Stereo Vision-based 3D Pose Estimation under Turbid Water for Underwater Vehicles

*Myo Myint, Khin Nwe Lwin, Naoki Mukada, Matsuno Takayuki, and Mamoru Minami
(Okayama University)

1. Introduction

Nowadays, autonomous underwater vehicles (AUVs) are essential in applications such as inspection of underwater structures (e.g., dams and bridges) and underwater cable tracking [1]. Docking operation is very useful not only for battery recharging applications but also other applications such as sleeping under the mother ship, and a new mission downloading [2]. Therefore, there are many studies on underwater docking [2]-[5]. However, a number of challenging issues hinder these applications, which require high accuracy and robustness against disturbances that occur in the underwater environment. To achieve these tasks in underwater vehicles, we have developed a vision-based docking system using stereo vision.

In an underwater vehicle with a lighting unit installed on it, especially, dynamic lighting environment addresses challenges when the own lighting system is dominant in a deep sea or during a night operation. Additionally, when an underwater vehicle approaches the sea bottom, water turbidity comes in picture as disturbance to be considered and solved for visual servoing. According to the authors' knowledge, there are few studies on the 3D pose estimation under turbidity for underwater vehicles. In [6], detection of interest points in turbid underwater images was reported. Collection of images acquired by a trinocular system under gradually increasing turbidity levels were used in that study. In [7], the robustness on underwater images local feature detection was reported. A new dataset, called TURBID, that is real seabed images with different amount of degradation, was used and the robustness of different feature detectors were analyzed in that report [7]. However, the studies in both reports [6],[7] do not include about the real-time visual servoing for underwater vehicles.

In previous works [8]-[10], different experiments to confirm the robustness of our vision-based system using two cameras and a known 3D marker were conducted. Sea trial docking using an ROV as a test bed was conducted in a real sea near Wakayama city in Japan successfully [11]. In [12], dual-eyes vision-based docking experiment was verified using an AUV "Tuna-Sand 2". Even though the robustness of the proposed system against different kinds of disturbances was confirmed in previous works, we have not confirmed the robustness of the proposed system against the effect of the water turbidity. Based on this motivation, we conducted some experiments as shown in Fig.1 to confirm the effectiveness of the proposed system under a certain level of turbidity.

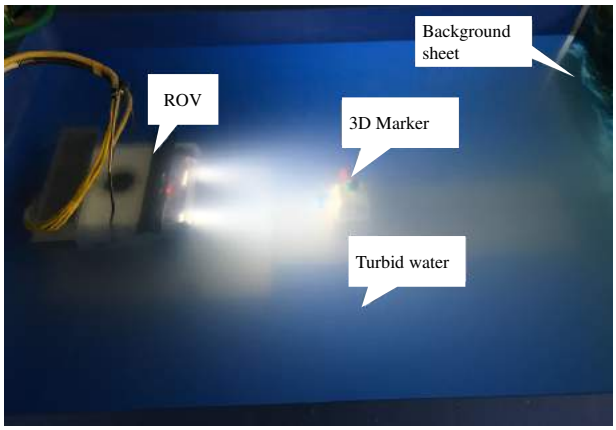

Fig.1 ROV and 3D marker in turbid water.

2. Stereo vision-based real-time 3D pose estimation

Figure 2 shows model-based pose estimation using the dual-eyes vision system. Knowing the information of the target such as shape, size and color, the solid model of the target is predefined in the computer system and projected onto 2D images. By comparing the projected solid model image with the captured 2D images by the dual-eye cameras, the relative pose difference is calculated. The target object is a 3D marker that consists of three spheres (40 mm in diameter) whose colors are red, green, and blue. In the pose estimation process, the main task is to define the number of solid models with different poses within the search space that is defined according to the field of view of the cameras. Then, models that match with the target in 2D images to a defined degree are searched for. Finally, the pose of the model that has the highest degree of matching with the target in 2D images is selected as the estimated relative pose. Figure 2 shows how a solid model is defined in 3D space and projected onto 2D images to match the captured real target in 2D images.

Real-time 3D pose estimation using 3D model-based recognition and Real-time Multi-step GA (RM-GA) was presented in detail in a previous paper [13]. The main task is to search the best model with an appropriate pose that is strongly correlated with the real 3D marker. Figure 3 shows the flowchart of RM-GA and how the best model is searched. Real time pose

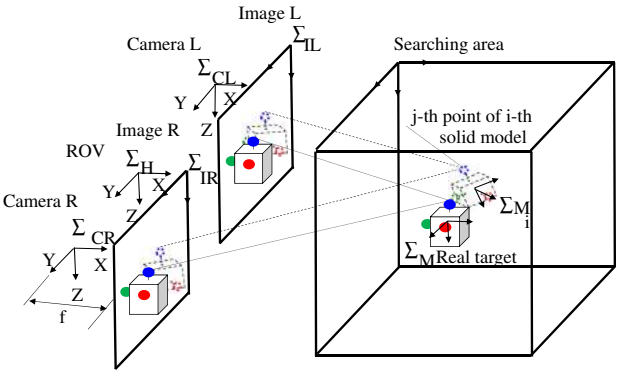

Fig.2 Model-based pose estimation using dual-eye vision system.

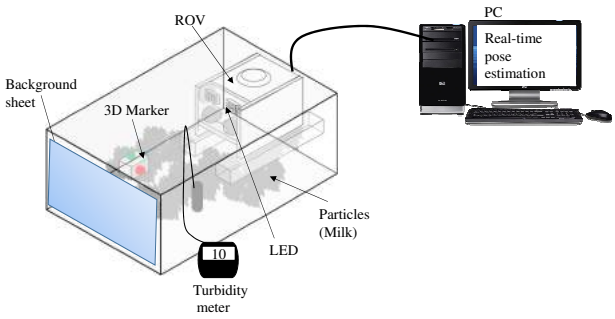

Fig.4 Experimental layout for 3D pose estimation against turbidity.

is estimated for every image with image frame rate of 30 fps. Please note that recognition and convergence are done in 3D space and evaluation is performed in 2D images.

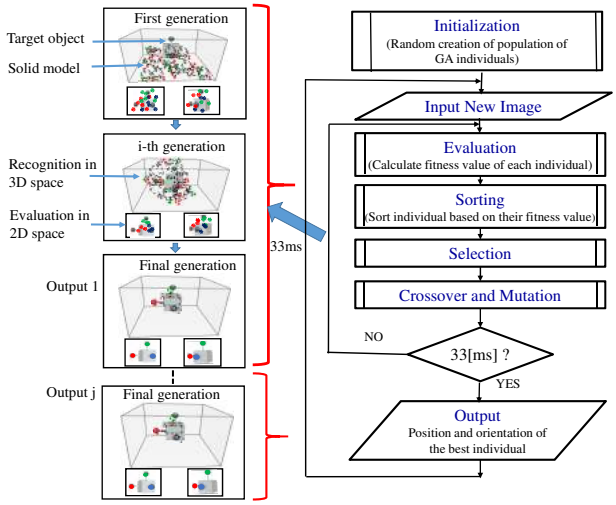

Fig.3 Pose estimation using Real-time Multi-step GA.

3. Experiment for 3D pose estimation against turbidity

3·1 Experiment layout

In this experiment, 3D pose recognition using the proposed system under different turbidity levels was conducted. Figure 4 shows the experimental layout for 3D pose estimation under different turbidity levels. In this experiment, the distance between the ROV and the 3D marker is fixed and illumination is constant by setting two LED units of the ROV to emit directly to the 3D marker as shown in Fig.1 with an illumination intensity of 200 Lx. The illumination intensity is measured by a Lux sensor LX-1010B when the sensor is set in front of the LED of the ROV with a distance of 600 mm. The experiments were conducted in the dark environment.

Water turbidity was simulated by adding milk. According to literature review [6],[7], the diameter range of milk molecules is from 10 to 600 nm. Particles with size of 10 nm do scattering of light with equal amount of lightforward and backward. The forward-scattering starts dominating for particles about 100 nm and, close to 1000 nm, there is strong small-angle forward scattering and weak backscattering. Therefore, milk is selected as main particles for turbidity since it can provide all types of scattering. Comparing to the maximum milk of 0.19 ml/l (190 ml for 1000 liter of water) in [6] and 1.5 ml/l in [7], we conducted experiment adding milk up to 0.12 ml/l of milk (95 ml for 800 liters of water). Please note that light sources used in [6] and [7] are different. Two fluorescent light strips were used in [6] and a halogen lamp was used in [6]. In this experiment, two LED units installed on the ROV as shown in Fig.6 was used as a light source. The position of the ROV is set in front of the 3D maker with a fixed distance of 600 mm as shown in Fig.4. To be similar to a real sea environment, a background sheet as shown in Fig.5 including sea environment patterns is placed behind the 3D marker as shown in Fig.4. The pool size is 1580 [mm] × 1100 [mm] × 590 [mm]. The amount of water filled up into the pool is 800 liters. Then, we added milk by adding 2 g for each time for different level of turbidity up to 30 g and added 4g of milk up to 98g. The turbidity of water is measured by a turbidity sensor TD-500 that has a measurement range of 0.0 to 500 FTU (Formazin Turbidity Unit).

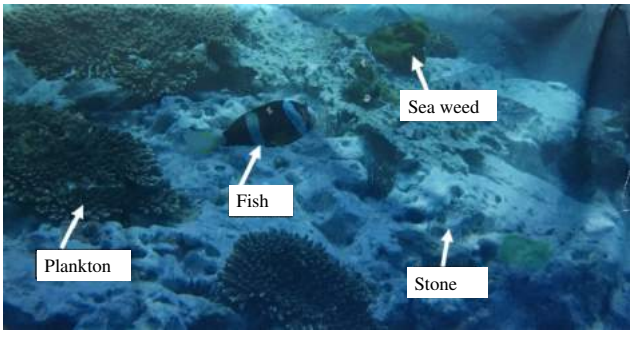

Fig.5 Photo of background sheet.

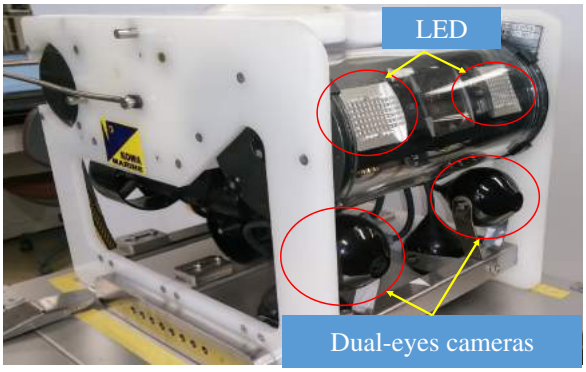

Fig.6 Photograph of ROV.

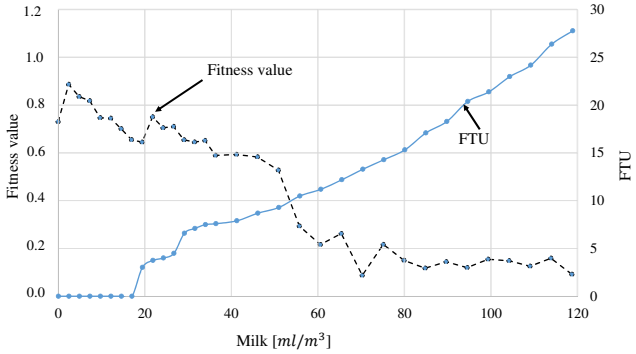

Fig.7 Fitness value and FTU measurement for different turbidity with milk.

3·2 Underwater Vehicle

The remotely controlled underwater robot (Kowa, maximum depth 50 m) used in this experiment is shown in Fig. 6. Two fixed forward cameras with the same specifications (imaging element CCD, pixel number 640 × 480, pixel focal length 2.9 mm, signal system NTSC, minimum illumination 0.8 lx, no zoom) are mounted on the ROV. These two fixed cameras are used for 3D object recognition. The thruster system of the ROV consists of two horizontal thrusters with a maximum thrust of 4.9 N each, and one vertical thruster and one lateral thruster with a maximum thrust of 4.9 N each. In this experiment, only recognition was conducted and the thrusters of the ROV were not controlled. The ROV is equipped with two units of LED lights (5.8 W) as an illumination source.

3·3 Evaluating 3D recognition

In this experiment, fitness value is used to evaluate the performance of the recognition under different turbidity levels. A correlation function of the real target projected in camera images with the assumed model, represented by poses in the chromosomes, is used as the fitness function in the GA process. We modified the fitness function based on the voting performance and the target's structure (color, size, and shape). Please refer to [14], [15] for a detailed definition of the fitness function. In this study, the averaged fitness value for the period of 60 s is used to verify the performance of the proposed system under different levels of turbidity.

4. Results and Discussion

The performance of the 3D pose estimation under different turbidity levels in terms of fitness value is shown in Fig. 7. It can be seen that the fitness value decreases from about 0.8 to 0.1 when the turbidity increases gradually from 0 FTU (0 ml/m^3) to 27.8 FTU(118.83 ml/m^3). Two parameters of the turbidity expressed in this study are FTU measured by turbidity sensor TD-500 and amount of milk in terms of ml/m^3. There is a defined fitness value to control the ROV in our system. According to the defined fitness value, the maximum value of turbidity, to which the proposed system can recognize the 3D marker to perform visual servoing, can be determined from the experimental results in this study. For example, if the fitness value of 0.4 is defined as a threshold to control the ROV, the maximum value of turbidity is about 10 FTU.

Figure 8 shows left and right camera images with recognized pose by RM-GA under different turbidity levels. The increase of turbidity level can be seen through the photos that are arranged from top to bottom in Fig.8. According to sensitivity of turbidity

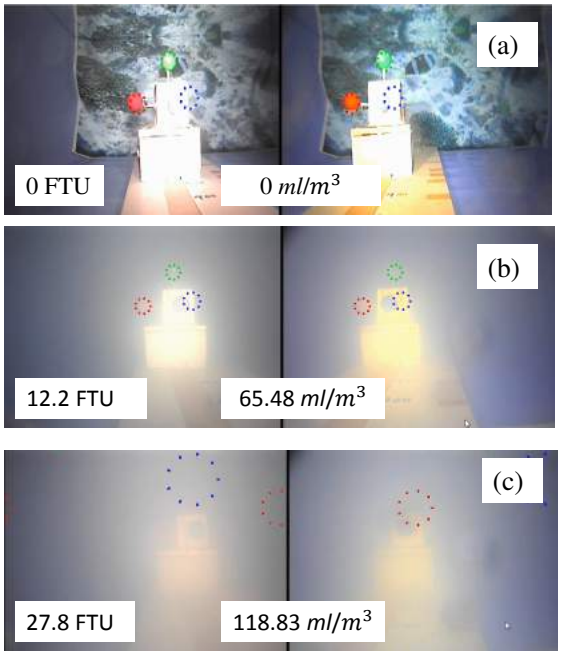

Fig.8 Left and right camera images with recognized pose for different turbidity (a) 0 FTU, (b) 12.2 FTU, and (c) 27.8 FTU when the distance between ROV and 3D Marker is 600 mm. Dotted circles in each photo is recognized pose by RM-GA. Water turbidity measured by TD-500 sensor is shown in term of FTU and added milk is expressed in term of ml/m^3.

sensor, there is no variation in measured FTU from 0 ml/m^3 to 16.98 ml/m^3. The increase of turbidity level can be also seen based on the appearance of the background sheet. For example, the photo in Fig. 8(a) is in the state of clear water and Fig. 8(c) shows the highest turbidity level in which the 3D marker can not be recognized at all in this experiment. Additional to an evaluating parameter that is fitness value, the accuracy of recognition can be visually determined by interpreting the dotted circles in each photo. When the dotted circles and 3D marker in images are matched well, it means that fitness value is high and the system can estimate the relative pose with high accuracy. As shown in Fig. 8, the system can recognize the 3D marker in the conditions of turbidity up to 12.2 FTU (Fig.8(b)).

5. Conclusion

In this work, real-time pose estimation using 3D model-based matching method and RM-GA under different turbidity levels was verified. Turbidity in a pool was simulated by adding milk. Fitness value that is a correlation function between model and captured 3D marker is used as an evaluating parameter. Experiments were conducted in a dark environment. According to experimental results, the performance of 3D pose estimation under different turbidity levels are analyzed and the maximum turbidity can be determined according to the defined threshold of fitness value.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP16K06183. The development of the ROV is cooperated by MITSUI Engineering and Shipbuilding Co.,LTD and Kowa cooperation. The authors would like to express their thanks to Professor Tamaki Ura, Professor Toshihiko Maki and Dr.Yuya Nishida for their help and support.

References

[1] B.A.A.P. Balasuriya, M. Takai, W.C. Lam, Tamaki 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] Robert S. McEwen, Brett W. Hobson, Lance McBride and James 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, October 2008.

[3] Ken Teo, Benjamin Goh and Oh Kwee Chai: " Fuzzy Docking Guidance Using Augmented Navigation System on an AUV," IEEE Journal of Oceans Engineering, Vol. 37, NO. 2, April 2015.

[4] Narc s Palomeras, Antonio Pe nalver, Miquel Massot-Campos, Guillem Vallicrosa, Pep Llu s Negre, J. Javier Fernndez, Pere Ridao, Pedro J. Sanz, Gabriel Oliver-Codina and Albert 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] J.-Y. Park, B.-H. Jun, P.-M. Lee and J. Oh: "Experiments on vision guided docking of an autonomous underwater vehicle using one camera," Ocean Eng., Vol. 36, No. 1, pp. 48-61, Jan. 2009.

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

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

[8] Myo Myint, Kenta YONEMORI, Akira YANOU, Shintaro ISHIYAMA and Mamoru MINAMI,: " Robustness of Visual-Servo against Air Bubble Disturbance of Underwater Vehicle System Using Three-Dimensional Marker and Dual-Eye Cameras," Proceedings of the International Conference OCEANS15 MTS/IEEE, Washington DC, USA, pp.1-8, 2015.

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

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

[11] Myo Myint, Kenta YONEMORI, Akira YANOU, Khin Nwe Lwin, Naoki Mukada and Mamoru MINAMI: "Dual eyes visual based sea docking for sea bottom battery recharging," Proceedings of the International Conference OCEANS16 MTS/IEEE, Monterey, USA, pp.1-7, 2016.

[12] Xiang Li, Yuya Nishida, Myo Myint, Kenta YONEMORI, Naoki Mukada, Khin Nwe Lwin, Matsuno Takayuki and Mamoru MINAMI: "Dual-eyes Vision-based Docking Experiment of AUV for Sea Bottom Battery Recharging," International Conference OCEANS17 MTS/IEEE, Aberdeen, Scotland, Jun. 2017. (The citable proceedings of the conference has not been delivered as of July 5th, 2017).

[13] Myo Myint, Kenta YONEMORI, Akira YANOU, Khin Nwe Lwin, Mamoru MINAMI and Shintaro Ishiyama : " Visual servoing for underwater vehicle using dual-eyes evolutionary real-time pose tracking," Journal of Robotics and Mechatronics, Vol. 28, No. 4, pp. 543-558, Aug. 2016.

[14] W. Song, M. Minami and S. Aoyagi: " On-line Stable Evolutionary Recognition Based on Unit Quaternion Representation by Motion-Feedforward Compensation," International Journal of Intelligent Computing in Medical Sciences and Image Processing (IC-MED), Vol. 2, no. 2, pp.127-139, 2008.

[15] Minami, M., Agbanhan, J. and Asakura, T.: " Evolutionary Scene Recognition and Simultaneous Position/Orientation Detection, In Soft Computing in Measurement and Information Acquisition," Springer Berlin Heidelberg, pp.178-207,2003.