The Twenty-Fourth International Symposium on Artificial Life and Robotics 2019 (AROB 24th 2019),
The Fourth International Symposium on BioComplexity 2019 (ISBC 4th 2019),
B-Con Plaza, Beppu, Japan, January 23-25, 2019

Stereo-vision visibility analyses in relation with turbidity and distance for AUV recharge docking

Horng-Yi Hsu[1], Naoki Mukada[1], Daiki Yamada[1], Khin Nwe Lwin[2], Myo Myint[3], Yuichiro Toda[1] , Takayuki Matsuno[1], Keigo Watanabe[1], and Mamoru Minami[1]

[1]Okayama University, Japan
(Tel: 81-86-251-8233, Fax: 81-86-251-8233)
[1]pcvv4h2h@s.okayama-u.ac.jp
[3]University of Technology, Myanmar
[3]Thanlyin Technological University, Myanmar

Abstract: Aiming at developing underwater battery recharging system, the author developed a docking system using stereo-vision-based visual servoing and a 3D marker. The 3D marker consists of red, green, blue spheres that do not emit the light. Real-time relative pose (position and orientation) estimation was implemented utilizing the 3D model-based matching method and real-time multi-step genetic algorithm (RM-GA). Given the situation that the docking aims for battery recharging in the deep-sea bottom, the pitch-dark and turbid environment should be considered as an inevitable condition for battery recharging. In our previous works, the docking experiments were conducted in the actual sea, having verified the effectiveness of the proposed system using the 3D marker in the daytime environment with turbid water condition. Since lighting 3D marker by light from the vehicle in turbid water environment results in a situation that the images taken by video cameras set on the vehicle were looked wholly white, some new idea seems to be required. To overcome this problem, first step is to conform the correctness of the proposed system. The main objective of this study is to check the feasibility area of the proposed system for the docking application in the simulated pool with different turbidity and distance.

Keywords: Stereo-vision, Visual servoing, Turbidity, Recognition, 3D marker

1 INTRODUCTION

In recent years, various underwater robots have been researched and developed for seabed exploration, submarine mapping, an ecological survey of underwater organisms, mining of underground resources and so on. To do this kind of works, the underwater robots need to be active for a long time in the seabed. There are many studies on underwater docking using different approaches. Depending on a docking station's structure for a specific application, different methods and sensors were utilized. Homing accuracy and robustness against disturbances are critical requirements for docking operation. To fulfill these demands, many studies have been conducted recently.

Among them, it is roughly divided into a method using monocular camera [1] and stereo camera [2] for homing and docking. Docking was conducted for the submersible power supply with an ultrasonic sensor and monocular camera with the control error of $\pm$100 mm [3].

In contrast, a vision-based docking system using two cameras and a 3D marker has been developed by our research group. In previous works, we especially confirmed the performance of 3D pose estimation and visual servoing, which are one of the docking steps in our approach using real-time pose estimation. In [4], real-time pose tracking ability with stereo-vision was confirmed when a target was moving even though there were some noise in captured images due to air bubbles in front of the cameras.

In that paper not only noise disturbances in images but also physical disturbances of water stream induced by floating motion of air bubbles were given to the ROV, having confirmed whether the proposed approach is robust enough to be able to operate in the actual sea. In [5], 3D pose estimation with partial occlusion was discussed. In [6], docking procedure was implemented and docking experiment was conducted when the ROV's starting positions were given arbitrarily in front of the 3D marker. As a follow-up work, we also checked the robustness of our system under a varying light environment in [7]. In [8], the sea docking experiment using a circular shaped docking hole was reported. Since this experiment was conducted in a relatively transparent sea in a part of the harbor, measurement of 3D marker for estimating the pose of ROV was successful.

In such a wide range of actual seabed, the vision-based vehicles have to face the difficulty for the recognition because the proliferation of mud will cause poor visibility by the thrust of the underwater robot and other disturbances such fish, seaweed, turbidity, lighting changing, and so on. The vehicle cannot avoid the turbidity and dark environment for the sea bottom battery recharging. According to the authors' knowledge, there are no studies using stereo-vision and 3D

marker for real-time visual servoing with the performance of turbidity tolerance and illumination varieties. In our previous works, docking experiments were conducted against turbidity and changing lighting environment. In [9], the performance of the 3D pose estimation system under different turbidity levels was analyzed by using 3D marker, and the turbidity tolerance of the system was examined experimentally.

In the case of using the 3D marker, the ROV's lighting was used to detect the 3D marker in the dark environment. Since lighting 3D marker by light from vehicles in turbid water environment results in the images being looked wholly white. To overcome this problem, the first step is to conform the correctness of the proposed system. The recognition experiments were conducted by simulating a turbid environment in the experimental pool using the 3D marker. The turbidity levels and the distance between the 3D marker and ROV were varied as variables, and the recognizable range of the 3D marker was confirmed by Real-time Multi-step genetic algorithm(RM-GA).

2 FITNESS FUNCTION

In this experiment, fitness value is used to evaluate the performance of the recognition under different turbidity levels and distance. A correlation function of the real target projected in camera images with the assumed model, represented by poses in the chromosomes, is used as the fitness function in the GA process. Fitness function has been modified based on the voting performance and the target's structure (color, size, and shape). The function used in the proposed system is called a fitness function that is shown by equation (1). ϕ means position and posture of 3D marker that relative to the fixed coordinate system of the ROV. The evaluation function that is used for ϕ and the image taken by cameras is defined by the previous research [10]. ϕ_i means pose of the i-th model that was given by the RM-GA.

$$F(\boldsymbol{\phi}_i) = \frac{1}{2}(F_L + F_R) \tag{1}$$

$$F_L = \frac{1}{N}\left(\sum_{{}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\in S_{L,in}(\boldsymbol{\phi}_i)} p\left({}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\right) + \sum_{{}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\in S_{L,out}(\boldsymbol{\phi}_i)} p\left({}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\right)\right) \tag{2}$$

$$F_R = \frac{1}{N}\left(\sum_{{}^{IR}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\in S_{R,in}(\boldsymbol{\phi}_i)} p\left({}^{IR}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\right) + \sum_{{}^{IR}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\in S_{R,out}(\boldsymbol{\phi}_i)} p\left({}^{IR}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\right)\right) \tag{3}$$

The left or right image of the solid 3D model to estimate the pose of the real 3D marker is shown in Fig. 1. This solid 3D model is made from the point cloud that made from the color and shape of 3D marker. The fitness calculation of equation (1)-(3) is described as below. In Fig. 4, fitness value is calculated by evaluating the inner sphere S_{in} overlap the hue value of the image. The function $p\left({}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)\right)$ is a function that gives a score "1" when the j-th model ${}^{IL}\boldsymbol{r}_j(\boldsymbol{\phi}_i)$ of the point group of the model determined by the i-th candidate $\boldsymbol{\phi}_i$ of the pose overlaps with the 3D marker. Similarly, the evaluation of the fitness value is calculated for the enveloping sphere S_{out} of the search model. Subtract the total value of S_{out} from the total value of the evaluation of inner sphere S_{in} and divide by the number of point N groups of the search model to obtain the total fitness value. The calculation of fitness value is averaged over the left and right camera images to obtain the final fitness value $F(\boldsymbol{\phi}_i)$ consists of pose information[x [mm], y [mm], z [mm], ε_1, ε_2, ε_3] (ε_1, ε_2, ε_3 are quaternions).

The 3D model was created by using the information of $\boldsymbol{\phi}_i$ that represented by genes is projected to the left and right camera images, and the correlation with the 3D marker is evaluated as the fitness value. The entered left and right camera images are used for evaluation without performing image processing. The search model matches the 3D marker when the relationship between various parameters of the camera and kinematics is perfectly matched and the search model accurately represents the shape of the recognition object. At this time, $F(\boldsymbol{\phi}_i)$ is configured to take the maximum value, and the pose $\boldsymbol{\phi}_i$ of the search model giving the maximum value represents the pose of the 3D marker.

3 COMPARISON OF RECOGNITION PERFORMANCE BY USING 3D MARKER

3.1 Experiment Environment

Figure 2 shows the experiment environment using 3D marker and coordinate system of the ROV and 3D marker. $\sum_H$ and $\sum_M$ are the coordinate system of the ROV and 3D marker. The position of ROV was fixed in the pool so that the pose between the ROV and the 3D marker is kept constant. The turbidity environment is created by putting milk in

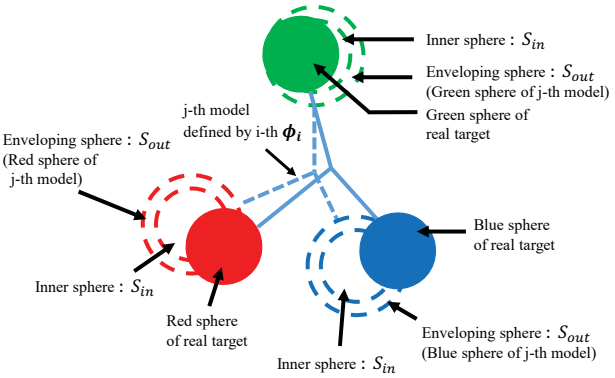

Fig. 1. Left or right camera's 2D image of the real 3D marker and model.

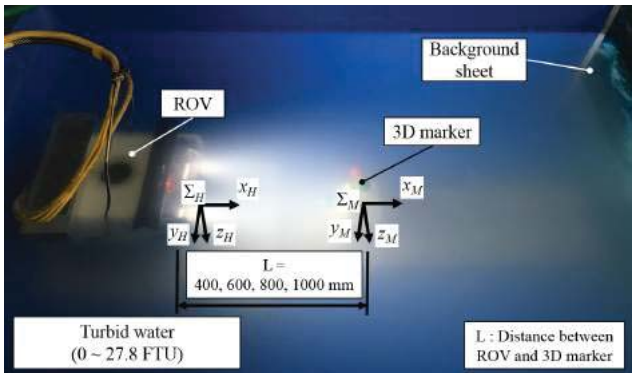

Fig. 2. Experimental environment using marker and coodinate systems of ROV and 3D marker. Photograph of ROV and marker in dark environment.

the pool. According to the other researchs [11, 12], there are 10 to 600 nm particles in milk. When light run through the size of 10 nm particles, light will be scattered equally in front and rear direction. When light run through the size of 100 nm particles, light will be scattered to forward. "Forward" means the traveling direction of light. In other words, the experiment using milk is possible to create the turbidity environment that considers the various light scattering. In [11], the maximum amount of milk 1.9×10^2 ml/m^3 have been added. In [12], the maximum amount of milk 1.5×10^2 ml/m^3 have been added. Those researches described the monocular image recognition. There is no discussion on pose estimation using the stereo-vision. In this study, the pose recognizable range using stereo-vision on different turbidity and using 3D marker was analyzed. Recognition experiment condition of using 3D marker will be described. The turbidity is measured by the portable turbidity sensor TD-M 500 manufactured by OPTEX Corporation. This turbidity sensor can measure turbidity in the range of 0.0 to 500 FTU (Formazin Turbidity Unit). The conditions of the recognition experiment are shown at below.

Light Environment 3D marker and ROV's LED lighting device

Turbidity Environment using the 3D marker : maximum 1.19×10^2 ml/m^3 milk have been added

- One times increase milk by 2.435 ml/m^3(2 g) between 0 and 3.64×10 ml/m^3
- One times increase milk by 4.870 ml/m^3(4 g) between 3.64×10 and 1.19×10^2 ml/m^3

Recognition Distance 400, 600, 800, 1000 mm

3.2 Recognition Experiments

In the recognition experiment using 3D marker, turbidity levels were divided into 33 parts from 0 to 1.19×10^2 ml/m^3 and the distance between the ROV and 3D marker divided into 4 parts from 400 to 1000 mm. The total of 132 (= 33×4) recognition experiments were conducted. Preliminary experiments[13] have already confirmed that the ratio of the amount of milk that was put into the pool (ml/m^3) and turbidity measurement value (FTU) is 0.25 FTU/ml/m^3. When the distance between the ROV and the 3D marker is 400 mm, the turbidity is taken as the maximum turbidity that 3D marker cannot be perceived completely.

The results of recognition experiments using 3D markers are shown in Table 1. The amount of milk that was put into the pool is shown in the first column. The second to fifth columns show the average of the fitness value on each turbidity level when the distances from the ROV to 3D marker is 400, 600, 800 and 1000 mm. All the fitness value in the table is the average of the fitness value obtained from the recognition experiment of 60 seconds. In Fig. 3, real-time and average fitness value under the conditions designated by Ⓐ in Table 1 is calculated by averaging above real-time measurement result (a) in Fig. 3, and Ⓑ and Ⓒ in Table 1 are calculated by (b) and (c) in Fig. 3.

Area I enclosed by the solid line at the upper left in Table 1 is the dockable range. If the average of fitness value is 0.6 or more from the experiment, it is confirmed that the ROV can perform the visual servo. Area II enclosed by the dotted line in Table 1 is an area in which the average of fitness value fall within the range of 0.22 to 0.6. It is dockable but sometimes fail. Area III is the rest area in Table 1. If the average of fitness value is lower than 0.22, it means the duel eye recognition system cannot recognize the 3D marker. The 3D search model does not converge on the 3D marker by RM-GA. Left and right camera images of the ROV in the areas I to III are shown in Fig. 4. The dotted line in the image shows the pose of the 3D search model by RM-GA. In the images of area I and II, it can be confirmed that the 3D search model converges to the 3D marker. But, the 3D search model cannot converge to the 3D marker in the Area III. It can be confirmed that the result of the pose measure-

ment of RM-GA is not correct. Figure 5 is the left and right camera images with the pose recognized by the pose estimation system at different turbidity levels and the distance of 600 mm between the ROV and 3D marker. The images are shown in three columns and the image described as 0 ml/m^3 at the top left was taken under the conditions indicated on the label A in Table 1. And, the images shown from the top to the bottom of the left column, from the top to the bottom of the center, and from the top to the bottom of the right column were taken under the conditions given in the order of the distance 600 mm column from the top to the bottom in Table 1.

Figure 6 shows the fitness value distributions confirming the robustness of the system at the distance of 600 mm. The fitness value shown at the top of each fitness distribution indicates the maximum and peak fitness value. Through analyzing the fitness value distributions using the 3D marker, it can be confirmed that the peak of fitness value is lowered as the turbidity increases. The position of the peak corresponding to the true pose of the 3D marker was maintained even though the height of the peak was reduced by increasing turbidity. The gradual reduction in the height of the peak shows the effect of turbidity on image recognition. The position given by the peak of fitness value has not changed. It means that the problems of solving the optimization problem and estimating the pose are not affected by the turbidity. This is the characteristics obtained by converting to the optimization problem.

Table 1. The fitness value, e.g. 0.730 with 0 ml/m^3 and 600 mm distance, represents average fitness value shown by Fig. 3(a). The average fitness values in this table are measured with different turbidity levels and distances between the ROV and 3D marker.

Milk ml/m^3(FTU) \ Distance mm	400	600	800	1000	
0(0)	0.687(0.048)	Ⓐ 0.735(0.025)	0.781(0.045)	0.751(0.032)	Area I
2.43(0)	0.881(0.028)	0.892(0.022)	0.761(0.023)	0.751(0.031)	
4.85(0)	0.776(0.037)	0.838(0.028)	0.724(0.026)	0.592(0.031)	
7.28(0)	0.924(0.028)	0.823(0.031)	0.663(0.026)	0.559(0.035)	
9.70(0)	0.881(0.029)	0.749(0.028)	0.576(0.033)	0.422(0.041)	
$1.21\times10(0)$	0.771(0.034)	0.746(0.029)	0.566(0.048)	0.491(0.051)	Area II
$1.46\times10(0)$	0.749(0.036)	0.701(0.029)	0.674(0.038)	0.459(0.047)	
$1.70\times10(0)$	0.712(0.028)	0.655(0.031)	0.530(0.033)	0.303(0.055)	
$1.94\times10(3.03)$	0.702(0.029)	0.646(0.031)	0.538(0.039)	0.284(0.059)	
$2.18\times10(3.75)$	0.674(0.025)	0.751(0.031)	0.329(0.034)	0.063(0.009)	
$2.43\times10(4)$	0.691(0.027)	Ⓑ 0.706(0.034)	0.397(0.044)	0.058(0.007)	
$2.67\times10(4.5)$	0.679(0.029)	0.712(0.032)	0.325(0.033)	0.134(0.020)	
$2.91\times10(6.6)$	0.685(0.030)	0.655(0.036)	0.381(0.050)	0.060(0.006)	
$3.15\times10(7.1)$	0.670(0.030)	0.646(0.034)	0.298(0.034)	0.052(0.013)	
$3.40\times10(7.5)$	0.671(0.031)	0.653(0.036)	0.215(0.035)	0.055(0.016)	
$3.64\times10(7.6)$	0.651(0.031)	0.589(0.037)	0.258(0.038)	0.058(0.012)	Area III
$4.12\times10(7.9)$	0.642(0.028)	0.595(0.038)	0.181(0.046)	0.185(0.046)	
$4.61\times10(8.7)$	0.613(0.029)	0.584(0.038)	0.161(0.044)	0.059(0.012)	
$5.09\times10(9.3)$	0.622(0.029)	0.529(0.044)	0.158(0.045)	0.056(0.015)	
$5.58\times10(10.5)$	0.583(0.058)	0.294(0.039)	0.124(0.046)	0.136(0.056)	
$6.06\times10(11.2)$	0.581(0.027)	0.216(0.039)	0.098(0.036)	0.153(0.038)	
$6.55\times10(12.2)$	0.547(0.028)	Ⓒ 0.262(0.056)	0.131(0.018)	0.119(0.039)	
$7.03\times10(13.3)$	0.568(0.032)	0.087(0.016)	0.139(0.024)	0.108(0.034)	
$7.52\times10(14.3)$	0.448(0.026)	0.217(0.027)	0.053(0.020)	0.060(0.010)	
$8.00\times10(15.3)$	0.551(0.033)	0.152(0.025)	0.059(0.011)	0.155(0.032)	
$8.49\times10(17.1)$	0.373(0.055)	0.117(0.015)	0.061(0.012)	0.167(0.034)	
$8.97\times10(18.3)$	0.489(0.041)	0.146(0.032)	0.150(0.020)	0.149(0.036)	
$9.46\times10(20.4)$	0.447(0.035)	0.118(0.049)	0.189(0.026)	0.226(0.028)	
$9.94\times10(21.4)$	0.426(0.037)	0.156(0.023)	0.150(0.024)	0.148(0.023)	
$1.04\times10^2(23)$	0.400(0.047)	0.149(0.027)	0.170(0.033)	0.061(0.013)	
$1.09\times10^2(24.2)$	0.100(0.010)	0.126(0.047)	0.130(0.024)	0.143(0.057)	
$1.14\times10^2(26.4)$	0.089(0.017)	0.161(0.033)	0.168(0.031)	0.172(0.048)	
$1.19\times10^2(27.8)$	0.174(0.022)	0.093(0.025)	0.212(0.023)	0.218(0.026)	

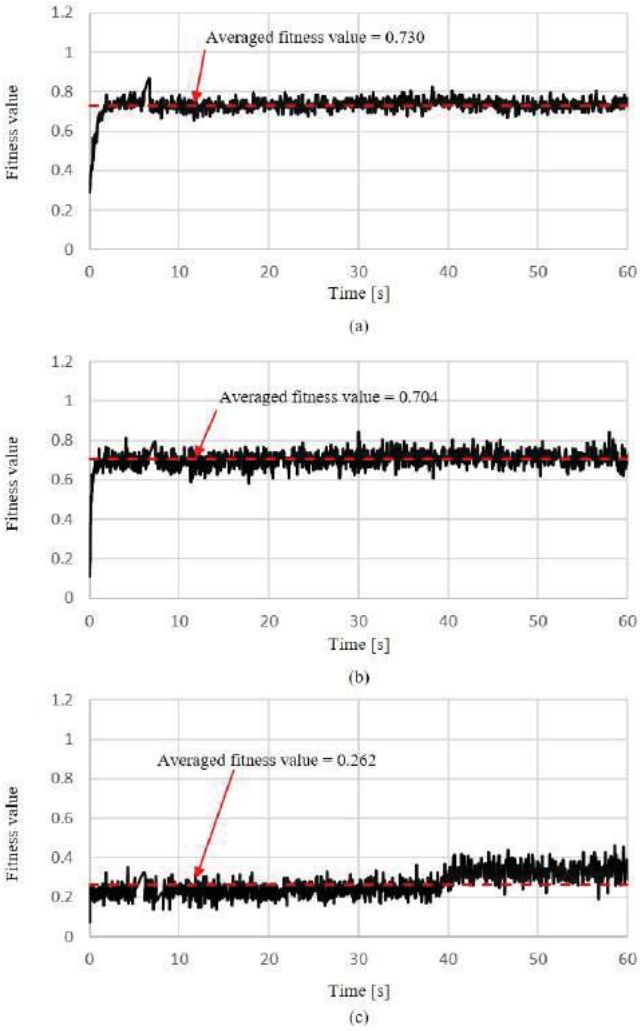

Fig. 3. Real-time and average fitness value 0.730 under the conditions designated by Ⓐ in Table 1 is calculated by averaging above real-time measurement result (a), and Ⓑ and Ⓒ in Table 1 are calculated by above (b) and (c).

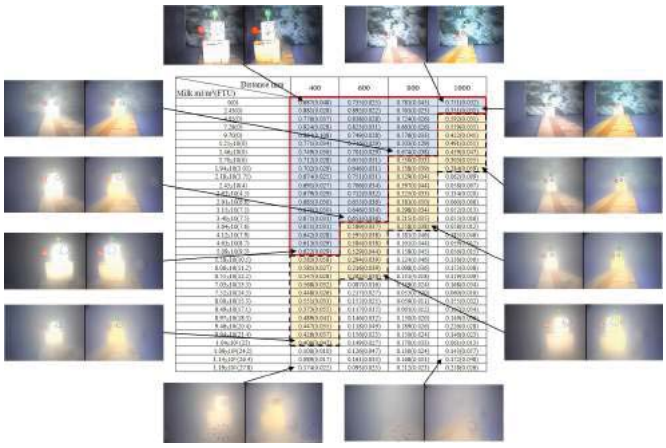

Fig. 4. Left and right camera images are taken under the different turbidity and distances, which are indicated by the arrows. Images taken at the maximum and minimum distances in clean water and the maximum turbidity, in which the 3D marker cannot be observed.

Fig. 5. Left and right camera images with the pose recognized by the pose estimation system at different turbidity levels and a distance of 600 mm between the ROV and 3D marker. The recognized pose is indicated by dotted circles in each photograph. The amount of added milk is given in units of milliliters per cubic meter.

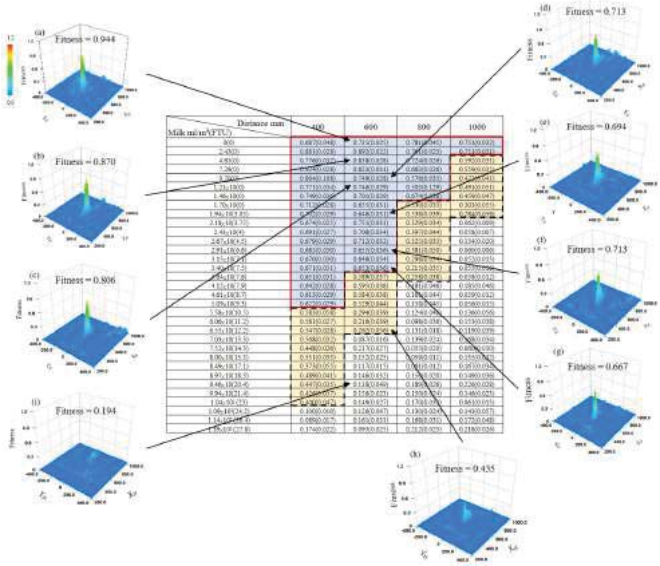

Fig. 6. Fitness value distributions confirming the robustness of the system at a distance of 600 mm. The position of the peak corresponding to the true pose of the 3D marker was maintained even though the height of the peak was reduced by increasing turbidity. The gradual reduction in the height of peak shows the effect of turbidity on image recognition. The fitness value shown at the top of each fitness distribution indicates the maximum and peak fitness value.

4 CONCLUSION

In this study, to confirm the effectiveness of the proposed system, milk was put into the pool to create turbid environments. According to the experimental results, it was confirmed the feasibility area of the proposed system for the docking application in the simulated pool with different turbidity and distance.

REFERENCES

[1] Park, J-Y., Jun, B-H., Lee, P-M. and Oh, J., Experiments on vision guided docking of an autonomous underwater vehicle using one camera, Ocean Engineering, Vol. 36, No. 1 (2009), pp.48–61.

[2] Ura, T., Kurimoto, Y., Kondo, H., Nose, Y., Sakamaki, T. and Kuroda, Y., Observation behavior of an AUV for ship wreck investigation, Proceedings of the OCEANS 2005 MTS/IEEE, Vol. 3 (2005), pp.2686–2691.

[3] Ishii, K., Sonoda, T., Nakanishi, R., Kawashima, S. and Hidaka, S., Research on docking control of autonomous underwater vehicle, ROBOMECH2015 in Kyoto (2015), 2A2-D06 (in Japanese).

[4] Myo Myint, Kenta YONEMORI, Akira YANOU, Shintaro ISHIYAMA and Mamoru MINAMI. Robustness of Visual-Servo against Air Bubble Disturbance of

Underwater Vehicle System Using Three-Dimensional Marker and Dual-Eye Cameras. Proceedings of the International Conference OCEANS15 MTS/IEEE. Washington DC, USA. pp.1-8. 2015.

[5] Myo Myint, Kenta Yonemori, Akira Yanou, Khin Nwe Lwin, Mamoru Minami and Shintaro Ishiyama. Visual Servoing for Underwater Vehicle Using Dual-eyes Evolutionary Real-time Pose Tracking. Journal of Robotics and Mechatronics. vol.28, no.4, pp.543-558. 2016.

[6] Myo Myint, Kenta Yonemori, Akira Yanou, Mamoru Minami and Shintaro Ishiyama. Visual-servobased Autonomous Docking System for Underwater Vehicle Using Dual-eyes Camera 3D-Pose Tracking. Proceedings of the 2015 IEEE/SICE International Symposium on System Integration. Nagoya, Japan. pp.989-994. 2015.

[7] Myo Myint, Kenta YONEMORI, Akira YANOU, Khin Nwe Lwin, Mamoru MINAMI and Shintaro ISHIYAMA. Visual-based Deep Sea Docking Simulation of Underwater Vehicle Using Dual-eyes Cameras with Lighting Adaptation. Proceedings of the International Conference OCEANS16 MTS/IEEE. Shanghai, China. pp.1-8. 2016.

[8] Kenta YONEMORI, Akira YANOU, Myo Myint, Khin Nwe Lwin, and Mamoru MINAMI. Docking experiment of underwater vehicle by dual-eye visual servoing in sea. Transactions of the JSME (in Japanese), Vol.83, No.848, 2017.

[9] Myo Myint, Khin Nwe Lwin, Naoki Mukada, Daiki Yamada, Takayuki Matsuno, Yuuichirou Yoda, Saitou Kazuhiro, Mamoru Minami, Experimental Verification of Turbidity Tolerance of Stereo-vision-based 3D Pose Estimation System, Journal of Marine Science and Technology, DOI 10.1007/s00773-018-0586-7, 2018 (24 pages).

[10] Myo, M., Kenta Y., Khin N. L., Akira Y. and Mamoru M., Dual-eyes Vision-based Docking System for Autonomous Underwater Vehicle: an Approach and Experiments, Journal of Intelligent and Robotic Systems (2017b), DOI:10.1007/s10846-017-0703-6.

[11] Garcia, R. and Gracias, N., Detection of interest points in turbid underwater images, In OCEANS, 2011 IEEE-Spain (2011), pp.1–9.

[12] Codevilla, F., Gaya, J.D.O., Duarte, N. and Botelho, S., Achieving turbidity robustness on underwater images local feature detection, International journal of computer vision (2004), 60 (2), pp.91–110.

[13] Myo M., Khin N. L., Naoki M., Matsuno T. and Mamoru M., Stereo Vision-based 3D Pose Estimation under Turbid Water for Underwater Vehicles, RSJ 2017 (in Japanese), (2017a), C1C3-02, 2017.9.11–14.