

# Autonomous Docking Experiment in the Sea for Visual-servo Type Underwater Vehicle Using Three-Dimensional Marker and Dual-eyes Cameras

Khin Nwe Lwin<sup>1†</sup>, Kenta Yonemori<sup>1</sup>, Myo Myint<sup>1</sup>, Akira Yanou<sup>1</sup> and Mamoru Minami<sup>1</sup>

<sup>1</sup>Department of Intelligent Robotics and Control Laboratory, University of Okayama, Okayama, Japan  
(Tel: +81-86-251-8233; E-mail: minami-m@cc.okayama-u.ac.jp)

**Abstract:** Nowadays, underwater docking application plays an important role in sea development. This paper studies visual-servo type docking system for underwater vehicle through real-time pose (position and orientation) tracking with stereo vision. Real-time estimation of vehicle's relative pose to 3D marker whose shape and color is predetermined and known is executed by 3D model-based matching utilizing Genetic Algorithms (GA). In this paper, unidirectional docking station is designed and effective docking strategy is proposed for recharging application of batteries. P controller is applied for visual servoing. Docking experiments were conducted in the sea at Wakayama prefecture, Japan. According to the experimental results, it can be confirmed that docking performance using proposed system is achieved successfully with millimetre level accuracy in recognition and visual servoing.

**Keywords:** AUV, Docking, Visual servoing, 3D-MoS, Dual-eyes cameras

## 1. INTRODUCTION

In recent years, underwater vehicles play an important role in the ocean science. An automatic docking for recharging battery is regarded as one of the most essential abilities for reliable long-term operation of underwater vehicle. While in docking state, the underwater vehicle can perform activities such as battery recharging, data downloading and uploading. These implementations enable the underwater robot to stay for the duration in water and work.

A large number of docking have been conducted in the underwater domain by applying various sensors and guiding methods [1]-[9]. They include the utilization of fuzzy guidance system [1], self-similar landmark [2], optical terminal guidance [3], sonar and video-based [4] and electromagnetic guidance [5]. Each aiming autonomous docking measures relative pose continuously in time and enable the vehicle to respond changing environment conditions. The cost of most of the guiding system using sensor-based methods is high. Even though expensive navigation sensors and guiding methods are able to provide high accurate position data, the final docking process is still difficult problem.

In order to construct reliable and inexpensive docking system, we have been studying for the purpose of implementation of the automatic control of underwater robot using a visual servoing. As a typical image recognition method relating to underwater robot, some researches opt to adopt monocular camera to acquire the distance between the target with respect to the vehicle [6], [7]. This method tend to be robust to variations in camera model and is difficult to perform the precise docking when being close into the target. In this work, we utilized the stereo vision system to estimate the relative pose between the vehicle and the target object. In proposed stereo vision system, the two cameras are used for pose estimation while seeing the same target object at the same time. We

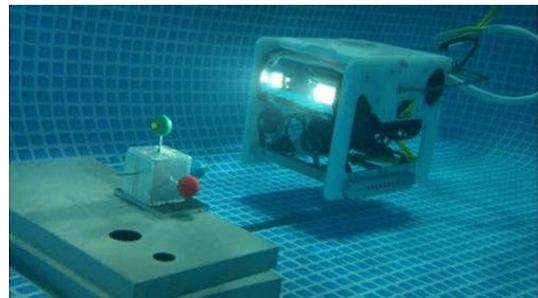


Fig. 1 Visual-Docking Underwater Vehicle

have developed visual-based underwater vehicle using dual-eyes cameras and 3-dimensional move on sensing (3D-MoS) as shown in Fig.1, which can perform automatically charging to extend the activity time duration in water. In this paper, real-time recognition of the relative pose of the target object utilizing model-based matching and Multi-step GA is proposed and we named the method as "Real-time Multi-step GA", since the method has been proved to be useful for real-time visual servoing purpose. Utilizing predefined information of shape and color of the 3D marker (shown in Fig.1), the underwater vehicle is regulated in the desired pose between the target and the vehicle by visual servoing. The proposed system of docking operation has been successful in previous work [8] at simulated pool with high homing accuracy. As a main contribution of this paper, we conducted docking experiment in the sea of Wakayama prefecture to evaluate how much our 3D-MoS system would be robust against natural sea environment. This paper shows the experimental result of docking performance.

Section 2 introduces the proposed docking system and explains the detail of the system. Section 3 describes the environment condition of the experiment. Underwater docking experiment results are reported in section 4 and conclude in section 5.

† Khin Nwe Lwin is the presenter of this paper.

## 2. PROPOSED DOCKING SYSTEM

To perform the docking operation, the proposed docking system includes four-stages:

1. **MANUAL OPERATION:** In this approach stage, the ROV approached by manually until the object is in the field of view of the dual-eyes cameras which is mounted in front of the robot. After approaching closer to the target object, the proposed system switches from manual to automatic control of visual servoing state.
2. **VISUAL SERVOING:** In this stage, the underwater robot detects the target object and relative pose of the vehicle through 3-D model-based recognition using multi-step GA.
3. **DOCKING (Fitting):** In docking state, the robot is performed the docking function utilizing a rod that is attached in ROV into the docking hole that is assembled with 3D marker in docking station. When the errors of tolerance of relative pose in image plane ( $x$  and  $y$  direction) is within  $\pm 20$  [mm] and the stable duration within desired relative pose error is more than 165 [ms], the rod is fitted into the docking hole by decreasing the desired value of  $x_d$ . However, if the docking process no longer meet the conditions of desired allowance error and stable condition, it return to the visual servoing.
4. **COMPLETION OF DOCKING:** In this stage, our mission is absolutely completed for docking. The robot is maintained constantly in final target pose for proposed recharging battery application.

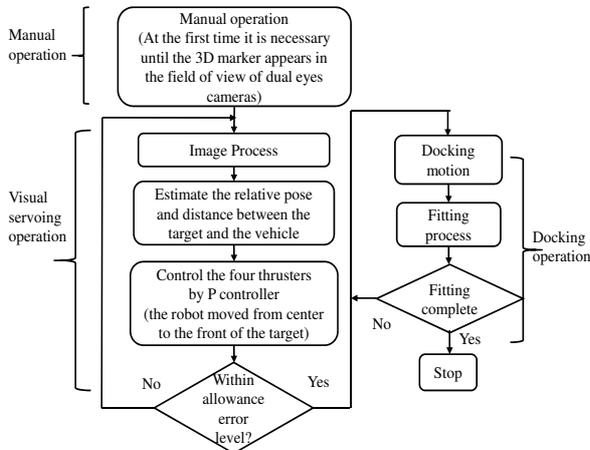


Fig. 2 Flowchart of the Docking System

### 2.1. Vision based Guidance for Underwater Vehicle

There are many guidance and control techniques to enable the underwater robot to perform useful tasks. Underwater robots required adequate guidance and control to perform these tasks. In this experiment, we used visual information to guide the robot's motion by using visual servo control. Visual servoing is the use of feedback information from camera to control the pose of the robot relative to the target object. Diagram of the visual servoing system is shown in Fig. 3. In this system, sequence of images are captured by the dual-eye cameras that is mounted on the underwater vehicle. The visual signal of left and right images are sent to the PC through the cable and then the relative pose of the vehicle is esti-

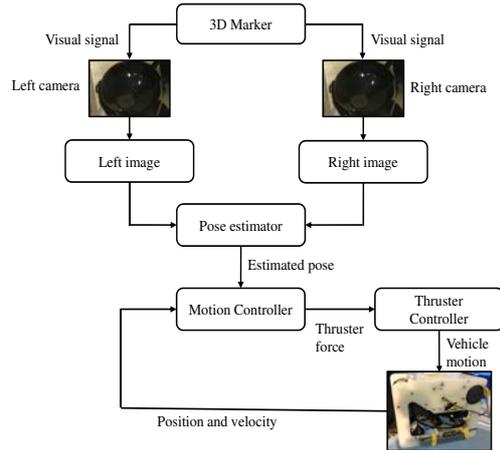


Fig. 3 Diagram of Visual Servoing System

mated by corresponding implemented software (pose estimator). The estimated pose of the current condition of underwater vehicle entered into the motion controller as the input signal to adjust the thrust force of the robot. By using feedback information of the P controller, we can eliminate the error between the estimated pose and the desired pose.

### 2.2. 3D Model-based Matching Method Using Dual-eyes Cameras and 3D Marker

For pose estimation, apart from other model-based recognition methods, we have developed 3D model-based matching method using dual-eyes cameras and 3D marker. This contribution was confirmed in previous works [13]. In this method, the target object named as 3D-marker that consists of three spheres (40[mm] diameter) whose colors are red, green and blue is designed. The position and orientation of the three-dimensional solid model is determined by six variable ( $x, y, z, \epsilon_1, \epsilon_2, \epsilon_3$ ), where ( $x, y, z$ ) are position in Cartesian coordinate system and ( $\epsilon_1, \epsilon_2, \epsilon_3$ ) are the orientation in quaternion avoiding singularity assume, where the pose is defined based on  $\Sigma_H$  in Fig4 and Fig.6. Fig. 4 shows the 3D model-based matching system with dual-eyes cameras and 3D-marker. In this figure,  $\Sigma_{CR}$  and  $\Sigma_{CL}$  are the reference coordinate frame of the right camera and left camera.  $\Sigma_H$  is the reference frame of the ROV.  $\Sigma_M$  is the reference frame

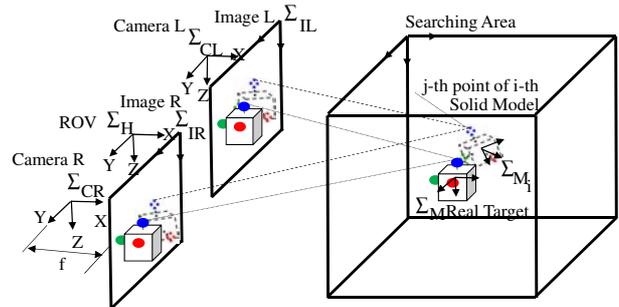


Fig. 4 3D Model-based Matching with Stereo Vision System

of the real target object. The solid model of the real target object in space is projected naturally to the dual-eyes cameras images and the dotted 3D marker model where pose is given by one of GA's genes is projected from 3D to 2D. We can calculate the correlation value between the projected real target 3D marker and the projected model. The correlation function that considers 3D marker shape and color is used for fitness function in GA process. Multiple models are initially located randomly within the search area that are the same information of 3D marker (color, shape, size) with different pose and projected from 3D space to 2D image plane. Then, the projected model (six variables) is matched with the image captured by dual eyes cameras (six variables). Finally, the best model that is generated by collecting the maximum fitness value is selected to represent truthful relative pose.

### 2.3. Fitness Function

Fitness function is very important to obtain the optimum recognition accuracy. In each population, the recognition accuracy of individual model is based on the fitness function. The good fitness function will support the optimum searching method effectively and efficiently. In this experiment, the following fitness function is represented the correlation between the searching model and the image input from the left and right video camera. In this function,  $F_R(\varphi)$  and  $F_L(\varphi)$  are the fitness function of the right camera and left camera image. Detail explanation about fitness function can be seen in previous work [11]. Fig. 5 shows the recognition regulation of the target object and the model.

$$F_R(\varphi) = \sum_{IR_{ri} \in F_{R,in}(\varphi)} \rho(IR_{ri}) - \sum_{IR_{ri} \in F_{R,out}(\varphi)} \rho(IR_{ri}) \quad (1)$$

$$F_L(\varphi) = \sum_{IL_{ri} \in F_{L,in}(\varphi)} \rho(IL_{ri}) - \sum_{IL_{ri} \in F_{L,out}(\varphi)} \rho(IL_{ri}) \quad (2)$$

$$F(\varphi) = (F_R(\varphi) + F_L(\varphi))/2 \quad (3)$$

There are two portions as shown in Fig. 5, the first portion is the inner one that is the same size with the target area and the second one is outer portion that is the background area. When the captured portion of the right camera image situated in the inner one, the fitness value will increase and it situated in the background area, the fitness value will decrease. Similarly, the good fitness value of the left camera image is calculated. The fitness value will maximum when the target and the model are identical. Finally, the optimum fitness value  $F(\varphi)$  is evaluated by

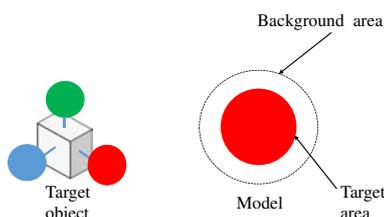
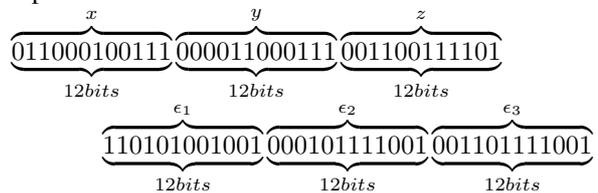


Fig. 5 Target Object and Model

adding the average value of the left and right image at the same time.

### 2.4. Optimization Method Using Real-time Multi-step GA

The proposed system can recognize the vehicle's pose through the (3D-Mos) by utilizing multi-step GA and model-based matching method. By using the fitness function described in the previous section, the problem of searching for the pose of the object can be replaced by optimum searching method for real time. GA is the optimization method to obtain the optimum solution in a short searching time. Position and orientation of the three-dimensional model (3D-Mos) is represented as individual of chromosome. The upper 36 bits of one chromosome represents the position coordinate of a three-dimensional model and the next 36 bits represent the orientation of the three-dimensional model.



Genetic algorithm searches a solution starting with

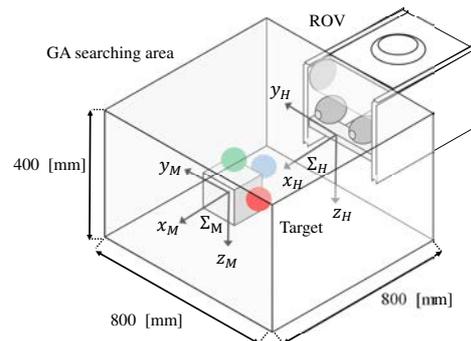


Fig. 6 GA Searching Area

all possible models which is called population. These possible models reproduce new models (new generation) by selection and recombination method which represents a better solution to the real target pose estimation. In each generation, every individual model is evaluated using the fitness function value. Then, again new models are formed from mutation or crossover. The procedure is performed continuously until 33 [ms] that is video rate. Multi-step GA is means the capable of real time recognition of the target object. Termination of GA generation is defined by video rate 33 [ms] in this work. Therefore, GA generation with 9 times evolution will perform for every image with video rate of 33 [ms]. Detection of target is defined by fitness value that is 0.6 in this work. It means that if the pose represented by the highest fitness value after 9 times evolution for one image is less than 0.6, this pose will not be used in feedback system telling that the target may not be in the field of view. Approaching step in which vehicle will approach the target to have the vehicle to be in field of view will be future work using other navigation systems. Therefore, assumption in this work is that the target exit

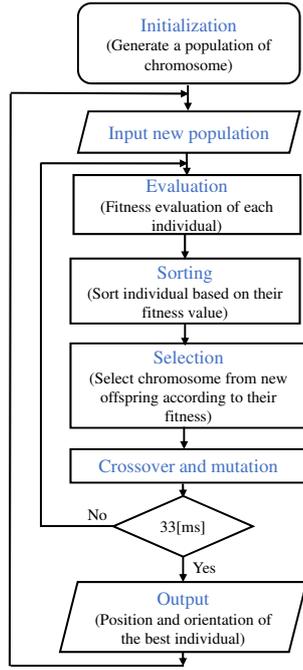


Fig. 7 Flowchart of Multi-step GA  
Table 1 Parameters of GA

Number of genes	60
Search area [mm]	$\{x,y,z\}=\{\pm 400, \pm 400, \pm 200\}$
Selection rate [%]	60
Mutation rate [%]	10
Crossover	2 point crossover
Evolution number of generations[Times/33ms]	9

in search space of GA as shown in Fig.6. Even though there are many powerful optimization methods, we selected GA and modified as Real-time Multi-step GA because of its simplicity and especially effectiveness in real-time performance. There is no story about powerful optimization that applied properly in real-time domain. Since our strategy is based on the thinking way that simple optimization method with high repetition is more effective than sophisticated method with low repetition and much calculation time when the calculation time used for the optimization is limited, we have been making effort to make the repetition number of simple GA increase so far. Then, detailed explanation about GA method and fitness function is referred to our previous paper[11]. Fig7. shows the flow chart of multi-step GA and indicated the parameters are shown in Table.1.

### 3. EXPERIMENT ENVIRONMENT

#### 3.1. Remotely Operated Vehicle

The ROV was designed and fabricated by KOWA cooperation as shown in Fig. 1. In this robot system, the eye visual sensor is used as a main sensor. There are totally four cameras (imaging element CCD, 380,000 pixel, signal system NTSC, Minimum Illumination 0.8[1X],

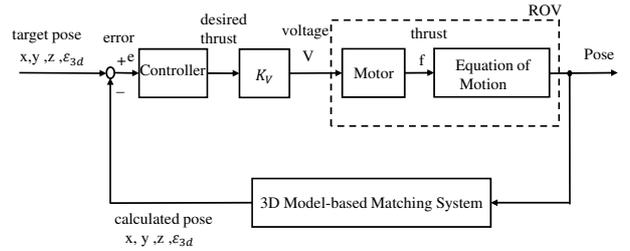


Fig. 8 Block Diagram of Visual Servo Control

without zoom) are mounted in this model and the two front cameras are used to perform a three-dimensional object recognition in visual servoing. Four thrusters are used in system, maximum thrust force is (9.8[N]) in horizontal and maximum vertical thrust force is (4.9[N]). In addition, the LED lights (5.8W) has been equipped for illumination ensure. The ROV obtains the camera image information and control signal from the PC through a tether cable (200[m]). It can operate in maximum water depth 50[m].

#### 3.2. Controller

Proportional controller is used to control the vehicle. The four thrusters that are mounted on the underwater robot are controlled by sending the command voltage based on the feedback relative pose between the underwater robot and the object ( $x_d$ [mm],  $y_d$ [mm],  $z_d$ [mm]). The block diagram of the control system is shown in Fig. 8. The control voltage of the four thrusters are controlled as the following equations.

$$\text{The depth direction} : v_1 = k_{p1}(x_d - x) + 2.5 \quad (4)$$

$$\text{Vertical axis rotation} : v_2 = k_{p2}(\epsilon_{3d} - \epsilon_3) + 2.5 \quad (5)$$

$$\text{Vertical direction} : v_3 = k_{p3}(z_d - z) + 2.5 \quad (6)$$

$$\text{Horizontal direction} : v_4 = k_{p4}(y_d - y) + 2.5 \quad (7)$$

Where  $v_1$ ,  $v_3$  and  $v_4$  are the control voltage of the four thrusters of x, z, y direction respectively.  $x_d$ ,  $y_d$ ,  $z_d$  are the desired relative pose between the vehicle and the target.  $\epsilon_{3d}$  is the rotation direction around the z-axis and it is expressed as the value of  $v_2$ . According to the experimental result, the gain coefficient is adjusted to perform the best condition for visual servoing.

#### 3.3. Structure of Docking Experiment

In our previous research, the autonomous docking system was conducted successfully in simulated pool. To confirm that docking performance using proposed system is achieved successfully in real sea, this experiment was conducted in the sea within Wakayama Prefecture, Japan, where the water depth is 3.5 [m] and there were some gentle waves. The buoyancy force was nearly 1.03 times the one of fresh water. In order to respond the environmental change in freshwater and sea water, we need to adjust the buoyancy balance of the ROV to ensure for the stable. Docking station was designed as shown in Fig. 9. The ROV and PC is connected with the tethered cable which has the length of 200 [mm]. Generally, the various sensors are used for long navigation until the target

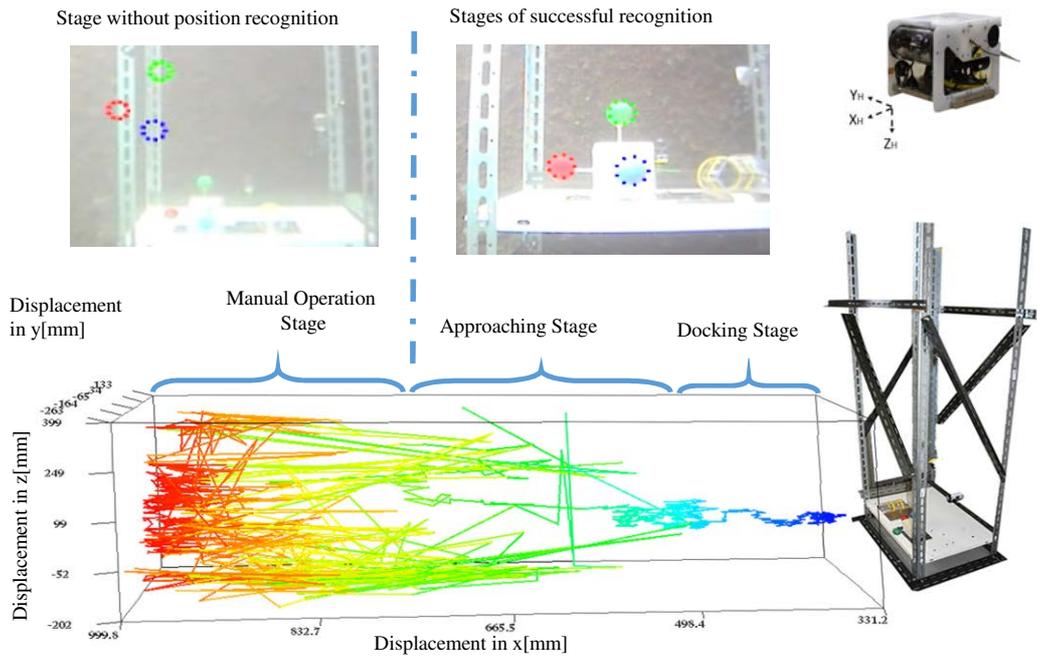


Fig. 11 3D Tracking chart of AUV

is in the field of view of camera in ROV. In this system, the manual operation is necessary to appear the image of the target object in the field of view of ROV's cameras due to the searching area is limited. The coordinate sys-

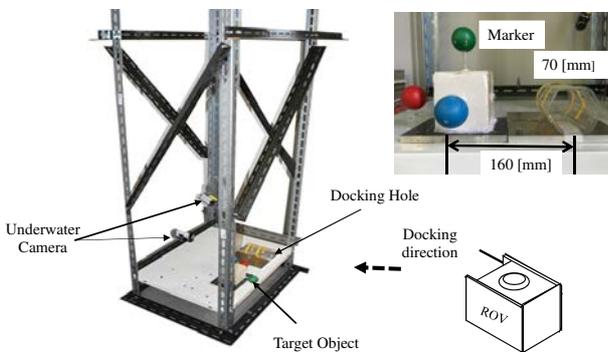


Fig. 9 Structure of Experiment Layout

tem of the underwater robot is shown in Fig. 10. The target object and the docking hole is fixed in docking station. The diameter of the docking hole is 70 [mm] and the center distance between the marker and the docking hole is 160 [mm]. The structure of docking experiment and the two underwater cameras are shown in Fig. 9. Regarding to computation speed of the system with respect to the speed of movement of ROV with maximum speed about 0.2 m/s, we selected the best GA parameter to converge the solution in real-time based on PC performance. According to experimental results, we confirmed that the system can perform real-time performance not only in pose estimation but also in controlling the vehicle.

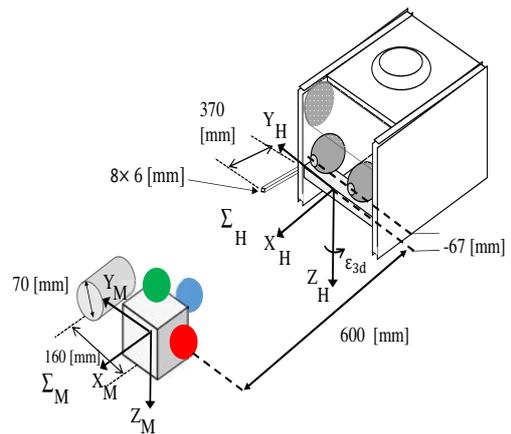


Fig. 10 Coordinate System of Docking Experiment

#### 4. UNDERWATER DOCKING EXPERIMENT

We conducted the underwater docking experiment to confirm the proposed system achieved successfully in real sea. The four operating states were conducted to perform the docking operation. These are Manual operation (Approach), Visual servoing (to keep the relative pose), Docking (fitting to the fixed docking hole) and Completion of docking (fully fitting into the docking station). The results of docking experiments are shown in Figs. (12-15). In each figure, Fig. (a) shows the fitness value recognized by multi-step GA with the time relationship and Fig. (b), (c) and (d) show the error of position in different axial during the docking experiment.

Fig. 11 shows the tracking chart of AUV. In this figure, the docking trajectory is simulated based on the pose which was estimated by multi-step GA. Figs. (16-19) is a sequence of continuously images taken by two underwa-

ter cameras. (a) Manual operation (ROV start), (b) Visual servoing (Automatic control), (c) Docking (fitting to the homing unit) and (d) Completion of docking (fully fitting into the homing unit).

In the first state of manual operation, the robot moved from the start point to the nearby region of the marker as can be seen in Fig. 12. During this state, the significant continuous oscillation occurs until 30 [s] because this duration is for long navigation of initial state. In this experiment, GA recognition accuracy needed to have fitness value of 0.5 or more for detection of 3D marker. In Fig. 16 (a), it can be seen that the vehicle did not correctly recognize the object at initial state meaning pose estimation during manual operation is not truthful and not applied in feedback control.

After finished manual state, the underwater robot goes forward and transit to a visual servoing by switching the automatic control from manual operation after 32 [s]. In this stage, the underwater robot moved to the desired relative target pose. When the range of error of the relative target position in image plane ( $y_d$  and  $z_d$ ) is within  $\pm 20$  [mm] and the stable duration within the desired relative pose error is more than 165 [ms], the rod is fitted into the docking hole by decreasing the distance in x-axis with the velocity of 30 [mm/s] until the desired value of ( $x_d$ ) is 350 [mm]. This is the proceeding state for the docking to perform the fitting to the docking station. It can be seen from Figs. (14-15) show the position in y-axis and z-axis direction. The robot maintained to perform the docking operation in the area between the dash line ( $-10 \leq y \leq 30$  and  $-80 \leq z \leq -40$ ). In x-axis direction in Fig. 13, it can be confirmed that the robot performed the docking operation after 35 [s]. However, if the underwater robot did not performed precisely fitting and right condition, it will keeps visual servoing process for executing the docking process again. The robot moved to reach ( $x_d = 600$ [mm],  $y_d = 0$ ,  $z_d = -67$ [mm],  $\epsilon_{3d} = 0$ ) by performing visual servoing in docking process.

In completion of docking state, the robot maintained constantly at desired relative pose between the vehicle and the object by visual servoing for application. In Fig. (15), it can be confirmed that the underwater vehicle have completed the docking within 40[s]. It can be said that successful underwater docking in real sea was conducted using proposed system.

## 5. CONCLUSION

Experiments of underwater docking were conducted the development of proposed system for underwater recharging application. In this paper, we proposed visual servo based underwater vehicle through real time pose docking with stereo vision system. The performance of real time 3-dimensional recognition of the relative pose of the 3D marker using multi-step GA was confirmed to be able to achieve recognition with an pose error of docking process is less than  $\pm 5$ [mm]. According to the experimental result, it can be confirmed that docking performance using proposed system is achieved successfully

with centimetre level accuracy in recognition and visual servoing.

## REFERENCES

- [1] Ken Teo , Benjamin Goh , and Oh Kwee Chai, "Fuzzy Docking Guidance Using Augmented Navigation System on an AUV", *IEEE Journal Of Oceanic Engineering*, Vol. 40, No. 2, April 2015.
- [2] Amaury N' egre, C edric Pradalier, Matthew Dunbabin, "Robust vision-based underwater homing using self-similar landmarks", *Journal of Field Robotics, Wiley-Blackwell, Special Issue on Field and Service Robotics*, 25 (6-7), pp. 360-377, October, 2008.
- [3] Jin-Yeong Park, Bong-Huan Jun, Pan-Mook Lee, Fill-Youb Lee and Jun-ho Oh, "Experiment on Underwater Docking of an Autonomous Underwater Vehicle 'ISiMI' using Optical Terminal Guidance" *Oceanic, Europe*, pp. 1-6, 2007.
- [4] Jonathan Evans, Paul Redmond, Costas Plakas, Kelvin Hamilton and David Lane, "Autonomous Docking for Intervention-AUVs using Sonar and Video-based Real-time 3D Pose Estimation", *Oceans*, Vol. 4, pp. 2201-2210, 2003.
- [5] Michael D. Feezor, Member, IEEE, F. Yates Sorrell, Paul R. Blankinship, and James G. Bellingham, "Autonomous Underwater Vehicle Homing/Docking via Electromagnetic Guidance", *IEEE Journal Of Oceanic Engineering*, Vol. 26, No. 4, October 2001.
- [6] Jin-Yeong Park, Bong-huan Jun, Pan-mook Lee, Junho Oh, "Experiments on Vision Guided Docking of an Autonomous Underwater Vehicle Using One Camera", *IEEE Ocean Engineering* 36, 48-61, 2009.
- [7] Kim, J., "Dual Control Approach for Automatic Docking using Monocular Vision", PhD dissertation, Stanford University, Stanford, CA, 2007.
- [8] Myo Myint, Kenta Yonemori, Akira Yanou, Mamoru Minami, Shintaro Ishiyama, "Visual-servo-based Autonomous Docking System for Underwater Vehicle Using Dual-eyes Camera 3D-Pose Tracking", *2015 IEEE/SICE International Symposium on System Integration (SII)*, Meijo University, Nagoya, Japan, December 11-13, 2015.
- [9] Palomeras, N., Ridao, P., Ribas, D. and Vallcrosa, G. "Autonomous I-AUV Docking for Fixed-base Manipulation", *Preprint of the International Federation of Automatic Control*, pp. 12160-12165, 2014.
- [10] Myo Myint, Kenta Yonemori, Akira Yanou, Shintaro Ishiyama and Mamoru Minami, "Robustness of Visual-Servo against Air Bubble Disturbance of Underwater Vehicle Using Three-dimensional Marker and Dual-eye Cameras", *International Conference OCEANS 15 MTS/IEEE, Washington DC, USA*, October 19-22, 2015.
- [11] Wei Song and Minami M., "On-line motion-feedforward pose recognition invariant for dynamic hand-eye motion", *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp.1047-1052, 2008.

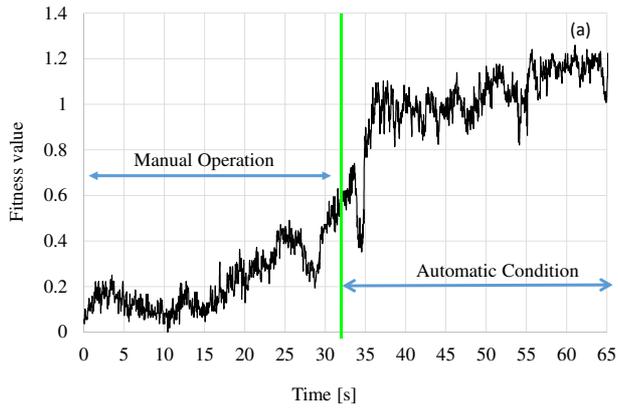


Fig. 12 Fitness value

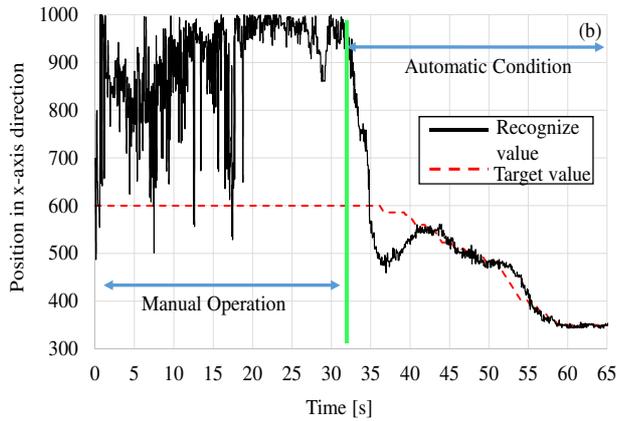


Fig. 13 Error in x-axis direction

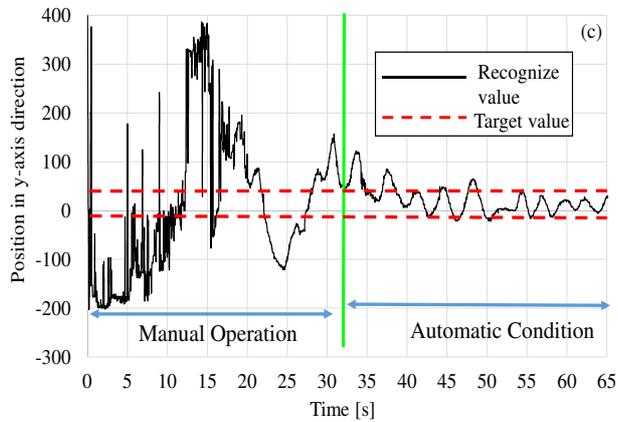


Fig. 14 Error in y-axis direction

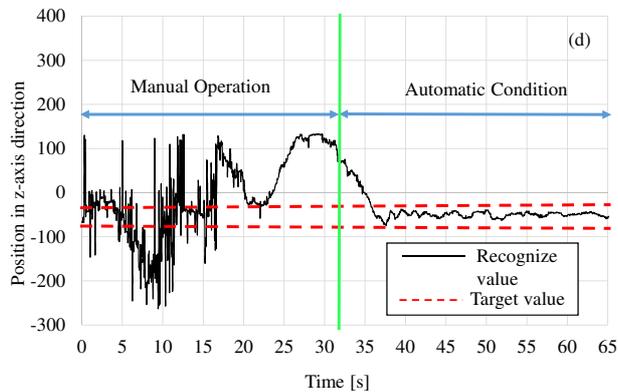


Fig. 15 Error in z-axis direction

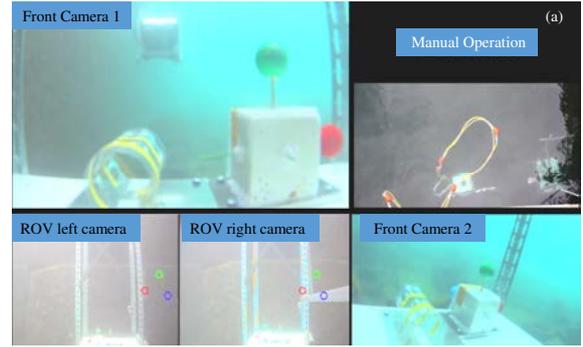


Fig. 16 Snapshot of manual Operation

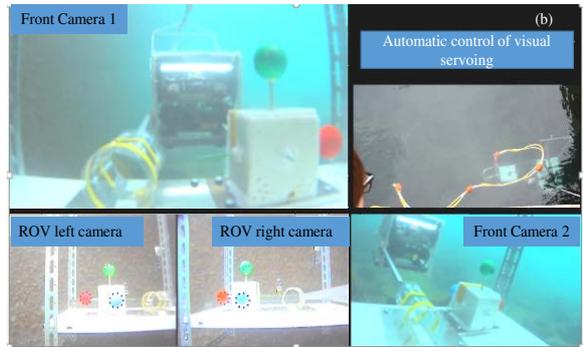


Fig. 17 Snapshot of automatic control of visual servoing

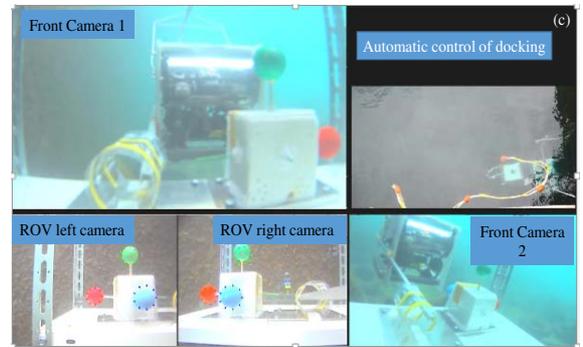


Fig. 18 Snapshot of automatic control of docking

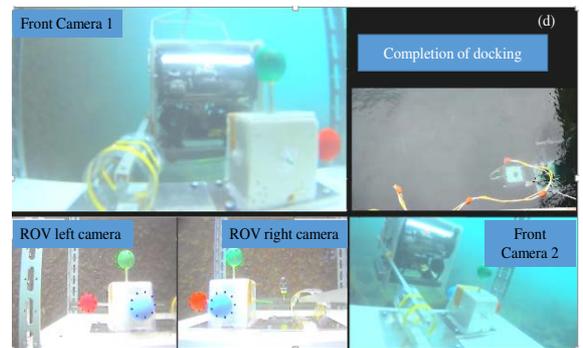


Fig. 19 Snapshot of completion of docking