The Twenty-Fifth International Symposium on Artificial Life and Robotics 2020 (AROB 25th 2020),
The Fifth International Symposium on BioComplexity 2020 (ISBC 5th 2020),
B-Con Plaza, Beppu, Japan, January 22-24, 2020

3D Simultaneous Localization and Mapping for AUV by RealTime 3-D Perception

Yuya Okada[1], Yuichiro Toda[1],Yoshiki Kanda[1], and Mamoru Minami[1]

[1]Okayama University, Japan
(Tel: 81-86-251-8233, Fax: 81-86-251-8233)
[1]phur9piu@s.okayama-u.ac.jp

Abstract: In order for autonomous underwater vehicle (AUV) to perform an unmanned exploration efficiently in an unknown environment, accurate self-localization is required. Simultaneous Localization and Mapping (SLAM) is useful for self-localization in an unknown environments. SLAM using laser scanner is best effective on the land because of high-speed and accurate measurement. However, since the attenuation of radio wave and light is large in the water, it is difficult to accurate estimate position. And, the sequential map update in real time is required for the autonomous navigation in the unknown environment of the robot. In the previous work, we proposed Projection-based Matching method as a new space recognition method only by image information of stereo-vision. Since Genetic Algorithm (GA) is used for detection of position and orientation in this method, real time space recognition is possible. In this paper, we apply space recognition by proposed Projection-based Matching method to SLAM and verify the effectiveness from the result of verification experiment.

Keywords: SLAM, Stereo-vision, Projection-based Matching method, Genetic algorithms

1 INTRODUCTION

Recently, research and development of autonomous underwater vehicle (AUV) has been made for various purposes such as resource excavation and geographical environmental research. For the efficient execution of such work, long-term unmanned search activity of AUV is indispensable. However, the research of the autonomous underwater robot is still in the developing stage, so there is a restriction in the continuous navigation time. In previous work, our research group has been studying the construction of an automatic charging system to extend the active period of AUV. In this system, the AUV equipped with the two cameras recognizes the power supply station in the sea and autonomously docks it. We succeeded in the automatic docking in the actual sea area even under turbid environment and dark environment [1][2]. Accordingly, we strive to have AUV autonomously return to the vicinity of the station by measuring the surrounding environment. The aim of this research is to extend the autonomous navigation of the AUV by combining SLAM route guidance and the docking system as shown in Fig. 1.

For autonomous navigation of AUV in unknown environment, accurate estimation of self position in real time is required. However, there is a problem that it is more difficult to estimate self position than on land because GPS is not available in the sea. Therefore, a map of the surrounding environment is important for AUV to know self position and determine its route in water. A solution of this problem is Simultaneous Localization and Mapping (SLAM). SLAM is a technology to carry out self position estimation and environment mapping sequentially and simultaneously. In SLAM on land, a laser scanner, a camera, and an encoder are generally used as sensors. A laser scanner is suitable for high-speed and accurate measurement, but it is not suitable in water because of large attenuation of visible light and radio waves. Therefore, Optical cameras are usually used to obtain information about the underwater environment.

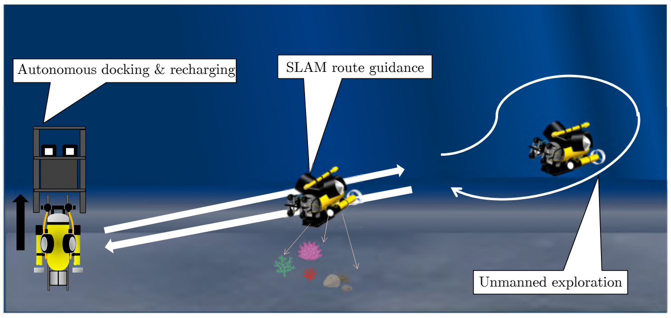


Fig. 1. AUV's approach to unmanned exploration combined with autonomous docking system and SLAM route guidance

In the previous work, we proposed Projection-based Matching method as a new space recognition method only by image information of stereo-vision [3]. Since Genetic Algorithm (GA) is used for detection of self position and orientation in this method, real time space recognition is possible. And, since no prior knowledge is required, there is an advantage that it can be operated even in an irregular environment such as water, which is named as "Real-Time 3-D Perception(RT-3DP)". In this paper, we apply space recognition by proposed Projection-based Matching method to SLAM and discuss the effectiveness from the result of verification experiment.

2 3D RECOGNITION BY PROJECTION-BASED MATCHING METHOD

Projection-based Matching method was proposed as a real time space recognition method using only image information by stereo-vision. Space recognition in video is realized by estimating the object in a still image continuously within a video rate (33[ms]). Therefore, we describe the summary of the recognition method for a still image in this section. Figure 2 shows a schematic diagram of the process of this recognition method. In this process, search model is generated from the center of the left camera image. The pose of 3D model $\boldsymbol{\phi} = (x, y, z, \varepsilon_3)$ (ε_3 is the angle about the z-axis.) is determined by the gene of Genetic Algorithms (GA) in real time. Based on the pose, the solid model is projected on to the left and right camera images. Comparing the projected solid model with the captured images by two cameras, the relative position and orientation difference can be calculated. In this comparison, the degree of correlation between the solid model and the real target in the image is evaluated by fitness function. Fitness value reaches maximum when the 3D pose $\boldsymbol{\phi}$ coincides that of the real object. Therefore, the problem that searching for the pose $\boldsymbol{\phi}$ of an object can be transposed to the problem that searches for the maximum of the fitness function. In this research, we use GA to get the maximum fitness value in the consecutively input dynamic image sequences by video rate. Hence, when the GA converges to some extent, the pose of the real target can be estimated by regarding the pose of the 3D model [3][4].

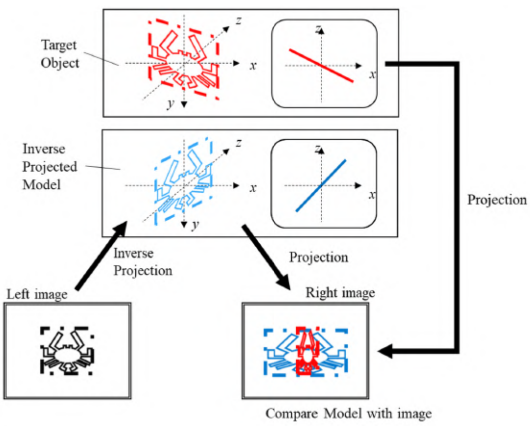


Fig. 2. Projection-based Matching Method process

2.1 Fitness function

Fitness function is a correlation function to evaluate how much degrees of matching between a projected model and a real target object in the images. Degree of correlation is calculated based on hue and brightness. Figure 3 shows a conceptual diagram of the calculation of the degree of correlation by the fitness function. Fitness value F_{union} is defined as the average of evaluation value of hue F_H and evaluation value of Brightness F_B as shown in equation(1). D_H is the sum of the differences of hue between each point of the projected model and the real object on the image. Similarly, D_B is the sum of the brightness's differences. Therefore, the fewer summation of differences of hue and brightness between projected model and real object on the image is, the higher fitness value is.

$$\begin{aligned} F_{union}(\boldsymbol{\phi}_M^j) &= \frac{F_H(\boldsymbol{\phi}_M^j) + F_B(\boldsymbol{\phi}_M^j)}{2} \\ &= \left\{ \frac{1}{1 + \sum_{{}^I\boldsymbol{r}_i^j \in S(\boldsymbol{\phi}_M^j)} D_H({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))} \right. \\ &\quad \left. + \frac{1}{1 + \sum_{{}^I\boldsymbol{r}_i^j \in S(\boldsymbol{\phi}_M^j)} D_B({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))} \right\} /2 \end{aligned} \tag{1}$$

$$D_H({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = |H_I({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_M({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \tag{2}$$

$$D_B({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = |B_I({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - B_M({}^I\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \tag{3}$$

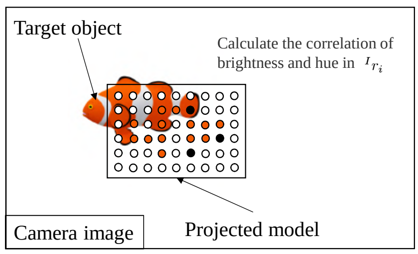


Fig. 3. Conceptual diagram of the calculation of the degree of correlation by the fitness function

3 PROPOSED SLAM

3.1 Algorithm of proposed SLAM

Figure 4 shows flowchart of SLAM by Projection-based Method. First, First, an image sequence is inputted by two cameras and the search model is generated from the center point of the left camera image. The 2D model is projected onto the map and searched where it exists on the map. Next, during the frame rate(33 [ms]), the position of the model on the 2D map is searched, and the self position of (x, y, ε_3) is estimated. At the same time, the self-position of depth direction(z) is estimated using the search model inverse projected on the right camera. This process is repeated for 33[ms], and the search model is updated after 33[ms]. The position and orientation by the estimation are converted to the world coordinate system. Finally, the 3D map is updated using the model based on the calculated self position.

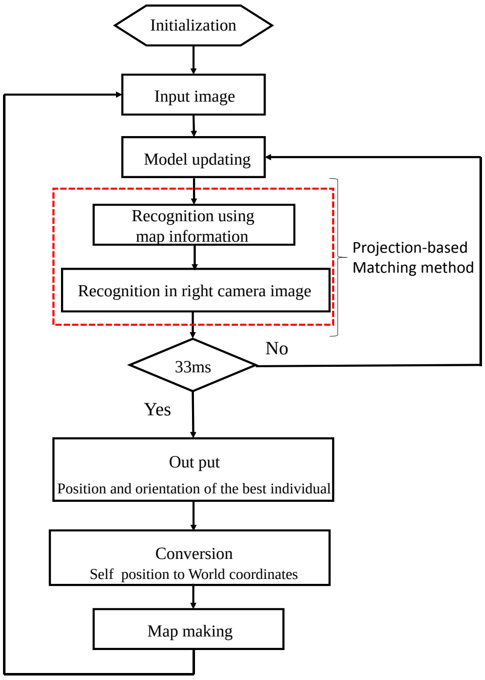


Fig. 4. The flowchart of SLAM using Projection-based Matching Method

3.2 Transformation to World Coordinate System

In order to calculate the position and orientation of the robot in the world coordinate system, it is necessary to accumulate the movement amount of the position and orientation obtained by the estimation. Figure 5 shows a conceptual diagram of a robot estimating its own position while moving. For simplicity, the robot coordinate system at n[s] is Σ_n. The homogeneous transformation matrix from Σ_{n-1} to Σ_n is defined as

$$ {}^{n-1}\boldsymbol{T}_n = \left[\begin{array}{c|c} {}^{n-1}\boldsymbol{R}_n & {}^{n-1}\boldsymbol{r}_{n-1,n} \\ \hline \boldsymbol{0} & 1 \end{array}\right] \tag{4} $$

where ${}^{n-1}\boldsymbol{R}_n$ is the rotation matrix calculated by the orientation change; ${}^{n-1}\boldsymbol{r}_{n-1,n}$ is translation vector; These are measured by two cameras at every model update.
The current position and orientation of the robot in the world coordinate system is calculated by the multiplication of the homogeneous transformation matrix from the initial position to the current as following equation:

$$ \left[\begin{array}{c|c} {}^{W}\boldsymbol{R}_n & {}^{W}\boldsymbol{r}_n \\ \hline \boldsymbol{0} & 1 \end{array}\right] = {}^{W}\boldsymbol{T}_0 \cdot {}^{0}\boldsymbol{T}_1 \cdots {}^{n-1}\boldsymbol{T}_n \tag{5} $$

where ${}^{W}\boldsymbol{R}_n$ is a rotation matrix indicating the orientation of the robot in the world coordinate system, and the orientation is calculated by back calculating this matrix. Additionally, ${}^{W}\boldsymbol{r}_n$ is the position of the robot in the world coordinate system. By performing this process at every model update, the robot's position and orientation in the world coordinate system can be calculated.

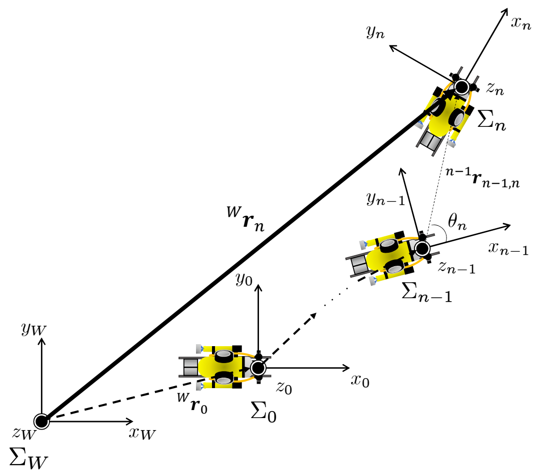


Fig. 5. Transformation to the world coordinate system by the homogeneous transformation matrix

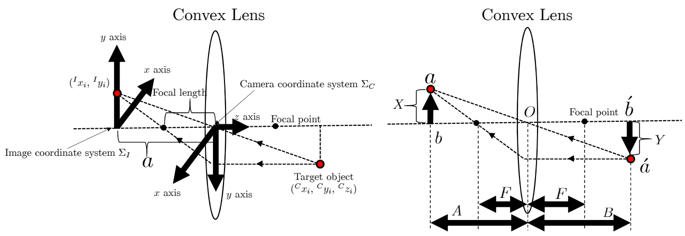
3.3 Update Map

When estimating the self position of (x, y, ε_3), the search model is compared with the 2D map while making sequential 2D map. To update 2D map, the position of each evaluation point of the projected model in the map coordinate system is calculated and is projected on the 2D map. The x and y coordinates of the model in the map coordinate system $({}^{M}x_i, {}^{M}yj)$ are given by the following equations:

$$ \begin{bmatrix} {}^{M}x_i \\ {}^{M}y_j \end{bmatrix} = \frac{1}{R_{map}} \left({}^{W}\boldsymbol{R}_n \boldsymbol{P} \begin{bmatrix} -120 + i * e_1 \\ -110 + j * e_2 \\ \hat{z} \\ 1 \end{bmatrix} + \begin{bmatrix} \hat{x} \\ \hat{y} \\ 0 \\ 0 \end{bmatrix} \right) \tag{6} $$

$$ \boldsymbol{P} = \frac{1}{\hat{z}} \begin{bmatrix} f/\eta_x & 0 & 0 & 0 \\ 0 & f/\eta_y & 0 & 0 \end{bmatrix} \tag{7} $$

where f is the focal length, η_x and η_y are the ratios of the pixel and actual length of the image sensor, and f = 4.2 [mm], η_x and η_y = 0.005365[mm/pixel]. $\boldsymbol{P}$is projective transformation matrix derived based on the formula of the lens as shown Fig.6; e_1, e_2 are the distances between the evaluation points of the model in the x and y directions [px]; $\hat{x}$,$\hat{y}$,$\hat{z}$ is an estimate of the self position; R_{map} is the scale of the map.



(a) (b)

Fig. 6. Projection schematic diagram

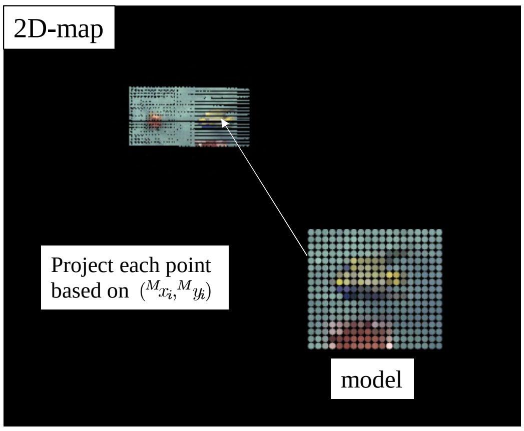


Fig. 7. Conceptual diagram of projection of a model onto a 2D map

4 VERIFICATION EXPERIMENT

4.1 Experimental environment

We reproduces the movement near the seabed by moving a robot manipulator equipped with two cameras with a cloth simulating the seabed as a background as shown Fig8. The center point of the two cameras is defined as the hand coordinate system Σ_H. The world coordinate system Σ_W is defined as the coordinate system at which the system is started. The camera has a focal length of 4.2 [mm], and both cameras are placed at an angle of 14 [$^\circ$] inward. The distance between cameras is $323.4[mm]$ and the distance to the object is 800 [mm]. Figure 9 shows the two kinds of backgrounds used in this experiment and the movement path of the manipulator against each. Move the manipulator in the following order: 250 [mm] in the positive x-axis direction, 150 [mm] in the positive y-axis direction, 250 [mm] in the negative x-axis direction, and 150 [mm] in the negative y-axis direction against the background (A). Furthermore, to confirm orientation estimation , rotate the manipulator in the following order; -20 [$^\circ$] around the z-axis, + 40 [$^\circ$] around the z-axis Without changing the position against background (B). We take above operations and reproduce the movement near the seabed. Using this video, we conduct localization and mapping 10 times each background .

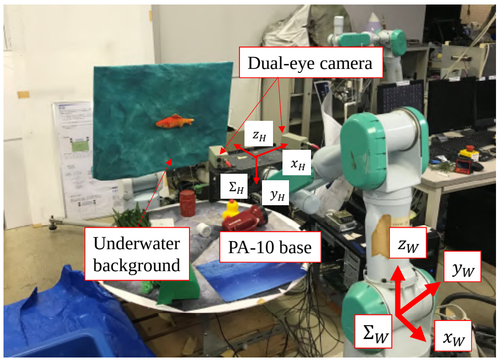
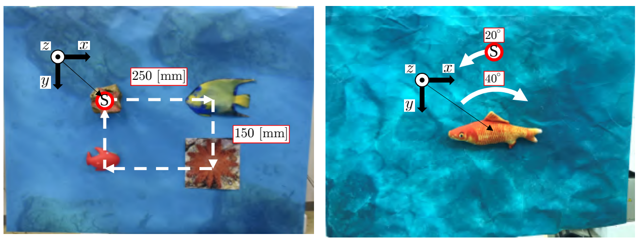


Fig. 8. Experimental environment



(A) (B)

Fig. 9. Background imitating the seabed used for experiments and movement path of the manipulator for each

4.2 Experiment Results

Figre 10 and 11 show the results of position and orientation estimation against background (A) and (B). In both figure, the black line is the trajectory obtained by estimation, and the red line is the actual trajectory. The estimated positions are consistent with the actual trajectories, although there is a steady-state error of about 20 $\sim$ 40 [mm]. Furthermore, on the orientation estimation, the estimated orientation converges around the actual orientation. Table.1 show the average errors and the standard deviations of (x, y, z, ε_3) from the actual values obtained from 10 experiments against each background. Comparing background (A) and (B), the average of error against background (A) having more complicated path than background(B) is more large. From this fact, it seems that the accumulation error of GA increases as the route becomes more complicated. This error is considered acceptable because the docking system is fitted from a carriage 600 [mm] away from the station, however it is necessary to consider using stochastic methods such as Bayesian filters to eliminate errors.

Next, the front and side views of the 3D map are shown in Fig.12 and 13. Mapping results can be used to determine the pattern of the seafloor and the color of the marine biological model. Moreover, It can also be seen that the positional relationship of each model is in agreement with the actual one. Therefore, it is possible to judge the pattern and landform of the seabed from the mapping result, and distinguish characteristic patterns of objects to be searched.

5 CONCLUSION

In this paper, we applied Projection-based Matching Method to SLAM. At first, we explained 3D Recognition by Projection-based Matching method. Next, we proposed SLAM using Projection-based Matching Method an explained the algorithm of it. In the verification experiment, it was confirmed that localization and mapping by the proposed method. Therefore, we will conduct the route guidance of AUV by the proposed method.

REFERENCES

[1] Khin Nwe Lwin, Naoki Mukada, Myo Myint, Daiki Yamada, Mamoru Minami, Takayuki Matsuno, Kazuhiro Saitou, Waichiro Godou, "Docking at Pool and Sea by Using Active Marker in Turbid and Day/Night Environment", Artificial Life and Robotics, DOI 10.1007/s10015-018-0442-1

[2] Daiki Yamada, Naoki Mukada, Myo Myint, Khin Nwe Lwin, Takayuki Matsuno, Mamoru Minami, "Docking Experiment in Dark Environment Using Active/Lighting Marker and HSV Correlation," Proceedings of IEEE/MTS Techno-Ocean Conference 2018, Paper No.171129-098, Kobe, Japan, May 29-31,2018

[3] Yejun Kou, Hongzhi Tian, Xiang Li, Mamoru Minami, "Projection-based Visual Sevroing System", The society of Computational Intelligence 13th, CI13-02, 2018.6.16-17

[4] Taro Yamamoto, Ryuki FUNAKUBO, Yejun Kou, Hongzhi Tian, Xiang Li,, Takayuki Matsuno, Mamoru Minami, "Verification of Estimation Performance of Pose Using Inverse Projection ", JSME Conference on Robotics and Mechatronics (ROBOMECH 2018), 2A1-M15, 2018.6.2-5

[5] Masahiro Tomono, SLAM nyumon（Introduction of SLAM）(2018), Ohmsha, Ltd.

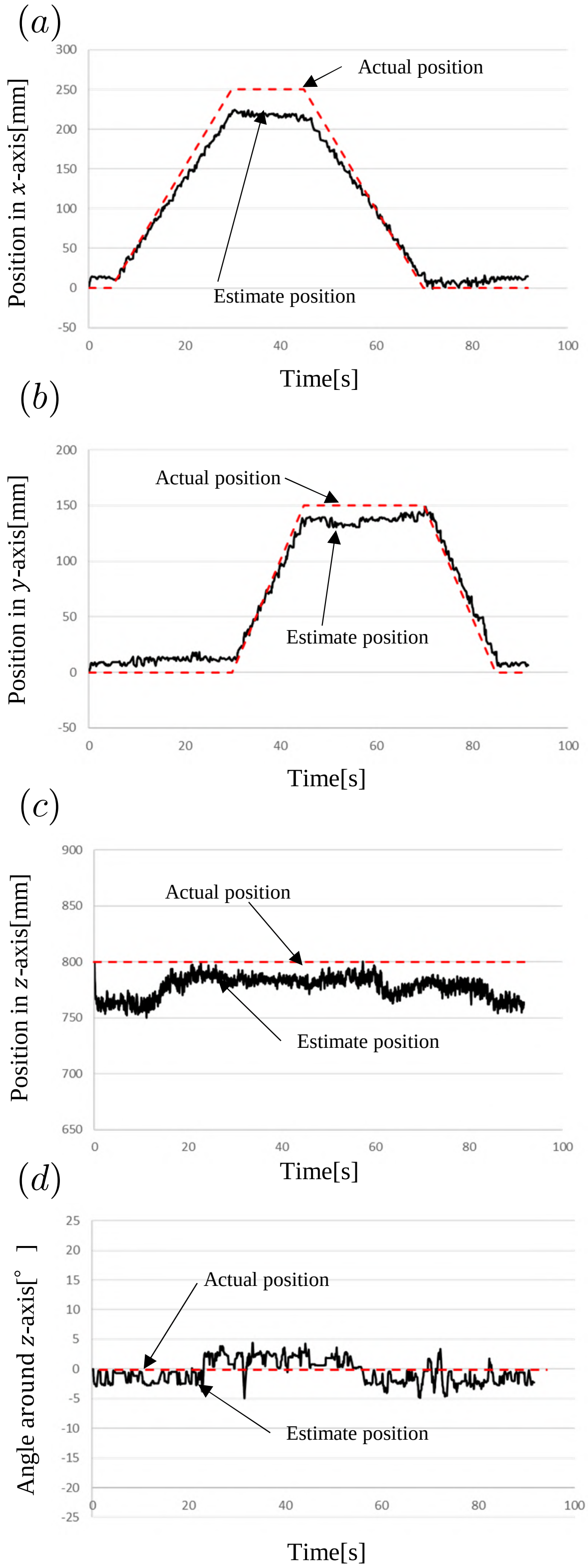


Fig. 10. Estimate position along x-axis(a),y-axis(b),z-axis(c), and orientation; ε_3(d) against background (A)

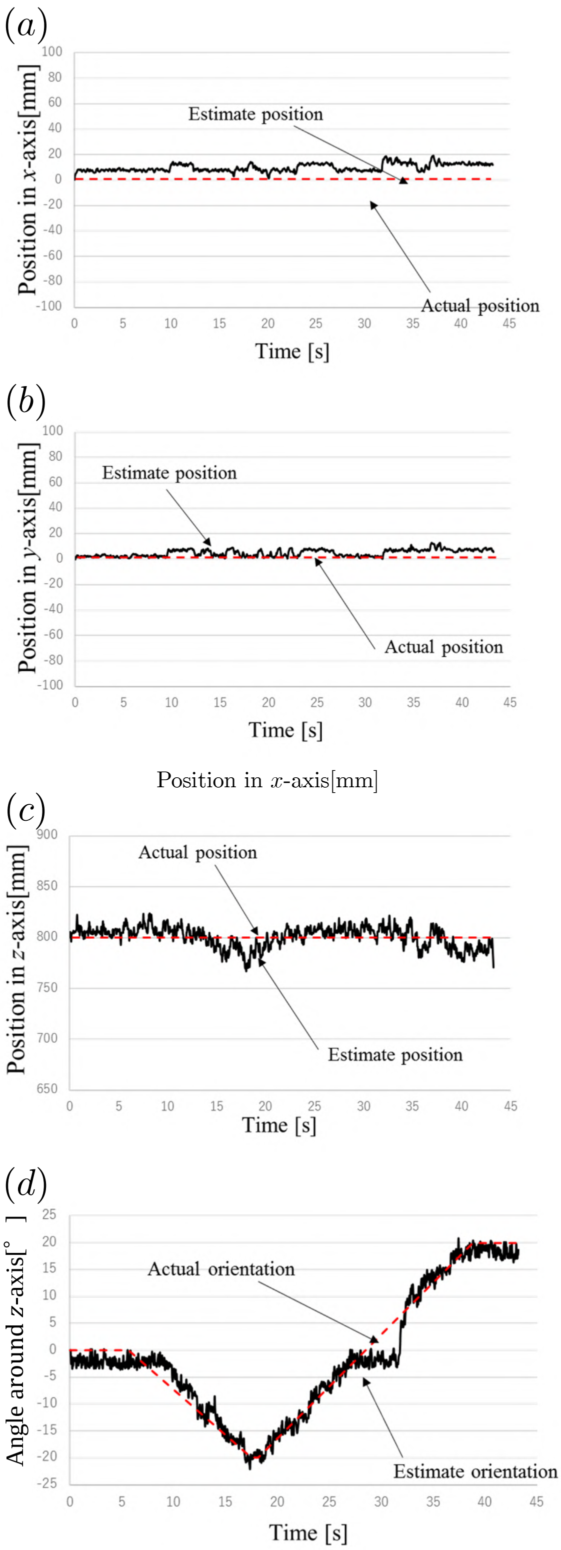
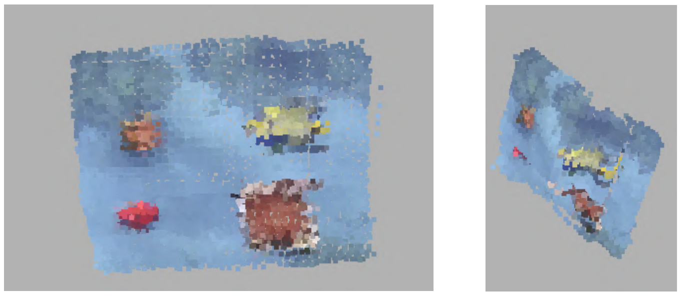


Fig. 11. Estimate position along x-axis(a),y-axis(b),z-axis(c), and orientation; ε_3(d) against background (B)

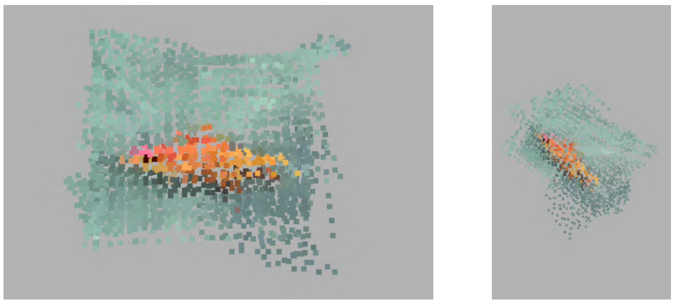
Table 1. Average of Error and standard deviation of 10 experiments conducted against each background

	Average of error		average of standard deviation of error	
	(A)	(B)	(A)	(B)
x-axis $[mm]$	15.5	8.8	11.5	3.5
y-axis $[mm]$	10.6	6.4	5.7	2.9
z-axis $[mm]$	20.7	11.1	9.7	9.8
ε_3 $[^\circ]$	2.9	2.0	1.7	1.7



(a) Front view (b) Side view

Fig. 12. Experimental result of construction 3D map using two camera images for background (A)



(a) Front view (b) Side view

Fig. 13. Experimental result of construction 3D map using two camera images for background (B)