Proposal of Stereo-vision Based Docking System for Battery Recharging in Real Sea

2018, March

MYO MYINT

Graduate School of

Natural Science and Technology

(Doctor's Course)

OKAYAMA UNIVERSITY

Proposal of Stereo-vision Based Docking System for Battery Recharging in Real Sea

Intelligent Robotics and Control Laboratory

Myo Myint (51427304)

Abstract :

Japan has many areas of sea from which future resources can be taken out using advanced technologies. Autonomous Underwater Vehicle (AUV) plays an important role in deep sea works such as oil pipe inspection, survey of sea floor, searching expensive metal, etc. To do such novel works that take a long period in deep sea, one of the main limitation of AUVs is limited power capacity. To solve this problem, underwater battery recharging unit with a docking function is one of the solutions to extend the operation time of AUVs. Most studies on docking for AUVs using visual information are based on monocular camera to acquire the pose between a target and a vehicle. The disadvantage is that the precision of distance measurement of the camera ' s depth direction is not enough for applications in which high homing accuracy is important. Therefore, as an initiated research to AUV environment, we have developed a 3D-Move on Sensing system using stereo vision to provide high homing accuracy.

This thesis proposes a stereo vision-based docking system that is new and noble for battery recharging in real sea. In the proposed approach, visual information is directly used in feedback control in real-time. Additionally, developed optimization method named Real-time Multi-step GA is implemented in accordance with the concept of optimization of dynamic images for real-time target tracking. Moreover, when AUVs must operate in unstructured environments such as near the seafloor, the most challenging and unavoidable problem with the autonomous operation of AUVs is turbidity limiting optical visibility. To the best of the author ' knowledge, no studies have yet been conducted on 3D pose estimation against turbidity for underwater vehicles. Therefore, the effect of

turbidity on the 3D pose estimation performance of underwater vehicles and a method of operating under turbid conditions were studied in this work. Experiments using a remotely operated vehicle (ROV) with dual-eye cameras and a passive 3D marker were conducted in the pool and real sea. The experimental results confirmed that the proposed system is able to provide high homing accuracy and robustness against disturbances that influence not only the captured camera images but also the movement of the vehicle. A successful docking operation using stereo vision that is new and novel to the underwater vehicle environment was achieved and thus proved the effectiveness of the proposed system for AUV.

In the first part of this dissertation, a new method of pose estimation scheme that is based on 3D model-based recognition is proposed for real-time pose tracking to be applied in Autonomous Underwater Vehicle (AUV). In this method, a 3D marker is used as a passive target that is simple but enough rich of information. 1-step Genetic Algorithm (GA) (later named as Real-time Multi-step GA) is utilized in searching process of pose in term of optimization, because of its effectiveness, simplicity and promising performance of recursive evaluation, for real-time pose tracking performance. The proposed system is implemented as software implementation and Remotely Operated Vehicle (ROV) is used as a test-bed. In simulated experiment, the ROV recognizes the target, estimates the relative pose of vehicle with respect to the target and controls the vehicle to be regulated in a desired pose. P control concept is adapted for proper regulation function. Finally, the robustness of the proposed system is verified in the case when there are physical disturbances and in the case when the target object is partially occluded. Experiments are conducted in an indoor pool. Experimental results show recognition accuracy and regulating performance with errors kept in centimeter level.

The second part of this dissertation presents docking performance using proposed docking strategy that was designed and demonstrated for underwater battery recharging. Among two common configurations of docking stations that are omnidirectional and unidirectional one, a simulated docking station with a unidirectional entry is designed for

underwater battery recharging in this work. Instead of integration with other sensors, a standalone dual-eye vision system was applied in this study. The overall target of this study is to check the functionality and practicality of proposed algorithm for an intended docking application. Docking strategy was designed and implemented experimentally. Experiments were conducted in a pool and real sea near Wakayama city in Japan using an ROV to confirm that the proposed approach is able to guide an ROV to insert a rod attached on the ROV into a docking hole with a radius of 35 mm attached with a 3D marker. A successful docking operation from different starting positions using stereo vision that is new and novel to the underwater vehicle environment was achieved and thus proved the effectiveness of the proposed system for AUV.

In the final part of this dissertation, the turbidity tolerance of proposed docking approach was analyzed and verified experimentally. Since underwater battery recharging units are supposed to be installed in deep sea to save the time consuming and work done from human beings in the case of returning surface vehicle for recharging, the deep-sea docking experiments cannot avoid turbidity. According to the author ' knowledge, there is no study on docking system using stereo-vision based real-time visual servoing with performance tolerance of turbidity. In this study, we conducted experiments to verify the robustness of the proposed docking approach in simulated pool where different levels of the turbidity of the water is simulated. The experimental results have confirmed the robustness of the docking system using stereo-vision based 3D pose estimation against turbidity. Finally, docking experiments in a real sea were conducted to verify the functionality and practicality of the proposed approach. A shallow sea area was selected as the docking area because the high turbidity in a shallow region would allow the verification of the robustness of the proposed system against turbidity. The experimental results verify the robustness of the system against turbidity, presenting a possible solution to a major problem in the field of robotics.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Recently, social demand for deep-sea resources, such as food and energy, has increased rapidly with the aid of developing technologies. At present, almost all metal and industrial mineral extraction is conducted onshore. Deep-sea extraction would expand the available resource pool and provide a level of independence from onshore resource extraction. Among the available deep-sea resources, expensive rare metals and methane hydrate in the seabed are of particular economical importance. Japan has been considering the deep-sea mining of methane hydrate, which is expected to be a future energy resource. Furthermore, the information that can be gathered from deep-sea exploration is useful in the prediction of disasters such as earthquakes and tsunamis and can help us understand how we are affecting and being affected by changes in Earth's climate and atmosphere. Therefore, the advancement of deep-sea research technologies would be highly beneficial in a number of applications.

To meet the above mentioned social demands for sea operations, underwater robots have been developed worldwide [1]-[8]. AUVs have become essential in deep-sea operations such as cable tracking [9], ocean bottom exploration and underwater surveying. Comparing to ROV, using AUV is effective for deep sea works because it has no tethered cable, it does not depend on a great deal of effort of the operator, and it can provide the efficient observation of wide area. In a conventional operation, (1) AUV is transported

to the desired working area by a mother ship, (2) it divides from the mother ship into the deep sea to do the tasks such as searching resources, and (3) it goes back to the sea surface. Then, (4) the AUV is collected by the mother ship for maintenance or taking new instruction for next task. The main problem of this operation method is that there is a limitation on the capacity of the battery. Since the electricity of AUVs is supplied by the battery for AUV's moving around the sea floor, AUVs have to float to the sea surface for recharging if the power capacity of AUV is not enough for tasks that take longer operation. Therefore, decreasing the working time and dropping the work efficiency in the deep sea became the problems for deep sea applications where operations take a couple of days.

To solve these problems, underwater battery recharging technology with docking function is one of the solutions even though challenges are still remained. In a docking- based battery recharging system, the power supply facility is installed on the seabed in which the AUV automatically charges without going to the sea surface and it can do tasks continuously for a long time. Moreover, docking function takes place as an important role not only for battery recharging but also for other advanced applications such as intervention using some manipulators.

1.1 Background and motivation

Research on docking operations using various homing sensors and techniques for the underwater robot has been conducted worldwide [10]-[46]. The optical terminal guidance technique was introduced in [10]. A docking guidance system was designed and implemented by the Sugeno fuzzy inference system (FIS) in [11]. In [12], an electromagnetic homing (EM) system for docking was proposed and tested. In [13], the AUV homed to a docking station by using an ultrashort base line (USBL) sonar transceiver mounted in the vehicle nose. The work in [14] proposed a robust AUV docking guidance and navigation approach that can handle unknown current disturbances without a velocity sensor. The

work in [15] addressed robust vision-based target recognition by presenting a scaled and rotationally invariant target design and a recognition routine based on self-similar landmarks (SSL). Generally, there are three steps in a docking operation; (1) long distance navigation step, (2) approaching step, and (3) final docking step. Among them, final docking step is a critical task when accurate homing accuracy and robustness against different disturbances are dominant. To fulfill this requirement, visual servoing is one of the solutions and that can be integrated with other sensor units.

Recently, due to the progress in computer vision, a vision-based system has been highlighted as a promising navigation system. As in land and space systems, numerous studies on underwater vehicles using visual servoing have recently been conducted worldwide. Each study has different merits and limitations depending on the intended application. Most research is based on monocular vision [20], [22]. In [21], features in a man-made plate were extracted and the relative pose was estimated from camera images. In [51], a vision system using two cameras and artificial underwater landmarks for autonomous operation was reported. Even though two cameras were installed in the vehicle in [21],[51], both cameras did not see the same object at the same time to estimate the relative pose. One camera detected the target and the second camera performed other tasks.

Apart from single camera based approach, as an initiated docking approach using stereo-vision, we have developed dual-eye vision-based docking system, especially for final docking step. In the proposed stereo vision system, both cameras seeing the same target object with parallactic displacement are used to estimate the pose of the target object through the proposed real-time 3D pose recognition method. The merit of the stereo camera is that the space recognition is superior than the monocular camera. Instead of measuring absolute position of vehicle using other non-contact sensors, estimation of the robot's relative position and posture (pose) using dual-eyes camera and 3D target object is proposed in this study. For relative pose estimation, model-based recognition approach is applied because of its real-time effective performance comparing to other methods such

as feature based recognition in which the information of the target object is determined by a set of image points, which can address complex searching the corresponding points and time consuming. According to effectiveness, simplicity and repeatable evaluation for real-time performance, Genetic Algorithm (GA) is utilized to evaluate the gene candidates which represent relative poses until getting the best gene with the most trustful pose in the proposed system. Fitness value which is correlation function of projected model against the real target in the image is used as the evaluation parameter of recognition process. P controller is used to control the vehicle for the desired pose using real-time images from dual-eyes camera. The proposed system is implemented in PC, and the Remotely Operated Vehicle (ROV) is tethered through the cable with 200 m length to receive image information and control signals. Since underwater environment is very complex, all possibilities that real underwater world might provide are considered. Therefore, several experiments were conducted using two cameras and a known 3D marker to confirm the robustness of the docking system against different disturbances. Finally, sea docking experiments were conducted to verify the robustness of the proposed system against real sea environment.

1.2 Aim and objectives

The overall aim of the research presented in this thesis is to develop stereo-vision based docking approach for underwater battery recharging application with high accuracy and robustness against different disturbances. To achieve this aim, the following objectives should be fulfilled:

- to develop real time 3D pose estimation with high accuracy and robustness against different disturbances that may occur in real sea environment.
- to develop docking approach through visual servoing using proposed real-time 3D pose estimation method.

- to verify the robustness of the proposed docking system against different disturbances in pool tests.
- to verify the turbidity tolerance of the proposed system since turbidity is the most challenging and unavoidable problem in deep sea where the underwater battery recharging unit is supposed to be installed.
- to verify the functionality and practicality of the proposed system by conducting docking experiment in a real sea environment.

1.3 Principal contributions

The principal contributions of this thesis are listed as follows:

- The main task in this docking experiment is 3D pose tracking in time by following dynamic images with a video rate of 30 frames per second (fps). There is no study that has achieved real-time 3D pose estimation by using dual-eye cameras for AUV in docking operation in which only visual information is directly used in a feedback loop. Therefore, as a main contribution of this study, a new method of real-time 3D pose estimation in successively input dynamic images from two cameras using 3D model-based recognition method utilizing Real-time Multi-step GA (RM-GA) is proposed.
- The overall target of this study is to check the functionality and practicality of our proposed algorithm for an intended docking application. Docking strategy was designed and implemented experimentally. Since underwater environment is more complex than space and ground, there are many disturbances for vision-based underwater vehicles. Therefore, it is important to consider the possible disturbances before testing the proposed approach in the sea. To verify the robustness of the proposed system, we conducted experiments with different disturbances including object occlusion, and physical disturbances. Experiments were conducted using an

ROV to confirm that the proposed approach is able to guide an ROV to insert a rod attached on the ROV into a docking hole attached with a 3D marker.

- Although AUVs do not need to descend to the sea bottom for some tasks, such as bottom topology surveys, working near the sea floor is necessary for most ocean exploration operations, including oil pipe inspection and the detection and extraction of precious metals. At this time, the most challenging and unavoidable problem in deep-sea operations is turbidity, which deteriorates the visual capabilities of AUVs. Therefore, the verification of the turbidity tolerance of an AUV and the development of a method to overcome disturbances caused by turbidity are important research questions not only for AUV development but also for the field of vision-based underwater systems. Because the intended application in this study is underwater battery recharging at the sea bottom to extend the operation time of AUVs, turbidity cannot be avoided by simply operating the AUVs in clean water. Therefore, the effect of turbidity on the 3D pose estimation performance of underwater vehicles and a method of operating under turbid conditions were studied in this work.

- Final contribution of this study is that it is the first experimental investigation of the practicality of undersea docking using two cameras under turbid conditions in an actual undersea environment. A continuous iterative docking experiment was conducted in a shallow sea region near the town of Ushimado, Japan.

1.4 Dissertation structure

This thesis is organized as follows:

Chapter 2 presents literature review on underwater docking approaches including pose estimation, visual servoing and optimization.

Chapter 3 describes the proposed system with detailed explanation on real-time 3D pose estimation, model based matching, fitness function, and real-time multi-step GA (RM-GA).

Chapter 4 presents experiments conducted in different environments for recognition, and regulating.

Chapter 5 describes the docking experiment using proposed docking strategy.

Chapter 6 presents the turbidity tolerance of the proposed system with the experimental results.

Chapter 7 concludes this thesis with the summary of contributions and provides some recommendations for the future work on this research field.

1.5 Publications

The research work presented in this thesis has resulted in the following publications.

Journals

1. 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, 2016.

2. Myo Myint, Kenta Yonemori, Khin Nwe Lwin, Akira Yanou, Mamoru Minami, "Dual-eyes Vision-based Docking System for Autonomous Underwater Vehicle: An approach and Experiments," *Journal of intelligent and robotic systems*,

DOI 10.1007/s10846-017-0703-6, 2017.

International Conferences

1. 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 ," in *Proceedings of International Symposium on System Integration (IEEE/SICE)*, pp.989-994, 2015.

2. 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," in *Proceedings of the International Conference on OCEANS (Washington DC, USA) (MTS/IEEE)*, 150521-002, 2015.

3. 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," in *Proceedings of the International Conference on OCEANS (Shanghai, China) (MTS/IEEE)*, 151114-025, 2016.

4. 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," in *Proceedings of the International Conference on OCEANS (Monterey, USA) (MTS/IEEE)*, 160425-018, 2016.

5. Naoki Mukada, Kenta Yonemori, Myo Myint, Khin Nwe Lwin, Akira Yanou and Mamoru Minami, " Tracking Trajectory Control of Dual-eyes Visual-based Underwater Vehicle," in *Proceedings of International Symposium on System Intergradation (IEEE/SICE)*, pp.748-755, 2016.

6. 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," in *Proceedings of the International Conference on OCEANS (Scotland, UK) (MTS/IEEE)*, 161130-048, 2017.

National Conferences

1. Myo Myint, Mamoru MINAMI, Kenta YONEMORI, Yukihiro SAKA and Akira YANOU, " Visual Servoing Experiments of Underwater Vehicle under Air Bubble Disturbances," in *Proceedings of Conference of the Robotics Society of Japan*, September 3-5, Tokyo, Japan, 2015.

2. Myo Myint, Kenta Yonemori, Akira Yanou, Shintaro Ishiyama, and Mamoru Minami, "Real-time 3D Pose Estimation and Tracking 3D Marker using Dual-eyes Camera for Underwater Vehicle," in *Proceedings of the 24th Annual Conference of the SICE Chugoku Chapter (Okayama,Japan)*, November 28, pp.62-63, 2015.

3. Myo Myint, Khin Nwe Lwin, Naoki Mukada, Matsuno Takayuki, and Mamoru Minami, "Stereo Vision-based 3D Pose Estimation under Turbid Water for Underwater Vehicles," in *Proceedings of the Robotics Society of Japan, Tokyo, Japan*, September 11-14, 2017.

4. Myo Myint, Kenta Yonemori, Khin Nwe Lwin, Naoki Mukada, Mamoru Minami, Takayuki Matsuno, "Vision-based Docking Simulation of Underwater Vehicle Using Stereo Vision under Dynamic Light Environment," in *Proceedings of the 9 th SICE Symposium on Computational Intelligence, Chiba, Japan*, July 8-9, 2017.

Chapter 2

Literature review

In this section, literature review on some background topics related to this study is presented. Related studies for each topic are discussed respectively. First, the studies related to underwater docking are described. The discussion on visual servoing with some related studies is presented. Then, 3D pose estimation that is the main approach in this study is explained with some references. After that, optimization methods especially genetic algorithm are discussed with some background reviews. Finally, robustness against disturbances that will be faced in sea underwater environment are discussed in this section.

2.1 Underwater docking

Firstly, the definition of underwater docking can be defined as the process whereby an AUV purposefully transitions from a state of free flight to being physically connected to another device or platform, called the docking station [48]. Generally, the docking process can be divided into three steps; (1) Long distance navigation, (2) approaching, and (3) docking as shown in Fig. 2.1.

The initial aim of early docking works was to operate AUVs without human servicing of the AUV especially for recovery the AUV. Recently, docking application area has been expanded wildly with the development of the technology. Underwater docking in this study is to provide an approach for the AUV to find the docking station, to physically

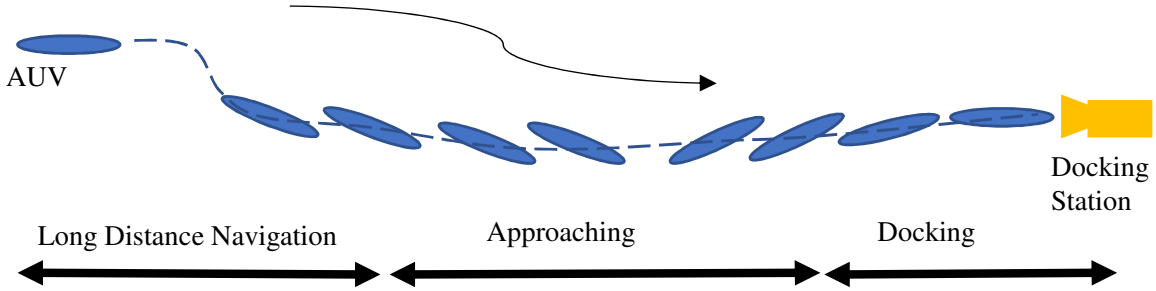
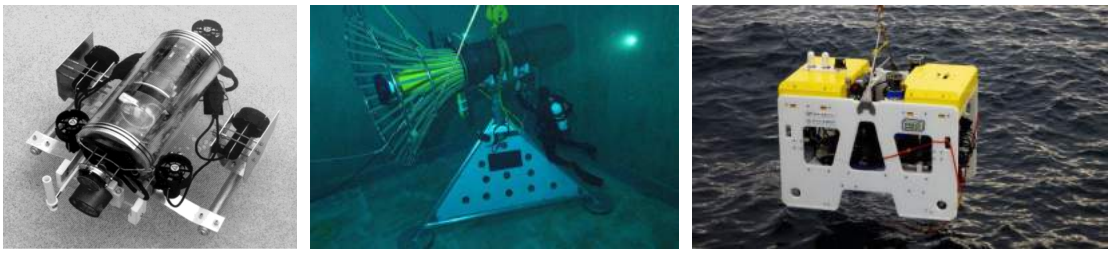


Fig. 2.1: Underwater AUV docks into the docking station. The docking process generally involves (1) long distance navigation, (2)approaching, and (3) docking.

attach, and to recharge AUV batteries by the power source that is supposed to be installed in underwater deep sea environment where the AUV is aimed to work for a long time operation without returning to the surface for recharging. Figure 2.2 shows the different AUVs those were used for docking experiments. There are different homing methods, docking station configurations, power transfer approaches, and communications links to achieve these respective applications. In this section, different approaches will be discussed in terms of homing methods, Docking station configuration, and Sensor Configuration.



(a) (b) (c)

Fig. 2.2: Different AUVs used for docking experiments: (a) FAU AUDREY AUV [14], (b) Dorado/Bluefin type AUV [17], and (c) Tuna-Sand 2 AUV [63].

2.1.1 Homing methods

The different methods of homing/docking that is connecting AUV to the docking station are proposed in [13],[17]-[19]. A method of capturing AUV straight to a funnel structure docking station was proposed [17],[18]. Normally, catching AUVs has been conducted by using big net mechanisms (as shown in Fig. 2.3 (a)) with appropriate homing accuracy. However, this kind of technique can occur any physical damage to AUVs as well as the

docking station. In research [13] using a torpedo-type vehicle, the lack of a hovering function reduced the docking efficiency, even though the speed of the vehicle was suitable. In [19], a docking method using manipulator (sample figure is shown in Fig. 2.2 (b)) is proposed. Instead, we proposed and used docking method that is fitting the docking pole attached in AUV into the docking hole installed in the docking station as shown in Fig. 2.3 (c). Since the power supply device and the AUV can be mechanically coupled to each other, automatic power recharging in this experiment is assumed to be enabled when the underwater vehicle fits its docking pole into the docking hole. This method can minimize equipment cost and can be applied to other methods [17] and [19].

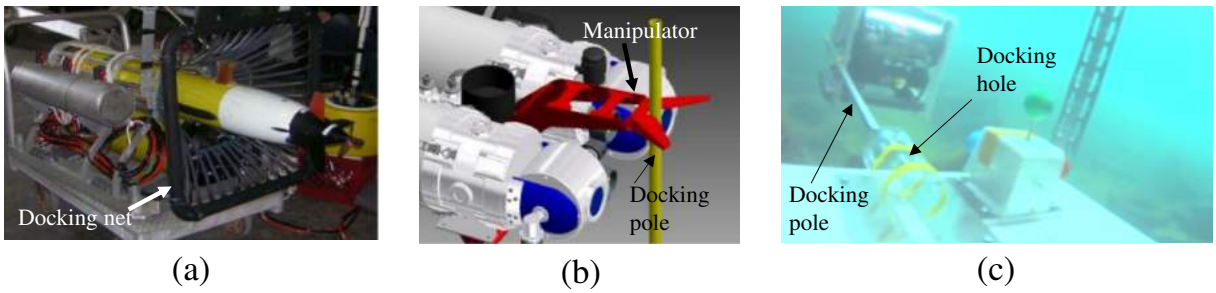
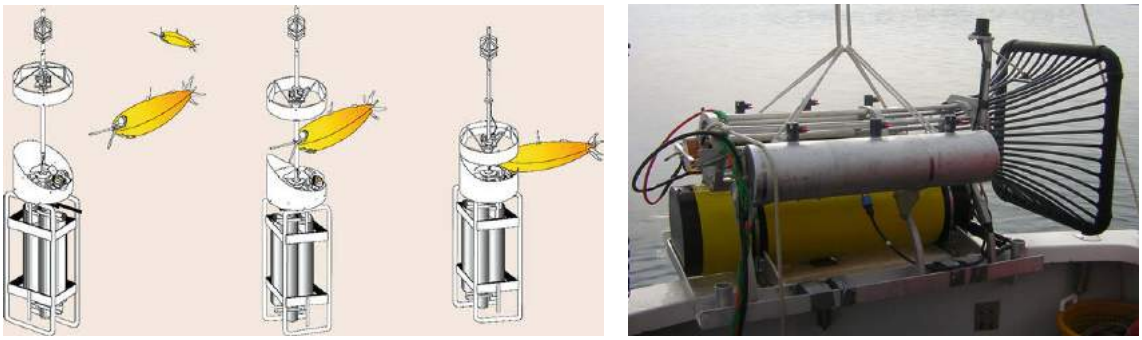


Fig. 2.3: Different homing methods: (a) homing using docking net mechanism, (b) homing using a manipulator, and (c) proposed homing method with docking pole and docking hole.

2.1.2 Docking station configuration

The two common configurations of docking stations as shown in Fig. 2.4 are omnidirectional [64], where the docking hole can rotate to allow a vehicle to approach and dock from any azimuth, and unidirectional [65], where the docking hole is oriented in a specific direction. The unidirectional station has been selected in many studies because of its robustness and simplicity. However, the final approach of unidirectional docking is a difficult task, even though expensive navigation sensor suites and large-scale dead-reckoning sensors are able to provide position data. In this work, a simulated docking station with an unidirectional entry is designed for underwater battery recharging. Therefore, the initial objective for successful docking is that the proposed system provide high homing

accuracy and robustness against disturbances.



(a) (b)

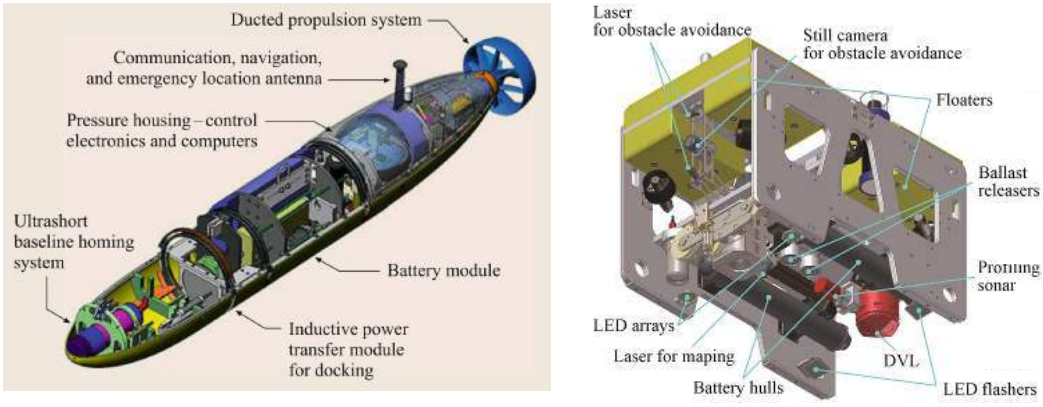
Fig. 2.4: Different docking structures : (a) Omnidirectional docking, (b) Unidirectional docking.

2.1.3 Sensor Configuration

The sensor unit installed either in AUV or both in AUV and docking station needs to provide the information for the AUV to find the docking station and perform docking/homing operation. In an ideal sensor unit, it would provide reliable relative location between the AUV and the docking station with high accuracy in high frequency. However, there is no such perfect unique sensor in practical world. Therefore, different sensor units with different techniques have been utilized for respective application as shown in Fig. 2.5. Normally, different types of sensor units are integrated for perfect docking operation. Among them, the common sensors and techniques are discussed in this section as follows.

Acoustic systems

Acoustic systems have been used for long distance navigation step in AUV docking. Acquisition distances ranging is from a fraction a kilometer to hundred of kilometers. In an acoustic system, one transducer is installed in the AUV and the another is attached in the docking station. Normally the transducer of AUV emits a signal and the one in docking station replies. By measuring the round trip travel time of the signal, the relative range is determined. According to the size of the transducer array, acoustic systems are



(a) (b)

Fig. 2.5: AUV with integrated sensor units : (a) torpedo-type AUV [17], (b) Hovering type AUV [63].

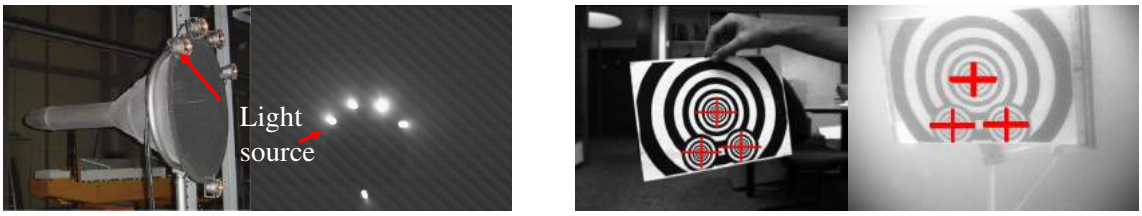
classified as long baseline, short baseline, and ultra-short baseline navigation systems. For acoustic-based sensors used in AUVs, however, the effects of undesirable acoustic reflections and attenuation may reduce the accuracy of AUV navigation. This is because speed of sound, sound refraction due to sound speed variations, interactions of sound with the seafloor and sea surface, ambient noise, and the absorption properties of seawater in complex environments, like shallow water may cause individual pulses in the received signal to fade or amplify it.

Most of the AUV studies are based on acoustic systems [16], [49], [50]. In [49], Doppler Velocity Log (DVL)and a Ultra-Short Base Line position system (USBL) were used for underwater vehicle localization.

Optical system

There are some studies using optical system to localize the underwater vehicle relative to the docking station. In an optical system, the target docking station is looked for using imaging systems and image processing algorithms. Generally, the target that is installed in the docking is detected by image processing technology. The target object can be man-made structured patterns [15] and active lighting unit [20]. Among them, the docking station in which light sources are installed the entrance of the tunnel shape

docking station is the most used type. Comparing to the acoustic systems, the acquisition range of optical systems is also limited in a real sea environment because optical signal is attenuated due to scattering and absorption of light. Figure 2.6 (a) shows the most simple and common used optical systems in which light sources are installed around the entrance of the docking station, and (b) shows the structured pattern that is used in [15] to provide the localization of the underwater vehicle relative to the docking station where the structured pattern is installed. The appearances of each system in the water are illustrated in the right side of each system in Fig. 2.6.



(a) (b)

Fig. 2.6: Different optical systems : (a) Using light sources [20], (b) Using structured patterns [15].

2.2 Visual servoing

Nowadays, visual servoing in which visual information is used to control the robot ' s motion plays an important role in different domains of application with the rapid progresses in computer vision technology. Generally, visual servoing techniques are divided into two categories; Image-Based Visual Servoing (IBVS) and Position-Based Visual Servoing (PBVS). In IBVS techniques, images from camera are used directly for control of robot. In PBVS, information of known object are extracted and interpreted from the images and used in controlling of robot in reference space rather than in image space as in IBVS [52]-[54]. Based on the location of camera, eye-in-hand and eye-to-hand configuration are considered according to the requirement of application. Then, the techniques are differentiated based on the number of cameras; from single to multi cameras. Even though there are some limitations for real-time applications in terms of image-acquisition-quantization

accuracy and processing rates, the role of visual information has been expanding rapidly in industry and human society in line with efforts of researchers [55], [56].

Recently, due to the progress in computer vision, a vision-based system has been highlighted as a promising navigation system. As in land and space systems, numerous studies on underwater vehicles using visual servoing have recently been conducted worldwide. Each study has different merits and limitations depending on the intended application. Most research is based on monocular vision [20], [22]. In our proposed system, we used position-based visual servoing in which the relative pose between the underwater vehicle and the target is estimated using images of two cameras and estimated pose is used as feedback in controlling the vehicle. The detailed of the vision-based pose estimation will be explained in the Chapter 3.

2.2.1 2D-to-3D reconstruction and 3D-to-2D projection

In conventional approaches, object recognition including relative pose information is implemented by feature-based recognition based on 2D-to-3D reconstruction. The information of the target object is determined by a set of image points in different images, and the process entails a time-consuming complex search of the corresponding points. A model-based pose estimation approach based on 3D-to-2D projection is applied in this work to avoid the effects of incorrect mapping points in images. Both 2D-to-3D reconstruction and 3D-to-2D projection are shown in Fig. 2.7. Points “B” in image 1 and “C” in image 2 are mapped incorrectly as a pair of points during 2D to 3D point-to-point reconstruction as shown in Fig. 2.7 (a). Consequently, the reconstructed 3D point “A” does not represent a true 3D object. In contrast, points including “A” and “B” are correctly projected in group from object in 3D in both images to 2D projection as shown in Fig. 2.7 (b). This is possible because the forward projection from 3D to 2D generates unique points in 2D images without errors. Based on this way of thinking, 3D model-based recognition is implemented. Other model-based approaches that are mostly based on template

matching have used a 2D model and evaluated 2D images. These kinds of techniques cannot be extended to 3D pose estimation. However, our method is based on the idea of recognition using a 3D model and evaluating 2D images from left and right cameras. The method evaluates how much the 3D model's pose overlaps the actual 3D target, that is, the correlation of the 3D model and the 3D target in 3D space through left and right projected 2D images.

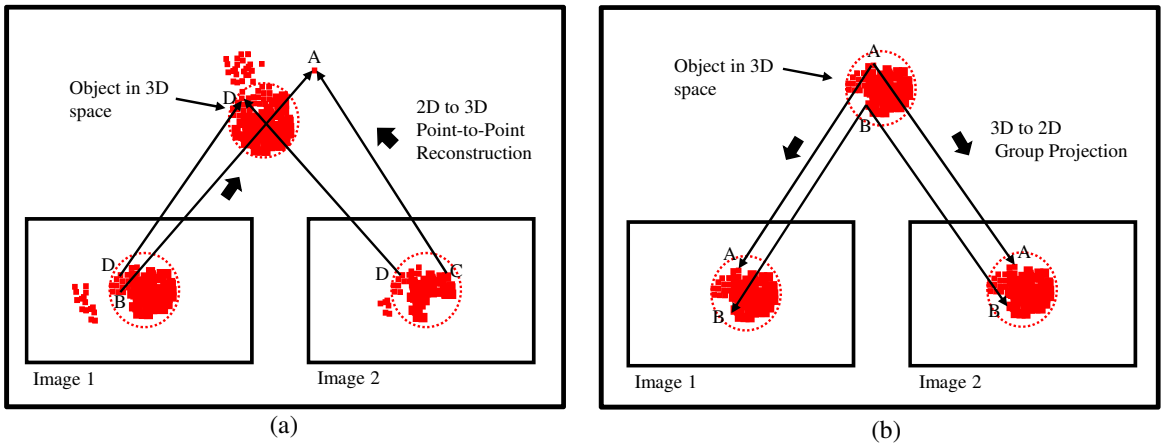


Fig. 2.7: (a) Mis-mapping in 2D-to-3D reconstruction, and (b) Pairing of points in 3D-to-2D projection.

2.2.2 Pose estimation using landmarks

In some docking experiments in other works, the relative pose is estimated by one camera and a known target or landmark [47],[66]-[69]. The pose estimation techniques reported in these works employ feature-based recognition. The work in [66] provided the relative position and distance from a geometric arrangement of lights set at docking station. Especially, the calculation of relative orientation was more complicated and difficult than detection of the position. In [69], the feature-based algorithm ORB was used. ORB is a combination of oriented features from the accelerated segment test (FAST) and rotated binary robust independent elementary features (BRIEF). ORB was applied in [86] for pose estimation of a man-made plate by using camera images to support the navigation system when position data from other sensors were no longer available. Even though the applied

vision-based docking detector algorithm in [86] utilized the ORB feature extractor, the derived relative information was input into a localization filter for information synthesis rather than a standalone vision sensor.

In [47], a signboard system including four colored balls was used as a passive target object. Based on the known information about the signboard system, the distance and orientation (heading angle only) of a vehicle to the target were calculated for position and heading error correction. The vehicle was assumed to be in a horizontal plane at the same level as three of the four balls, and so the accuracy was very dependent on other sensors, such as the altimeter, as well as stability control of the vehicle. Apart from above studies, 3D model-based matching based on 3D-to-2D projection has been developed in this study for real time pose estimation using real-time multi step GA that will be explained in detail in Chapter 3.

2.3 Optimization

In the proposed system, the best chromosome that represents the most trustful pose is the chromosome with the highest fitness function value for correlation between the model, whose pose is defined by the chromosome, and real target in the input images. The best chromosome has to be evaluated by an optimization technique instead of evaluating all possible chromosomes. Many kinds of powerful advanced optimization methods are available. However, almost all focus on accuracy rather than real-time application merit. In contrast, with the goal of constructing a pose feedback control system for docking, the two criteria of accuracy and real-time performance are indispensable to extract pose estimation in dynamic images input by video rate. Therefore, instead of comparing different optimization methods, GAs, which have a long history of usefulness, are selected and utilized in the form of the Real-time Multi-step GA for the proposed system. In other words, our strategy is utilizing the Real-time Multi-step GA (RM-GA), which has simplified optimizing calculations with reasonable performance in one loop and increasing

accuracy with repeatability within a real-time video frame rate, that is 30 fps. In this section, the concept and background of GA will be discussed.

2.3.1 Genetic algorithm

A genetic algorithm (GA) is an adaptive solution-search and optimization algorithm. GA has been used in many studies [70]-[82]. In the GA process, the candidates including optimal solution is coded in binary string that is known as a chromosomes. The GA begins its search from a randomly generated population of chromosomes that evolve over successive generations (iterations). Generally, there are three main operators in the GA process.

Selection

In the selection step, the chromosomes that are randomly generated are evaluated by a process of fitness-based selection. According to the designed fitness function, the fitter chromosomes are selected as parent chromosomes.

Crossover

The selected parent chromosomes are used to reproduce the next generation. During the process of second operator that is " Crossover, " a defined portion of each parent chromosome is changed between them to generate a new generation. The crossover operator propagates features of good surviving designs from the current population into the future population, which will have better fitness value on average.

Mutation

The third operation of GA process is Mutation that is the last operator in reproduction of next generation together with Crossover operator. Basically, the defined bits of each chromosomes after crossover process are inverted into the another state meaning that the

binary 1 becomes 0 and 0 becomes 1 in the string. Mutation operator promotes diversity in population characteristics. The mutation operator allows for global search of the design space and prevents the algorithm from getting trapped in local minima.

The whole process is being iterated until the specified stopping criterion is satisfied, and the best solution for the given problem is obtained. Figure 2.8 shows the evolution process in which the chromosome generation is evoluted from the generation to next generation through the GA operators (Selection, Crossover, Mutation).

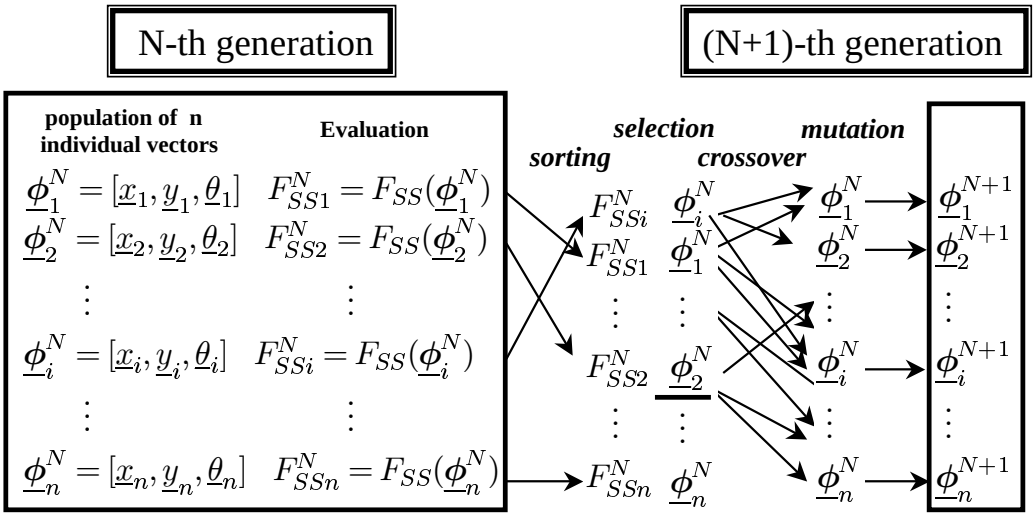


Fig. 2.8: Evolution process in GA.

GAs are widely used to tune the parameters of controllers. In [70], GA is used to tune the parameters of fuzzy based motor controller. In [71], FLC is utilized by GA tuning for steering control of underwater vehicle. PID controller is tuned by GA in [72]. Apart from them, GA optimization is utilized in the feedback of the controller in the present study. The designed GA is named as Real-time Multi-step GA (RM-GA) in this study and explained in detail in Chapter 3.

2.4 Robustness against Disturbances

The underwater world gives complexity to underwater vehicle operation due to disturbances [14] rather than the space environment. Because the proposed system is a vision-based system, not only the physical disturbances of ocean currents but also the noise in recognized images should be considered in the experiments. By completing the experimental tasks while including these considerations, the proposed docking system demonstrates its effectiveness against different disturbances. The common disturbances for visual servoing are as follows:

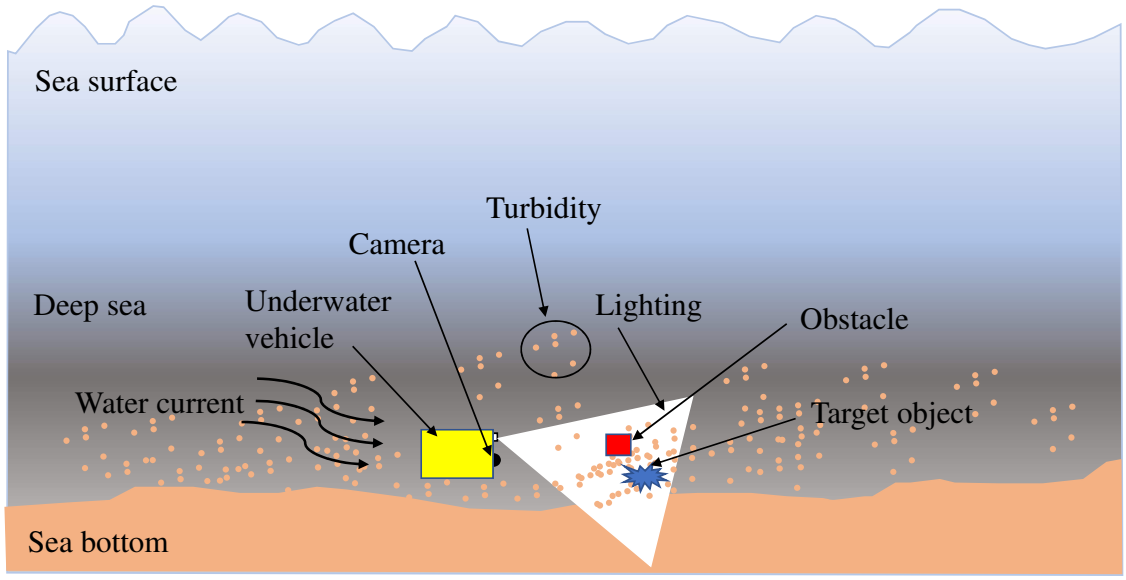


Fig. 2.9: Visual servoing in a deep-sea environment with disturbances such as current wave, turbidity, illumination (natural light and vehicle's light), and obstacle (such as fish).

2.4.1 Water Current

The ocean current can make the motion of the vehicle as well as motion of imaging while visual servoing. Since the vehicle's lighting unit also moves along with the vehicle, the illumination of variation can disturb recognition of the target with the dynamic images. Therefore, visual servoing with sensing unit and control unit has to be robust against the water current. The robustness of the proposed system against physical disturbances is

verified in this study.

2.4.2 Occlusion

Since there are many living things such as fish and particles such as seaweed in the sea, the target object can be partially occluded. Therefore, it is necessary to verify the robustness of the proposed system against occlusion. To fulfill this objective, some experiments were conducted and discussed in this study.

2.4.3 Turbidity and illumination variation

When a vision-based underwater vehicle approaches the sea bottom, a combination of water turbidity and fluctuations in the lighting direction produces artifacts in camera images and acts as a disturbance for visual servoing, as shown in Fig. 2.9. To the best of the author' knowledge, there have been no studies on the 3D pose estimation of underwater vehicles under turbid conditions. The detection of points of interest in turbid underwater images has been investigated using a collection of images acquired by a trinocular camera system under gradually increasing turbidity [83]. In [84], the robustness of local feature detection in underwater images was analyzed using a new dataset called TURBID, which consists of real seabed images with different amounts of degradation. Additionally, methods of underwater image quality assessment, visibility enhancement, and disparity computation under turbid conditions have been proposed in [85]. None of the above studies consider image recognition in real-time dynamic images, which is an indispensable technology for visual servoing in underwater vehicle docking. In this study, the robustness of the proposed system is verified against different disturbances including turbidity and illumination variation.

Chapter 3

3D MoS system with stereo-vision based real-time 3D pose estimation

In this section, a proposed 3D Moving on Sensing (MoS) system with stereo-vision based real-time 3D pose estimation for underwater docking is explained in detail including 3D model based matching method, kinematics of stereo vision, projection matrix and fitness function. First, how the relative pose of a 3D marker based on the vehicle coordinate system is estimated by 3D model based matching method is described. Second, the kinematics of stereo vision system including the robot (underwater vehicle) and a 3D marker is presented. Then, the basic concept of projection matrix is described. After that, the fitness function that is used as evaluation function in the pose estimation process is presented with the designed fitness function for this study. Finally, the main contribution of RM-GA that is a novel pose estimation method is described with the detailed discussion.

3.1 3D MoS

A robotic system that uses three dimensional measurement with solid object recognition and target tracking is based on visual servoing technology has been developed and named as 3D-MoS (Three Dimensional Move on Sensing). Figure 3.1 shows a 3D-MoS robotic

system. The robot estimates the real time relative pose of the target using stereo vision. According to the estimated pose, the robot moves to the target sensing the free space between it and the target as shown in Fig.3.1.

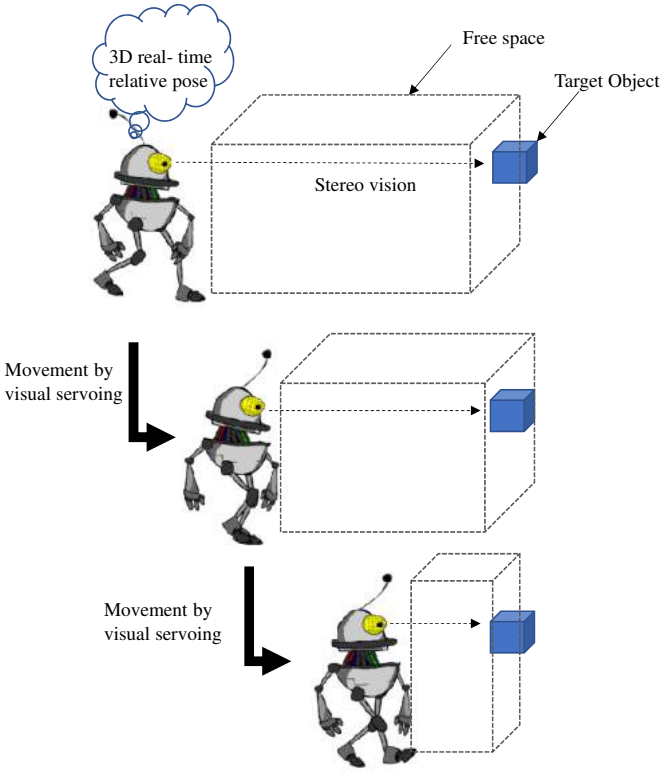


Fig. 3.1: A 3D MoS based robotic system in which the free space is estimated for every movement by sensing the relative pose using stereo vision.

3.2 3D Model-based matching using stereo-vision

Apart from image-based visual servoing, position-based visual servoing has been developed for the vision-based docking approach proposed by our research group. Instead of localizing the vehicle and target in an absolute pose in world coordinates to address the requirement of measurements using sensors such as GNSS and INS, localizing the vehicle relative to the target through recognition with a known target's information is implemented in feedback control using standalone dual cameras and a 3D marker.

A model-based matching method was used to recognize the 3D marker and estimate its pose in real time using stereo vision. Figure 3.2 shows the 3D marker coordinate system

Σ_M, the ith model coordinate system Σ_{M_i}, the left and right camera coordinate systems Σ_{CL} and Σ_{CR}, and the left and right image coordinate systems Σ_{IL} and Σ_{IR}. The origins of Σ_M and Σ_{M_i} are the intersections of the three lines perpendicular to the faces to which the spheres are attached. The target is a 3D marker as shown in Fig. 3.3 that consists of three spheres of 40 mm in diameter colored red, green, and blue.

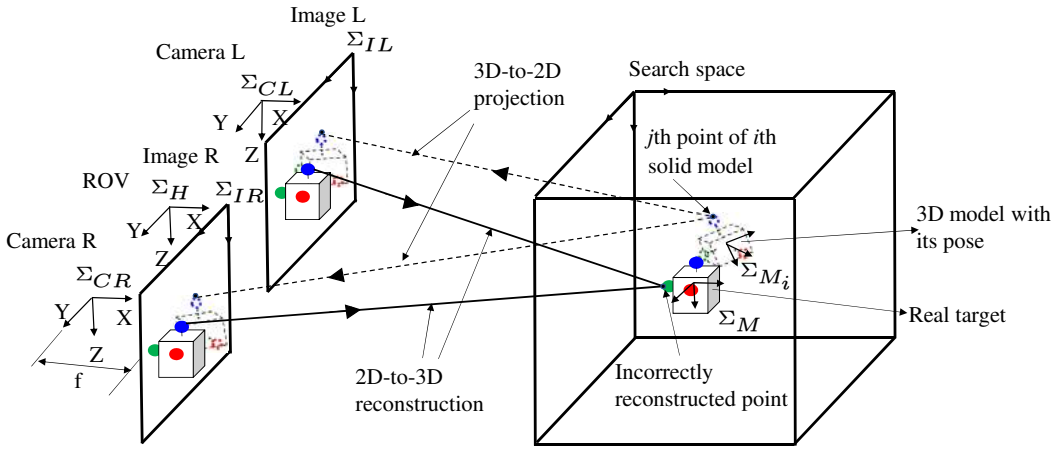


Fig. 3.2: Model-based pose estimation using the dual-eye vision system in the coordinate systems of the left and right cameras, the object (solid object), and the model (represented by a dotted box and dotted spheres). The jth point on the model in 3D space can be described in each coordinate system using these coordinates and homogeneous transformations. Similarly, a 3D model with its pose defined as a group of points in 3D space is projected onto the left and right cameras images through 3D-to-2D projection.

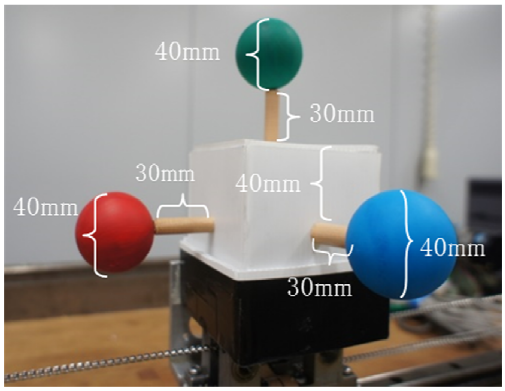


Fig. 3.3: 3D Marker that consists of three spheres which color of each are red, green and blue. The selected color RGB are chosen based on their distance in hue space and according to the experimental verification.

In conventional approaches, object recognition including relative pose information is implemented by feature-based recognition using 2D-to-3D reconstruction calculations, in

which the information of the target object is determined from a set of points in different images, generally using epipolar geometry. The main challenge in this type of approach is ensuring that points are correctly mapped. If a point in one image is incorrectly mapped to a point in another image, the pose of the reconstructed object does not represent that of the real 3D object. Figure 3.2 shows the incorrect reconstruction of a point through 2D-to-3D reconstruction resulting from incorrect mapping. To avoid incorrect mapping, which results from the original problem being ill posed, a model-based pose estimation approach based on 3D-to-2D projection was applied in this study because the forward projection from 3D to 2D generates unique points in 2D images without any errors, meaning incorrect mapping is avoided. As shown in Fig. 3.2, the jth point on the ith model in 3D space is projected onto the left and right camera images correctly. With this approach, 3D model-based recognition is implemented in this study.

3.3 Kinematics of stereo-vision

Here is a description of the kinematics of stereo-vision before the explanation of proposed system in detail. Figure 3.4 shows a perspective projection of the dual-eyes vision system. The coordinate systems of dual-eyes cameras and the target object (3D marker) in Fig. 3.4 consist of jth model coordinate system Σ_{M_j}, vehicle coordinate system Σ_H, camera coordinate systems as Σ_{CL} and Σ_{CR}, and image coordinate systems as Σ_{IL} and Σ_{IR}.

In Fig. 3.4, the position vectors of an arbitrary ith point of the jth 3D model Σ_{M_j} based on each coordinate system are as follows:

- ${}^{M}\boldsymbol{r}_i^j$: position of an arbitrary ith point on jth 3D model in Σ_{M_j}, where ${}^{M}\boldsymbol{r}_i^j$ is constant vector
- ${}^{CR}\boldsymbol{r}_i^j$ and ${}^{CL}\boldsymbol{r}_i^j$: position of an arbitrary ith point on jth 3D model based on Σ_{CR} and Σ_{CL}
- ${}^{IL}\boldsymbol{r}_i^j$ and ${}^{IR}\boldsymbol{r}_i^j$: projected position of an arbitrary ith point on jth 3D model based

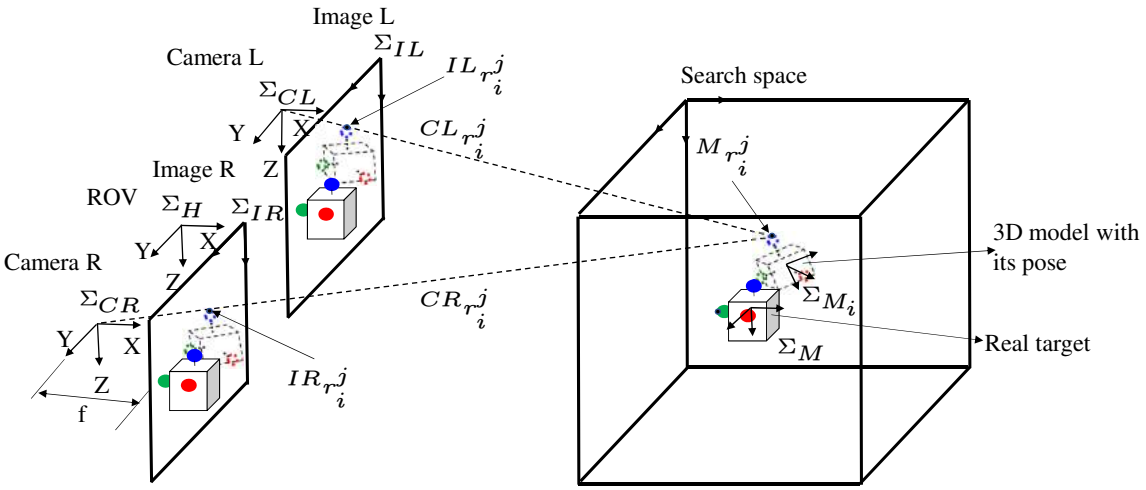


Fig. 3.4: Perspective projection of dual-eyes vision-system: In the searching area, a 3D solid model is represented by dotted point (jth photo-model). The coordinate systems of photo-model, camera and image are represented by Σ_{M_j}, Σ_{CL}, Σ_{CR}, Σ_{IL} and Σ_{IR} respectively. A 3D solid model that is assumed to be in the searching area is projected from 3D space to 2D left and right camera images.

on Σ_{IL} and Σ_{IR}

The homogeneous transformation matrix from the right camera coordinate system Σ_{CR} to the target object coordinate system Σ_M is defined as ${}^{CR}\boldsymbol{T}_M(\boldsymbol{\phi}_M^j)$, where $\boldsymbol{\phi}_M^j$ is j-th model's pose. Then, ${}^{CR}\boldsymbol{r}_i^j$ can be calculated by using Eq. (3.1),

$$ {}^{CR}\boldsymbol{r}_i^j = {}^{CR}\boldsymbol{T}_M(\boldsymbol{\phi}_M^j)\ {}^{M}\boldsymbol{r}_i^j. \tag{3.1} $$

where ${}^{M}\boldsymbol{r}_i^j$ is predetermined as fixed vectors since Σ_{M_j} is fixed on the jth model. ${}^{CL}\boldsymbol{r}_i^j$ that represents the same ith point on jth model based on Σ_{CL} is also calculated by using ${}^{CL}\boldsymbol{T}_M(\boldsymbol{\phi}_M^j)$. Equation (3.2) represents the projection transformation matrix $\boldsymbol{P}_C$. The projection transformation matrix is explained in next section 3.2.2.

$$ \boldsymbol{P}_C = \frac{1}{{}^{C}z_i} \begin{bmatrix} \frac{f}{\eta_x} & 0 & {}^{I}x_0 & 0 \\ 0 & \frac{f}{\eta_y} & {}^{I}y_0 & 0 \end{bmatrix}. \tag{3.2} $$

where,

- $C = CL, CR$,
- ${}^{C}z_i$; position of the ith point in the camera sight direction in Σ_{CR} and Σ_{CL} (See Fig. 3.5),
- f ; focal length,
- $\eta_x; [mm/pixel]$ in x-axis,
- $\eta_y; [mm/pixel]$ in y-axis.

The position vector of the ith point in the right and left camera image coordinates ${}^{IR}\boldsymbol{r}_i^j$ can be described by using $\boldsymbol{P}_C$ as,

$$ {}^{IR}\boldsymbol{r}_i^j = \boldsymbol{P}_C \, {}^{CR}\boldsymbol{r}_i^j = \boldsymbol{P}_C {}^{CR}\boldsymbol{T}_M(\boldsymbol{\phi}_M^j) {}^{M}\boldsymbol{r}_i^j \tag{3.3} $$

${}^{IL}\boldsymbol{r}_i^j$ can also be described as the same manner like ${}^{IR}\boldsymbol{r}_i^j$,

$$ {}^{IL}\boldsymbol{r}_i^j = \boldsymbol{P}_C \, {}^{CL}\boldsymbol{r}_i^j = \boldsymbol{P}_C {}^{CL}\boldsymbol{T}_M(\boldsymbol{\phi}_M^j) {}^{M}\boldsymbol{r}_i^j \tag{3.4} $$

Then, ${}^{IR}\boldsymbol{r}_i^j$, ${}^{IL}\boldsymbol{r}_i^j$ can be described as,

$$ \begin{cases} {}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j) = \boldsymbol{f}_R(\boldsymbol{\phi}_M^j, {}^{M}\boldsymbol{r}_i^j) \\ {}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j) = \boldsymbol{f}_L(\boldsymbol{\phi}_M^j, {}^{M}\boldsymbol{r}_i^j) \end{cases} \tag{3.5} $$

${}^{CR}\boldsymbol{r}_i^j$ and $\boldsymbol{P}_C$ in Eq. 3.3 are derived in detail in section 3.3.1 and section 3.3.2 respectively.

3.3.1 Homogeneous transformation matrix

${}^{CR}\boldsymbol{r}_i^j$ that consists of the homogeneous transformation matrix, ${}^{CR}\boldsymbol{T}_M$ in Eq.3.3 will be explained in this section. This matrix represents the relation between the right camera image coordinates Σ_{CR} and the model coordinate system Σ_M including the ROV coordinate system Σ_H as shown in Fig.3.4.

An arbitrary ith point on the target object defined on the model in Σ_{CR} (right camera) is as Eq.3.6.

$$ {}^{CR}\boldsymbol{r}_i^j = {}^{CR}\boldsymbol{T}_M \, {}^{M}\boldsymbol{r}_i^j. \tag{3.6} $$

$$ {}^{CR}\boldsymbol{T}_M {}^{M}\boldsymbol{r}_i^j = {}^{CR}\boldsymbol{T}_H \, {}^{H}\boldsymbol{r}_i^j. \tag{3.7} $$

$$ {}^{CR}\boldsymbol{r}_i^j = {}^{CR}\boldsymbol{T}_H \, {}^{H}\boldsymbol{r}_i^j. \tag{3.8} $$

where

- ${}^{CR}\boldsymbol{T}_M$:Homogeneous transformation matrix from Σ_{CR} to Σ_M
- ${}^{CR}\boldsymbol{T}_H$:Homogeneous transformation matrix from Σ_{CR} to Σ_H
- ${}^{H}\boldsymbol{r}_i$:The object is viewed from the search point ith on the model in Σ_H

According to the inverse homogeneous transformation matrix,

$$ \boldsymbol{T}\boldsymbol{T}^{-1} = I_4 \tag{3.9} $$

$$ \boldsymbol{T}^{-1} = \begin{bmatrix} \boldsymbol{R}^T & & & -\boldsymbol{R}^T\boldsymbol{r} \\ & & & \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.10} $$

$$
\begin{aligned}
({}^{H}\boldsymbol{T}_{CR})^{-1} &= {}^{CR}\boldsymbol{T}_{H} \\
&= \begin{bmatrix} ({}^{H}\boldsymbol{R}_{CR}{}^{T}) & & & -({}^{H}\boldsymbol{R}_{CR})^{T}\ {}^{H}\boldsymbol{r}_{CR} \\ & & & \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{3.11}
$$

${}^{H}\boldsymbol{T}_{CR}{}^{-1}$ is substituted ${}^{CR}\boldsymbol{T}_{H}$ in Eq.3.8 and also ${}^{H}\boldsymbol{T}_{M}\ {}^{M}\boldsymbol{r}_{i}$ is substituted ${}^{H}\boldsymbol{r}_{i}$ in Eq.3.8 then Eq.3.12 becomes as follow;

$$
{}^{CR}\boldsymbol{r}_{i}^{j} = {}^{H}\boldsymbol{T}_{CR}{}^{-1}\ {}^{H}\boldsymbol{T}_{M}\ {}^{M}\boldsymbol{r}_{i}^{j}. \tag{3.12}
$$

$$
{}^{H}\boldsymbol{T}_{M} = \begin{bmatrix} {}^{H}\boldsymbol{R}_{M}(\epsilon_1, \epsilon_2, \epsilon_3) & & & {}^{H}x_{M} \\ & & & {}^{H}y_{M} \\ & & & {}^{H}z_{M} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.13}
$$

$$
{}^{M}\boldsymbol{r}_{i}^{j} = \begin{bmatrix} {}^{M}x_{i}^{j} \\ {}^{M}y_{i}^{j} \\ {}^{M}z_{i}^{j} \\ 1 \end{bmatrix} \tag{3.14}
$$

Using Eqs.3.11, 3.13, 3.14, Eq. 3.12 can be derived as follow:

$$
{}^{CR}\boldsymbol{r}_i^j = \begin{bmatrix} ({}^H\boldsymbol{R}_{CR}{}^T) & & & -({}^H\boldsymbol{R}_{CR})^T\ {}^H\boldsymbol{r}_{CR} \\ & & & \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^H\boldsymbol{R}_M(\epsilon_1,\epsilon_2,\epsilon_3) & & & {}^Hx_M \\ & & & {}^Hy_M \\ & & & {}^Hz_M \\ 0 & & 0\ 0 & 1 \end{bmatrix} \begin{bmatrix} {}^Mx_i^j \\ {}^My_i^j \\ {}^Mz_i^j \\ 1 \end{bmatrix} \tag{3.15}
$$

Similarly, ${}^{CL}\boldsymbol{r}_i^j$: an arbitrary ith point on the target object defined on the model in Σ_{CL} (left camera) is received by doing same procedure. Then, ${}^{CL}\boldsymbol{r}_i^j$ become as Eq.3.16 and Eq. 3.17.

$$
{}^{CL}\boldsymbol{r}_i^j = {}^{CL}\boldsymbol{T}_M\ {}^M\boldsymbol{r}_i^j. \tag{3.16}
$$

$$
{}^{CL}\boldsymbol{r}_i^j = \begin{bmatrix} ({}^H\boldsymbol{R}_{CL}{}^T) & & & -({}^H\boldsymbol{R}_{CL})^T\ {}^H\boldsymbol{r}_{CL} \\ & & & \\ & & & \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} {}^H\boldsymbol{R}_M(\epsilon_1,\epsilon_2,\epsilon_3) & & & {}^Hx_M \\ & & & {}^Hy_M \\ & & & {}^Hz_M \\ 0 & & 0\ 0 & 1 \end{bmatrix} \begin{bmatrix} {}^Mx_i^j \\ {}^My_i^j \\ {}^Mz_i^j \\ 1 \end{bmatrix} \tag{3.17}
$$

3.3.2 Projection matrix

In this section, how the Projection transformation matrix, $\boldsymbol{P}$ is derived will be discussed in detail. Figure 3.5 shows the projection of a point from 3D space to 2D images. Since the explanation of projection in this section is for a camera, the projection transformation matrix $\boldsymbol{P}$ here will stand for $\boldsymbol{P}_C$ that is used in previous section.

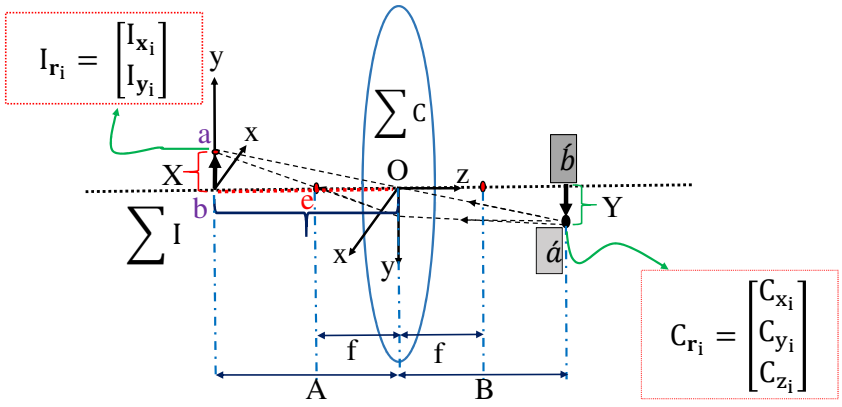


Fig. 3.5: Projection Matrix.

In Fig. 3.5, a point is situated in front of the camera len Σ_C with the position of ${}^{C}x_i, {}^{C}y_i, {}^{C}z_i$. The corresponding point will appear at the position ${}^{I}x_i, {}^{I}y_i$ in image coordinate plane according to the projection matrix. The projection matrix can be derived as follow:

From Fig.3.5, the Δ oab and the Δ o$\acute{a}\acute{b}$, we get the ratio as the following;

$$ab : \acute{a}\acute{b} = bo : \acute{b}o \tag{3.18}$$

$$X : Y = A : B \tag{3.19}$$

$$\frac{X}{Y} = \frac{A}{B} \tag{3.20}$$

By using this equations, the position of the object can translate 3D to 2D by the

projection,

$$\frac{({}^{I}y_i - {}^{I}y_0)\eta_y}{{}^{C}y_i} = \frac{e}{{}^{C}z_i} \tag{3.21}$$

$$\frac{({}^{I}x_i - {}^{I}x_0)\eta_x}{{}^{C}x_i} = \frac{e}{{}^{C}z_i} \tag{3.22}$$

(e = f) is assumed because of small value, and then Eq.3.23 becomes as follow;

$$\begin{bmatrix} {}^{I}x_i \\ {}^{I}y_i \end{bmatrix} = \frac{1}{{}^{C}z_i} \begin{bmatrix} \frac{f}{\eta_x} & 0 & {}^{I}x_0 & 0 \\ 0 & \frac{f}{\eta_y} & {}^{I}y_0 & 0 \end{bmatrix} \begin{bmatrix} {}^{C}x_i \\ {}^{C}y_i \\ {}^{C}z_i \\ 1 \end{bmatrix} \tag{3.23}$$

$$\boldsymbol{P} = \frac{1}{{}^{C}z_i} \begin{bmatrix} \frac{f}{\eta_x} & 0 & {}^{I}x_0 & 0 \\ 0 & \frac{f}{\eta_y} & {}^{I}y_0 & 0 \end{bmatrix}. \tag{3.24}$$

where

- $\begin{bmatrix} {}^{I}x_i \\ {}^{I}y_i \end{bmatrix}$: coordinate of the position of the image
- $\begin{bmatrix} {}^{I}x_0 \\ {}^{I}y_0 \end{bmatrix}$: coordinate of central position of the image
- e :distance of two coordinates of origin
- f :distance of two coordinates of origin is approximately equal to the distance of focal length
- $\eta_x = \eta_y = \frac{mm}{pixel}$
- ${}^{C}z_i$: distance between origin and camera image

3.4 Fitness function

The fitness function was designed to be used as an evaluation parameter in the pose estimation process. It is defined as the correlation between a projected model and a real target in the image. The highest peak of the fitness distribution is at the value equal to the true pose of the 3D marker. This can be stated in another way: the correlation function used for the fitness function and the target 3D model should be designed to have a dominant peak at the true pose of the target. In this subsection, the detailed design of fitness function will be explained in detail.

3.4.1 Design of fitness function

In Fig. 3.6(a), the three solid circles and the three circles outlined with dotted lines represent the spheres on the real target and those on the jth model obtained from 3D-to-2D projection, respectively for the right camera image. The pose $\boldsymbol{\phi}_M^j$ of the 3D model is an unknown variable composed of six parameters (x, y, z, ϵ_1, ϵ_2, and ϵ_3) where the first three are position in Cartesian coordinate frame and the latter are orientation by unit quaternion avoiding singularity issues [30]. These six parameters are determined in the pose estimation process.

The 2D projection of each sphere in the model is divided into two regions, as shown by the dashed circles in Fig. 3.6(b). Instead of evaluating the positions of all of the points in the model, only selected points are considered, as shown in Fig. 3.6(b). When the jth model is projected onto the 2D images of the left and right cameras, the fitness value for that model is calculated. Portions of the target object that lie inside the inner and outer regions of each corresponding sphere of the projected model proportionally increase and decrease the fitness value, respectively. Therefore, the value of fitness function is maximum (ideal value is 1.667) when the pose of the model fits that of the target object depicted in the images of the left and right cameras.

The correlation between the projected models including a pose of $\boldsymbol{\phi}_M^j$ and captured

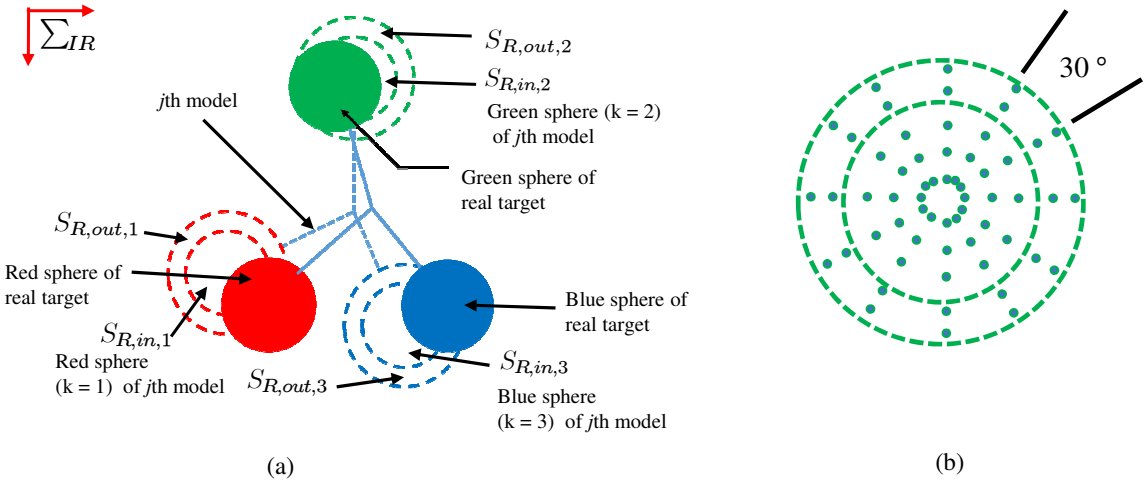


Fig. 3.6: (a) Real target (solid circles) and projected 3D model (circles with dashed outlines) in a 2D image obtained by the right camera. (b) Projection of the green sphere of a model with selected sample points. There are a total of 60 points (36 and 24 points in the inner and outer regions, respectively) in the projection, and the diameter of the inner region is same as that of the actual sphere. Note that k stands for each one of RGB; $k = 1$ for red color, $k = 2$ for green color, and $k = 3$ for blue color.

images with actual 3D marker that were projected on the left and right 2D searching areas is calculated by Eqs. (4.1) - (4.2). $F(\boldsymbol{\phi}_M^j)$ is calculated by averaging the fitness functions of both left camera image $F_L(\boldsymbol{\phi}_M^j)$ and right camera image $F_R(\boldsymbol{\phi}_M^j)$ as shown in Eq. (4.1).

$$
\begin{aligned}
F(\boldsymbol{\phi}_M^j) &= \frac{1}{\wedge}\sum_{k=1}^{m}\Bigg\{\Bigg(\sum_{\substack{{}^{IR}\boldsymbol{r}_i^j \in \\ S_{R,in,k}(\boldsymbol{\phi}_M^j)}} p_{R,in,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - \sum_{\substack{{}^{IR}\boldsymbol{r}_i^j \in \\ S_{R,out,k}(\boldsymbol{\phi}_M^j)}} p_{R,out,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))\Bigg) \\
&\quad + \Bigg(\sum_{\substack{{}^{IL}\boldsymbol{r}_i^j \in \\ S_{L,in,k}(\boldsymbol{\phi}_M^j)}} p_{L,in,k}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - \sum_{\substack{{}^{IL}\boldsymbol{r}_i^j \in \\ S_{L,out,k}(\boldsymbol{\phi}_M^j)}} p_{L,out,k}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))\Bigg)\Bigg\}/2 \\
&= \left\{F_R(\boldsymbol{\phi}_M^j) + F_L(\boldsymbol{\phi}_M^j)\right\}/2
\end{aligned}
\tag{4.1}
$$

$$p_{R,in,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = \begin{cases} 1, & \text{if}(|H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \leq l_k); \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

$$p_{R,out,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) = \begin{cases} 1, & \text{if}(|H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)) - H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))| \leq l_k); \\ 0, & \text{otherwise.} \end{cases} \tag{4.3}$$

where, $H_{ML}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j) = H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j) = b_k$ is the defined hue value of each color sphere and l_k is the defined hue range of each color. Please see the definition of b_k, l_k in Fig. 3.7. In this study, we defined b_k, l_k experimentally. Note that k stands for each one of RGB; $k = 1$ for red color, $k = 2$ for green color, and $k = 3$ for blue color as illustrated in Fig. 3.6 and Fig. 3.7.

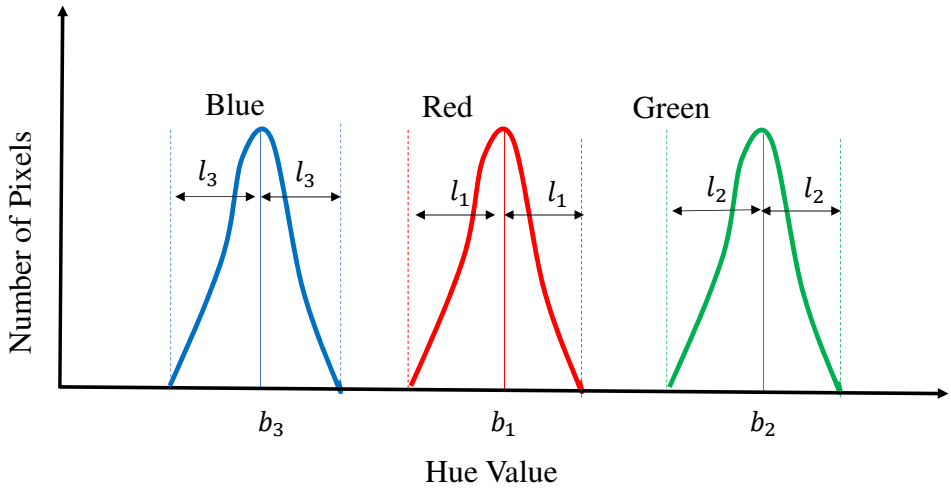


Fig. 3.7: Histogram of RGB in hue space with defined parameters: b_k is defined hue value of each color of RGB, and l_k is defined hue range of each color of RGB. Note that k stands for each one of RGB; $k = 1$ for red color, $k = 2$ for green color, and $k = 3$ for blue color.

Since the calculation of fitness function for left and right camera images are same, the explanation of the fitness function will be discussed using Eqs. (4.1) to (4.3) and Fig. 3.6 based on right camera images in this section. The projected points on $S_{R,in,k}(\boldsymbol{\phi}_M^j)$ and $S_{R,out,k}(\boldsymbol{\phi}_M^j)$ to the right camera image are described as ${}^{IR}\boldsymbol{r}_i^j \in S_{R,in,k}(\boldsymbol{\phi}_M^j)$ and

${}^{IR}\boldsymbol{r}_i^j \in S_{R,out,k}(\boldsymbol{\phi}_M^j)$ respectively. For detailed explanation of Eq. (4.1), the following definitions should be stated here.

- $S_{R,in,k}$; the inside area projected to right image plane,
- $S_{R,out,k}$; the space on a strip area surrounding $S_{R,in}$,
- $H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$; the hue value of the right camera image at the point ${}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)$,
- $H_{ML}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$; the hue value of the model at the point ${}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j)$ (i-th point on the j-th model),

The next Eqs. (4.2) and (4.3) is used for calculating $p_{R,in,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ and $p_{R,out,k}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ that are included in Eq. (4.1).

Equations (4.2) and (4.3) are designed to provide a peak in the fitness value distribution, $F(\boldsymbol{\phi}_M^j)$ when $\boldsymbol{\phi}_M^j$ coincides with the true pose of the target 3D marker. The evaluation values in the equations are tuned experimentally. In Eq. (4.2), if the hue value of each point of captured images, $H_{IR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$, which lies inside the surface model frame $S_{R,in,k}$, and the hue value of corresponding point in a model, $H_{MR}({}^{IR}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$, are same with a tolerance less than l_k that is 20 in this study, then the fitness value will increase with the voting value of "+1." The fitness value will decrease with the value of "−1" for every point of 3D marker in the right camera image that lie in the outer area of projected model. Similarly, functions $p_{L,in,k}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ and $p_{L,out,k}({}^{IL}\boldsymbol{r}_i^j(\boldsymbol{\phi}_M^j))$ follow suit.

3.4.2 Properties of fitness function

The properties of fitness function as shown in equations (4.2) and (4.3) are discussed in this section. A fitness function, which is a shape-based integration/differentiation calculation, is modeled to calculate the correlation between a model and images captured by two cameras using hue value of images. In other words, the intention of the designed fitness function is to have a dominant peak at the true pose of the target. Here is the brief

explanation on why designed fitness function is a shape-based integration/differentiation calculation. Since a model has spheres with quantitative diameters rather than a point, shape information is used when calculating the correlation between the model and the target object. Therefore, it is said to be a shape-based approach. As a group of image points that lie inside the inner area and outer area (see Fig. 3.6) of the projected model are evaluated together and added all together in the area, it is said to be in terms of integration. Integration operation can reduce the noise that appears in images like spike noise. To increase the sensitivity, differentiation operation is also considered in the construction of fitness function. The evaluation value is calculated by subtracting the values for points that lie in the outer area from the ones that are overlapped with the inner area of the model. It is therefore said to be in terms of differentiation. The total fitness value is calculated from averaging two fitness functions of the left and right camera. Please note that there is no individual evaluation of left and right images. Finally, the fitness function will have a maximum value when the pose of the searching model fits the one of the target object being imaged in the right and left cameras' images. The evaluation parameters of the objective function (that is fitness function in this study) are designed to reduce the noise (noise in here means some peak points that represent incorrect poses of the target).

The fitness distribution with respect to a position in the XY plane based on Σ_H (see Fig.3.4) is illustrated in Fig. 3.8. Because the pose of the target is composed of six parameters (three for position and three for orientation), the fitness distribution with a peak at the true pose can be seen in 3D space, including the fitness value and any pairs of dimensions of pose parameters, as shown in Fig. 3.8. In the plot in Fig. 3.8, there is a large peak that corresponds to the true pose, and some additional peaks that correspond to other incorrect poses are present. The proposed system can be considered robust as long as the highest peak of the fitness distribution represents the true pose and the effect of the noise that represents incorrect pose is significantly less than this peak. The shape of the fitness distribution will change in a dynamic image with a video rate of 30 fps.

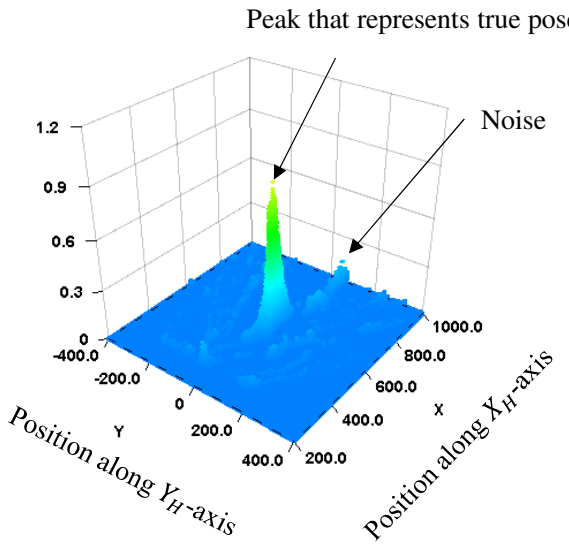


Fig. 3.8: Fitness distribution. The peak represents the true pose detected by the designed fitness function. The noise, which represents incorrect poses, is generated in the fitness distribution as a result of image deformation caused by environmental effects.

Please refer to [94], [98] for a detailed definition of the fitness function. The concept of fitness function in this study can be said to be an extension of the work in [98] in which different models including a rectangular shape surface-strips model were evaluated using images from a single camera. Since the designed fitness function make sure to exist a peak that represents the true pose in fitness distribution within search space, searching the peak within the video rate that is 33 ms is the main task for the proposed system. To solve this searching problem, we utilized the real-time multi-step GA as explained in next section.

3.5 Real-time Multi-step GA

In the proposed 3D model-based recognition method, searching for all possible models is time consuming for real-time recognition. Therefore, the problem of finding/recognizing the 3D marker and detecting its pose is converted into an optimization problem with a multi-peak distribution, which can be confirmed directly by calculating the distribution of a fitness function against the 3D pose [94], [98].

3.5.1 What is Real-time Multi-step GA?

The optimization should be completed in the changing condition, for example, dynamic images input by a video rate when visual feedback is required. In this situation, we have two options: (1) A simple optimization method that needs a short time to complete with appropriate accuracy and repeats the optimizing procedure, and (2) Sophisticated optimization methods that might provide better accuracy but require large calculation time. Here the question is which of these options is better for pose tracking in dynamic images. The first one, simple optimization, was chosen for the underwater docking experiment based on its simple logic and effectiveness. The following is one of the reasons for choosing (1). Fast evolution due to shorter life spans, such as mouse evolution, can enable an animal to adapt itself to a changing environment faster than slower evolution of animals such as an elephant, which can live for several decades with fewer chances to adapt. For example, support that the life span of a mouse is 1 year and that of an elephant is about 80 years. Therefore, the mouse has 80 times more chances in the time domain than does the elephant to adapt to a changing environment.

The discussions on optimization performance in other studies such as [100]-[102] are based on the speed measured by iteration times. We think that performance evaluation of different methods based on iteration times is unfair because operation time for one iteration of each method in different systems may not be the same. For example, one method may take one hour for one iteration while another method may finish one iteration within one minute. It is difficult to find performance comparison in time domain. According to authors' knowledge, our work is the first one to measure the optimization performance in time domain. On the other hand, other discussions in [103],[104] are concentrated on finding out valuables' numbers to give a maximum function number that are used for the controller's parameters for tuning up the system's performances. In contrast, optimization is directly used in a feedback of the control system in the present paper.

We have developed Real time Multi-step GA, formerly known as the 1-step GA, which

can deal with non-differential distribution with a multi-peak, for this underwater docking experiment, although it may not be the best GA in comparison to other optimization methods. We did not compare GA with other optimization methods in this study. Real-time Multi-step GA evolves the chromosomes with as many generations as possible within the video frame rate for each image; in our experimental system, nine generations are possible. The practical performance of the Real-time Multi-step GA was confirmed in a previous work [93]. In [93], the Real-time Multi-step GA was used to estimate the pose of a fish in real time.

3.5.2 How does Real-time Multi-step GA work?

In the proposed Real-time Multi-step GA, each chromosome as shown in Fig. 3.9 encodes 12 bits for each of six parameters: three for position and three for orientation. Fig. 3.10 shows the 3D model-based recognition process in 3D space that evolves through the evaluations of chromosomes by forward projection from the 3D marker onto 2D images. The defined number of chromosomes that represent the different relative poses of the 3D model to the ROV in back-projection is initiated randomly, as shown in the first generation in Fig. 3.10. Fitter chromosomes as evaluated by a fitness function have a higher chance to be selected for the reproduction of offspring by using designed operators (selection, crossover, and mutation). After reaching the predefined number of generations, the chromosome that has the best fitness value is selected to represent the actual pose of the object. Because the main objective is real-time performance, termination of GA evolution is determined by the video frame rate, which is 30 fps in this work.

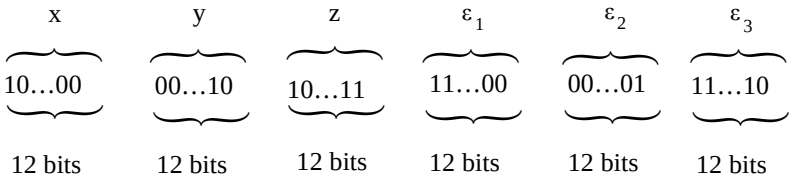


Fig. 3.9: Gene representing for position and orientation.

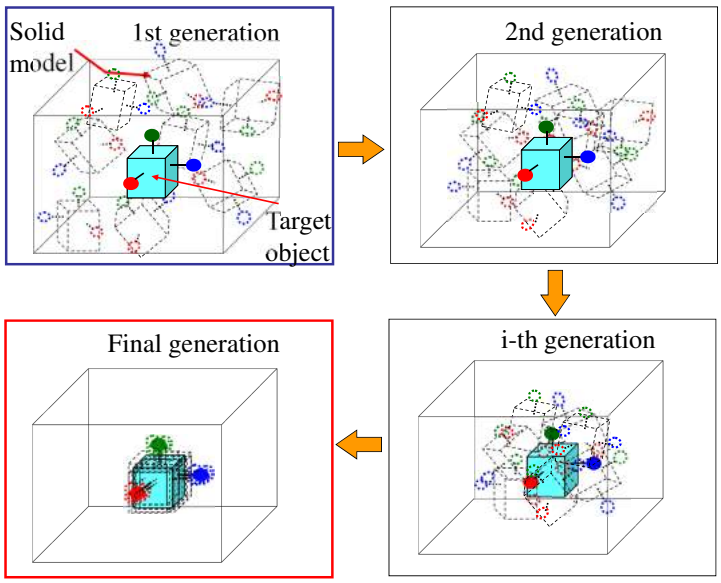


Fig. 3.10: 3D model-based recognition process.

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). As shown in Fig. 3.6, two spaces in the model object can have a scored fitness value: the inner space that is the same size as the target sphere and the other space that is the background area. The portion of the captured target that lies inside the inner area of the model will score a higher fitness value and the portion that lies inside the background area will score a lower value. Therefore, the fitness value is maximum when the poses of the target and the model are coincident. Note that the evolution of models to the real target is in 3D space and the evaluation between projected model and real target in terms of the fitness function is done in 2D images. The color information in hue space is used to search for the 3D marker in images, because hue space is less sensitive to the lighting condition [60]. The effectiveness of this method was confirmed in our previous research [95]-[97]. The time-convergence performance of the Real-time Multi-step GA as a dynamic evaluation function was approved mathematically by a Lyapunov analysis in [91].

Fig. 3.11 shows the flowchart of the Real-time Multi-step GA (right sub figure) and

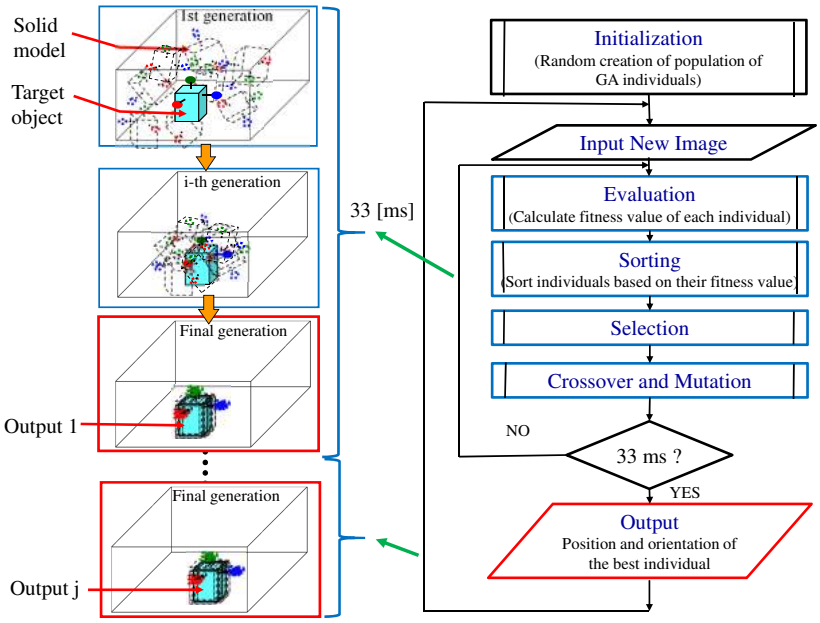


Fig. 3.11: Flowchart of Real-time Multi-step GA.

recognition process in 3D space (left sub figure). Please note that a solid model in 3D space represents a GA individual. GA operations such as Selection, Crossover and Mutation are performed to reproduce the next generation through evaluation by a fitness function (explained in previous section). Several solid models that represent different relative poses converge to the target object through GA evolution process within 33 ms as shown in Fig. 3.11 (left sub figure). The solid model (Output j in Fig.3.11) that represents the true pose with the highest fitness value is searched for every 33 ms. Then, these fit models are forwarded to the next step as the initial models for the next new images in real time.

3.5.3 Optimal Real-time Multi-step GA

The relative pose estimation by 3D model-based recognition is assumed to be executed in the GA search area set in front of the underwater robot, as shown in Fig. 3.12. We considered the visibility range in real seawater as about 1 m. Therefore, the searching space is defined as shown in TABLE 3.1. The searching space depends on the camera lens specification, which has a focal length of 2.9 [mm]. TABLE 3.1 shows the conditions of the GA.

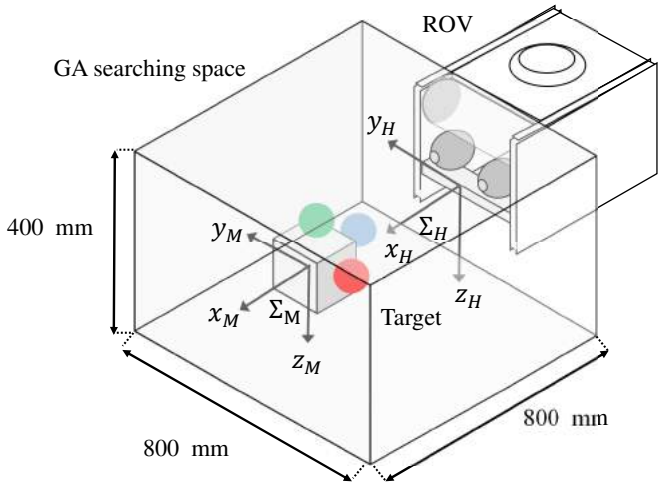


Fig. 3.12: Underwater target and GA searching space.

Table 3.1: Parameters of Real-time Multi-step GA.

Items	Specification
Number of genes	60
Evolved pose (position and orientation)	$(x, y, z, \epsilon_1, \epsilon_2, \epsilon_3)$, all genes are coded by binary 12 bits $(\epsilon_1, \epsilon_2, \epsilon_3)$ are represented by quaternion
Pose used for controlling (Position, Orientation)	Position (x mm, y mm, z mm) Orientation (ϵ_3) around z-axis of Σ_H in Fig.3.12
Searching space defined by Σ_H in Fig.3.12	{x,y,z,ϵ_3}={$\pm$400, $\pm$400, $\pm$200, $\pm$0.15 (equal to $\pm 17.3 degree$)}
Control period [ms]	33
Number of gene evolution [times/33ms]	9
Selection rate [%]	60
Mutation rate [%]	10
Crossover	Two-point
Evolution strategy	Elitism preservation

Chapter 4

3D pose estimation and visual servoing

4.1 3D Pose tracking using two cameras and 3D marker

The main task in the docking operation is 3D pose tracking in time by following dynamic images with a video rate of 30 frames per second. There is no study that has achieved real-time 3D pose estimation by using dual-eye cameras for AUV in docking operation in which only visual information is directly used in a feedback loop. Therefore, as a main contribution of this study, a new method of real-time 3D pose estimation in successively input dynamic images from two cameras as shown in Fig. 4.1 using 3D model-based recognition method utilizing Real-time Multi-step GA is verified by conducting some experiments.

The block diagram of the proposed control system is shown in Fig. 4.1. Visual information is used in the feedback as position-based visual servoing. The images acquired from the dual-eye cameras are sent to the PC in which real-time multi step GA is implemented. Then, the real-time recognition of the 3D pose of the target object by the Real-time Multi-step GA is executed by software installed in the PC. Based on the error between the target value and the recognized value, command signals generated from

calculating the voltage value gained by the controller for the thrusters are input into the ROV.

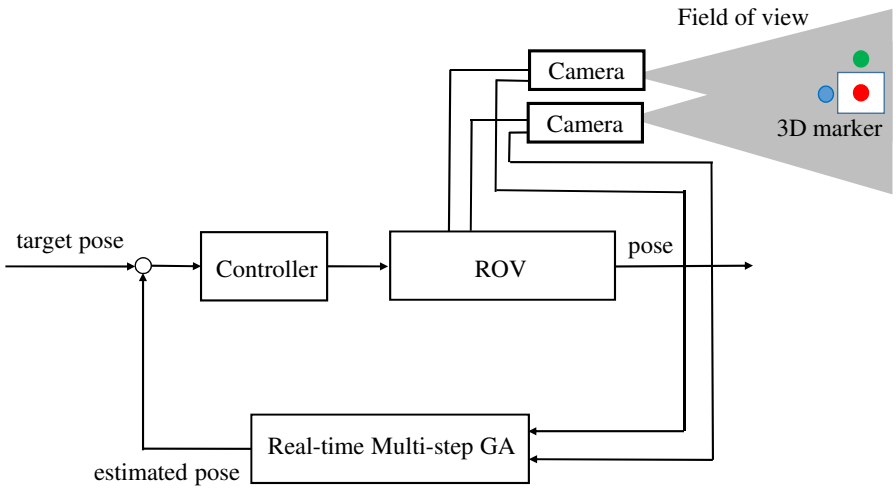


Fig. 4.1: Block diagram of the proposed system with Real-time Multi-step GA.

4.2 Experiment of Visual Servoing

Experiments were conducted in simulated environment in order to verify the effectiveness of the proposed visual servoing. Firstly, the experiment in which the underwater robot keeps the relative pose with fixed target, was conducted while setting the experimental conditions approving that the robot is regulated to the final pose against with the target object. Secondly, the robustness against the physical disturbance that were simulated as water current in the sea was verified while visual servoing. Finally, the experiments were conducted in the case when the target is seen partially.

4.2.1 Underwater Vehicle

Remotely controlled underwater robot used in this experiment (manufactured by KOWA, maximum depth 50m) is shown in Fig.4.3. Two fixed forward cameras with the same specification (imaging element CCD, pixel number 380,000 pixel, signal system NTSC, minimum Illumination 0.8 [lx], no zoom) are mounted on the ROV. The two fixed for-

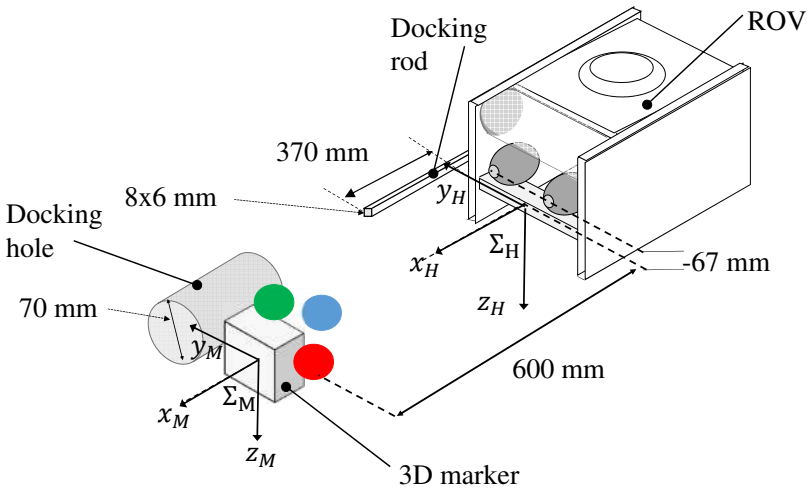


Fig. 4.2: Layout of the ROV with a docking rod and the 3D marker with a docking hole.

ward cameras are used for three-dimensional object recognition in visual servoing. In the thruster system of ROV, 2 horizontal thrusters with maximum thrust of 9.8 [N], 1 vertical thruster with maximum thrust of 4.9 [N] and 1 traverse thruster with 4.9 [N] are installed. In addition, it has been equipped with two units of LED lights (5.8W) for illumination source. The ROV used in this experiment is actuated in 4 degrees of freedom (DoF) (x, y, z and ϵ_3). The specifications of main hardware components are summarized in Table.4.1.

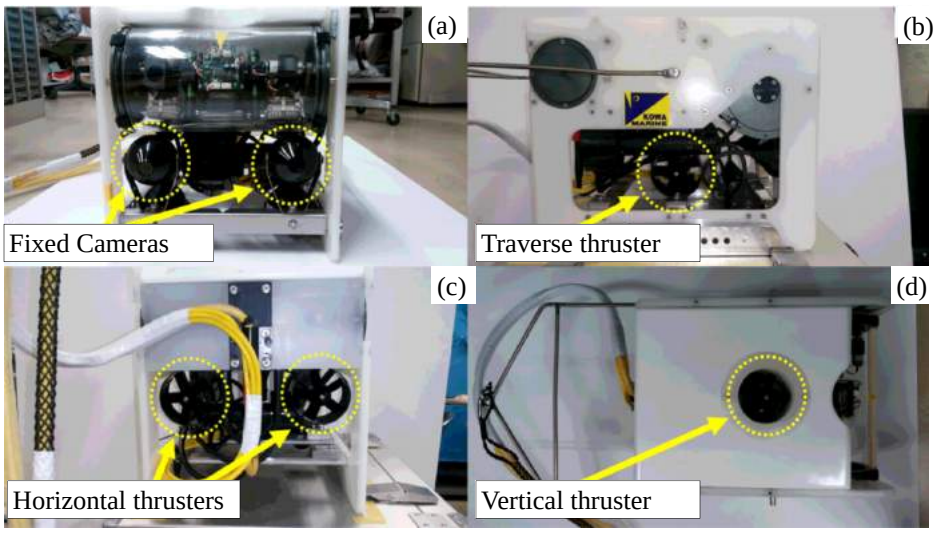


Fig. 4.3: Overview of ROV (a) front view (b) side view (c) top view (d) back view.

Table 4.1: Specification of ROV

Max: operating depth [m]	50
Dimension [mm]	280 (W) × 380 (L) × 310 (H)
Dry weight [kg]	15
Number of Thrusters	2 (Horizontal), 1 (Vertical), 1 (Traverse)
Number of Cameras	2 (Front, fixed), 2 (Downward, fixed), 1 (Tilting and zooming)
Number of LED lights	2 (5.8 [W])
Number of Line lasers	2 (2 [mW])
Tether cable [m]	200
Structial materials	Aluminum alloy and acrylate resin
Maximum thrust force [N]	9.8 (horizontal), 4.9 (vertical, Traverse)

4.2.2 Experimental Environment

A pool (length × width × height, 2 [m] × 3 [m] × 0.75 [m]) filled with tap water was used as an experimental tank for underwater vehicle experiments. ROV is tethered through an cable with 200m length to receive image information and control signals as shown in Fig.4.4. Based on the images which are given by binocular camera, the 3D information is calculated through Model-based matching method and Genetic Algorithm (GA). For physical disturbance to disturb the movement of the vehicle, abrupt external forces are applied to the vehicle by pushing the vehicle in different direction using a rod. In order to perform experiments to confirm the robustness of the system in case when the target is seen partially, one of the three spears is hidden for some times.

4.2.3 System Configuration

Adjustment of dead zone

For manual operation by joystick instructions of the controller in the remote-operated ROV, it has to have a certain amount of dead zone voltage that makes thrusters with no thrust in order to prevent malfunction due to the motion of a human finger. On the other hand, in this study, when approaching to the object by thruster propulsion, the realization of the movement as well as the attitude control performance of high accuracy

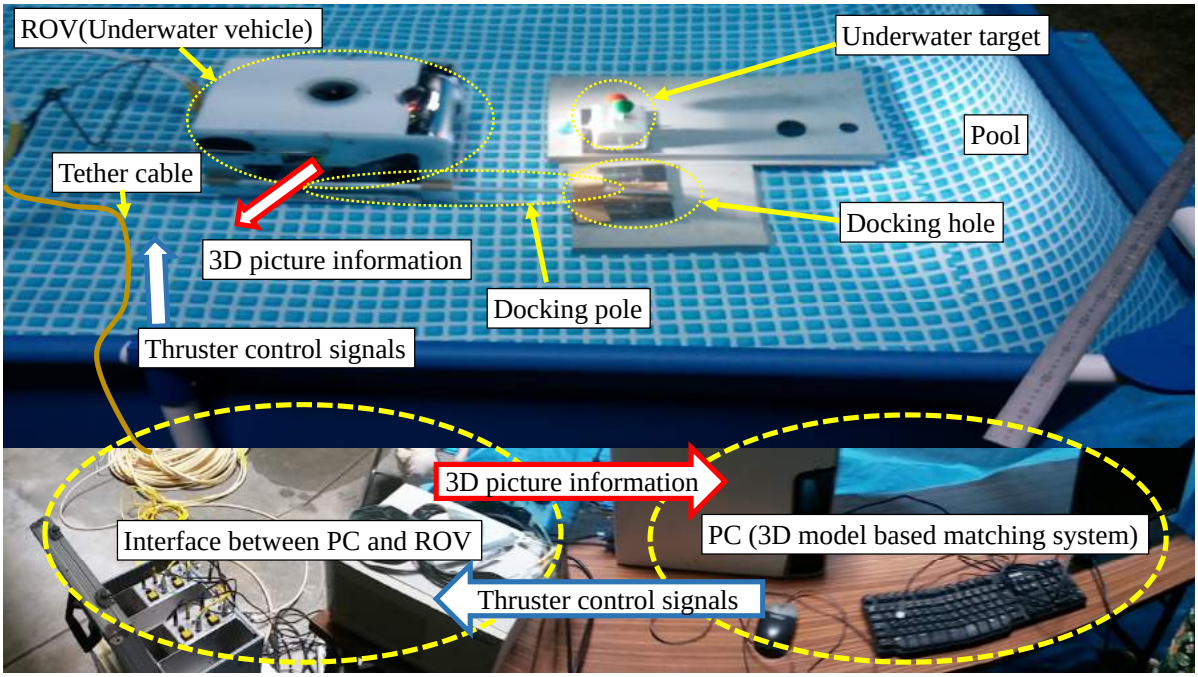


Fig. 4.4: Layout of underwater experimental devices.

is in mm. Therefore, the characteristic of thrust of each thruster that changes with respect to the dead zone in the control voltage can be easily configured by using formulation in thrust approximation. Specifically, the dead band characteristics of the ROV which was confirmed in preliminary experiments are removed by mean of linear approximation as a solid line and thrust was configured in the control software so as to generate.

Interfacing

Figure 4.5 shows the interfacing between proposed system implemented in PC and the camera mounted in the robot. As the resolution of pose is in 12 bits, Digital to Analog Converter with high resolution is used for precise control. Even though there is initial delay time about 9 ms in receiving dynamic video images with 33[ms], it does not come in picture as issues for real-time operation.

Adjustment of cable tension

As the experimental environment is indoor pool, there is space limitation for vehicle that is tethered through an cable. Therefore, cable tension may cause sometimes disturbance in controlling the vehicle that should be in 4DoF. However, careful adjustment of cable position eliminates these small issues while conducting experiments.

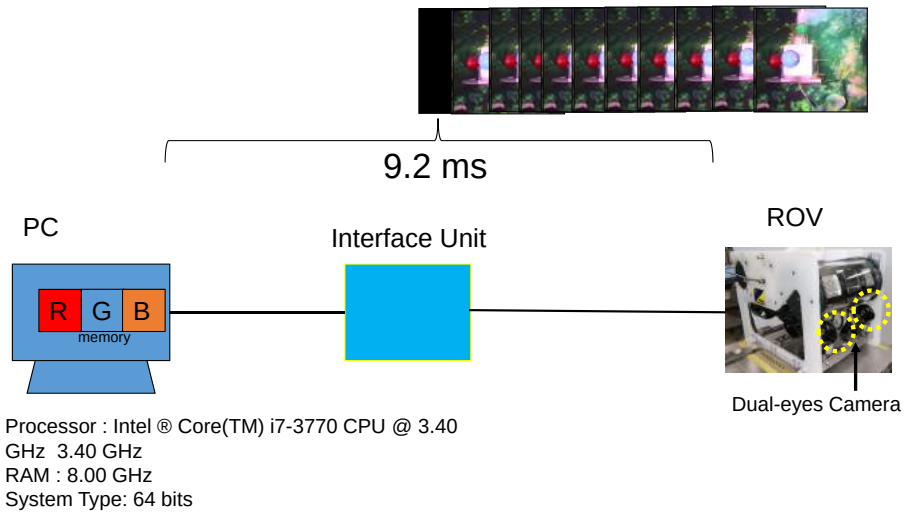


Fig. 4.5: Interface between robot and PC.

Desired pose setting

According to the range of camera for recognition and experimental environment scale, the desired pose is set as below. The negative distance in z-axis is the difference between the origin of the camera and vehicle frame Σ_H .

$$x_d = {}^H x_M = 600[mm] \ , \quad y_d = {}^H y_M = 0[mm],$$
$$z_d = {}^H z_M = -67[mm] \ , \quad \epsilon_{3d} = 0[deg]$$

where, x[mm], y[mm], z[mm], ϵ_3[deg] represent the position and orientation of the target object recognized by RM-GA. In order to regulate the underwater robot with this desired relative pose to the target, the command voltage value $v_1 \sim v_4$ fed to respective thrusters are calculated from P controller.

4.2.4 Controller

To eliminate the error in relative pose to the target, conventional P controller is used as shown in Equations (1)-(4). Even though the proposed system estimates all six variables of pose, torque around x-axis and y -axis are neglected in control system because the x-axis and y-axis rotations are naturally restored to zero by the restoration torque made by the z-axis distance between the center of buoyance and the one of gravity. Therefore, only four degree of freedom is considered in control system. ON-OFF control is used for control in left and right direction (y-axis) and the other three degree of freedom are controlled by p-controller. According to the recognition accuracy that is millimeter level and experimental results in movement of ROV, the threshold for ON-OFF control is defined to be 5 mm as shown in Equation (2). The proportional gain for each thruster is tuned according to the experimental results. The block diagram of the proposed control system is shown in Fig.4.6.

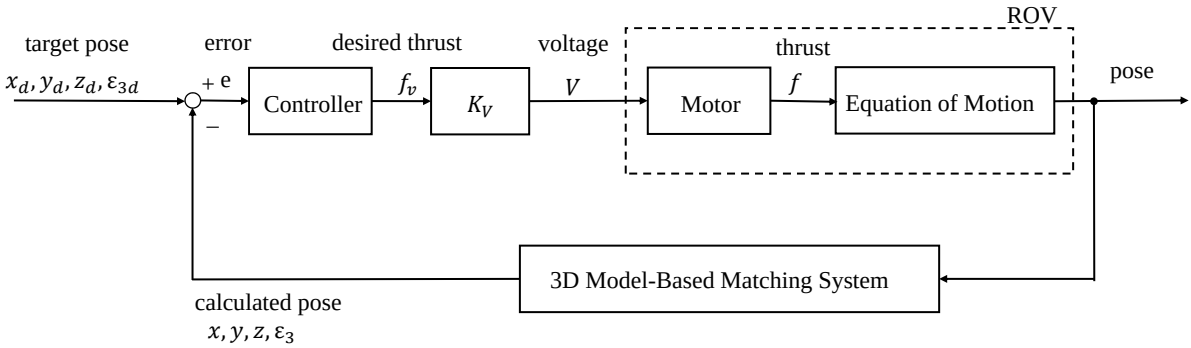


Fig. 4.6: Control logic for the proposed system.

$$\begin{aligned} \text{Back and Forth} \quad &: \quad v_1 = k_{p1}(x_d - x) + 2.5 \qquad (2.1) \\ \textit{direction} \quad &\quad (v_1 = 0\text{[V] for thrust } 9.8\text{[N] in } X_H \\ (X_H \, axis \, in \, Fig.4.2) \quad &\quad \text{of } \Sigma_H,\ v_1 = 5\text{[V] for } -9.8\text{[N]}) \end{aligned}$$

$$\text{Left and right : direction } (Y_H \, axis \, in \, Fig.4.2) \quad v_2 = \begin{cases} 5[V] \; ((y_d - y < -5[\text{mm}]) \\ \quad \text{for thrust in } Y_H \text{ of} \\ \quad \Sigma_H \text{ is } -4.9[\text{N}]) \\ 2.5[V] \; ((-5 \leq y_d - y \leq 5) \\ \quad \text{meaning thrust} \\ \quad \text{equals to zero)} \\ 0[V] \; ((y_d - y > 5[\text{mm}]) \\ \quad \text{for thrust in } Y_H \text{ of} \\ \quad \Sigma_H \text{ is } 4.9[\text{N}]) \end{cases} \tag{2.2}$$

$$\text{Rotation } (around \, Z_H \, axis \; in \, Fig.4.2) \; : \; v_3 = k_{p2}(\epsilon_{3d} - \epsilon_3) + 2.5 \quad (v_3 = 0[\text{V}] \text{ for } 0.882[\text{N}] \text{ in } Z_H \text{ of } \Sigma_H, \; v_3 = 5[\text{V}] \text{ for } -0.882[\text{N}]) \tag{2.3}$$

$$\text{Vertical direction } (Z_H \, axis \, in \, Fig.4.2) \; : \; v_4 = k_{p3}(z_d - z) + 2.5 \quad (v_4 = 0[\text{V}] \text{ for } -4.9[\text{N}] \text{ in } Z_H \text{ of } \Sigma_H, \; v_4 = 5[\text{V}] \text{ for } 4.9[\text{N}]) \tag{2.4}$$

where, v_1 is input voltage for each of two horizontal thrusters (shown in Fig.4.3 (c)) for movement of ROV in back and forth direction (X_H in Fig.4.2); v_2 is input voltage for traverse thruster (shown in Fig.4.3 (b)) for movement of ROV in right and left direction (Y_H in Fig.4.2); v_3 is input voltage for thrusters for rotation movement of ROV around z-axis (Z_H in Fig.4.2), and v_4 is input voltage for vertical thruster (shown in Fig.4.3 (d)) for movement of ROV in vertical direction (Z_H in Fig.4.2). Note that the rotation of vehicle is controlled by two horizontal thrusters that rotate in opposite direction.

4.3 Results and Discussion

4.3.1 Linearization of Dead Zone Voltage in Thrusters

To remove the dead zone in control voltage, the linearization of control voltage is performed according to the experimental adjustment as shown in Fig.4.7. Fig.4.7(a),(c),(e) shows the initial characteristics of thrust and torque control voltage observed in experimental measurements and (b),(d),(f) are their corresponding adjusted ones.

4.3.2 Regulating Performance

Fig. 4.8 (a) shows the time variation of the fitness value at the time of GA recognition of underwater robot that was regulated in $x_d = 600[mm]$, $y_d = 0[mm]$, $z_d = -67[mm]$, $\epsilon_{3d} = 0[deg]$. According to the experiment result, it can be seen that the fitness value is maintained above 0.8 within a few seconds from the recognition start. In general, when performing precise visual servoing, GA recognition accuracy is thought to be necessary 0.5 or more. It was confirmed that the object recognition accuracy in water using GA was almost the same degree of fitness in comparison with the case in the air. This result addresses huge benefit of reducing frequency of doing experiments in water for testing every advanced step in recognition process. Generally, it is difficult to confirm how much speed of moving target can be detected by the speed of evaluation of GA. However, two contributions in the proposed system provides promising performance considering this fact. The first one is that RM-GA forwards its best candidates to the next generations. On the other hand, the regulating control algorithm keeps field of view of both cameras after initial recognition. For example, when the relative pose makes the invisible of right camera for certain reasons such as the target is moving or the vehicle is physically disturbed after recognition, the control algorithm makes priority to turn left direction while visual servoing.

The regulating performance by mean of visual servoing is shown in Fig.4.8. Figure

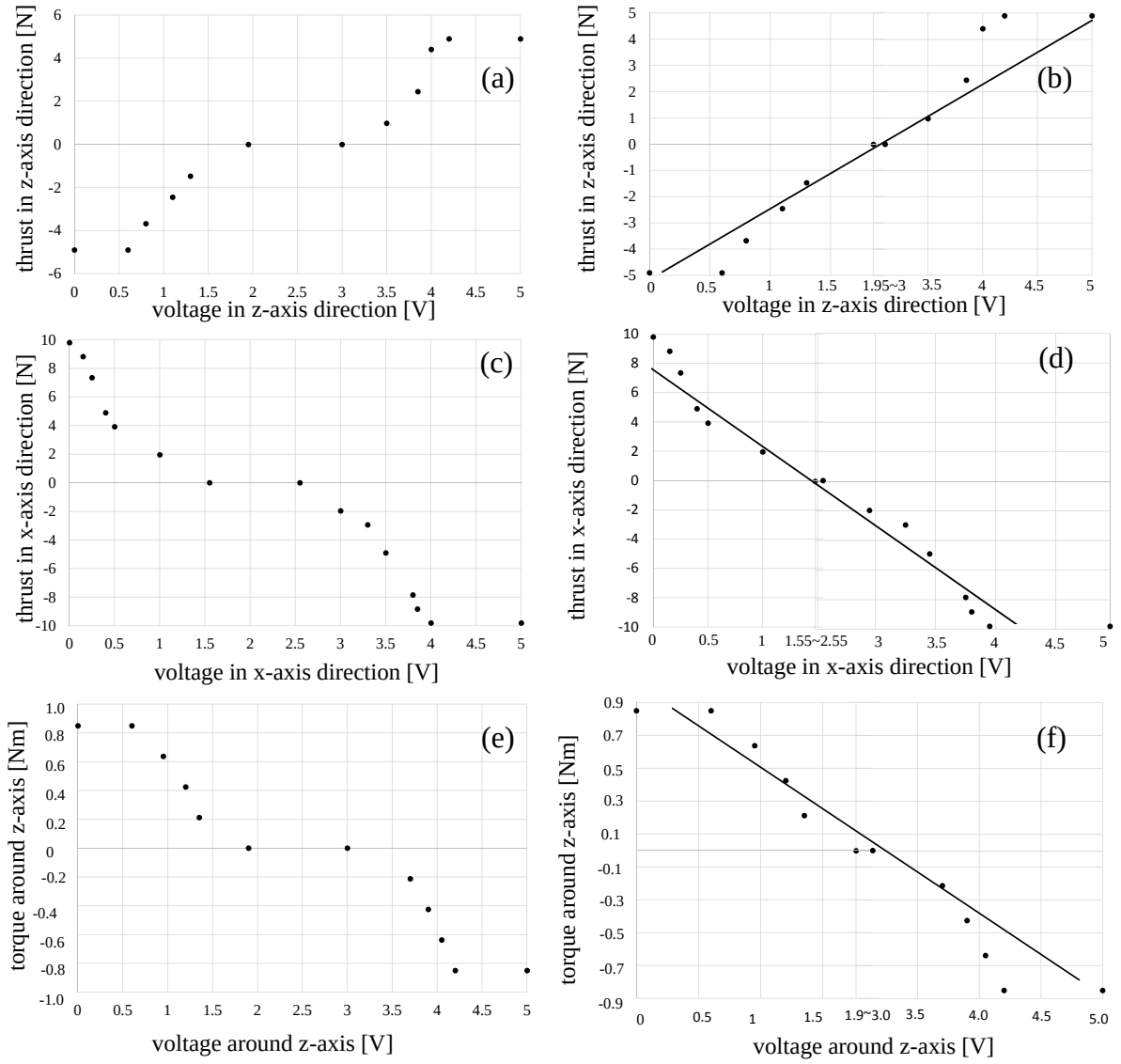


Fig. 4.7: Initial characteristics of thrust and torque control voltage and adjusted ones by removing dead-band and linearization: (a)initial characteristics in z-axis direction, (b)characteristics removing dead-band (solid line) and adjusted one (black dot) in z-axis direction, (c)(d) characteristics in x-axis direction and (e)(f) characteristics around z-axis.

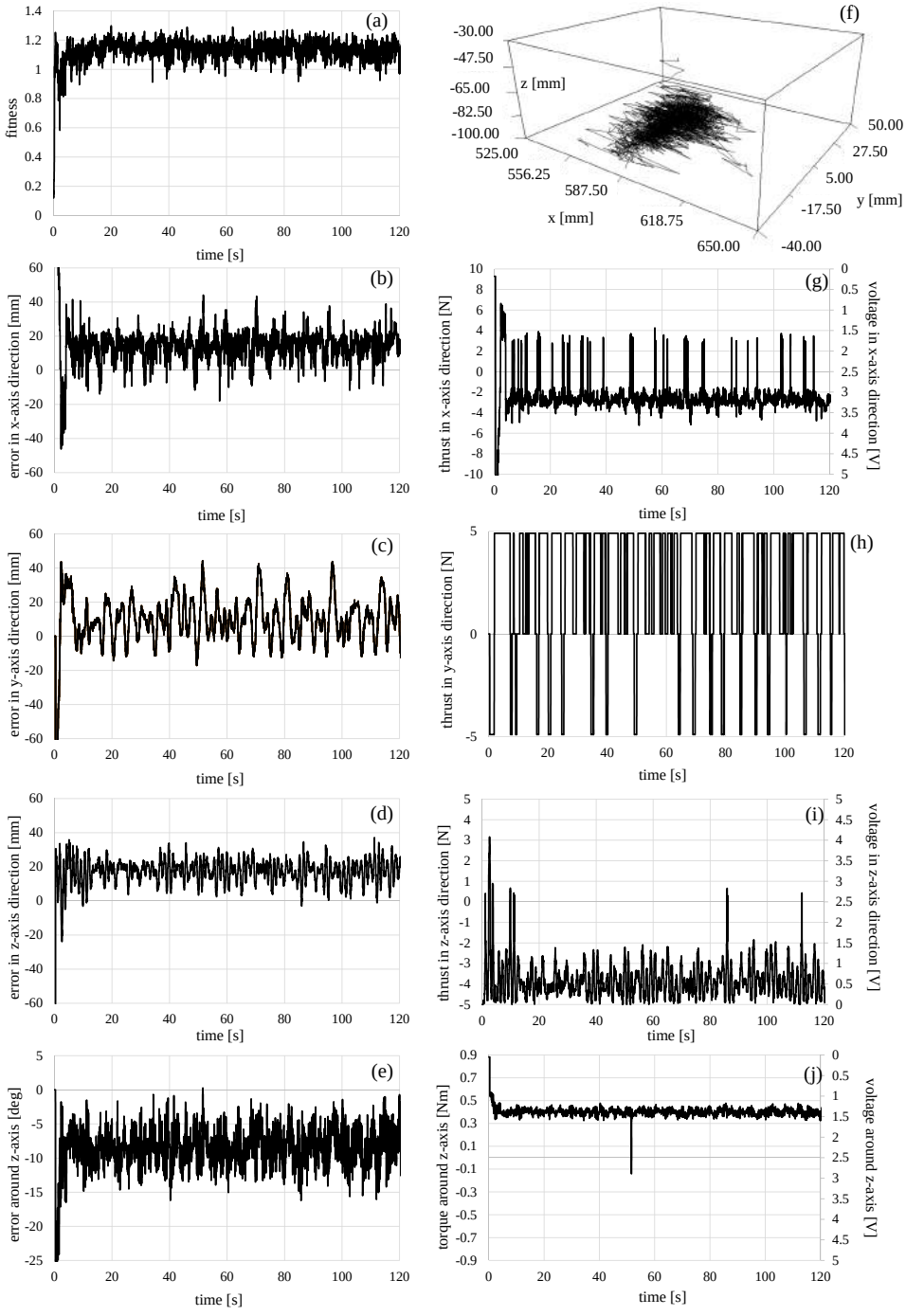


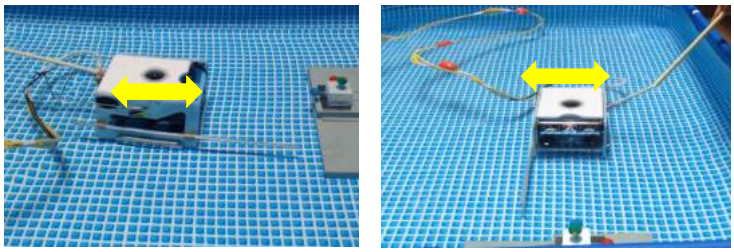
Fig. 4.8: Regulating performance : (a) fitness value, (b) error in x-axis direction, (c) error in y-axis direction, (d) error in z-axis direction, (e) error around z-axis, (f) 3D trajectory of underwater vehicle (g) thrust in x-axis direction, (h) thrust in y-axis direction, (i) thrust in z-axis direction and (j) torque around z-axis.

4.8 (a) shows the fitness value recognized by RM-GA. Figure 4.8(f) shows the position of underwater robot in the regulation as measured in RM-GA and Fig. 4.8(b)∼(e), (g)∼(j) represent the errors between the relative pose of the target and that of the underwater robot, and the torque to restore it, respectively. Although error in the relative target pose appears constantly and the four thrusters operate simultaneously, there are some pose fluctuation according to the cable tension during robot movement and reflected waves from the pool sides that occur due to water pressure changes with robot movement. However, the proposed system is able to regulate the relative pose by canceling these disturbance elements. As the lateral thruster control is on-off logic, the position error in this axis may significant comparing to the others. Therefore, P controller will be adopted for that thruster in next experiment.

4.3.3 Robustness Against Physical Disturbances

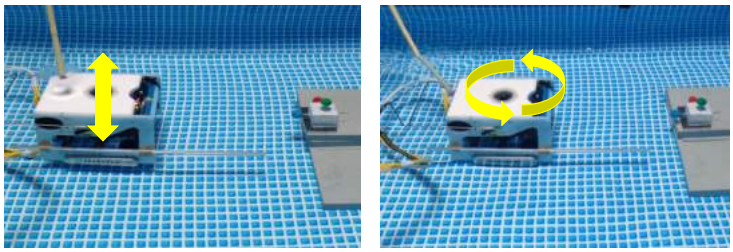
The robustness of the proposed system should be verified considering all possibilities that the real environment could provides the vehicle. Therefore, physical disturbances are simulated for proposed system. Abrupt external forces to move the vehicle for distance of 150∼200 [mm] between 1.5 to 2.0 [s] in difference directions and to rotate degree 15 [deg] per 1 [s] for rotation around a vertical axis by mean of a rod from the outside of pool are applied as shown in Fig.4.9 and the robustness of the visual servoing is analyzed.

The regulating performance with a disturbance in each direction is shown in Figs. 4.10 to 4.13. The term “stability” in this study means the property in which the underwater robot can be restored to the target pose relative to the 3D marker, even when a disturbance is given to the ROV. Fig. 4.10 to Fig. 4.13 show (a) the fitness of GA recognition, (b) the error between the relative pose between the target object and the underwater robot recognized for each variable, and (c) the same results of (b) enlarged during a disturbance. Furthermore, (d) represents the thrust force (torque) applied to the thruster. The disturbance was applied twice around 20 s and 60 s from the beginning



(a) Disturbance in x-axis direction

(b) Disturbance in y-axis direction



(c) Disturbance in z-axis direction

(d) Disturbance in around z- axis

Fig. 4.9: Physical disturbance in different directions: (a)x-axis, (b)y-axis, (c)z-axis (d)around z-axis.

of the experiment. In the section shown by (A) and (B) in (a), (b), (c), it is found that varying the thrust (torque) applied to the thrusters in response to an error of the relative target pose maintains the relative pose during visual servoing, although the fitness is temporarily lowered when a disturbance is applied. In other words, it is possible to confirm that an operation for correcting the error has occurred and consequently will change the pose of the ROV to restore the relative target pose. From the above results, the proposed system could restore the robot to the original position within a few to several tens of seconds for all of these disturbances.

Regulating performance with disturbance in vertical axis is shown in Fig.4.12. Fig.4.12 shows (a) the fitness of GA recognition, (b) the error between the relative pose of the object target and underwater robot in z-axis direction and , (c) the same results of (b) enlarged view from 55[s] to 65[s]. The disturbance has been added in each of the figures after 20 [s] and 60 [s] from the beginning of the experiment. Fig.4.12 (c) shows enlarged view during regulating response against prodding to be seen how the system behaved in real-time. In the period shown with (A) and (B) in Fig.4.12(d), it is found that varying the thrust (torque) is applied to the thrusters in response to error from the relative target

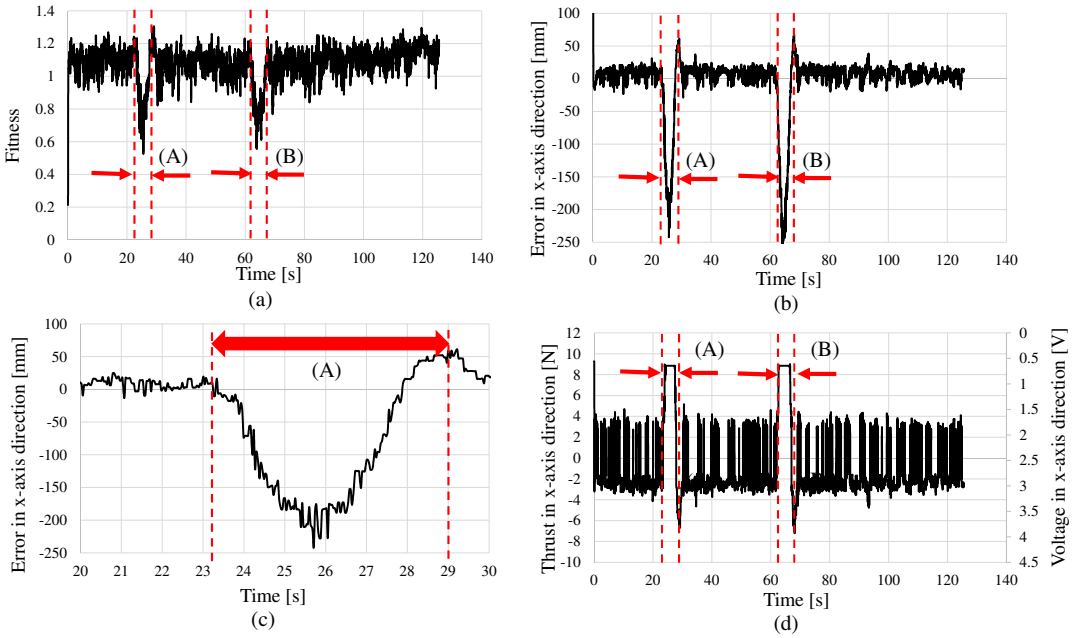


Fig. 4.10: Regulating performance with disturbance in x-axis direction: (a) fitness value, (b) error in x-axis direction, (c) error in x-axis direction (enlarged view from 20 s to 30 s), and (d) thrust in x-axis direction.

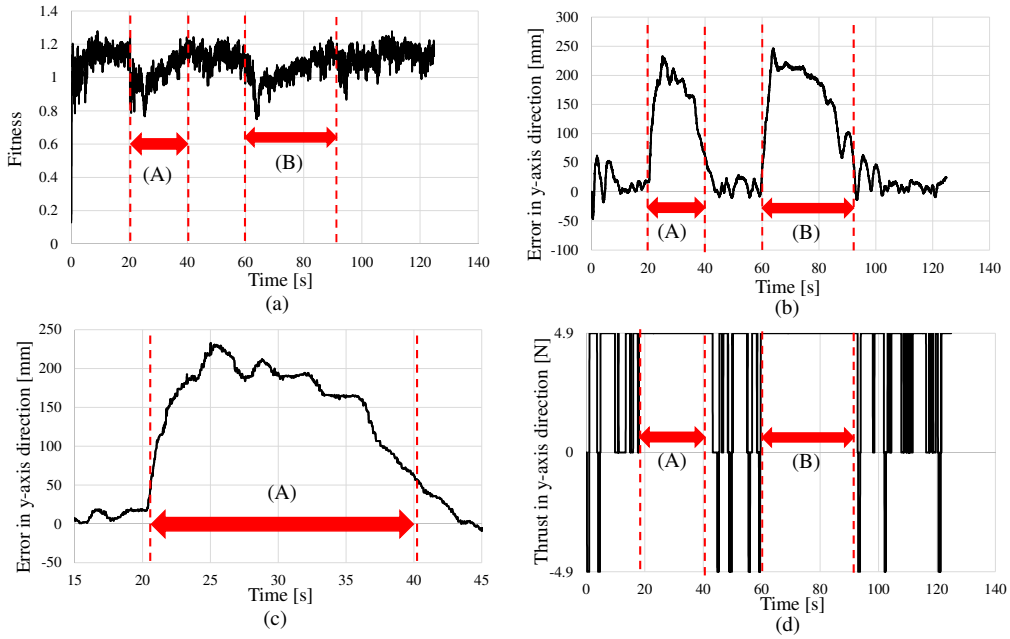


Fig. 4.11: Regulating performance with disturbance in y-axis direction: (a) fitness value, (b) error in y-axis direction, (c) error in y-axis direction (enlarged view from 15 s to 45 s), and (d) thrust in y-axis direction.

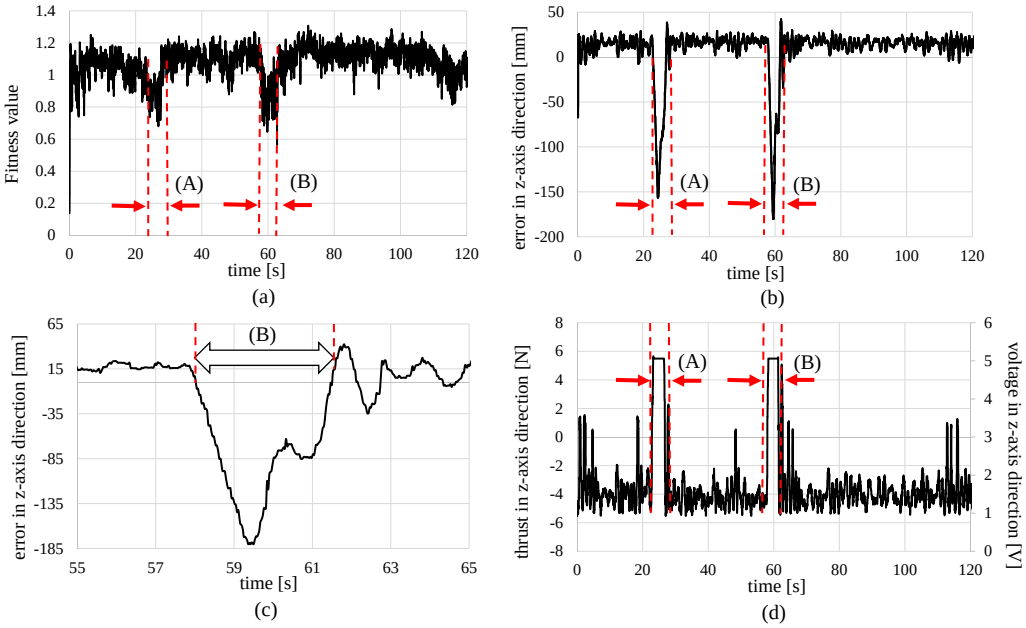


Fig. 4.12: Regulating performance with disturbance in z-axis direction: (a)fitness value, (b)error in z-axis, (c)error in z-axis(enlarged view from 55[s] to 65[s]) and (d)torque in z-axis.

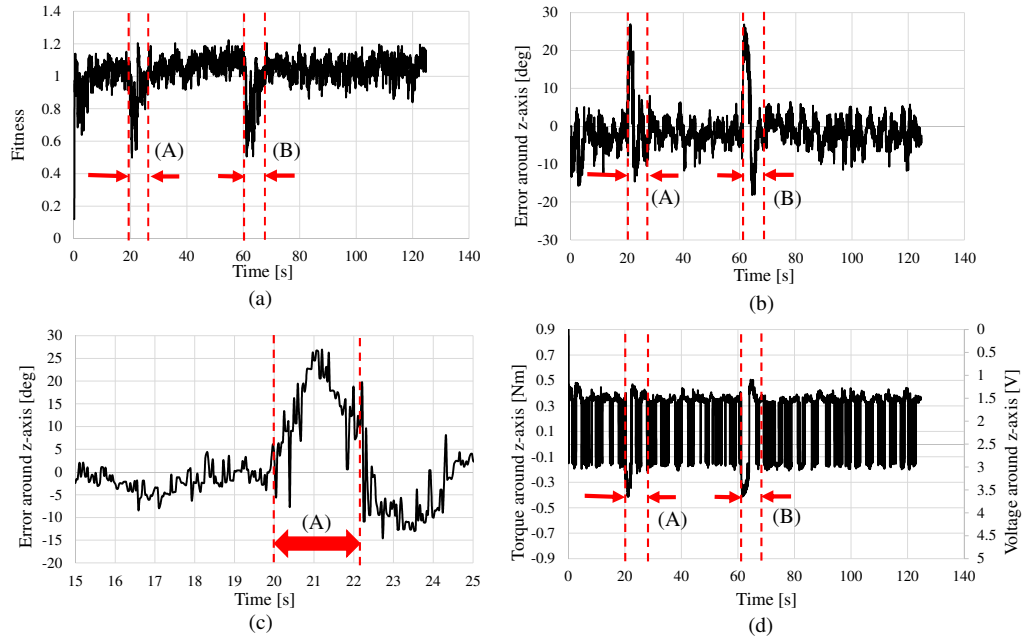


Fig. 4.13: Regulating performance with disturbance around z-axis: (a) fitness value, (b) error around y-axis, (c) error around y-axis (enlarged view from 15 s to 25 s), and (d) torque around z-axis.

pose while maintaining the visual servo although fitness is temporarily lowered when a disturbance is applied. From the above results, the proposed system can be restored to its desired pose within a few [s]$\sim$ several 10 [s] for all of these disturbances. Therefore, it was confirmed that the system is robust against external disturbances.

4.3.4 Robustness Against Target Occlusion

To verify the robustness of the system in term of target visibility, visual servoing when the object is partially seen was conducted. In simulated environment, one of the spheres was hidden for certain period and the visual servoing performance is analysed. Fig.4.14 shows recognition performance when the red sphere of target object is invisible for some period shown as (A), and Fig.4.15 shows the result when green sphere is being hidden from the view of the vehicle's camera. As shown in figures, the fitness value is reduced for the period when target is partially seen. To evaluate the proposed system utilizing RM-GA, recognition results are compared to the system in which the best gene is searched throughout all possible poses without using GA. For instance, the recognized positions in z-y plane with respect to fitness are shown in Fig. 4.14 and Fig.4.15. The poses evaluated in full searching are represented as contour and the ones evaluated by RM-GA are shown as black doted ones. These results highlight the promising optimization performance of GA to find the best gene using only some selected candidates for real-time. Fig.4.17 and Fig.4.18 are snap photos of conducting the visual servoing when the target object is partially hidden. It can be seen that the system is able to estimate the relative pose even one ball is hidden. The estimated model of the hidden ball can be seen as doted circle in cover white plane.

Fig.4.16 shows the regulating performance when the red ball is hidden during 20-30 seconds and 40-50 seconds from the view of the vehicle. Firstly the system recognizes the object with 1.2 fitness value in few seconds. Then, it can be seen that the fitness value reduces to around 0.8 from 1.2 for the period in which the red ball is hidden. According

to the experimental results as shown in Fig.4.16(b),(c),(d), however, the proposed system can maintain pose estimation accuracy and regulating performance even the object is partially occluded. The position error in y-axis direction is significant comparing to others because of the on-off control in transverse direction thruster. According to the several experiments, it was confirmed that the proposed system is robust not only for physical disturbances but also when the object itself is partially seen. Therefore, the proposed passive 3D marker with known color, size and especially structure, and RM-GA which forwards the best genes to the next generation might make this robustness come true in picture.

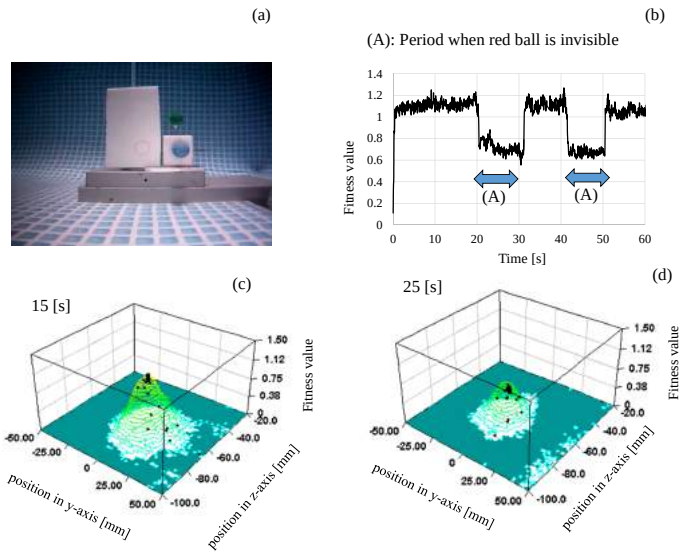


Fig. 4.14: Recognition performance : (a) Recognized model and real target, (b) fitness value when red ball is hidden in some period, (c)comparison of full search and GA search when all three balls are visible, (d) comparison of full search and GA search when red ball is invisible.

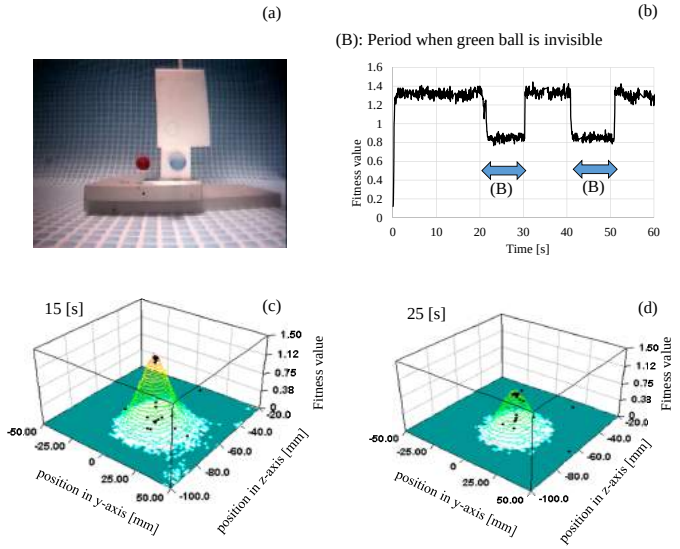


Fig. 4.15: Recognition performance : (a) Recognized model and real target, (b) fitness value when green ball is hidden in some period, (c)comparison of full search and GA search when all three balls are visible, (d) comparison of full search and GA search when green ball is invisible.

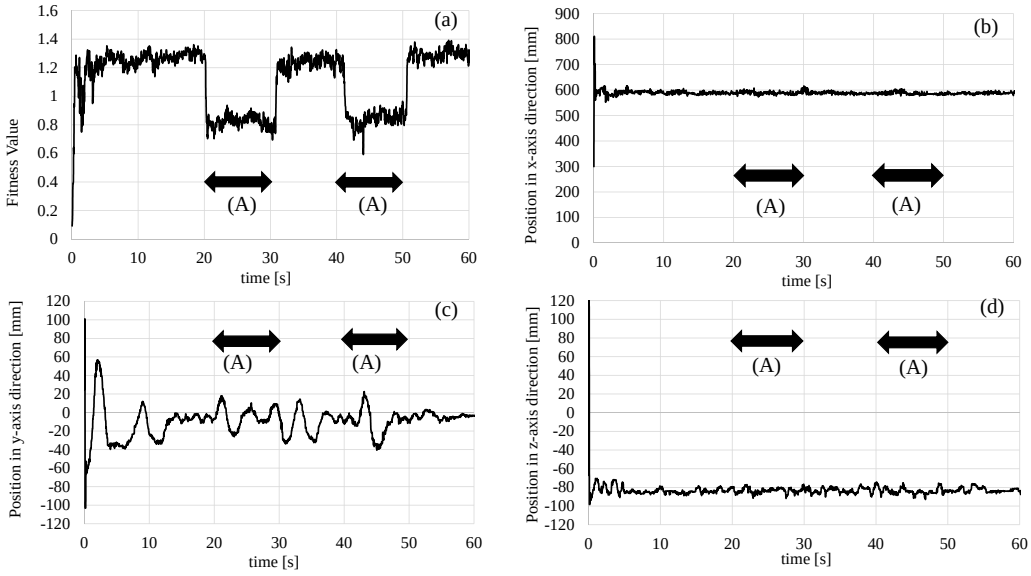


Fig. 4.16: Regulating performance when the object is partially seen : (a) fitness value, (b) position in x-axis direction, (c) position in y-axis direction, (d) position in z-axis direction. Corresponding photos of left and right camera images are shown in Fig. 4.17.

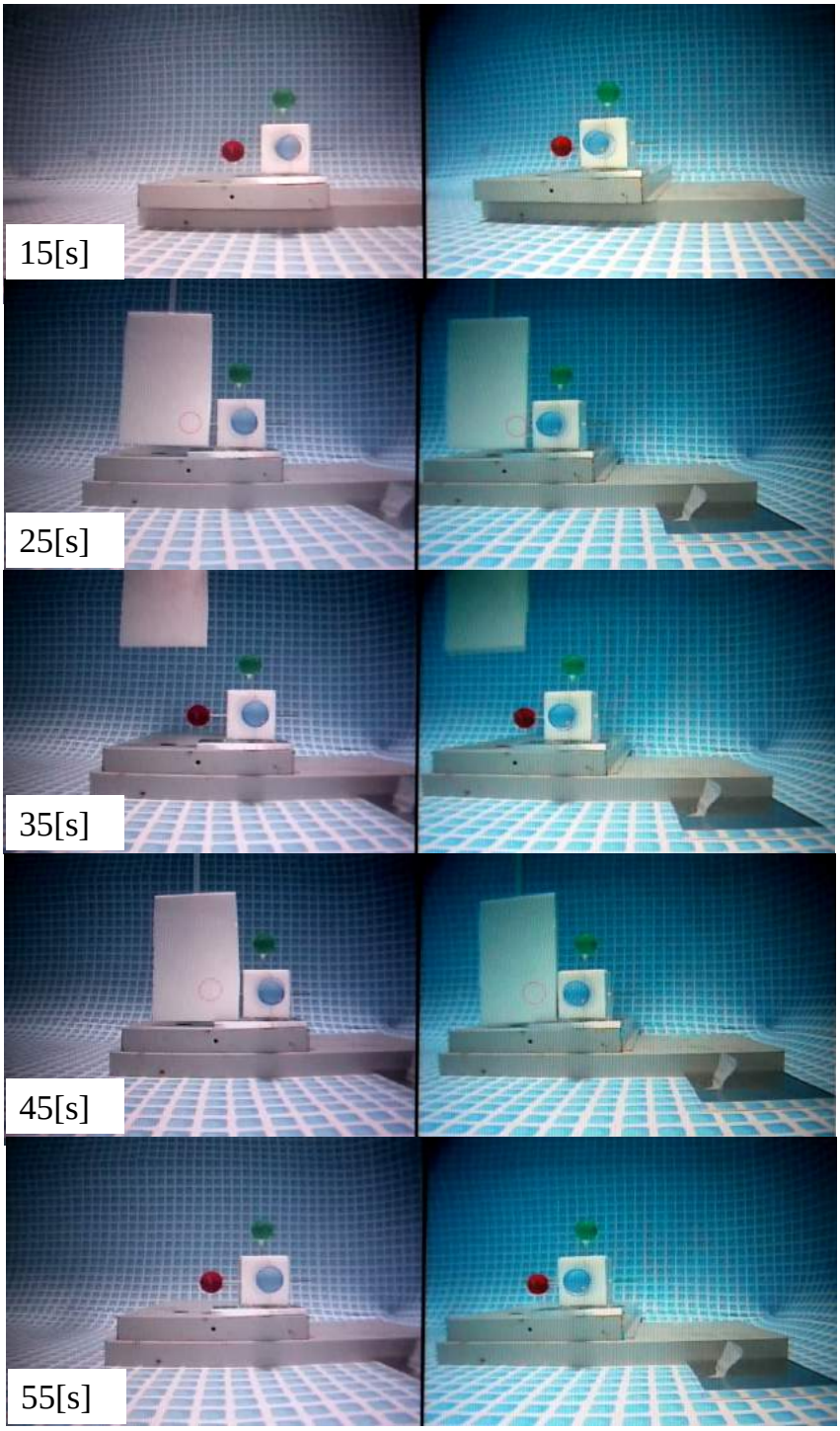


Fig. 4.17: Left and right camera images when the red ball is invisible between 20[s] to 30[s] and 40[s] to 50[s].

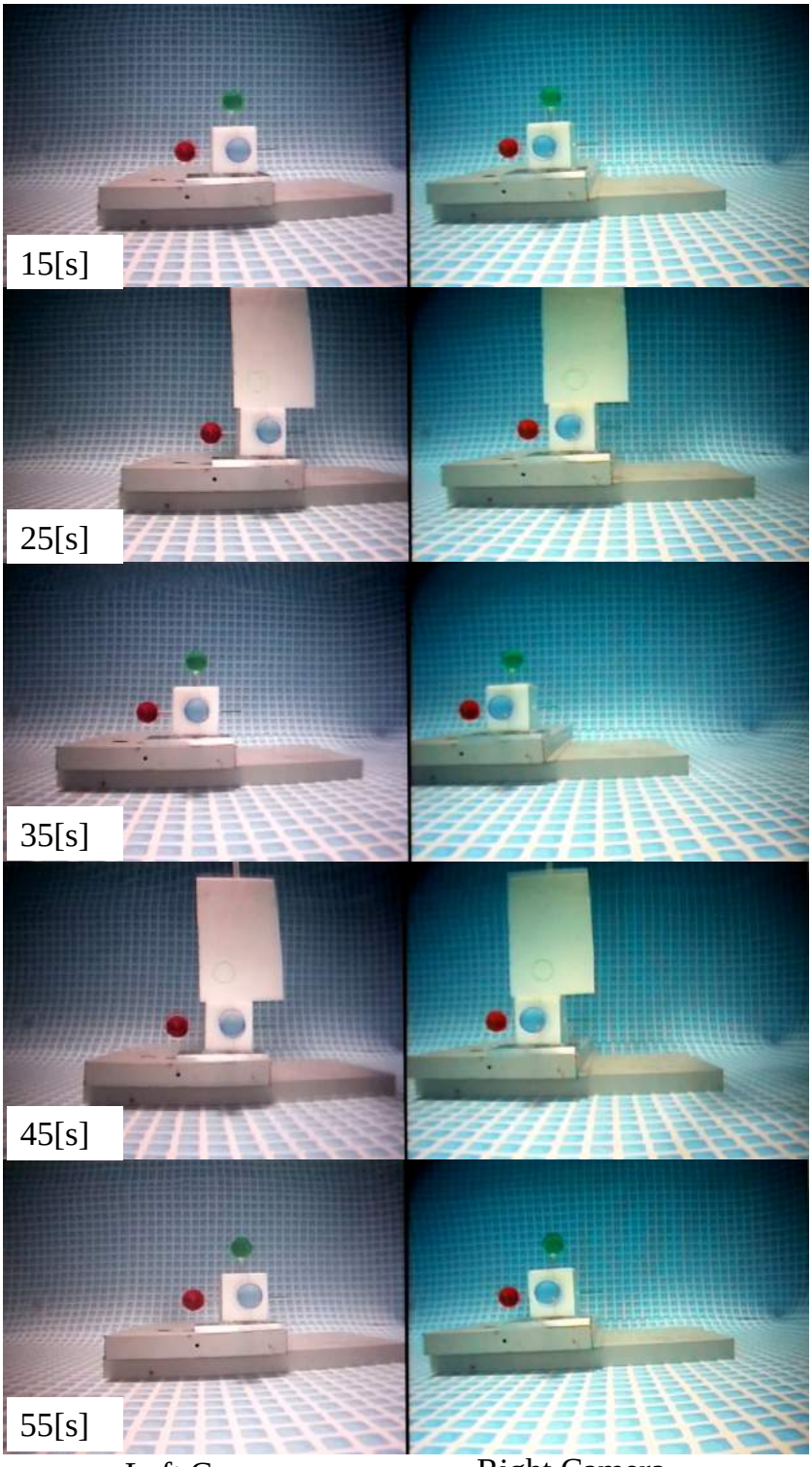


Fig. 4.18: Left and right camera images when the green ball is invisible between 20[s] to 30[s] and 40[s] to 50[s].

Chapter 5

Docking performance using proposed docking strategy

This section presents a vision-based docking system consisting of a 3D model-based matching method and Real-time Multi-step Genetic Algorithm (GA) for real-time estimation of the robot ' s relative pose. Experiments using a remotely operated vehicle (ROV) with dual-eye cameras and a separate 3D marker were conducted in an indoor pool. The experimental results confirmed that the proposed system is able to provide high homing accuracy and robustness against disturbances that influence not only the captured camera images but also the movement of the vehicle. A successful docking operation using stereo vision that is new and novel to the underwater vehicle environment was achieved and thus proved the effectiveness of the proposed system for AUV.

Figure 5.1 shows overall block diagram of the proposed system. Images from the dual-eye cameras installed on the underwater vehicle are sent to the GA-PC. Real-time pose estimation using the 3D model based matching method and real-time multi-step GA (RM-GA) is implemented as software implementation in GA-PC. Based on the real-time estimated relative pose between the AUV and the docking station, and designed docking strategy that will be explained in detail in this section, GA-PC sends command signal that is control voltage for each thruster to the ROV.

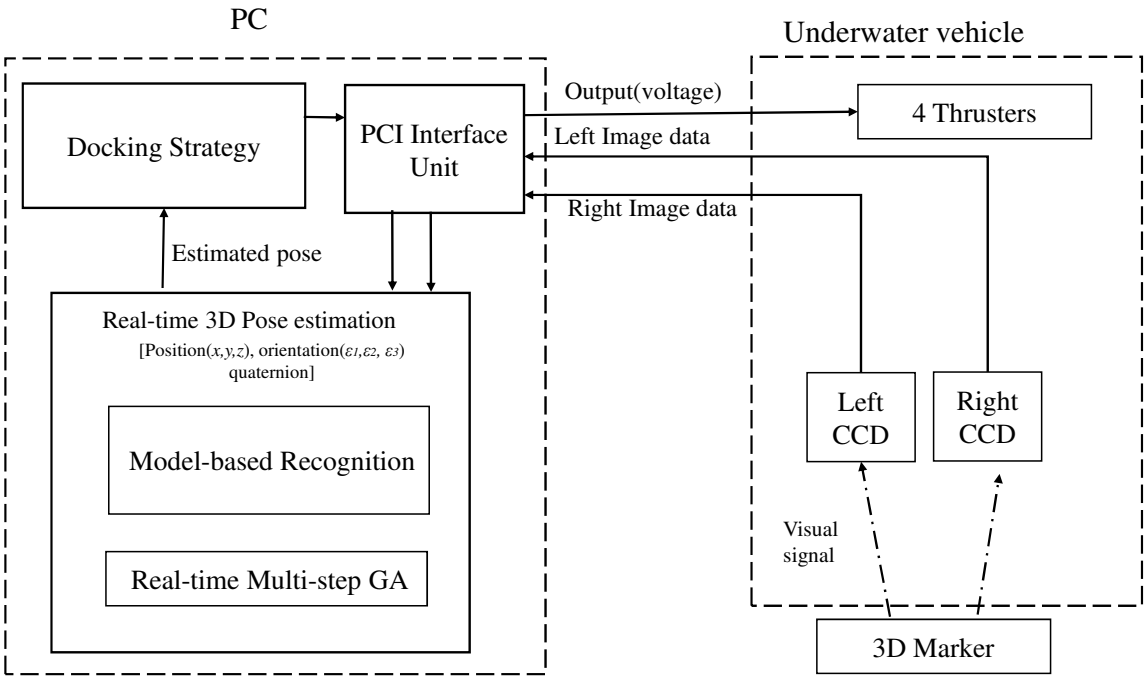


Fig. 5.1: Block diagram of the proposed system including designed docking strategy.

5.1 Docking Strategy

The proposed docking strategy consists of three steps. First, the ROV has to approach the 3D target until the target is in its field of view. Second, detecting the object and regulating the vehicle to the defined relative pose of the target is performed in the visual servoing step. Third, the docking operation is completed. The flowchart of the docking strategy is shown in Fig. 5.2. The originality of this work is concentrated on the dual-eye visual servoing as a possible new docking strategy rather than conventional docking methods. Therefore, the main contribution in the present paper is focused on the second and the third steps of Fig. 5.2 to demonstrate the effectiveness of the proposed docking system.

The first step can be extended for real-world application by using a long-distance navigation sensor to guide the vehicle into the field of view of the cameras. In [69], a state machine was proposed to generate a waypoint around the estimated target position and inside the vehicle's field of view, but that discussion was limited to the approaching step in Fig. 5.2. In this study, after approaching with constant speed and a constant proceeding

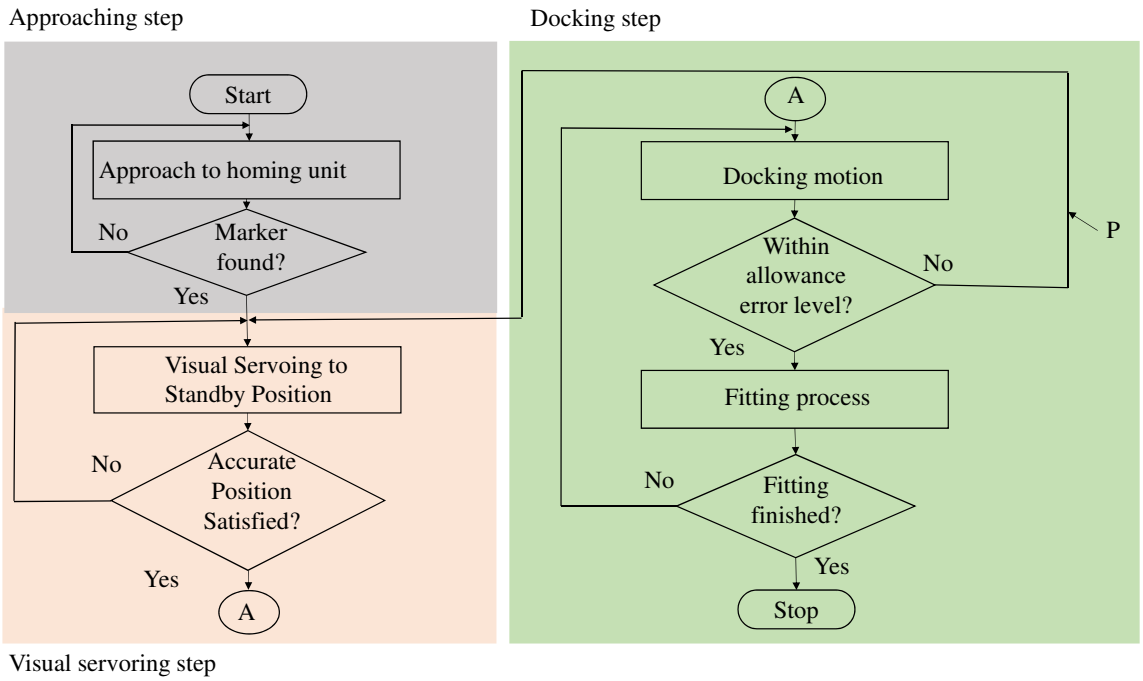


Fig. 5.2: Flowchart of docking strategy.

direction while trying to detect the 3D marker, the vehicle is stabilized in the visual servoing step and controlled to keep the ROV with a defined pose relative to the target. In the docking step, when the vehicle is stable within the tolerance of the position error for the defined time period, the forward thrust that enables the docking pole attached to the ROV to fit into the dock is generated by gradually decreasing the distance between the vehicle and the target object. Switching between the visual servoing mode and the docking mode by using the continuous pose feedback in the docking strategy (see "P" in Fig. 5.2) makes the system robust with little surfacing of the dock and minimizes the mechanical aspect as well.

5.2 Desired pose

The following relative pose between the ROV and the 3D marker (x_d [mm], y_d [mm], z_d [mm], ϵ_{3d} [deg]) is controlled according to the visual servoing step in Fig. 5.2.

$$x_d = {}^{H}x_M = 600\ (350)\, mm, \quad y_d = {}^{H}y_M = 0\ (0)\, mm,$$

$$z_d = {}^H z_M = -67\ (-67)\, mm, \quad \epsilon_{3d} = 0\ (0)\, deg$$

Each number in the above formulas is a target value for regulating the underwater robot immediately after recognizing the object in the visual servoing step as shown in Fig. 5.3 (a). Hx_M represents the x position of the origin of Σ_M in reference to Σ_H, where Σ_H and Σ_M are defined in Fig. 5.3. It should be noted that the numbers in parentheses are the defined target values at the time of completion of the fitting in the docking experiment as shown in Fig. 5.3 (b).

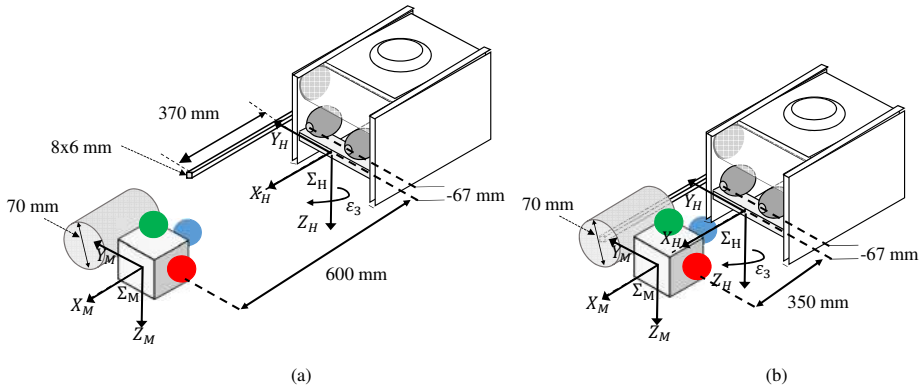


Fig. 5.3: Layout of the docking experiment showing the process of aligning the ROV with the 3D marker. (a) Desired pose in the visual servoing step. (b) Desired pose at the completion of the docking step.

5.3 Pool Docking Experiments

Experiments were carried out with different start positions: (a) in front of the 3D marker, (b) on the left side of the pool relative to the 3D marker, and (c) on the right side of the pool relative to the 3D marker, as shown in Fig. 5.4. The docking experiments were carried out as shown in Figs. 5.5, 5.7, and 5.8 following the four states (A) approaching the object (approach), (B) visual servoing to keep the relative pose to the object (visual servoing), (C) fitting to the fixed homing unit (docking), and (D) fully fitting into the homing unit (completion of docking) as shown in Fig. 5.5(a).

In the approaching step (A) in Figs. 5.5, 5.7, and 5.8, the robot's speed is low. This is the state until the underwater robot finds the 3D marker (recognition) under the

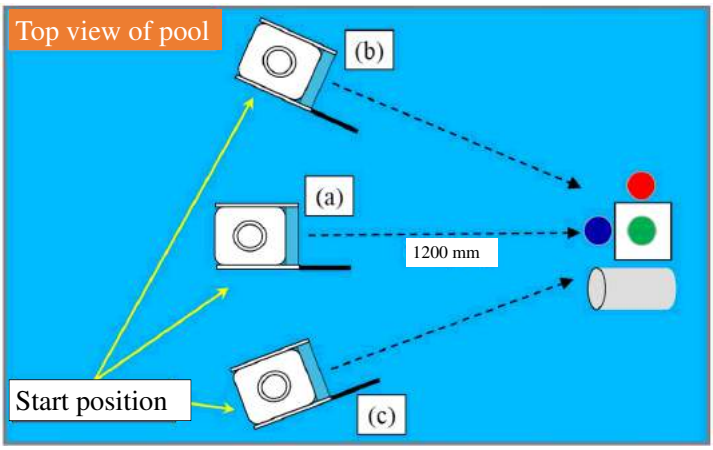


Fig. 5.4: Start position of underwater vehicle: (a) in front of 3D marker, (b) on the left side of pool relative to 3D marker, and (c) on the right side of pool relative to 3D marker.

assumption that the object is presented in front of the underwater robot. In other words, the underwater robot does not know the relative pose to the object in the initial condition, and then goes forward and transits to the state of visual servoing after discovering the object (judged by the fitness function rising to 0.5). Then, visual servoing is the state in which the underwater robot is regulated in the desired pose. As described above, the control process is performed to maintain the robot in the relative pose to the 3D marker. After transition to docking step, the underwater robot moves forward. The relative target position x_d decreases by 30 mm/s in the x_H-axis direction when the error of the relative position of the robot (y_d and z_d) with respect to the object is stable within $\pm$20 mm for the minimum period of 165 ms (control loop $\times$ 5). Then the docking state performs the fitting to the homing unit.

When the range of the abovementioned errors exceeds a defined value in the docking process, the underwater robot suspends the docking process and goes back to the visual servoing state to execute the docking process again, as shown by the arrow "P" in Fig. 5.2. In the state of visual servoing, when the error of the relative target position and the posture between the object and the underwater robot reaches below the abovementioned threshold, the process transits to docking, in which the robot is expected to fit into the homing unit while recognizing the target object. The docking process is done by

performing visual servoing until the robot moves to ($x_d = 350$ mm, $y_d = 0$ mm, $z_d = -67$ mm, $\epsilon_{3d} = 0$ deg). In completion of the docking state, the underwater robot is expected to keep the relative target pose of the object in the connected state by visual servoing. Figs. 5.5, 5.7, and 5.8 (b) to (j) show the result of docking experiments with different start positions: (b) is fitness, (c) (e) (g) (i) are the position and orientation of the underwater robot, and (d) (f) (h) (j) are the thrust and torque in each axis. Fig. 5.6 represents the error of the relative pose with respect to Fig. 5.5. From each of the figures, the transition of the state to (A) approach, (B) visual servoing, (C) docking, and (D) completion of docking can be seen. In other words, the experiments confirm that the underwater robot can achieve docking by using the proposed system. Figs. 5.9, 5.10, and 5.11 show the trajectories of the ROV recognized by the Real-time Multi-step GA when the ROV starts from different positions. Please note that these trajectories start by using visual servoing and end at the point where docking is completed. From the trajectories in Figs.5.10 (c) and 5.11 (c), surge motion appeared faster than sway motion.

5.4 Sea Docking Experiment

A docking experiment using the proposed approach in a real sea was the main task to be conducted to confirm the functionality and possibility of the proposed approach for AUVs for the sea docking application. Finally, to fulfill this aim, we have conducted docking experiments in the sea near Wakayama city in Japan aiming at evaluating the practicability of the system after developing the proposed approach based on previous works. The sea docking experimental results are discussed as the main contribution in this paper. To the best of the authors' knowledge, this work is new and the first trial about docking operation in the real sea using standalone dual-eyes cameras.

5.4.1 Docking experiment in a turbid water before sea trial docking experiment

To check the functionality of the proposed approach before the sea trial experiment and to extend the research in terms of robustness against water turbidity in future, a docking experiment in a turbid pool as shown in Figure 5.12 was conducted before the sea trial docking experiment. The pool is situated near the sea and was filled with seawater since a few months ago. There were a lot of particles inside the pool such as seaweeds, dry leaves that can provide disturbance to visual servoing in terms of noise in captured images. Because of the shadow of the trees beside the pool and the particles inside the pool, the inside environment of the pool is the natural condition with poor light and turbidity. The docking station was landed on the floor of the pool in this experiment. The detailed analyses on the robustness of the proposed system against water turbidity and illumination variation is in line with our on-going work.

Figure 5.13 shows the experimental result of the docking operation in a pool in which turbid water was filled up. Figure 5.13 (a) illustrates the time variation of fitness value. Figure 5.13 (c)(e)(g)(i) show the relative position of desired and recognized in each direction measured by RM-GA. As shown in Figure 5.13, the ROV was controlled manually until the 3D marker was detected. At the operation time of 35 s, automatic control by visual servoing was started with the minimum fitness value of 0.6 being confirmed by an operator. It can be seen in Figure 5.13 (a) that the fitness value increases from about 0.8 to more than 1 when the distance between the ROV and the 3D marker decreases. When the operation time is 70 s, docking is completed and the ROV is kept at the desired pose (See Figure 5.13 (b)). Even though the water is turbid, the recognition of 3D marker using defined information of 3D marker (color in hue value, size, and shape) utilizing designed fitness function was confirmed experimentally to be operational. According to the experimental result, the docking operation was conducted successfully within 60 s after the automatic control was started.

5.4.2 Sea docking experiment

We conducted docking operations in the sea four times. Figure 5.14 shows the experimental results of the first docking operation. Figure 5.14 (a) shows the fitness function. The recognized poses in manual control are not considered as truthful ones. According to the results, the visual servoing step started when the fitness value was above 0.6 and the distance between the vehicle and the station was about 800 mm. According to the experimental results, it can be seen that the ROV is controlled precisely in the docking step by visual servoing to perform the docking operation successfully. The time profile of the fitness value is shown in Figure 5.14 (a). It can be seen in Figure 5.14 (a) that the fitness value increases slightly when the ROV approaches the 3D marker. The recognized positions of the vehicle in x, y, and z axes are illustrated in Figure 5.14 (c),(e),(g). It can be seen that the docking step was performed when the position errors in y-axis and z-axis were within the allowable range that was ± 20 mm. As shown in Figure 5.14 (e), the position of the ROV in the y-axis direction is fluctuated and sometimes it is out of the allowance error until the operation time is 50 s. This is the condition of switching from docking step to visual servoing step shown by " P " in Figure 5.2. At that time, the control process is performed to maintain the robot in the desired pose. When the position in the y-axis and z-axis direction are stable for 165 ms continuously within the allowance error (shown by dotted lines in Figure 5.14 (e), (g), the vehicle moves ahead until the position in the x-axis is 350 mm from the 3D marker as shown in Figure 5.14 (c). It can be seen that the docking operation was completed successfully within 40 s after detecting the 3D marker.

The success of docking operation was also confirmed by checking recorded video images from two cameras of the ROV and underwater cameras installed on the docking station. Figure 5.18 shows the periodically grabbed images of the dual-eye cameras showing that the ROV approaches the 3D target marker by manual operation, visual servoing and docking. The right column images are taken by one of the underwater cameras that were

installed in the docking station to observe docking operation. The recognized pose by RM-GA are projected as dotted circles (see the top photo in Figure 5.18 (a)) on the images for visual observation to confirm how much recognized pose is matched with the real target 3D marker on images. Please note that the recognition poses in manual operation are not trustful ones and they were not used in the feedback control. This is why the dotted circles (see the top photo in Figure 5.18 (a)) are not totally overlapped with the target 3D marker in the initial stage of manual control. The images from two cameras and the underwater camera taken when the operation time is 32 s shown in Figure 5.18 (a) show the state when the automatic control was started after the 3D marker was detected with the fitness value increasing beyond 0.6. Images taken when the operation time is 32 s show the condition of visual servoing in which the ROV is controlled to be stable to transit to docking step. The images taken when the operation time are 35 s, 45 s, 55 s, 65 s illustrate the docking step.

The other docking results are shown in Figures 5.15 to 5.17. It can be seen in Figures 5.15 to 5.17 (a) that the fitness value was above 1 when the 3D marker was detected by the proposed system. All the docking operations were finished successfully within 40 s after the automatic control was started. There were some pose fluctuations because of natural disturbances like sea current. However, the vehicle can be maintained by visual servoing using the proposed system and finally the docking operations were performed successfully.

Please note that the recognized poses during manual control are not truthful ones and they are not used in the feedback system because the detection of the 3D marker was defined by a fitness value that is 0.6. Therefore, automatic control was started when the fitness value is above 0.6 and docking operation was performed in automatic control. Since whether the RM-GA has detected the 3D marker or not can be judged by the value of fitness function, the switching from manual control to automatic control could be done by autonomously. Regarding accuracy, it was confirmed experimentally that both recognition and docking accuracy is centimeter level because the docking hole radius

is 35 mm and the allowance error is ± 20 mm. Figure 5.19 shows docking steps while conducting sea trials.

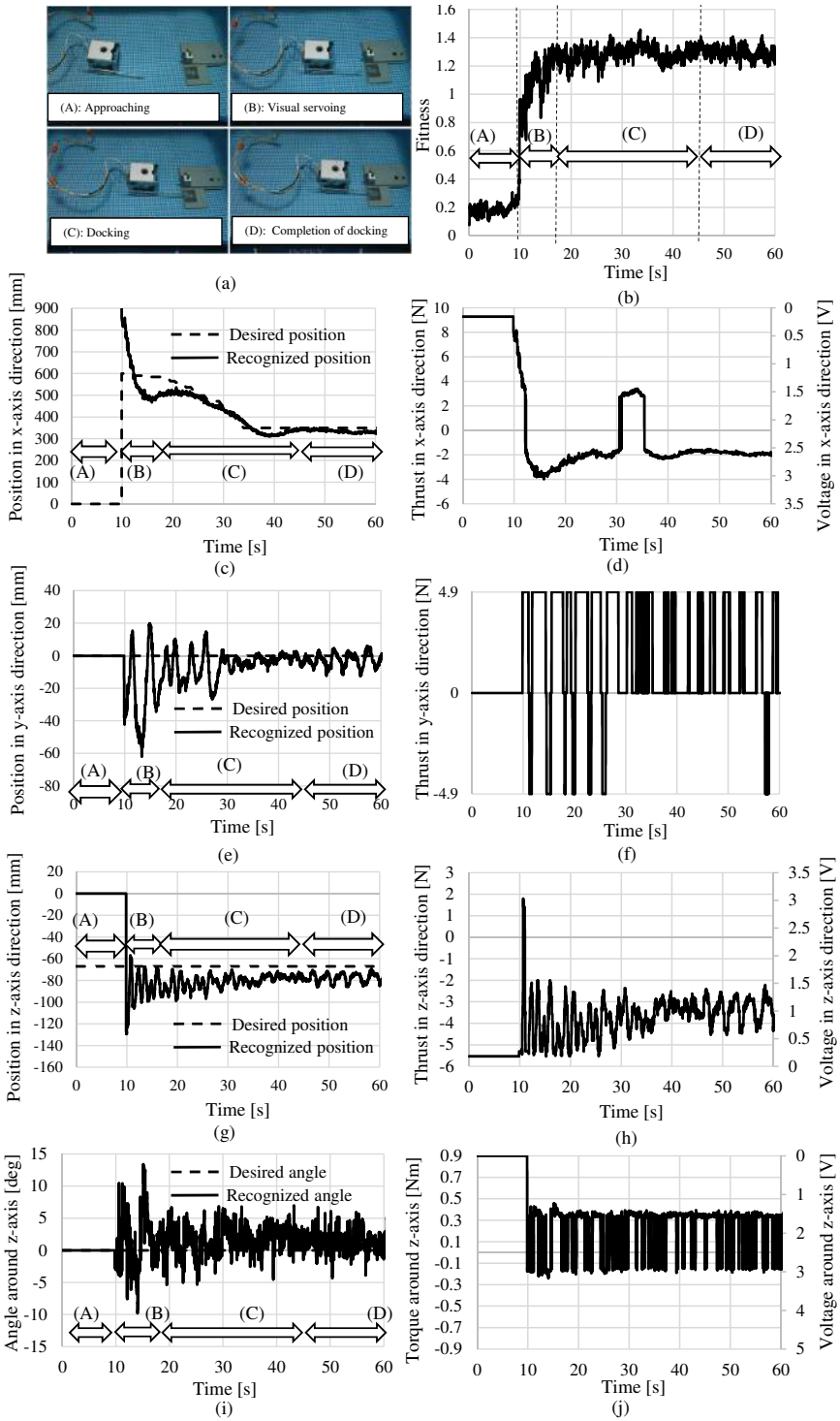


Fig. 5.5: Docking experimental results for start position of underwater vehicle in front of 3D marker, position (a) in Fig. 5.4 : (a) photo of docking experiment, (b) fitness value, (c) position in x-axis direction, (d) thrust in x-axis direction, (e) position in y-axis direction, (f) thrust in y-axis direction, (g) position in z-axis direction, (h) thrust in z-axis direction, (i) angle around z-axis, and (j) torque around z-axis.

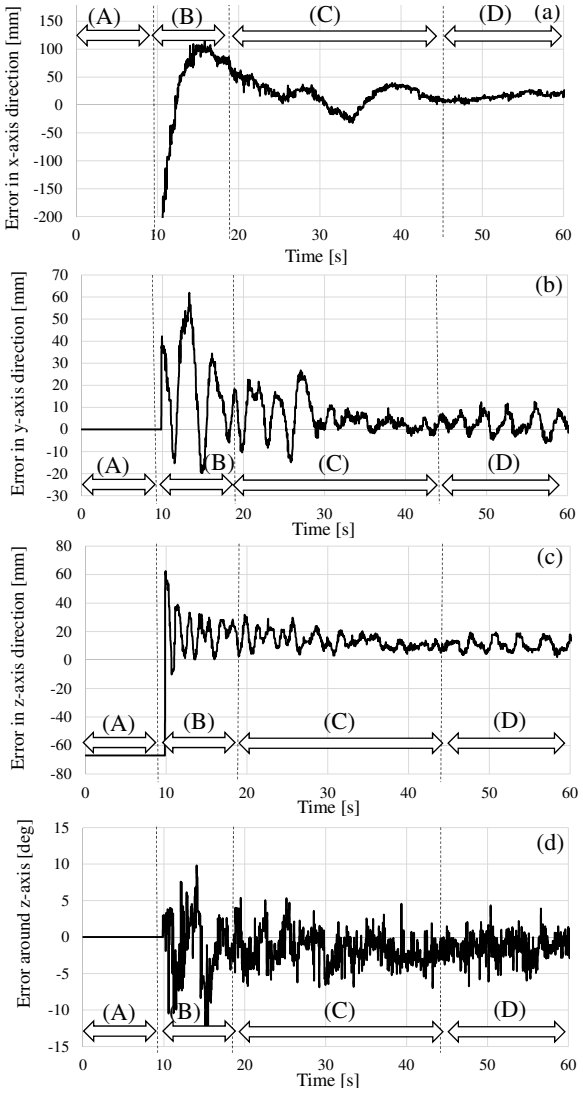


Fig. 5.6: Further docking experimental results for start position of underwater vehicle as in front of 3D marker, position (a) in Fig. 5.4 : (a) error in x-axis direction, (b) error in y-axis direction, (c) error in z-axis direction, and (d) error around z-axis.

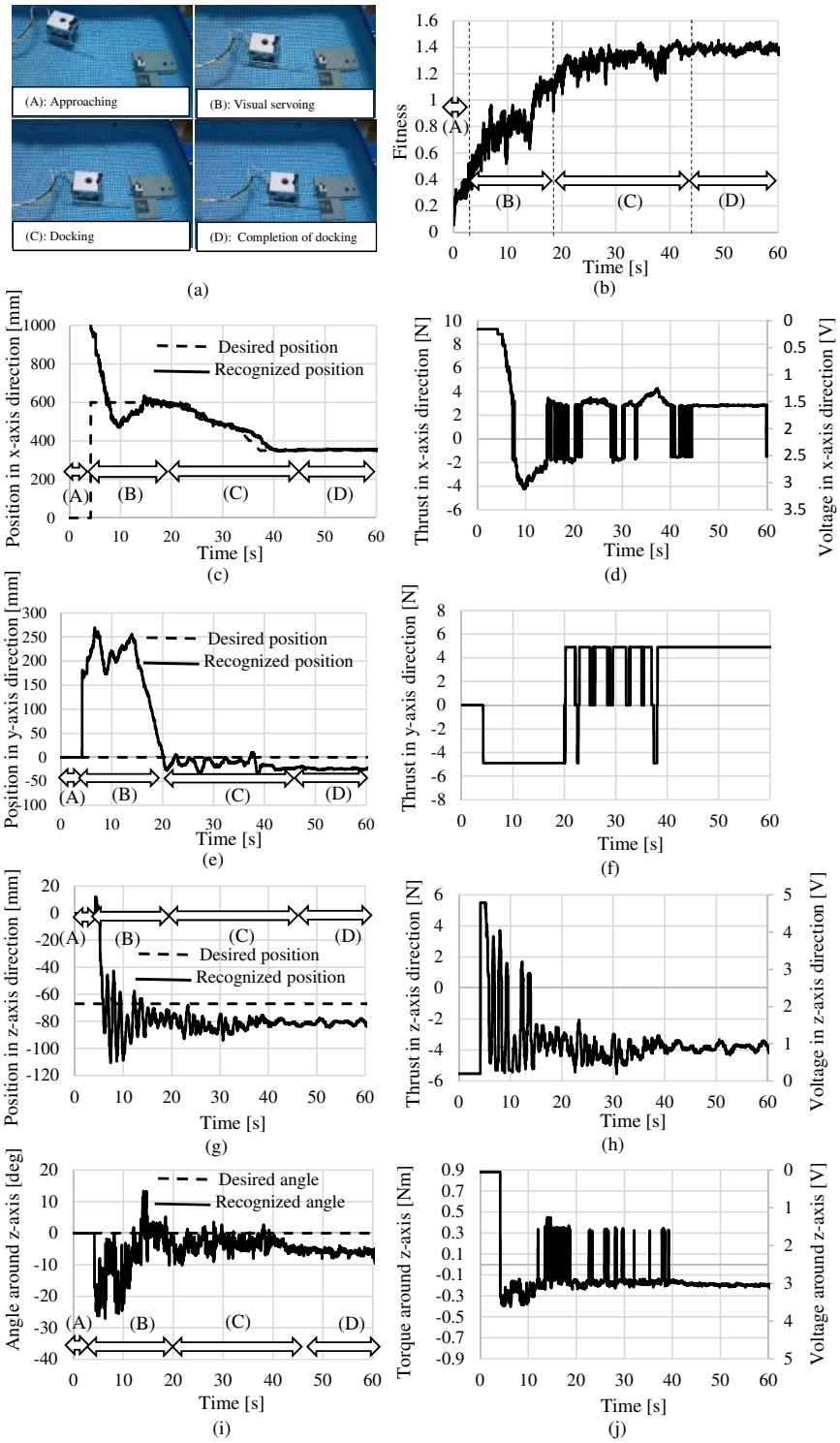


Fig. 5.7: Docking experimental results for start position of underwater vehicle on the left side of pool relative to 3D marker, position (b) in Fig. 5.4 : (a) photo of docking experiment, (b) fitness value, (c) position in x-axis direction, (d) thrust in x-axis direction, (e) position in y-axis direction, (f) thrust in y-axis direction, (g) position in z-axis direction, (h) thrust in z-axis direction, (i) angle around z-axis, and (j) torque around z-axis.

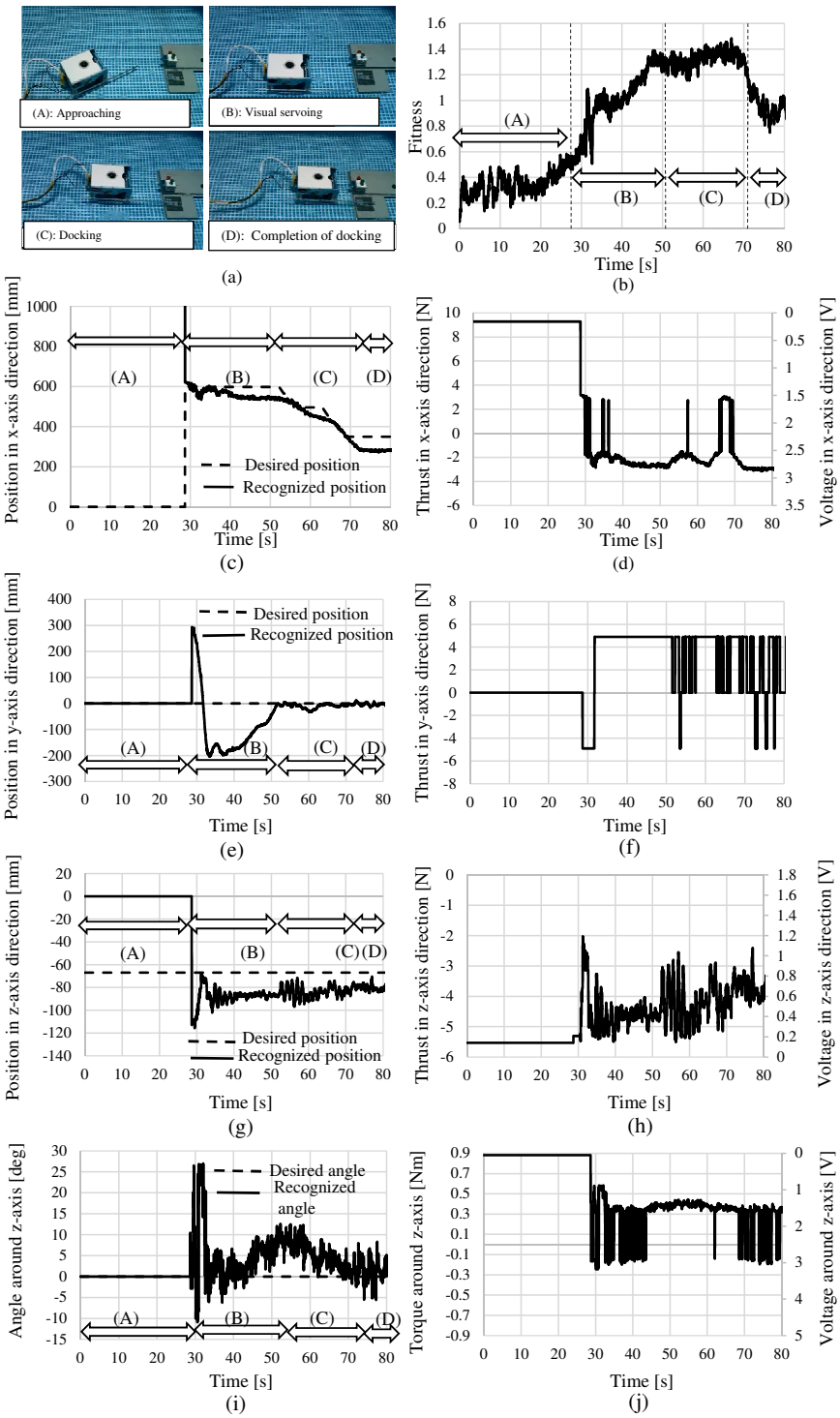


Fig. 5.8: Docking experimental results for start position of underwater vehicle on the right side of pool relative to 3D marker, position (c) in Fig. 5.4: (a) photo of docking experiment, (b) fitness value, (c) position in x-axis direction, (d) thrust in x-axis direction, (e) position in y-axis direction, (f) thrust in y-axis direction, (g) position in z-axis direction, (h) thrust in z-axis direction, (i) angle around z-axis, and (j) torque around z-axis.

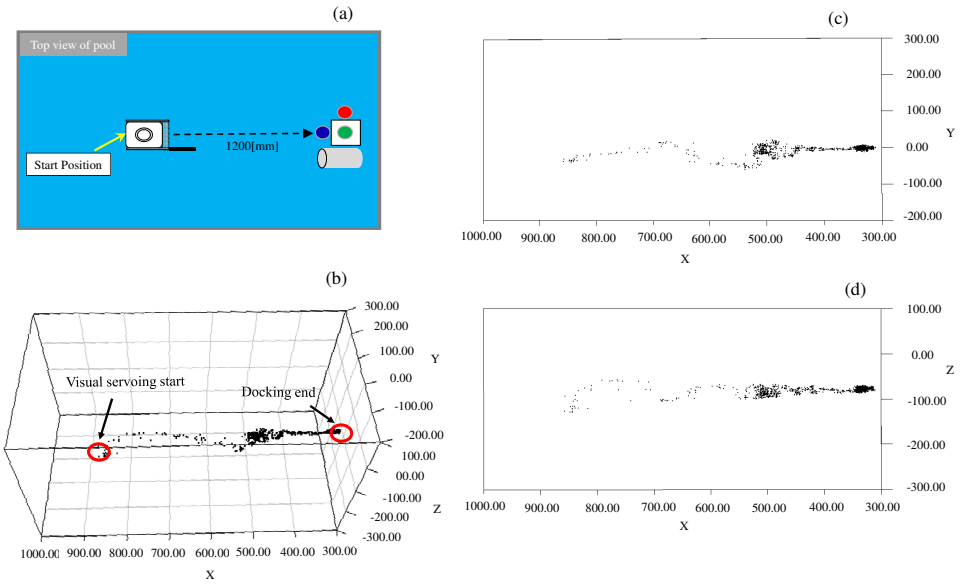


Fig. 5.9: Recognized trajectory for start position of underwater vehicle in front of 3D marker, position (a) in Fig. 5.4: (a) Start position of underwater vehicle, (b) recognized position in 3D by Real-time Multi-step GA, (c) recognized position in x-axis and y-axis, and (d) recognized position in x-axis and z-axis.

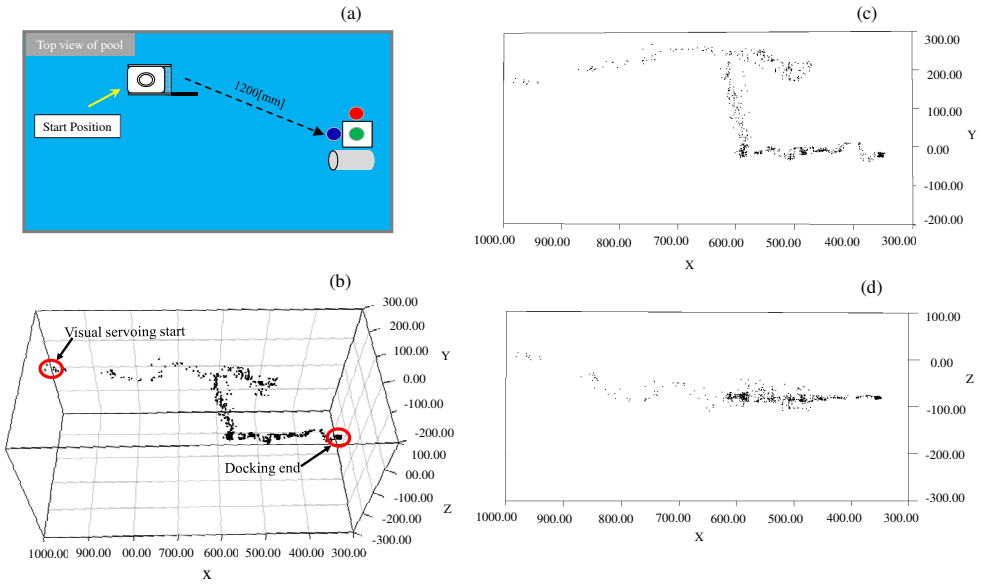


Fig. 5.10: Recognized trajectory for start position of underwater vehicle on the left side of pool relative to 3D marker, position (b) in Fig. 5.4: (a) Start position of underwater vehicle, (b) recognized position in 3D by Real-time Multi-step GA, (c) recognized position in the xy plane, and (d) recognized position in the xz plane

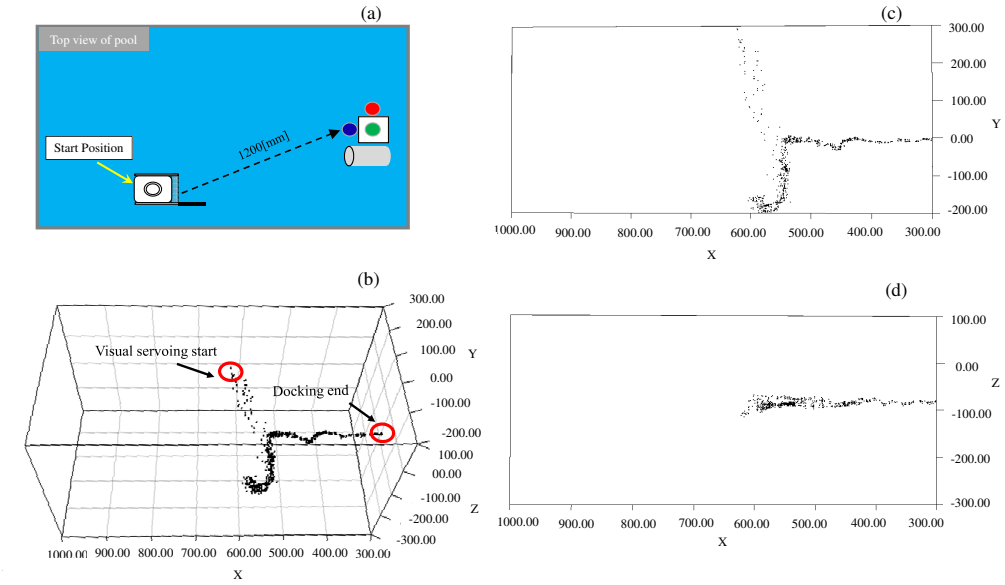


Fig. 5.11: Recognized trajectory for start position of underwater vehicle on the right side of pool relative to 3D marker, position (c) in Fig. 5.4: (a) Start position of underwater vehicle, (b) recognized position in 3D by Real-time Multi-step GA, (c) recognized position in the xy plane, and (d) recognized position in the xz plane.

(a) (b)

Fig. 5.12: (a) Turbid water in the pool, and (b) ROV and docking station.

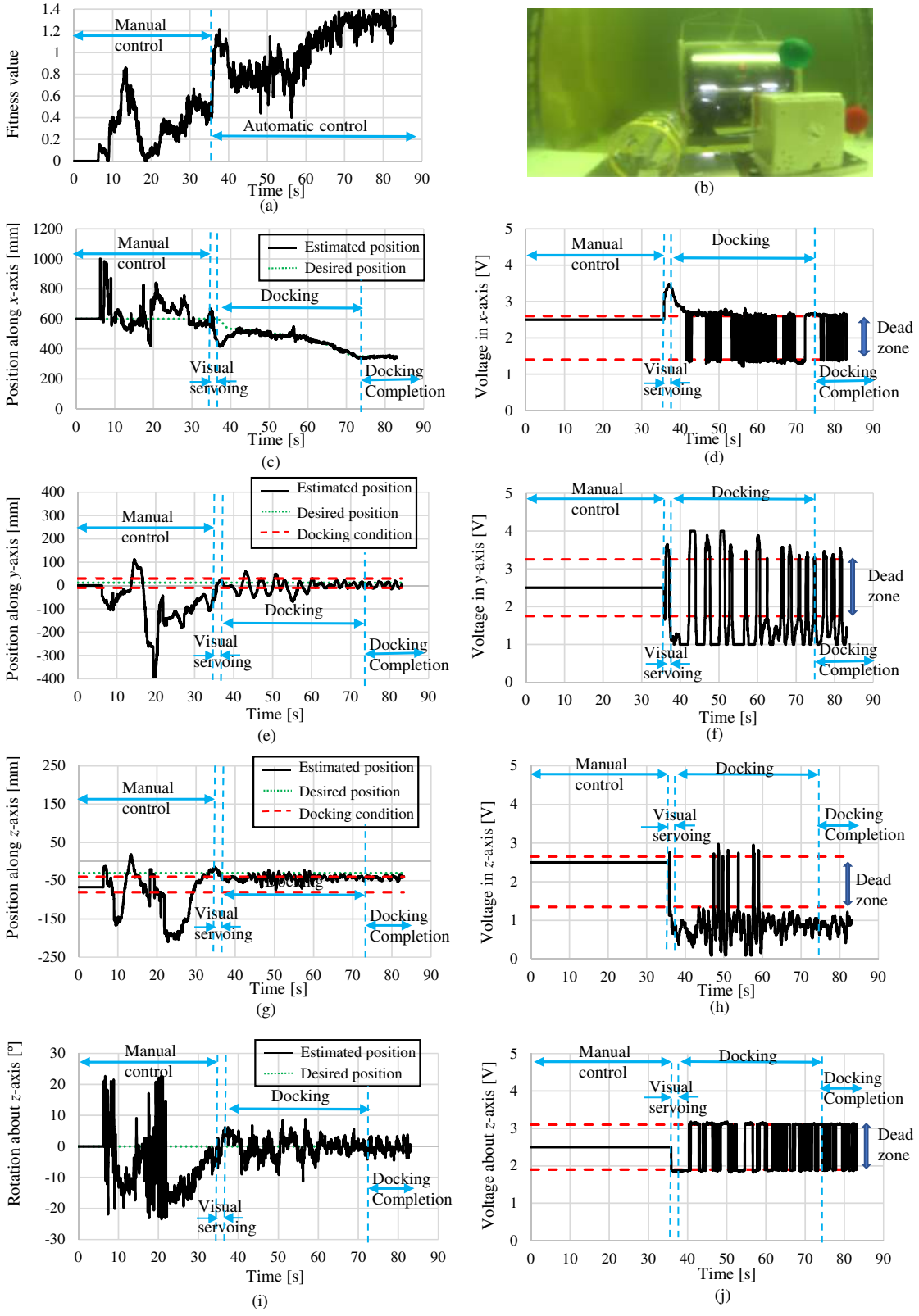


Fig. 5.13: Docking result at pool with turbid water using a circular shaped docking hole : (a) fitness value, (b) Photograph of ROV in docking process, ((c), (e), (g), (i)) recognized positions in x,y,z axes directions and rotation around z-axis, ((d), (f), (h), (j)) recognized positions in x,y,z axes directions and rotation around z-axis.

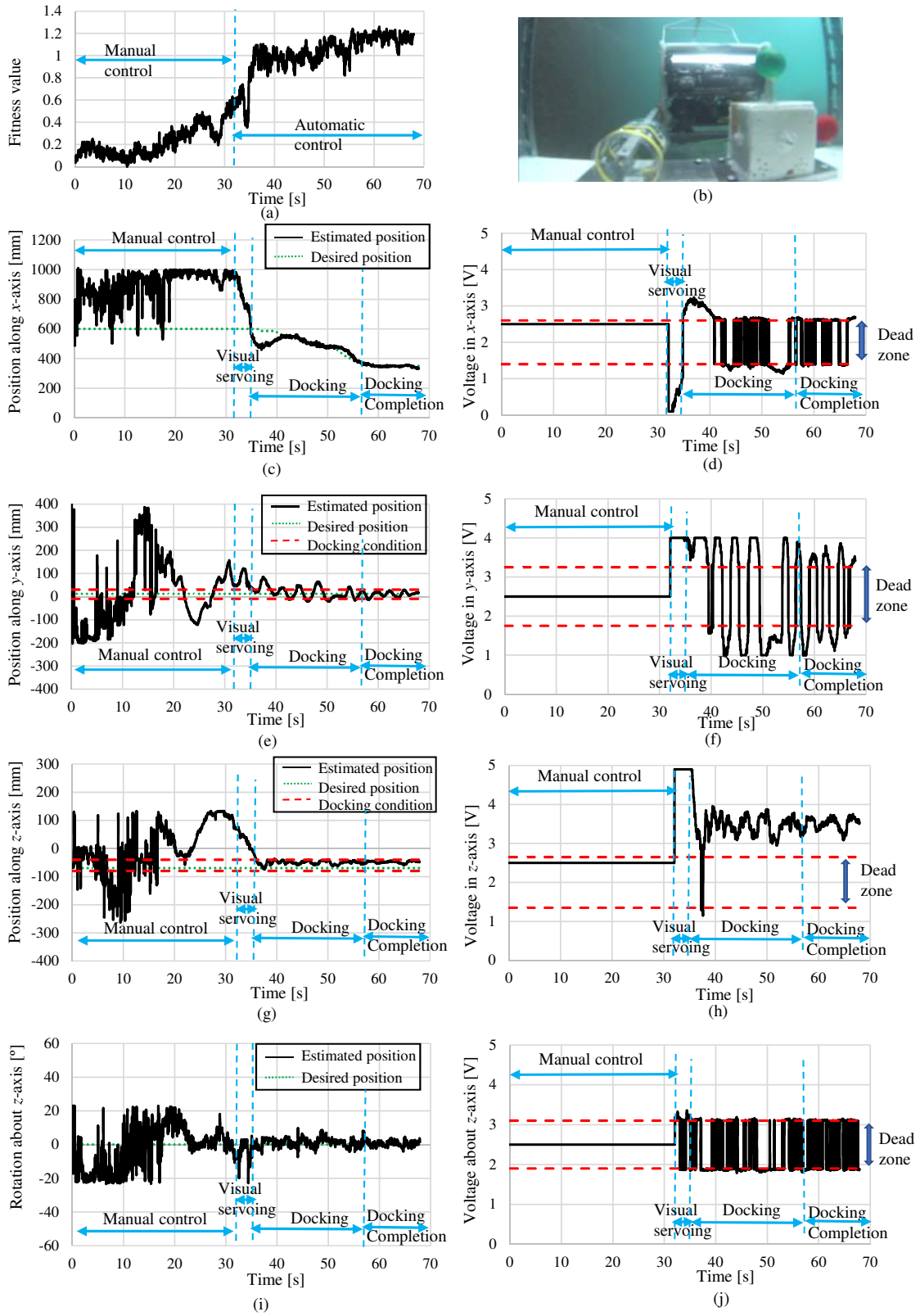


Fig. 5.14: Sea docking result 1 using a circular shaped docking hole : (a) fitness value, (b) Photograph of ROV in docking process, ((c), (e), (g), (i)) recognized positions in x,y,z axes directions and rotation around z-axis, ((d), (f), (h), (j)) recognized positions in x,y,z axes directions and rotation around z-axis.

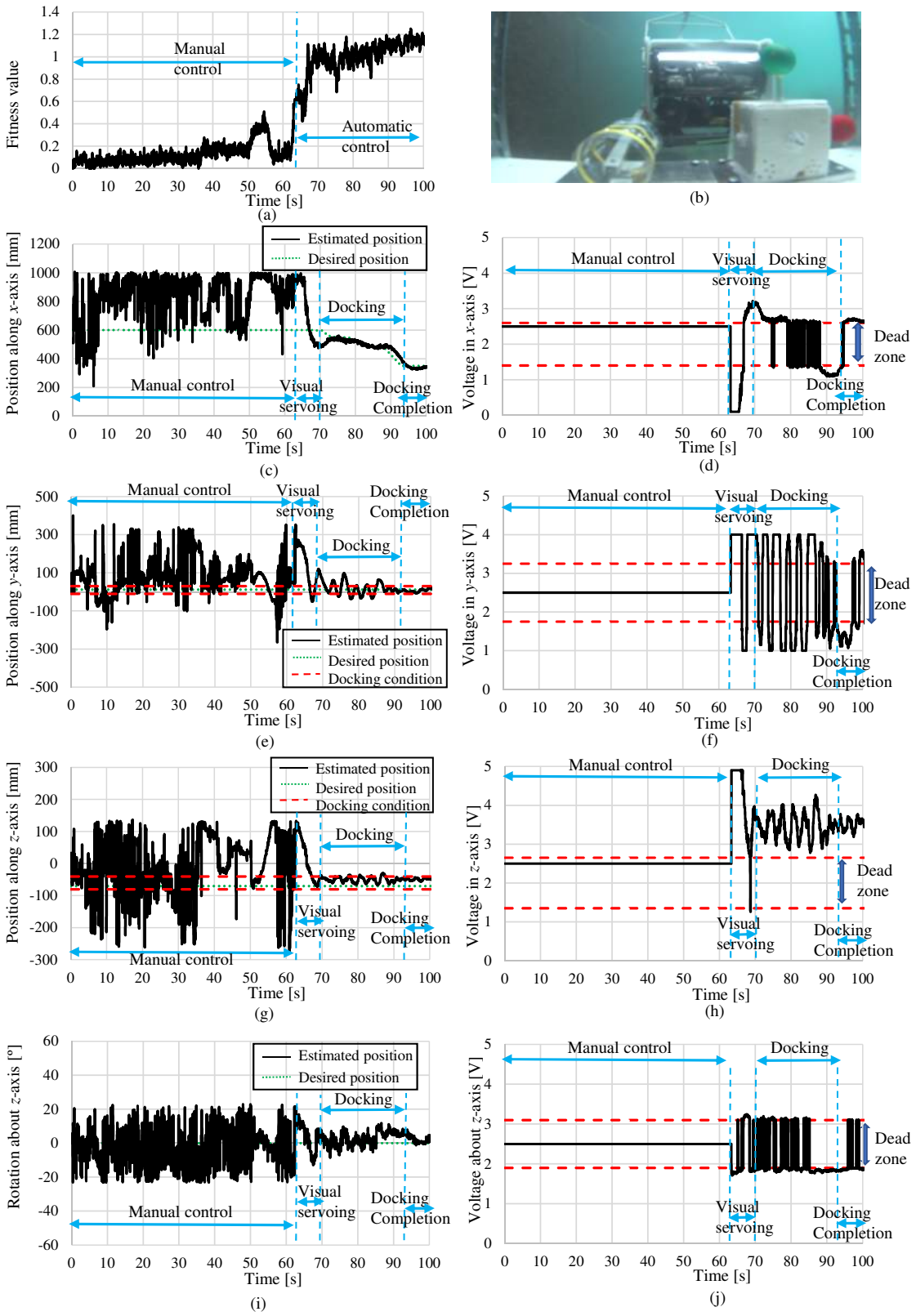


Fig. 5.15: Sea docking result 2 using a circular shaped docking hole : (a) fitness value, (b) Photograph of ROV in docking process, ((c), (e), (g), (i)) recognized positions in x,y,z axes directions and rotation around z-axis, ((d), (f), (h), (j)) recognized positions in x,y,z axes directions and rotation around z-axis.

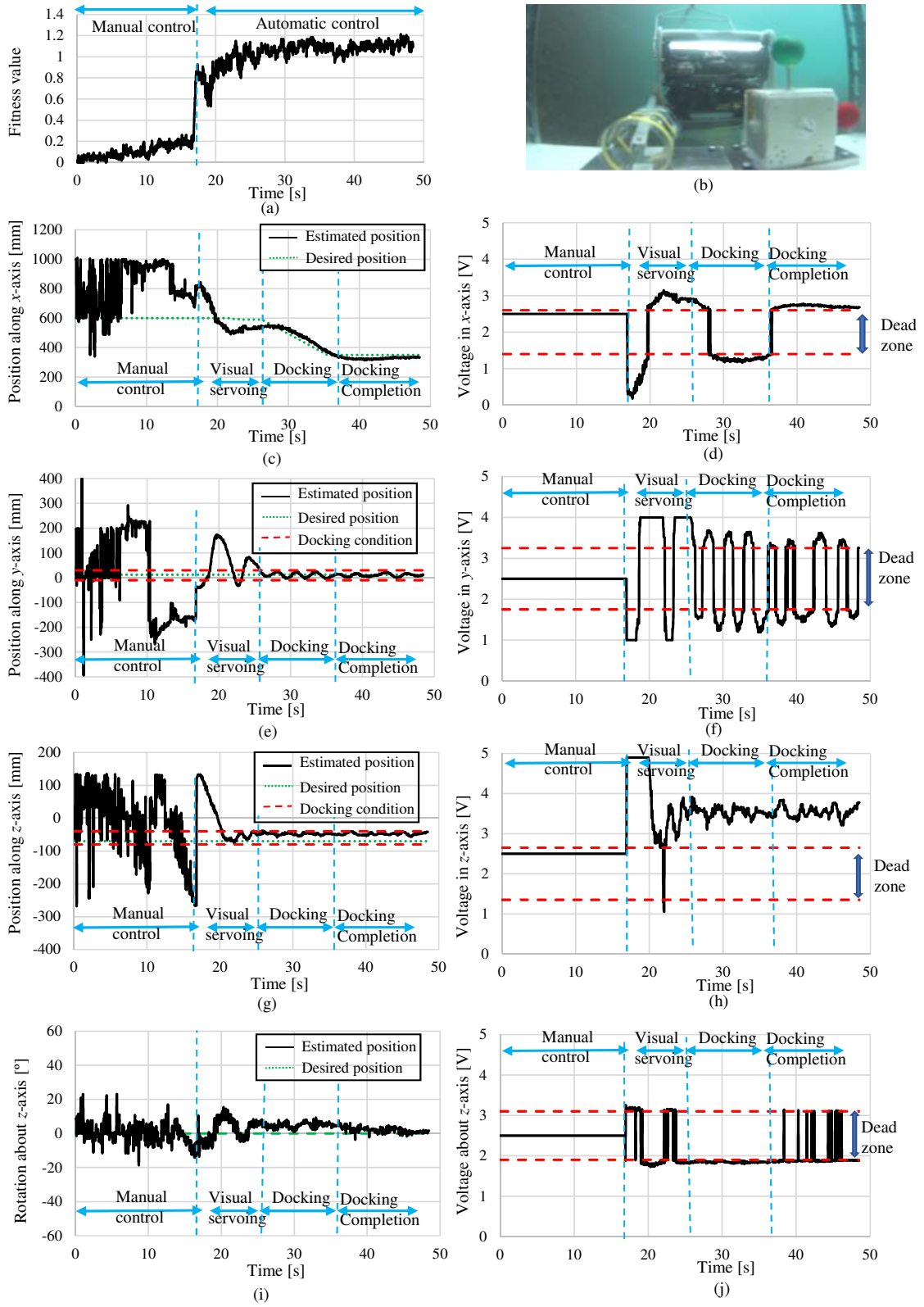


Fig. 5.16: Sea docking result 3 using a circular shaped docking hole : (a) fitness value, (b) Photograph of ROV in docking process, ((c), (e), (g), (i)) recognized positions in x,y,z axes directions and rotation around z-axis, ((d), (f), (h), (j)) recognized positions in x,y,z axes directions and rotation around z-axis.

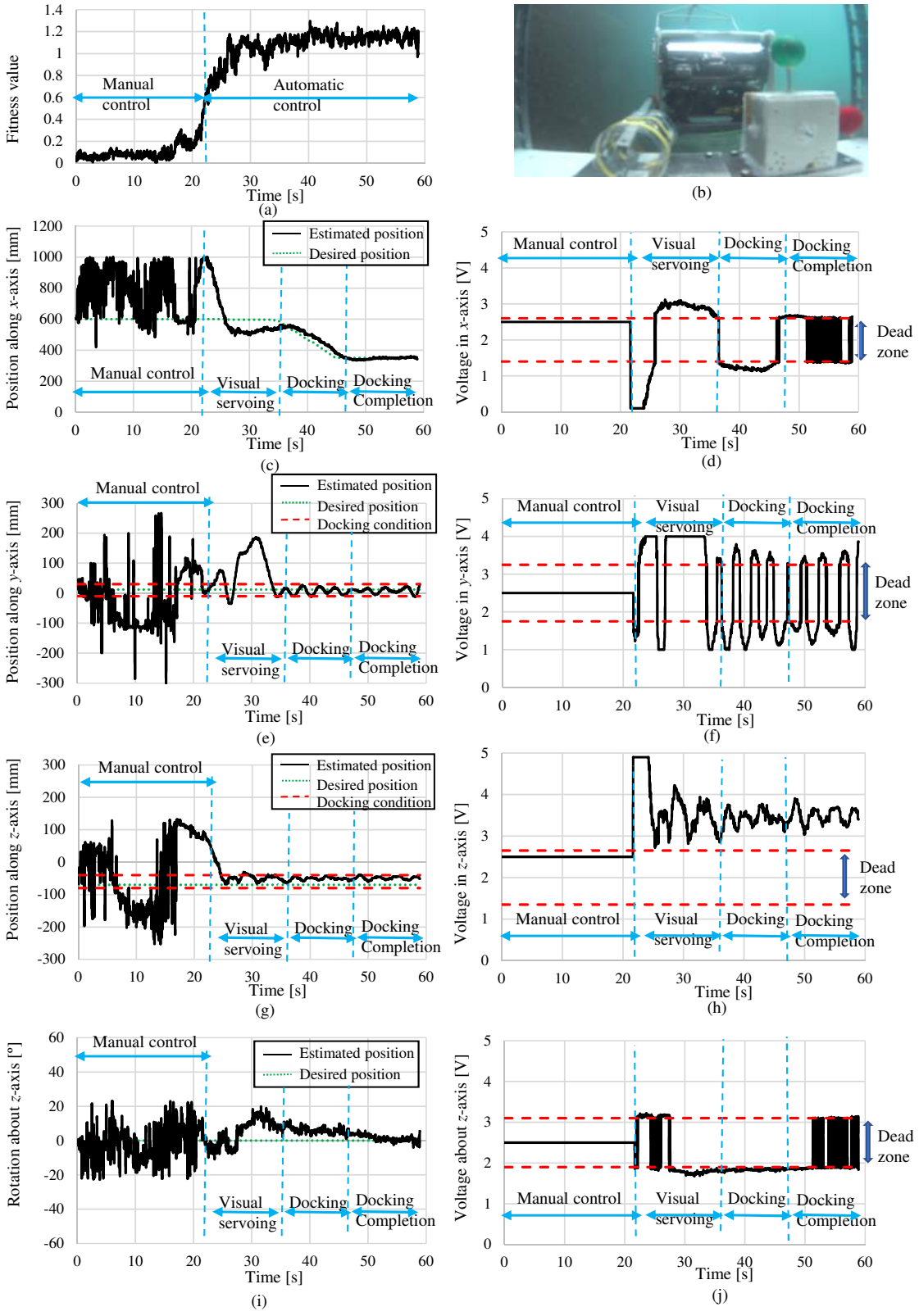


Fig. 5.17: Sea docking result 4 using a circular shaped docking hole : (a) fitness value, (b) Photograph of ROV in docking process, ((c), (e), (g), (i)) recognized positions in x,y,z axes directions and rotation around z-axis, ((d), (f), (h), (j)) recognized positions in x,y,z axes directions and rotation around z-axis.

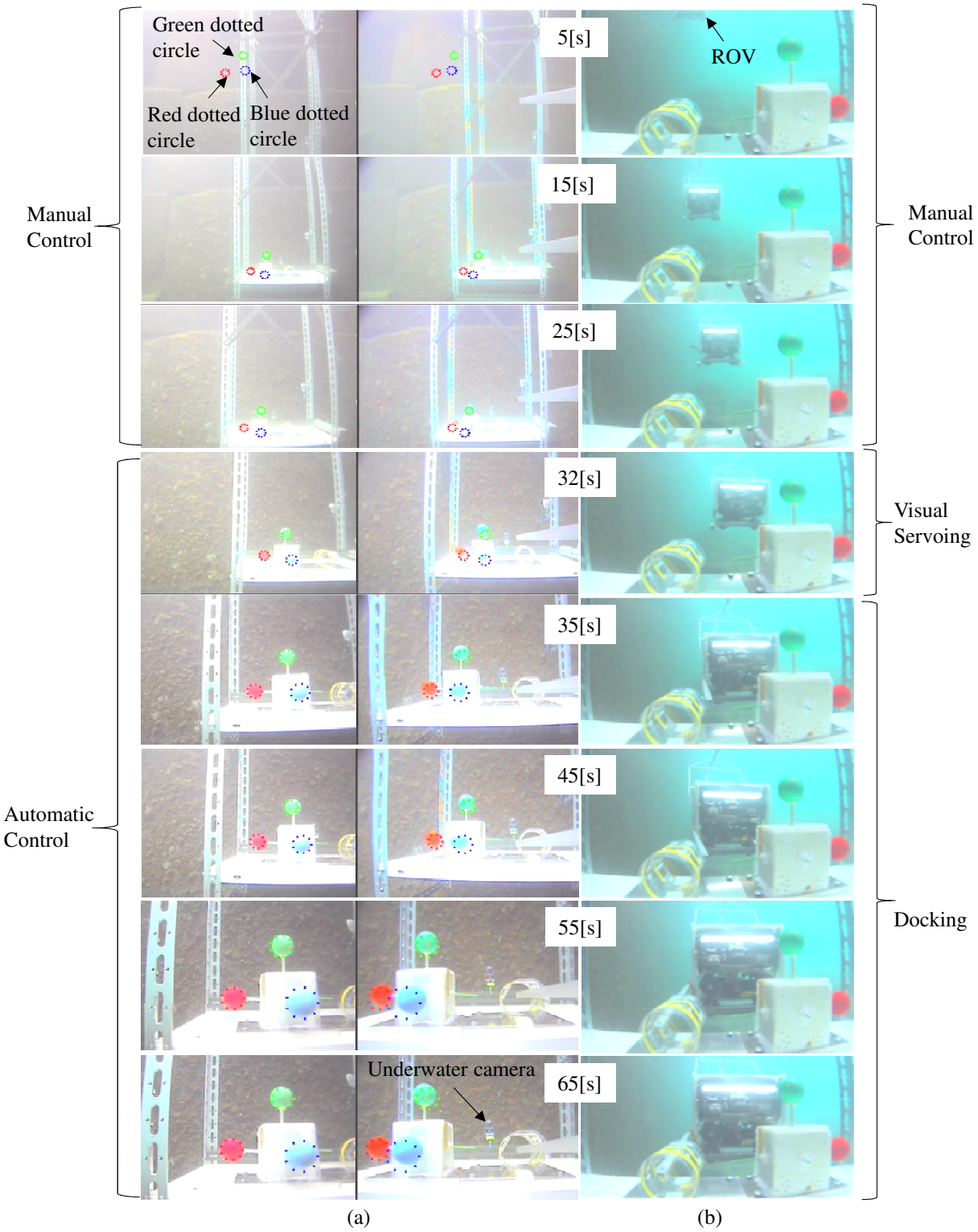


Fig. 5.18: Periodically grabbed images during the fourth time docking operation: (a) images of dual-eye cameras of ROV (b) image of underwater camera that was installed in the docking station to observe docking operation. Dotted cycles in dual-eye cameras images are the recognized poses by RM-GA.

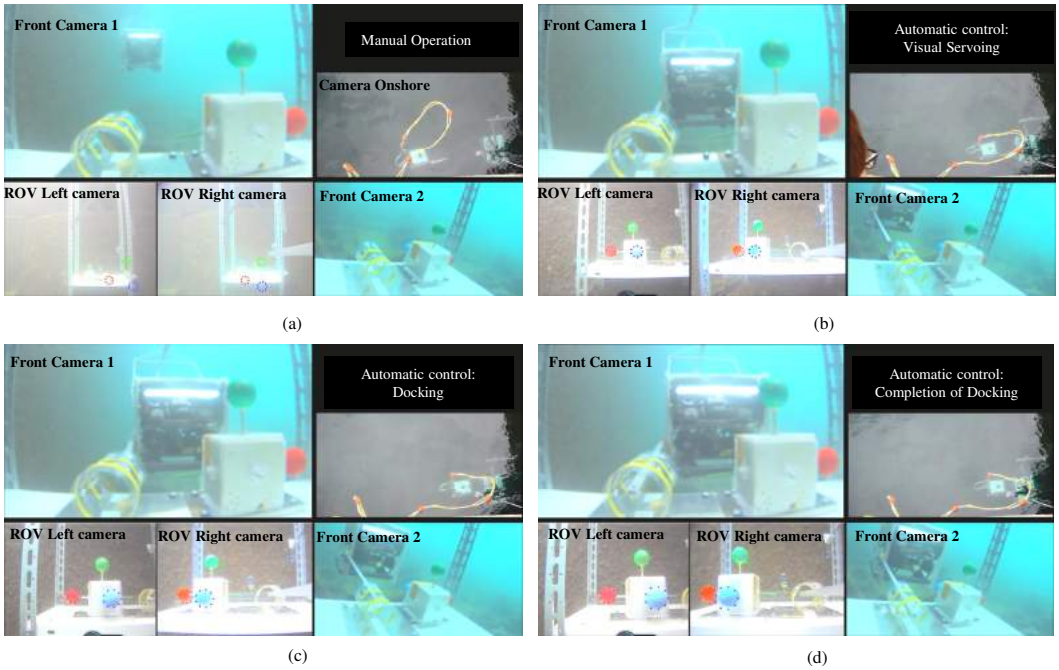


Fig. 5.19: Docking process : (a) approaching step by manual control, (b) visual servoing step, (c) docking step, and (d) docking completion.

Chapter 6

Verification of turbidity tolerance of the proposed system

Since underwater environment is more complex than space and ground, there are many disturbances for vision-based underwater vehicles. Therefore, it is important to consider the possible disturbances before testing the proposed approach in the sea. The common disturbances for vision-based underwater vehicle is turbidity. Since underwater battery recharging units are supposed to be installed in deep sea to save the time consuming and work done from human beings in the case of returning surface vehicle for recharging, the deep-sea docking experiments cannot avoid turbidity. According to the authors ' knowledge, there is no study on docking system using stereo-vision based real-time visual servoing with performance tolerance of turbidity. In this study, we conducted experiments to verify the robustness of the proposed docking approach in simulated pool where different levels of the turbidity of the water is simulated, and finally conducted sea docking in the turbid sea. The experimental results have confirmed the robustness of the docking system using stereo-vision based 3D pose estimation against turbidity.

6.1 Turbidity tolerance

The fitness distributions with respect to a position in the XY plane based on Σ_H (see Fig.3.4) are illustrated in Fig. 6.1(c) (d) corresponding to the left and right images in (a) (b). Because the pose of the target is composed of six parameters (three for position and three for orientation), the fitness distribution with a peak at the true pose can be seen in 3D space, including the fitness value and any pairs of dimensions of pose parameters, as shown in Fig. 6.1. The values shown in (a), (b) are FTU measured by a turbidity sensor and the added amount of milk in ml/m^3.

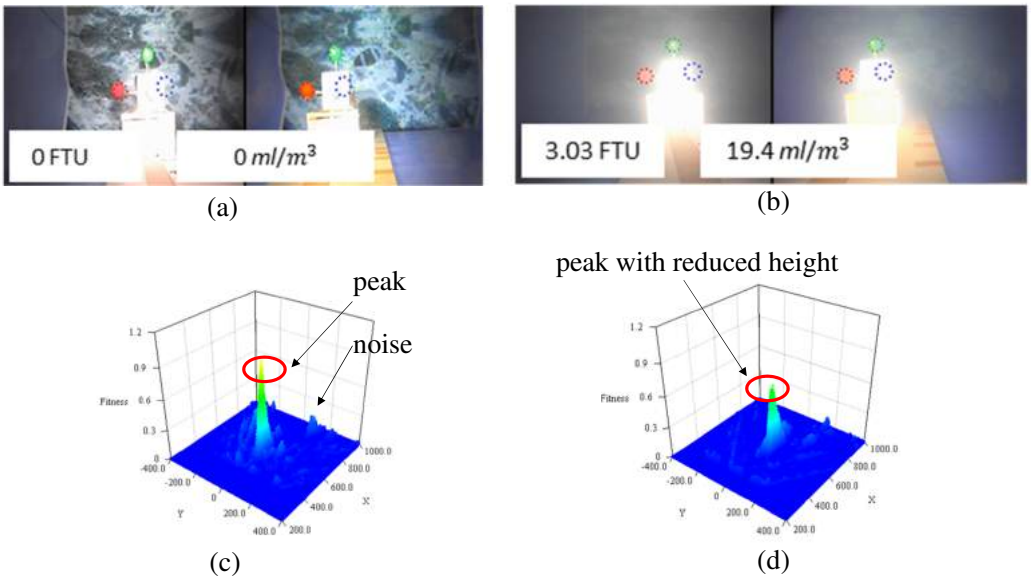


Fig. 6.1: Turbidity tolerance: photo of left and right images (a) in clean water, (b) in turbid water, (c) fitness distribution of (a), and fitness distribution of (b). The peak represents the true pose detected by the designed fitness function. The noise, which represents incorrect poses, is generated in the fitness distribution as a result of image deformation caused by environmental effects.

In the plot in Fig. 6.1 (c), there is a large peak that corresponds to the true pose, and some additional peaks that correspond to other incorrect poses are present. The proposed system can be considered robust as long as the highest peak of the fitness distribution represents the true pose and the effect of the noise that represents incorrect pose is significantly less than this peak. As shown in Fig. 6.1 (c)(d), the height of the peak is reduced when the water is turbid. However, the pose represented by a peak as

shown in Fig. 6.1 (d) is maintained. It means the system is robust against turbidity up to certain level. The shape of the fitness distribution will change in a dynamic image with a video rate of 30 fps. In the sea, turbidity is one of the greatest disturbances to visual servoing and a major source noise. When the turbidity level is high enough to render the designed fitness function ineffective, there will be no peak that represents the true pose of the target. The reason the proposed system can be considered robust against turbidity is that the problem of finding the target object and detecting its pose has been converted to an optimization problem. Therefore, it is necessary to verify the turbidity tolerance of the proposed system. To do this, it is critical to simulate the turbidity levels using a suitable medium and assess the performance of the proposed system against different turbidity levels. The turbidity tolerance of the proposed system was verified in this study, and the experimental results provide an assessment of the system performance against turbidity and demonstrate the potential of the proposed approach for actual sea docking applications.

6.2 Real-time 3D pose estimation against turbidity

6.2.1 Experimental layout

In this experiment, 3D pose recognition was conducted using the proposed system under different turbidity levels. Figure 6.2 shows the experimental layout for 3D pose estimation under different turbidity levels. In this experiment, the ROV was fixed at the same distance from the 3D marker, and illumination was kept constant by setting the two light-emitting diode (LED) units of the ROV to emit light aimed directly at the 3D marker, with an illuminance of 200 lx. The illuminance was measured using a lux sensor (model: LX-1010B manufactured by Milwaukee) placed 600 mm in front of the LED of the ROV. The experiments were conducted in a dark environment where the LED of the ROV is the dominant light source.

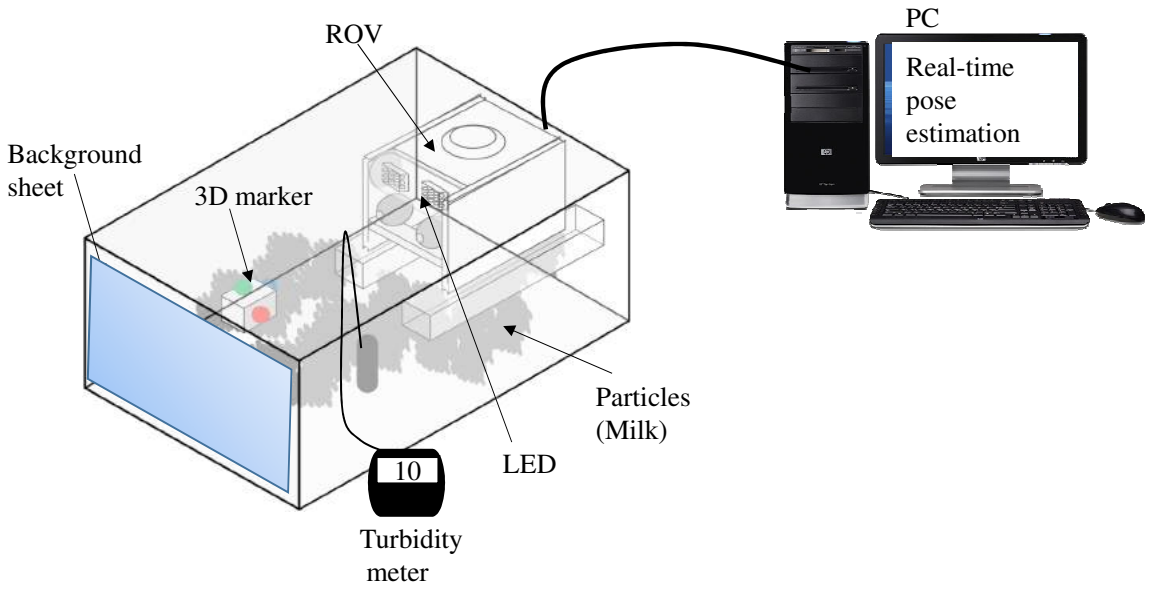


Fig. 6.2: Experimental layout for 3D pose estimation against turbidity. The system was implemented in a PC (Intel ® Core ^{TM} i7-3770 CPU 3.40 GHz, 8.00 GHz RAM, 64 bits).

Water turbidity was simulated by adding milk to the water in which the system was submerged. According to previous literature reviews [83],[84], the diameter of milk molecules ranges from 10 to 600 nm. Particles with a diameter of 10 nm scatter equal amounts of light forward and backward. The forward scattering begins to dominate for particles with diameters of approximately 100 nm, and close to 1000 nm, there is strong small-angle forward scattering and weak backscattering. Therefore, milk was selected to model the turbidity because it can provide all types of scattering. On the basis of the maximum milk concentrations of 0.19 ml/l (190 ml milk in 1000 l water) in [83] and 1.5 ml/l in [84], the experiments in this study were conducted with milk concentrations of up to 0.12 ml/l (95 ml milk in 800 l water). Note that the light sources used in [83] and [84] are different from that used here. Two fluorescent light strips were used in [83], and a halogen lamp was used in [84]. In the present experiment, two LED units installed on the ROV were used as a light source. The ROV was positioned in front of the 3D marker at a fixed distance ranging from 400 to 1000 mm. During an actual docking operation, the ROV approaches the 3D marker from a distance of approximately 1000 mm. It then performs visual servoing and the final docking stage from distances of approximately 600

and 400 mm, respectively. Details of this docking process are provided in Section 5.1.

To provide an experimental environment similar to a real undersea environment, a background sheet printed with an image similar to what would be observed in a real sea environment was placed behind the 3D marker, as shown in Fig. 6.2. The pool size is 1580 mm × 1100 mm × 590 mm, and the pool was filled with 800 l of water. Milk was added to the water in increments of 2 and 4 g up to a total of 30 and 98 g, respectively, to run the experiment at different levels of turbidity. The turbidity of the water was measured using a turbidity sensor (model: TD-500 manufactured by OPTEX) with a measurement range of 0.0 to 500 FTU (Formazin Turbidity Unit). The relationship between the measured turbidity and the milk concentration is illustrated in Fig.6.3.

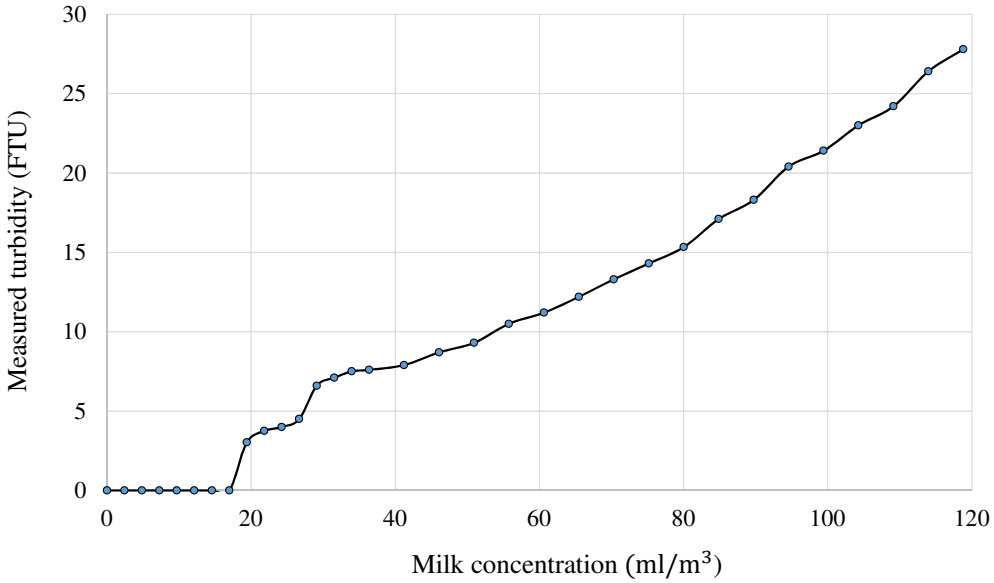


Fig. 6.3: The relationship between the measured turbidity and the milk concentration.

6.2.2 Evaluation of 3D recognition

In this experiment, the fitness value was used to evaluate the performance of the proposed pose recognition method at different turbidity levels. The correlation function of the real target projected onto the camera images and the assumed model, which was represented by poses in the chromosomes, is used as the fitness function in the GA process, in which selection, crossover, and mutation are performed to reproduce the next generation via

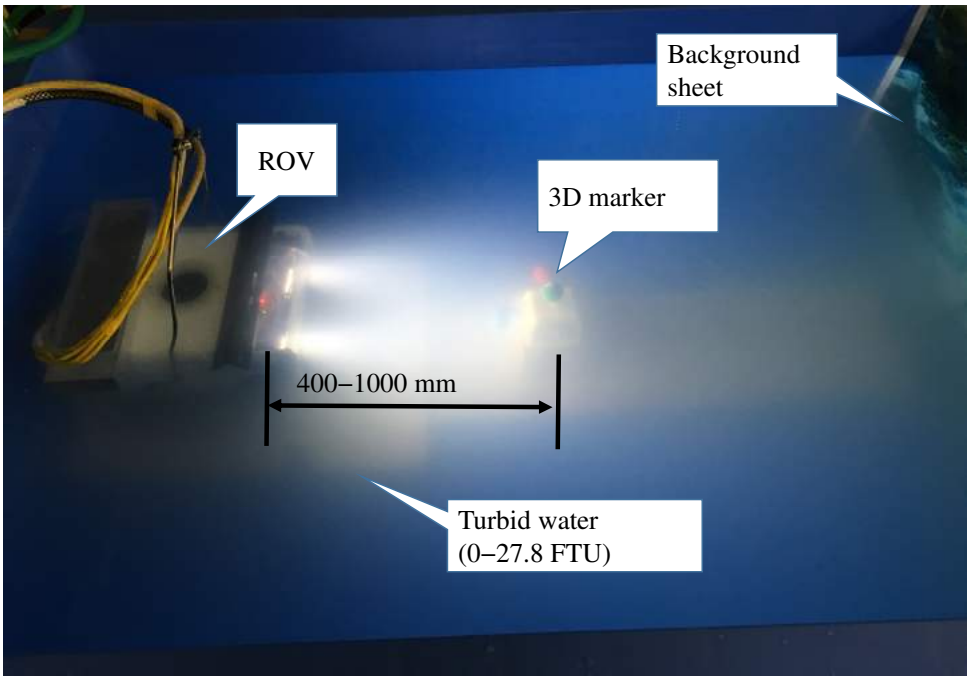


Fig. 6.4: ROV and 3D marker in turbid water. The turbidity range was from 0 to 27.8 FTU, as measured by the turbidity sensor, and the distance between the ROV and the 3D marker was varied from 400 to 1000 mm.

evaluation by the fitness function. The fitness function was modified based on the voting performance and the characteristics of the target (color, size, and shape). In this study, the fitness value averaged over 60 s was used to verify the performance of the proposed system under different turbidity levels. In addition to using the fitness value to evaluate the recognition performance, the recognized pose was visually evaluated by projecting the spheres on the target onto the left and right camera images, as shown by the dotted circles in Fig. 6.5. Moreover, to evaluate the accuracy of RM-GA, a full search method in which the vehicle searches globally for the 3D marker is performed using left and right cameras images for off-line analysis.

6.3 Result and discussion

A total of 132 iterations of this experiment were conducted at different turbidity levels and distances between the ROV and the 3D marker. The turbidity tolerance in terms of the fitness value under each set of conditions was analyzed. The maximum amount of

milk added to the 800 l of water in the pool was 98 g (118.825 ml/m^3); this corresponds to a maximum turbidity of 27.8 FTU, as measured by the turbidity sensor. The maximum amount of milk was selected as the value at which the system could not detect the 3D marker from a distance of 400 mm. Table 6.1 gives the fitness values at all considered turbidity levels and distances between the ROV and the 3D marker. Here, the fitness distribution is named as TD parameter space (T and D stand for turbidity and distance respectively). The first column gives the turbidity level measured by the turbidity sensor in units of FTU and the corresponding amount of milk in units of milliliters per cubic meter. The remaining columns give the fitness values at each of the given turbidity levels with the distance between the ROV and the 3D marker ranging from 400 to 1000 mm. The fitness values were calculated by averaging the fitness values over a real-time recognition period of 60 s in each case. Figure 6.6 shows the time profiles of the real-time and average fitness values for some illustrative cases.

There exists a fitness value threshold below which the ROV cannot reliably recognize the 3D marker. Thus, in an actual docking scenario, the ROV would not be controlled by visual servoing when the fitness value is less than the threshold. Area I, shown in blue in Table 6.1, represents the control area ($F \geq 0.6$), which is the area in which the ROV can be controlled by visual servoing. This upper fitness value threshold of 0.6, hereafter called the control threshold, was determined experimentally. Area II, shown in yellow, is the

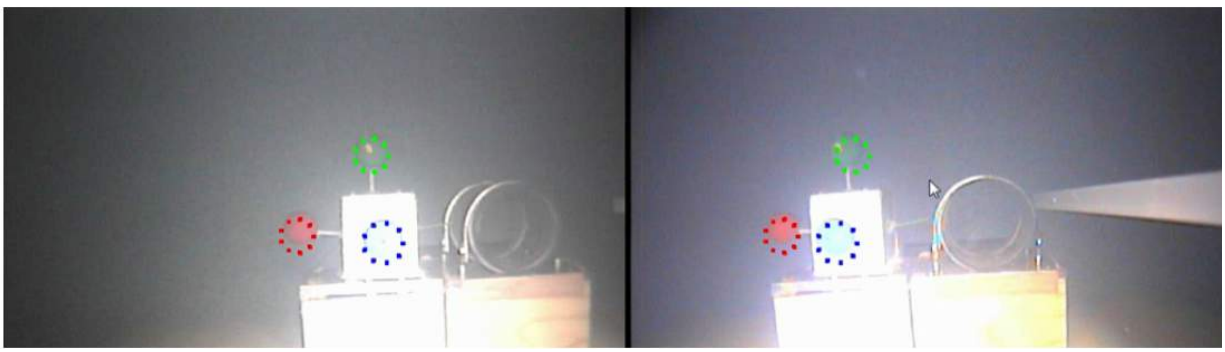


Fig. 6.5: Projection of the recognized pose onto images taken by the left and right cameras with dotted spheres indicating the positions of the three spheres for user visualization during experiments and analysis.

Table 6.1: TD parameter space: Average fitness value distribution for different turbidity levels and distances between the ROV and 3D marker (The first column gives the turbidity level measured by the turbidity sensor (FTU) and the corresponding amount of milk (ml/m^3). The fitness values are given at each of the considered turbidity levels and with the distance between the ROV and 3D marker ranging from 400 to 1000 mm. Area I in blue represents the control area ($F \geq 0.60$). Area II in yellow represents the recognition area ($0.22 \leq F < 0.60$). The remaining white area, Area III, is the loss of recognition area ($F < 0.22$). The labels A–F represent the conditions in which docking experiments were conducted in another pool.)

Turbidity \ Distance	400 mm	600 mm	800 mm	1000 mm
0.00 FTU (0.00 ml/m^3)	0.680	(A) 0.730	0.775	0.748
0.00 FTU (2.43 ml/m^3)	0.873	0.886	0.759	0.750
0.00 FTU (4.85 ml/m^3)	0.766	0.834	0.721	0.589
0.00 FTU (7.28 ml/m^3)	0.921	0.817	0.660	0.555
0.00 FTU (9.70 ml/m^3)	0.849	0.746	0.573	0.420
0.00 FTU (12.13 ml/m^3)	0.764	(B) 0.744	0.498	0.490
0.00 FTU (14.55 ml/m^3)	0.739	0.700	0.672	0.455
0.00 FTU (16.98 ml/m^3)	0.708	0.654	0.528	0.301
3.03 FTU (19.40 ml/m^3)	0.697	0.644	0.535	0.283
3.75 FTU (21.83 ml/m^3)	0.671	0.750	0.327	0.0619
4.00 FTU (24.25 ml/m^3)	0.687	(C) 0.704	0.395	0.0575
4.50 FTU (26.68 ml/m^3)	0.673	0.710	0.325	0.133
6.60 FTU (29.10 ml/m^3)	0.682	0.654	0.380	0.0596
7.10 FTU (31.53 ml/m^3)	0.667	0.645	0.298	0.298
7.50 FTU (33.95 ml/m^3)	0.666	(D) 0.651	0.214	0.0556
7.60 FTU (36.38 ml/m^3)	0.646	0.589	0.257	0.0572
7.90 FTU (41.23 ml/m^3)	0.639	0.593	0.182	0.183
8.70 FTU (46.08 ml/m^3)	0.606	0.582	0.159	0.0579
9.30 FTU (50.93 ml/m^3)	0.618	(E) 0.527	0.157	0.0552
10.50 FTU (55.78 ml/m^3)	0.578	0.294	0.124	0.135
11.20 FTU (60.63 ml/m^3)	0.578	(F) 0.216	0.098	0.151
12.20 FTU (65.48 ml/m^3)	0.545	0.262	0.130	0.118
13.30 FTU (70.33 ml/m^3)	0.565	0.0876	0.137	0.167
14.30 FTU (75.18 ml/m^3)	0.446	0.217	0.0529	0.0599
15.30 FTU (80.03 ml/m^3)	0.549	0.150	0.0582	0.153
17.10 FTU (84.88 ml/m^3)	0.371	0.118	0.0604	0.166
18.30 FTU (89.73 ml/m^3)	0.485	0.144	0.149	0.152
20.40 FTU (94.58 ml/m^3)	0.444	0.120	0.189	0.223
21.40 FTU (99.43 ml/m^3)	0.422	0.155	0.148	0.145
23.00 FTU (104.28 ml/m^3)	0.396	0.149	0.167	0.0599
24.20 FTU (109.13 ml/m^3)	0.100	0.126	0.129	0.140
26.40 FTU (113.98 ml/m^3)	0.0892	0.159	0.166	0.169
27.80 FTU (118.83 ml/m^3)	0.172	0.0919	0.210	0.215

recognition area ($0.22 \leq F < 0.6$), in which the ROV can still recognize the 3D marker but can no longer be reliably controlled using visual servoing. According to the experimental results, the system cannot recognize the 3D marker when the fitness value is less than this lower threshold of 0.22, hereafter called the recognition threshold. Area III in Table 6.1, the loss of recognition area, represents the cases in which the fitness value is below the recognition threshold. Figure 6.7 shows examples of the left and right camera images with the recognized pose of the 3D marker under conditions near the control and recognition thresholds. Example images at the minimum and maximum considered distances and turbidity levels are also shown for comparison. Figure 6.8 shows examples of the left and right camera images under different turbidity levels at a distance of 600 mm between the ROV and 3D marker. According to the experimental results, the system can recognize the 3D marker up to a turbidity of 12.2 FTU at this distance, which corresponds to the recognition threshold (Table 6.1). Similarly, the maximum turbidity at which the marker can be recognized for each distance is given in Table 6.1.

Figure 6.9 shows the fitness distributions at different turbidity levels for a distance of 600 mm between the ROV and 3D marker. The height of the peak of the fitness distribution decreased with increasing turbidity. However, in areas I and II, the pose represented by the peak corresponded with the true pose even though the height of the peak was reduced by increasing turbidity. These experimental results confirm that the proposed system is robust against turbidity up to 12.2 FTU at a distance of 600 mm. When the turbidity level reached 20.4 FTU, there was no peak at the true pose (Fig. 6.9), indicating that the 3D marker could not be recognized. In a real undersea environment, it is necessary for the ROV to be able to determine whether the turbidity level is too high to conduct the docking operation. To enable this independent determination, the control threshold could be defined based on the experimental results. The recognition results given in terms of the fitness are useful as a turbidity indicator during docking operations. For example, the ROV can return to the sea surface or wait until the turbidity level is low enough to dock and recharge its battery.

6.4 Docking experiment under turbidity in a pool

The docking experiment was conducted under turbid conditions in a large pool (length 2870 mm × width 2000 mm × height 1000 mm) filled with 4000 l of clear water. Experiments were conducted in a dark environment, and turbidity was simulated by adding milk in amounts ranging from 0 to 250 g in increments of 50 g, corresponding to the concentration conditions labeled Ⓐ–Ⓕ in Table 6.1; these docking experiments are hereafter referred to as experiments A–F, respectively. A photograph of this experiment is shown in Fig. 6.10. This experiment was conducted to verify whether each of the considered conditions were in the control area. In this system, the images acquired from the dual-eye camera are sent to the PC. The real-time recognition of the 3D pose of the target is then executed using the model-based matching method and the RM-GA in the PC software. Finally, based on the error between the actual and recognized poses, command signals generated by a position controller for the thrusters are input into the ROV to ensure it maintains the target pose. In the docking experiments, the ROV is placed at a distance of approximately 600 mm in front of the 3D marker with an arbitrary initial orientation. The docking alignment process is shown in Fig. 5.3 along with the marker and ROV coordinate systems.

Among the six docking experiments conducted in the large pool, experiments B and E (milk concentrations of 12.125 and 48.5 ml/m^3, respectively) are discussed in detail. Figures 6.11 and 6.12 show the fitness values for experiments B and E, respectively. The results obtained using the full-search and RM-GA methods were compared for some sample points during the visual servoing and docking steps and after docking completion. The poses estimated by all of the genes in the RM-GA are represented by black points, and the peak fitness value represents the estimated pose. The fitness distribution for each pose was searched using the full search method, which involves scanning all planes to find the true pose. A comparison of the fitness values obtained in experiments B and E, especially those obtained during visual servoing, reveals that increasing the turbidity

reduces the fitness value.

Figures 6.13 and 6.14 show the desired pose of the 3D marker and the pose estimated by the RM-GA. Note that the ROV is controlled manually until the 3D marker is recognized by the system. For example, the recognized poses during the first 10 s in Fig. 6.12 are not meaningful because the fitness value is less than 0.22. Therefore, visual servoing started at approximately 8 s in this case as a result of the system switching from manual mode to visual servoing mode. Visual servoing continued until the y- and z-components of the estimated position were within the error allowances of the desired pose. When these docking criteria were satisfied, the docking step was initiated. During the docking process, the rod is inserted into the docking hole by gradually decreasing the desired value of the position along the x-axis. Note that the desired position along the x-axis reduces until it reaches a distance of approximately 350 mm from the 3D marker, at which point docking is complete. In experiments A–E, the docking operation was completed successfully within 60 s after the start of visual servoing. The longest time required for completion among the five successful docking experiments was 60 s in experiment E. In experiment F, in which the turbidity level was 11.2 FTU (60.6 ml/m^3 of milk), the 3D marker could not be recognized by the system, as shown in Fig. 6.15, because the fitness value was less than 0.2. Thus, docking was successfully executed under the turbidity levels in experiments A–E, and the system failed to recognize the target under the conditions of experiment F.

6.5 Continuous iterative docking experiment in the sea

Continuous iterative at-sea docking trials were conducted near Ushimado, Japan, as shown in Fig. 6.16. The docking station (length 600 mm × width 450 mm × height 3000 mm) was oriented with its long sides perpendicular to the pier. Underwater cameras were installed in the docking station to observe the performance of the ROV during operation,

as shown in Fig. 6.17. Docking tests began with the vehicle at a distance of 1.5 m in front of the dock. A shallow sea area was selected as the docking area because the high turbidity in a shallow region would allow the verification of the robustness of the proposed system against turbidity. The turbidity level measured by the turbidity sensor during the experiment was 7.7 FTU; as indicated by Table 6.1, a turbidity of approximately 7 FTU is within the control area. The turbidity was measured at the position of 600 mm in front of the 3D marker in the sea. The depth of the sea floor in the docking area is 2.1 m. Natural waves in the sea continued while the experiments were conducted. The ROV was tethered to an onshore platform with a cable of 200 mm in length. To demonstrate the underwater battery recharging operation, a docking rod was attached to vehicle, and a docking hole affixed with a 3D marker was designed. The main task for the vehicle was to automatically insert the docking rod into the docking hole under visual servoing control. First, the vehicle was guided to the dock by manual control until the 3D marker was in the field of view (at a distance of approximately 600 mm from the target). In the visual servoing step, the vehicle took the desired pose for docking. When the vehicle stably achieved the position within an error of ± 30 mm in the image plane (y, z) for 165 ms, it began to insert the docking rod by gradually decreasing the distance between the vehicle and target in the x-direction until it reached 350 mm. After the docking operation was complete, the vehicle returned to a distance of 600 mm from the target in the x-direction for the next docking iteration.

Continuous iterative docking was conducted successfully for 19 iterations. The fitness function and desired position in the x-direction in this experiment are shown in Fig. 6.18. Among the 19 iterations, docking iteration 3, which was one of the shortest docking operations, and docking iteration 7, which was one of the longest, were analyzed in detail; the results of these two iterations are shown in Figs. 6.19 and 6.20, respectively. Figure 6.19(a), (b), and (c)–(f) shows the fitness function, the vehicle trajectory in 3D space, and the components of the recognized and desired poses, respectively, for docking iteration 3. The same results are shown in Fig. 6.20 for docking iteration 7. Docking iteration 3 was

completed successfully within 30 s. In contrast, the completion of docking iteration 7 took more than 60 s. The position along the y-axis and the rotation about the z-axis fluctuated significantly, which delayed docking completion. This fluctuation seems to have been an effect of the waves. Therefore, the vehicle trajectory in docking iteration 7 (Fig .6.20(b)) shows much larger variations than that of docking iteration 3 (Fig. 6.20(b)). As shown in Fig. 6.20(c), there was a gap between the desired and estimated positions along the x-axis because the error allowance for the docking operation is defined for only the positions along the y- and z-axes and the rotation about the z-axis. Additionally, the desired position along the x-axis remained constant for some periods during the docking step because of some fluctuations in the position along the y-axis and especially the rotation about the z-axis that exceeded the error allowance, as shown in Fig. 6.20(d) and (f). This condition triggers a switch from the docking step to the visual servoing step, as shown by the path labeled "P" in Fig. 5.2.

During the undersea docking experiments, all data were stored for offline analysis. However, the left and right camera images were stored only up to docking iteration 7 because of limitations to the memory of the PC. As shown by the experimental results of the docking iterations, the docking operations conducted in the sea at turbidity levels below 7.7 FTU were executed successfully with good agreement between the analysis of the recognition accuracy in the pool under turbid conditions and the experimental docking results; the turbidity limit of 7.7 FTU agrees well with the set of conditions labeled E in Table 6.1. A comparison of the docking performance in the sea in docking iteration 7 with that in the pool in experiment E reveals that the docking period in the sea docking experiment was nearly twice that in the pool docking experiment and the fluctuation in the pose in the sea docking experiment, especially regarding the position along the y-axis and the rotation about the z-axis (Fig. 6.20(d) and (f)), was larger than that in the pool docking experiment (Fig. 6.14). Therefore, the turbidity tolerance described in Table 6.1 for the proposed system in a pool environment was verified experimentally in a real sea environment. The control and recognition areas (areas I and II in Table 6.1) can be

expanded by improving the system in future.

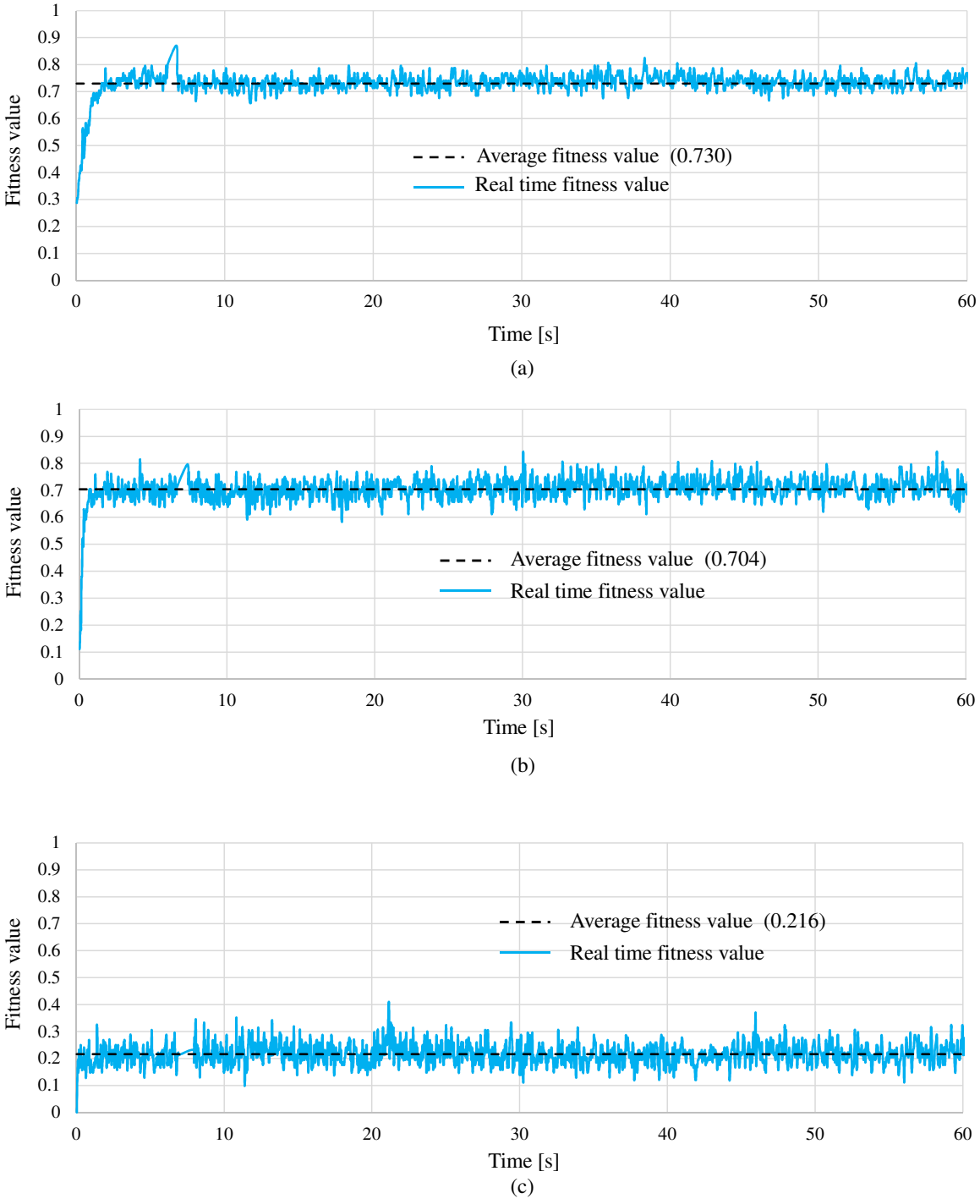


Fig. 6.6: Real-time and average fitness values under the conditions labeled (a) A, (b) C, and (c) F in Table 6.1.

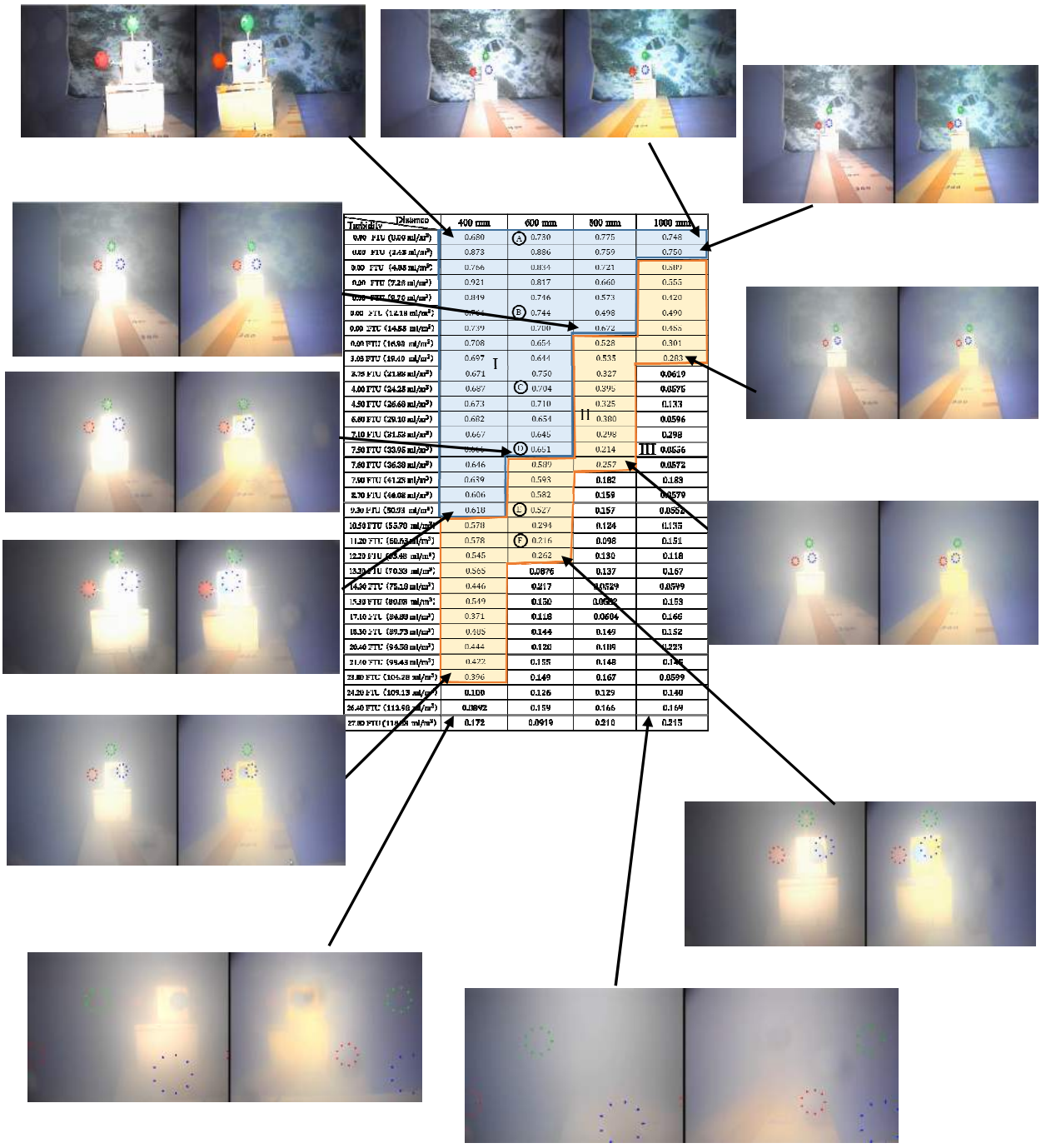
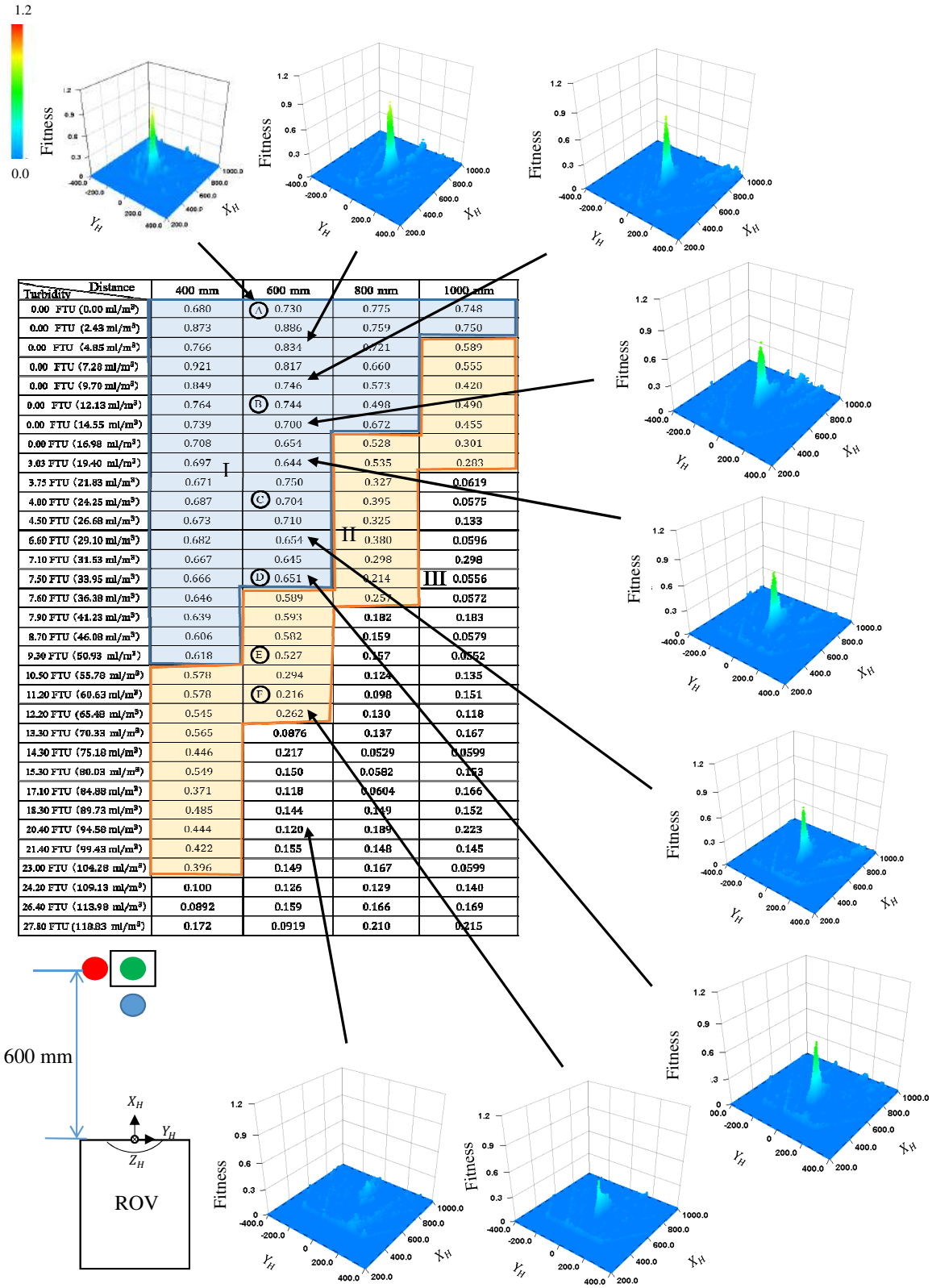


Fig. 6.7: Left and right camera images under the maximum turbidity conditions in the control and recognition areas at each considered distance. Images taken at the maximum and minimum distances in clean water and at the maximum turbidity, in which the 3D marker is not observable, are also shown at the top and bottom, respectively.

Fig. 6.8: Left and right camera images with the pose recognized by the pose estimation system at different turbidity levels and a distance of 600 mm between the ROV and 3D marker. The recognized pose is indicated by dotted circles in each photograph. The water turbidity measured by the turbidity sensor is shown in units of FTU, and the amount of added milk is given in units of milliliters per cubic meter.



Turbidity \ Distance	400 mm	600 mm	800 mm	1000 mm
0.00 FTU (0.00 ml/m³)	0.680	A 0.730	0.775	0.748
0.00 FTU (2.43 ml/m³)	0.873	0.886	0.759	0.750
0.00 FTU (4.85 ml/m³)	0.766	0.834	0.721	0.589
0.00 FTU (7.28 ml/m³)	0.921	0.817	0.660	0.555
0.00 FTU (9.70 ml/m³)	0.849	0.746	0.573	0.420
0.00 FTU (12.13 ml/m³)	0.764	B 0.744	0.498	0.490
0.00 FTU (14.55 ml/m³)	0.739	0.700	0.672	0.455
0.00 FTU (16.98 ml/m³)	0.708	0.654	0.528	0.301
3.03 FTU (19.40 ml/m³)	0.697	0.644	0.535	0.203
3.75 FTU (21.83 ml/m³)	0.671	0.750	0.327	0.0619
4.00 FTU (24.25 ml/m³)	0.687	C 0.704	0.395	0.0575
4.50 FTU (26.68 ml/m³)	0.673	0.710	0.325	0.133
6.60 FTU (29.10 ml/m³)	0.682	0.654	0.380	0.0596
7.10 FTU (31.53 ml/m³)	0.667	0.645	0.298	0.298
7.50 FTU (33.95 ml/m³)	0.666	D 0.651	0.214	0.0556
7.60 FTU (36.38 ml/m³)	0.646	0.589	0.257	0.0572
7.90 FTU (41.23 ml/m³)	0.639	0.593	0.182	0.183
8.70 FTU (46.08 ml/m³)	0.606	0.582	0.159	0.0579
9.30 FTU (50.93 ml/m³)	0.618	E 0.527	0.157	0.0552
10.50 FTU (55.78 ml/m³)	0.578	0.294	0.124	0.135
11.20 FTU (60.63 ml/m³)	0.578	F 0.216	0.098	0.151
12.20 FTU (65.48 ml/m³)	0.545	0.262	0.130	0.118
13.30 FTU (70.33 ml/m³)	0.565	0.0876	0.137	0.167
14.30 FTU (75.18 ml/m³)	0.446	0.217	0.0529	0.0599
15.30 FTU (80.03 ml/m³)	0.549	0.150	0.0582	0.153
17.10 FTU (84.88 ml/m³)	0.371	0.118	0.0604	0.166
18.30 FTU (89.73 ml/m³)	0.485	0.144	0.149	0.152
20.40 FTU (94.58 ml/m³)	0.444	0.120	0.189	0.223
21.40 FTU (99.43 ml/m³)	0.422	0.155	0.148	0.145
23.00 FTU (104.28 ml/m³)	0.396	0.149	0.167	0.0599
24.20 FTU (109.13 ml/m³)	0.100	0.126	0.129	0.140
26.40 FTU (113.98 ml/m³)	0.0892	0.159	0.166	0.169
27.80 FTU (118.83 ml/m³)	0.172	0.0919	0.210	0.215

Fig. 6.9: Fitness value distributions confirming the robustness of the system at a distance of 600 mm. The position of the peak corresponding to the true pose of the marker was maintained even though the height of the peak was reduced by increasing turbidity. The gradual reduction in the height of peak shows the effect of turbidity on image recognition.

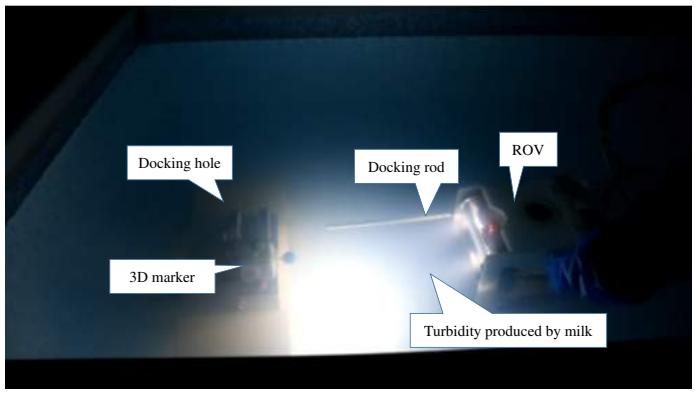


Fig. 6.10: Photograph of the docking experiment under turbid conditions in a dark environment.

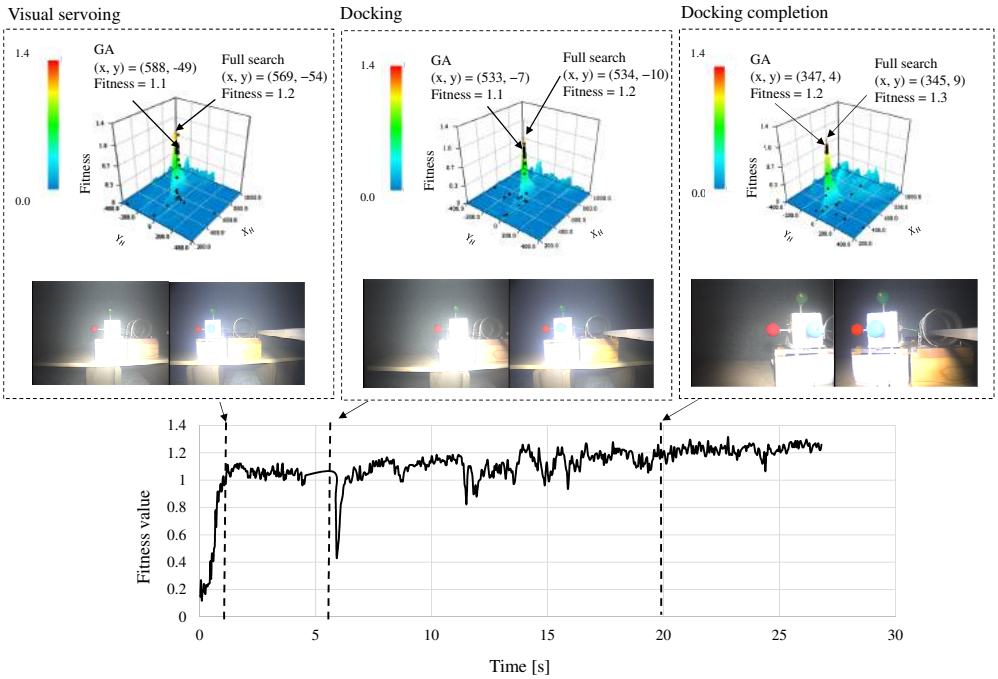


Fig. 6.11: Fitness value results for experiment B. The photographs show examples of the left and right camera images from which the pose was estimated using the RM-GA. From left to right, the photographs show selected images from the visual servoing step, the docking step, and after the completion of the docking step. The poses estimated using the RM-GA and the full-search method are indicated in the fitness value distributions for each of these docking steps. The area around the peak of the fitness distribution was searched by scanning all planes of the images. The presence of a peak in the distribution indicates the robustness of the recognition method against turbidity, and the correspondence between the peak and the black points indicates the accuracy of the RM-GA results. The black point represents each gene of RM-GA. The pose yielded by the RM-GA is shown in Fig. 6.13.

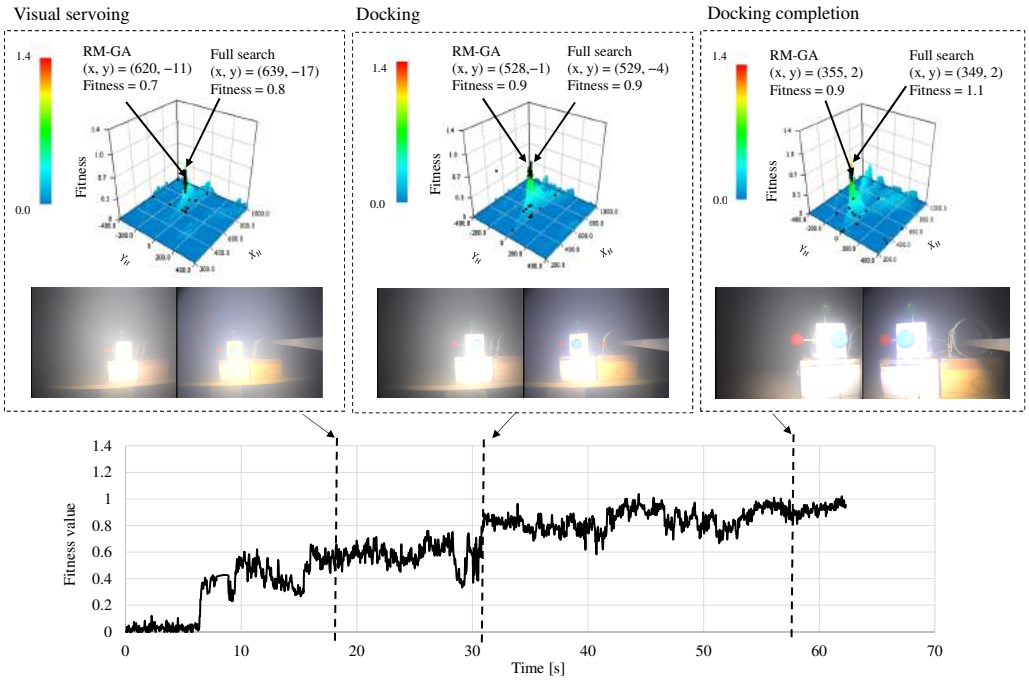


Fig. 6.12: Same as Fig. 6.11 for experiment E. The pose yielded by the RM-GA is shown in Fig. 6.14.

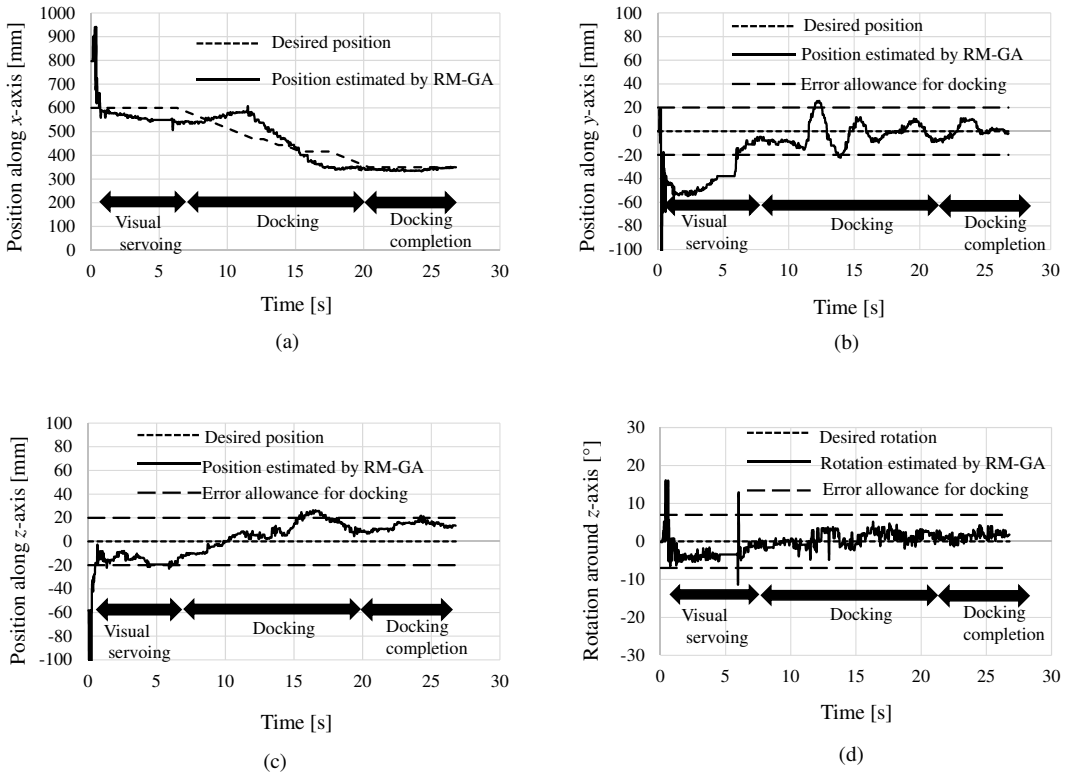


Fig. 6.13: Position along the (a) x-, (b) y-, and (c) z-axes and (d) rotation about the z-axis estimated using the RM-GA in the docking experiment for experiment B. In this case, the control threshold is 0.6.

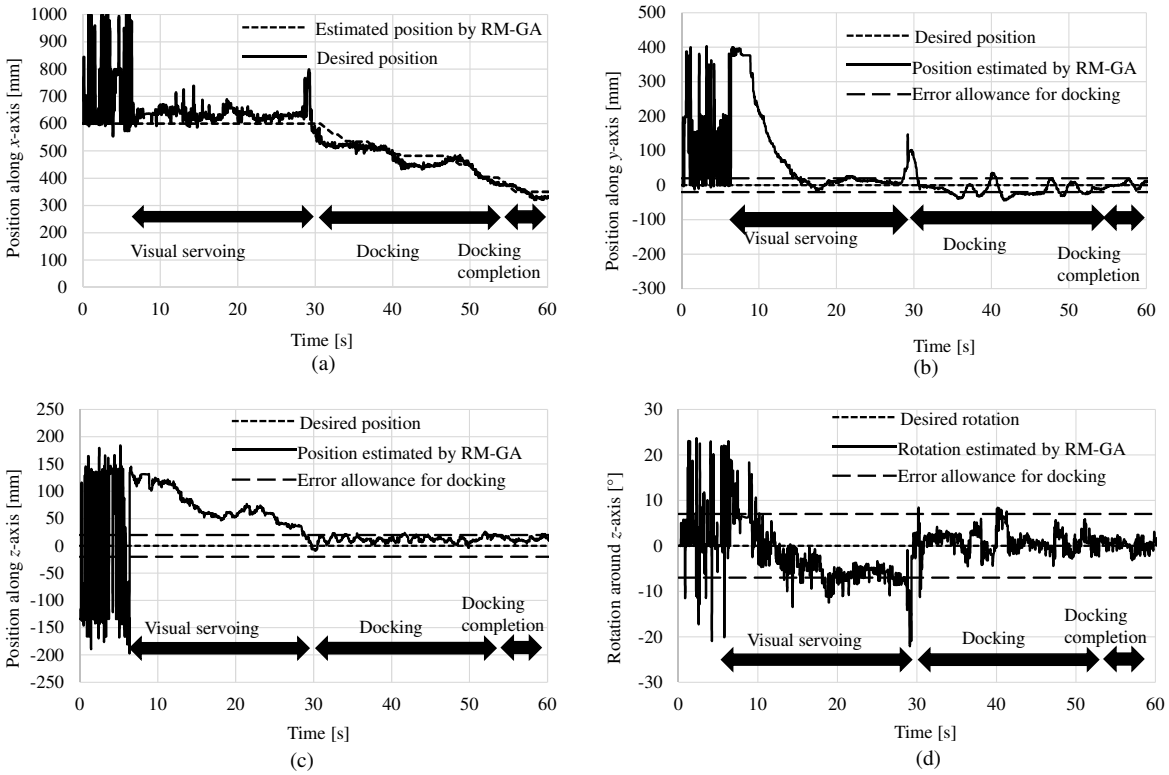
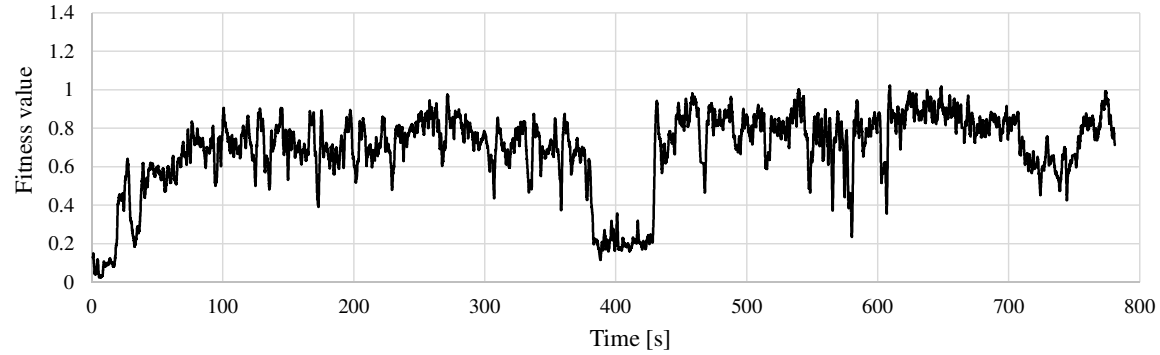


Fig. 6.14: Same as Fig. 6.13 for experiment E. In this case, the control threshold is 0.4.

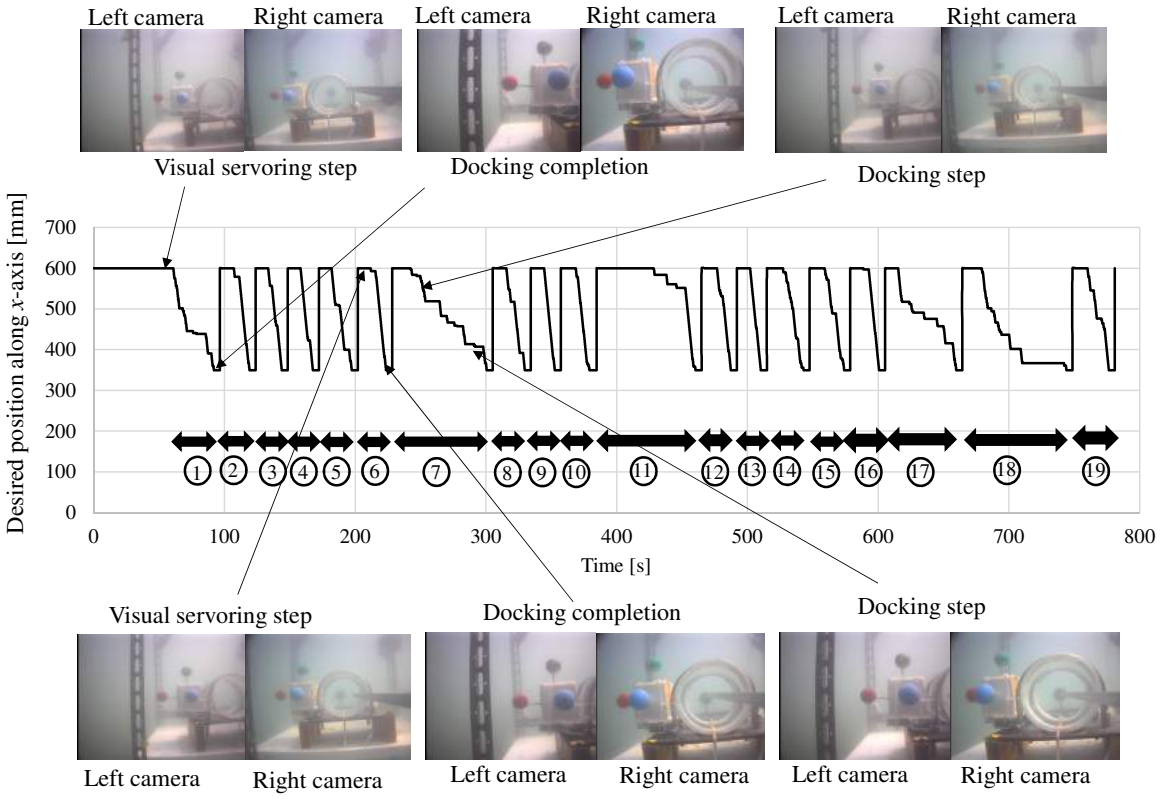
Fig. 6.15: Left and right camera images in experiment F. The dashed circles, which are not aligned with the target, represent the system's failure to recognize the target.

Fig. 6.16: ROV and docking station in the sea.

Fig. 6.17: Continuous iterative docking experiments in the sea. These photographs were taken by two underwater cameras installed in the docking station and from a pier.



(a)



(b)

Fig. 6.18: Results of continuous iterative docking experiment. (a) Fitness value plotted against time. (b) Desired position in the x-direction during 19 docking iterations in the sea. The numbers along the bottom of the plot represent the docking iteration number, and the duration of each docking iteration is represented by the length of the corresponding arrow. Examples of the left and right camera images taken during the visual servoing and docking steps and after docking completion are shown above and below the plot. Detailed results for docking iterations 3 and 7 are presented in Figs. 6.19 and 6.20, respectively.

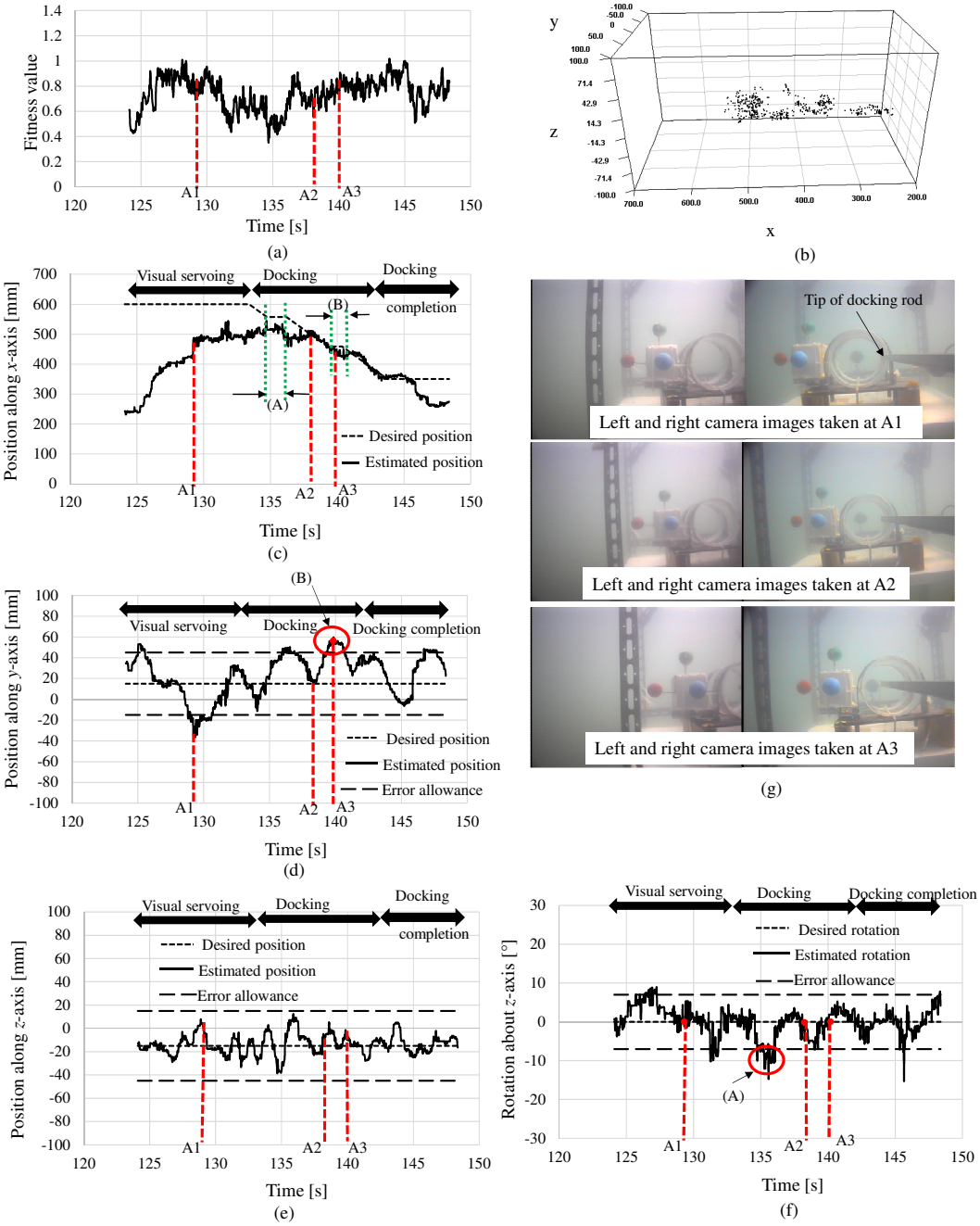


Fig. 6.19: Results for docking iteration 3. (a) Fitness value plotted against time. (b) Vehicle trajectory in 3D space. (c)–(f) Recognized position along the x-, y-, and z-axes and rotation about the z-axis obtained by the RM-GA. The desired position along the x-axis remained constant for the periods labeled (A) and (B) in (c) during docking because the rotation error about the z-axis labeled (A) in (e) and the position error in the y-direction labeled (B) in (d) respectively surpassed the error allowance. (g) Left and right camera images taken at the times labeled A1, A2, and A3 in the time profiles. These images show the movement of the ROV in the y-direction when the rotation of the ROV about the z-axis was almost zero.

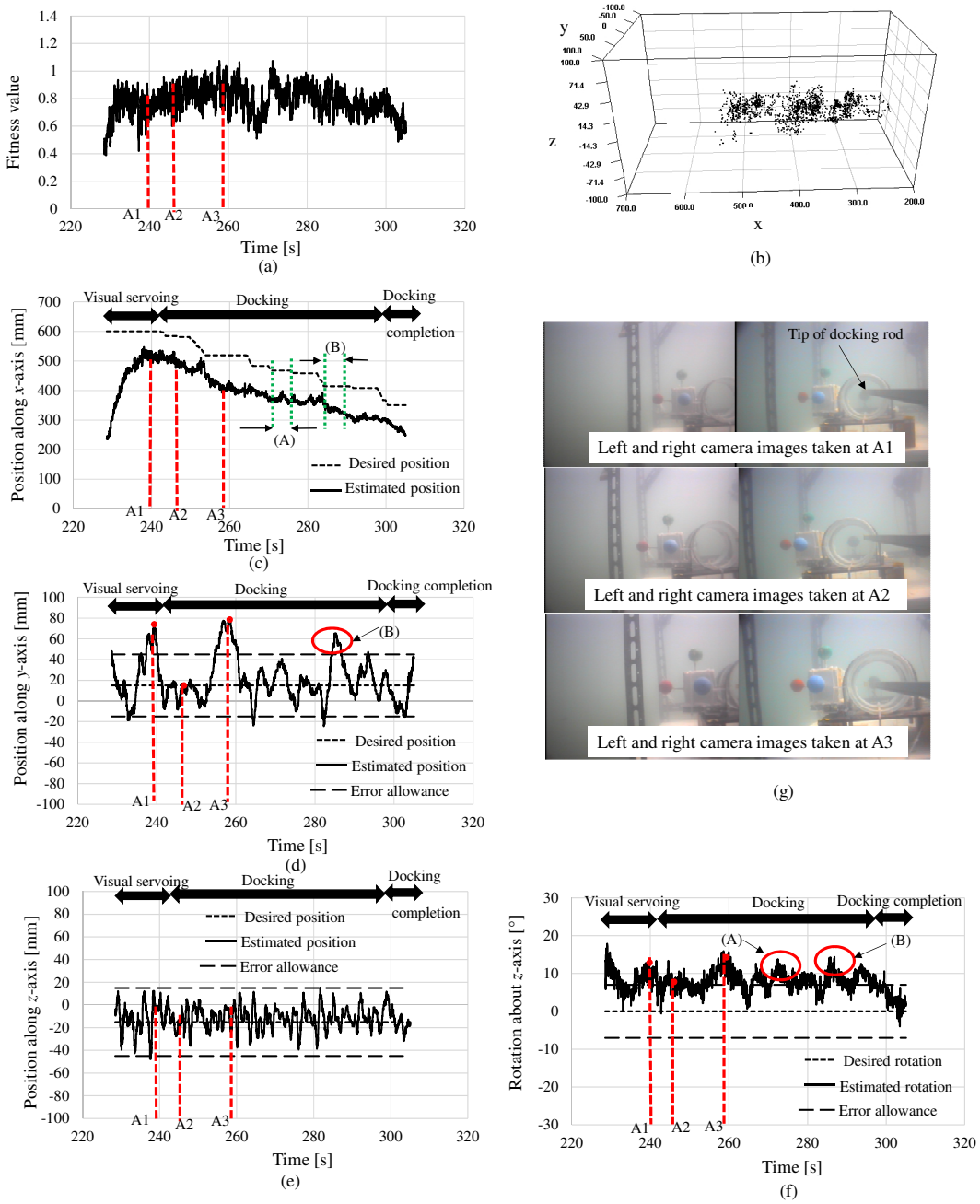


Fig. 6.20: Same as Fig. 6.19 for docking iteration 7. The desired position along the x-axis remained constant for the periods labeled (A) and (B) in (c) because the rotation error about the z-axis labeled (A) in (e) and both the position error along the y-axis and the rotation error about the z-axis labeled (B) in (d) and (e) respectively surpassed the error allowance. At the time labeled A2 in the time profiles, the ROV is at the desired position along the y-axis, and the rotation angle about the z-axis is within the error allowance. At A1 and A3, the position along the y-axis and the rotation about the z-axis both surpassed the error allowance. This indicates that the rotation about the z-axis and the position along the y-axis are coupled. Therefore, the tip of docking rod appears to be within the allowed area in the images taken at A1 and A3 even though there are some deviations in the position along the y-axis and the rotation about the z-axis.

Chapter 7

Conclusion

In this work, vision-based docking approach by using two cameras for an underwater vehicle was designed and implemented for underwater battery recharging. First, 3D pose estimation approach using RM-GA was proposed and verified in 3D pose recognition experiment. Second, the recognition accuracy and regulation performance was verified in pool tests in which regulating experiments were conducted. Since the real sea environment addresses different disturbances, the robustness against object occlusion and physical disturbances were experimentally verified. Third, docking experiment through designed docking strategy was conducted in the pool. Then, sea docking experiment was conducted using an ROV in the sea near Wakayama city in Japan. After achieving sea docking experiment, the proposed system was verified for turbidity tolerance since it is the main challenging and unavoidable issue in the sea floor where the intended underwater battery recharging unit with docking function is supposed to be installed. Therefore, experimental verification of turbidity tolerance of the proposed system was conducted and presented in this study. Finally, sea docking experiment in the turbid sea in coastal area was conducted to verify the functionality and practicality of the proposed system against real sea disturbances especially turbidity. As future works, even through the parameters of RM-GA are turned experimentally in this work to have enough accuracy, optimal parameters can be selected based on some analysis on their performance to improve the

proposed system especially in term of convergence time and recognition accuracy. Additionally, the turbidity tolerance of the proposed system using passive 3D marker is limited in some level of turbid sea environment, the system can be expanded to be able to work in higher turbid environment in future work.

Acknowledgement

This work would not have been possible without the sustained effort of the entire my team, and the commitment and management of my supervisor Professor Mamoru Minami. This work was supported by JSPS KAKENHI Grant Number JP16K06183. The development of the ROV is cooperated by MITSUI Engineering and Shipbuild-ing Co.,LTD and Kowa cooperation. Thanks to all the people in the Underwater vehicle Group for making it such an delightful and enjoyment environment.

Most unforgettably, I would like to thank my family for letting me chose my own part and always being reasonable and supportive. My final thanks to EEHE project for their scholarship support throughout theses years.

2018

Myo Myint

Bibliography

[1] Amat, J., J. Batlle, A. Casals, and J. Forest. GARBI, " a low cost ROV, constrains and solutions," In: 6me Seminaire IARP en robotique sous-marine, pp. 1-22 (1996)

[2] Ridao, P., Batlle, J., Amat, J. and Roberts, G. N, " Recent Trends in Control Architectures for Autonomous Underwater Vehicles," International Journal of Systems Science, Vol.30, No. 9, pp 1033-1056 (1999)

[3] Ridao, P., Batlle, J., Amat, J. and Roberts, G. N, " Recent Trends in Control Architectures for Autonomous Underwater Vehicles," International Journal of Systems Science, Vol.30, No. 9, pp 1033-1056 (1999)

[4] Valavanis, P.K., Gracanin, D., Matijasevic, M., Kolluru R., and Demetriou, G.A, " Control Architectures for Autonomous Underwater Vehicles," IEEE Control Systems Magazine, pp.48-64 (1997)

[5] Tena-Ruiz I., S. Raucourt, Y. Petillot, D.M. Lane, " Concurrent Mapping and Localization Using Sidescan Sonar," IEEE Journal of Oceanic Engineering, 29(2), pp. 442-456 (2004)

[6] Palomeras N, Garca JC, Prats M, Fernndez JJ, Sanz PJ, Ridao P," A distributed architecture for enabling autonomous underwater intervention missions," Systems Conference, 2010 4th Annual IEEE, pp. 159-164 (2010)

[7] Sanz PJ, Prats M, Ridao P, Ribas D, Oliver G, Orti A ,“ Recent progress in the RAUVI project. a reconfigurable autonomous underwater vehicle for intervention,” 52-th International Symphosium ELMAR-2010, pp. 471-474 (2010)

[8] Warren, D. J., Church, R. A., and Eslinger, K. L.,“ Deepwater archaeology with autonomous underwater vehicle technology,” In Offshore Technology Conference, Houston, TX. (2007)

[9] 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)

[10] Steve Cowen, Susan Briest and James Dombrowski, “Underwater Docking of Autonomous Undersea Vehicle using Optical Terminal Guidance,” Proc. MTS/IEEE OCEANS Conf., vol.2, pp.1143-1147 (1997)

[11] 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, pp.349-361 (2015)

[12] Michael D. Feezor, 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, pp.515-521 (2001)

[13] 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 (2008)

[14] Ken Teo, E. An and P.-P. J. Beaujean, “A robust fuzzy autonomous underwater vehicle (AUV) docking approach for unknown current disturbances,” IEEE Journal of Oceanic Engineering, vol. 37, no. 2, pp. 143-155 (2012)

[15] Amaury Negre, Cedric Pradalier and 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 (2008)

[16] B. Allen, T. Austin, N. Forrester, R. Goldsborough, A. Kukulya, G. Packard, M. Purcell, R. Stokey: "Autonomous docking demonstrations with enhanced REMUS technology," Proc. IEEE OCEANS'06 (2006) pp. 1-6

[17] 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 .

[18] 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," OCEANS 2007, Europe, pp 1-6, 2007.

[19] M. Wirtz, M. Hildebrandt, and C. Gaudig, "Design and Test of a Robust Docking System for Hovering AUVs," OCEANS 2012, pp 1-6, 2012.

[20] 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," IEEE Journal of Oceanic Engineering, vol. 36, no. 1, pp. 48-61 (2009)

[21] Palomeras, N., Ridao, P., Ribas, D. and Vallicrosa, G., "Autonomous I-AUV docking for fixed-base manipulation," in Proc. International Federation of Automatic Control, vol.47, no.3, pp.12160-12165 (2014)

[22] Foresti, G.L., Gentili, S., Zampato, M.: "A vision-based system for autonomous underwater vehicle navigation," Proc. MTS/IEEE OCEANS Conf. 1, 195-199 (1998)

[23] J. Evans, P. Redmond, C. Plakas, K. Hamilton, and D. Lane, "A Autonomous docking for intervention-AUVs using sonar and video-based real-time 3D pose estimation," MTS/IEEE OCEANS Conf.,vol. 4, pp. 2201-2210 (2003)

[24] L. Brignone, M. Perrier, and C. Viala, "A fully autonomous docking strategy for intervention AUVs," Proc. MTS/IEEE OCEANS Conf.-Eur., , June 2007, DOI: 10.1109/OCEANS.2007.4302290.

[25] X. Bian, X. Wang, W. Zhang, and Z. Van, "Design of a visual guiding system for IADV docking," Proc. IEEE Int. Conf. Mechatron. Autom., 2009, pp. 2107-2112 (2009)

[26] P. B. Sujit, A. J. Healey, and J. B. Sousa, "AUV docking on a moving submarine using a K-R navigation function," Proc. IEEE/RSJ Int.Conf. Intell. Robots Syst., pp. 3154-3159 (2011)

[27] G. J. S. Rae and S. M. Smith, "A fuzzy rule based docking procedure for autonomous underwater vehicles," Proc. OCEANS Conf., 1992, vol. 2, pp. 539-546 (1992)

[28] P. Jantapremjit and P. A. Wilson, "Control and guidance for homing and docking tasks using an autonomous underwater vehicle," Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst., pp. 3672-3677 (2007)

[29] J.-Y. Park, B.-H. Jun, P.-M. Lee, and Y.-K. Lim, "Docking problem and guidance laws considering drift for an underactuated AUV," Proc. IEEE OCEANS Conf., DOI: 10.1109/Oceans-Spain.2011 (2011) 6003574.

[30] J.-Y. Park, B.-H. Jun, P.-M. Lee, and Y.-K. Lim, "Modified linear terminal guidance for docking and a time-varying ocean current observer," Proc. IEEE Symp. Underwater Technol., DOI: 10.1109/UT.2011.5774141 (2011)

[31] G. J. S. Rae and S. M. Smith, "A fuzzy rule based docking procedure for autonomous underwater vehicles," Proc. MTS/IEEE OCEANS Conf., vol. 2, pp. 539-546 (1992)

[32] T. Fukasawa, T. Noguchi, T. Kawasaki, and M. Baino, "Marine bird - A new experimental AUV with underwater docking and recharging system," Proc. MTS/IEEE OCEANS Conf., San Diego, CA, pp. 2195-2200. (2003)

[33] H. Singh, J. G. Bellingham, F. Hover, S. Lerner, B. A. Moran, K. Heydt, and D. Yoerger, "Docking for an autonomous ocean sampling network," IEEE J. Ocean. Eng., vol. 26, no. 4, pp. 498-513 (2001)

[34] R. Coulson, J. Lambiotte, and E. An, "A modular docking system for 12.75-inch class AUVs," Sea Technol., pp. 49-54 (2005)

[35] A. Martins, J. Almeida, H. Ferreira, H. Silva, N. Dias, A. Dias, C. Almeida, and E. Silva, "Autonomous surface vehicle docking manouvre with visual information," Proc. International Conference on Robotics and Automation, pp. 4994-4999 (2007)

[36] J. G. Bellingham, R. S. McEwen, and B. W. Hobson,"The development of a docking station for a regional cabled observatory," in Proc. 4th Int. Workshop Scientific Use of Submarine Cables, Dublin, Ireland, Feb. 2006, pp. 150-155 (2006)

[37] R. Coulson, J. Lambiotte, and E. An, "A modular docking system for 12.75-inch class AUVs," Sea Technol., pp. 49-54 (2005)

[38] H. Singh, M. Bowen, F. Howver, P LeBas, D. Yoerger, "An Intelligent Dock for an Autonomous Ocean Sampling Network," Proc Oceans 97, MTS/IEEE, Halifax, (1997)

[39] Deltheil, C., Didier, L., Hospital, E., Brutzman, D.P., "Simulating an optical guidance system for the recovery of an unmanned underwater vehicle," IEEE Journal of Oceanic Engineering 25 (4), 568-574 (2000)

[40] Oh, M.H., Oh, J.H., " Homing and docking control of AUV using model predictive control," Proceedings of the Fifth ISOPE Pacific/Asia Offshore Mechanics Symposium, pp. 138-142 (2002)

[41] Oh, K.H., Kim, J.Y., Park, I.W., Lee, J.H., Oh, J.H., " A study on the control of AUV's homing and docking," Proceedings of the Ninth IEEE Conference on Mechatronics and Machine Vision in Practice, pp. 45-52 (2002)

[42] Wettergreen, D., Gaskett, C., and Zelinsky, A, " Development of a visually-guided autonomous underwater vehicle," In Proc. of OCEAN'98 Conference, volume 2, pp. 1200-1204 (1998)

[43] P. Weiss, J. Catret Mascarell, D. Grosset, L. Brignone, D. Labbe, and P. Wilson, " Freesub: Dynamic stabilization and docking for autonomous underwater vehicles," pp. 1-6 (2003)

[44] R. Stokey, B. Allen, T. Austin, R. Goldsborough, N. Forrester, M. Purcell, and C. von Alt, "Enabling Technologies for REMUS Docking: An Integral Component of an Autonomous Ocean Sampling Network," IEEE Journal of Oceanic Engineering, Vol 26, pp. 487-497 (2001)

[45] J.C.Evans, K.M.Keller, J.S.Smith, P.Marty, and O.V.Rigaud, "Docking techniques and evaluation trials of the SWIMMER AUV: An autonomous deployment AUV for work-class ROVs," in Proc. Oceans MTS/IEEE, vol. 1, pp. 520-528 (2001)

[46] Young-Hwa Hong, Jung-Yup Kim, Pan-Mook Lee, Bong-Hwan Jeon, Kyu-Hyun Oh and Jun-ho Oh, "Development of the Homing and Docking Algorithm for AUV," ISOPE 2003. International Offshore and Polar Engineering Conference, 25-30, pp. 205-212 (2003)

[47] Son-Cheol Yu, Ura T. , Fujii T. and Kondo H., "Navigation of autonomous underwater vehicles based on artificial underwater landmarks," Proc. MTS/IEEE OCEANS Conf., vol.1, pp 409-416 (2001)

[48] Bellingham J.G. (2016) "Autonomous Underwater Vehicle Docking." In: Dhanak M.R., Xiros N.I. (eds) Springer Handbook of Ocean Engineering. Springer, Cham

[49] Vasilijevic, A., Borovic, B., Vukic, Z., "Underwater vehicle localization with complementary filter: performance analysis in the shallow water environment," J. Intell. Robot. Syst. 68(3-4), 373-386 (2012)

[50] R.D. Light, J. Morison: "The autonomous conductivity-temperature vehicle: First in the sea shuttle family of autonomous underwater vehicle's for scientific payloads," Proc. IEEE OCEANS'89, Vol. 3 (1989) pp. 793-798

[51] Ura, T., Kurimoto, Y., Kondo, H., Nose, Y., Sakamaki, T. and Kuroda, Y., "Observation behavior of an AUV for ship wreck investigation," Proc. MTS/IEEE OCEANS Conf., vol.3, pp.2686-2691 (2005)

[52] F. Chaumette and S. Hutchinson," Visual servo control I: Basic approaches," IEEE Robot. Autom. Mag., vol. 13, no. 4, pp. 82-90, Dec.2006.

[53] F. Chaumette and S. Hutchinson, "Visual servo control II: Advanced approaches," IEEE Robot. Autom. Mag., vol. 14, no. 1, pp. 109-118, Mar.2007.

[54] Farrokh Janabi-Sharifi, Lingfeng Deng, and William J. Wilson, "Comparison of basic visual servoing methods," IEEE/ASME TRANSACTIONS ON MECHATRONICS, VOL. 16, NO. 5, OCTOBER 2011.

[55] Yasutake Okuda, Hiroki Kamada, Satoru Takahashi, Shum'ichi Kaneko, Kuniaki Kawabata, and Fumiaki Takemura, "Method of dynamic image processing for ecology observation of marine life," Journal of Robotics and Mechatronics, vol. 25, no. 5, pp. 820-829, 2013.

[56] Daljae Lee, Xiaodong Tao, Hyungsuck Cho, and Youngjun Cho, "A dual imaging system for flip-chip alignment using visual servoing," Journal of Robotics and Mechatronics, vol. 18, no. 6, pp. 779-786, 2006.

[57] 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," Proc. MTS/IEEE OCEANS Conf., pp.1-8 (2015)

[58] 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 (2016)

[59] 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," Proc. 2015 IEEE/SICE International Symposium on System Integration (SII), pp.989-994 (2015)

[60] 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," Proc. MTS/IEEE OCEANS Conf., pp.1-8 (2016)

[61] Myint, M., Yonemori, K., Yanou, A., Lwin, K.N., Mukada, N. and Minami, M., "Dual-eyes visual-based sea docking for sea bottom battery recharging," Proc. MTS/IEEE OCEANS Conf. Monterey, pp. 1-7(2016)

[62] 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.

[63] Yuya Nishida, Takashi Sonoda, Shinsuke Yasukawa, Jonghyun Ahn, Kazunori Nagano, Kazuo Ishii and Tamaki Ura , "Development of an Autonomous Underwater Vehicle with Human-aware Robot Navigation," Proc. MTS/IEEE OCEANS Conf. Monterey, pp. 1-7(2016)

[64] Stokey, R., Purcell, M. Forrester, N., Austin, T., Goldsborough, R.,Allen, B., von Alt, C., "A Docking System for REMUS, an Autonomous Underwater Vehicle," Proc. MTS/IEEE OCEANS Conf., Vol. 2, pp. 1132-1136, (1997)

[65] Park, J. Y., Jun. B. H., Lee, P. M., Oh, J. H. and Lim Y. K., "Underwater docking approach of an under-actuated AUV in the presence of constant ocean current," IFAC Control Applications in Marine Systems, vol.43, no.20, pp. 5-10 (2010)

[66] 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," OCEANS'2007-Europe, pp 1-6 (2007)

[67] Matthew Dunbabin, Brenton Lang, and Brett Wood, "Vision-based Docking Using an Autonomous Surface Vehicle, IEEE International Conference on Robotics and Automation," Pasadena, CA, USA, pp 26-32 (2008)

[68] Pedro Batista, Carlos Silvestre, and Paulo Oliveira, "A Two-step Control Strategy for Docking of Autonomous Underwater Vehicles," American Control Conference, Fairmont Queen Elizabeth, Montrreal, Canada, pp 5395-5400 (2012)

[69] Palomeras, N., Penalver, A., Massot-Campos, M., Vallicrosa, G., Negre, P.L., Fernandez, J.J., Ridao, P., Sanz, P.J., Oliver-Codina, G. and Palomer, A., "I-AUV Docking and Intervention in a Subsea Panel, IEEE/RSJ International Conference on Intelligent Robots and Systems," Chicago, IL, pp 2279-2285 (2014)

[70] M. A. Rahman and M. N. Uddin. "A novel genetic algorithm based fuzzy logic controller for ipm synchronous motor drive," In Industrial Electronics, IEEE International Symposium on, pp 1007-1010 (2003)

[71] K. N. Tiwari, L. M. Waghmare, and P. Krishnankutty. "Single input fuzzy logic controller tuning for steering control of autonomous underwater vehicle: Genetic algorithm approach," In India Control Conference (ICC), pp 335-340 (2016)

[72] F. Ding and Y. Zhang. "The ga-bp algorithm based controller of pid neural network and its application," In Natural computation (ICNC), Seventh International Conference on, pp 1-4 (2011)

[73] John McCall, "Genetic algorithms for modeling and optimization," Journal of Computational and Applied Mathematics, 184(1), pp 205-222 (2005)

[74] M. N. Uddin, M. A. Abido, and M. A. Rahaman, "Real-time performance evaluation of a genetic-algorithm-based fuzzy logic controller for ipm motor drives," IEEE Transactions on Industry Applications, 41(1), pp 246-252 (2005)

[75] M. E. D. Mandour, E. S. Ali, and M. E. Lotfy, "Robust load frequency controller design via genetic algorithm and h∝," In Modern Electric Power systems (MEPS), Proceedings of the International Symposium, pp 1-6 (2010)

[76] M. E. D. Mandour, E. S. Ali, and M. E. Lotfy, "Robust load frequency controller design via genetic algorithm and h∝," In Modern Electric Power systems (MEPS), Proceedings of the International Symposium, pp 1-6 (2010)

[77] T. Teng, J. Shieh, and C. Chen, "Genetic algorithm applied in online autotuning pid parameters of a liquid-level control system," Transactions of the Institute of Measurement and Control, 25(5), pp 433-450 (2003)

[78] J. Q. Puma and D. G. Colome, "Parameters identification of excitation system models using genetic algorithms," IET Generation, Transmission Distribution, 2(3), pp 456-467 (2008)

[79] Q. Li, S. Xie, X. Tong, and G. Liu, "Path planning algorithm for vehicles based on time-dependent optimization criterion," In Control and Automation, ICCA. IEEE International Conference, pp 2360-2364 (2007)

[80] R. N. Banu and D. Devaraj, "Optimal power flow for steady state security enhancement using genetic algorithm with facts devices," In Industrial and Information Systems, IEEE region 10 and the Third international Conference, pp 1-6 (2008)

[81] N. Metni, " Neuro-control of an inverted pendulum using genetic algorithm," In Advances in Computational Tools for Engineering Applications, International Conference, pp 27-33 (2009)

[82] S. An, T. Lu, and Y. Ma, " Simple adaptive control for siso nonlinear systems using neural network based on genetic algorithm," In Machine Learning and Cybernetics (ICMLC), International Conference on, pp 981-986 (2010)

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

[84] 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.

[85] Roser, M., Dunbabin, M. and Geiger, A., May., "Simultaneous underwater visibility assessment, enhancement and improved stereo," In Robotics and Automation (ICRA), 2014 IEEE International Conference, pp. 3840-3847, 2014.

[86] Rublee, E.; Rabaud, V.; Konolige, K.; Bradski, G., "ORB: An efficient alternative to SIFT or SURF," IEEE International Conference on Computer Vision (ICCV), pp 2564-2571 (2011)

[87] Kasaei, S.H., Oliveira, M., Lim, G.H., Lopes, L.S. and Tom, A.M., "Interactive open-ended learning for 3D object recognition: An approach and experiments," Journal of Intelligent and Robotic Systems, 80(3-4), pp.537-553 (2015)

[88] Wu, P., Liu, Y., Ye, M., Li, J. and Du, S., "Fast and Adaptive 3D Reconstruction with Extensively High Completeness," IEEE Transactions on Multimedia, (2016)

[89] Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S.M. and Szeliski, R., "Building rome in a day," In Communications of the ACM, 54(10), pp.105-112 (2011)

[90] Lwin, K.N., Yonemori, K., Myint, M., Naoki, M., Minami, M., Yanou, A. and Matsuno, T., "Performance Analyses and Optimization of Real-time Multi-step GA for Visual-servoing Based Underwater Vehicle," In Techno-oceans (2016)

[91] Wei Song, Yu Fujia, and Mamoru 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)

[92] Wei. Song, M. Minami, Fujia Yu, Yanan Zhang and Akira Yanou, "3-D Hand and Eye-Vergence Approaching Visual Servoing with Lyapunouv-Stable Pose Tracking," IEEE Int. Conf. on Robotics and Automation (ICRA), pp. 5210-5217 (2011)

[93] Suzuki H. and Minami M., "Visual Servoing to catch fish Using Global/local GA Search," IEEE/ASME Transactions on Mechatronics, vol.10, no.3, pp.352-357 (2005)

[94] 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)

[95] 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)

[96] Wei Song and Minami M., "Stability / precision improvement of 6-DoF visual servoing by motion feedforward compensation and experimental evaluation," IEEE International Conference on Robotics and Automation, pp.722-729 (2009)

[97] Wei Song and Minami M., "Hand and eye-vergence dual visual servoing to enhance observability and stability," IEEE International Conference on Robotics and Automation, pp.714-721 (2009)

[98] 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)

[99] Suzuki, H. and Minami, M., "Real-Time Recognition of Multiple Pedestrians Using Car-Mouted Camera," Electronics and Communications, Part 3, Vol.89, no.4, pp.21-33 (2006)

[100] Trelea, I.C., "The particle swarm optimization algorithm: convergence analysis and parameter selection," Information processing letters, 85(6), pp.317-325 (2003)

[101] Lin, H.I., "A fast and unified method to find a minimum-jerk robot joint trajectory using particle swarm optimization," Journal of Intelligent and Robotic Systems, 75(3-4), p.379 (2014)

[102] Polden, J., Pan, Z., Larkin, N. and van Duin, S., "Adaptive partial shortcuts: Path optimization for industrial robotics," Journal of Intelligent and Robotic Systems, pp.1-13 (2015)

[103] Mousavian, S.H. and Koofigar, H.R., "Identification-Based Robust Motion Control of an AUV: Optimized by Particle Swarm Optimization Algorithm," Journal of Intelligent and Robotic Systems, pp.1-22 (2016)

[104] Ma, X.M., "Application of ant colony algorithm in PID parameter optimization for mining hoist direct torque control system," In Advanced Computer Control, ICACC'09 International Conference, pp. 632-636 (2009)