Proceedings of the 2015
IEEE Conference on Robotics and Biomimetics
Zhuhai, China, December 6-9, 2015

G-ICP SLAM: An Odometry-Free 3D Mapping System with Robust 6DoF Pose Estimation

Ryo Kuramachi[1,2], Akihito Ohsato[1,2], Yoko Sasaki[2], Hiroshi Mizoguchi[1,2]

Abstract— **The paper proposes an odometry-free 3D mapping system that combines a LIDAR and a inertial sensor. The proposed system achieved robust 6DoF pose estimation for arbitrary motion and is implemented as a hand-held unit to make use of simplified mobile mapping applications. The pose estimation algorithm is based on "Velodyne SLAM" which is a state of the art ICP based SLAM (Simultaneous Localization and Mapping) method using only point cloud data. We added 3DoF inertial information to process the point cloud correction and the position prediction. Compared to the previous method, the proposed method is robust to rotary motion and works for fast and large change of sensor position and orientation. The results demonstrate effective operation in various environments and we confirmed the improvement of the self-position estimation and mapping performance.**

I. Introduction

In recent years, research on autonomous mobile robots has made progress. Robots work in environments where humans are living, for example, the home, office, and city. They are increasingly required to operate in varied environments and interact safely and effectively with humans, objects and their environment. For such autonomous robots, environment mapping is a vital function for its localization and path planning.

There are amount of existing methods for 3D data reconstruction, such as LIDAR or radar, stereo vision, imaging sonar, and structured triangulation techniques. It is known as Simultaneous Localization and Mapping (SLAM) algorithms[1][2] in robotics: generating the environment map and performing self-position estimation at the same time. There are plenty of existing techniques and differs in many ways of primary relevance: accuracy, resolution, field of view, operating environment, computing cost and robot configuration.

As for algorithms, there are two main approaches: using Bayesian theorem and scan matching. In the former approach, the sensor trajectories are estimated stochastically. Extended Kalman Filters (EKFs)[3] performs condition estimation for non-linear models using linear approximation. Rao-blackwellized Particle Filters (RBPFs)[4] performs condition estimation using probability distribution. These algorithms are well established and widely used in many robotics applications. But if the environment is large, the number of landmarks used for calculation increases; thus, calculation becomes difficult. As for latter approach, ICP algorithm[5][6] estimates the movement amount of the robot and ensures that it is equal to the movement amount of the scan data by using overlapping scan information obtained by the sensor each time.

Many SLAM frameworks have been developed in recent years; for example, MCL[7] and GMapping[8] provide 2D localization and mapping. "ethzasl icp mapping", is a state of the art ICP-based 3D SLAM. These are well established and widely used for many applications. On the other hand, these SLAM systems require wheel odometry or robot motion model. "Velodyne SLAM"[9] is the latest study of 3D mapping without odometry information. The system is designed for car applications in urban area, i.e., two-dimensional motion on the ground plane. It is a state-of-the-art ICP based mapping method using only point cloud data obtained from the LIDAR sensor mounted on the car.

This paper proposes a 3D mapping system named "G-ICP SLAM" (Gyro integrated ICP SLAM) using a LIDAR and a tri-axial inertial sensor. The proposed system is based on ICP algorithm and implemented as a hand-held unit. The main contribution is that G-ICP SLAM is worked without odometry and robust to arbitrary 6-DoF movements. In Section II, we describe a particular implementation of the proposed G-ICP SLAM based on performance evaluation of Velodyne SLAM. Some experiments demonstrating G-ICP SLAM are presented in Section III. We evaluate the performance in various environments including motion capture area.

II. G-ICP SLAM

G-ICP SLAM is based on "Velodyne SLAM" which copes with two-dimensional movement on the ground e.g. a car moving on the ground. We investigated the performance of "Velodyne SLAM" with local rotary movement by experiments. In accordance with the result of the experiments, we could cope with 6-DoF movement with fast and large change of sensor position and orientation e.g. using the hand-held unit sensor in the proposed method.

A. Algorithm of Velodyne SLAM

"Velodyne SLAM" uses a laser scanner manufactured by Velodyne, HDL-64E. The sensor was implemented on the car, and it was turned at 10 [Hz] to scan the environment while driving. The provided data was saved as a range image, and we display the point cloud data in 3D space in accordance with the range information from the range image. Finally, a 3D map of the environment was generated. For each frame, the system predicted the position of the next frame from the movement between the previous frame and the current frame. And the common point cloud of the current frame with a map that was generated before the frame

[1]Dept. of Mechanical Engineering, Tokyo University of Science
[2]Human Informatics Research Institute, National Institute of Advanced Industrial Science and Technology

978-1-4673-9675-2/15/$31.00 © 2015 IEEE

was calculated using an ICP algorithm. The ICP algorithm defines the corresponding points a_i and b_i: i-th points at two different points cloud A and B, which are included two point clouds at different positions. Then, the square sum of the distance $d(a_i, b_i) = \min ||a_i - b_i||$, between the corresponding points was calculated. Now, a movement vector T:

$$T = \begin{pmatrix} R_{1,1} & R_{2,1} & R_{3,1} & t_1 \\ R_{1,2} & R_{2,2} & R_{3,2} & t_2 \\ R_{1,3} & R_{2,3} & R_{3,3} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (1)$$

which minimize the square sum. Matrix R is a 3×3 rotary movement matrix, and matrix t is a 1×3 translation movement matrix. The point cloud A moved in accordance with the movement vector T

$$A' = T(A) \quad (2)$$

The previous method predict the movement vector T by using only the translation movement matrix calculated from the velocity of the before frame. The ICP algorithm fix the predicted vector T at first trial, and thereafter, make harmonizes point cloud A to B by performing these processes repeatedly. The previous method matches the point clouds from the sensor by using the ICP algorithm; with the movement of the sensor, it can generate a 3D environment map.

B. Performance inspection experiment of "Velodyne SLAM"

We first evaluated the performance of the previous method, " Velodyne SLAM " . To implement the hand-held unit, we use a laser scanner HDL-32E, and investigate sensor pose estimation performance of large motion. The unit is put on a tripod horizontally to control rotation angles as shown in Fig.1. And yaw, pitch and roll direction turn are showed on the figure.

Two scan data image is used, the horizontal scan data and added only a single axis turn scan data. We performed 3D mapping by using the two scan data with "Velodyne SLAM". By increasing the rotation angle ($\Delta\theta$) difference between scan data pair, we evaluate the valid angle difference between two frames; the error value of the difference between the estimated rotary angle and the added true rotary angle. The laser sensor scans surrounding area in a 360 [deg] rotation. And the initial orientation is horizontal, so the turn of Pitch direction and turn of Yaw direction are considered to be the same turn. Therefore, we added the Yaw direction turn and Pitch direction turn to the experiment unit and matched the provided data.

The experimental results of indoors and outdoors are shown in Fig.2 and Fig.3. In these graphs, the y-axis shows the angle estimation error, and the x-axis shows the rotary angle. We changed the rotary angle by 0.05 [rad] every time, which is equal to about 2.8[deg] indoors in the motion capture area. We changed the rotary angle every 3 [deg] outdoors in the car parking lot.

As a result, for both the Pitch direction turn and the Yaw direction turn, the error values of outdoors are larger than

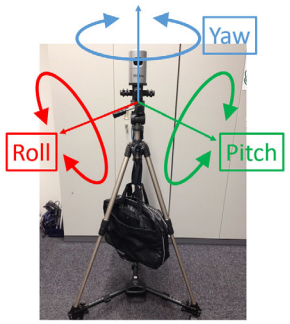


Fig. 1. The hand-held unit on a tripod

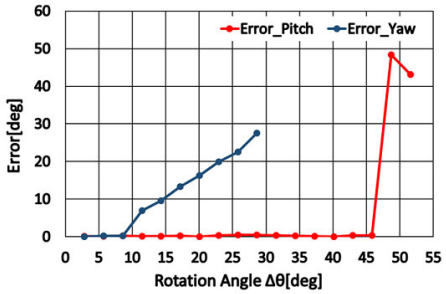


Fig. 2. The sensor motion estimation error of indoors

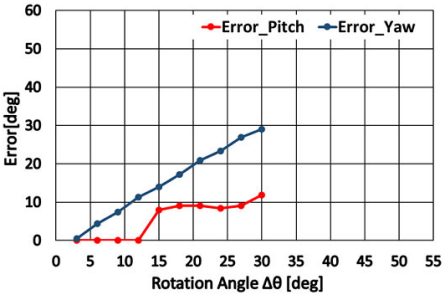


Fig. 3. The sensor motion estimation error of outdoors

those of indoors. And the rotation angle became larger, the error value became larger too. It was shown that the previous method can't match the two scan data when the rotary angle became larger. So, correspondence to rotary movement was necessary for the 3D mapping method so that it is robust to rotary motion and works for fast and large changes of sensor position and orientation.

C. Approach of the proposed method

To improve the ability to larger motion for the hand-held system, we propose adding the 3-axis gyro at the bottom of the laser scanner. The inertial information of the gyro is used for two functions; de-skewing each frame scan data as pre-processing, and position prediction of the next frame.

Compared to car applications, the hand-held system is dynamically changing its motion in a three-dimensional space. The proposed system performs 3D mapping from arbitrary sensor motion using estimated sensor position and scanned point cloud data. And it improves the precision for local rotary movement. The proposed method has following three steps.

1) De-skew: The skew of the point cloud data affects accuracy, so the skew should be minimized as much as possible. However, especially in the case of rapid rotation movement, the skew of the point cloud becomes large. So, we de-skew each point $\boldsymbol{p}_{i,t}$: i-th point at time t between the scan start time t_1 to end t_2, in the current scan $\boldsymbol{P}_{scan}$: has n points, by transforming it with a gyro rotation measurement $\Delta\boldsymbol{R}_{gyro,t}$ at time t. Here, $\boldsymbol{p}'_{i,t}$ is a raw point data:

$$\boldsymbol{p}_{i,t} = \Delta\boldsymbol{R}_{gyro,t}\ \ \boldsymbol{p}'_{i,t} \tag{3}$$

2) Position estimation: The ICP registration method requires a correctly estimated initial pose $\boldsymbol{x}'_k$ at frame k so as to avoid a local optimum solution. It searches the nearest neighbor (NN) points $\boldsymbol{P}_{NN}$ in the map with the current scan transformed into the initial pose. It is not possible to choose the correct NN points when there is a large difference between the scan and the map, so the accuracy of $\boldsymbol{x}'_k$ is important. In the previous method, the initial pose is generated with a constant velocity model. In the proposed method, $\boldsymbol{x}'_k$ is determined by the integrated value of the gyro rotation measurements ΔR_{gyro}:

$$\Delta\boldsymbol{R}_{gyro} = \int_{t_1}^{t_2} \Delta\boldsymbol{R}_{gyro,t} \tag{4}$$

and making a constant velocity translation $\Delta\boldsymbol{t}_{cv}$:

$$\boldsymbol{x}'_k = \Delta\boldsymbol{R}_{gyro}\ \boldsymbol{x}_{k-1} + \Delta\boldsymbol{t}_{cv} \tag{5}$$

3) Registration: With the current scan $\boldsymbol{P}_{scan}$ and initial pose $\boldsymbol{x}'_k$, we can calculate the ICP of a 6DoF self pose $\boldsymbol{x}_k$ and a 3D environment map $\boldsymbol{P}$. The ICP algorithm is the same as the one described in [9] and is similar to the one in [10]; $\boldsymbol{x}_k$ and $\boldsymbol{P}$ are estimated by minimizing the energy:

$$E(\boldsymbol{x}'_k) = \sum_{i=0}^{n} (\boldsymbol{N}_i^{NN}\ (T(\boldsymbol{p}_i,\ \boldsymbol{x}'_k) - \boldsymbol{P}_i^{NN}))^2 \tag{6}$$

where $\boldsymbol{P}_i^{NN}$ denotes the nearest point to $T(\boldsymbol{p}_i,\ \boldsymbol{x}'_k)$ in $\boldsymbol{P}$, and $\boldsymbol{N}_i^{NN}$ is the estimated normal vector of $\boldsymbol{P}_i^{NN}$. The 3D environment point cloud map $\boldsymbol{P}$ is created by repeating these three steps.

III. Performance Evaluation Experiments of the Proposed Method

We generate the 3D map by both of the previous method and the proposed method at various environment, and evaluated the effectiveness of the proposed method by compared the generated 3D maps.

A. Experimental setup

Fig.4 shows the hand-held sensor unit used for the experiments. The unit is connected to a notebook PC, and performed the experiment with the hand-held sensor unit by carrying them in a bag. The 3D-LIDAR is a HDL-32E. The sensor has a column of 32 laser sensors. The unit can turn around 360 [deg] and cover a pitch range of 40 [deg]. The IMU (Inertial Measurement Unit) is a NavChip manufactured by InterSense. Four MoCap marker balls are attached to measure the true sensor position and orientation. A commercial motion capture system (Motion Analysis Eagle) with 18 cameras measures robot position in 240[Hz] as a ground truth. Standard deviation of sensor position measured by this MoCap system is 0.042[mm] in translation and 1.09e-5[deg] in rotation.

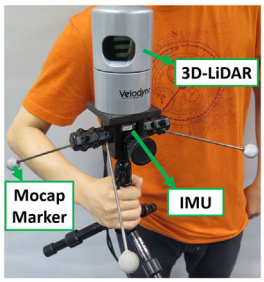


Fig. 4. Image of the sensor unit

The following section examine the performance of the proposed system in two ways. The first experiment is evaluation of the sensor position estimation using MoCap as a ground truth. The second experiment is for generated 3D map and evaluate the distortion of the known ground plane. The comparison results between Velodyne SLAM and the G-ICP SLAM (proposed) are performed in both experiments. Finally, some demonstrations in different environments are also shown.

B. Performance evaluation of sensor position estimate

This section evaluates the mapping accuracy as sensor position estimation error compared to the MoCap measurement. We walked in the motion capture area in circle while adding 6DoF rotary movement to the hand-held sensor unit like as shown in Fig.5. Added 6DoF rotary movement increased with time.

We show the generated 3D environment map by the previous method in Fig.6(a) and show that of the proposed method in Fig.6(c). The gradation of the color shows the difference of the height, and the blue part shows the floor area. The map is generated by the previous method and shows that floor part intersects, and they are reflected diagonally; the form of the passage is not generated correctly. This problem meant that the position estimate for rotary movement was not performed correctly and continued to reflect point cloud data to the map with the wrong sensor position estimate. On the other hand, with the map generated by the proposed method, point cloud data is included walls of the straight passage are reflected correctly.

Fig.7 shows the error of estimated sensor orientation. The y-axis shows the angle difference between the predicted direction and the ground truth of MoCap measurement. The x-axis shows time. They show, the error became larger with time, and that of the previous method is larger than the proposed method in after part that the movement increased. The main reason is that the system gives wrong sensor position as a initial prediction, when the sensor position prediction is not performed correctly on the rotary movement. The difference

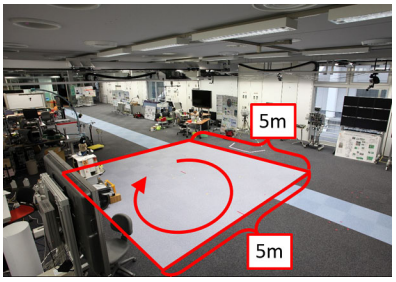
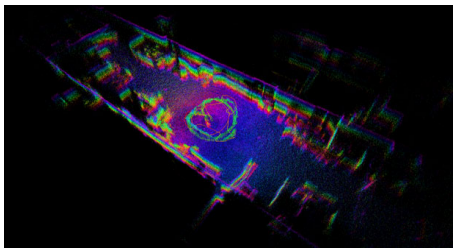
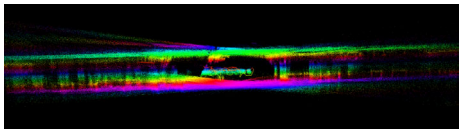


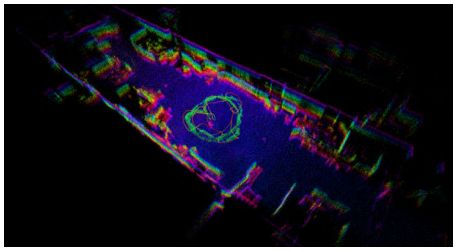
Fig. 5. Image of MoCap area



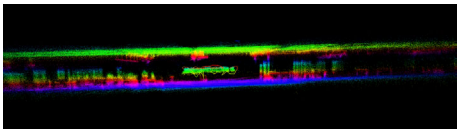
(a) The bird view (Conventional)



(b) The cross-section (Conventional)



(c) The bird view (Proposed)



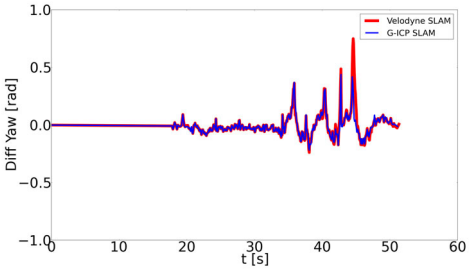
(d) The cross-section view (Proposed)

Fig. 6. Generated 3D map image at MoCap area

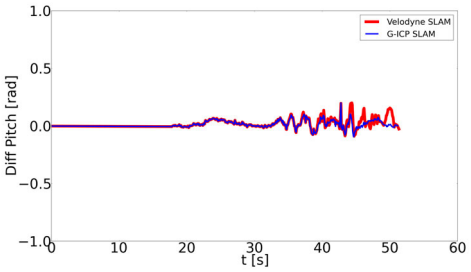
made the wrong reflection of the point cloud data on the generated map. It is considered that the proposed method enables the performing of a robust self-position estimate of rotary motion and works for fast and large change of sensor position and orientation.

C. Evaluation of point distribution belonging to plane

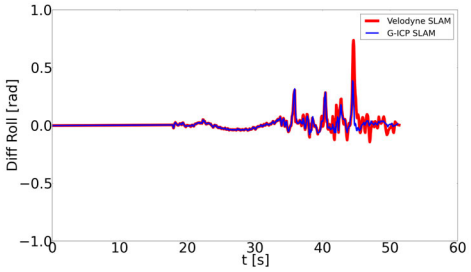
This section evaluates the mapping accuracy as the point distribution belonging to the plane. If the estimated sensor position and orientation are performed correctly, we can reflect its surrounding environment to a map correctly. But if the sensor position and orientation estimate are performed incorrectly on an original plane part, the additional frame is updated diagonally. Then, the point distribution belonging to the plane may be dispersed. If the thickness of the plane



(a) Yaw direction turn



(b) Pitch direction turn



(c) Roll direction turn

Fig. 7. The estimated sensor motion error compared to MoCap measurement

is smaller, we can consider that the sensor position and orientation estimate may have been performed correctly. It is because the ideal thickness of the plane should be zero. In this paper, for the plane part with the least irregularities, we extract floor part. And we calculated the standard deviation σ[mm] of the thickness of the extracted floor part as a point distribution belonging to the plane, and we evaluate whether the sensor position and orientation estimate was performed correctly.

To verify the effect of the proposed method, we calculated the thickness of the plane and evaluated the generated maps of motion capture area by the previous method, the proposed method, and the true value obtained by motion capture. This data is same as the data of preceding paragraph. In addition, we walked and ran in the car parking lot with the hand-held sensor unit. Outdoor experiment data is easily influenced by external factor e.g. the moving cars, the working human, and the trees swaying by the wind. And the hand-held unit may move faster to rotary motion between running experiment than working. We show the experimental parking in Fig.9. We calculated the thickness of the plane of the yellow area in Fig.9 in the same way as the motion capture area experiment.

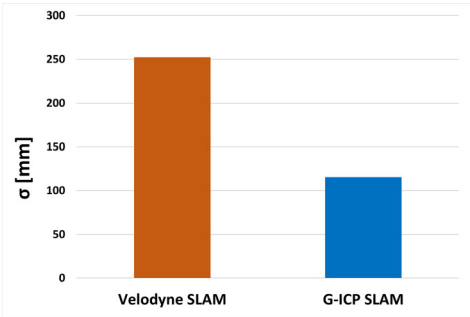


Fig. 8. Standard deviation of plane thickness of MoCap area experiment

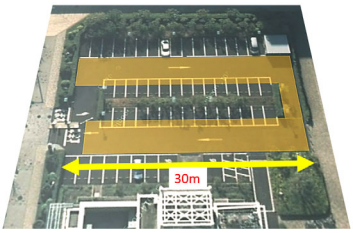


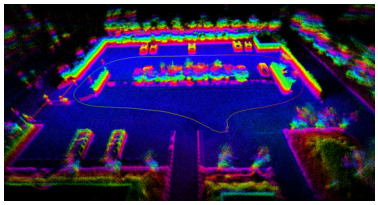
Fig. 9. Image of the car parking lot

The cross-section image of the motion capture area of the 3D map generated by the previous method is shown in Fig.6(b), by the proposed method in Fig.6(d). The calculated thickness of the plane indoors is shown in Fig.8.

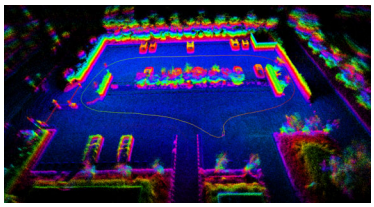
The 3D map of the walking experiment generated by the previous method is shown in Fig.10(a) and by the proposed method in Fig.10(b). The cross section-image of the parking area by walking for the 3D map generated by the previous method is shown in Fig.11(a) and by the proposed method in Fig.11(b). The 3D map of the running experiment generated by the previous method is shown in Fig.12(a) and that generated by the proposed method is shown in Fig.12(b). The cross-section image of the parking area by the running experiment of the 3D map generated by the previous method is shown in Fig.13(a) and by the proposed method in Fig.13(b).The calculated thickness of the plane outdoors is shown in Fig.14.

As a result, for both indoors and outdoors, the thickness of the down floor by the proposed method became less than that by the previous method. It shows that the proposed method enables the performing sensor position and orientation estimate correctly by using a laser sensor and 6DoF IMU sensor at a place without a large machine to measure sensor position and orientation correctly e.g. motion capture.

At the outdoor experiment, for both of the walking and running experiments, the thickness of the down floor decreases. In particular, the result for running is a large reduction. For the walking experiment, it is difficult for fast and large change of sensor position and orientation to occur. The previous method intended to perform SLAM with a sensor on a car, so the change of sensor position and orientation when walking is able to be dealt with by the previous method. On the other hand, for the running experiment, fast and large rotary movement must occur. Therefore, the proposed method enables a decrease of the thickness by improving the performance of the sensor position prediction.

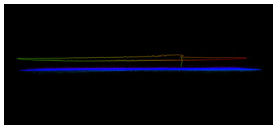


(a) Velodyne SLAM (Conventional)

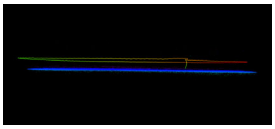


(b) G-ICP SLAM (Proposed)

Fig. 10. Generated 3D map image by walking

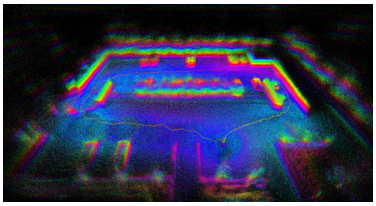


(a) Velodyne SLAM (Conventional)

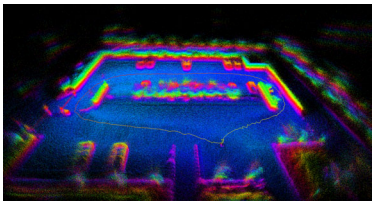


(b) G-ICP SLAM (Proposed)

Fig. 11. The cross-section image by the walking experiment

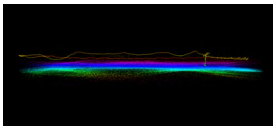


(a) Velodyne SLAM (Conventional)

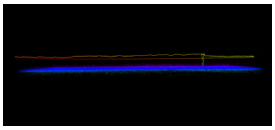


(b) G-ICP SLAM (Proposed)

Fig. 12. Generated 3D map image by running



(a) Velodyne SLAM (Conventional)



(b) G-ICP SLAM (Proposed)

Fig. 13. The cross-section image by the running experiment

D. Distance of origin of sensor

In addition, for the outdoors experiment, we calculated the coordinates of the sensor position estimate at each time and calculated difference of the distance of the origin of the sensor between the previous method and the proposed method for both the walking and running experiment. The origin

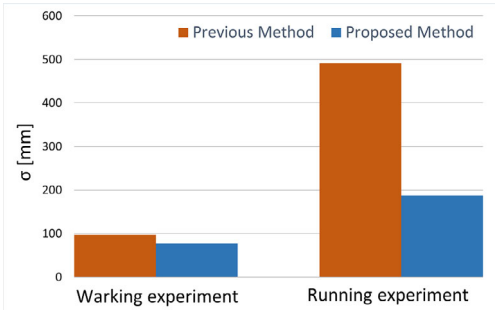


Fig. 14. Standard deviation of plane thickness of car parking lot experiment

of the sensor is the position at the start of the experiment. The result of the walking experiment is 270 [mm] at most, and the average is 170 [mm]. However, the result of the running experiment is 2320 [mm] at most, and the average is 1090 [mm]. The difference of the running experiment is clearly larger than that of the walking experiment. In the context of taking into account whether the sensor position estimate is equal to the true value, we can consider that the proposed method enables improvement of the performance of the sensor position estimate as can be seen in the results of preceding paragraph.

E. Generated map at various situations

This section shows the 3D maps we generated by the proposed method in various situations. The map of the paths of our laboratory is shown in Fig.15. And a map of the steps of our laboratory is shown in Fig.16. For visibility, the roof in Fig.15 and part of wall in Fig.16 have been removed. The shape of the objects in scanned space and steps could be reflected clearly by the proposed method.

IV. Conclusion and Future Works

This paper proposes a 3D mapping system using a laser scanner and a triaxial inertial sensor. The proposed method can correspond to 6DoF movement by hand-held sensor unit is robust to rotary motion, and works for fast and large changes of sensor position and orientation. In accordance with the result of the performance inspection experiment of the previous study "Velodyne SLAM", we show by examination that the previous method cannot perform matching point cloud data to local rotary movement.

In addition, we compared the generated 3D map by the previous method without inertial information and the proposed method with added inertial information. As a result, we confirmed the improvement of the self-position estimate and mapping performance, especially for fast and large rotary movement outdoors.

We consider that our study has large extensibility because that a 3D map can be generated by using only a hand-held sensor unit and notebook PC regardless of the person and environment. Realization of a 3D map that is generated by the proposed method to assist path planning for autonomous mobile robots is likely. Now the proposed method uses all of the point cloud data that is obtained by the laser sensor. Thus, moving objects in scanned space e.g. humans

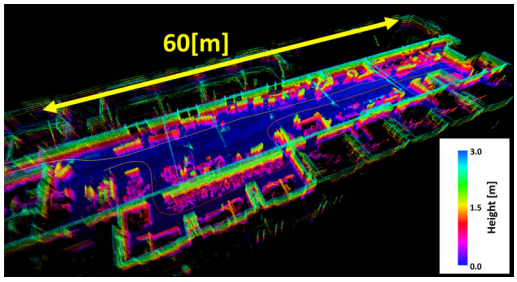


Fig. 15. Generated 3D map at paths of our laboratory

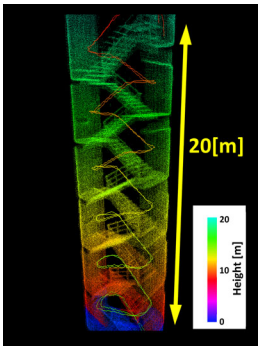


Fig. 16. Generated 3D map at steps of our laboratory

and cars are shown on the generated map for each frame. Therefore, if classify moving objects and static objects is enabled and only static objects are shown on the map, we can generate a 3D environment without noise. This means that if the effects of moving objects are removed to perform self-position estimate, then the performance will be improved. That is our next task.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 24220006.

References

[1] Durrant-Whyte, Hugh, and Tim Bailey. "Simultaneous localization and mapping: part I." Robotics and Automation Magazine, IEEE 13.2 (2006): 99-110.
[2] Leonard, J. J., and Feder, H. J. S., A computationally efficient method for large scale concurrent mapping and localization, In Robotics Reserch-International Symposium, Vol. 9, (2000) , pp. 169-178.
[3] Grisetti, G., Stachniss, C. and Burgard. W., Improved techniques for grid mapping with rao-blackwellized particle filters, Robotics, IEEE Transactions, Vol. 23, No. 1, (2007), pp. 34-46.
[4] Rusinkiewicz, S. and Levoy, M., Efficient variants of the icp algorithm, 3-D, Digital Imaging and Modeling, 2001, Proceedings Third International Conference, IEEE, (2001), pp. 145-152.
[5] Nchter, A., Lingemann, K., Hertzberg, J. and Surmann, H. 6D SLAM —3D mapping outdoor environments. J. Field Robotics, vol. 24, 699722, (2007).
[6] S. Thrun, W. Burgard and D. Fox. "Probabilistic Robotics," The MITPress, 2005.
[7] Grisetti, Giorgio, Cyrill Stachniss, and Wolfram Burgard. "Improving grid-based slam with rao-blackwellized particle filters by adaptive proposals and selective resampling." Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on. IEEE, 2005.
[8] Moosmann, F. and Stiller, C., Velodyne slam, 2011 IEEE Intelligent Vehicles Symposium(IV), IEEE, (2011), pp. 393-398.
[9] Chen, Yang, and Grard Medioni. "Object modeling by registration of multiple range images." Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on. IEEE, (1991).