SICE Journal of Control, Measurement, and System Integration, Vol. 12, No. 2, pp. 047–055, March 2019

Dual-Eye Vision-Based Docking Experiment in the Sea for Battery Recharging Application

Khin Nwe Lwin *, Myo Myint *, Kenta Yonemori *, Naoki Mukada *, Yoshiki Kanda *, Akira Yanou **, and Mamoru Minami *

Abstract: This paper presents a stereo-vision-based approach for sea-bottom docking of autonomous underwater vehicles (AUVs) for battery recharging. According to the intended application, a unidirectional docking station was designed in which the AUV has to dock from a specific direction. Real-time relative pose (position and orientation) estimation was implemented utilizing three-dimensional model-based matching to the actual target and a real-time multi-step genetic algorithm. Using the proposed approach, we conducted the experiments in which an AUV docked to a simulated underwater battery recharging station in the sea near Wakayama City, Japan. The experimental results confirmed the functionality and potential of the proposed approach for sea-bottom docking of AUVs. Although similar sea trials were reported previously, detailed discussions and performance analyses were not presented, especially regarding the relations among pose estimation, output control voltage, and photographic records. The analyses confirmed that the successful docking was realized and that the method has tolerance against turbulence applied to a remotely operated vehicle near the docking station.

Key Words: sea docking, stereo vision, visual servoing, genetic algorithm, autonomous underwater vehicle.

1. Introduction

Currently, autonomous underwater vehicles (AUVs) are being used in many applications, such as sea-bottom surveying, inspection of underwater structures (dams and bridges), and underwater cable tracking [1]. However, underwater vehicle operations are limited to activities that can be completed with the duration supported by the power capacity of their batteries. Even though advanced technology related to power devices provides long operation periods, underwater vehicles have to come back to a surface vessel for recharging when operations take more than a couple of days. To overcome this issue, underwater recharging stations with a docking function have been proposed using various approaches. Indeed, the docking of AUVs has become indispensable for advanced applications such as "sleeping" under a surface ship [2], downloading new mission instructions [3], and interventions using a manipulator installed on an AUV [4].

Because docking is thought to be an inevitable procedure for enabling battery recharging of AUVs at the sea bottom, many studies of underwater docking have used various approaches [2]–[10]. Depending on a docking station's structure for a specific application, different methods and sensors have been utilized. For simplicity and effectiveness of the intended application, the authors selected and designed a unidirectional station to which an underwater vehicle has to approach from a specific direction and perform docking at a single point of entry. Therefore, homing accuracy and robustness against disturbances, such as water currents, entered the picture as critical requirements for docking operations. To fulfill these demands, several studies have been conducted recently. The localization of a vehicle and a station was implemented using different kinds of sensors and techniques. Required navigation and homing accuracy varies from less than the meter level to the centimeter level, depending on the application. For precise docking accuracy for underwater battery recharging, visual information is more important than information collected from other sensors [5],[6].

Cameras used in docking by most of the researches are based on monocular vision system [7],[8]. The work in [7],[8] obtained the relative position and distance from a geometric arrangement of lights at the docking station. In such an approach, the calculation of relative orientation was more complicated and difficult than detection of the position. However, stereo vision can use parallactic displacement, which is effective for camera depth estimation in three-dimensional (3D) pose detection. On the other hand, monocular vision estimation methods cannot use parallactic displacement, the precision of distance measurement in the camera depth direction is not precise enough for applications in which high homing accuracy of AUVs is important. Even though two cameras were mounted on the vehicle in [9],[10], both cameras did not see the target object simultaneously, as one camera detected a target marker while the second camera was looking at something else to perform other tasks. Though efforts have been increasing to construct a practical docking system, there is still not a practical implementation useful for battery recharging at the actual sea bottom. The stereo-vision-based real-time control approach seems to have been researched only by the authors' group.

* Department of Intelligent Mechanical System, Okayama University, Okayama 700-8530, Japan
** Faculty of Health Science and Technology, Kawasaki University of Medical Welfare, 288 Matsushima, Kurashiki, Okayama 701-0193, Japan
E-mail: pdoj8yez@s.okayama-u.ac.jp, minami-m@cc.okayama-u.ac.jp
(Received April 18, 2018)
(Revised July 10, 2018)

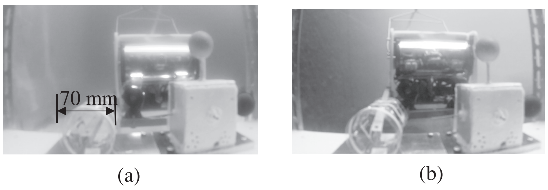

Fig. 1 Docking experiment for underwater battery recharging. (a) Docking experiment in the pool and (b) docking experiment in the sea.

In previous work [11]–[14], the authors concentrated especially on the performance of 3D pose estimation and visual servoing to achieve what are critical steps for real-time pose estimation and maneuvering for docking operations. In [11], the ability of dual-eye cameras to perform real-time pose tracking was confirmed when a target was moving, even when the captured images contained some noise due to air bubbles in front of the cameras. In that study, the remotely operated vehicle (ROV) was subjected to not only noise disturbance in images but also physical disturbance in the water stream induced by the buoyancy of the air bubbles. This confirmed that the proposed approach is robust enough to operate in actual sea conditions. In [12], 3D pose estimation when the target was partially obscured was discussed. The docking procedure was implemented, and its performance was evaluated in a pool in [13]. As follow-up work, the robustness of the system under a changing lighting environment was checked in [14]. However, this work [11]–[14] was conducted in an indoor pool for each individual objective. Therefore, a docking experiment using the proposed approach in the actual sea was the main task remaining to confirm the functionality and practicality of our approach for sea docking. Docking tests were conducted using the circular hole shown in Fig. 1 in both pool and sea tests. Finally, in this study, the docking experiments were conducted in the sea near Wakayama City, Japan, to evaluate the practicality of the system after further development of the proposed approach based on previous work. The specific contribution of this study is described in the following section.

2. Contribution of This Study

An ROV that can dive up to 50 m deep was used in this study. The feasibility of docking in an actual sea environment with lighting variation and turbidity disturbances as destabilizing factors for visual servoing was verified firstly by conducting docking tests in a pool. In previous work [15]–[17], sea docking experiments were conducted. The docking experiments were conducted on the coast of Wakayama Prefecture to evaluate how much robustness the proposed 3D move-on-sensing (3D-MoS) system would have in the natural sea environment [16]. (3D-MoS is what we call the dual-eye visual feedback motion control system). In [17], results of sea docking experiments were analyzed with a fitness distribution related to the real-time multi-step genetic algorithm (RM-GA) behavior. However, the study in [17] did not provide a detailed discussion of the time profile of the recognized pose with the respective output voltages to control the thrusters and the related photographs during the docking motion. The purpose of the present paper is to report these aspects rigorously. This point is important because sway and yaw motion has been observed several times when the ROV was positioned just in front of the docking station. The influences of these transient motion responses were stabilized by visual servoing, resulting in docking completion. The contributions of this study that are different from author's previous presentations [15]–[17] are as follows:

(1) A docking experiment in the actual sea was conducted to confirm the functionality of the proposed system.
(2) The effectiveness of the proposed dual-eye control system was evaluated from the viewpoint of practicality and functionality in natural environments by conducting a docking experiment, which included
- The time profile of the recognized pose by RM-GA,
- The relevant thruster output voltage, and
- Recording ROV poses, as depicted by photos exhibiting sway and yaw motion in turbulence.

Based on the above comprehensive analyses of docking behavior, practical docking performance was confirmed in the actual sea.

3. 3D-MoS System

In the proposed approach, only visual information is directly used for feedback control of 3D pose tracking in real time using dynamic images from two cameras input at a video rate of 30 frames per second. Figure 2 is a block diagram of the proposed system, which shows how images from dual-eye cameras mounted on the ROV are sent to and processed by the PC. Based on the error between the desired pose and estimated pose, the 3D motion controller outputs signals to direct thruster outputs through an interface unit to control the vehicle. The interface unit is for image capture and digital-to-analog conversion to output voltages for the vehicle thrusters.

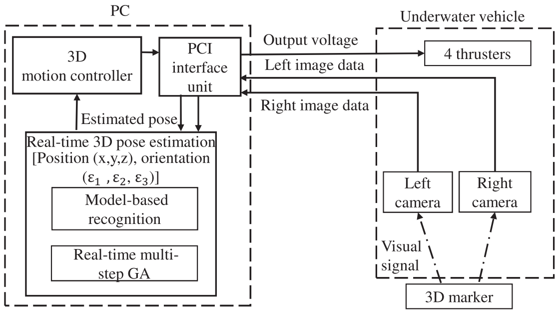

Fig. 2 Block diagram of proposed system with real-time 3D pose estimation and 3D motion controller implemented on a PC.

3.1 Docking Station

A unidirectional docking station was designed to simulate underwater battery recharging. The size of the docking station was 60 cm (L) × 45 cm (W) × 180 cm (H). Because a docking hole with a diameter of 7 cm and a 3D marker were fixed in the docking station, the vehicle has to approach and perform docking operation from one direction precisely. Two underwater cameras were installed in the docking station to record the docking operation.

3.2 Docking Procedure

The docking procedure was designed with three steps: approaching, visual servoing, and docking.

3.2.1 Approaching step

Normally, this step is performed using a long-distance navigation sensor unit. In this work, the vehicle was controlled manually to approach the docking station until the 3D marker was detected. This step will be extended in the future by using a long-distance navigation sensor to guide the vehicle until the 3D target marker comes into the field of view of the cameras.

3.2.2 Visual servoing step

After detecting the 3D marker, the relative pose between the vehicle and the 3D marker is estimated in real time using a 3D-model-based matching method and RM-GA. The vehicle was controlled automatically using the estimated pose to assume the desired pose.

3.2.3 Docking step

When the vehicle is stable at the desired pose with allowable error levels for docking operation by visual servoing, the docking step is performed in which the vehicle inserts its docking pole into the docking hole. Errors in the x, y, and z directions (surge, sway, and heave, respectively) are defined as less than ±20 mm, and the yaw angle θ error is defined as less than 3°, where pitch and roll angles are naturally stabilized and not controlled. Note that whenever the relative pose error exceeds the allowance range, the process switches to visual servoing. This switching process between the visual servoing and docking steps is intended to avoid any physical damage made by contact between the docking pole and hole.

3.3 3D Motion Controller

Because of self-stabilization, roll and pitch angles are neglected in controlling the movement of the vehicle. Therefore, 4 degrees of freedom (x, y, and z in mm and ε_3 in °) are variables in the 3D pose tracking control, as shown in Fig. 3. Even though a proportional–derivative (PD) controller appears to be more efficient when the damping characteristic of water is not sufficient for stabilizing control, it was found experimentally that the damping characteristic was sufficient when using a proportional (P) controller. Therefore, the simple P controller was used with appropriate gains to control all thrusters of the ROV instead of comparing the performance using the different controllers, such as PD and proportional–integral–derivative (PID) controllers with different gains. The P controller was applied in the control system with feedback using the pose estimated by RM-GA. In previous work [11]–[13], thrust in the y-axis direction was given by an on-off controller, and this has been upgraded to a P controller. Thus, in this study, the control voltages of all thrusters were calculated by the following proportional control laws:

$$\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{bmatrix} = \begin{bmatrix} k_{p1} & 0 & 0 & 0 \\ 0 & k_{p2} & 0 & 0 \\ 0 & 0 & k_{p3} & 0 \\ 0 & 0 & 0 & k_{p4} \end{bmatrix} \begin{bmatrix} x_d - x \\ y_d - y \\ z_d - z \\ \varepsilon_{3d} - \varepsilon_3 \end{bmatrix} + \begin{bmatrix} 2.5 \\ 2.5 \\ 2.5 \\ 2.5 \end{bmatrix}, \quad (1)$$

where x_d, y_d, z_d, and ε_{3d} are desired relative values based on Σ_H (see Fig. 5 shown later) against the 3D marker; v_1, v_2, and v_3 are the voltages to produce thrust in the x-axis, y-axis, and z-axis directions, respectively; ε_{3d} is the desired rotation angle around the z-axis; and v_4 is the voltage for torque around the z-axis. According to the thruster characteristics configured to stop at 2.5 V, the control voltage for each thruster is the differentiated value gained by proportional gain value and added by offset value, 2.5. Based on the experimental results, gain coefficients were tuned to perform better during visual servoing.

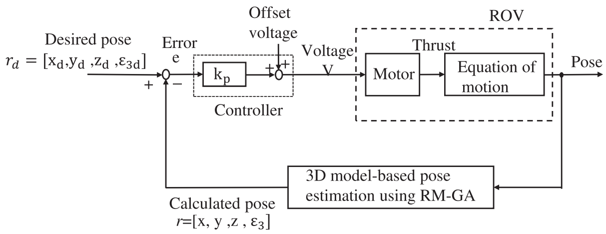

Fig. 3 Block diagram of real-time 3D pose tracking.

3.4 Real-Time 3D Pose Estimation

3.4.1 Projection direction

This section briefly explains the proposed 3D pose estimation method for the reader's convenience; a more detailed discussion can be found in [18]. Instead of calculating the absolute positions of the vehicle and the target at the docking station, the estimated relative pose between them is input as feedback to the control system. This avoids the limitation [19],[20] of feature-based stereo vision, whose visual servoing technique is called image-based, 3D model-based recognition based on 3D-to-2D projection was applied in the proposed system. The 3D-to-2D projection has the merit of avoiding the ill-posed problem of the 2D-to-3D reconstruction approach that hinders the 3D pose estimation when trying to measure the pose from the decomposed 2D image information of dual-eye cameras.

3.4.2 3D-model-based real-time pose estimation

Many 3D models use the same 3D information, such as shape, color, and size, with different poses allocated randomly in the search area. The real target object in 3D space is captured by dual-eye cameras and the poses of the model, each of which is represented by one RM-GA gene, are projected as 2D images. The difference in relative pose is calculated by comparing the projected 2D image model and the real target image that is captured by the dual-eye cameras. The estimated pose of the best model—that is, the model possessing the highest overlap between the target 3D marker and the model in space—is assumed to be a true estimated pose. Therefore, the problem of finding/recognizing the 3D marker and detecting its pose is converted into an optimization problem with a multi-peak distribution. To perform posing in real-time, GA is modified as RM-GA for searching for the best model in dynamic images whose poses represent the target 3D marker's pose. Note that the evaluation is performed in 2D space and the convergence occurs in 3D space.

3.4.3 Correlation function used as a fitness function

A fitness function is used to calculate the correlation between the model and the actual target in the images captured by two cameras using hue values. In other words, the intention of the fitness function is to have a dominant peak at the true pose of the target object. This fitness function with the highest peak at true pose of the 3D marker allows to change the problem to find a true pose into another problem to find the highest peak and the pose giving the peak, that is an optimization problem. This conversion rendered the RM-GA effective for real-time pose estimation [18].

Here the fitness function is briefly explained. Because the model has spheres with quantitative diameters rather than a point, shape, and color (red, green, and blue), information for these spheres is used when calculating the correlation between the model and the target object, as shown in Fig. 4. Figure 4 (a) shows the actual 3D marker. Figure 4 (b) is a model of the 3D marker with an enlarged view of the green ball, which consists of two portions. The first portion is the inner area that is the same size as the actual 3D marker, and the second portion is the outer area that is the background. The dots in Fig. 4 (b) represent points for calculating the correlation degree, that is, how much the inner area overlaps the green ball and the outer area does not overlap the green ball [21].

When the hue value of the point on the model is similar to the same point on the actual target, it just increases the fitness value. If the model's pose coincides with the pose of the actual 3D marker in 3D space, then all points in the inner area of models are on the 3D marker's circle, and all points in the outer area are not on the marker's circle. In this case, the fitness function has the highest value at the estimated true pose.

Finally, the fitness function has a maximum value when the pose of the searching model fits the one of the target object being imaged in the left and right camera images. The total fitness value is calculated from averaging two fitness functions based on the left and right cameras. A detailed explanation of the fitness function is given in [22],[23]. The concept of the fitness function in this study is an extension of the work in [23], in which different models, including a rectangular surface-strip model, were evaluated using images from a single camera.

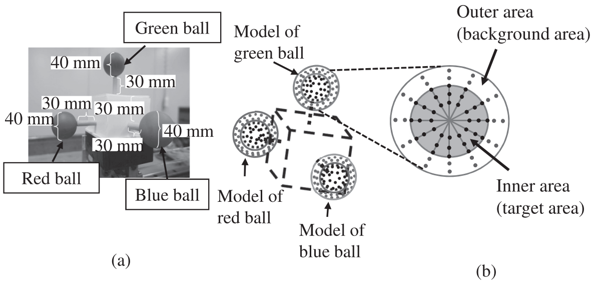

Fig. 4 Real 3D marker and model. (a) Real 3D marker and (b) model with enlarged view of the green ball model, where the inner area is the same size of the actual target object (green ball) and the outer area is the background area. The dots in panel (b) are points for calculating the correlation degree, that is, how much the inner area overlaps the green ball and the outer area does not overlap the green ball.

4. Experiment

4.1 Experimental Layout for Docking

The main task in this experiment was to insert a docking pole into the docking hole automatically by visual servoing. Figure 5 illustrates the relation of the ROV before and after docking. In the visual servoing step, the vehicle goes to the desired pose as shown in Fig. 5 (a), which is the condition ready for docking. When the vehicle is continuously stable within the position error range of ± 20 mm in the desired pose for 165 ms, the vehicle proceeds to insert the docking pole into the docking hole by decreasing the desired distance between the vehicle and target in the x-axis direction gradually until it reaches a distance of 350 mm from the 3D marker, as shown in Fig. 5 (b). The desired position and orientation during visual servoing are expressed mathematically below, with the corresponding values for docking completion given in parentheses.

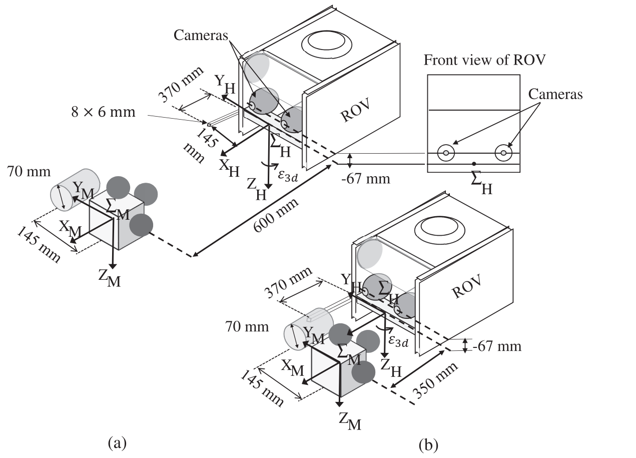

Fig. 5 Layout of docking experiment and description of the alignment process between the ROV and 3D marker. (a) Desired pose in visual servoing state and (b) desired pose in docking completion state.

Fig. 6 Appearance of the testing pool.

$$x_d = {}^H x_M = 600\ (350)\ \text{mm}, \tag{2}$$

$$y_d = {}^H y_M = 0\ (0)\ \text{mm}, \tag{3}$$

$$z_d = {}^H z_M = -30\ (-30)\ \text{mm}, \tag{4}$$

$$\varepsilon_{3d} = 0\ (0)^\circ, \tag{5}$$

where ${}^H x_M$ represents the x position of the origin of Σ_M in reference to Σ_H, and Σ_M is as defined in Fig. 5.

4.2 Docking Experiment in a Pool

To check the functionality of the proposed approach before the sea trial experiment, a docking experiment was conducted in a pool, as shown in Fig. 6. The pool is near the sea, so it was filled with seawater for a few months prior to the experiment. The pool water contained particles and debris, such as seaweed and dry leaves, that created noise in images captured for visual servoing. In addition, trees next to the pool cast shadows. Thus, the environment inside the pool naturally had poor lighting, which is consistent with our ongoing work to perform docking under turbid conditions with poor lighting. The docking station was placed on the bottom of the pool in this experiment.

4.3 Docking Experiment in the Sea

The docking experiment was conducted near Wakayama City, Japan. A total of five dockings were conducted. The

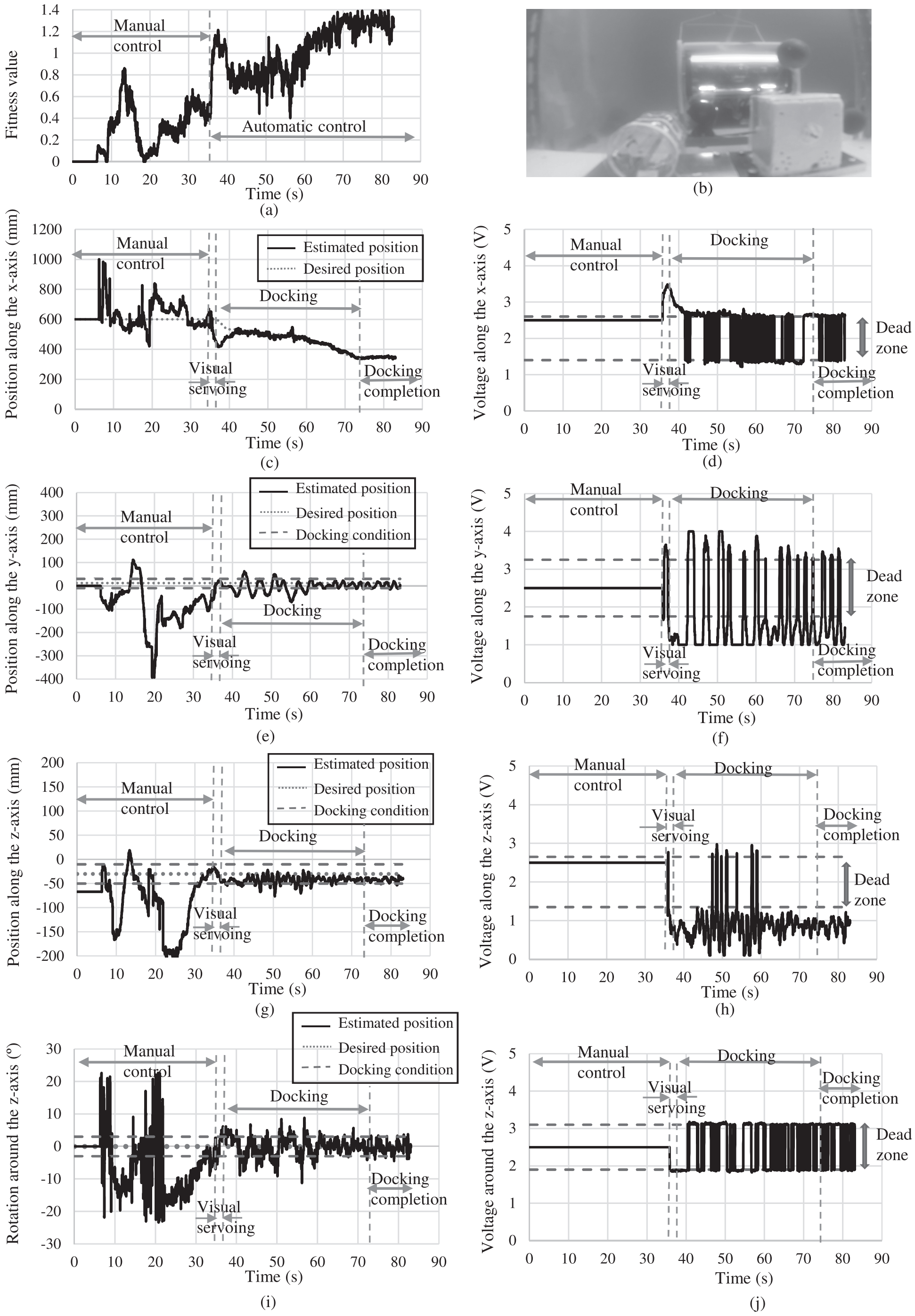

Fig. 7 Pool docking results. (a) Fitness value; (b) photograph of ROV after docking completion; (c), (e), (g), (i) recognized positions along the x-, y-, and z-axis directions and rotation around the z-axis, respectively, where the desired values for visual servoing state are given as x_d = 600 mm, y_d = 0 mm, z_d = -30 mm, ε_{3d} = 0°; and (d), (f), (h), (j) voltage commanding thrust in the x-, y-, and z-axis directions and rotation around the z-axis, respectively.

docking station was oriented with the long sides perpendicular to the shore. Docking tests began with the vehicle placed in front of the docking station at a distance of about 3.5 m. The buoyancy force was nearly 1.035 times that of freshwater, and gentle waves were rolling in during the experiments. The ROV was tethered and connected by a 200 m cable to the onshore platform.

5. Results and Discussion

5.1 Docking Experiment in a Pool

Figure 7 shows the experimental results of the docking operation in the pool. Figure 7 (a) illustrates the time profile of the fitness value. Figure 7 (b) is a photograph of the ROV as it completed docking. Figures 7 (c)-7 (j) show the relative position of the desired and recognized poses in each direction, as measured by RM-GA, and the output voltage for each thruster of the ROV. As shown in Fig. 7, the ROV was controlled manu-

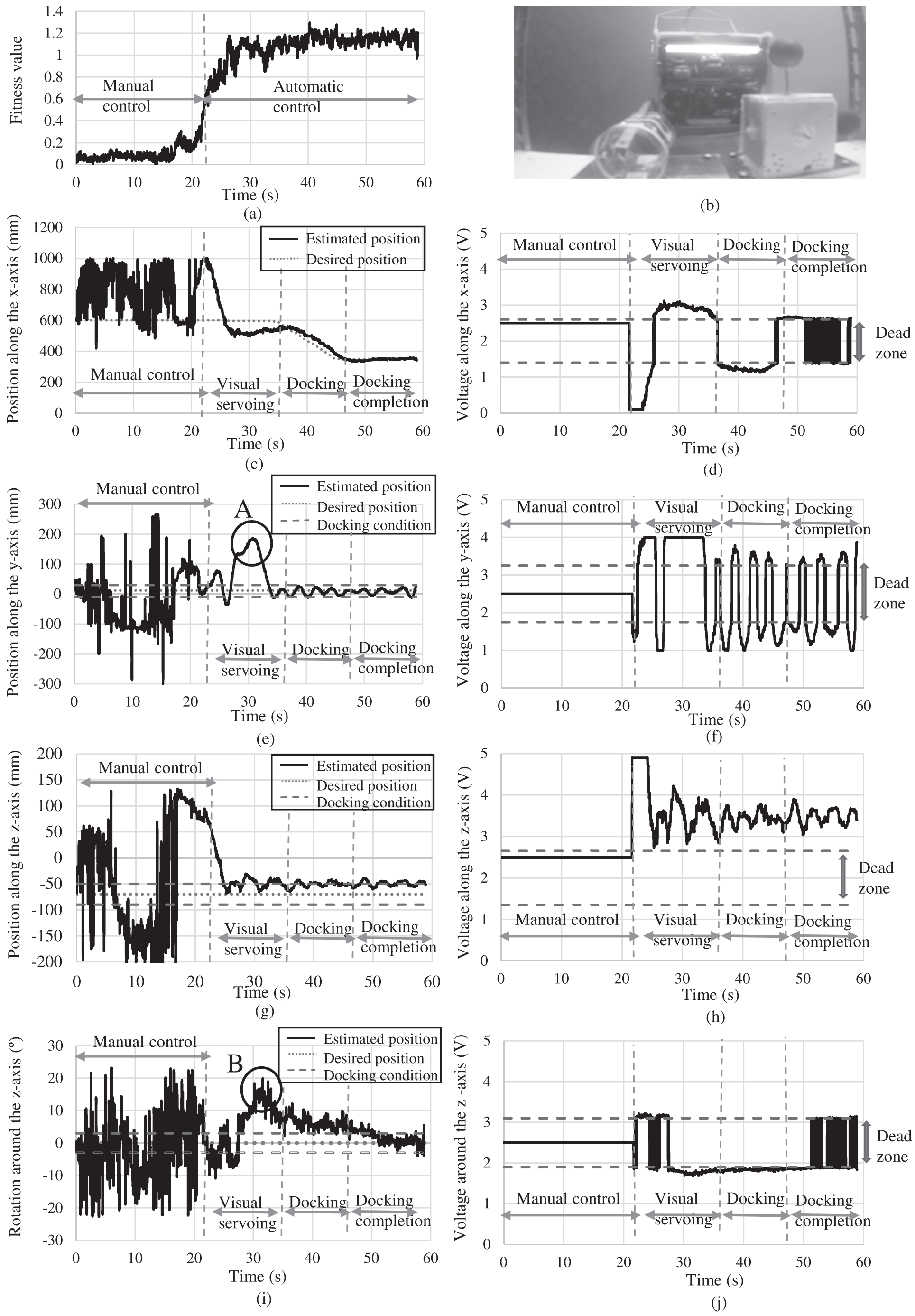

Fig. 8 Sea docking results. (a) Fitness value; (b) photograph of ROV in docking completion; (c), (e), (g), (i) recognized positions along the x-, y-, and z-axis directions and rotation around the z-axis, respectively, where the desired values for visual servoing state are given as x_d = 600 mm, y_d = 12 mm, $z_d = -70$ mm, $\varepsilon_{3d} = 0°$; and (d), (f), (h), (j) voltage along the x-, y-, and z-axis directions and rotation around the z-axis, respectively. A and B correspond to Fig. 9 (d), (e), indicated as C.

ally until the 3D marker was detected. Note that the recognized poses during manual control are not accurate, and they were not used in the feedback system. Automatic control was started when the fitness value increased above 0.6, and then docking was performed under automatic control. Because detection of the 3D marker by the RM-GA can be determined by the value of the fitness function, manual control could be switched to automatic control autonomously. In the manual operation, the controller was not operational when the output voltages were 2.5 V, as shown in panels (d), (f), (h), and (j) of Fig. 7. At the operation time of 35 s in Fig. 7 (a), the fitness value increases from about 0.4 to more than 1 when the distance between the ROV and the 3D marker decreased. Theoretically, the maximum fitness value is about 1.67. During the visual servoing stage, the controller starts to output the thruster commands in each direction (x-, y-, and z-axes) and orientation around the z-axis with the relative voltage values shown in panels (d), (f), (h), and (j), respectively, of Fig. 7.

When the operation time was about 70 s, docking was completed, and the ROV maintained the desired pose, as shown in

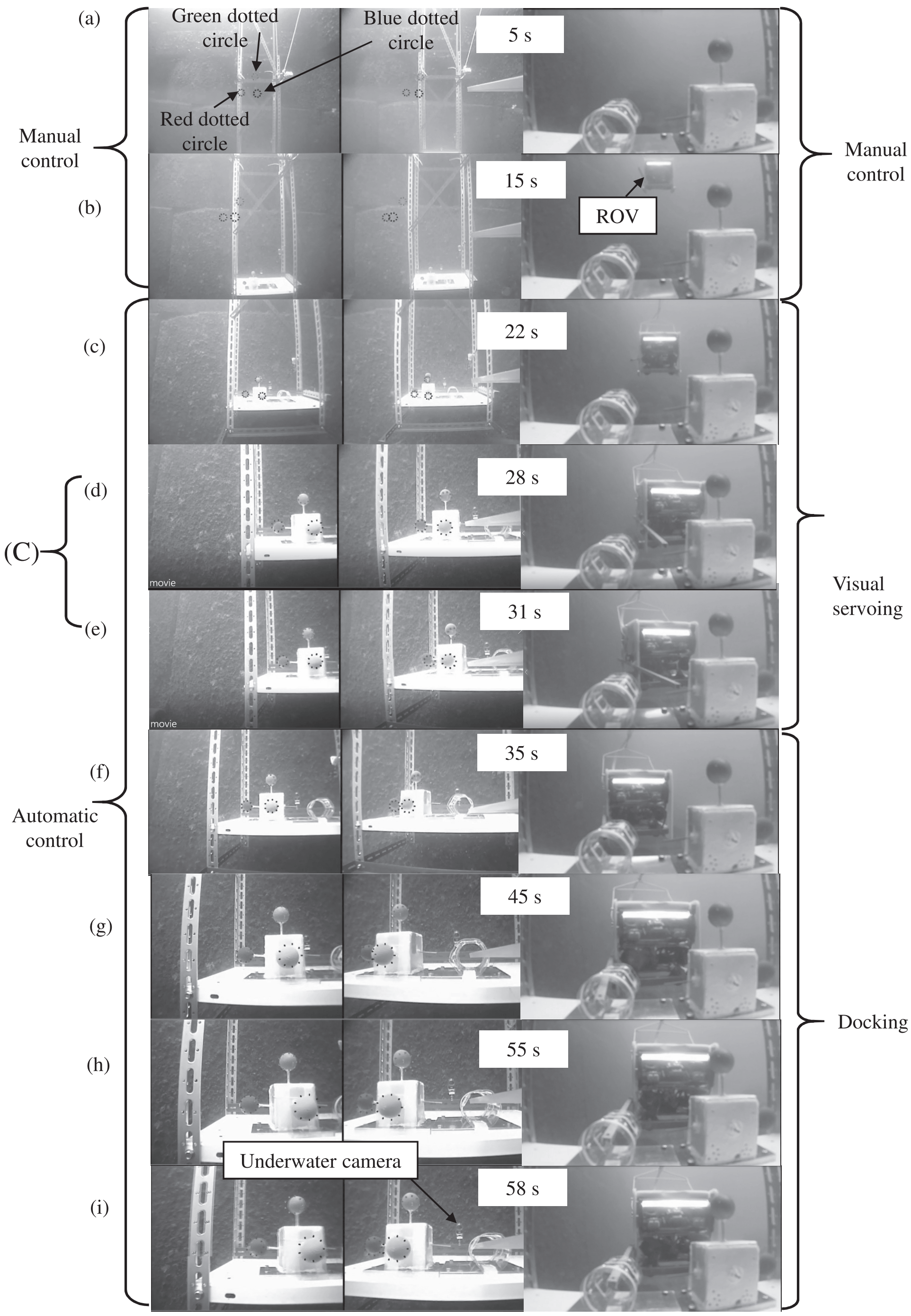

Fig. 9 Periodically captured images from dual-eye cameras during the docking operation and corresponding images from underwater camera installed in the docking station. (a), (b) Images of manual control, (c)–(e) images of visual servoing, and (f)–(i) images of docking. Dotted circles in dual-eye camera images are the poses recognized by RM-GA. The images labeled with (C) correspond to the results shown in Fig. 8 (e), (i) that are indicated by A and B, respectively.

Fig. 7 (b). The recognition results of the 3D marker using such information as hue value, size, and shape were confirmed experimentally to be reliable. According to the experimental result, the docking operation was completed successfully within 40 s after initiation of automatic control.

5.2 Sea Docking Experiment

Of the five sea docking tests, one docking failed, but the other four were successful. This paper provides a detailed explanation of the third docking test. The experimental results of this sea docking operation are depicted in Fig. 8. Initially, the ROV was controlled manually until it was 1.5 m from the docking station. The desired pose was $x_d = 600$ mm, $y_d = 12$ mm, $z_d = -70$ mm, and $\varepsilon_{3d} = 0°$. The visual servoing started after 22 s. Figure 8 (a) shows fitness values, and panels (c)-(j) show the recognized pose and respective output voltages for each thruster. Figure 8 (b) is a photograph taken by the underwater camera installed in the docking station. Figure 9 shows the periodic snapshots from the dual-eye cameras and the underwater camera as the ROV approached the 3D target marker by manual operation, visual servoing, and docking.

At the start of manual control, the fitness value before 17 s was below 0.2 in Fig. 8 (a), the output voltage for each thruster was 2.5 V in panels (d), (f), (h), and (j) of Fig. 8, which means that the controller was not controlling the ROV thrusters. The dotted circles do not entirely overlap the target 3D marker in the initial stage of manual control, as shown in Figs. 9 (a) and (b), which were taken from the two cameras of the ROV and the underwater camera set at the docking station at the operation times of 5 s and 15 s.

In the visual servoing step, the ROV approached from a distance of about 900 mm in the x-axis direction to the 3D marker by visual servoing. After 22 s, the 3D marker was detected, and the fitness value increased to more than 0.6, as shown in Fig. 8 (a). The controller commanded the forward thruster with an output voltage of about 0.1 V in the x-axis direction, as shown in Fig. 8 (d). At the same time, the controller commanded each thruster in the y- and z-axis directions and the orientation around the z-axis with the output voltages shown in panels (f), (h), and (j), respectively, in Fig. 8. The positions in the y- and z-axes are outside of the error allowance range at the time 30 s, as denoted by "A" and "B" in Figs. 8 (e), (i). This is a large deviation in the y-axis and orientation around the z-axis during the third docking operation of the four succeeded docking. The deviation of sway can be confirmed by Figs. 9 (d), (e), which are denoted by (C). However, the sway motion and rotation error stabilized at about 35 s, as shown in Fig. 9 (f). At that time, the dotted circle recognized by the RM-GA overlapped with the real 3D marker.

At the time of 35 s, the position in the y- and z-axes was within the error allowance range, as shown in Figs. 8 (e), (g) as visual servoing transitioned to docking. Note that, as shown by the z-axis voltage in Fig. 8 (h), the controller tried to direct the ROV in a downward direction with a positive output voltage because of the buoyancy of the sea water. The success of the docking operation was also confirmed by checking video images captured by the two ROV cameras and the underwater cameras installed at the docking station, as shown in Figs. 9 (f)-(i). The images taken when the operation times were 35 s, 45 s, 55 s, and 58 s illustrate the docking step. Regarding accuracy, it was confirmed experimentally that both recognition and docking accuracy was at the centimeter level because the docking hole radius was 35 mm and the allowance error was ± 20 mm. According to the experimental results, the sea docking operations were completed successfully within 40 s after control transitioned to automatic.

In the other successful docking tests, the fitness value was above 1 when the 3D marker was detected by the RM-GA system. All successful docking operations were completed within 40 s from the start of automatic control. Among the four successful sea docking experiments, the second docking operation had the shortest completion time, and the third docking operation had the longest completion time after automatic control was initiated. The sway motion of the first and third docking experiments was larger during the visual servoing state than in the other docking trials.

6. Conclusion

In this work, sea docking experiments of an underwater vehicle using two cameras and a 3D marker were implemented. A docking experiment in a pool was conducted first to confirm the functionality of the system. A docking station was designed and deployed at the sea bottom to verify the proposed approach for docking to a battery recharging station. Some pose fluctuations occurred because of natural disturbances like sea currents. However, the vehicle could be controlled by visual servoing using the proposed system, and the final docking operations were performed successfully. The docking experiment in the pool and natural sea environment proved the functionality and practicality of the proposed approach using stereo vision for docking to a battery recharging station in an actual sea. The docking trials were conducted in an actual sea to evaluate the proposed docking effectiveness. The performance of real-time pose tracking using standalone cameras and a 3D marker was confirmed to provide real-time 3D pose recognition, and it produced successful docking. The robustness of the proposed system under different illumination variation and water turbidity is our on-going research and will be discussed in future work.

References

[1] B.A.A.P. Balasuriya, M. Takai, W.C. Lam, T. Ura, and Y. Kuroda: Vision based autonomous underwater vehicle navigation: Underwater cable tracking, *Proc. MTS/IEEE OCEANS Conf.*, Vol. 2, pp. 1418–1424, 1997.

[2] M. Dunbabin, B. Lang, and B. Wood: Vision-based docking using an autonomous surface vehicle, *IEEE International Conference on Robotics and Automation*, pp. 26–32, 2008.

[3] R.S. McEwen, B.W. Hobson, L. McBride, and J.G. Bellingham: Docking control system for a 54-cm-diameter (21-in) AUV, *IEEE Journal of Oceanic Engineering*, Vol. 33, No. 4, pp. 550–562, 2008.

[4] N. Palomeras, A. Nalver, M. Massot-Campos, G. Vallicrosa, P.L. Negre, J.J. Fernndez, P. Ridao, P.J. Sanz, G. Oliver-Codina, and A. Palomer: I-AUV docking and intervention in a subsea panel, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL., pp. 2279–2285, 2014.

[5] S. Cowen, S. Briest, and J. Dombrowski: Underwater docking of autonomous undersea vehicle using optical terminal guidance, *Proc. IEEE Oceans Engineering*, Vol. 2, pp. 1143–1147, 1997.

[6] M.D. Feezor, F.Y. Sorrell, P.R. Blankinship, and J.G. Bellingham: Autonomous underwater vehicle homing/docking via electromagnetic guidance, *IEEE Journal of Oceans Engineering*, Vol. 26, No. 4, pp. 515–521, October 2001.

[7] J.Y. Park, B.H. Jun, P.M. Lee, F. Youb Lee, and J.H. Oh: Experiment on underwater docking of an autonomous underwater vehicle isimi using optical terminal guidance, *OCEANS 2007, Europe*, pp. 1–6, 2007.

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

[9] T. Ura, Y. Kurimoto, H. Kondo, Y. Nose, T. Sakamaki, and Y. Kuroda: Observation behavior of an AUV for ship wreck investigation, *Proc. MTS/IEEE OCEANS Conf.*, Vol. 3, pp. 2686–2691, 2005.

[10] N. Palomeras, P. Ridao, D. Ribas, and G. Vallicrosa: Autonomous I-AUV docking for fixed-base manipulation, *Proc. International Federation of Automatic Control*, Vol. 47, No. 3, pp. 12160–12165, 2014.

[11] M. Myint, K. Yonemori, A. Yanou, S. Ishiyama, and M. Minami: Robustness of visual-servo against air bubble disturbance of underwater vehicle system using three-dimensional marker and dual-eye cameras, *Proceedings of International Conference OCEANS 2015 MTS/IEEE*, Washington DC, USA, pp. 1–8, 2015.

[12] M. Myint, K. Yonemori, A. Yanou, K.N. Lwin, M. Minami, and S. 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.

[13] M. Myint, K. Yonemori, A. Yanou, M. Minami, and S. Ishiyama: Visual-servo-based autonomous docking system for underwater vehicle using dual-eyes camera 3D-pose tracking, *Proceedings of 2015 IEEE/SICE International Symposium on System Integration*, Nagoya, Japan, pp. 989–994, 2015.

[14] M. Myint, K. Yonemori, A. Yanou, K.N. Lwin, M. Minami, and S. Ishiyama: Visual-based deep sea docking simulation of underwater vehicle using dual-eyes cameras with lighting adaptation, *Proceedings of International Conference OCEANS 2016 MTS/IEEE*, Shanghai, China, pp. 1–8, 2016.

[15] M. Myint, K. Yonemori, A. Yanou, K.N. Lwin, N. Mukada, and M. Minami: Dual-eyes visual-based sea docking for sea bottom battery recharging, *Proceedings of International Conference OCEANS 2016 MTS/IEEE*, Monterey, USA, pp. 1–7, 2016.

[16] K.N. Lwin, K. Yonemori, M. Myint, A. Yanou, and M. Minami: Autonomous docking experiment in the sea for visual-servo type underwater vehicle using three-dimensional marker and dual-eyes cameras, *2016 55th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, pp. 1359–1365, 2016.

[17] K. Yonemori, A. Yanou, M. Myint, K.N. Lwin, and M. Minami: Docking experiment of underwater vehicle by dual-eye visual servoing in sea, *Transactions of the JSME*, Vol. 83, No. 848, DOI:10.1299/transjsme.16-00410, 2017.

[18] M. Myint, K. Yonemori, K.N. Lwin, A. Yanou, and M. Minami: Dual-eyes vision-based docking system for autonomous underwater vehicle: An approach and experiments, *Journal of Intelligent and Robot Systems*, Vol. 92, No. 1, pp. 159–186, 2018 (in Japanese).

[19] C. Chen and D. Schonfeld: Pose estimation from multiple cameras based on Sylvester's equation, *Computer Vision and Image Understanding*, Vol. 114, No. 6, pp. 652–666, 2010.

[20] R. Laganiere, S. Gilbert, and G. Roth: Robust object pose estimation from feature-based stereo, *IEEE Transactions on Instrumentation and Measurement*, Vol. 55, No. 4, pp. 1270–1280, 2006.

[21] W. Song, Y. Fujia, and M. Minami: 3-D visual servoing by feedforward evolutionary recognition, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, Vol. 4, No. 4, pp. 739–755, 2010.

[22] H. Suzuki and M. Minami: Visual servoing to catch fish using global/local GA search, *IEEE/ASME Transactions on Mechatronics*, Vol. 10, No. 3, pp. 352–357, 2005.

[23] M. Minami, J. Agbanhan, and T. Asakura: Evolutionary scene recognition and simultaneous position/orientation detection, *Soft Computing in Measurement and Information Acquisition*, pp. 178–207, Springer Berlin Heidelberg, 2003.

Khin Nwe Lwin

She received her M.E. degree in mechatronic engineering from Mandalay Technological University, Myanmar, in 2011. In 2012, she joined University of Technology (Yatanarpon Cyber City), Pyin Oo Lwin, Myanmar, where she is currently an Lecturer of the Department of Faculty of Precision Engineering. She is a Ph.D. student in Intelligent Robotics and Control, Okayama University. Her research interests include visual servoing, control of autonomous underwater vehicles, robotics, and genetic algorithms. She is a student member of IEEE.

Myo Myint

He received his M.E. and Ph.D. degrees in electronic engineering from Yangon Technological University, Mandalay Technological University, Myanmar, in 2006 and 2009, respectively. He received a Ph.D. degree in intelligent robotics and control from Okayama University, Japan in 2018. He worked as an Associate Professor in Yangon Technological University from 2011 to 2015. His research interests include autonomous underwater vehicles, robotics, electronics, and renewable energy.

Kenta Yonemori

He received his B.E. and M.E. degrees in intelligent robotics and control from Okayama University, Japan in 2014 and 2016, respectively. Since 2016, he has worked at Nabtesco Corporation, Japan. His research interests include 3D pose estimation, robotics, and optimization.

Naoki Mukada

He received his B.E. and M.E. degrees in intelligent robotics and control from Okayama University, Japan in 2016 and 2018, respectively. He has worked at DISCO Corporation, Japan. His research interests include 3D pose estimation, robotics, and optimization.

Yoshiki Kanda

He received his B.E. degree in intelligent robotics and control from Okayama University, Japan in 2018. His research interests include 3D pose estimation, robotics, and optimization. He is a member of JSME.

Akira Yanou (Member)

He received his Ph.D. degree in engineering from Okayama University, Japan in 2001. He worked at School of Engineering, Kinki University from 2002 to 2008, and Graduate School of Natural Science and Technology, Okayama University from 2009 to 2016. He joined Department of Radiological Technology, Kawasaki College of Allied Health Professions in 2016, and Faculty of Health Science and Technology, Kawasaki University of Medical Welfare in 2017, as an Associate Professor. His research interests include adaptive control, strong stability systems, and estimation. He is a member of the IEEE, JSME, ISCIE, and JSRT.

Mamoru Minami (Member)

He obtained his Ph.D. degree from Kanazawa University in 1990. He worked for University of Fukui from 1994 to 2002. He has worked as a Professor in the Department of Intelligent Mechanical System, Okayama University, Japan since 2010. His research interests include adaptive control, visual servoing, robotics, autonomous underwater vehicles, cloth recognition, and underwater docking systems. He is a member of JSME, IEEE, RSJ, and others.