.: Indoor Positioning Method Integrating Pedestrian Dead Reckoning with Magnetic Field and WiFi Fingerprints, *Proc. 8th International Conference on Mobile Computing and Ubiquitous Networking (ICMU2015)*, pp.169–174 (2015).
- [4] Gao, R., Zhao, M., Ye, T., Ye, F., Wang, Y., Bian, K., Wang, T. and Li, X.: Jigsaw: Indoor Floor Plan Reconstruction via Mobile Crowdsensing, *The 20th Annual International Conference on Mobile Computing and Networking (MobiCom2014)*, pp.249–260 (2014).
- [5] Jiang, Y., Yun, X., Pan, X., Li, K., Lv, Q., Dick, P.R., Shang, L. and Hannigan, M.: Hallway Based Automatic Indoor Floorplan Construction Using Room Fingerprints, *The 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp2013)*, pp.315–324 (2013).
- [6] Hardegger, M., Troster, G. and Roggen, D.: Improved ActionSLAM for Long-term Indoor Tracking with Wearable Motion Sensors, *International Semantic Web Conference (ISWC2013)*, pp.1–8 (2013).
- [7] Alzantot, M. and Yousef, M.: CrowdInside: Automatic Construction of Indoor Floorplans, *The 20th International Conference on Advances in Geographic Information Systems (SIGSPATIAL'12)*, pp.99–108 (2012).
- [8] Shin, H., Chon, Y. and Cha, H.: Unsupervised Construction of an Indoor Floor Plan using a Smartphone, *IEEE Trans. Systems, Man, and Cybernetics, Part C: Applications and Reviews*, Vol.42, No.6, pp.889–898 (2012).
- [9] Kaji, K. and Kawaguchi, N.: Estimating 3D Pedestrian Trajectories using Stability of Sensing Signal, *Proc. 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN2016)*, pp.1–8 (2016).
- [10] Sugimoto, S., Ito, N., Naito, K., Chujo, N., Mizuno, T. and Kaji, K.: Partial Matching Estimation Method of Walking Trajectories for Generating Indoor Pedestrian Networks, *International Conference on Mobile Computing and Ubiquitous Networking (ICMU2018)*, pp.1–6 (2018).

- [11] Jin, Y., Toh, H.S., Soh, W.S. and Wong, W.C.: A robust dead-reckoning pedestrian tracking system with low cost sensors, *Proc. 2011 IEEE International Conference on the Pervasive Computing and Communications (PerCom)*, pp.222–230 (2011).
- [12] Lan, K.C. and Shih, W.Y.: Using simple harmonic motion to estimate walking distance for waist-mounted PDR, *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pp.2445–2450 (2012).
- [13] Niino, S., Nakamura, S., Suzuki, M. and Komatsu, T.: Average handwritten shapes are also beautiful, *Entertainment Computing 2015 (EC2015)*, pp.469–478 (2015) (in Japanese).
- [14] Murata, Y., Kaji, K., Hiroi, K. and Kawaguchi, N.: Pedestrian Dead Reckoning Based on Human Activity Sensing Knowledge, *The 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp2014)*, pp.797–806 (2014).
- [15] Yotsuya, K., Iwase, H., Ito, N., Naito, K., Chujo, N., Mizuno, T. and Kaji, K.: Detection of Half-turn Stairs from Walking Trajectories Estimated by Pedestrian Dead Reckoning, *International Conference on Mobile Computing and Ubiquitous Networking (ICMU2017)*, pp.61–62 (2017).
- [16] Kaji, K., Watanabe, H., Ban, R. and Kawaguchi, N.: HASC-IPSC: Indoor Pedestrian Sensing Corpus with a Balance of Gender and Age for Indoor Positioning and Floor-plan Generation Researches, *International Workshop on Human Activity Sensing Corpus and Its Application (HASCA2013)*, pp.605–610 (2013).
- [17] Rosenblatt, M.: A CENTRAL LIMIT THEOREM AND A STRONG MIXING CONDITION, *Proc. National Academy of Sciences of the United States of America*, Vol.42, No.1, pp.43–47 (1956).
- [18] Abe, M., Kaji, K., Hiroi, K. and Kawaguchi, N.: PIEM: Path Independent Evaluation Metric for Relative Localization, *Proc. 7th International Conference on Indoor Positioning and Indoor Navigation (IPIN2016)*, pp.1–8 (2016).
- [19] Geospatial Information Authority of Japan: Overview of “Guidelines about installing BLE beacons for indoor positioning”, available from <http://www.gsi.go.jp/common/000198739.pdf> (accessed 2019-03) (in Japanese).
- [20] Real Estate Fair Trade Council of Federation: Fair competition agreement on display of real estate, available from <http://www.rftc.jp/webkanri/kanri/wp-content/uploads/2018/08/h.kiyaku.pdf> (accessed 2019-03) (in Japanese).
- [21] Lamberg, M.E. and Muratori, M.L.: Cell phones change the way we walk, *Gait & Posture*, Vol.35, No.4, pp.688–690 (2012).

Appendix

A.1 Pseudo-code of the Proposal Method

Algorithm 1 Selection of Reliable Trajectories (Section 3.1)

Require:

- n : The number of identical route
- i : Route ID
- m_i : The number of trajectories of route i
- t_j^i : The j -th trajectory of route i
- T : A set of all trajectories

Ensure:

```

 $T_{reliable}$ : A set of the most reliable trajectories for each identical route
function LENGTH_WHOLE( $t$ )
    return length of trajectory  $t$ 
end function
function LENGTH_STABLE( $t$ )
    return total length of stable-sensing zones in trajectory  $t$ 
end function
function RELIABILITY( $t$ )
    return length_stable( $t$ )/length_whole( $t$ )
end function
procedure SELECT_RELIABLE_TRAJECTORIES
    for all route  $i$  do
        pick up all trajectories of route  $i$  from  $T$ 
        find  $j$  that reliability( $t_j^i$ ) becomes the largest
        put  $t_j^i$  into  $T_{reliable}$ 
    end for
end procedure

```

Algorithm 2 Correction of Trajectory Length (Section 3.2)

Require:

- a : Scaling factor
- s : Partial trajectory of stair section
- o : Partial trajectory other than stairs

Ensure:

```

 $T'_{reliable}$ : A set of processed trajectories in  $T_{reliable}$ 
function SCALING( $t, a$ )
    for all coordinates  $(x, y, z)$  in  $t$  do
        replace the coordinates to  $(a \cdot x, a \cdot y, z)$ 
    end for
    return scaled trajectory
end function
function PARALLEL_SHIFT( $t_1, t_2$ )
    parallelly shift  $t_2$  so that the end coordinates of  $t_1$  and the start coordinates of  $t_2$  match
    return shifted trajectory
end function
procedure CORRECT_TRAJECTORY_LENGTH
    for all trajectory  $t$  in  $T$  do
        divide  $t$  into a sequence of partial trajectories  $(o_1, s_1, \dots, s_k, o_{k+1})$ 
        for all partial trajectory  $c_i$  in  $t$  do
            if  $c_i$  is a partial trajectory other than stairs then
                 $c_i = \text{scaling}(c_i, a)$ 
                 $c_i = \text{parallel\_shift}(c_{i-1}, c_i)$ 
            else
                 $c_i = \text{parallel\_shift}(c_{i-1}, c_i)$ 
            end if
        end for
        put modified trajectory  $t$  into  $T'$ 
    end for
    for all trajectory  $t$  in  $T_{reliable}$  do
        same process as above
        put modified trajectory  $t$  into  $T'_{reliable}$ 
    end for
end procedure

```

Algorithm 3 Generation of Average Trajectories (Section 3.3)**Require:**

c_i : Reference point of trajectory t
 C : A set of reference points of trajectories
 o_i : Number of stable-sensing zones in i -th route
 t_{median} : Median stable sensing zone
 a : Scaling factor of expansion/contraction

Ensure:

$T''_{reliable}$: A set of processed trajectories in $T'_{reliable}$

function GET_PARTIAL_TRAJECTORIES(t, j)

pick up c_i from C

generate t_{stable} which is a partial trajectory of j -th stable-sensing zone of trajectory t

generate t_{prev} which is a partial trajectory from reference point c_i to the start point of j -th stable-sensing zone

generate t_{next} which is a partial trajectory from the end point of j -th stable-sensing zone to the end of the trajectory

return [$t_{stable}, t_{prev}, t_{next}$]

end function

function DISTANCE(c, t_{stable})

return vertical distance between c and the line segment t_{stable}

end function

function EXPAND_CONTRACT($c_i, t_{stable}, t_{prev}, t_{median}$)

$a = \text{distance}(c_i, t_{median}) / \text{distance}(c_i, t_{stable})$

Return expanded/contracted t_{prev} using scaling factor a in the perpendicular direction of c_i and t_{stable}

end function

function GET_MEDIAN_STABLE_SENSING_ZONE(T'^i, j)

define S as a set of j -th stable-sensing zones of each trajectory in T'^i

calculate each distance of stable-sensing zones in S

find median stable-sensing zone t_{median} according to the distances

Return t_{median}

end function

procedure GENERATE_AVERAGE_TRAJECTORIES

for all trajectory t in T' **do**

c_i = starting point of t

put c_i into C

end for

for all route i **do**

generate T'^i that is a set of all trajectories of i -th route from T'

for $num=1$ to o_i **do**

$t_{median} = \text{get_median_stable_sensing_zone}(T'^i, num)$

for all trajectory t in T'^i **do**

[$t_{stable}, t_{prev}, t_{next}$] = get_partial_trajectories(t, num);

pick up c_i from C

$t'_{prev} = \text{expand_contract}(c_i, t_{stable}, t_{prev}, t_{median})$

$t'_{stable} = \text{parallel_shift}(t'_{prev}, t_{stable})$

$t'_{next} = \text{parallel_shift}(t'_{stable}, t_{next})$

t = sequence of partial trajectories ($t'_{prev}, t'_{stable}, t'_{next}$)

c_i = coordinates of the end point of t'_{stable}

end for

end for

find j that reliability(t'_j) becomes the largest from T'^i

put t'_j into $T''_{reliable}$

end for

end procedure

Editor's Recommendation

This paper proposes a method to improve the accuracy of indoor Pedestrian Dead Reckoning (PDR) trajectories by using a large number of trajectories. The proposed method extracts reliable trajectories for step length correction and generation of average trajectories. The method greatly mitigates trajectory error, which is also useful in various actual environments. Therefore, the paper is selected as a recommended paper.

(Chief examiner of SIGMBL Nobuo Kawaguchi)



Kosuke Yotsuya is a graduate student at Graduate School of Business Administration and Computer Science, Aichi Institute of Technology. His research focuses on indoor positioning. He is also interested in the Internet of Things.



Katsuhiro Naito received his B.S. degree in Electronics Engineering from Keio University, Japan in 1999, and received his M.S. and Ph.D. degrees in Information Engineering from Nagoya University, Japan in 2001 and 2004, respectively. From 2004 to 2014, Dr. Naito was an assistant professor in the electrical and electronic engineering department of Mie university. He was a visiting scholar in the computer science department of University of California, Los Angeles (UCLA) in 2011. Since 2014, he has been an associate professor in the information science department of Aichi Institute of Technology. His research interests include 5G technologies, vehicular communication systems, Internet of Things (IoT) and Machine to Machine (M2M) systems, overlay networks, and network protocols and architectures.



Naoyuki Chujo received his B.E. degree in applied physics and his M.S. degree in information science and his Ph.D. degree in electrical engineering from Nagoya University in 1980, 1982 and 2004. He joined Toyota Central R&D Labs. in 1982. He has been a professor at Aichi Institute of Technology since 2010. His research interests are in the area of embedded system and automotive electronics. He is a member of IEEE, IPSJ, IEICE, IEEJ, and Informatics Society.



Tadanori Mizuno received his B.E. degree in Industrial Engineering from the Nagoya Institute of Technology in 1968 and received his Ph.D. degree in Computer Science in 1987 from the Kyushu University in Fukuoka. In 1968, he joined Mitsubishi Electric Corporation in Kamakura. From 1993 to 2011, he had

been a Professor of Informatics at Shizuoka University in Hamamatsu. From 2011 to 2016, he had been a Professor of Information Science, and since 2016, he is an Affiliate Professor of Information Science at the Aichi Institute of Technology in Toyota. His research interests include mobile computing, distributed computing, computer networks, broadcast communication and computing, and protocol engineering. He is a member of Information Processing Society of Japan, the Institute of Electronics, Information and Communication Engineers, IEEE, and Informatics Society.



Katsuhiko Kaji received his Ph.D. in information science from Nagoya University in 2007. He became a RA at NTT Communication Science Laboratories in 2007 and an assistant professor in Nagoya University in 2010. Currently, he is associate professor of Faculty of Information Science, Aichi Institute of Technol-

ogy from 2015. His research interests include indoor positioning and remote interaction. He is a member of IPSJ and JSSST.