

Docking Performance Against Turbidity Using an Active Marker Under Changing Lighting Environment

○非 Khin Nwe Lwin (岡山大) 非 Myo Myint (岡山大)
非 向田 直樹 (岡山大) 非 山田 大喜 (岡山大)
正 松野 隆幸 (岡山大) 正 矢納 陽 (川崎短大)
正 見浪 護 (岡山大)

Khin Nwe Lwin, Okayama University, pdoj8yez@s.okayama-u.ac.jp
Myo Myint, Okayama University
Naoki Mukada, Okayama University
Akira Yanou, Kawasaki University of Medical Welfare
Daiki Yamada, Okayama University
Takayuki Matsuno, Okayama University
Mamoru Minami, Okayama University

Vision-based underwater exploration and exploitation require a robust computer vision system, particularly for operation in turbid water that may govern the poor visibility of the underwater environment. Additionally, visual servoing is largely dependent on not only turbidity level but also lighting changing in the real sea environment. This paper presents the turbidity tolerance of the proposed dual-eye based docking system using an active 3D marker under changing lighting condition. Pool docking experiment was conducted against turbidity using an ROV. In proposed system, Real-time Multi-Step GA (RM-GA) and model-based matching method are applied for real time 3D pose estimation. The experimental results have confirmed the docking performance of the proposed system against turbidity under different lighting conditions.

Key Words: Active marker, Turbidity, Docking, Lighting, Dual-eye cameras

1 Introduction

The visual servoing based underwater vehicles have been developed for many applications such as inspection, repair oil and gas, pipeline tracking, docking task, scientific studies of the deep ocean, etc [1], [2] all over the world in recent year. Some studies are based on single eye camera to estimate the pose of the target object [3]-[5]. A binocular vision was used in some of these studies to estimate the relative pose of the target object [6], [7]. Even though two cameras were used in [6], one was facing downward for shooting the sea-floor images and the second camera was pointing forward for the purpose of obstacle avoidance. In [7], Girona 500 AVU was developed for inspection and intervention tasks for the seafloor survey. In that work, several sensors such as sonar, GPS, pressure sensor, velocity, fiber optic gyro, and two video cameras are used to localize an object. But the two cameras look at different targets, then these approaches do not materialize parallaxic nature.

In the real sea areas, investigating using visual servoing inevitably has the difficulties in recognition of object when detected scene images are disturbed by lighting effect, water current and turbidity, etc. Therefore, the technologies for observing underwater situations correctly and robustly from cameras of the visual servoing systems are needed.

The role of turbidity should be considered in the underwater environment because it can degrade the visual quality. In [8], [9], two cameras and three cameras are used to increase the image recognition robustness by analyzing degradation factors in turbid water. But the discussion about turbidity is not enough in these references. They evaluated the multiple features of detectors and compared the performance of detec-

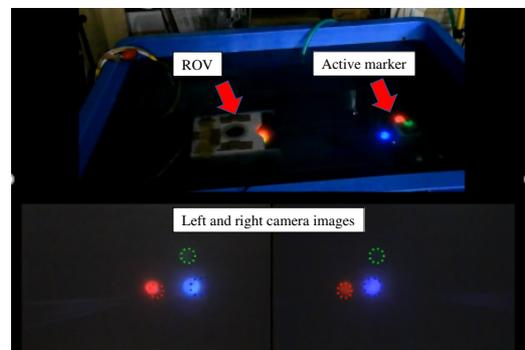


Fig.1 Docking experiment in turbid water using an active marker under changing lighting environment.

tors on images degraded by turbidity. They did not perform the pose estimation. According to the authors' knowledge, there is no study on turbidity tolerance of 3D pose estimation for real-time vision-based underwater vehicle.

The authors have proposed a new 3D pose estimation method with dual-eye cameras that exploits the parallaxic nature that enables reliable 3D pose estimation in real-time. Visual servoing using stereo vision and parallaxic character for the underwater vehicle utilizing 3D model-based recognition and Real-time Multi-step Genetic Algorithm (RM-GA) has been developed, and have confirmed the effectiveness of the proposed system in [10]-[12]. The dual-eye cameras based on perception means solving the corresponding points problem. If the corresponding points in the real object are not con-

nected with the corresponding points in images during 2D to 3D reconstruction, the true 3D object cannot be represented because of the wrong reconstructed 3D points. Instead, the point is projected from 3D real target object into the both 2D images correctly using the forward projection from 3D to 2D generates unique points in 2D images without errors. Based on this point, 3D model-based recognition is implemented. Furthermore, solid object that is represented a group of points on the sphere of the 3D marker is projected rather than the individual pixel.

In previous works, the robustness of the proposed system was confirmed using passive marker by conducting the different experiments [10]-[12]. In this study, the active marker was developed to improve the 3D pose estimation in high turbidity. Illumination variation is one of the challenging problems to be solved for robust recognition systems in unknown environment. Turbidity and changing illumination can disturb the recognition of the target object in real sea area. Additionally, this study presents the docking performance using an active marker under the condition whereby the turbidity and changing illumination degrade the visual quality in order to confirm the performance of the proposed system.

The remainder of the paper is organized as follows: Section 2 describes the method of 3D pose estimation. Experiment results are reported in section 3 with discussion and conclude in section 4.

2 3D Pose Estimation Method

2.1 Model-based Matching Method

In proposed system, the model-based matching method is used to estimate the relative pose between the vehicle and a known 3D marker. In another conventional methods, the pose estimation method is implemented using the feature-based recognition based on 2D to 3D reconstruction. In that approach, the set of image points in different images is used to determine the information of the target object. The main drawback is complex for searching the corresponding points and time taken. Apart from this, the model-based pose estimation approach based on 3D to 2D projection is applied in this work avoiding the effects of wrong mapping points in images using dual-eye cameras.

Figure 2 shows the model-based matching method using dual-eye cameras for 3D pose estimation and Fig. 3 shows the GA searching area. In Fig. 2, Σ_{CR} and Σ_{CL} are the reference coordinate frame of the right camera and the left camera. Σ_H is the reference frame of the ROV. Σ_M is the reference frame of the real target object. The solid model of the real target object in space is projected naturally to the dual-eyes cameras images and the dotted 3D marker model, where the pose is given by one of GA's genes, is projected from 3D to 2D. The different relative pose is calculated by comparing the projected 2D image and the solid model captured by the dualeye cameras. Finally, the best model of the target object that represents the true pose can be obtained based on its highest fitness value. There are some works done on visual-servoing experiments concerning hand-eye manipulator in the air using 3D model-based matching method utilizing genetic algorithms and dual-eyes camera [13], [14], which are used as fundamental knowledge for this research.

2.2 Fitness Function

The fitness function is constructed to evaluate the matching degree between the projected model and the captured image. The optimum searching and the GA convergence speed were

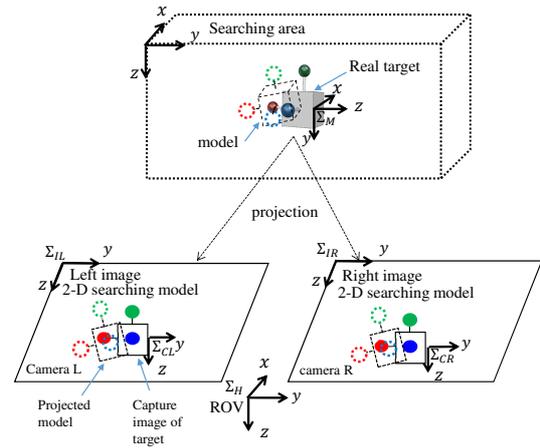


Fig.2 Model-based matching method using dual-eye cameras and 3D marker

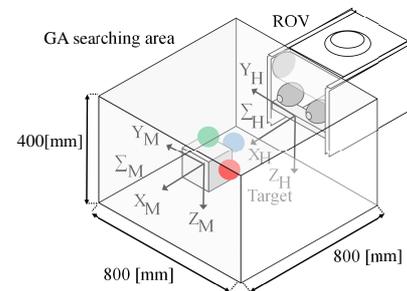


Fig.3 GA search space.

effected by the construction of the fitness function. The calculation of fitness value of each individual model can be done repeatedly in genetic algorithm. Figure 4 shows the real target and model of 3D marker. Only hue value is used for recognition of 3D marker because of less sensitive to the environment. Each model consists of three spherical balls (red, green and blue). Each spherical ball consists of two portions, where the inner portion is the same size as the real target object and the outer area is the background area. The dots in each ball mean points to calculate the correlation degree on how much the inner area overlaps the target object and the outer area does not overlap the target object. The captured image (pixel) is detected in 2D image as (green or blue or red) in hue space. If the captured image (pixel) situated in inner portion, the fitness value will be increased and the captured image (pixel) is situated in the outer portion, the fitness value will be decreased. Therefore, the fitness value will be maximum when the model and the real target are identical. Finally, the pose of the model with the maximum fitness value is thought to represent the pose of the real target 3D marker. Detailed explanation about the fitness function is referred to our previous paper [15].

2.3 Real-time Multi-step GA

The genetic algorithm is used as a search and an optimization method to estimate the relative pose between the ROV and 3D marker. Real-time multi-step GA means the capable of real time recognition of the true pose of the target object within 33 ms. Figure 5 shows an individual of GA population and Fig. 6 shows the flowchart of the RM-GA. Position and orientation of the three-dimensional model are represented as

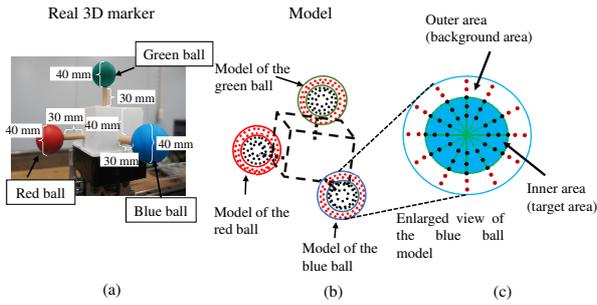


Fig.4 Real 3D marker and model: (a) real 3D marker, (b) model, (c) enlarged view of the blue ball model, where the inner area is the same size as the real target object (blue ball) and the outer area is the background area. The dots in (c) mean points to calculate the correlation degree on how much the inner area overlaps the blue ball and the outer area does not overlap the blue ball.

individual of the chromosome. The former 36 bits represent the position of the 3D marker and the later 36 bits describe the orientation defined by a quaternion.

Firstly, a random population of the chromosome is generated. A new pair of left and right images is input. The RM-GA procedure is performed within 33 ms. The real-time multi-step GA evolves the chromosomes with as many generations as possible within the video frame rate for each image. In the present study, the number of iterations of the RM-GA is chosen to be nine, which is a maximum that the computer used in the present study could calculate within 33 ms (determined by the video frame rate) during the GA evolution process. The RM-GA find repeatedly the solutions to get the best pose of the target object within the video frame rate to deal with time varying distribution for newly input images. The fitness function is designed to get the maximum value when the model and the real target exactly coincide. The true pose of the target object is expressed with the peak of the mountain shape in the fitness distribution. Instead of catching the peak point with a long time, the RM-GA find the true pose of the target object using the group point within the short time. The RM-GA quickly evaluated each gene in the fitness distribution by changing the group of point for finding the optimum solution to represent the true pose of the target object with the peak of the mountain shape. Finally, the best pose of the individual can be made to approach the real target's pose. Although the pose of the target object is evaluated in 2D, convergence occurs in 3D. For the next input, a new video image is used. The convergence performance to an optimum value of GA's evolution function used as fitness function has been proved mathematically by a Lyapunov analysis in a previous work [16]. The effectiveness of the GA was demonstrated in a previous study on visual servoing for catching fish using a GA search [17].

2.4 Controller

The 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. 7. The control voltage of the

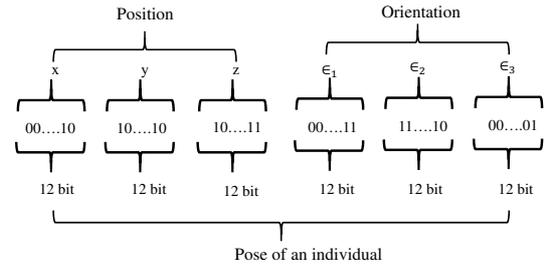


Fig.5 An individual of GA population: 12 bits for each x , y , z represents the position coordinate of the three dimensional model of the gene and 12 bits for each ϵ_1 , ϵ_2 , ϵ_3 describes the orientation defined by a quaternion.

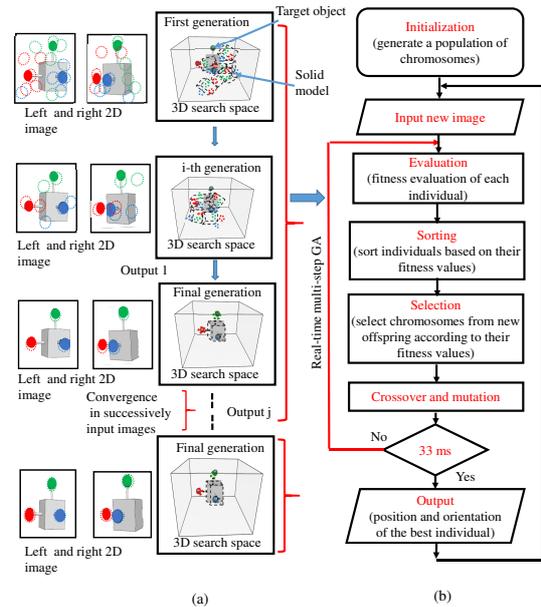


Fig.6 Flowchart of the real-time multi-step GA: (a) the recognition process of the target pose is evaluated in 2D, convergence occurs in 3D (b) the flowchart of the RM-GA, the best solution is evaluated within 33 ms through the GA process.

four thrusters is controlled as the following equations.

$$\text{The depth direction} : v_1 = k_{p1}(x_d - x) + 2.5 \quad (1)$$

$$\text{Vertical axis rotation} : v_2 = k_{p2}(\epsilon_{3d} - \epsilon_3) + 2.5 \quad (2)$$

$$\text{Vertical direction} : v_3 = k_{p3}(z_d - z) + 2.5 \quad (3)$$

$$\text{Horizontal direction} : v_4 = k_{p4}(y_d - y) + 2.5 \quad (4)$$

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. ϵ_{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.

2.5 Active Marker

In our previous research, the passive marker was used to conduct the experiment. In the present study, the active

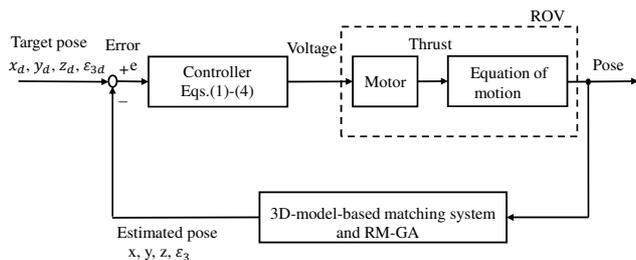


Fig.7 Block Diagram of Visual Servo Control.

marker was designed and constructed to improve the pose estimation at high turbidity level in day and night time. Figure 8 shows the appearance of the active marker. The circuit was created by combining the variable resistors, resistors, and the light emitting diodes such as red, green, and blue. The resistance value of the variable resistors, and the number of resistors are determined by trial and error. The 3D marker is constructed from a water proof box (100 mm × 100 mm × 100 mm) and the white spheres (diameter: 40 mm) are attached to the water proof box. The red, green and blue LED were installed into the white spherical ball and covered by the color balloon as shown in Fig. 8. This marker can be used as the passive marker when the light is switched off. The active marker allowed the ROV to recognize in day and night time by emitting the light of LED. The effectiveness of the active marker will be discussed in the next section.

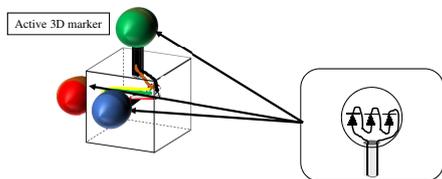


Fig.8 Active 3D marker: Red, green and blue LED were installed into the white spherical ball and covered by colour balloon.

3 Experimental Results and Discussion

3.1 3D Pose Estimation in Turbid Water

The 3D pose estimation was performed when the ROV and 3D marker was fixed in position at 600 mm under different turbidity levels in day and night time. The amount of turbidity is controlled by adding mud in water in the tank. Mud is chosen in order to simulate the natural condition. In this experiment, the turbidity level (Formazin Turbidity Unit, FTU) is measured by using a portable turbidity monitoring sensor TD-M500.

The ROV performed the visual servoing at about 600 mm in docking operation. It is the aware distance for docking operation to recognize the target object. Therefore, we give prominence to discuss 600 mm distance for recognition performance. Figure 9 shows the fitness value and turbidity using mud and the ROV and the 3D marker were fixed in distance 600 mm. The horizontal axis is described by the amount of

mud (ml/m^3) and the vertical axis is expressed in terms of fitness values and FTU values.

According to graphical results, the fitness value is decreased from 1.3 to 0.1 in the case of day time and from 0.6 to 0.1 in the case of night time when the turbidity is gradually increased from 0 FTU ($0 ml/m^3$) to 50.2 FTU ($375.875 ml/m^3$). The fitness value is nearly same at day and night time above 30 FTU. According to the experimental results, the performance of 3D pose estimation under different turbidity levels is analyzed and the maximum turbidity can be determined according to the defined threshold of fitness value.

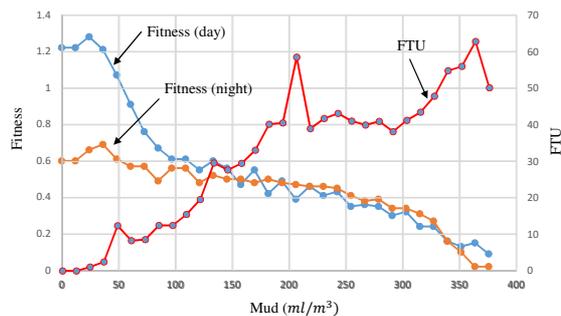


Fig.9 Fitness values against turbidity at the distance 600 mm between the ROV and 3D marker. The fitness value is nearly same in day and night time above 30 FTU.

3.2 Docking Performance Against Turbidity Under Changing Lighting Condition

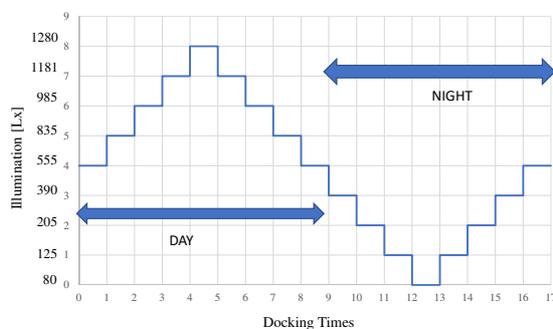


Fig.10 Illumination simulated for each docking time. Lighting condition is changed from day time to night time gradually.

This experiment was conducted in an indoor pool and the turbidity was created by adding mud (40 FTU). The desired pose ($x_d = 600$ mm, $y_d = 15$ mm, $z_d = -20$ mm, and $\epsilon_{3d} = 0$ deg) between the target and the ROV is predefined so that the ROV performs stationary hovering through visual servoing. The P controller is used to control the four thrusters of the vehicle in this docking experiment. The totally 17 times continuous docking was performed successfully by changing lighting from day time to night time as shown in Fig.10. Figure 10 shows the lighting simulation for each docking time. The horizontal direction is described by the number of docking times and the vertical direction is expressed with the illumination [Lx]. In lighting condition, we used the lighting

that can adjust from maximum to minimum condition. By adjusting the lighting condition, the maximum illumination is 1280 Lx in day time and minimum illumination is 80 Lx in night time.

The results of docking performance against turbidity at the maximum illumination 1280 Lx (day time) are shown in Fig. 11, and the results of docking performance at the minimum illumination 80 Lx are shown in Fig. 12. In Fig. 11 (a), the fitness value is above 0.8 for the few seconds of the recognition process and then increased to 1, which means that the system could recognize the 3D pose of the active marker well. Figures 11 (b), (c), (e), and (f) represent the relative pose between the desired pose and the estimated pose of the active marker recognized by RM-GA. Figure 11 (d) indicates the trajectory of the underwater robot based on Σ_H in Fig. 13 during the docking process.

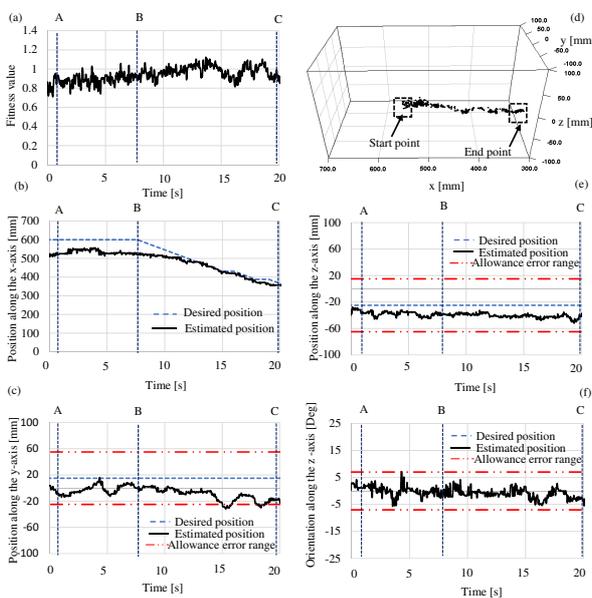


Fig.11 Docking performance against turbidity (40 FTU) in the case of mud using dual-eye images recognition at 1280 Lx (maximum lighting condition): (a) fitness value, (b) position along the x-axis, (c) position along the y-axis, (d) 3D trajectory of the underwater vehicle, (e) position along the z-axis, and (f) orientation along the z-axis.

In docking strategy, visual servoing starts when the 3D marker is detected, which means the fitness value is above a defined threshold (0.4 in the present study). When the pose of the vehicle is within the allowable error range of ± 40 mm of the desired pose, as shown in Figs. 11(b), (c), and (e), and the orientation around the z-axis (f) is controlled to within 7 deg for the desired period (165 ms, which is equal to five times the control loop period) in this experiment, docking starts by decreasing the distance between the ROV and the 3D marker from 550 mm to 350 mm, as shown in Fig. 11(b). The dotted line labeled A in each subfigure of Fig. 11 indicates the visual servoing state, where the desired position along the x-axis is 600 mm, and the desired position along the y-axis is within the allowance error range, as shown in Fig. 11(c). Visual servoing continues until the desired pose is within the error

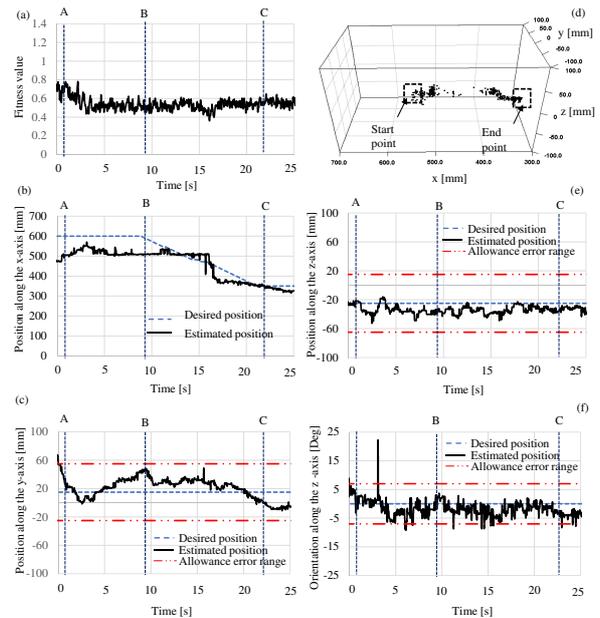


Fig.12 Docking performance against turbidity (40 FTU) in the case of mud using dual-eye images recognition at 80 Lx (minimum lighting condition): (a) fitness value, (b) position along the x-axis, (c) position along the y-axis, (d) 3D trajectory of the underwater vehicle, (e) position along the z-axis, and (f) orientation along the z-axis.

range for the y and z directions and the orientation around the z-axis, as shown in Figs. 11(c), (e), and (f). At time B, as shown in Figs. 11(b), (c), and (d), the docking criteria are satisfied and docking operation starts. Note that the position in the x direction at point B is approximately 500 mm because only the positions in the y and z directions and the orientation around the z-axis are considered in the docking criteria. The docking operation started approximately 7 s after starting the experiment. Finally, the docking operation was successfully completed approximately 20 s after starting the experiment. The dotted line labeled C in each sub-figure of Fig. 11 indicates the state whereby the docking is complete.

In the case of 80 Lx (night time), the fitness value is about 0.8 in recognition of active 3D marker at the start of the experiment and then decreased to about 0.5 as shown in Fig. 12(a). The ROV could recognize the active marker even though the environment is dark. The desired position along the orientation around the z-axis are out of error range at 3 s as shown in Fig.12 (f). Therefore, visual servoing continues until the desired pose of other direction y, z, and orientation around the z-axis is within allowance error range. The time for docking from the start of the experiment is 25 s in this case. The underwater robot was confirmed to maintain the desired pose while docking was performed under changing lighting condition at high turbidity, as shown in Figs. 11 and 12(a) through (f). According to the experimental results, even though the lighting condition was changed from day to night in high turbidity, the relative pose of the 3D marker can keep recognize well and the docking has been done successfully against turbidity under changing lighting condition.

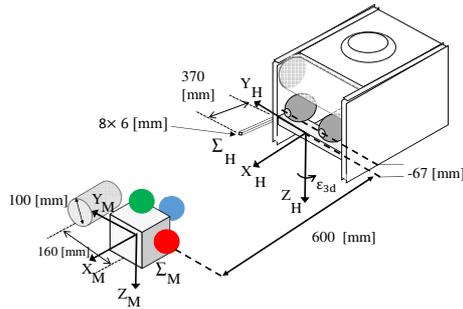


Fig.13 Docking coordinate system.

4 Conclusion

In the present study, visual-servoing based 3D pose estimation and docking performance against turbidity for underwater vehicle in changing lighting environment is presented. A real-time pose detection scheme was implemented by means of 3D model-based recognition and real-time multi-step GA using dual-eye cameras and active 3D marker. According to the experimental results, the proposed system using RM-GA can keep recognizing the relative pose of the 3D marker and can perform the docking operation against turbidity under changing lighting condition.

References

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

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

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

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

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

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

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

[8] Codevilla, F., Gaya, J. D. O., Duarte, N., & Botelho, S., "Achieving turbidity robustness on underwater images local feature detection." *International journal of computer vision*, 60(2), 91-110, 2004.

[9] Garcia, R., & Gracias, N., "Detection of interest points in turbid underwater images." In *OCEANS, 2011 IEEE-Spain*, pp.1-9, 2011.

[10] Myint, M., Yonemori, K., Yanou, A., Ishiyama, S., Minami, M., "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*, pp.1-8, Washington DC, USA 2015.

[11] Myint, M., Yonemori, K., Yanou, A., Lwin, K. N., Minami, M., Ishiyama, S., "Visual servoing for underwater vehicle using dual-eyes evolutionary real-time pose tracking." *Journal of Robotics and Mechatronics*, 28(4), pp.543-558, 2016.

[12] Myint, M., Yonemori, K., Yanou, A., Lwin, K.N., Minami, M. and Ishiyama, S., "Visual-based deep sea docking simulation of underwater vehicle using dual-eyes cameras with lighting adaptation." In *OCEANS 2016-Shanghai*, pp. 1-8, IEEE, 2016.

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

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

[15] Song, W., Minami, M., Aoyagi, S., "On-line stable evolutionary recognition based on unit quaternion representation by motion-feedforward compensation." *International Journal of Intelligent Computing in Medical Sciences & Image Processing*, 2(2), pp.127-139, 2008.

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

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