2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN), 24-27 September 2018, Nantes, France

Review of PDR Challenge in Warehouse Picking and Advancing to xDR Challenge

Ryosuke Ichikari[1], Ryo Shimomura[12], Masakatsu Kourogi[1], Takashi Okuma[1], and Takeshi Kurata[123]
[1]Human Informatics Research Institute, AIST, Ibaraki, Japan
[2]University of Tsukuba, Ibaraki, Japan
[3]Sumitomo Electric Industries, Ltd., Osaka, Japan

***Abstract*— In this paper, we present a review of the competition track organized by us at IPIN 2017. Our competition track was named as "PDR challenge in warehouse picking". Our pedestrian dead-reckoning (PDR) challenge focused on testing the PDR methods under a realistic scenario, which is tracking employees during the picking operation in a warehouse. Firstly, we introduce an overview of the PDR challenge including detailed information of all evaluation metrics. Secondly, we summarize the results and findings from the competition and list the challenges that should be tackled in the next competition. Thirdly, we introduce the design concept of an xDR challenge by reflecting the findings and feedback of design of the competition. The target of the PDR challenge was only the employees who move by foot during working in the warehouse.. The main update in the xDR challenge is dealing with forklifts driven by employees as well as walking employees. We term dead-reckoning of a vehicle as vehicle dead-reckoning (VDR). The term xDR is derived from PDR plus VDR. Another update is providing reference data for the reconstructing situation of picking operation in the warehouse by using an IMU-based motion capture system.**

Keywords—PDR challenge; xDR challenge; competitions; pedestrian dead-reckoning; vehicle dead-reckoning; motion capture

I. INTRODUCTION

Owing to the advances of the indoor localization and Internet of Things (IoT) technologies, we have many choices for indoor localization such as pedestrian dead-reckoning (PDR) based, Bluetooth low-energy based, Wi-Fi based, and ultrasonic signal based methods. Although, the variety of choices increases the freedom in selection of the methods, the characteristics of the methods are sometimes not clear enough for distinguishing between them, especially for technology users or integrators who are not experts of indoor localization.

PDR is one of the relative tracking methods mainly used for indoor localization. Comparing PDR methods is very difficult due to the nature of relative tracking. Performance of an integrated location system with PDR depends on many aspects such as accuracy of the relative tracking by PDR, accuracy of the initial position and direction, and accuracy of the error correction methods for reducing accumulated error.

In order to fairly evaluate existing localization methods, tracking competitions have been held by preparing shared data or controlled environments for fair comparison. Similar to the camera tracking competitions in the research community on the subjects of augmented reality (AR) and mixed reality (MR) [1], there are two types of competitions in the research community on indoor localization. One type of competition is an on-site competition. The organizers of this type of competition prepare controlled environments for testing that are measured carefully in advance. The competitors actually run their algorithm or system in the testing environment. The other type is an off-site competition wherein, instead of physically testing in the test environment, the organizers prepare test data, which are shared with the competitors. The competitors apply their algorithms or methods to the test data and submit the results to the organizers. Microsoft have organized tracking competitions at the International Conference on Information Processing in Sensor Networks (IPSN) since 2014 [2]. The target of the method was not limited to PDR or localization methods on smartphones. NIST also organized a PerfLoc competition [3]. The IPIN/EvAAL competitions are also traditionally held on-site/off-site during the conference [4][5]. The competitions of the IPIN were more focused on PDR or methods without relying on infrastructures. In the International Symposium of Wearable Computing (ISWC) /UbiComp 2015 PDR challenge was first held for competing PDR methods by gathering actual performance data measured under a specific scenario [6]. The scenario of the first PDR challenge was indoor navigation with a smartphone. We decided to hold another PDR challenge by adhering to the original concept and name with due permission in 2017. Our PDR challenge was named as "PDR challenge in Warehouse Picking", which is also called "PDR Challenge 2017" for short.

We believe these competitions should be continually held for putting the potential users on notice and encouraging many competitors to join the competition. In 2018, we will hold another competition as a sequel to the PDR challenges. We name new competition as "xDR challenge for warehouse operations 2018" and "xDR challenge" for short.

One of the important updates in the xDR challenge over the PDR challenge 2017 is that, the xDR challenge adds dead-reckoning of forklifts as the tracking target of the competition. In order to help the competitors learn what warehouse operations look like, we prepare motion capture data as reference data. This data can reconstruct an employee's whole body posture during typical warehouse operations. Furthermore, scale of the data used for the competition will be extended in terms of the number of employees and time duration.

978-1-5386-5635-8/18/$31.00 ©2018 IEEE

In this paper, we summarize the PDR challenge 2017 and the upcoming xDR challenge as an example approach for open innovation, which aims to promote development of an indoor tracking method with a realistic scenario in the industry.

II. FINDINGS FROM PDR CHALLENGE IN WAREHOUSE PICKING

A. Overview of PDR Challenge in Warehouse Picking

The PDR challenge in warehouse picking (PDR challenge 2017) was held as an official tracking competition during the IPIN 2017 conference. The purpose of the PDR challenge 2017 was to evaluate PDR based indoor localization methods under a realistic scenario. The PDR challenge 2017 was adhered to the original PDR challenge held at UbiComp/ISWC 2016. The assumed scenario of the original PDR challenge was indoor navigation. On the contrary, the assumed scenario of the PDR challenge 2017 was tracking employees during the picking operation in the warehouse.

The data used for the PDR challenge 2017 were measured during the actual picking operation in an actual warehouse. Therefore, the measured behaviors are not limited to walking. They included other operations than walking, such as pushing carts, picking objects, driving forklifts, and using handy terminals. The duration of the data measurement was approximately three hours. However, the competitors were able to grasp the entire data with unveiled correct reference points and apply global optimization algorithm to them for improving accuracy of the localization.

The sensor data were measured by having employees carry a smartphone in a pouch. The logging application installed in smartphones recorded the following data.

- Raw sensor data for PDR: angular velocity, acceleration, magnetism, and atmospheric pressure (at 100 Hz)
- Received signal strength indicator (RSSI) from BLE beacons (at 2 Hz)

Additionally, we shared the picking log data from the warehouse management system (WMS). The WMS is often operated in warehouses for managing items and displaying work instructions for employees. In the case of a target warehouse, the barcodes are attached on each shelf, and the employees scan the barcodes with handy terminals for confirming items according to the WMS's instruction. We utilize the logs of the barcode scan at each shelf as the ground truth of employee's positions. As subsidiary information about the warehouse, we share detailed layout of the warehouse (Fig. 1.), correspondence table of shelf IDs and their locations, and locations of BLE beacons.

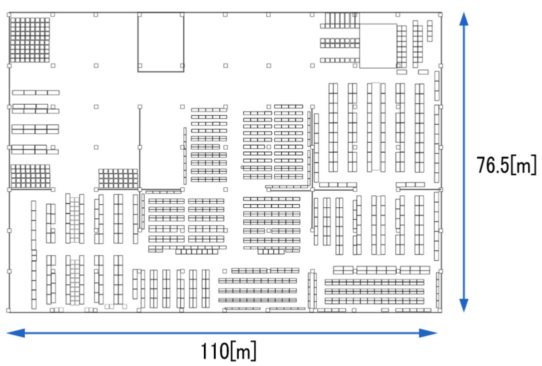

Fig. 1. Layout of the warehouse

Some parts of the ground truth of location from WMS are shared for correcting the error of localization. The rest of ground truth data are hidden and utilized for evaluating submitted trajectories from the competitors. The ratio between hidden and unveiled ground truth can control the difficulty of integrated localization, because it can change frequency of availability of the unveiled ground truth. The competitors can utilize WMS picking logs and RSSIs of BLE beacons and other information for improving accuracy of employees' location. Therefore, this competition can be regarded as the competition for integrated localization method based on PDR.

B. Results of PDR Challenge 2017

1) Evaluation metrics for PDR Challenge 2017

The PDR Challenge 2017 aimed to be a valuable competition of PDR, which is a relative tracking method in a realistic situation. In order to evaluate the relative tracking method, relative accuracy and error accumulation other than absolute accuracy need to be considered. The evaluation metrics of the PDR Challenge 2017 consist of not only the positional error, but also many possible evaluation metrics required for the warehouse scenario. In the realistic scenario, an ideal method should simultaneously meet comprehensive metrics such as positional error metric and obstacle interference metric. We added subsidiary metrics such as frequency metric for increasing practicality of the submitted trajectories. According to the above concept, we defined following six elemental evaluation metrics categorized into three namely, metrics related to accuracy, trajectory naturalness, and special metrics for warehouse picking scenario, as shown in Fig. 2. The elemental metrics are weighted summed for a comprehensive evaluation for reflecting all evaluation metrics.

Evaluation metric for absolute positional error (E_d)

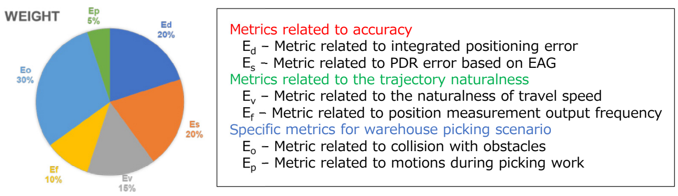

Fig. 2. Evaluation metrics and their weights\

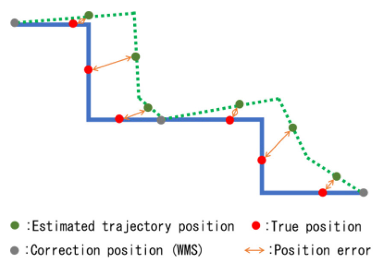

Fig. 3. Conceptual image of evaluation of positional error

This metric is an error metric of absolute positional error. Distance between the estimated and true position is measured at all checking points when the ground truth positions are available from WMS as shown in Fig. 3. We calculate median of the errors in the checking points. Detailed equation of E_d is shown in Appendix A.

Evaluation metric for error accumulation of PDR (E_s)

This metric indicates speed of error accumulation caused by PDR. It is known that, error accumulation is one of the main concerns for PDR. We partially provided the ground truth of the target's positions for error correction. All competitors can correct error to 0 at the correction points. We assume that the error increases as time proceeds from the correction points and quantify the speed of error accumulation from the correction points. As shown in Fig. 4, we adopt linear regression between the time from the closest correction points and error; taking inspiration from Abe et.al. [7]. We assume the intercept of the linear model to be 0. A detailed equation of E_s is shown in Appendix A. We term the slope of the error accumulation as error accumulation gradient (EAG).

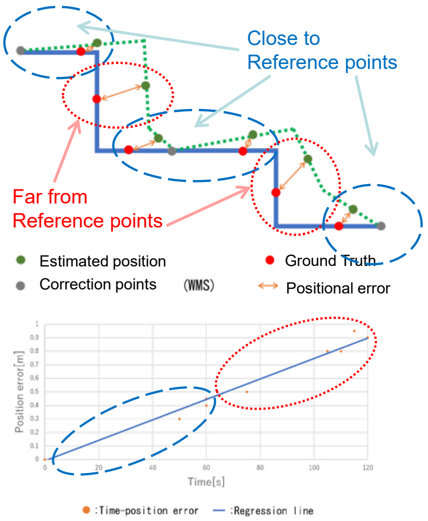

Fig. 4. Conceptual image of evaluation of error accumulation

Evaluation for naturalness during picking (E_p)

This metric is one of the dedicated metrics for warehouse operation. This metric evaluates naturalness of the trajectories during picking. As observed in the warehouse, employees stop in front of the shelves when they pick items from the shelves. Therefore, this should be reflected in the estimate at the time of picking. We check whether the target of the trajectories stop or not with this metric. More specifically, we check the total length of the movement measured from 1.5 sec before the picking to 1.5 sec after the picking. If the movement is more than 1 meter, index of this metric is deducted. The final E_p is calculated as the percentage of the checking points without the movement out of all the checking points. The detailed equation for E_p is shown in Appendix A.

Evaluation for velocity naturalness (E_v)

This metric checks whether the velocity in the trajectory is within range of natural human movements. Some algorithm may be able to accurately estimate the position with a noticeable frequent jitter. We prefer a natural and smooth trajectory. In order to deduct such artifact by a jitter, we check the velocity of all submitted points. We define 1.5 m/sec as the threshold of natural human speed. The E_v is calculated as the percentage of the checking points within the speed limit. The detailed equation for E_v is shown in Appendix A.

Evaluation for obstacle interference (E_o)

This metric checks whether the trajectories enter forbidden areas. We assume that ideal trajectories do not enter the area where employees cannot walk inside or pass over because of the existence of obstacles such as shelves, poles, and walls. Inconsistency between the trajectory and environments also affects the analysis of the trajectories such as traffic line analysis in warehouses. We introduced a metric E_o for quantifying the degree of incursion of the trajectories into the forbidden area as shown in Fig. 5. , We shared the location and size of shelves, poles, wales as reference information about the warehouse. Therefore, the competitors were informed about the forbidden areas. E_o is the ratio of the length of trajectories which enters the forbidden area. The detailed equation for calculating E_o is shown in Appendix A. Note that we defined a tolerance area with 0.17 m width around borders of the forbidden area for ignoring small amount of the incursion.

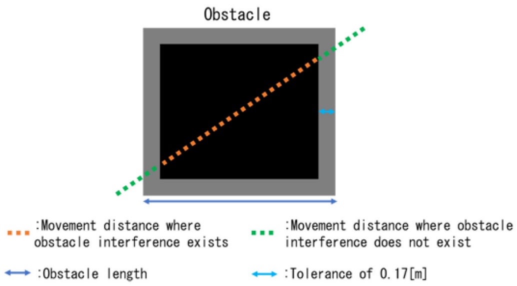

Fig. 5. Conceptual image of evaluation of obstacle interference

Evaluation for update frequency (E_f)

This metric checks the update frequencies of the trajectories submitted by the competitors. If there are two trajectories with same accuracy of the positional estimation and different update frequencies, the trajectory with the higher frequency would have a greater value than the one with a lower frequency. Given the application of localization for the employees during the work, we define the minimum update frequency as 1 Hz. To calculate the update frequency for every frame, the elapsed time from the previous submitted frame is measured for all submitted frames. The detailed equation for calculating E_f is shown in Appendix A.

Comprehensive evaluation (C.E.)

In order to rank the submitted trajectories by quantifying the performance as an integrated indicator, the elemental evaluation metrics are integrated with the following weights. We term the integrated indicator as "comprehensive evaluation" (C.E.), calculated by equation (1).

$$C.E. = 0.2E_d + 0.2E_s + 0.05E_p + 0.15E_v + 0.3E_o + 0.1E_f \quad (1)$$

We believed that the error metrics (E_d, E_s) and the metric for obstacle interference (E_o) were more important than others. Therefore, we added larger weights for them.

2) Results of Evaluation

We announced the call for participation for the competition by releasing the detailed information of the regulation and information about the measured data and target warehouse.

Consequently, we accepted applications from 5 teams, which hailed from four countries. The submitted trajectories were evaluated by using the evaluation metrics and the comprehensive metric. In this section, we discuss about the evaluation results. As a reference for the comparison, we add our result with our integrated localization method [8] based on our PDR method [9].

Tab. 1 lists the results of all evaluation metrics and the comprehensive evaluation. As seen in the column of C.E., Team2 is ranked 2nd in the table. Team2 did not receive the highest scores in all elemental evaluation metrics. However, they received the highest score for metric E_o, which has the highest weight. The competitors could calculate the evaluation metrics E_o, E_v and E_f by themselves because all information regarding the metrics was communicated to them. Team2 got almost perfect scores for those metrics. Therefore, Team2 seemed to have prepared their algorithm very well by tuning it for the regulations.

An example of the comparison of the submitted trajectories is shown in Fig. 6. The trajectory of Team2 is colored yellow. As shown in the figure, Team2's trajectories consist of straight lines on the aisles in the warehouse. Their presentation during the special session of the competition indicates that they adopt some sort of path matching algorithm. These trajectories do not look like natural human trajectories, but there is no specific metric which deducts such trajectories in the regulation. The deduction in E_p might be related to this; however, the weight for E_p was very small.

Tab. 1. Results of the competition

Team	E_d	E_s	E_p	E_v	E_o	E_f	50% eCDF [m]	50% eCDF [m/s]	C.E.
Team1	66.876	93.692	97.195	99.998	51.821	11.323	10.606	0.173	68.652
Team2	71.524	94.872	43.545	100	99.876	100	9.258	0.150	90.419
Team3	76.459	95.333	72.719	87.835	93.549	99.271	7.827	0.141	89.161
Team4	51.934	90.769	84.965	95.657	59.623	99.239	14.939	0.230	74.948
Team5	78.386	96.308	97.484	99.093	45.530	100	7.268	0.122	78.336
Team6	80.272	96.718	81.057	98.711	89.968	95.879	6.721	0.114	90.836

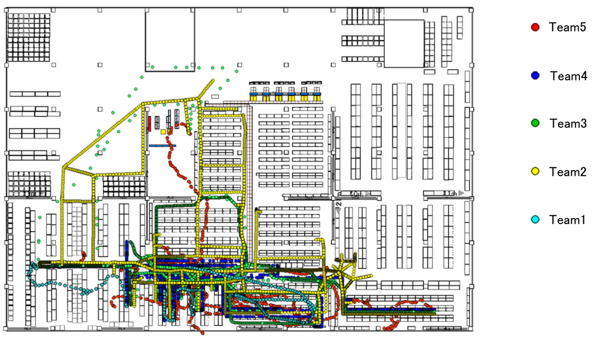

Fig. 6. Example of the submitted trajectories

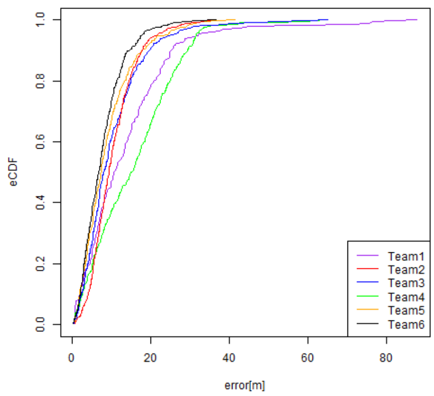

Fig. 7. eCDF for positional error

One of the most important evaluation metrics is evaluation metric for absolute positional error, given by E_d. Team6 won in this elemental evaluation. Inspired by other competition tracks in IPIN [4][5], we also generate an empirical cumulative distribution function (eCDF) of absolute error for understanding the distribution of errors. Fig. 7 shows an example of eCDF of absolute positional errors. By using eCDF, we can easily calculate 25 percentile error, 75 percentile error, and median error.

One of the unique characteristics of this competition is adoption of the evaluation metric for error accumulation of PDR as E_s. As introduced in the previous section, E_s is proportional to median of EAG. Fig. 8 shows a comparison of the slopes of error accumulation obtained by applying simple linear regression. As shown in the graph, there are many points regarded as outliers, which are the points far from the diagonal line. This result gives us an opportunity for discussing the application the simple linear regression for obtaining EAG. However, we believe that this is a good starting point for quantifying error accumulation for PDR. EAG can be utilized for planning the infrastructure for integrated localization. For example, if the given EAG is equal to 0.1 m/sec (6 m/min.) and target accuracy of the localization is less than 4 m, then we can estimate that the target accuracy can be accomplished if positional corrections by an absolute positioning method with an error less than 1 m are available at least 2 times per minute (4.0-1.0=0.1*60/2).

According to the simulation, we can control the density of the arrangement of the infrastructure. Additionally, for the infrastructure whose sampling frequency is related to power consumption, the simulation result can be a rough indication for determining the setting of the update frequency.

C. Findings from PDR Challenge 2017

The results of the PDR Challenge 2017 introduced in the previous session can be summarized as follows:

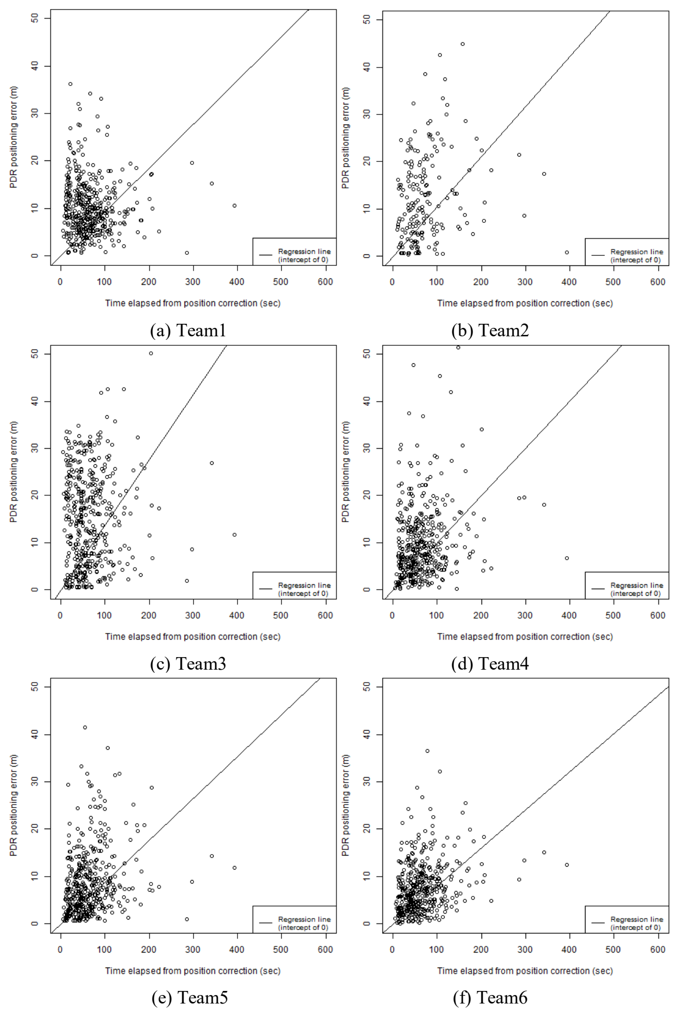

Fig. 8. Comparison of slope of error accumulation

We have successfully held a special tracking competition for warehouse picking by adopting multi-faceted evaluation metrics required for evaluating practicality of PDR based localization methods in such scenario.

We have successfully encouraged the competitors to develop practical localization methods which can fulfill the requirements for analyzing employees' trajectories in a warehouse by adopting multi-faceted evaluation metrics.

The results of scores calculated for all competitors indicate that the weight for obstacle interference (E_o) was too large. The winner of the competitions did tuning for this metric by adopting some sort of path fitting algorithm instead of simply using the result of integrated localization. However, this tuning is completely fair because the weighs and regulation had been informed in advance. In order to emphasize the naturalness of the PDR trajectories, the weight for E_o should be deducted. Moreover, other metrics and reference data for evaluating naturalness of PDR trajectories are required.

We proposed EAG for evaluating error accumulation. The EAG is not just an indicator for evaluation, but the physical meaning can also be interpreted from EAG. More specifically, the EAG is able to quantify the degree of speed of error accumulation, which is one of most important problems of

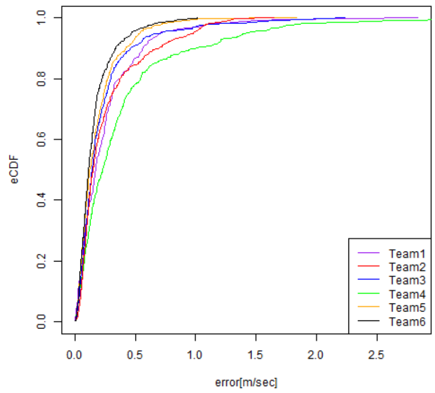

Fig. 9. eCDF for EAG

Tab. 2. Comparison of methods for determining the representative EAG

Team	Regression line	Robust regression_line	P25eCDF	P50eCDF	P75eCDF
Team1	0.106	0.128	0.069	0.173	0.311
Team2	0.139	0.169	0.095	0.230	0.441
Team3	0.089	0.100	0.064	0.122	0.240
Team4	0.100	0.107	0.077	0.141	0.262
Team5	0.093	0.106	0.085	0.150	0.309
Team6	0.081	0.087	0.062	0.114	0.186

PDR. The EAG can not only quantify the performance of PDR, but can be utilized for planning the integrated localization and error correction by other absolute positioning methods. Particularly, the EAG can be used for helping a developer in determining sampling frequency of the correction methods and/or their density of arrangement. We adopted simple linear regression for obtaining EAG. There is a room for discussion whether applying linear regression can accurately estimate EAG or not.

D. *Discussion about the findings*

As discussed in the previous sections, simple linear regression is easily affected by outliers. Additionally, we assumed that the competitors can propagate the partial ground truth in chronological order and reverse chronological order. However, it is not guaranteed that all competitors can utilize the ground truth in such a manner. If a competitor propagates it only in a chronological manner, it increases errors even close to the correction points and further increases EAG.

In order to investigate how to robustly estimate EAG, we first test linear regression with robust estimation. We adopt "rlm" function for robust linear model fitting provided by statistical computing language R, given that the intercept is 0. As other options, we can generate eCDF for EAG also. EAG can be calculated for every checking point. Fig. 9 shows an eCDF of EAG for the all checking points in all trajectories. By referring the graph of eCDF, we can easily calculate 25 percentile of EAGs; 50 percentile of EAGs, which is equal to median of EAGs; and 75 percentile of EAGs. Tab. 2 compares methods for estimating EAG in original linear regression, linear regression with robust estimation, 25 percentile of EAGs, median of EAGs, and 75 percentile of EAGs.

As seen in Tab. 2, the first ranked team in terms of EAG is always Team6 and the second ranked is Team3; however, the lower ranked teams vary in the methods. It is impossible to determine the best methods from these results only. However,

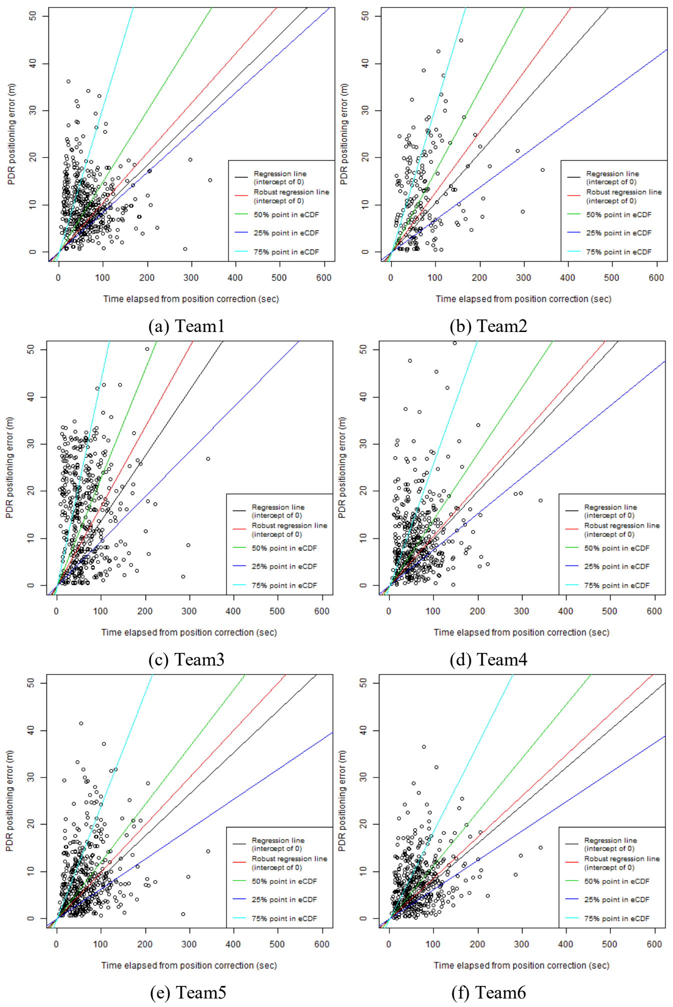

Fig. 10. Comparison of the representative EAGs with plots

Fig. 11. Teaser image for xDR Challenge

it is worth checking multiple results for estimating EAG. Fig. 10 shows the updated comparison of the methods for estimating the representative EAGs on top of the plotted data. The EAGs are varied among the methods. It seems to be better to check and compare results from the multiple methods for better understanding of the error accumulation.

During the exchange of e-mails with competitors and Q&A of the special session at IPIN 2017, we received fruitful feedback for improving the competition and its evaluation. We summarize the feedback including self-review as follows.

A team requested us to provide them sample videos of the recording of the picking work in the warehouse, because they were not familiar with the warehouse operation.

We have not implemented a fully automated system for sharing and receiving the data; and analyzing the submitted trajectories. It was possible to deal with five contestants, but automated system is required for facilitating organizers works.

There are some employees driving forklift rather than walking during the operation. PDR challenge did not deal with the data of employees who drove forklift because the methodology for tracking a forklift is different from methodologies for tracking a pedestrian. Tracking forklift has been added because tracking targets is required for future competition.

III. Design of xDR Challenge

A. *Overview of xDR challenge*

In 2018, we will hold a new competition called “xDR challenge for warehouse operations” as a sequel to the "PDR Challenge in Warehouse Picking" by reflecting the feedback and self-reviews. The PDR Challenge 2017 was targeted for tracking of walking employees during the picking operation. The xDR challenge will add forklifts driven by employees as its tracking targets. A forklift is supposed to be tracked by the dead-reckoning methodologies as well. We term dead-reckoning of the vehicle including forklifts as vehicle dead-reckoning (VDR in short). Therefore, this year’s competition is renamed as xDR challenge, which means a competition for PDR and VDR. We started promotions and uploaded the first call for participation on our web-site [10]. Fig. 11 is the teaser image that appeared on the website.

B. *Updates of xDR Challenge*

Main updates from the PDR Challenge 2017 are as follows.

Adding VDR as the competition’s target

The xDR challenge adds the tracking of the forklift as the tracking target in the competition. In general, there are many moving objects with wheels such as forklifts, hand forks, carts, dollies, and wagons. If we could track such kind of moving objects, we could track almost any kind of moving object including people. In the xDR challenge, we pick forklifts as representative moving objects with wheels, because their length of time in service seems to be the longest.

Sharing reference information about warehouse operations

In order to reconstruct the situation of the warehouse operation, we will share the motion capture data measured during the operation. We adopted Xsens’s Awinda system [11] for measuring the whole body posture. The Awinda is an IMU sensor based motion capture. Therefore, the whole body posture can be measured while working naturally and moving freely in the warehouse. The confidential information in the warehouse and privacy of the employee can be protected by sharing only the animation of the computer graphics model and whole body motion data instead of real videos captured during the operation. The competitors can understand whole body movement during the operation and can analyze the movement for specific parts of the body.

Extension of the scale of data measurement

The scale of data measurement for the PDR challenge 2017 was seven employees for about three hours in total. The scale of the data measurement is also upgraded. We measured about 30 employees and six forklifts for one week (from morning to evening, five business days).

Refinement for evaluation metrics

Data format for sharing sensor data and the evaluation metric of PDR Challenge 2017 will be adhered in the xDR challenge. The detailed regulation will be determined by discussing with PDR benchmarking standardization committees in Japan, and the competitors. Current plans for refinement for the evaluation metrics are reconsideration of determining the representative EAGs and reduction of the weight for E_o.

C. *Vehicle Dead-Reckoning*

As described in the previous session, we term the dead-reckoning of vehicles such as forklifts, carts, and dollies as VDR. In addition to PDR, VDR can increase the coverage of moving objects by dead-reckoning methodologies. This session describes technical details of VDR.

Initially we introduce the technical background of PDR, because VDR is developed as a derivative of a kind of PDR. Approaches for PDR can be roughly separated into two types according to where the sensor is attached. One type is attaching IMUs onto the foot. This type can utilize the so-called zero velocity update (ZUPT), which can correct the error of velocity estimation when feet land [12]. Another type is attaching an IMU sensor module including smartphone onto other parts or holding by hands. For estimating position and direction of the pedestrian, the common approach is called step-heading system (SHS), which detects pattern of walking from the sensor data. Here we assume that the SHS-type of PDR is used.

The elemental technologies of SHS-type PDR can be separated into four parts. As shown in Fig. 12., they are pose estimation, direction correction by magnetism, velocity estimation, and heading direction estimation. The first two items can be applied to all moving objects. The last two items are dedicated for pedestrians by considering their characteristics. In order to utilize the elemental technologies for VDR, only the last two items need to be replaced by considering the characteristics of the vehicle. For velocity estimation, we are developing a method based on the fact that high frequency pattern of the acceleration of the vehicle and velocity are linearly related. For heading direction estimation, we can assume that the angle between the heading direction and sensor attached pose is static for a vehicle.

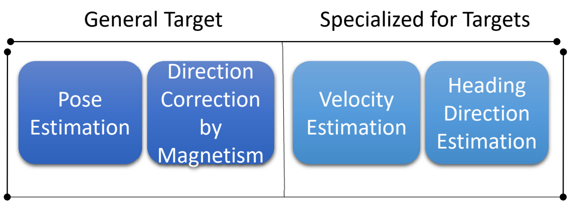

Fig. 12. Components of PDR and VDR

D. *Data Measurement for xDR Challenge*

Data measurement of the xDR challenge was conducted at an actual warehouse for a week (five business days) in the middle of March 2018. Sensor data for PDR and VDR were measured using dedicated application on smartphones. On the first day of the measurement, we conducted instruction session for all participants of the measurements and asked them to start and end the measurements by themselves during the one-week duration. In the warehouse, we arranged about 100 BLE beacons for correcting accumulated errors. The total number of forklifts that operated during the measurement was nine. The smartphones for measuring forklifts were taped on around the pillar of the forklifts. In order to check when the forklifts were operated, we attached a BLE beacon onto each forklift. The employees driving the forklifts also carried a smartphone. Therefore, the approach of the driver could be detected by checking RSSI from the BLEs mounted on the forklifts. We plan to inform model numbers and wheel base of the forklift for allowing the contestant to tune the parameters for each model of the forklift.

The whole body posture of an employee was measured as the reference data for picking operation for about two hours by using the Awinda system. The employee wears a motion capture suit with 17 small IMU sensor modules. Sensor data form the modules can be automatically converted into whole body posture by Xsens's solver. We plan to release the whole body posture data as animation video of human CG model as shown in Fig. 13 and joints information data such as FBX data.

As a background model of the referential motion, we plan to reconstruct the warehouse as a computer graphics model with texture. In order to reduce the amount of obstruction for the warehouse operation, we only recorded 4K omnidirectional movies by Ricoh's Theta V. The recorded omnidirectional movies are assumed to be used for reconstruction.

In order to control the privacy of valuable data in the warehouse, the introduced data for the competition will be released only for the participant who registered for the competition. We uploaded the call for participation on our website. We will inform updates through this website.

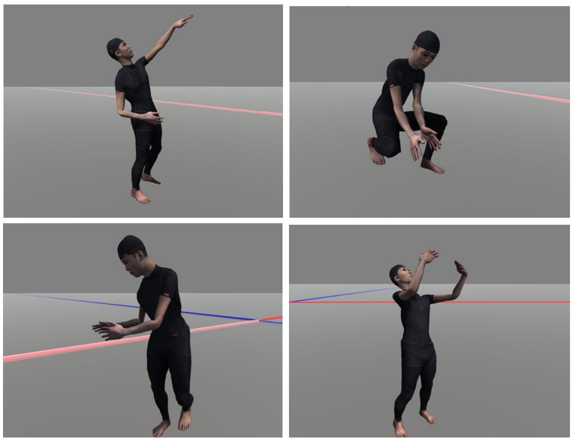

Fig. 13. Examples of reference data for whole body posture

IV. CONCLUSION

In this paper, we presented a review of the competition called PDR challenge in warehouse picking conducted by us in 2017. The PDR challenge 2017 proved to be a valuable competition under a realistic scenario, which was tracking employees during the picking operation in an actual warehouse. We adopted a comprehensive evaluation integrating multiple metrics required for the scenario in a warehouse and evaluation of integrated localization based on PDR. One of the most important error metrics used for the competition was the metric of error accumulation. We adopted an indicator namely, EAG for quantifying the speed of error accumulation. The important finding is that EAG can be used for helping the developer to determine sampling frequency of the correction methods and/or density of the arrangement of the correction methods.

Currently, we are preparing the next competition namely, xDR challenge for warehouse operations as a sequel to the PDR challenge 2017. The xDR challenge will be held by reflecting the findings and feedback. We will add forklifts as the tracking target by VDR. We will also provide reference data for learning how warehouse operations appear.

ACKNOWLEDGMENT

This research is partially supported by JST's Program on Open Innovation Plat-form with Enterprises, Research Institute and Academia (OPERA).

REFERENCES

[1] R. Ichikari, T. Kurata, K. Makita, T. Taketomi, H. Uchiyama, T. Kondo, S. Mori and F. Shibata: Reference framework on vSRT-method benchmarking for MAR, Proc. of Int. Conf. on Artificial Reality and Telexistence, Eurographics Symposium on Virtual Environments (ICAT-EGVE), 2017.

[2] D. Lymberopoulos and J. Liu: The microsoft indoor localization competition: Experiences and lessons learned, IEEE Signal Processing Magazine, Vol.34, Issue.5, pp.125-140, 2017.

[3] N. Moayeri, M. Ergin, F. Lemic, V. Handziski and A Wolisz: PerfLoc (Part 1): An extensive data repository for development of smartphone indoor localization apps, Proc. of Personal, Indoor, and Mobile Radio Communications (PIMRC), 2016.

[4] J. Torres-Sospedra, A. R. Jiménez, S. Knauth, A. Moreira, Y. Beer, T. Fetzer, Vi. Ta, R. Montoliu, F. Seco, G. M. Mendoza-Silva, O. Belmonte, A. Koukofikis, M. J. Nicolau, A. Costa, F. Meneses, F. Ebner, F. Deinzer, D. Vaufreydaz, T. Dao and E. Castelli: The Smartphone-Based offline indoor location competition at IPIN 2016: Analysis and future work, Sensors 2017, vol.17, No.3, pp. 557.

[5] F. Potortì, S. Park, A. R. J. Ruiz, P. Barsocchi, M. Girolami, A. Crivello, S. Lee, J. Lim, J. Torres-Sospedra, F. Seco, R. Montoliu, G. M. Mendoza-Silva, M. D. C. P. Rubio, C. Losada-Gutiérrez, F. Espinosa, and J. Macias-Guarasa: Comparing the performance of indoor localization systems through the EvAAL framework. Sensors, Vol. 17, No. 10, 2017.

[6] K. Kaji, M. Abe, W. Wang, K. Hiroi and N. Kawaguchi: UbiComp/ISWC 2015 PDR challenge corpus. UbiComp Adjunct 2016 (Proceeding of HASCA 2016), pp. 696-704, 2016.

[7] M. Abe, K. Kaji, K. Hiroi and N. Kawaguchi: PIEM: Path Independent Evaluation Metric for Relative Localization, Proc. of Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN2016), 2016.

[8] R. Ichikari, C.-T. Chang, K. Michitsuji, T. Kitagawa, S. Yoshii and T. Kurata: Complementary integration of PDR with absolute positioning methods based on time-series consistency, Proc. of Int. Conf. on Indoor Positioning and Indoor Navigation (IPIN2016) Work-in-Progress, 2016.

[9] M. Kourogi and T. Kurata: Personal Positioning Based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera, Proc. ISMAR 2003, pp. 103-112, 2003.

[10] https://unit.aist.go.jp/hiri/xDR-Challenge-2018/

[11] https://www.xsens.com/products/mtw-awinda/

[12] A. R. Jiménez: IPIN 2017 Pedestrian Dead-Reckoning (PDR) Tutorial http://www.car.upm-csic.es/lopsi/static/downloads/Pedestrian%20Dead-Reckoning%20(PDR)%20Tutorial_IPIN2017.pdf.

APPENDIX A

The detailed equations for calculating evaluation metrics:

【Metric for integrated positioning error: E_d】

$$E_d = 100 \quad \text{(median(PosErr) < 1)}$$

$$E_d = 100 - \frac{100}{29}(median(PosErr) - 1) \quad \text{(30>median(PosErr) > 1)}$$

$$E_d = 0 \quad \text{(median (PosErr) > 30)}$$

, where median of positional errors is median(PosErr)

【Metric for PDR error based on EAG:E_s】

$$E_s = 100 \quad \text{(SlopeErr < 0.05)}$$

$$E_s = 100 - \frac{100}{1.95}(SlopeErr - 0.05) \quad \text{(2.0> SlopeErr > 0.05)}$$

$$E_s = 0 \quad \text{(SlopeErr > 2.0)}$$

, where slope by linear regression of errors is SlopeErr.

【Metric for motions during picking work: E_p】

Given, $Flag_i^p$ is assigned 1 when target stops and is assigned 0 when target does not stop at the i^th checking points.

$$E_p = \frac{\sum_{i=1}^{N_{pick}} Flag_i^p}{N_{pick}} \times 100$$

, where N_{pick} is the number of all checking points.

【Metric for naturalness of travel speed :E_v】

Given, $Flag_i^v$ is assigned 0 when target moves faster than the threshold 1.5 m/sec and is assigned 1 when it is below the threshold at the i^th frame in the trajectory's frames.

$$E_v = \frac{\sum_{i=1}^{N_{traj\_frame}} Flag_i^v}{N_{traj\_frame}} \times 100$$

, where N_{traj_frame} is the number of total frames of the trajectory.

【Metric for collision with obstacles: E_o】

Given, $Flag_i^o$ is assigned 1 when the target pixel on the trajectory is not in obstacle areas, and is assigned 0 when the target pixel is in the obstacle areas at the i^th pixels in the trajectory.

$$E_o = \frac{\sum_{i=1}^{N_{traj\_pixel}} Flag_i^o}{N_{traj\_pixel}} \times 100$$

, where the total number of pixels in trajectory is N_{traj_pixel}.

【Metric for position measurement output frequency :E_f】

Given, $Flag_i^f$ is assigned 1 when the temporal frequency is more than 1 Hz and assigns 0 when the frequency is less than 1 Hz at the i^th frames in the submitted trajectory.

$$E_f = \frac{\sum_{i=1}^{N_{traj\_frame}} Flag_i^f}{N_{traj\_frame}} \times 100$$