2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Mobile Robot Localization System in Frequent GPS-denied Situations

Takato Saito, Kentaro Kiuchi and Yoji Kuroda
Meiji University, Department of Mechanical Engineering,
1-1-1 Higashimita, Tama-ku, Kawasaki, Kanagawa, Japan
Email: {ce22023, ce32012, ykuroda}@meiji.ac.jp

Abstract— **In this paper, we propose a mobile robot localization system in frequent GPS-denied situations. We utilize multiple observations that are obtained from sequential appearance-based place recognition and GPS. Using GPS observations has still some challenging problems such as multipath or signal lost under environments where there are tall buildings nearby. The appearance-based place recognition that is combined with positional information has capability to overcome the issue. Nevertheless, GPS observations which are obtained in the situation sometimes have better quality (e.g. Precision or accuracy) than positional information from the place recognition because those coordinates always have some errors. We apply both of observations to a mobile robot localization for the sake of achieving robust localization. Moreover sequential appearance-based place recognition makes it possible to recognize their own position even when we navigate a robot at night. Our system uses not only multiple observations but also dead reckoning with the gyrodometry model. Our experiments are performed over aggregate 5300 m trajectory approximately that contains three times trials through a 1600 m outdoor route in different seasons and at different times, and once trail through a 500 m short-range route to verify its validity.**

I. Introduction

RECOGNIZING a robot's current position is very important and inevitable ability in order to navigate robots autonomously or manually anytime and anywhere.

Dead reckoning system that fuses data coming from odometry and gyro sensors has been widely used for tracking the local position; nevertheless, the accumulated error of it is in proportion to the running distance. It's necessary to acquire the robot's position in global coordinates to reduce the error. Many researchers have employed Global Positioning Systems (GPS) as the means to recognize the robot's global position[1]. Especially in urban environments, there are buildings and trees that give rise to the fatal position error known as multipath error, though we should deny those GPS observations to prevent false recognition of the robot's position. However, detecting those in outliers by using only cheap GPS receiver would be difficult[1].

On the other hand, it is a well-known approach that appearance-based place recognition (PR) methods are efficient to recognize the robot's position in urban areas[2]. Fast Appearance-Based Mapping[3] that is a kind of feature-finding based approach has attained high-speed processing and stable performance by image information only. It employs Bag-of-Words representation, and assumes a strong Bayesian framework with taking account of relations of co-occurring visual words to alleviate the adverse effect of disturbances.

In many papers which are represented by[4], these PR methods usually have been used as a loop-closure detection merely in SLAM (Simultaneous localization and mapping). Additionally, many researchers such as [5] have proposed a localization system based on multiple sensors.

PR methods enable to recognize the current position accurately in environments where there are surrounded by many artificial objects in contrast to GPS. We can grasp a robot's current position from reference images with attached global position information. As far as we know, there has never been applied the position information directly as an observation in a sensor fusion. However, it could fail to function properly in natural environments like grassy areas and leafy streets, or when a captured image is entirely occupied by disturbances.

We have proposed a localization system using not only GPS observations, but also PR so as to attain a robust system everywhere in urban environments[6]. Obtaining a robot's global position from multiple observations constantly, we can reduce the accumulated error of dead reckoning and prevent the divergence of the estimation even in urban environments.

Nevertheless, feature-based methods are not able to overcome transitions between daytime and nighttime since the performance of it depends on the extraction stability of interest points from objects in images[7]. If we cannot obtain it constantly, these methods do not function properly. As a result, we cannot observations from PR for our localization.

In this study, we employ SeqSLAM[7][8] as a sequential appearance-based PR. The approach utilizing sequential images does not need interest points. It surmounts severe environmental changes such as a sunny summer day, stormy winter night, and seasonal changes through a long journey[7][8][9][10]. In addition, we also use GPS observations because it has higher accuracy and precision at a place with good visibility although appearance-based PR has some troubles under the condition. Those observations are integrated to dead reckoning with the gyrodometry model[11] using Divided Difference Filter[12] which is a kind of Kalman Filter. To verify the validity of our system, we conducted experiments through outdoor courses.

II. Proposed System Architecture

We propose a localization system using multiple observations to solve disadvantages of these observations. We use GPS observations which are evaluated by a pre-filter. Our pre-filter detects its outlier based on the number of satellites and the measurement mode such as differential mode or not.

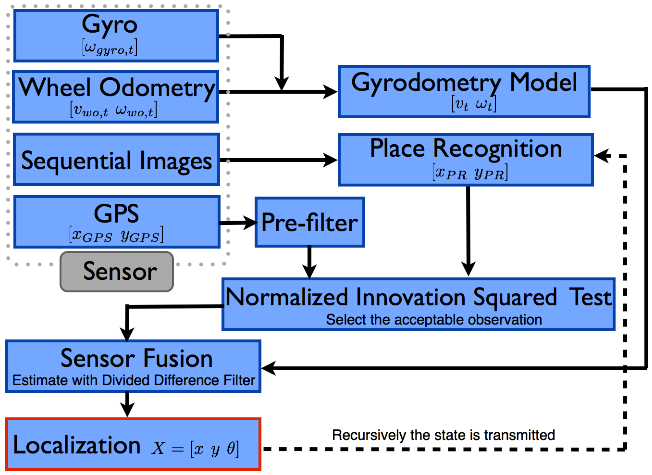


Fig. 1. Proposed System : ω is a yaw angular velocity, v is a forward velocity with the time t. x and y indicate these coordinates.

It is difficult to detect outliers perfectly despite Horizontal Dilution of Precision (HDOP) also is monitored in the filter.

Appearance-based PR has three main problems. The first is that its performance depends on interest point detector methods. The second that is attributed to the first problem is the robustness against closely-similar places (e.g. Grassy areas in a park and similar buildings in urban environments). The third is that disturbances such as people and cars might occupy a whole image area. Even if the robot is in a known place, the likelihood between images becomes low. It may cause a false-positive. To prevent false-positive completely is difficult even with the use of state-of-the-art techniques.

To ameliorate above issues we use a sequential-image based PR method which does not need an extraction of interest point, and it has the ability to work in a night even if we use cheap camera. It can recognize the true-positive even when an image is occupied by disturbances. Disadvantages of GPS observations would be compensated by the method.

There are some situations with no observations regardless of multiple observations. Therefore, we fuse these observations with dead reckoning using Divided Difference Filter (DDF)[12]. The robot state is transmitted to PR, then it restricts the searching scope in reference images. PR finds local best matching image sequences in the scope and informs us coordinates recursively.

The system is shown in Fig. 1 and combined as follows :

1) GPS Observations : Using Pre-filter to evaluate the number of satellites, its mode, and HDOP
2) Place Recognition : Using SeqSLAM and a frontal camera with a fisheye lens
3) Dead Reckoning : Using Wheel Odometry and a gyro sensor with the gyrodometry model
4) Sensor Fusion : Using DDF for integrating observations and dead reckoning

Our localization becomes more stable and accurate by fusing dead reckoning even when these observations are not obtained. The most plausible observation could be employed using Normalized Innovation Squared (NIS) Test based on a robot motion model[13].

III. Observations from Place Recognition

We employ SeqSLAM as an appearance-based PR. In this section, we describe about our using PR system briefly. Following it, we explain how combine SeqSLAM with localization system.

A. Appearance-based Place Recognition

SeqSLAM[8] has been proposed by M. Milford et al., and it has several steps to find the best matching image as follows : Preprocessing of input images, Sum of Absolute Difference (SAD) Matrix Calculation, Contrast Enhancement of Difference Matrix, and Find Local Best Matchings.

Firstly, we convert input images to grayscale, and create those thumbnail images as a preprocess. These both reference images and experimental images should be applied a patch normalization to compensate for illumination change[10].

Secondly, we calculate SAD values by subtracting all of reference images and recent experimental sequential images, then create the matrix $\mathbf{M}$ that consists of those SAD values. The matrix $\mathbf{M}$ owns a set of some 1D difference vectors from the time $T-d_s$ to the current time T. After that, we apply the contrast enhancement which is a kind of 1D normalization to the difference matrix.

Finally, SeqSLAM searches for it throughout $\mathbf{M}$ to find local best matching image sequences. Taking into account of total scores S of image sequences, we work out a robust matching. It is expressed as follows:

$$S = \sum_{t=T-d_s}^{T} D_k^t \tag{1}$$

where D_k^t indicates a SAD value in $\mathbf{M}$ at time t with index k. k is represented by the searching range parameter V within a range of V_{max} from V_{min} at an interval of V_{step} as follows:

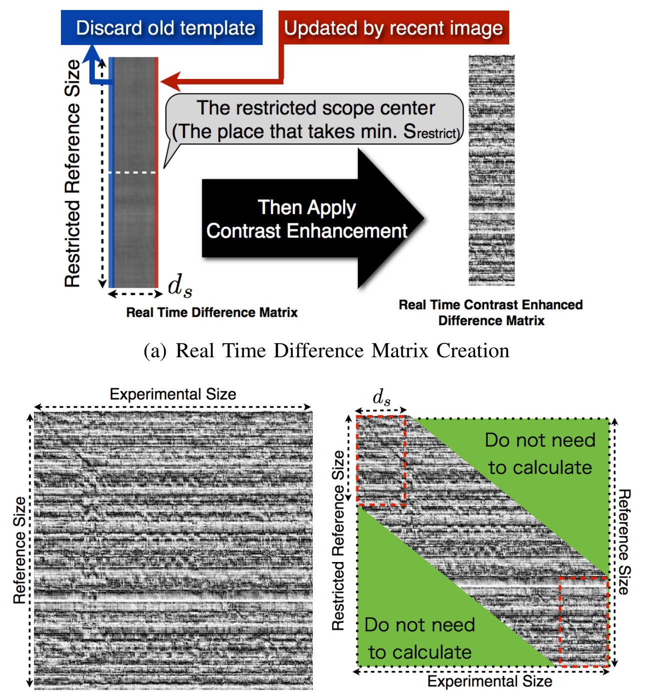
$$k = s + V(d_s - t + T) \tag{2}$$

where s gives a start index in reference images. We distinguish the image sequence that takes a minimum score as the matching image. For more details about SeqSLAM, please refer to their papers[7][8].

B. Restrict the searching scope

We can find local best matching image sequences and obtain coordinates, as long as the reference database size and length of d_s are modest. Unfortunately, we have to spend a long time to find it when these sizes become larger. SeqSLAM needs mainly four steps to find images. The difference matrix calculation and finding local best matching images especially take time. Thus we search for it in the scope which is confined by the robot's state derived from DDF estimation. Reference images have own coordinates that are attached previously using GPS, or by adjusting hand. We define a score for restricting as follow :

$$S_{restrict} = d/cos\theta \tag{3}$$



(a) Real Time Difference Matrix Creation

(b) The Entire Matrix

(c) The Calculated Matrix

Fig. 2. Real-time Derivation of Differential Matrix : The horizontal axis and the vertical axis mean experimental image sequences and reference image sequences respectively. (a) shows the real-time differential matrix in this experiment. (b) means the entire contrast enhanced difference matrix through a trajectory. (c) is the restricted contrast enhanced matrix from (b). We do not need to set up green regions in the entire matrix. Red-framed regions indicate examples of (a).

d means Euclidean distance from attached coordinates of reference images to current robot's position. θ indicates an angle between the yaw of reference and that of the current. If the denominator takes negative value, we impose a penalty value on the score since it means the robot faces the opposite direction. The place that takes minimum value is selected for the center of restricted searching scope.

A real time difference matrix is composed of recent d_s images that would be updated with sustaining d_s images as shown in Fig. 2. Fig. 2(a) shows the real time process. We find the local best matching image sequence in this matrix.

IV. Sensor Fusion

A. Sensor Fusion with Divided Difference Filter

We estimate the 2D robot state that is defined as $\mathbf{X} = [x\ y\ \theta]$. The observation that is either GPS or PR is fused with the gyrodometry model[11] by DDF.

Gyrodometry model is a way to integrate wheel odometry and a gyro sensor that measures yaw angles, this method is well-known to prevent the accumulated error of dead reckoning. Then the state of the dead reckoning is considered as input for DDF. It can be anticipated to be a non-systematic odometry and a gyro error (e.g. Slip, bump and gyro bias).

The motion model in our system employs commonly-used two wheel differential model. As a consequence, the dead reckoning represents the robot state as $[x_t\ y_t\ \theta_t]$ at the time t. The observation of GPS or that of PR is deemed as the observation of DDF, and the vector is defined as $z_t = [x_t\ y_t]$.

DDF is a kind of Kalman Filtering algorithm that has some advantages more than other algorithms for mobile robot localization. It does not need Jacobian matrices unlike Extend Kalman Filter especially, but the divided difference is used in place of it. That is because DDF is based on Stirling's interpolation formula[12]. It uses divided difference for estimating the gradient of an arbitrary nonlinear function with the set of sampled points.

The most important advantage for us is that DDF estimation is more accurate and robust than the other algorithms such as methods using Taylor expansion when the behavior of observations is asynchronous or noncontiguous chiefly[12]. It is suited for our localization system, since we often switch the observation, which is obtained by GPS or PR, used for the estimation in urban environments. Thus DDF is very useful for us.

There are two types of DDF. That is DDF 1 and the other is DDF 2. The former limits its interpolation in first order, and the latter limits in second order. Our localization system uses DDF 1. The more details about the algorithm can be found in [12].

B. A fault detection and selection based on motion model

We use the observation that is derived from either GPS or PR. However, GPS observations which are influenced by multipath indicate false positions. False positive of appearance-based PR is often triggered by the existence of similar places. A place several meters away from ground truth might be matched.

These faults should be detected for achieving a high accurate and consistent localization. In this study, the observation passed through Normalized Innovation Squared (NIS) Test[13] is allowed to fuse with the input of DDF in the sake of preventing these mistakes. It calculates Mahalanobis distance recursively by the innovation matrix as a normalizer as follows:

$$(\bar{x}_t - z_t)P_{xy}^{-1}(\bar{x}_t - z_t)^t < \delta_t \tag{4}$$

where P_{xy} is the innovation matrix in the correction step of DDF. z_t means an observation of GPS or PR. $\bar{x}_t$ is the state vector before update, and δ_t is threshold. NIS Test values are calculated for each observation, and selected the observation that takes the minimum value to fuse with localization. It is possible that localization becomes stable and consistent based on a motion model by this selection. The constant threshold may reject a good observation if the estimated robot state is in the distance from ground truth, hence we predefine thresholds for each. These thresholds vary at an exponential rate, according to the distance $\mathcal{D}$ between the estimated current place and the most recent place where obtained each observation. We define the threshold in the form of $\alpha exp(\beta \times \mathcal{D})$. α and β are constant values. If an observation is rejected by NIS Test or an observation could not be obtained, DDF executes only the prediction step, and then DDF is not updated.

V. Experimental Results

We conduct some experiments where are a lot of buildings, leafy streets and so on. Fig. 3 explains these trajectories: the distance of one lap of trajectory 1 is a 1600 m, and that of trajectory 2 is approximately a 500 m. Three times experiments are conducted in trajectory 1 at different times of different seasons. Then an assessment of our proposed method is quantified by an experiment in the trajectory 2. Typical scenes in this route are shown in Fig. 4.

We store images that are captured every 1 m based on wheel odometry, and attach its coordinate to create reference database. The coordinate is adjusted by hands when it is hard to get it, though we collect reference images and attach its coordinate by using GPS. Our reference database for PR has been build with data on Apr. 30, 2012, a cloudy day. We can approximately obtain 1600 images and 500 images along trajectory 1 and 2 respectively. We assess our result by comparing it to the ground truth derived from the experiment on Mar. 15, 2012. The evaluate experiment includes peculiar disturbances such as scattered flower petals(e.g. Cherry blossom) on roads. Some experimental images on Jul. and Aug. have been taken against the sun. Moreover, the condition on Aug. 23 is darkening.

Qcam for Notebooks Pro (QCAM-200V) made by Logicool with a fisheye lens is mounted at the height of 0.9 m on the robot, and it faces to frontal direction. We do not take account of lens distortion. The original image resolution is VGA (640×480 pix). The IMU (NAV 420) made by Crossbow Co. is used as a gyro sensor. DGPS A-100 made by Hemisphere is used in our system. We also obtains reference position data for PR from this GPS. The observation accuracy of position is 50 cm (95%). Fiber Optic Gyro (FOG, JG-35FD made by Japan Aviation Electronics Industry) has the ability to get highly accurate and precise angular variation. This sensor is used when we derive metric ground truth to evaluate the localization result. Its drift angle is less than 3 degrees per hour. All of our system (shown in Fig. 1) is computed on Intel®Core$^{\text{TM}}$2 Quad CPU Q9650 3.00GHz.

A. Result of Place Recognition

Our main parameters for SeqSLAM and NIS Test are listed on Table I. We have defined the restricted reference size as $2 \times d_s$ from the center that is determined by utilizing the estimated state. The parameter $\alpha = 2.0$ means that we can accept the observation that its significance probability surpasses 36.5% as default. Perhaps it is a lower level than general confidence level. Nevertheless we have found that it is suitable for experiments in urban environments by multiple observations.

Performances of PR are shown in Fig. 5. Plots on the diagonal line mean that PR is performed correctly when the reference and experimental trajectory are closely similar. There is not plot until place number 99 to store images, because we had defined d_s as 100 images. Fig. 5(a) and (b) indicate promising results for enough accuracy and stability to estimate the state. PR finds the corresponding image within ±5 frames with a probability of 90% at the worst.

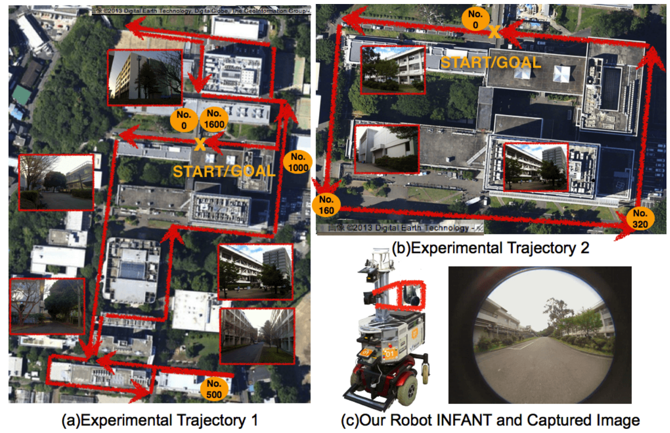


Fig. 3. Our Mobile Robot and Experimental Field: (a)(b) The x-mark means start and end of the track. The red arrow shows both experimental trajectories 1 and 2. The number on the yellow circle indicates a ballpark place number. Images with the fence-line represent buildings in this environment. [14] (c) Our robot named INFANT that has a camera with a fisheye lens. This robot has been developed in our lab by using Pride Inc. Wheel Chair Jet 3.

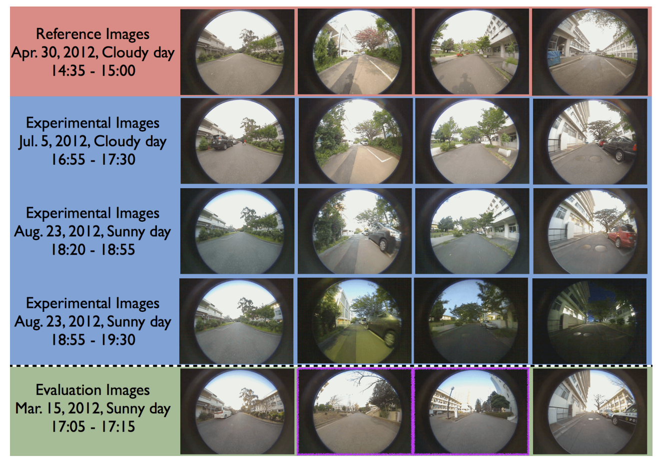


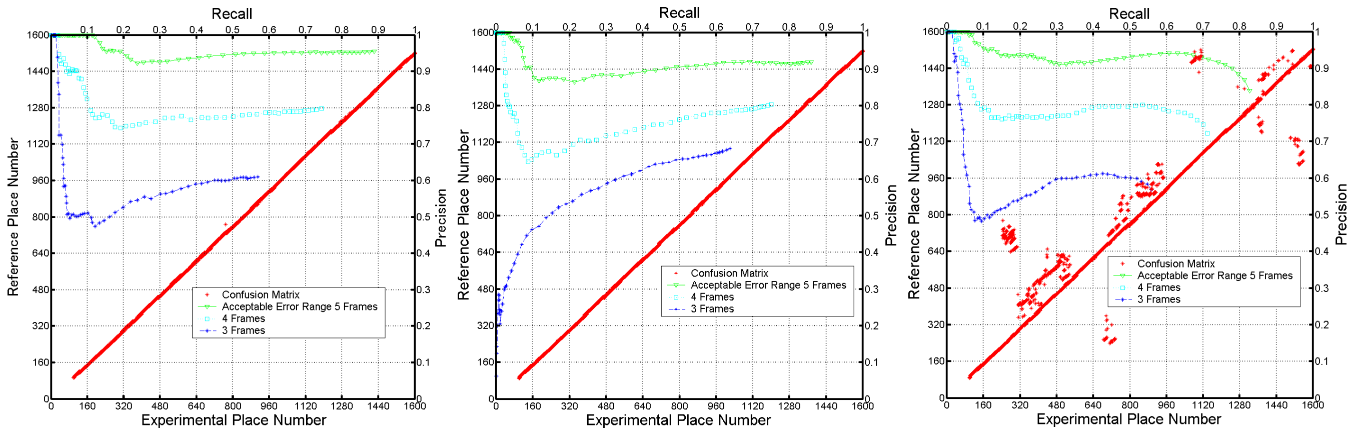
Fig. 4. Experimental Conditions and Typical Scenes: The purpose of use is described in the left column. Reference and experimental images are indicated on a red and blue background respectively. All of those above the dotted line are captured through a 1600 m route (Trajectory 1). Images on the bottom of the dotted line on green background represent the scene of the evaluation route. Two images with purple fence-line express the image not being contained in our reference database. There is correspond to the region at the bottom left in Fig. 3(b).

Some false positives shown in Fig. 5(c) have been occurred. The major contribution is that subtle differences between buildings in images could not be distinguished. It can be improved by changing the direction of the cameras[7]. We do not discard the matching result with using the uniqueness parameter[8], since our next step calculates the reliabilities of observations.

The acceptable range should be small for the sake of utilizing with a mobile robot localization, but our localization system discriminates whether the observation of PR should be employed or not by using NIS test. We set the thresholds of NIS test as large values, but outliers have been discarded adequately. It proves that PR has functioned properly under the condition that our robot works at daytime, sunset, or night even though reference data have been collected in different seasons.

TABLE I
PARAMETERS FOR EXPERIMENTS OF SEQSLAM AND NIS TEST

For Preprocessing				For SeqSLAM					For NIS Test			
Raw Image Size	Cropped Size	Subsampled Size	Patch Size	d_s [frames]	R_{window} [frames]	V_{max}	V_{min}	V_{step}	α_{GPS}	β_{GPS}	α_{PR}	β_{PR}
640×480(VGA)	480×240	120×60	10×10	100	10	1.2	0.8	0.1	2.0	0.28	2.0	0.20



(a)On Jul. 5. (b) On the sunset of Aug. 23. (c) On the night of Aug. 23.

Fig. 5. Matching Performance of each experiment: Confusion Matrix shows the matching result along the bottom horizontal axis and right vertical axis. Precision-Recall curves are depicted as lines with dots. The corresponding image within the acceptable error range is considered as true-positive.

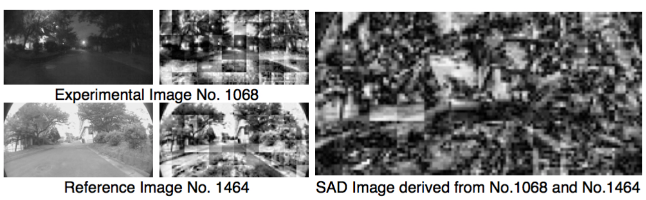


Fig. 6. False Positive of Place Recognition on Aug. 23: Both captured image and normalized image are shown with its place number. The bottom left image is the answer image by SeqSLAM. SAD image is shown on the right. The darker pixel means high similarity.

B. Result of Localization

All of the reference data have definitely been the same. That had been collected through trajectory 1. Observations of PR have been capable of processing at approximately 1 Hz. The rate is equal to the frequency of our DGPS. Conventional localization is the same system as proposed system except that it does not employ PR. Thus it fuses GPS observations, which is declined outliers by the judgements of the NIS test, with dead reckoning based on wheel odometry and IMU.

Localization result is shown in Fig. 7 to Fig. 9. To assess these results quantitatively we calculate positions as the ground truth by using the dead reckoning that employs FOG and high accuracy GPS observations. The accuracy of ground truth depends on those of sensors though we employed a short route to avoid accumulated errors. Our localization system can estimate positions more accurate estimation than conventional localization. Even when observation-denied or intermittent conditions, jumps of localization are converged by making up for absent of each observation and estimating with DDF.

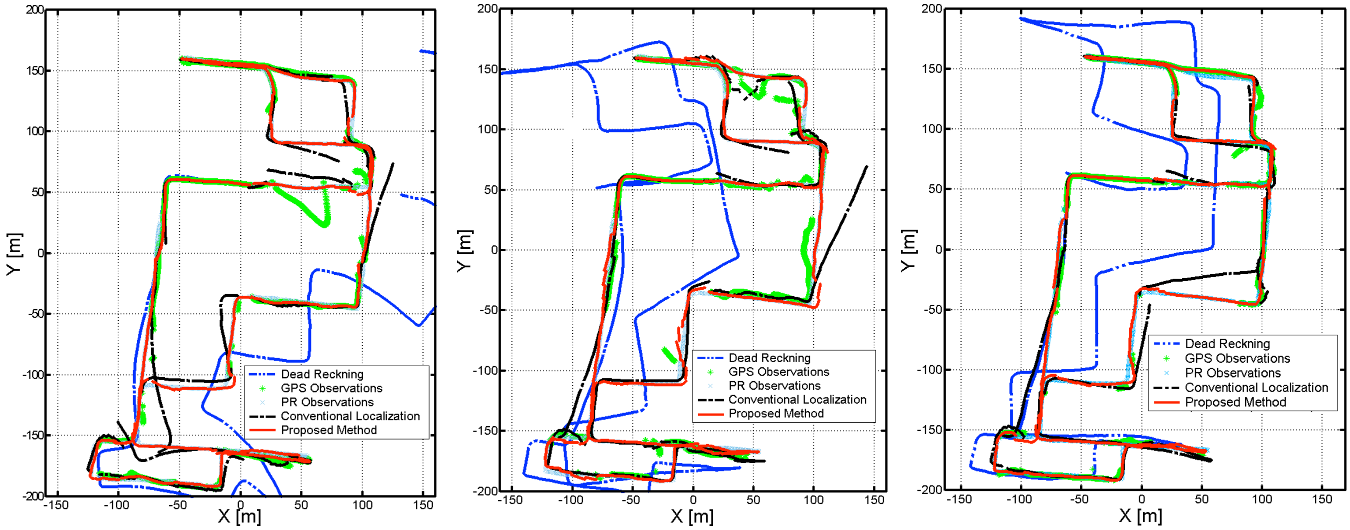
The remarkable point is that our system functions properly, even if the experimental route includes an unknown-route in Fig. 8 and 9. Fig. 9 shows 2D localization errors along the processing time, and its performance of PR. It shows that accumulated error does not increase, and our positional error have not been diverged. We compare errors shown in Table II, and have led to a decrease in 40% for RMSE, 24% for avg. error and 64% for max. error respectively. The main reason why the error has been caused is the difference between reference and experiment. The performance of PR in this evaluation is indicated in Fig. 9(b). Although it seems that we had got some doubtful observations due to false positives in this evaluation, almost of the matching can be considered as adequate for localization. Our system would become stable and accurate as far as the experimental route is extremely-different from the reference route, but narrow alley would have a harmful influence on our system. From these results, the effectiveness and the redundancy of our proposed system is demonstrated.

TABLE II
ERROR COMPARISON

Approach	Proposed Method	Conventional Method
RMSE [m]	1.218	1.976
Avg. Error [m]	2.125	2.792
Max. Error [m]	4.829	13.47

VI. CONCLUSION AND FUTURE WORKS

In this paper, we have proposed a localization system using multiple observations. To show the validity of our system, experiments have been performed over a 5300 m route in different seasons environments. The system could achieve robust localization. In the near future we must estimate the pose such as the yaw angle from the PR by using



(a) Jul. 5, 2012. (b)On sunset of Aug. 23, 2012. (c) On night of Aug. 23, 2012.

Fig. 7. A comparison of Proposed Localization with Conventional Localization

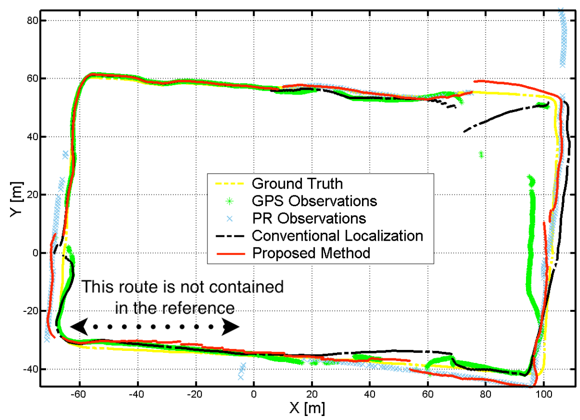
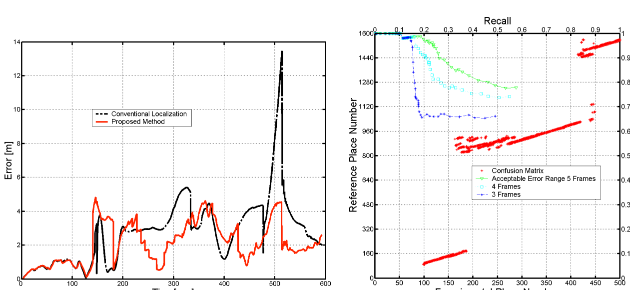


Fig. 8. Localization Results for Evaluation: Although there are observations due to false positives of PR, these's no significant effects for the estimation.



(a) Localization Error (b) Matching Result

Fig. 9. The Localization Error and Performance of PR in the evaluation

omnidirectional camera though we focus on the horizontal position of the state in this paper. Large-scale experiments have to be conducted using the Internet Services(e.g. Google Street Views). Moreover, we must to be able to deal with unknown places in PR in real-time.

References

[1] D. Maier and A. Kleinar, “Improved GPS Sensor Model for Mobile Robots in Urban Terrain,” in Proc. of the IEEE International Conference on Robotics and Automation, pp. 4385-4390, 2010.

[2] K. Konolige, J. Bowman, J. D Chen, P. Mihelich, M. Calonder, V. Lepetit and P. Fua, “View-Based Maps,” in Proc. of The International Journal of Robotics Research, 29(8),pp. 941-957, 2010.

[3] M. Cummins, P. Newman, “FAB-MAP : Probabilistic localization and mapping in the space of appearance,” in Proc. of The International Journal of Robotics Research, 27(6), pp. 647-665, 2008.

[4] P. Newman, G Sibley, M. Smith, M. Cummins, A. Harrison, C. Mei, I. Posner, R. Shade, D. Schroeter, L. Murphy, W. Churchill, D. Cole, and I. Reid, “Navigating, Recognizing and Describing Urban Spaces With Vision and Lasers,” in Proc. of The International Journal of Robotics Research, 28(11-12), pp.1406-1433, 2009.

[5] S. Thrun and M. Montemerlo, “The GraphSLAM algorithm with applications to large-scale mapping of urban structures,” in Proc. of The International Journal of Robotics Research, 25(5), pp. 403-430, 2006.

[6] T. Saito and Y. Kuroda, “Mobile Robot Localization using Multiple Observations based on Place Recognition and GPS,” in Proc. of the IEEE International Conference on Robotics and Automation, pp. 1540-1545, 2013.

[7] Michael J. Milford, “Vision-based place recognition: how low can you go?,” in Proc. of the International Journal of Robotics Research, 32(7), pp. 766-789, 2013.

[8] Michael J. Milford and Gordon. F. Wyeth, “SeqSLAM: Visual Route-Based Navigation for Sunny Summer Days and Stormy Winter Nights,” in Proc. of the IEEE International Conference on Robotics and Automation, pp. 1643-1649 , 2012.

[9] N. Sünderhauf, P. Neubert, and P. Protzel, “Are We There Yet? Challenging SeqSLAM on a 3000 km Journey Across All Four Seasons,” Workshop on Long-Term Autonomy held in conjunction with the International Conference on Robotics and Automation, 2013.

[10] Alan M. Zhang and L. Kleeman, “Robust appearance based visual route following for navigation in large-scale outdoor environments,” in Proc. of the International Journal of Robotics Research, 28(3), pp. 331-356, 2009.

[11] J. Borestein and L. Feng, “Gyrodometry: A New Method for Combining Data from Gyros and Odometry in Mobile Robots,” in Proc. of the IEEE International Conference on Robotics and Automation, Vol. 1, pp.423-428, 1996.

[12] Alexandre N. Ndjeng, D. Gruyer, A. Lambert, B. Mourllion, S. Glaser, “Experimental Comparison of Bayesian Outdoor Vehicle Localization Filters,” in Proc. of the IEEE International Conference on Robotics and Automation, pp.4724-4729, 2009.

[13] Y. Morales, A. Carballo, E. Takeuchi, A. Aburadani, and T. Tsubouchi, “Autonomous Robot Navigation in Outdoor Cluttered Pedestrian Walkways,” in Proc. of the Journal of Field Robotics, 26(8), pp.609-635, 2009.

[14] GoogleEarth Image ©2012 DigitalGlobe, GeoEye, Digital Earth Technology, maps©ZENRIN. ©2012 Google