

# Visibility improvement in relation with turbidity and distance and application to docking

Horng-Yi Hsu · Yuichiro Toda · Keigo Watanabe · Mamoru Minami

Received: date / Accepted: date

**Abstract** Recently, Autonomous Underwater Vehicles (AUVs) are being used in many applications. There is a limitation for underwater vehicle's operation that takes a longer duration than the power capacity of underwater vehicles. Since underwater battery recharging station are supposed to be installed in the deep-sea bottom, the deep-sea docking experiments cannot avoid turbidity and dark environment. In this study, we propose a newly designed active 3D marker to improve the visibility of the 3D marker. A docking experiment apparatus was built-up. Two kinds of the 3D marker were used in the experiment, which was the active (light) and passive (non-light) 3D marker. The experimental results show the active 3D marker can be more recognizable in turbidity and dark environment than the passive 3D marker.

**Keywords** Stereo-vision · Visual servoing · Turbidity · 3D marker

## 1 Introduction

Japan has a huge sea area from which resources can be taken out using advanced technologies. The Autonomous Underwater Vehicle (AUV) plays an important role in the deep-sea such as oil pipe inspection, a survey of the seafloor and searching rare metal [1]. Moreover,

the Japanese government is now seriously considering searching for methane hydrate as a future energy solution. To do such novel works that need long duration time in the deep sea, one of the main limitations of AUVs is power capacity. Even though advanced technology related to power provides long operation periods, underwater vehicles have to come back to a surface vessel for recharging. For solving this problem, underwater battery recharging with a docking function is one of the solutions to extend the operation time of AUVs. But, a challenging issue occurs in actual sea environments such as turbidity made by mad blown up by the AUV, which hinders their application to recharging. To overcome the problem, some idea should be added to the current docking system.

Nowadays, a number of studies using different sensors have been conducted on underwater vehicle operation [2–6]. Normally, long navigation is performed by using sonar sensors, which is not suitable for docking operation because of its low resolution. Vision-based navigation is one of the possible approaches, especially if high accuracy positioning is required.

Studies related to vision-based navigation for the underwater vehicle are based on a single camera to estimate the position and orientation of the target object [7,8]. The reliable pose estimation for underwater docking through a scene invariant approach by using a single camera was proposed in [9]. However, this methodology has not been applied to docking, and it is thought to be difficult for real-time pose estimation. In [10], a torpedo-shaped autonomous underwater vehicle was optically guided by the lights that mounted around the entrance of a docking station using single camera. In this approach, because of the complicated estimation of the geometry calculation, it is difficult to estimate the information about the orientation. Even

---

Horng-Yi Hsu  
Okayama University, 1-1-1 Tsushimanaka, Kita Ward,  
Okayama, 700-8530, Japan  
Tel.: +81-86-251-8233  
Fax: +81-86-251-8233  
E-mail: pcvv4h2h@s.okayama-u.ac.jp

Yuichiro Toda, Keigo Watanabe and Mamoru Minami  
Okayama University, 1-1-1 Tsushimanaka, Kita Ward,  
Okayama, 700-8530, Japan

though the attitude keeping control is used for the final docking step, the accuracy of the orientation information is still not high. An approach to docking using an electromagnetic guidance system was proposed in [11]. But, the presence of magnetic abnormalities near a station limits the accuracy of this approach. The merit of the monocular camera is that the configuration is simple, and processing time seems to be less than the multi-cameras. The disadvantage is that the precision of distance measurement of the camera's depth direction is not enough.

For overcoming the limitations of a single camera system [12], there are some studies developed by two cameras systems [13, 14]. In [13], two cameras have been mounted on the vehicle. One was looking forward to estimating the panel pose and the other was pointing down to help the mounted manipulator conduct some tasks. In [14], the same with [13], two cameras were used, one was looking forward, and the other was facing downward for the purpose of obstacle avoidance and docking. In two of the studies mentioned above [13] and [14], the binocular vision was used to detect the position of the vehicle and dock with a station. Even though two cameras were set, "stereo-vision" was not used in the system. Although [15] uses stereo-vision with structured triangular point marker, the visibility tends to be vulnerable to floating obstacles and turbidity.

The underwater environment is complex, and there are many disturbances for vision-based underwater vehicles. Therefore, it is important to make a robust vision system against possible disturbances. The common disturbances for the vision-based underwater vehicle are a lighting condition and turbidity. Since underwater battery recharging units are supposed to be installed in the deep-sea bottom, the deep-sea docking cannot avoid the turbidity and low-light environment. To the best of the author's knowledge, no existing study has conducted the docking using stereo-vision-based real-time visual servoing with performance tolerance of turbidity and illumination varieties.

Our research group has conducted a number of studies on our stereo-vision based docking system in [16–19]. In our approach, the relative pose between the underwater vehicle and a known 3D marker is estimated by using the Real-time Multi-step Genetic Algorithm (RMGA) that is a real-time 3D pose estimation method. The docking experiments were conducted using two cameras and a known 3D marker set at docking station to confirm the robustness of the docking system, having verified the effectiveness of the proposed system in the daytime in an environment with less turbid water [16]. The robustness of the 3D pose (position and orientation) estimation system against air bubbles [17]

and target occlusion [18] has been verified experimentally, respectively. In [17], the stability of an underwater vehicle system using a 3D marker in the presence of air bubbles was investigated in an indoor pool. Experiments were conducted to verify the robustness of the proposed system against physical disturbances in different situations in [18]. In [19], a deep-sea docking simulation experiment was conducted under a variable lighting environment. From the result, the lighting direction from the underwater vehicle affects the pose estimation.

In the turbid water environment and nighttime condition, the conventional idea is that the vehicle's LED illuminates the target. In this situation, the input image become white since the turbid particles in the water reflect the lighting from the vehicle. The lighting from the vehicle has been confirmed not to be an effective method to detect something in turbid and dark conditions. On the other hand, point light marker with no lighting from the vehicle has been sometimes used for pose estimation [20]. Predefined marker is geometrical information enables single camera to detect pose values. However, stereo-vision allows to exploit parallax displacement, which contributes to accurate position detection in camera depth direction. Despite that stereo vision research concept that reconstructs 3D information from 2-D stereo-vision images has been studied so far by many researchers, no practically useful results have not been reported for the underwater vehicles. This seems to be derived from the vulnerability that the point matching on geometric feature points in stereo-vision like spire of marker may be disturbed by small things in real sea, resulting in erroneous 3D reconstruction.

Since the 3D marker with being lighted 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. For overcoming this difficulty, the proposed system with a newly designed active 3D marker has been developed to improve the visibility of the marker from an underwater vehicle, especially in turbid water [22]. The lighting marker [20] that has been proposed uses LED point marker, whereas our 3D marker composed of three spheres has larger size of diameter. This leads to a merit that near positioned sphere looks bigger than the other sphere positioned far. The difference of size of spheres can help increase the measurement accuracy concerning camera depth direction.

Let us clarify the differences between the previous research reported by authors and this paper. The remotely operated vehicle (ROV) system utilizing the stereo vision and 3D marker for docking operation has

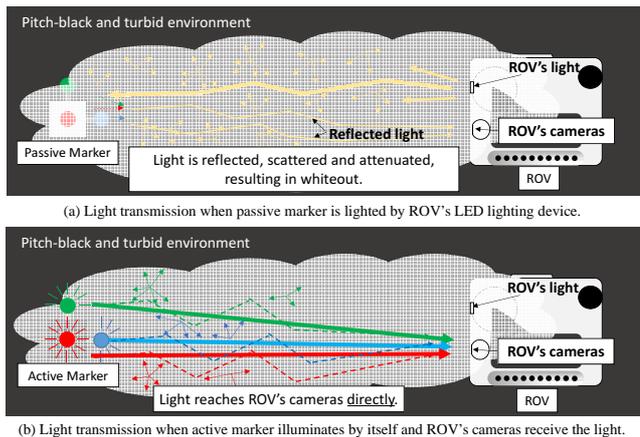
been used in [16]-[18], but those 3D markers are all without lighting abilities and the experiments are under daytime condition. [19] has discussed adaptive image recognition method and the effectiveness in changing ROV's lighting condition caused by the lighting direction changes, where the lightning variation is derived from the movement of the ROV. Based on the above researches, the authors have turned the directions of how to deal with pitch-dark environment of deep sea and the turbid water, that is a proposal of lighting 3D marker in [22], [25]. However, those reports [22], [25] have only discussed the docking experiments in pool with lighting condition varieties, which mean preparations for docking operations in real oceans.

What subjects have been left as future works are evaluation of practical usefulness of our original docking strategy, including actual recharging of batteries installed at underwater vehicles. Since the previous researches cling to a single camera strategy [1], [4]-[10], [13]-[15], insufficiency of space recognition accuracy and space estimation has resulted in the fact that there exists no practically reliable docking system in real sea.

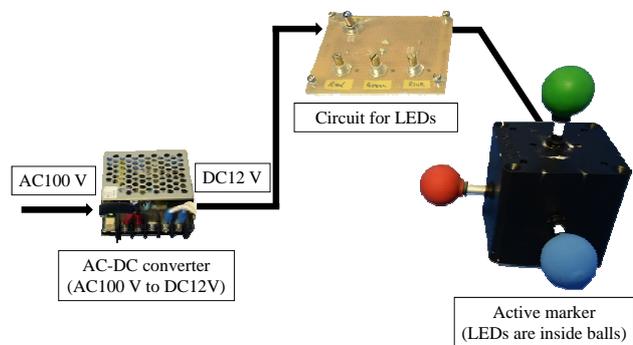
This paper, based on above preparations, discusses newly the tolerance against turbid water. The reason why the authors have conducted experiment to compare the effectiveness between passive/active markers, is the intention of how much the active marker outperforms the conventional passive marker on the viewpoint of "numerical evaluation." Especially, how far the markers could be visible in numerically classified turbidity on condition of stereo-vision strategy, which seems to indicate usability of the marker for docking in deep sea bottom with turbidity. The effectiveness of the proposed system has been proved by docking-experiment at the real sea.

## 2 Active 3D marker

Figure 1 shows the illustration of light transmission using a passive and an active 3D marker in pitch-dark and the turbid environment. Figure 1(a) shows the light transmission of the passive marker from the ROV's LED light. Figure 1(b) shows the light transmission of the active marker illuminates, and ROV's cameras receive the light. In the case of the passive 3D marker, the camera image is affected by a diffused reflection because of the particles of seabed mud. As a result, the captured camera image is more likely to be whiteout. On the other hand, some light beams from the active 3D marker reach the cameras directly because of the diffused light. In addition, the distance of the light beam in the active 3D marker is half of the passive 3D marker. From these points of view, it is considered that the



**Fig. 1** Illustration of light transmission using a passive and an active marker in a pitch-dark and turbid environment



**Fig. 2** Structure of active 3D marker

camera image is clearer than that of the passive 3D marker. Therefore, we developed the active 3D marker composed of red, green, and blue in each sphere. Figure 2 shows the appearance of the active 3D marker. The 3D marker is constructed with a water proof box (100mm × 100mm × 100mm) and the spheres (diameter; 40mm) are attached to the water proof box. The red, green and blue LEDs were installed into the spherical ball that is covered by the color balloon.

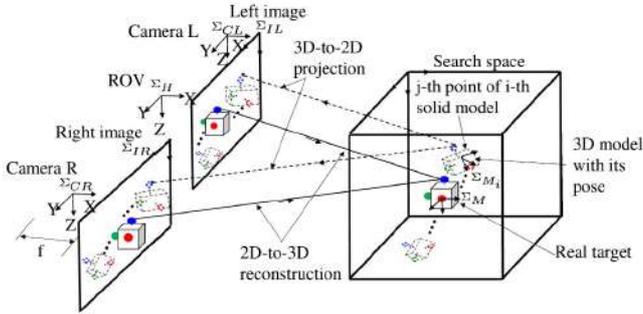
## 3 Proposed stereo-vision-based control system

### 3.1 3D model-based matching method

In generally approaches, object recognition is implemented by the feature-based recognition using 2D-to-3D reconstruction calculations, in which the information of the target object is determined from a set of points in different images, generally. If a point in one image is incorrectly mapped to a point in another image, the pose of the reconstructed object does not rep-

resent that of the real 3D object. A pose estimation approach based on 3D-to-2D model projection was applied in this study because the forward projection from 3D-to-2D generates unique points in 2D images without any errors [16–19],[21,22].

With the 3D-to-2D approach, a model-based matching method was used to recognize the 3D marker and estimate its pose in real-time. A model-based matching method is based on the known shape, color, and size of the 3D marker, models with assumed poses are predefined and distributed in the 3D search space in front of the cameras. And then, each model is projected onto the two camera images, as shown in Fig. 3, where  $\Sigma_M$  is the 3D marker coordinate system used in the proposed system.  $\Sigma_{M_i}$  is the  $i$ -th model coordinate system.  $\Sigma_{CL}$  and  $\Sigma_{CR}$  are the left and right camera coordinate systems.  $\Sigma_{IL}$  and  $\Sigma_{IR}$  are the left and right image coordinate systems. The origins of  $\Sigma_M$  and  $\Sigma_{M_i}$  are the intersections of the three lines perpendicular to the faces to which the spheres of 3D marker face are attached. The  $j$ -th point on the  $i$ -th model in 3D space is projected onto the left and right camera images correctly, and these positions are calculated from a camera projection geometry.



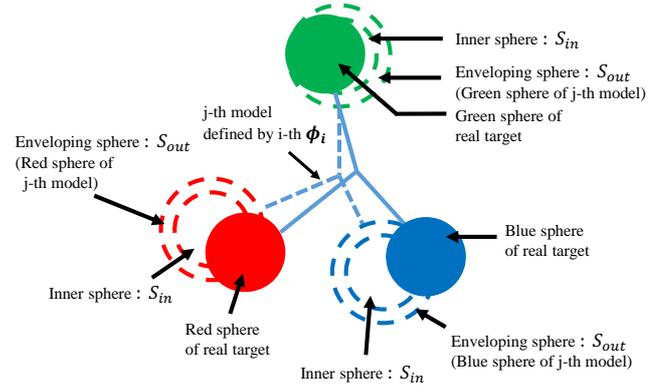
**Fig. 3** 3D model based matching system with dual-eye camera, including concepts of 3D-to-2D projection and 2D-to-3D reconstruction

### 3.2 Fitness value

In this study, the fitness function is used to measure the matching degree between the captured image and the projected models. The fitness function evaluates the correlation between the  $i$ -th search model with its assumed pose  $\phi_i$  and the real 3D marker in the left and right images.  $\phi_i$  means the pose of the  $i$ -th model given by the RM-GA. The RM-GA evolution process will be described in section 3.3.

Figure 4 shows the real 3D marker projected onto the image plane and the dotted circle that is the  $i$ -th

model obtained from the 3D-to-2D projection to the same plane. Each model consists of the red, green, and blue ball. Each ball of the model comprises an inner sphere  $S_{in}$  and an enveloping sphere  $S_{out}$ . The inner sphere  $S_{in}$  is intended to evaluate the ball area of the real target, and the enveloping sphere  $S_{out}$  is for the background area.



**Fig. 4** Left or right camera's 2D image of the real 3D marker and model

The fitness function used in the proposed system is defined as the following Eq. (1).

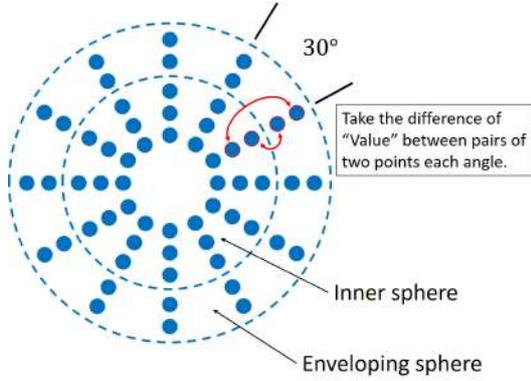
$$F(\phi_i) = \frac{1}{2} (F_L + F_R)$$

$$F_L = \frac{1}{N} \left( \sum_{{}^{IL}\mathbf{r}_j(\phi_i) \in S_{L,in}(\phi_i)} p({}^{IL}\mathbf{r}_j(\phi_i)) - \sum_{{}^{IL}\mathbf{r}_j(\phi_i) \in S_{L,out}(\phi_i)} p({}^{IL}\mathbf{r}_j(\phi_i)) \right)$$

$$F_R = \frac{1}{N} \left( \sum_{{}^{IR}\mathbf{r}_j(\phi_i) \in S_{R,in}(\phi_i)} p({}^{IR}\mathbf{r}_j(\phi_i)) - \sum_{{}^{IR}\mathbf{r}_j(\phi_i) \in S_{R,out}(\phi_i)} p({}^{IR}\mathbf{r}_j(\phi_i)) \right) \quad (1)$$

The fitness function for the model  $F(\phi_i)$  with assumed pose  $\phi_i$  is calculated by averaging the fitness functions of both the left camera image  $F_L$  and right camera image  $F_R$ . The summation in Eq. (1) is concerning the  $j$ -th point  ${}^{IL}\mathbf{r}_j(\phi_i)$  in the left camera image, defined on the  $i$ -th 3D model, whose pose is  $\phi_i$ . The

score is evaluated for each point by using the Eq. (1), where  $N$  represents the number of points to be evaluated. As shown in Fig. 5, there is a total of 60 points ( $36(S_{in})$  points in the inner sphere and 24 ( $S_{out}$ ) points in the enveloping sphere).

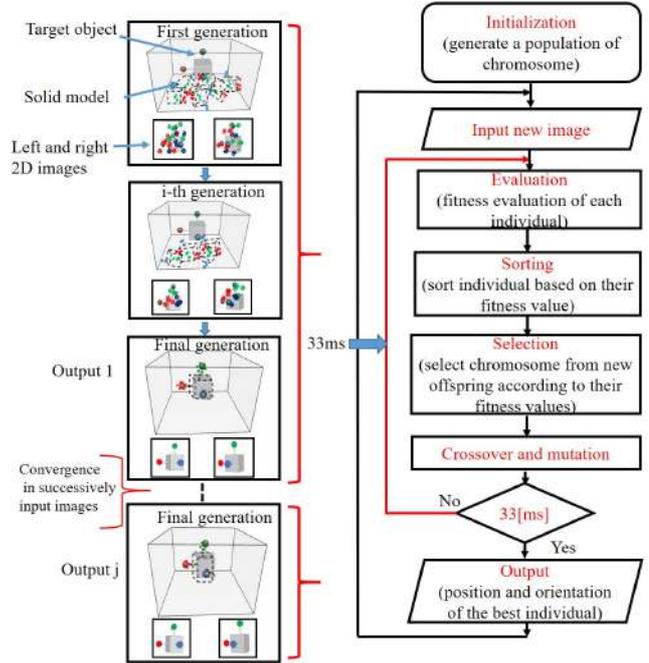


**Fig. 5** Projection of the blue sphere of a model with selected sample points

In the projection, the diameter of the inner sphere is the same as that of the actual sphere. If the  $j$ -th point of  $i$ -th model defined by  $\phi_i$  in  $S_{in}$  overlaps with real 3D marker and the model's hue color value coincides with the projected real 3D marker's hue value, in this case, the fitness function has a value of  $p({}^L\mathbf{r}_j(\phi_i)) = +1$  in Eq. (1), making the fitness value increase. Otherwise, if  $p({}^L\mathbf{r}_j(\phi_i)) = -1$ , then the fitness value decreases. Therefore, the fitness value is maximum if the model and the real target exactly coincide, and the system obtains the true pose of the real target 3D marker that gives maximum point in the fitness function distribution.

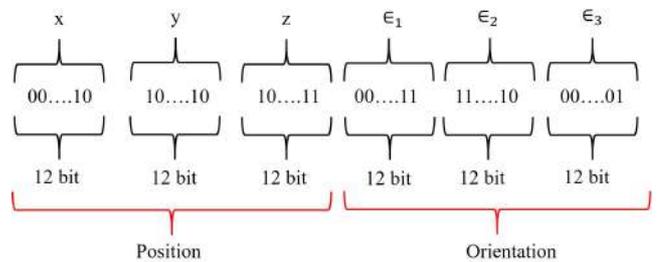
### 3.3 Real-time multi-step genetic algorithm

In the process of the 3D pose estimation, it is assumed that there are many models in the search area, as shown in the left of Fig. 6. In the right of Fig. 6 shows the flowchart of the RM-GA and how the best model is obtained. The previous research [16] explained why and how the RM-GA developed for the real-time 3D pose estimation. For determining which model is closest to the actual target, the fitness function defined in the previous section was used to quantify the correlation between the models and the target. The main task of the pose estimation process is to search for the optimal model with the pose that is most strongly correlated with that of the real 3D marker.



**Fig. 6** Flowchart of the RM-GA. In the flowchart, the terminal condition is defined as 33ms because of the video frame rate, 30 frames per second, of our camera system.

As shown in Fig. 7, the pose of the model is expressed as a 72 bits string to represent six pose parameters ( $x$ ,  $y$ ,  $z$ ,  $\epsilon_1$ ,  $\epsilon_2$ , and  $\epsilon_3$ ). The first 36 bits (12 bits each for  $x$ ,  $y$ , and  $z$ ) represent the position coordinates of the position of the 3-D marker. The last 36 bits (12 bits each for  $\epsilon_1$ ,  $\epsilon_2$ , and  $\epsilon_3$ ) describe the orientation defined by a quaternion.



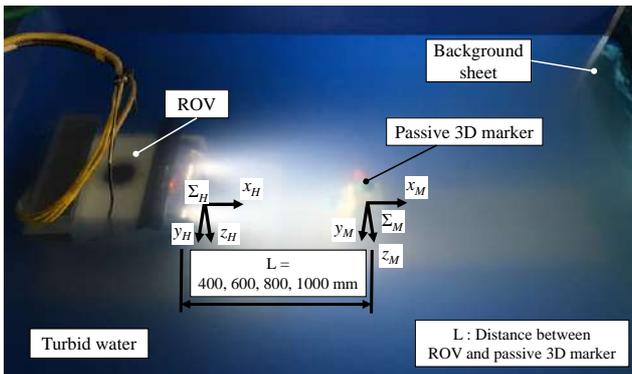
**Fig. 7** Structure of position and orientation represented by a gene

#### 4 Comparison experiment of recognition performance using the active and passive 3D marker

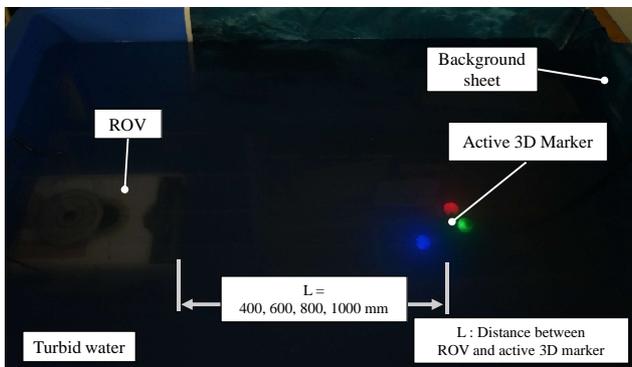
This chapter shows comparison experiments between the passive and active 3D marker for discussing the tolerance against turbid water in the simulated pool.

##### 4.1 Experiment environment

The experimental layouts of the recognition experiments are shown in Figs. 8 and 9. In Figs 8,  $\Sigma_H$  and  $\Sigma_M$  are the coordinate systems of the ROV. The ROV is fixed in the pool so that the pose between the dual-camera of ROV and the 3D marker is kept at constant. The distance between the dual-camera and ROV is set to 400, 600, 800 and 1000 mm.



**Fig. 8** Experimental environment by using the passive marker



**Fig. 9** Experimental environment by using the active marker.

##### 4.2 Turbidity environment

The turbidity environment is created by putting milk in the simulated pool. According to the other researches [23,24], there are 10 to 600 [nm] particles in milk. When the light runs through the size of 10 nm particles, light will be scattered equally in front and rear direction. When the light runs through the size of 100 nm particles, the light will be scattered to forward. “Forward” means the travelling direction of light. In other words, the experiment using milk is possible to create a turbidity environment that considers the various light scattering. In [23], the maximum amount of milk using the experiment was  $1.9 \times 10^2$  ml/m<sup>3</sup>. In [24], the maximum amount of milk was  $1.5 \times 10^2$  ml/m<sup>3</sup>. Those researches described the monocular image recognition. There is no discussion on pose estimation using the dual-eye cameras. In this study, the pose recognizable range using dual-eye cameras and passive/active 3D markers on different turbidity was analyzed.

In the experiment using the passive 3D marker, the amount of milk added each time is 2.435 ml/m<sup>3</sup>(2g) from 0 to  $3.64 \times 10$  ml/m<sup>3</sup> and 4.870 ml/m<sup>3</sup>(4g) from  $3.64 \times 10$  to  $1.19 \times 10^2$  ml/m<sup>3</sup>. In the experiment using the active 3D marker, the amount of milk added each time from 0 to  $1.19 \times 10^2$  is same as the experiment using the passive 3D marker. The amount of milk added each time is 9.740 ml/m<sup>3</sup>(8g) from  $1.19 \times 10^2$  to  $2.45 \times 10^2$  ml/m<sup>3</sup>. The reason for why it has a different range of turbid environment between the active and passive 3D marker is that the recognition ability using the active 3D marker is possibly better than using the passive 3D marker.

##### 4.3 Light environment

The passive marker was lighted from the ROV’s LED. On the other hand, the active marker was not lighted from any light device in this experiment. According to the previous experiments [25], the LED current of the active 3D marker has been confirmed that the red, green and blue LEDs are set to 5.7 mA, 3.4 mA and 3.4 mA, respectively.

##### 4.4 Evaluation of recognition result

For the evaluation of recognition performance on different turbidity environments, the fitness value is calculated by evaluating the correlation between the 3D marker and 3D search model. The recognition level is evaluated by using the average of the fitness value between 60[s] from the start of recognition. Based on the

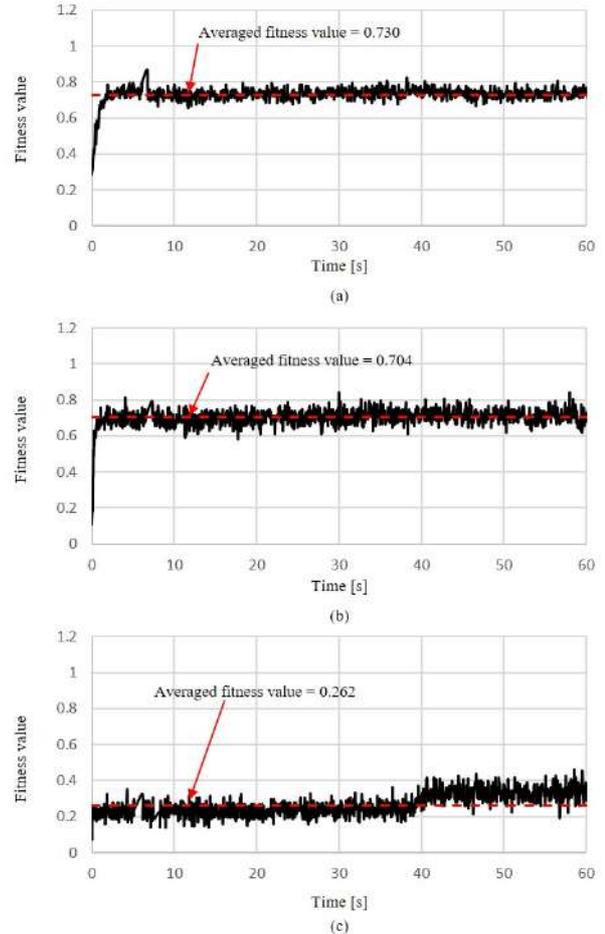
average of the fitness value, Area I is defined as possible docking range ( $F \geq 0.60$ ). Area II is a recognizable range ( $0.22 \leq F < 0.60$ ) that means dockable but not as reliable as Area I. The remaining area (Area III) ( $F < 0.22$ ) represents that the 3D marker cannot be detected by the RM-GA.

#### 4.5 Experimental results

The results of recognition experiment using the passive 3D marker are shown in Table 1. The amount of milk 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 in each condition. Figure 10 shows example results of the time-series and average fitness values in Table 1 (A), (B) and (C). It can be confirmed that the fitness value decreases as the distance between the ROV and 3D marker and the turbidity in the simulation pool increase. The left and right camera images of the ROV are shown in Fig. 11. The dotted line in the image represents the pose estimation results of the 3D marker. The images of Area I and II show that the 3D search model converges to the 3D marker. However, in Area III, the 3D search model cannot converge to the 3D marker because the images were almost whiteout. Therefore, we confirmed that the estimated pose of the RM-GA is not correct in Area III.

Next, the results of recognition the experiments using the active 3D marker are shown in Table 2. Figure 12 shows example results of the time-series and average fitness values in Table 2 (A), (B) and (C). The left and right camera images are shown in Fig. 13. Unlike the result of the passive 3D marker, the turbidity with the highest fitness value is not 0, but a low turbidity level as shown in figure 12(b). The LED current in the active 3D marker is set for having the best performance when the environment has a low turbidity level. On the other hand, the fitness value also decreases as the distance between the ROV and 3D marker and the turbidity in the simulation pool increases.

Compared with Table 1, the active 3D marker with no-lighting from a vehicle is more recognizable in the dark and turbid environment than the passive 3D marker lighted from the vehicle because the Area I and Area II in the active 3D marker are expanded. In addition, the camera images using the active 3D maker are clearer than using the passive 3D maker lighted from the vehicle. In this way, our proposed active 3D marker expands the docking ability from the viewpoints of the turbidity and distance.



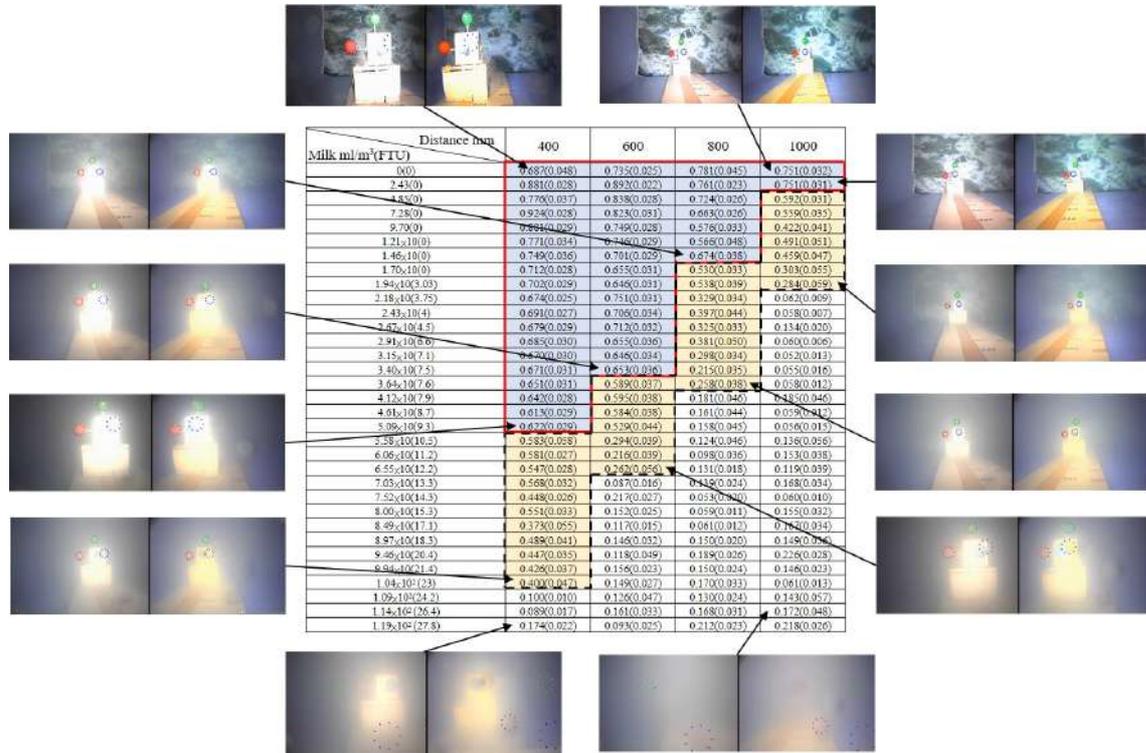
**Fig. 10** Real-time and average fitness values 0.730 under the conditions designated by (A) in Table 1 are calculated by averaging above real-time measurement result (a), and (B) and (C) in Table 1 are calculated by above (b) and (c).

#### 5 Docking experiment in the actual sea

In this chapter, we demonstrate the docking experiment in real sea environment with dark and turbid condition. At first, we explain our docking system using an ROV with dual-camera. Figure 14 shows block diagram of the proposed system with real-time 3D pose estimation and 3D motion controller implemented on a PC. Images from the dual-eye camera installed on the ROV are sent to the PC. Real-time pose estimation using 3D model-based matching method and RM-GA is implemented in PC. Based on the error between the desired pose and estimated pose, 3D motion controller outputs control signals to control the thrusters of the vehicle. The interface unit is for image capturing and digital-to-analog converting between the vehicle and PC.

**Table 1** The experimental results by using the passive 3D marker

Milk ml/m <sup>3</sup> (FTU)	Distance mm				
	400	600	800	1000	
0(0)	0.687(0.048)	<b>(A)</b> 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)	
1.21×10(0)	0.771(0.034)	0.746(0.029)	0.566(0.048)	0.491(0.051)	Area II
1.46×10(0)	0.749(0.036)	0.701(0.029)	0.674(0.038)	0.459(0.047)	
1.70×10(0)	0.712(0.028)	0.655(0.031)	0.530(0.033)	0.303(0.055)	
1.94×10(3.03)	0.702(0.029)	0.646(0.031)	0.538(0.039)	0.284(0.059)	
2.18×10(3.75)	0.674(0.025)	0.751(0.031)	0.329(0.034)	0.062(0.009)	
2.43×10(4)	0.691(0.027)	<b>(B)</b> 0.706(0.034)	0.397(0.044)	0.058(0.007)	
2.67×10(4.5)	0.679(0.029)	0.712(0.032)	0.325(0.033)	0.134(0.020)	
2.91×10(6.6)	0.685(0.030)	0.655(0.036)	0.381(0.050)	0.060(0.006)	
3.15×10(7.1)	0.670(0.030)	0.646(0.034)	0.298(0.034)	0.052(0.013)	
3.40×10(7.5)	0.671(0.031)	0.653(0.036)	0.215(0.035)	0.055(0.016)	
3.64×10(7.6)	0.651(0.031)	0.589(0.037)	0.258(0.038)	0.058(0.012)	Area III
4.12×10(7.9)	0.642(0.028)	0.595(0.038)	0.181(0.046)	0.185(0.046)	
4.61×10(8.7)	0.613(0.029)	0.584(0.038)	0.161(0.044)	0.059(0.012)	
5.09×10(9.3)	0.622(0.029)	0.529(0.044)	0.158(0.045)	0.056(0.015)	
5.58×10(10.5)	0.583(0.058)	0.294(0.039)	0.124(0.046)	0.136(0.056)	
6.06×10(11.2)	0.581(0.027)	0.216(0.039)	0.098(0.036)	0.153(0.038)	
6.55×10(12.2)	0.547(0.028)	<b>(C)</b> 0.262(0.056)	0.131(0.018)	0.119(0.039)	
7.03×10(13.3)	0.568(0.032)	0.087(0.016)	0.139(0.024)	0.168(0.034)	
7.52×10(14.3)	0.448(0.026)	0.217(0.027)	0.053(0.020)	0.060(0.010)	
8.00×10(15.3)	0.551(0.033)	0.152(0.025)	0.059(0.011)	0.155(0.032)	
8.49×10(17.1)	0.373(0.055)	0.117(0.015)	0.061(0.012)	0.167(0.034)	
8.97×10(18.3)	0.489(0.041)	0.146(0.032)	0.150(0.020)	0.149(0.036)	
9.46×10(20.4)	0.447(0.035)	0.118(0.049)	0.189(0.026)	0.226(0.028)	
9.94×10(21.4)	0.426(0.037)	0.156(0.023)	0.150(0.024)	0.146(0.023)	
1.04×10 <sup>2</sup> (23)	0.400(0.047)	0.149(0.027)	0.170(0.033)	0.061(0.013)	
1.09×10 <sup>2</sup> (24.2)	0.100(0.010)	0.126(0.047)	0.130(0.024)	0.143(0.057)	
1.14×10 <sup>2</sup> (26.4)	0.089(0.017)	0.161(0.033)	0.168(0.031)	0.172(0.048)	
1.19×10 <sup>2</sup> (27.8)	0.174(0.022)	0.093(0.025)	0.212(0.023)	0.218(0.026)	



**Fig. 11** The left and right camera images are taken under different turbidity conditions and distances, which are indicated by the arrow. Images taken at the maximum and minimum distances in clean water and at the maximum turbidity, in which the passive 3D marker is not observable, are also shown at the top and bottom, respectively.

**Table 2** The experimental results by using the active 3D marker

Distance mm \ Milk ml/m <sup>3</sup> (FTU)	400	500	800	1000
0(0)	0.792(0.040)	(A)0.777(0.042)	0.803(0.044)	0.829(0.052)
2.43(0)	0.795(0.040)	0.863(0.043)	0.859(0.047)	0.802(0.051)
4.85(0.57)	0.811(0.040)	0.885(0.043)	0.837(0.044)	0.770(0.045)
7.28(0.77)	0.801(0.039)	0.929(0.038)	0.824(0.042)	0.797(0.045)
9.70(1.03)	0.845(0.036)	0.847(0.038)	0.838(0.040)	0.776(0.047)
1.21×10 <sup>1</sup> (1.2)	0.865(0.038)	0.932(0.037)	0.849(0.040)	0.786(0.042)
1.46×10 <sup>1</sup> (1.6)	0.852(0.034)	0.917(0.039)	0.842(0.038)	0.767(0.042)
1.70×10 <sup>1</sup> (1.77)	0.862(0.032)	0.931(0.034)	0.816(0.040)	0.722(0.044)
1.94×10 <sup>1</sup> (1.7)	0.854(0.032)	0.927(0.032)	0.818(0.039)	0.705(0.045)
2.18×10 <sup>1</sup> (2.47)	0.854(0.031)	0.916(0.035)	0.806(0.038)	0.707(0.039)
2.43×10 <sup>1</sup> (2.6)	0.85(0.033)	0.917(0.033)	0.781(0.041)	0.673(0.040)
2.67×10 <sup>1</sup> (3.03)	0.841(0.032)	0.890(0.032)	0.765(0.039)	0.703(0.044)
2.91×10 <sup>1</sup> (3.5)	0.830(0.030)	0.885(0.032)	0.772(0.037)	0.608(0.038)
3.15×10 <sup>1</sup> (3.57)	0.817(0.031)	0.850(0.029)	0.750(0.037)	0.613(0.043)
3.40×10 <sup>1</sup> (3.9)	0.816(0.029)	0.875(0.034)	0.783(0.034)	0.610(0.040)
3.64×10 <sup>1</sup> (4.07)	0.803(0.028)	(B)0.840(0.033)	0.781(0.031)	0.607(0.040)
4.12×10 <sup>1</sup> (4.67)	0.814(0.031)	0.818(0.030)	0.744(0.035)	0.531(0.039)
4.61×10 <sup>1</sup> (5.17)	0.798(0.030)	0.803(0.032)	0.687(0.034)	0.447(0.045)
5.09×10 <sup>1</sup> (6.87)	0.799(0.031)	0.805(0.035)	0.628(0.038)	0.380(0.045)
5.58×10 <sup>1</sup> (7.47)	0.779(0.029)	0.780(0.034)	0.617(0.031)	0.314(0.046)
6.05×10 <sup>1</sup> (8.03)	0.792(0.030)	0.762(0.033)	0.560(0.031)	0.197(0.049)
6.55×10 <sup>1</sup> (8.67)	0.782(0.033)	0.741(0.034)	0.475(0.042)	0.168(0.043)
7.03×10 <sup>1</sup> (9.17)	0.793(0.033)	0.724(0.027)	0.491(0.035)	0.171(0.040)
7.52×10 <sup>1</sup> (9.8)	0.786(0.031)	0.683(0.028)	0.423(0.043)	0.089(0.033)
8.00×10 <sup>1</sup> (10.57)	0.763(0.031)	0.674(0.028)	0.452(0.040)	0.220(0.049)
8.49×10 <sup>1</sup> (11.33)	0.756(0.032)	0.645(0.032)	0.403(0.039)	0.187(0.048)
8.97×10 <sup>1</sup> (11.93)	0.732(0.033)	0.614(0.031)	0.345(0.041)	0.154(0.045)
9.46×10 <sup>1</sup> (12.23)	0.719(0.031)	(C)0.580(0.032)	0.244(0.028)	0.094(0.035)
9.94×10 <sup>1</sup> (12.73)	0.706(0.032)	0.572(0.034)	0.251(0.043)	0.108(0.037)
1.04×10 <sup>2</sup> (13.13)	0.703(0.033)	0.505(0.033)	0.218(0.039)	0.051(0.024)
1.09×10 <sup>2</sup> (14.07)	0.674(0.032)	0.469(0.036)	0.163(0.036)	0.049(0.024)
1.14×10 <sup>2</sup> (14.6)	0.670(0.033)	0.472(0.035)	0.167(0.037)	0.053(0.028)
1.19×10 <sup>2</sup> (15.03)	0.653(0.034)	0.401(0.036)	0.142(0.036)	0.041(0.020)
1.29×10 <sup>2</sup> (18.33)	0.625(0.034)	0.289(0.036)	0.108(0.037)	0.049(0.023)
1.38×10 <sup>2</sup> (18.5)	0.563(0.033)	0.272(0.038)	0.065(0.028)	0.054(0.027)
1.48×10 <sup>2</sup> (19.07)	0.552(0.033)	0.145(0.023)	0.070(0.031)	0.046(0.022)
1.58×10 <sup>2</sup> (20.33)	0.524(0.033)	0.181(0.034)	0.065(0.031)	0.049(0.027)
1.67×10 <sup>2</sup> (20.93)	0.471(0.032)	0.089(0.028)	0.049(0.025)	0.049(0.025)
1.77×10 <sup>2</sup> (22.8)	0.387(0.030)	0.109(0.022)	0.052(0.025)	0.046(0.022)
1.87×10 <sup>2</sup> (23.67)	0.428(0.040)	0.092(0.021)	0.050(0.027)	0.047(0.024)
1.96×10 <sup>2</sup> (24.97)	0.508(0.031)	0.091(0.023)	0.053(0.027)	0.072(0.028)
2.06×10 <sup>2</sup> (25.77)	0.236(0.021)	0.077(0.027)	0.073(0.033)	0.042(0.020)
2.16×10 <sup>2</sup> (27.73)	0.218(0.024)	0.067(0.024)	0.052(0.027)	0.040(0.019)
2.26×10 <sup>2</sup> (28.5)	0.198(0.025)	0.065(0.029)	0.034(0.019)	0.038(0.019)
2.35×10 <sup>2</sup> (29.47)	0.168(0.027)	0.059(0.027)	0.049(0.024)	0.047(0.023)
2.45×10 <sup>2</sup> (30.63)	0.167(0.025)	0.069(0.025)	0.045(0.024)	0.044(0.020)

Area I

Area II

Area III

## 5.1 Underwater vehicle

Figure 15 shows Hovering-type underwater vehicle (manufactured by Ko-wa cooperation). Two fixed cameras installed at the front of the vehicle are used for real-time pose tracking. In the thruster unit, four thrusters are controlled with maximum thrust force of 4.9N. The vehicle can dive up to 50m, and two LED light sources are also installed on the vehicle.

## 5.2 Docking procedure

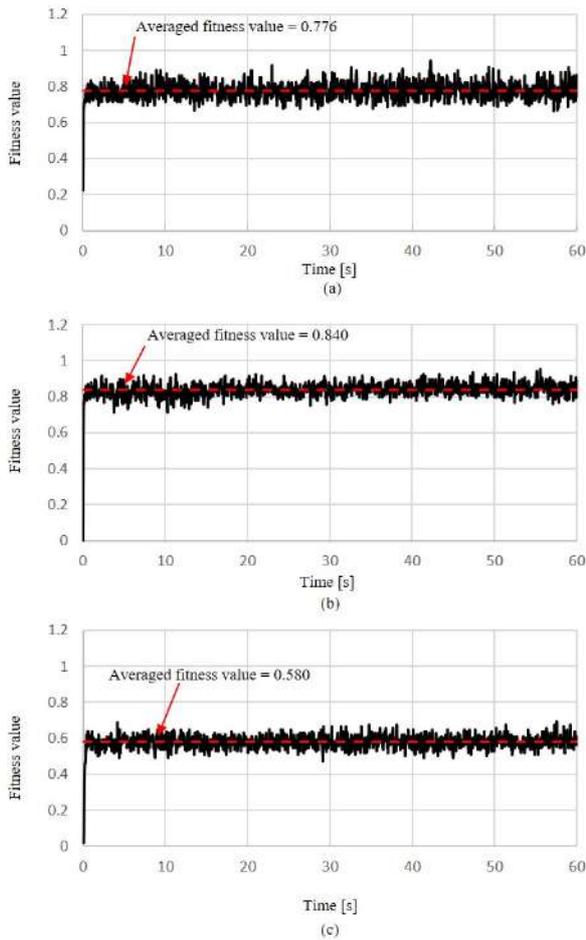
We designed a docking procedure, as shown in Fig.16. There are three steps to complete the docking operation.

(1) Approaching step: Normally, this step is performed by using a long-distance navigation sensor unit. In this

work, the vehicle was controlled by manually to approach the docking station till 3D marker was detected by the proposed system.

(2) Visual servoing step: After detecting 3D marker, relative pose between the vehicle and 3D marker is estimated by using the RM-GA. The vehicle was automatically controlled to the desired pose.

(3) Docking step: When the vehicle is stable in a defined position for defined period for a docking operation while visual servoing, docking step is performed in which the vehicle inserts its docking pole into the dock hole. Please note that whenever the relative pose error exceeds allowance range, the process switches to the visual servoing as shown as P in Fig.16.



**Fig. 12** Real-time and average fitness values 0.776 under the conditions designated by (A) in Table 2 are calculated by averaging above real-time measurement result (a), and (B) and (C) in Table 2 are calculated by above (b) and (c).

### 5.3 Sea trial experiment

We conducted the sea docking experiment at Setonai-kai in Japan for verifying the active 3D marker. The docking station was a rectangle of 60[cm]  $\times$  45[cm] and oriented with the long sides perpendicular to the shore. The ROV was tethered and connected by a cable with 200[mm] length to the onshore platform. Figures 17 and 18 show the layout of the docking experiment. In this experiment, the main task of the vehicle is to insert the docking pole into the docking hole automatically by our proposed visual servoing. Firstly, the vehicle manually approaches the docking station until the distance between the vehicle and 3D marker is less than 1.5[m] distance. In the visual servoing step, the vehicle is controlled to the desired pose for the docking. If the vehicle is satisfied with the desired pose, the vehicle goes ahead

with decreasing the desired position in x-axis direction gradually until it reaches 350[mm].

### 5.4 Experimental results

We conducted docking operation successfully by using the active 3D marker in the condition that was failed to the docking operation using the passive 3D. Figure 19 shows the fourth docking experiment of the vehicle during the docking process in the night time. Time profile of fitness value is shown in Fig. 19 (a). Recognized position of the vehicle in x, y and z axes is illustrated in Fig. 19 (b)-(d). From these results, the docking operation was successful since the position errors in y-axis and z-axis were within the predefined range that was  $\pm 40$ [mm] and the estimated position in x-axis was reached to 350[mm]. In addition, the experimental results show that the docking operation was a success within 50[s] by using the active 3D marker. Figure 20 shows a docking experiment in the night time. Figure 21 shows the overlay images in Front Camera 2 and the ROV Left Camera of Fig. 21 for emphasizing the result of the docking. In this way, the active 3D marker enables the ROV to perform docking to the station in the pitch-dark and turbid environment.

## 6 Conclusion

In this study, an active 3D marker was proposed to overcome the problem of difficulty to recognize the target in turbid environments. For confirming the effectiveness of the proposed active 3D marker, milk was put into the pool to create turbid environments. The recognition performance was compared between the active and passive 3D marker. According to the experimental results, it can be confirmed that using the active 3D marker is more recognizable under the high turbidity environment than using the passive 3D marker. In addition, the docking operation in the actual sea by using the active 3D marker is successful that originally failed in the same conditions by using passive 3D marker. In future work, the accuracy of the pose estimation using the active 3D marker will be verified for indicating the effectiveness and limitation of our stereo-vision-based pose estimation method.

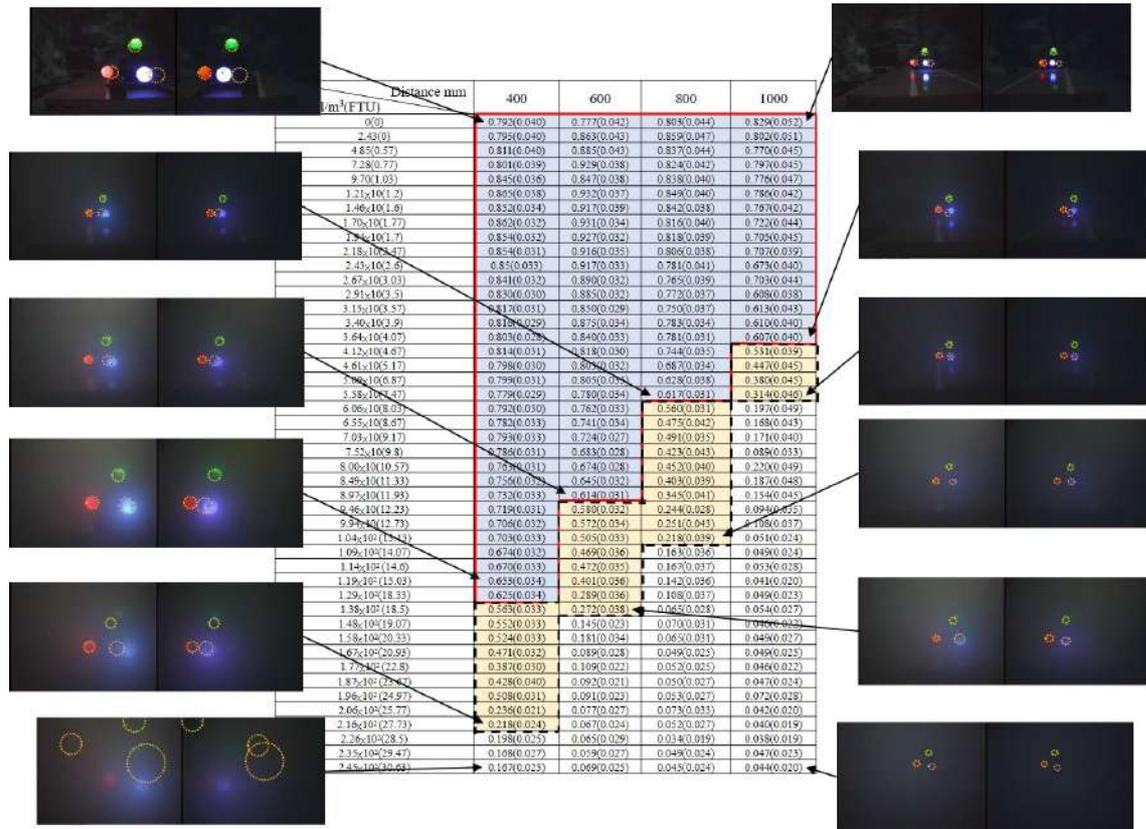


Fig. 13 The left and right camera images are taken under different turbidity conditions and distances, which are indicated by the arrow. Images taken at the maximum and minimum distances in clean water and at the maximum turbidity, in which the active 3D marker is not observable, are also shown at the top and bottom, respectively.

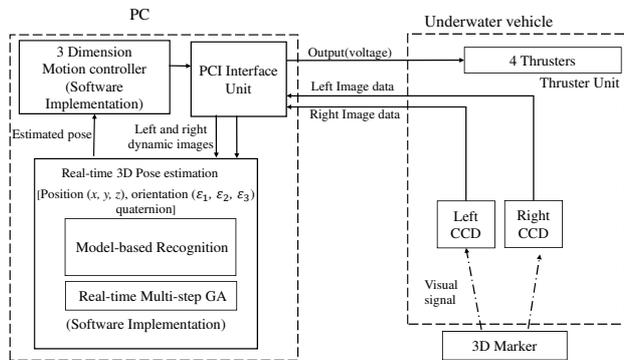


Fig. 14 Block diagram of the proposed system with real-time 3D pose estimation and 3D motion controller implemented on a pc.

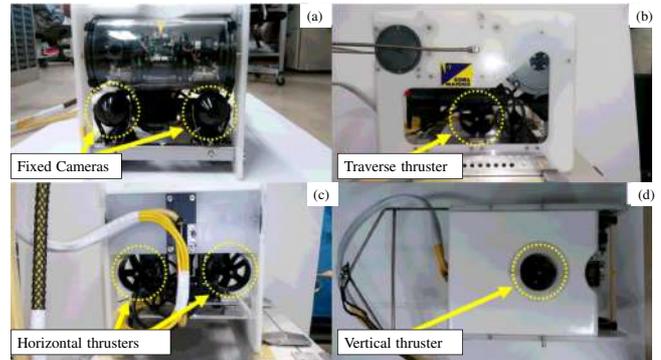


Fig. 15 Overview of the ROV (a)Front view (b)Side view (c)Back view (d)Top view

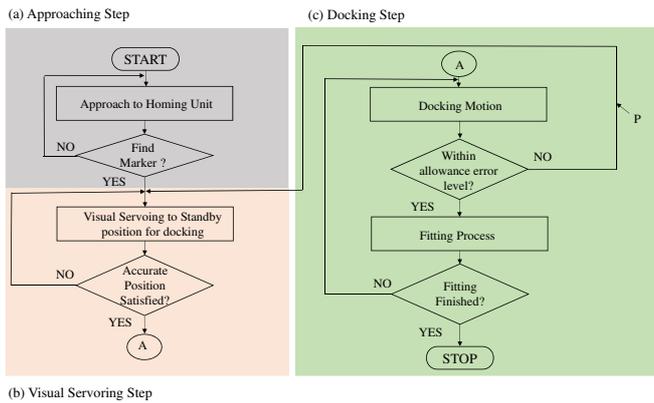


Fig. 16 Flowchart of Docking Strategy

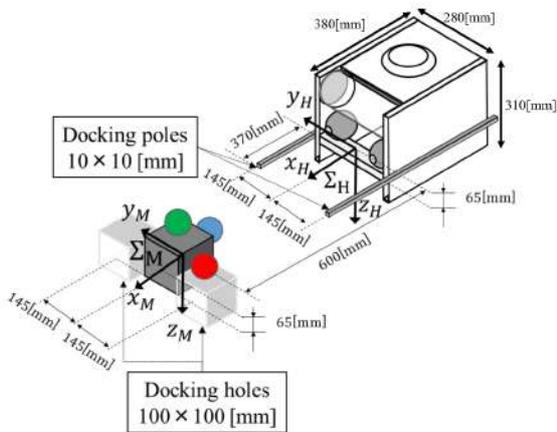


Fig. 17 Coordinate system of ROV and 3D marker

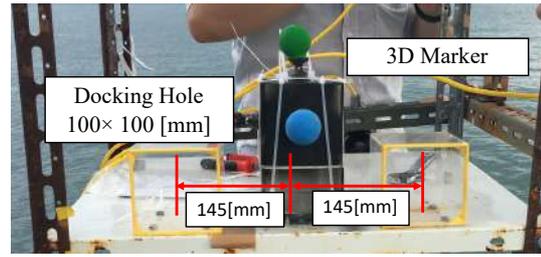


Fig. 18 Docking station

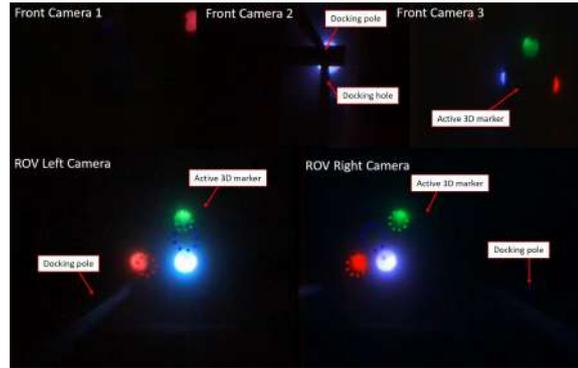


Fig. 20 Docking experiment in the actual sea environment

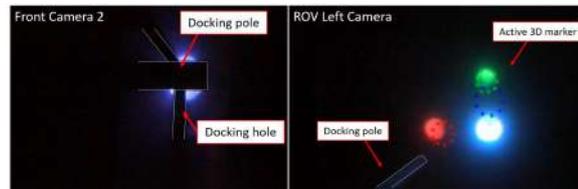


Fig. 21 Images were adding the support lines in Front Camera 2 and the ROV Left Camera of Fig. 20

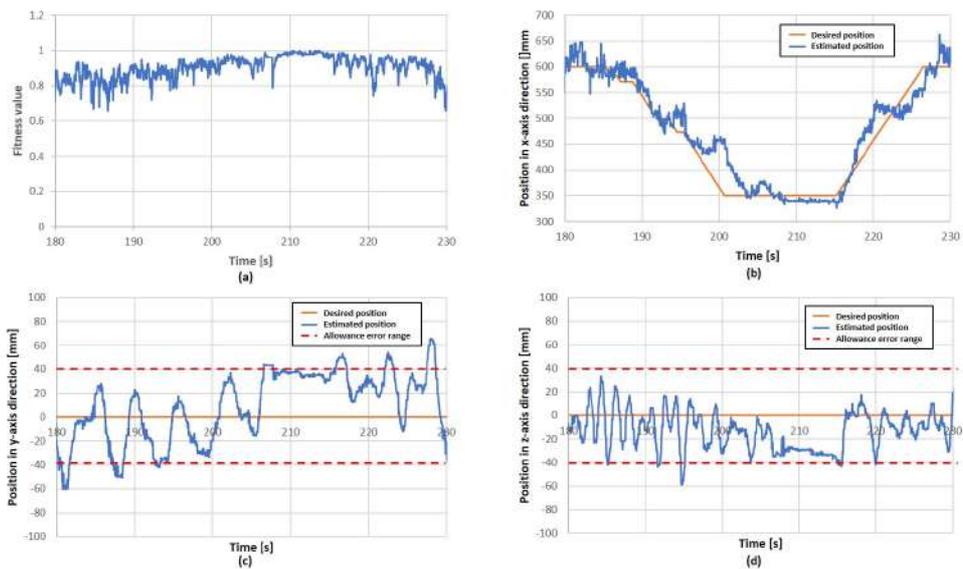


Fig. 19 Fourth-time trial docking experiment

## References

1. Jasper A (2012), Oil/Gas pipeline leak inspection and repair in underwater poor visibility conditions: challenges and perspectives. *Journal of Environmental Protection*, 3(05), 394.
2. Kume A, Maki T, Sakamaki T, Ura T (2013), A Method for Obtaining High-Coverage 3D Images of Rough Seafloor Using AUV-Real-Time Quality Evaluation and Path-Planning-. *JRM*, 25(2), 364-374.
3. Ribas D, Palomeras N, Ridao P, Carreras M, Mallios A (2011), Girona 500 auv: From survey to intervention. *IEEE/ASME Transactions on mechatronics*, 17(1), 46-53.
4. Krupinski S, Allibert G, Hua MD, Hamel T (2012), Pipeline tracking for fully-actuated autonomous underwater vehicle using visual servo control. In 2012 American control conference (ACC) (pp. 6196-6202). IEEE.
5. Yu SC, Ura T, Fujii T, Kondo H (2001), Navigation of autonomous underwater vehicles based on artificial underwater landmarks. In *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conference Proceedings* (IEEE Cat. No. 01CH37295) (Vol. 1, pp. 409-416). IEEE.
6. Cowen S, Briest S, Dombrowski J (1997), Underwater docking of autonomous undersea vehicles using optical terminal guidance. In *Oceans' 97. MTS/IEEE Conference Proceedings* (Vol. 2, pp. 1143-1147). IEEE.
7. Eustice RM, Pizarro O, Singh H (2008), Visually augmented navigation for autonomous underwater vehicles. *IEEE Journal of Oceanic Engineering*, 33(2), 103-122.
8. Jung J, Choi S, Choi HT, Myung H (2016), Localization of AUVs using depth information of underwater structures from a monocular camera. In 2016 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI) (pp. 444-446). IEEE.
9. Ghosh S, Ray R, Vadali SRK, Shome SN, Nandy S (2016), Reliable pose estimation of underwater dock using single camera: a scene invariant approach. *Machine Vision and Applications*, 27(2), 221-236.
10. Park JY, Jun BH, Lee PM, Oh J (2009), Experiments on vision guided docking of an autonomous underwater vehicle using one camera. *Ocean Engineering*, 36(1), 48-61.
11. Feezor MD, Sorrell FY, Blankinship PR, Bellingham JG (2001), Autonomous underwater vehicle homing/docking via electromagnetic guidance. *IEEE Journal of Oceanic Engineering*, 26(4), 515-521.
12. Negahdaripour S, Xu X, Khamene A (1998), A vision system for real-time positioning, navigation, and video mosaicing of sea floor imagery in the application of ROVs/AUVs. In *Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98* (Cat. No. 98EX201) (pp. 248-249). IEEE.
13. Palomeras N, Peñalver A, Massot-Campos M, Negre P, Fernandez J, Ridao P, ..., Oliver-Codina G (2016), I-AUV docking and panel intervention at sea. *Sensors*, 16(10), 1673.
14. Maire FD, Prasser D, Dunbabin M, Dawson M (2009), A vision based target detection system for docking of an autonomous underwater vehicle. In *Proceedings of the 2009 Australasian Conference on Robotics and Automation*. Australian Robotics and Automation Association.
15. Sagara S, Ambar RB, Takemura F (2013), A Stereo Vision System for Underwater Vehicle-Manipulator Systems-Proposal of a Novel Concept Using Pan-Tilt-Slide Cameras-. *JRM*, 25(5), 785-794.
16. Myint M, Yonemori K, Lwin KN, Yanou A, Minami M (2018), Dual-eyes vision-based docking system for autonomous underwater vehicle: an approach and experiments. *Journal of Intelligent & Robotic Systems*, 92(1), 159-186.
17. Myint M, Yonemori K, Yanou A, Ishiyama S, Minami M (2015), Robustness of visual-servo against air bubble disturbance of underwater vehicle system using three-dimensional marker and dual-eye cameras. In *OCEANS 2015-MTS/IEEE Washington* (pp. 1-8). IEEE.
18. Myint M, Yonemori K, Yanou A, Lwin KN, Minami M, Ishiyama S (2016), Visual Servoing for Underwater Vehicle Using Dual-Eyes Evolutionary Real-Time Pose Tracking. *JRM*, 28(4), 543-558.
19. Myint M, Yonemori K, Yanou A, Lwin KN, Minami M, Ishiyama S (2016), Visual-based deep sea docking simulation of underwater vehicle using dual-eyes cameras with lighting adaptation. In *OCEANS 2016-Shanghai* (pp. 1-8). IEEE.
20. Maki T, Shiroku R, Sato Y, Matsuda T, Sakamaki T, Ura T (2013), Docking method for hovering type AUVs by acoustic and visual positioning. In 2013 IEEE international underwater technology symposium (UT) (pp. 1-6). IEEE.
21. Lwin KN, Myint M, Mukada N, Yamada D, Matsuno T, Saitou K, ..., Minami M (2019), Sea Docking by Dual-eye Pose Estimation with Optimized Genetic Algorithm Parameters. *Journal of Intelligent & Robotic Systems*, 1-22.
22. Lwin KN., Mukada N, Myint M, Yamada D, Minami M, Matsuno T, ..., Godou W (2018), Docking at pool and sea by using active marker in turbid and day/night environment. *Artificial Life and Robotics*, 23(3), 409-419.
23. Garcia R, Gracias N (2011), Detection of interest points in turbid underwater images. In *OCEANS 2011 IEEE-Spain* (pp. 1-9). IEEE.
24. Codevilla F, Gaya JDO, Duarte N, Botelho S (2004), Achieving turbidity robustness on underwater images local feature detection. *International journal of computer vision*, 60(2), 91-110.
25. Kanda Y, Mukada N, Yamada D, Lwin K, Myint M, Yamashita K, ..., Minami M (2018), Development and Evaluation of Active/Lighting Marker in Turbidity and Illumination Variation Environments. In 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO) (pp. 1-7). IEEE.