Improvement of 3D Pose Estimation Abilities by Light-Emitting-3D Marker for AUV Docking

1st Kohei Yamashita
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
p9mk0tyb@s.okayama-u.ac.jp

2nd Hsu Horng Yi
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
pcvv4h2h@s.okayama-u.ac.jp

3nd Daiki Yamada
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
p7kw2h27@s.okayama-u.ac.jp

4rd Naoki Mukada
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
pola9zud@s.okayama-u.ac.jp

5th Khin New Lwin
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
pdoj8yez@s.okayama-u.ac.jp

6th Myo Myint
Technological University Thanlyin
Yangon, Myanmar
ecmyomyint@gmail.com

7th Yuichiro Toda
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
ytoda@s.okayama-u.ac.jp

8th Takayuki Matsuno
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
matsuno@cc.okayama-u.ac.jp

9th Mamoru Minami
Natural Science and Technology
Okayama University
OKAYAMA, JAPAN
minami-m@cc.okayama-u.ac.jp

***Abstract*—Disturbances of turbidity and low illuminance are problems in real sea areas when recognizing objects with cameras. Therefore, the recognition target was made to emit light so that it can be recognized correctly even in that environment. However, a suitable light intensity of the target was not decided and it is obvious that recognition results was changed by light intensity of the target. This paper presents the analysis of recognition accuracy of the Real-time 3D estimation system by changing the current value of each color LED (red, green, blue) under turbid and low illuminance. Recognition experiments were conducted at the distance 600 [mm] between the ROV and 3D marker. The turbidity level was set constant value. The current value was changing from 0 [mA] to 16 [mA] for each LED individually. The best current for each LED was optimized by the fitness value and estimation value of position and orientation. The results showed that the recognition accuracy of the proposed system was improved by using optimized lighting intensity.**

***Index Terms*—Pose Estimation, Turbidity, Active/Lighting 3D marker, LED's current value**

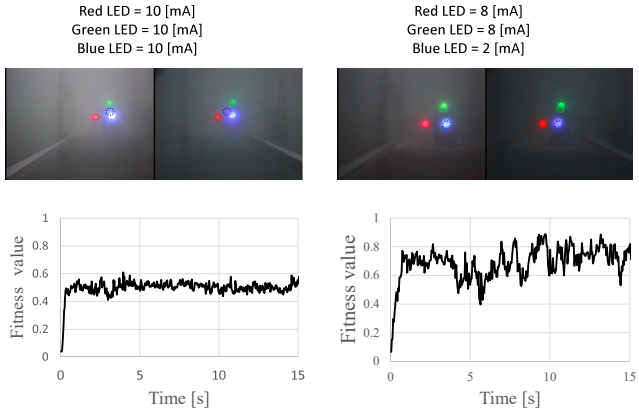

Fig. 1. Images from ROV cameras in two different current value condition.

I. Introduction

Autonomous Underwater Vehicles (AUVs) play an important role in deep sea works such as oil pipe inspection, survey of sea floor, searching expensive metal, scientific studies, etc [1], [2]. These works can be made more effective by long-term continuous operation of AUVs. However, the operation time of AUVs is still limited because a recharging method for AUVs has not been established. Landing and docking operation accompanied by highly accurate control are necessary for charging underwater robot.

The studies on visual servoing based underwater vehicle have been conducted all over the world in recently years. Some references are based on single eye camera to estimate the pose of the target object [3]- [5]. The disadvantage of single eye camera is that the precision of distance measurement of the camera's depth direction is not enough for applications in which high homing accuracy is important. A binocular vision was used in some of these studies in order to estimate the relative pose of the target object in [6], [7]. However, the two cameras look at difference target, then these approaches do not have enough materialize parallactic nature. Therefore, as an initiated research to AUV environment, a 3D-Move on sensing system using stereo vision to provide high homing accuracy have been conducted by our research group. In previous work

978-1-5386-4814-8/18/$31.00 ©2018 IEEE

[8], docking experiments for AUV using stereo vision in simulated pool for sea bottom battery recharging application were succeed. However, the experiments were conducted in relatively clear water. Recognition ability using ROV's light is limited to low turbidity environment. In the real sea area, turbidity of water is made by mad, plankton and marine snow. To expand tolerance against turbidity and low illuminance, a active 3D marker (recognition target) was designed and constructed [9]. Figure 1 shows the recognition image from stereo cameras and transition of fitness value at two different light intensities of active 3D marker (current values). The fitness function expresses the degree of agreement between the search model and the recognition target. This figure shows that the light intensities have a large influence on the recognition result. Then, these have been a room to improve the pose estimation more accurate by adjusting the appropriate light intensity for image recognition process. The remainder of the paper is organized as follow: Section II describes 3D recognition and control. Section III describes Methods of current value determination and experiment results. The final section concludes this paper.

II. Real-time 3D Pose Estimation Method

A. Model-based Matching Method

In this section, recognition system based on 3D MoS using dual-camera for underwater docking especially 3D pose estimation method has been discussed for reader convenience. Please refer to [10] for detailed explanation. In proposed system, the model-based matching method is used to estimate the matching degree between the projected model and the captured images. In other conventional methods, the pose estimation method is implemented by using feature-based recognition based on 2D to 3D reconstruction. In that approach, the set of image points in different images is used to determine the information of the target object. The main drawback is complex for searching the corresponding points and time taken. Apart from this, the model-based pose estimation approach based on 3D to 2D projection is applied in this work avoiding the effects of wrong mapping points in images using dual-eye cameras. Figure 2 shows the model-based matching method using dual-eye cameras for 3D pose estimation. In Fig. 2, Σ_{IR} and Σ_{IL} are the reference coordinate frames of the right and left camera images. Σ_H is the reference frame of the ROV. Σ_M is the reference frame of the real target object. The solid model of the real target object in space is projected naturally to the dual-eye camera images and the dotted 3D marker model where pose is given by one of GA's genes projected from 3D to 2D. The different relative pose is calculated by comparing the projected model and the captured images by the dual-eye cameras. Finally, the best model of the target object that represents the true pose can be obtained based on its highest fitness value. There are some works done on visual-servoing experiments concerning hand eye manipulator in the air using 3D model-based matching method utilizing genetic algorithms and dual-eyes camera [11], [12], which are used as fundamental knowledge for this research.

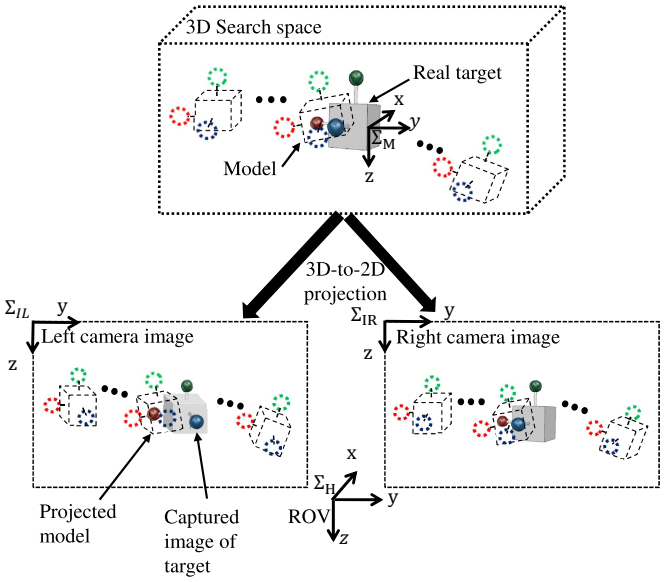

Fig. 2. Model-based matching method using dual-eye cameras and 3D marker

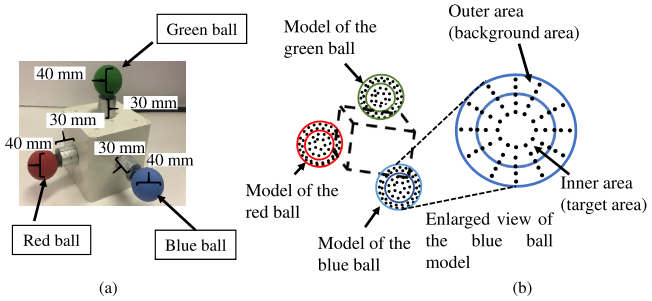

Fig. 3. Real 3D marker and model: (a) active/lighting marker, (b) model with enlarged view of the blue ball model, where the inner area is the same size as the real target object (blue ball) and the outer area is the background area. The dots in enlarged view of blue ball mean points to calculate the correlation degree on how much the inner area overlaps the blue ball and the outer area does not overlap the blue ball.

B. Fitness Function

The fitness function is constructed to evaluate the matching degree between the projected model and the captured image. The good fitness function affects GA to explore the search space and convergence speed more effectively and efficiently. In previous research, only hue value was used for recognition in this system. However, due to the diffused light of the active 3D marker, the range of color extends to the outside of the marker sphere. Therefore, color spread decreases recognition accuracy. Equation (1) shows the fitness function used in this paper.

$$F = F_{hue} + F_{Bd} - F_{hue}F_{Bd} \tag{1}$$

F_{hue} : fitness value from hue information

F_{Bd} : fitness value from Brightness information

In the present studies, besides hue, brightness value was also used for evaluating fitness function. Brightness is the weighted value of each RGB, which is strictly different from Value of HSV. However, the method of calculating fitness value is almost the same as method in our previous study [13]. Figure 3 shows the real target and model of 3D marker. Each model consists of three spherical ball (red, green and blue). Each spherical ball consists of two areas, where the inner area is the same size as the real target object and the outer area is the background area. If a color ball of 3D marker in captured image fits in the inner area of searching model, the fitness value will be increased. On the contrary, if a color ball of 3D marker in captured image overlap in the outer area of searching model, the fitness value will be decreased. Therefore, when the real target matches the model, fitness value is maximized. A concept of the fitness function can be found in our previous study [13].

C. Real-time Multi-step GA (RM-GA)

The genetic algorithm is used as a search and an optimization method to estimate the relative pose between the ROV and 3D marker. Even though there are many powerful optimization methods, we selected GA and modified as Real-time Multi-step GA (RM-GA) because of its simplicity and especially effectiveness in real-time performance. The RM-GA is capable of real time recognition of the true pose of the target through successively input images. Figure 4 shows a genotype of GA population and Fig. 5 (a) shows the flowchart of the RM-GA and Fig. 5 (b) illustrates the behavior of GA convergence from the first generation to the final generation. Please note that although the pose of the target object is evaluated in 2D, convergence occurs in 3D. Position and orientation of the three-dimensional model are represented as 72 bits string of length as shown in Fig. 4. The former 36 bits represent the position of the 3D marker and the later 36 bits represent the orientation defined by a quaternion.

Firstly, a random population of model is generated. A new pair of left and right images captured by ROV's cameras is input. The fitness value of each model is evaluated by using the fitness function. Each model is sorted and selected the better model from the current generation according to the fitness value to reproduce the new generation. Then, again new generations are formed from the two-point crossover and mutation operation of GA. The RM-GA evolves the models with as many generations as possible within the video frame rate for each image. In the present study, the number of evolution times of the RM-GA is chosen to be nine, which is a maximum that the computer used in the present study could calculate within 33 ms (determined by the video frame rate) during the GA evolution process. The RM-GA find repeatedly the solutions to get the optimum value that indicates the best pose of the target object. The convergence performance to an optimum value of the GA's evolution function used as fitness function has been proved mathematically by a Lyapunov analysis in a previous work [14]. The effectiveness of the GA

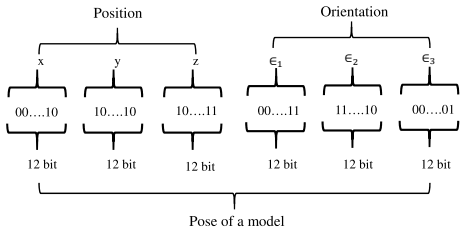

Fig. 4. A genotype of GA population: 12 bits for each x, y, z represents the position coordinate of the three dimensional model of the gene and 12 bits for each $\epsilon_1, \epsilon_2, \epsilon_3$ describes the orientation defined by a quaternion.

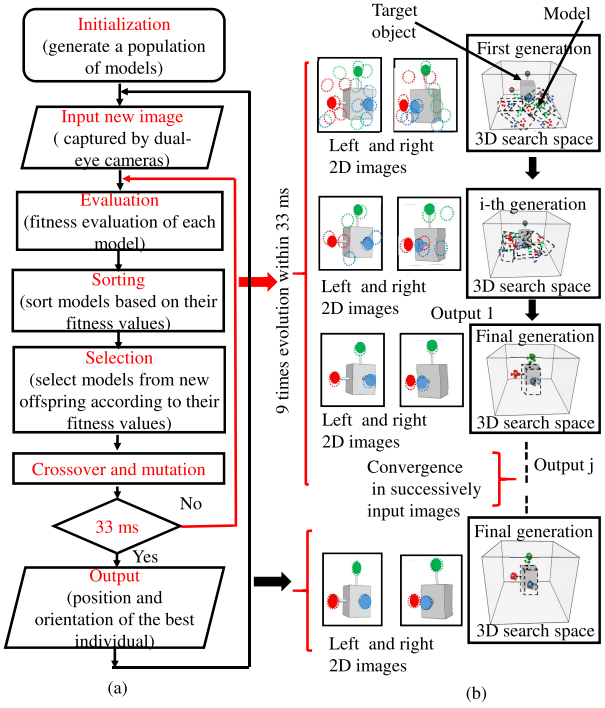

Fig. 5. Flowchart of the real-time multi-step GA: (a) the flowchart of the RM-GA, the true pose of the target object is evaluated within 33 ms through the GA process (b) the convergence behavior of RM-GA from the first generation to the final generation to represent the true pose of the target object in which evaluation is performed in 2D and convergence occurs in 3D.

was demonstrated in a previous study on visual servoing for catching fish using a GA search [15].

D. Active/Lighting Marker

The recognition target is called as a 3D marker in our research group. In our previous research, the passive (not lighting) 3D marker was used to conduct experiments. Therefore, in order to recognize the 3D marker by using dual-camera in the dark turbid environment, it was necessary to illuminate with ROV's light. Then, the whole recognition images became whitish, and recognition accuracy decreased (Fig. 7 ROV's Light). In the present study, the active (lighting) 3D marker was designed and newly constructed to improve the

pose estimation at dark turbid environment (Fig. 7 3D marker Light). The value of each variable resisters and resisters are shown in Fig. 6. Red, green and blue LEDs were installed into the white spherical ball and covered by three color balloon (red, green, blue). It can take in only the light necessary for recognition even without ROV's light in dark turbid environment. However, in previous studies, the lighting intensity for each color LED in 3D marker has not been determined quantitatively. Therefore, it is necessary to experimentally determine suitable current values for recognition.

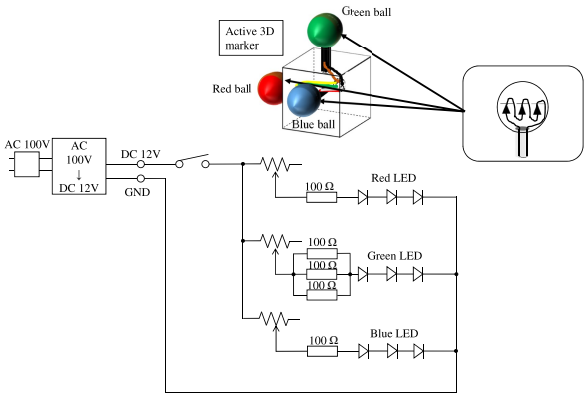

Fig. 6. Active 3D marker and its internal circuit.

III. Experimental Results and Discussion for Detarmination of Suitable Current Value

A. Determined by Fitness Value

1) Experimental Environment: Recognition experiments were conducted in order to quantitatively determine suitable current value of our pose estimation system. Figure 8 shows the experimental layout in an indoor pool (length 750 [mm] × width 570 [mm] × height 490 [mm]) which was filled with 800 liters water. Table I shows the experimental condition. The methods of measuring illuminance and turbidity are shown in [16]. Assuming that AUVs work on the sea bottom, illumination in the water was set to 0 [lx]. The turbidity of water was set to 8.6 [FTU]. The distance from ROV's camera to active 3D marker was set to 600 [mm]. Current value of the LEDs in one color was changed by 1 [mA] from 0 [mA] to 16 [mA], and current values of LEDs in other two color balls were fixed at initial current value (Table I). The initial current value with little diffused light enabled the color and shape to be visible by the human eye. Recognition experiments were conducted for a minute at each current value.

2) Experimental Results : Figure 9 shows the fitness value when current of red LED changed by 1 [mA] from 0 [mA] to 16 [mA]. Figure 9 (b) shows the fitness value of F_{hue}. In these results, although the fitness value of red also increases as the current value increases, it gradually decreases over 4 [mA]. Figure 9 (c) shows the fitness value of F_{Bd}. In this graph, although the fitness value increases as the current value increases, the fitness value converges about 9.5 when the current value exceeds 7 [mA]. Therefore, the red light intensity of active 3D marker is considered to be suitable for recognition when the current value for red LED is 7 [mA]. Figure 10 shows the fitness value when current of green LED

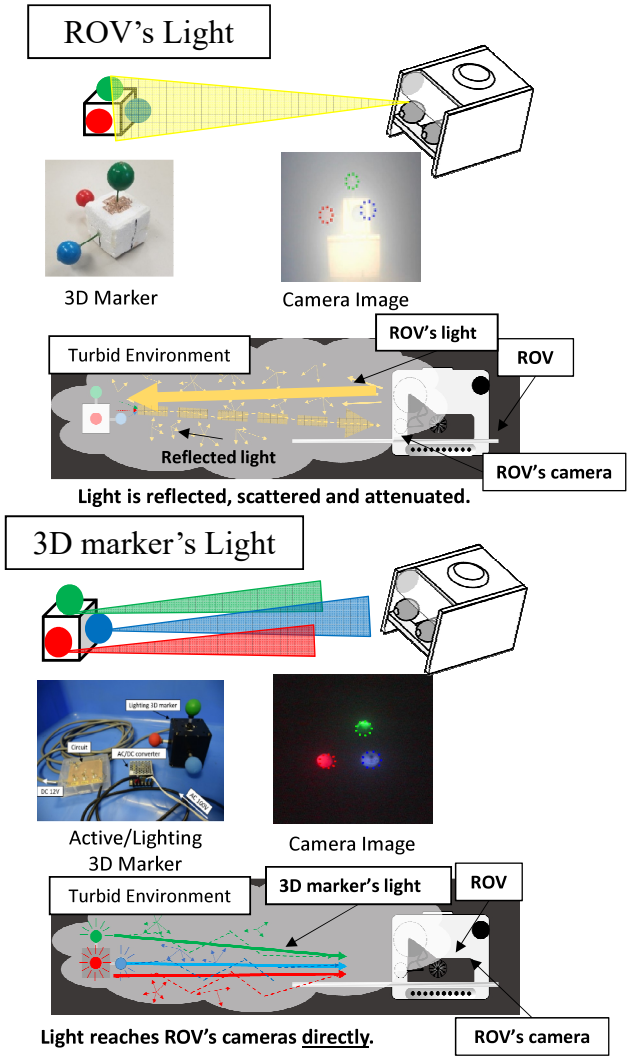

Fig. 7. Compare ROV's light and 3D markers light

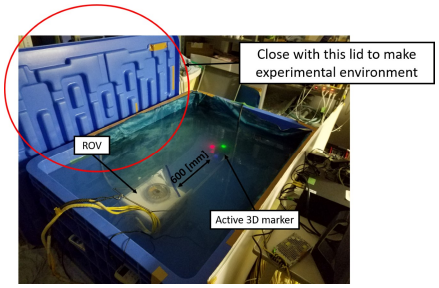

Fig. 8. Experiment environment

TABLE I
EXPERIMENTAL CONDITION

Turbidity [FTU]		8.6
Illumination [lx] (in the water)		0
Distance x [mm]		600
Initial current value [mA]	Red	6
	Green	6
	blue	1
Range of current value to be changed [mA]		0 ~ 16

was changed for every 1 [mA] from 0 [mA] to 16 [mA]. For the green LED as well, considering fitness value of F_{hue} and F_{Bd} separately, the green light intensity of active 3D marker is considered to be suitable for recognition when the current value for green LED is from 9 [mA] to 11 [mA]. Figure 11 shows the fitness value when current of blue LED is changed by 1 [mA] from 0 [mA] to 16 [mA]. For the red and green LEDs as well, the green light intensity of active 3D marker is considered to be suitable for recognition when the current value for blue LED is from 2 [mA] to 4 [mA]. From the above results, the suitable current values for recognition were in the range (red : 7 [mA], green : 9∼11 [mA], blue : 2∼4 [mA]).

B. Optimal Combination of LED's Currents

1) Experimental Environment: The suitable current values established on previous subsection had range with current values of green LED and blue LED. Therefore, in order to determine the suitable current values uniquely from the estimated value and the distribution of fitness value, a recognition experiment was conducted again. Figure 12 shows the experimental layout. Experimental condition is shown in Table II. Illuminance was set the same condition as section III (A) and turbidity was set the close value to section III (A). The distance between the camera and the active 3D marker was set to be 600 [mm] in the x axis direction and 0 [mm] in the y axis and z axis direction. The current values (red : 7 [mA], green : 9∼11 [mA], blue : 2∼4 [mA]) determined in section III (A) were used. Camera images were acquired for each current value, and recognition accuracy was verified.

2) Experimental Results: Figure 13 (I) shows an example of the image obtained by recognition experiments. Figure 13 (II) shows an example of the distribution of fitness value in the x-y plane. Table III shows the fitness values of the peak position in fitness distribution for each current value. There was no significant difference in the peak position for each current value in table III. Therefore, the current values (red : 7 [mA], green : 11 [mA], blue : 2 [mA]) that give the highest fitness value (table III) were decided as the most suitable current value for the present recognition system.

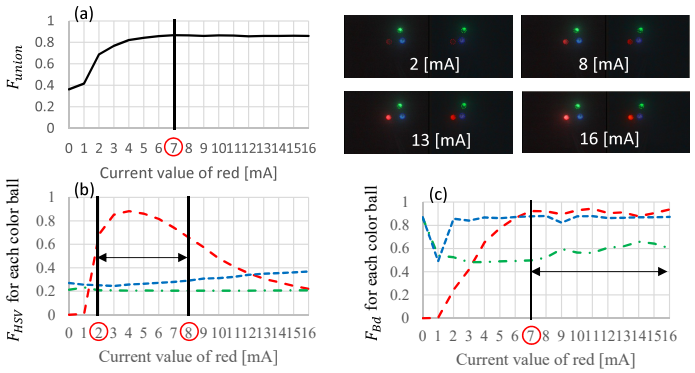

Fig. 9. The average of fitness value for 1 minute when changing the current value of red LED and captured images. (a) Final calculated fitness value (b) Fitness value calculated from hue information . (c) Fitness value calculate from brightness information

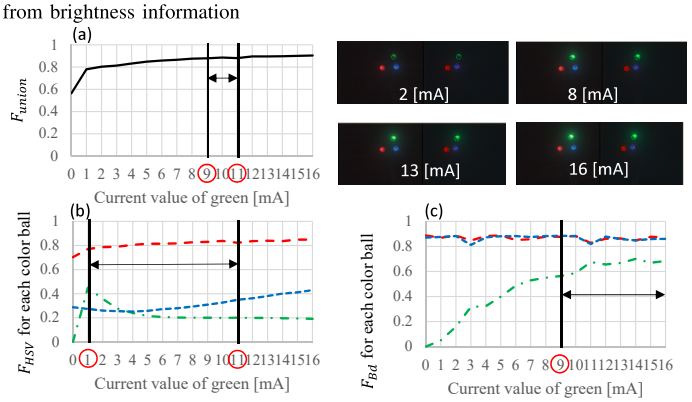

Fig. 10. The average of fitness value for 1 minute when changing the current value of green LED and captured images. (a) Final calculated fitness value (b) Fitness value calculated from hue information . (c) Fitness value calculate from brightness information

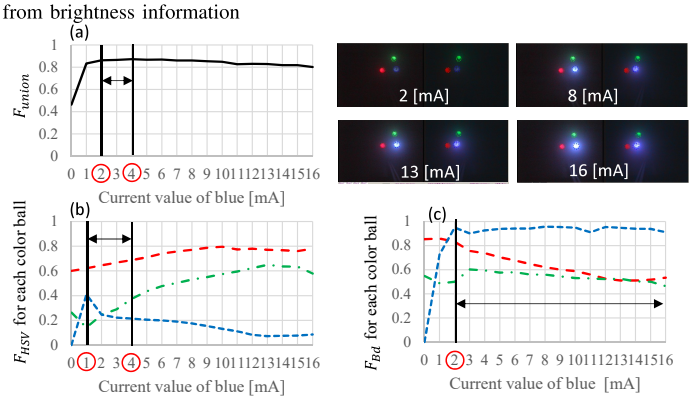

Fig. 11. The average of fitness value for 1 minute when changing the current value of blue LED and captured images. (a) Final calculated fitness value (b) Fitness value calculated from hue information . (c) Fitness value calculate from brightness information

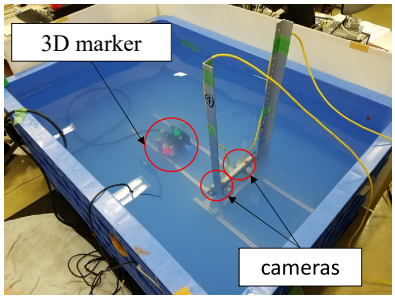

Fig. 12. Experiment environment

TABLE II
THE TABLE OF EXPERIMENTAL CONDITION

Turbidity [FTU]		8.9
Illumination [lx] (in the water)		0
Distance x [mm]		600
current value [mA]	Red	7
	Green	9~11
	blue	2~4

IV. Conclusion

In this paper, the suitable current values to recognition system proposed by our research group were decided from experimental data. In the previous research, robustness to the turbidity of our recognition system was discussed. Furthermore, using the current values determined in this paper (red : 7 [mA], green : 11 [mA], blue : 2 [mA]) will improve recognition accuracy and stability of recognition system.

V. Acknowledgment

This work is supported by JSPS KAKENHI Grant Number 16K06183 and KOWA CORPORATION for the development of ROV.

References

[1] Agbakwuru J. "Oil/Gas Pipeline Leak Inspection and Repair in Underwater Poor Visibility Conditions: Challenges and Perspectives," Journal of Environmental Protection, 2012.

[2] Krupi r ki, S., Allibert, G., Hua, M.D. and Hamel, T., "Pipeline tracking for fully-actuated autonomous underwater vehicle using visual servo control," In American Control Conference (ACC),pp.6196-6202. IEEE, 2012.

[3] Eustice, R.M., Pizarro, O. and Singh, H., Visually augmented navigation for autonomous underwater vehicles, IEEE Journal of Oceanic Engineering, 33(2), pp.103-122, 2008.

[4] Jung, J., Choi, S., Choi, H.T. and Myung, H., Localization of AUVs using depth information of underwater structures from a monocular camera, In Ubiquitous Robots and Ambient Intelligence (URAI), 2016 13th International Conference, pp. 444-446, IEEE, 2016.

[5] Ghosh, S., Ray, R., Vadali, SR., Shome, SN., Nandy, S, Reliable pose estimation of underwater dock using single camera: a scene invariant approach," Machine Vision and Applications, 27(2), pp. 221-36, 2016.

[6] Kume, A., Maki, T., Sakamaki, T. and Ura, T., A method for obtaining high-coverage 3D images of rough seafloor using AUV-realtime quality evaluation and path-planning, Journal of Robotics and Mechatronics, 25(2), pp.364-374, 2013.

[7] Ribas D, Palomeras N, Ridao P, Carreras M, Mallios A, Girona 500 auv: From survey to intervention, IEEE/ASME Transactions on Mechatronics 17(1), pp.46-53 2012

[8] Li X., Nishida Y., Myint M., Yonemori K., Mukada N., Lwin K. N., Matsuno T., Minami M.,"Dual-eyes Vision-based Docking Experiment of AUV for Sea Bottom Battery Recharging," MTS/IEEE OCEANS' 17, Aberdeen Scotland, June 19-22, 2017.

[9] Yamada D., Mukada N., Myint M., Lwin K. N., Matsuno T., Minami M., "Improvement of 3D-Pose Real-time Estimation by Active Marker and HSV-evaluated Function," 23rd International Symposium on Artificial Life and Robotics, B-Con Plaza Beppu, January 18-20, 2018

[10] Myint M., Yonemori K., Lwin K. N., Yanou A., Minami M., Dual-eyes Vision-based Docking System for Autonomous Underwater Vehicle: an Approach and Experiments, J Intell Robot Syst , DOI 10.1007/s10846-017-0703-6 2017

[11] Yu F, Minami M, Song W, Zhu J, Yanou A, On-line head pose estimation with binocular hand-eye robot based on evolutionary model-based matching, Journal of Computer and Information Technology, 2(1), pp.43-54, 2012.

[12] Song, W., Minami, M. and Aoyagi, S., Feedforward on-line pose evolutionary recognition based on quaternion. Journal of the Robot Society of Japan, 28(1), pp. 55-64, 2010.

[13] Yamada D., Mukada N., Myint M., Lwin K. N., Matsuno T., Minami M., "Docking Experiment in Dark Environments Using Active/Lighting Marker and HSV Correlation," MTS/IEEE OCEANS' 18, Kobe Japan, May 28-31, 2018.

[14] Song W, Fujia Y, Minami M, 3D visual servoing by feedforward evolutionary recognition, Journal of Advanced Mechanical Design, Systems, and Manufacturing, 4(4), pp.739-755, 2010.

[15] Suzuki H, Minami M, Visual servoing to catch fish using global/local GA search, IEEE/ASME Transactions on Mechatronics, 10(3), pp.352-357, 2005.

[16] Khin Nwe Lwin, Myo Myint, Naoki Mukada, Daiki Yamada, Takayuki Matsuno, Mamoru Minami, "Docking Performance Against Turbidity Using Active Marker Under Day and Night Environment," 23rd International Symposium on Artificial Life and Robotics, B-Con Plaza Beppu, January 18-20, 2018

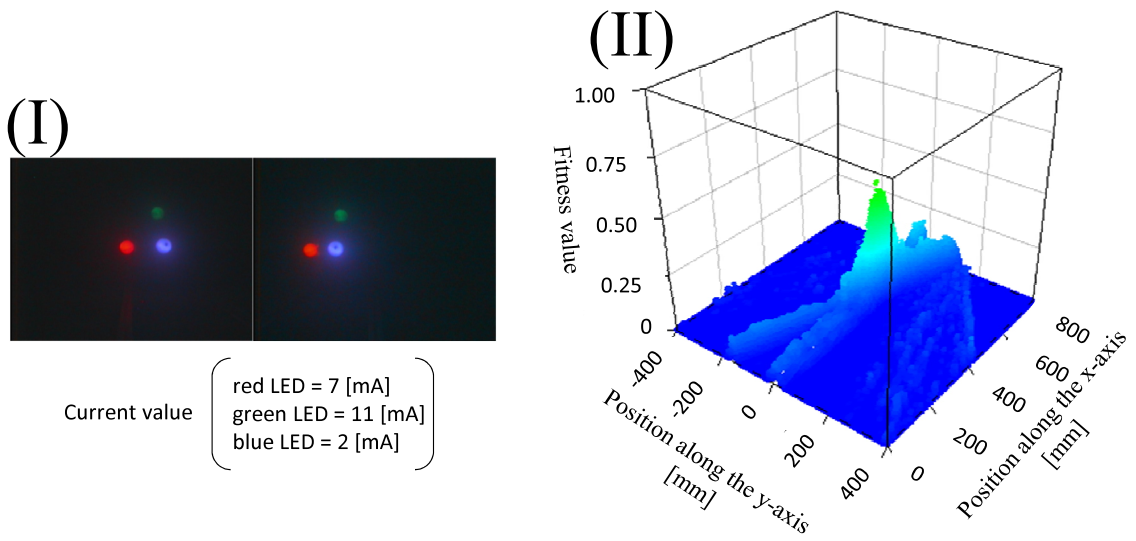

Fig. 13. Results obtained in the experiment. (I) an example of obtained image (II) an example of 3D graph of fitness distribution in the x-y plan

TABLE III
FITNESS VALUE AT PEAK ON 3D GRAPH AND PEAK POSITION FOR EACH CURRENT VALUE

	Current value [mA]			Estimation value		
	Red	Green	Blue	Fitness	x [mm]	y [mm]
(A)	7	9	2	0.5361	499	-4
(B)	7	9	3	0.5028	512	-4
(C)	7	9	4	0.5192	498	-2
(D)	7	10	2	0.5194	498	-2
(E)	7	10	3	0.4778	499	-5
(F)	7	10	4	0.5	523	-1
(G)	7	11	2	0.5472	499	-4
(H)	7	11	3	0.5306	497	-4
(I)	7	11	4	0.5139	499	-4