IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 10, 2021, accepted March 31, 2021, date of publication April 27, 2021, date of current version May 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3076110

Collection of Marine Debris by Jointly Using UAV-UUV With GUI for Simple Operation

NAOKI SHIRAKURA, (Graduate Student Member, IEEE), TAKUYA KIYOKAWA, (Member, IEEE), HIKARU KUMAMOTO, JUN TAKAMATSU, (Member, IEEE), AND TSUKASA OGASAWARA, (Member, IEEE)

Nara Institute of Science and Technology (NAIST), Ikoma 630-0192, Japan

Corresponding author: Naoki Shirakura (shirakura.naoki.se8@is.naist.jp)

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant 20H00611.

ABSTRACT In recent years, floating debris in the sea has damaged the environment and harmed wildlife. Systems for collecting floating marine debris in the open ocean by installing a huge net or using a vacuum device have been proposed. However, to the best of our knowledge, there is no system for collecting marine debris in complex topographies, such as ports wherein marine debris tends to accumulate, and large vessels cannot enter. This study proposes a system to locate and collect floating marine debris by jointly using an unmanned aerial vehicle (UAV) and an unmanned underwater vehicle (UUV). The UAV discovers marine debris that cannot be observed by the UUV owing to its limited view, and the UUV is used for collection. However, for application of such a system, problems related to target detection, UUV navigation, and target collection must be solved. In the proposed system, the following components are developed: a graphical user interface to select the collection target, controller for the UUV to navigate to the target, and a collecting motion generator. We tested and evaluated the proposed cooperative system using a dynamic simulator and physical robots. As a result, we confirm that our system can achieve efficient debris collection in complex topographies, is easy to use, will reduced workload, and can be applied practically.

INDEX TERMS Aerial view, cooperative system, marine debris cleaning, multi-view, teleoperation.

I. INTRODUCTION

Marine debris is a serious environmental problem. In particular, plastic waste floating on the water surface can be accidentally ingested by marine organisms as well as acts as obstructions for ships, thus leading to economic losses for marine industries. The current solution to this problem is manual collection of marine debris using ships or divers. However, since water is a harsh environment for humans, more efficient methods are required to reduce the labor and cost of collection.

An environmental organization Ocean Cleanup proposed a method to collect floating debris by installing a net larger than 100 m [1]. Although this method can efficiently collect floating debris in the open ocean, it is not viable in complex topographies such as ports characterized by marine debris accumulation. To address this problem, floating debris can be collected using an unmanned underwater vehicle (UUV), which is suited to the complex topography of ports. As a UUV can be operated underwater via remote control, it can reduce the labor required for underwater work and be used in various environments, such as docks and on the seabed [2], [3].

However, UUVs have a very limited view, making it difficult to understand and control the situation around them. In sea rescues, aircrafts or unmanned aerial vehicles (UAVs) are often used to locate targets [4]. Searching from the air provides a wider view and makes it possible to locate objects more easily. This paper proposes an efficient system for collecting marine debris by combining the use of a UAV and a UUV. The following capabilities are required to achieve our proposed system: 1) detection of floating marine debris as a collection target, 2) navigation of the UUV to the target using actual scaled position data from the UAV's viewpoint, and 3) motion generation of the UUV for debris collection.

A convolutional neural network (CNN) can automatically detect floating debris on the water surface, but more datasets are required to enable the detection of marine debris with different appearances [5]. Using teleoperation, which skips automatic detection, we can greatly reduce the workload by automating operational elements, except for recognition, such

The associate editor coordinating the review of this manuscript and approving it for publication was Shafiqul Islam.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

as the control of the UUV. Hence, we design a graphical user interface (GUI) controller, where a user selects the target object from the UAV viewpoint, and then the UUV automatically collects the object. Using this GUI, even people with no experience in operating UUVs can collect floating debris easily and efficiently through teleoperation.

However, the relative position between the UUV and collection target must be known for navigating the UUV. For sea rescues, a system for navigating an unmanned surface vehicles (USV) toward victims using homographic translated images has been proposed [4], In our study, we evaluated the provision of an appropriate view to improve usability. Furthermore, we proposed a simple method to measure the three-dimensional position (3D position) of debris using a monocular camera mounted on a general UAV. By assuming that both the debris and the UUV are on the water surface, the normal and position of the water surface can be estimated based on the UUV pose, and the 3D position of the debris can be calculated from these values.

In considering the collection of marine debris, we focus on the motion generation and control of the UUV and use a simple mechanism that can collect floating objects by the motion of the UUV only, such as by scooping. This simplicity enables the installation of various types of UUVs and does not require additional power supply. The motion of the UUV to collect debris after approaching it is generated by teaching playback in advance; the generation does not require user intervention. In our study, we only focus on the debris collection process using the UUV; post debris collection processes will be addressed in future work.

To evaluate the collection efficiency of our system, we experimented with the collection of floating debris using a simulation and compared the time required to collect all debris using our proposed method with those of a method using only the view from the UUV as well as from other related work [4]. We also evaluated the usability and workload of our system using evaluation indices such as the system usability scale (SUS) [6] and the NASA Task Load Index (NASA-TLX) [7]. Furthermore, our system successfully collected floating debris in a pool using real robots; thus, our experiments demonstrate the practicality of the proposed method.

The following contributions were achieved in this study.

- Comparison of collection time and traveling distance demonstrates the high efficacy of the proposed cooperative system.
- SUS indicates high usability of the proposed GUI, which provides the user with the original natural image instead of the image after homographic transformation used in previous research [4].
- NASA-TLX indicates a low workload when using the proposed semi-autonomous system.

II. RELATED WORK

USVs have been used for the collection of floating marine debris on water surfaces using remote operation. FRED [8] and Trashbot [9] are examples of USVs. These USVs are manually controlled by a remote operator using an on-board camera. The main focus of these projects was the collection mechanisms and the use of suction and a conveyor belt, while reducing the workload was not investigated.

UAVs and other robots have been used in combination in various ways. For example, cooperation between a UAV and an unmanned ground vehicle (UGV) has been proposed for rescue and inspection in several studies [10]–[14]. In these studies, UAVs were used to assist the UGV by leveraging their wide field of view and 3D mobility. Since UUV control has six degrees of freedom (DoFs), the difficulty of manual operation is higher than that of a UGV, which has three DoFs. Therefore, in this study, we simplify the operation of the UUV to reduce the workload.

Similarly, UAV-UUV systems have been proposed for rescue operations [4] and underwater inspection [15]. Keila *et al*. proposed a framework for cooperation between a UUV and a UAV, and as an example of an application, an underwater inspection was performed [15]. Since the inspection environment is fixed, the inspection can be performed by moving the UUVs along a pre-programmed path. The UAV is used to collect data measured by UUVs, and thus the views of the UUVs and UAV are not shared. Similar to our work, Xuesu *et al*. proposed a joint UAV-USV system that uses a UAV viewpoint to guide the USV to a rescue target [4]. Their study aimed to reach the rescue target and perform position adjustment after approaching the target; otherwise, the scooping motion would not be performed. Their method provided the view of the UAV after homographic translation. In contrast, our proposed system directly presents the view of the UAV. We evaluated the advantages and disadvantages of the two types of views experimentally.

Detection of floating marine debris using UAVs has been performed for an environmental assessment [16]–[18]. These studies do not aim at collecting debris but collecting statistical data, such as the distribution, the amount, and the classification of debris. Estimation of the position of debris on an image is required to collect debris.

Underwater objects have been detected in many studies using neural networks [19]–[23]. Lili *et al*. achieved a mean average precision of 83.7% for detecting floating object areas on the water surface using Faster R-CNN [5]. Since the existing dataset includes only a single category of debris, marine debris with various appearances, such as plastic bottles, cannot be detected completely. For their future work, they mentioned that additional datasets are required to detect plastic bottles and other debris. Once these methods improve sufficiently with enough datasets to detect more marine debris, our method's manual detection can be replaced, allowing full automation.

A method of estimating the pose of the UUV using its shape has been proposed [24]. Since the method depends on the shape of the hardware, effective implementation is difficult if the UUV shape has few features. As debris collection

is the primary purpose of our research, we have first used augmented reality (AR) markers to recognize the UUV and estimate its pose. If necessary, markerless pose estimation using shape and appearance may be considered in the future.

III. JOINT UAV-UUV DEBRIS COLLECTION

A. OVERVIEW

Fig. 1 shows the overview of the proposed system. In the proposed system, the UAV shown in Fig. 1 (A) is used to search for marine debris. Searching from the air with a UAV provides a 360-degree view to locate debris despite the limited view of the UUV. The UUV then initiates the scooping motion shown in Fig. 1 (C) to collect the debris found with the UAV. The operator teaches the scooping trajectory and the initial condition (e.g., relative position $\boldsymbol{p}_0$ between debris and a UUV) once. The UUV then playbacks the trajectory when the initial condition is satisfied.

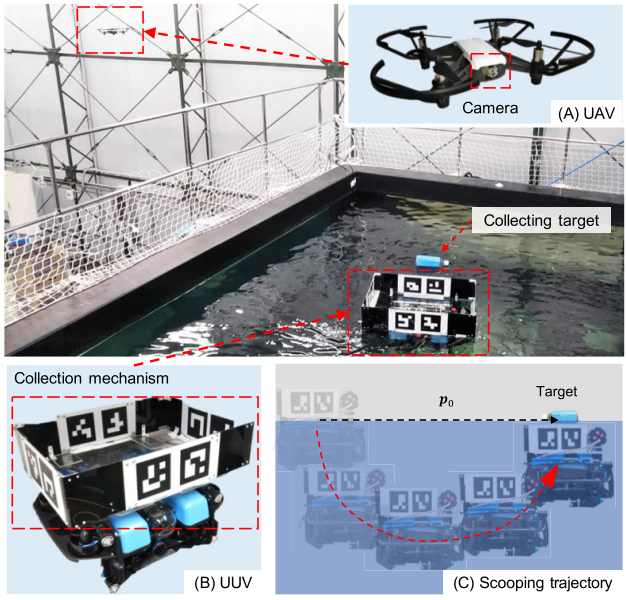


FIGURE 1. Overview of our joint UAV-UUV system for collecting marine debris (A) The UAV with a camera to locate debris. (B) The UUV with a mechanism for debris collection. (C) The scooping trajectory to collect debris.

Fig. 2 shows the process of the proposed system. The system consists of autonomous and manual sequence parts. By using the image from the UAV's camera, the pose of the UUV in the UAV frame can be measured using AR markers. The distance and normal of the water surface are estimated from the pose of the UUV in the UAV frame. The target debris in the camera image is selected manually by an operator. The 3D position of the target can be calculated as the intersection between the view direction and the water surface. The UUV will approach the target based on the calculated 3D position. After the UUV is sufficiently close to the target, the UUV collects the target with the scooping motion shown in Fig. 1 (C).

Fig. 3 shows the GUI used to control the system. A camera image from the UAV is displayed on the left side of the GUI. The operator selects the target by pressing and holding on one piece of debris on the GUI (Fig. 3 (a)). Since debris is not stationary due to the water flow, the position of the target debris can be updated by dragging. The red circle in Fig. 3 (a) is displayed at the relative position $\boldsymbol{p}_0$ between the debris and the UUV recorded at the same time as the scooping motion. When the target debris enters the red circle, the circle turns green to inform the operator that the UUV is sufficiently close to the target for collection (Fig. 3 (b)). Then, by releasing the target on the GUI, the scooping motion is started. The right side of the GUI shows an image of the upward-facing camera on the UUV to observe the target object during scooping motion. The operator can adjust the scooping trajectory of the UUV by pressing the target on the right side of the GUI (Fig. 3 (c)).

B. PROJECTION TO THE WATER SURFACE

Fig. 4 shows the perspective projection from points on the camera image to the 3D position on the water surface. The 3D position of the selected debris in the UAV frame ${}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{t}} = [\ {}^{\mathrm{a}}p_{\mathrm{tx}}\ {}^{\mathrm{a}}p_{\mathrm{ty}}\ {}^{\mathrm{a}}p_{\mathrm{tz}}\ 1\ ]^T$ is constrained on a line as follows:

$$ {}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{t}} = a\boldsymbol{v} = a\boldsymbol{M}^{-1}\ {}^{\mathrm{ac}}\boldsymbol{p}_{\mathrm{t}}\,, \tag{1} $$

where $\boldsymbol{M}$ is the camera matrix of the UAV's camera, ${}^{\mathrm{ac}}\boldsymbol{p}_{\mathrm{t}}$ is the position of the target debris in the UAV camera frame, $\boldsymbol{v}$ is a vector from the UAV camera's center to the image point of the selected debris, and a is any scalar value. Since debris floats on the water surface, the following constraint is given:

$$ \boldsymbol{n} \cdot ({}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{t}} - \boldsymbol{b}) = 0\,, \tag{2} $$

where $\boldsymbol{n}$ is a normal vector of the water surface, and $\boldsymbol{b}$ is any point on the water surface. If the UUV is horizontal on the water surface, $\boldsymbol{n}$ and $\boldsymbol{b}$ can be calculated as follows:

$$ \boldsymbol{b} = {}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{u}}\,, \tag{3} $$

$$ \boldsymbol{n} = {}^{\mathrm{u}}\boldsymbol{T}_{\mathrm{c}}\boldsymbol{p}_{\mathrm{uz}}\,, \tag{4} $$

where ${}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{u}}$ is the position of the UUV in the UAV frame, $\boldsymbol{p}_{\mathrm{uz}} = [\ 0\ \ 0\ \ 1\ ]^T$ is a vertically upward unit vector in the UUV frame, and ${}^{\mathrm{u}}\boldsymbol{T}_{\mathrm{c}}$ is the transformation from the UUV to the UAV camera frame. As mentioned above, ${}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{u}}$ and ${}^{\mathrm{u}}\boldsymbol{T}_{\mathrm{c}}$ are obtained from the pose of the UUV in the UAV frame. We obtained them using AR markers.

Using Eq. (1) and Eq. (2), a is calculated as follows:

$$ a = \frac{\boldsymbol{n} \cdot \boldsymbol{b}}{\boldsymbol{v} \cdot \boldsymbol{n}}\,. \tag{5} $$

The 3D position of the debris in the UUV frame ${}^{\mathrm{u}}\boldsymbol{p}_{\mathrm{t}}$ is calculated, using the transformation from the camera to the UUV ${}^{\mathrm{a}}\boldsymbol{T}_{\mathrm{u}}$ as follows:

$$ {}^{\mathrm{u}}\boldsymbol{p}_{\mathrm{t}} = {}^{\mathrm{a}}\boldsymbol{T}_{\mathrm{u}}\,{}^{\mathrm{a}}\boldsymbol{p}_{\mathrm{t}}\,. \tag{6} $$

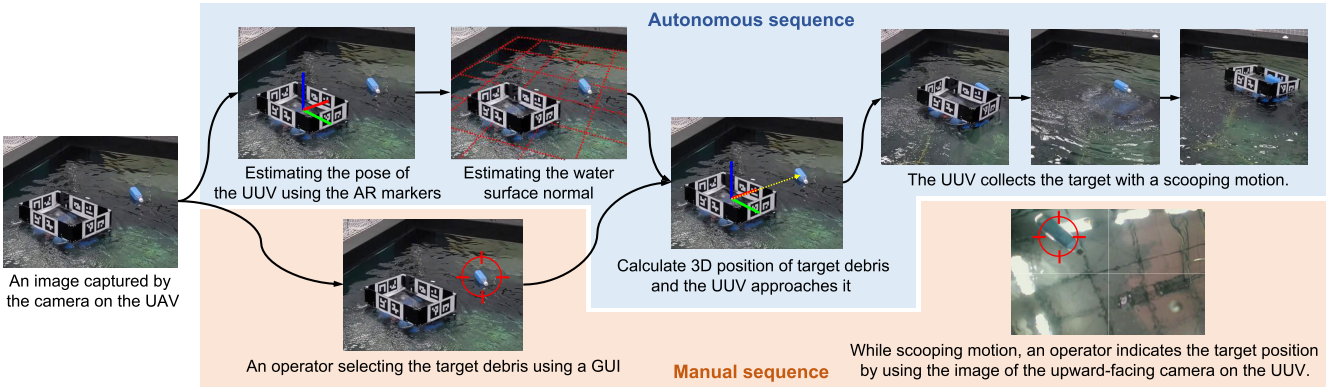
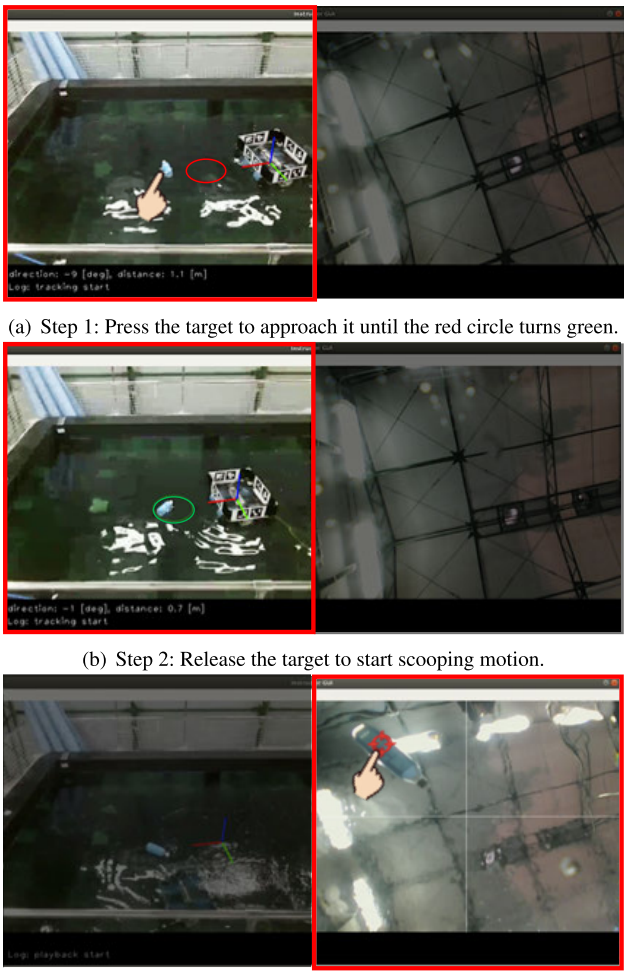


FIGURE 2. Complete process of the proposed system.



(a) Step 1: Press the target to approach it until the red circle turns green.

(b) Step 2: Release the target to start scooping motion.

(c) Step 3: Correct the trajectory by pressing the target on the right side of the GUI screen.

FIGURE 3. Semi-automatic collection procedure of our GUI controller.

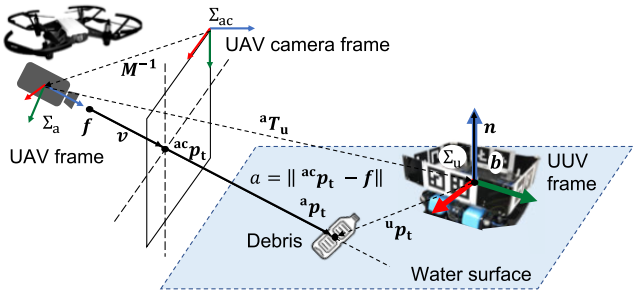


FIGURE 4. Perspective projection from a point on the camera image to a 3D position on the water surface.

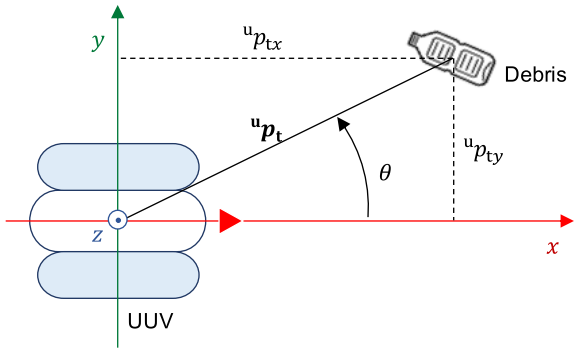


FIGURE 5. Example of the relative position of debris to the UUV.

C. APPROACHING TARGET WITH UUV

To successfully collect debris, the UUV first needs to be sufficiently close to the target debris. Fig. 5 shows a simple situation of the relative position of the UUV and debris. The UUV coordinate system is defined as a right-handed coordinate system in which the upward and forward directions are the $+z$ and $+x$, respectively. In this case, since the movement of the UUV is limited to the water surface, the control of the UUV can be considered as a two- dimensional control (2D control). Therefore, a control input $\boldsymbol{u}$ of the UUV is defined as follows:

$$\boldsymbol{u} = \begin{bmatrix} u_x \\ u_y \\ u_r \end{bmatrix}, \tag{7}$$

where u_x and u_y are the horizontal control inputs, and u_r is the rotational control input. The UUV state $\boldsymbol{x}$ is defined as

the positional difference between the UUV and the target. The state is formulated as follows:

$$\boldsymbol{x} = -\begin{bmatrix} {}^{\mathrm{u}}p_{\mathrm{tx}} \\ {}^{\mathrm{u}}p_{\mathrm{ty}} \\ \tan^{-1}({}^{\mathrm{u}}p_{\mathrm{ty}}/{}^{\mathrm{u}}p_{\mathrm{tx}}) \end{bmatrix}, \tag{8}$$

where ${}^{\mathrm{r}}p_{\mathrm{tx}}$ and ${}^{\mathrm{r}}p_{\mathrm{ty}}$ are the x-y components of the target's position. The deviation is defined using $\boldsymbol{x}$ as follows:

$$\boldsymbol{e} = \hat{\boldsymbol{x}} - \boldsymbol{x}\,, \tag{9}$$

where $\hat{\boldsymbol{x}}$ is the target state of the UUV, that is, the state for starting the scooping motion. When $\boldsymbol{e}$ is less than a threshold value, it indicates that the UUV is sufficiently close to the target debris. To ensure that $\boldsymbol{e}$ is less than the threshold value, the input is calculated using a proportional-integral-differential (PID) controller as follows:

$$\boldsymbol{u} = \boldsymbol{K}_{\mathrm{p}}\boldsymbol{e} + \boldsymbol{K}_{\mathrm{d}}\frac{d}{dt}\boldsymbol{e} + \boldsymbol{K}_{\mathrm{i}}\int \boldsymbol{e}dt\,, \tag{10}$$

where $\boldsymbol{K}_{\mathrm{p}}$, $\boldsymbol{K}_{\mathrm{d}}$, and $\boldsymbol{K}_{\mathrm{i}}$ are the gains of the PID controller and are determined by trial and error using a simulator and a real robot to converge quickly and not diverge.

D. SCOOPING MOTION FOR COLLECTION

Fig. 6 shows the overview of the collecting motion generation. The motion generator is divided into a teaching step that records a scooping motion manually and a playback step that automatically generates the scooping motion based on the recorded motion. In the teaching step, debris was collected manually using the proposed mechanism. At the initial starting point before trajectory teaching, the relative 2D position and orientation angle on the water surface between the UUV and the target debris $\boldsymbol{p}_0$ were recorded. $\boldsymbol{p}_0$ was used to decide the goal position when the UUV approached the target, as follows:

$$\hat{\boldsymbol{x}} = -\boldsymbol{p}_0\,. \tag{11}$$

The control input during the manually controlled trajectory was also recorded as time series data with a time stamp. The control input consists of four channels: up-down, front-rear, left-right, and a continuous value of the vertical rotational angle. The motion of the UUV to collect the target debris was generated by playing back the recorded data according to the time stamp.

E. TEACHING PLAYBACK WITH FEEDBACK

The target debris may move during the scooping motion owing to the waves and water flow created by the UUV. Therefore, an upward-facing camera was installed on top of the UUV. By using the position of the target on the upward-facing camera image, the scooping trajectory was sequentially adjusted, and a robust collection was achieved. To collect the target with the collection mechanism, it is ideal to approach the target from directly below it. Therefore, using

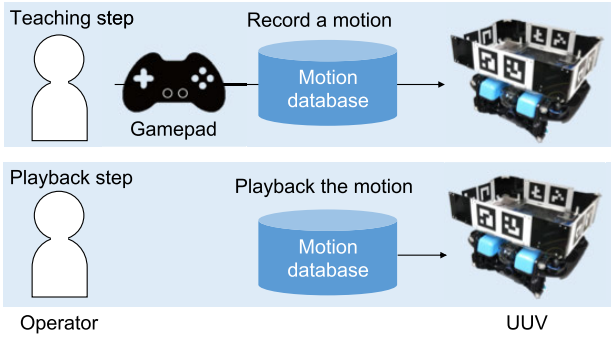


FIGURE 6. Overview of the motion generator using teaching-playback.

the target's position from the UUV's upward-facing camera frame ${}^{\mathrm{uc}}\boldsymbol{p}_{\mathrm{t}}$, the deviation $\boldsymbol{e}_{\mathrm{u}}$ is defined as follows:

$$\boldsymbol{e}_{\mathrm{u}} = \boldsymbol{A}({}^{\mathrm{uc}}\boldsymbol{p}_{\mathrm{t}} - \boldsymbol{c})\,, \tag{12}$$

where $\boldsymbol{A}$ is the transformation matrix from the upward-facing camera image frame to the UUV frame, and $\boldsymbol{c}$ is the center of the camera image. To approach from directly below the target, the scooping trajectory of the UUV was corrected using a PID controller. Hence, using the original recorded trajectory $\boldsymbol{u}(t)$, the corrected trajectory $\tilde{\boldsymbol{u}}(t)$ is calculated as follows:

$$\tilde{\boldsymbol{u}}(t) = \boldsymbol{u}(t) + \boldsymbol{K}_{\mathrm{up}}\boldsymbol{e}_{\mathrm{u}} + \boldsymbol{K}_{\mathrm{ud}}\frac{d}{dt}\boldsymbol{e}_{\mathrm{u}} + \boldsymbol{K}_{\mathrm{ui}}\int \boldsymbol{e}_{\mathrm{u}}dt\,, \tag{13}$$

where $\boldsymbol{K}_{\mathrm{up}}$, $\boldsymbol{K}_{\mathrm{ud}}$, and $\boldsymbol{K}_{\mathrm{ui}}$ are the gains for the trajectory correction.

IV. EXPERIMENT

To evaluate our semi-automatic floating debris collecting system, the following two experiments were conducted: an experiment to evaluate the collecting efficiency in a simulator, and a debris collection test with real robots to evaluate the practicality of our system.

A. IMPLEMENTATION

Fig. 7 shows the hardware configuration of the proposed system. The *Blue Robotics BlueROV2* was used as a UUV. The collecting mechanism is a 42 cm × 57 cm × 19 cm ($X \times Y \times Z$) cage-like mechanism with AR markers and was installed on top of the UUV. The upward-facing camera on the UUV *BUFFALO BSW200MBK* is connected to a *Raspberry Pi 4* and installed on the UUV in a *Blue Robotics Watertight Enclosure (3 Series)*. A micro quadrotor UAV *DJI Tello* was used. The camera was mounted on the front of the UAV at a downward angle of 45 degrees.

The UUV is controlled by Ubuntu 18.04.3 LTS PC *iiyama W350SS* equipped with an Intel(R) Core (TM) i7-4710MQ processor and 7.6 GB-RAM. The UAV is controlled by Ubuntu 18.04.1 LTS PC *VAIO SVS13A2AJ* equipped with an Intel(R) Core (TM) i7-3520M processor and 11.6 GB-RAM. The UAV and its control PC communicate through Wi-Fi. The UUV, its control PC and the *Raspberry Pi 4* for the

upward-facing camera were connected through a *Fathom ROV Tether*, which acts as a network and traction cable. We used *Fathom-X Tether Interface Boards* to convert *Fathom ROV Tether* to an ethernet cable. A gamepad *dualshock 4*, which was used to teach the scooping motion was connected to the UUV control PC. All hardware except *Tello* was connected to the same network through a switching hub.

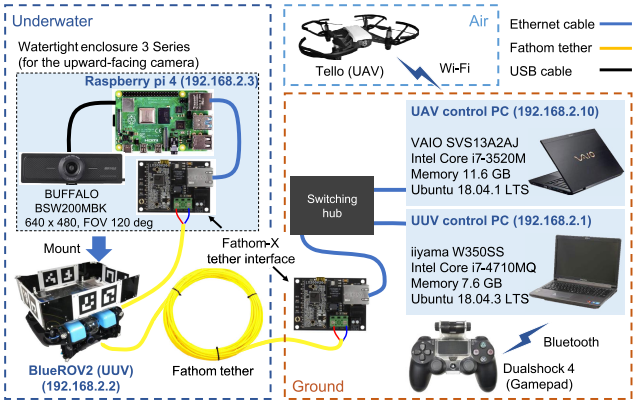


FIGURE 7. Hardware configuration of the proposed system.

Fig. 8 shows the overview of nodes running on PCs and data communication. The proposed system runs on a robot operating system (ROS). The *DJI* official software development kit *Tello-Python* was used to control *Tello* and acquire the camera image. *Mjpeg-streamer* ran on *Raspberry Pi 4* to share the image of the UUV's upward-facing camera through the network. The proposed GUI controller node ran on the UAV control PC and acquired the latest frame from the camera by multi-thread processing. The main thread of the GUI controller node calculated and sent the 3D position of the target to the UUV controller. The UUV controller received the 3D position of the target debris ${}^{r}\boldsymbol{p}_{\mathrm{t}}$, using a motion database for teaching playback. Control commands were sent to the UUV through a wired network using a *MAVLink* protocol. The *MAVLink* interface for ROS, *MAVROS* was used. To display the approaching target circle (red and green circles in Fig. 3), the UUV controller sent the control target position $\hat{\boldsymbol{x}}$ and the current state $\boldsymbol{x}$ to the GUI.

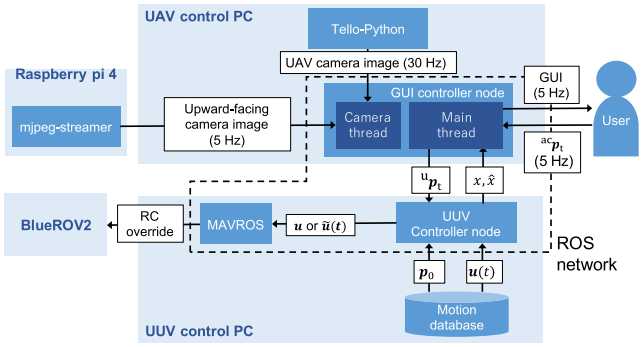


FIGURE 8. Software configuration of the proposed system.

B. EVALUATION OF THE COLLECTING EFFICIENCY, USABILITY, AND WORKLOAD

To evaluate the efficiency, usability, and workload, our system was tested in a simulated environment.

- To evaluate the collection efficiency of the UAV-UUV cooperation, total collection time and travel distance were compared.
- To confirm that the proposed GUI was easy to use, SUS was used as an index to evaluate usability.
- To evaluate labor reduction when using the proposed semi-automation of the collection process, NASA-TLX was used as an index to evaluate the difference in workload.

Fig. 9 shows the environment of the simulation experiment. We used Gazebo as a dynamic simulator. To simulate the underwater environment, a plugin and an environment provided by *Blue Robotics* and *VRX Simulation* were used. In the simulation environment, five plastic bottles with a radius of 6 cm and a height of 26 cm were randomly located in a 10 m $\times$ 10 m area, and the UUV was located at the center of the area. The UAV was placed at an altitude of $+3.5$ m with a -11.0 m offset from the UUV in the y-direction to overlook the entire 10 m $\times$ 10 m area. Buoyancy and vertical sway due to waves were simulated, but the movement of the plastic bottles due to water flow was not simulated. This study assumed that the collection was completed when the debris entered the collection box. Therefore, in the simulation environment, the bottle disappeared when it entered the box.

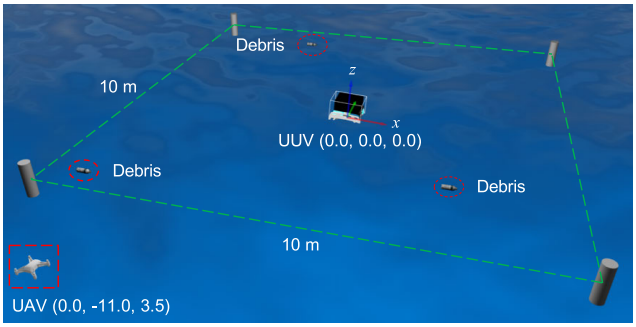


FIGURE 9. Environment of the simulation experiment.

Ten Japanese subjects performed the simulation experiment. For comparison, each subject controlled the UUV under the following four conditions:

1) *UUV Alone:* Subjects used the GUI shown in Fig. 3 where the left side screen was replaced with a forward camera image on the UUV. The operation and control procedures were the same as those shown in Fig. 3.
2) *Gamepad:* Subjects controlled the UUV using the joystick of a gamepad as shown in Fig. 10 to collect debris while looking at the GUI shown in Fig. 3.
3) *Proposed GUI:* Subjects collected the debris with the GUI shown in Fig. 3 (proposed method).

4) *Homographic GUI:* Subjects used the GUI shown in Fig. 3, where the screen on the left side was replaced with the image after homographic transformation.

Moving from condition 1 (*UUV alone*) to condition 2 (*Gamepad*) verifies the effect of cooperation with the UAV, using to *Proposed GUI* verifies the usability of our GUI and the effect of reducing the workload, and *Homographic GUI* corresponds to the previous study [4], which was used for rescue.

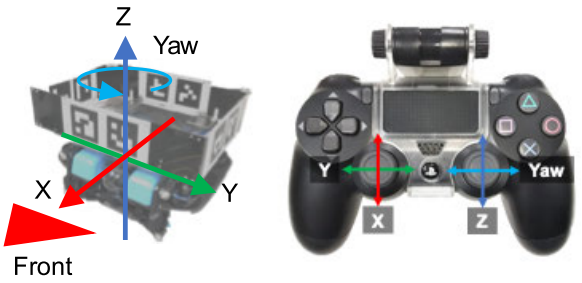


FIGURE 10. Configuration of the manual control.

The flow of the experiment for each subject was as follows: First, we only provided the subjects with the information required to complete the experiment, such as the usage of the controllers and the experimental procedures. Subsequently, the subjects practiced collecting one plastic bottle under each of the four conditions. After practice, the subjects collected five plastic bottles under each of the conditions in order from conditions 1 (*UUV alone*) to 4 (*Homographic GUI*). The time, route, and travel distance from the starting to the finishing point of the operation was recorded. The subjects answered the NASA-TLX and SUS questionnaires after completing the experiment under all conditions so that the contents of the questionnaire did not affect the subjects' intentions. The Japanese version of the NASA-TLX [25] was used as it is the subjects' native language. The original English text and Japanese translations of the SUS questionnaire were used together. In the SUS questionnaire, we also accepted free-form answers for each condition.

To ensure consistency, we asked the subjects to adhere to the following rules:

- Subjects must select the plastic bottle that seems to be the closest as the next target.
- Collection must be done as quickly and accurately as possible.
- If the collection fails, the subjects must try again until they succeed.
- Under condition *UUV alone*, if no plastic bottle was visible, the subjects had to rotate the UUV counterclockwise using the gamepad. As soon as a plastic bottle was found, the subjects controlled the UUV with the GUI.
- In conditions using the GUI, *i.e.*, *UUV alone*, *Proposed GUI*, and *Homographic GUI*, the subjects collected debris using steps 1-3 in Fig. 3.

1) RESULTS OF COLLECTION TIME AND TRAVELING DISTANCE

Table 1 shows the total collection time T and traveling distance D for each subject under each condition. Under condition *UUV alone*, which only considered the UUV viewpoint, T was the largest, but there was almost no difference in the travel distance D among the other conditions. *Proposal GUI* and *Homographic GUI* showed the same tendency, but the average collection time and average total distance traveled were slightly shorter in *Homographic GUI*, as shown in Fig. 11. This may be because the homographic transformation made the positional relationship clearer and made it easier to select the shortest path and adjust the position. In addition, under the manual operation of *Gamepad*, the difference between the subjects was large, and T's variance was the largest among the four conditions. When comparing each condition's average, *UUV alone* required a 1.7-times longer duration than the others, and there was almost no difference between the other conditions.

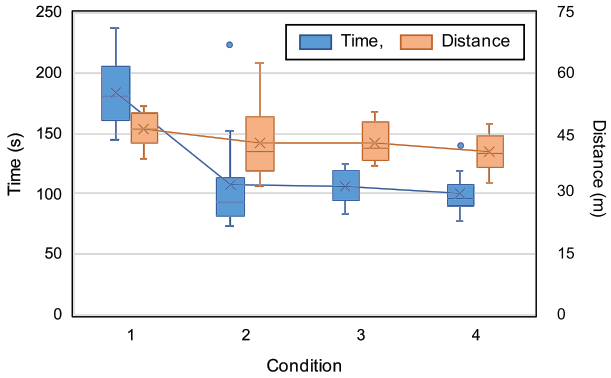


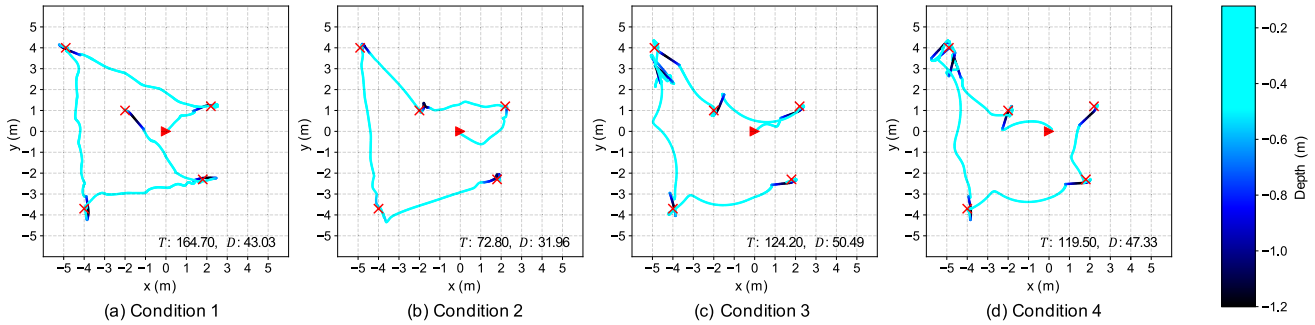
FIGURE 11. Comparison of collection time and traveling distance in the simulation experiment.

Fig. 12 shows the trajectories of the UUV for subjects 5 and 6, who had different tendencies for *Gamepad*, as examples. Since the field of view of *UUV alone* was limited, the order of collection was different from that of the other conditions. Even though there is no difference in D, there is a large difference in T because the UUV had to rotate on the spot after collecting one target until the next target was found. With the other conditions, the UAV provided a view of all the plastic bottles at all times; thus, the UUV could immediately move on to the next target after collecting the previous one.

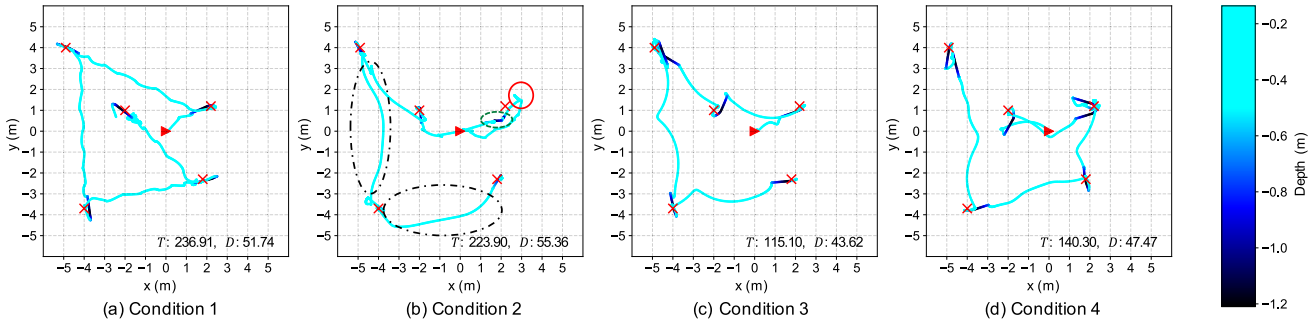
For *Gamepad*, the order of collecting targets and D were similar to the other conditions, but T of subject 6 was approximately three times larger than subject 5. For *Gamepad* in Fig. 12 (b), the solid red circle indicates that the UUV had passed the target once before it was collected. The green dotted circle indicates that the collection failed, and the scooping motion was repeated. The black dash-dotted circle indicates that the UUV could not head straight toward the next target. These facts indicated that the operation

TABLE 1. Results of the user experiment.

Subjects	Conditions							
	1. UUV alone		2. Gamepad		3. Proposed GUI		4. Homographic GUI	
	T/[s]	D/[m]	T/[s]	D/[m]	T/[s]	D/[m]	T/[s]	D/[m]
1	207.10	49.50	99.99	34.11	99.20	38.69	103.20	43.11
2	176.79	46.69	93.80	42.05	92.60	37.10	77.80	32.73
3	204.50	51.24	152.69	62.32	117.30	48.00	88.70	38.22
4	181.20	47.30	92.70	42.49	123.20	47.62	91.10	38.07
5	164.70	43.03	72.80	31.96	124.20	50.49	119.50	47.33
6	236.91	51.74	223.90	55.36	115.10	43.62	140.30	47.47
7	188.00	45.88	82.50	36.19	82.90	37.21	100.60	42.51
8	150.60	41.70	87.41	39.24	103.59	41.53	93.20	36.72
9	145.48	38.82	97.39	46.76	94.60	40.43	99.20	41.90
10	181.38	45.51	77.80	37.20	107.40	41.24	90.30	36.66
Average	183.67	46.14	108.10	42.77	106.01	42.59	100.39	40.47
SD	26.09	3.92	43.89	9.14	13.17	4.47	16.89	4.57



(a) Subject 5



(b) Subject 6

FIGURE 12. Trajectories of the UUV's during the simulation experiments. The five red crosses in each graph indicate the location of the five pieces of debris. The red triangle indicates the initial pose of the UUV. The line color indicates the depth of the UUV. The results of subjects 5 and 6 are shown as examples due to the different tendencies shown during condition *Gamepad*.

proficiency level using the gamepad affected the results in *Gamepad*.

As a result of the comparison of T and D between different conditions, it was confirmed that the time required to collect multiple floating debris could be shortened, and the collection efficiency could be improved by cooperation with the UAV. In addition, in the case of manual operation using a gamepad, it was confirmed that the collection time varied greatly among individuals, and some subjects could collect debris more efficiently.

2) RESULTS OF THE NASA-TLX QUESTIONNAIRE

Fig. 13 shows the results of the NASA-TLX questionnaire. The TLX score expresses the workload as a value from 0 to 100; larger values indicate a higher workload. As a result, the proposed method shows a significant difference ($p < 0.05$) from *Gamepad*. From the comparison of *Gamepad* and *Proposed GUI* in Fig. 13, it is evident that the workload of *Proposed GUI* was lower than that of *Gamepad*. Furthermore, even subjects 5 and 10, who achieved better collection times under condition *Gamepad* than *Proposed*

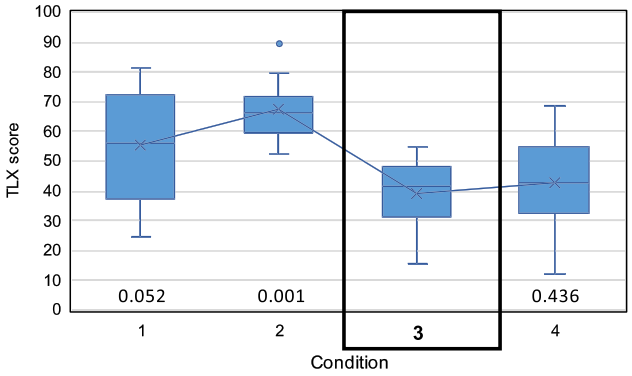


FIGURE 13. TLX score of each condition. Outliers were eliminated by constraining the minimum and maximum values within 1.5 times the box length. The number in the graph shows the *p*-value of the *t*-test between each condition and the *Proposed GUI*. The significance level was 0.05.

GUI, confirmed that the workload was reduced in *Proposed GUI*. *UUV alone* showed a slight improvement in the workload, even though the value was close to the significance level ($p = 0.052$). Therefore, using the viewpoint from the UAV may reduce the workload. There was no significant difference when compared to *Homographic GUI*, but the mean and variance were small in the proposed method. In addition, the TLX score for each element in Fig. 14 indicated that *Proposed GUI* had a lower mental demand (MD) than *Homographic GUI*. Since MD represents how much mental and perceptual thinking is required for the task, the difference in MD shows that our GUI is intuitive and does not require extra thinking to use.

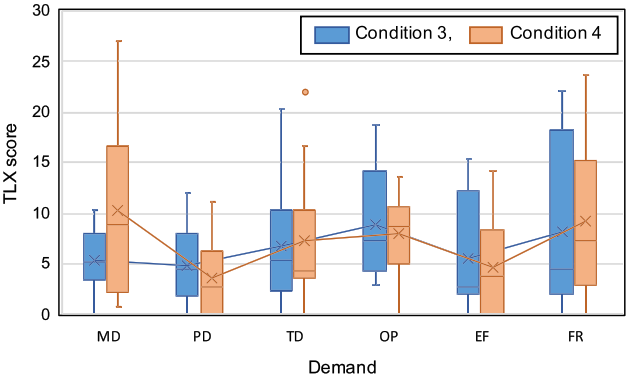


FIGURE 14. TLX score comparison of *Proposed GUI* and *Homographic GUI* for each element.

3) RESULTS OF THE SUS QUESTIONNAIRE

Fig. 15 shows the results of the SUS questionnaire. The SUS score expresses usability as a value from 0 to 40, where a large value indicates high usability. The results confirmed that *Proposed GUI* was significantly different from the other conditions. Fig. 15 shows that the proposed method had the highest SUS score. The cause of the difference can be considered as follows from the contents of the free-form answers in the SUS questionnaire:

- *UUV Alone:* The field of view is narrow and hence, locating the target is time consuming. It is not easy to track the target because the viewpoint changes with the motion of the UUV.
- *Gamepad:* If the user is familiar with the operation, debris can be collected efficiently, but it is not easy at first. The operation becomes confusing when the viewpoint changes.
- *Homographic GUI:* The positional relationship of plastic bottles is easy to grasp but may seem unnatural. Therefore, the original image is easier to use.

UUV alone and *Gamepad* showed significant differences in both efficiency and load of the proposed method, which may also have affected usability. Regarding *Homographic GUI*, there was no significant difference in efficiency or load, but a significant difference was confirmed in the case of usability. This may be because of the unnaturalness of the deformed image. In the proposed GUI, since the original image is given to the user as a viewpoint, the usability does not deteriorate.

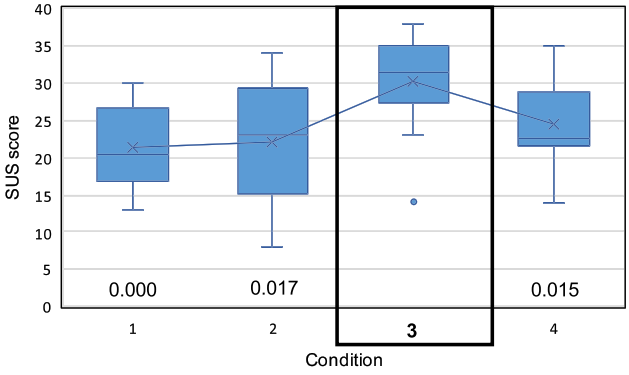


FIGURE 15. SUS score of each condition. Outliers were eliminated by constraining the minimum and maximum values within 1.5 times the box length. The number in the graph shows the *p*-value of the *t*-test between each condition and the *Proposed GUI*. The significance level was 0.05.

C. DEBRIS COLLECTING TEST WITH REAL ROBOTS

To evaluate the practical application of our system, we attempted debris collection using real vehicles. A pool with a length of 2.5 m, width of 4.0 m, and depth of 1.5 m was used as an experimental environment. A plastic bottle with a radius of 6 cm and a height of 26 cm was placed in the pool. We carried out 17 separate trials to collect the plastic bottle using the proposed system. In each trial, targeting using the GUI, approaching the target, and generating the scooping motion were executed, and the number of successful collections was recorded. The initial positions of the UUV and plastic bottle were not too close, and the plastic bottle was located away from the wall of the pool to prevent the UUV from colliding with the wall. Regardless of the collection result, the position of the UUV and plastic bottle was reset for each trial.

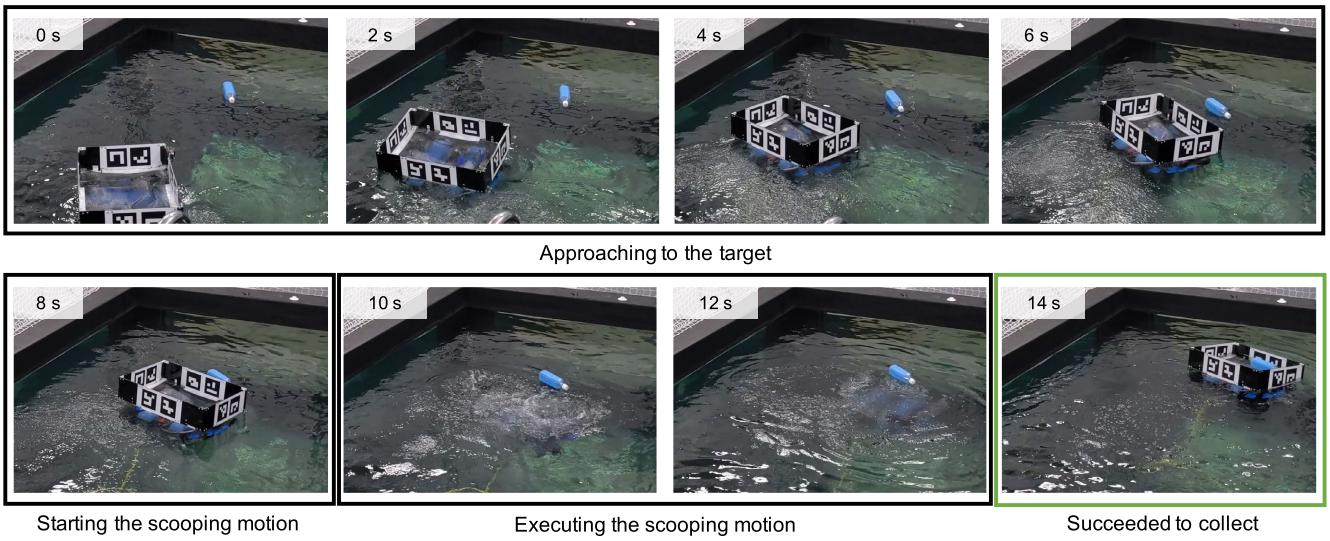


FIGURE 16. Motion of the UUV in a success case of the experiment using a real robot.

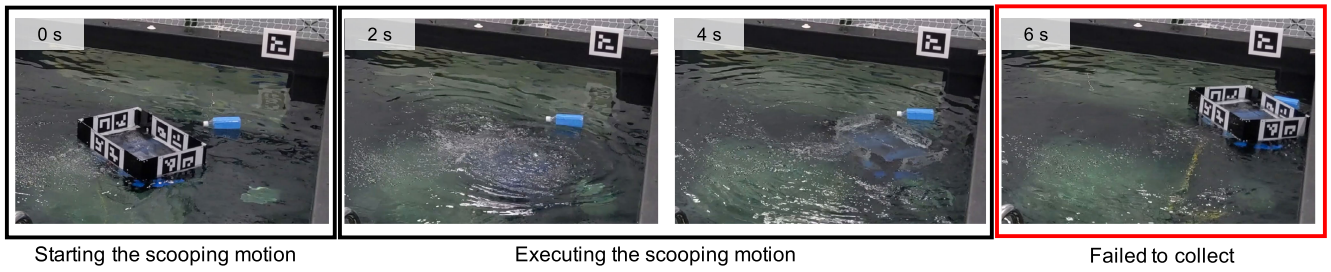


FIGURE 17. Motion of the UUV in a failure case of the experiment using a real robot.

Table 2 shows the success rate of the debris collection experiment using real robots. The results show that the collection of one piece of debris was successful in 15 of 17 trials.

TABLE 2. Results of the experiment using real robots.

Success	Failure	Total	Success rate
15	2	17	$15/17 \approx 88.23\%$

Fig. 16 and Fig. 17 show the motions of the UUV when the collection succeeded and failed, respectively. In the case of a successful collection, the UUV approached the target (within 0 to 6 s), UUV was sufficiently close to the target, the operator triggered the execution of the scooping motion (at 8 s), completed the scooping motion, and succeeded to collect the target (at 14 s). In the case of a failed collection, the UUV did not come sufficiently close to the target, and the target could not be tracked using the upward-facing camera on the UUV. Therefore, if the target can be captured with an upward-facing camera, the collection will almost certainly succeed.

V. CONCLUSION

In this study, we proposed a semi-automatic collection system for floating marine debris by jointly using a UAV and a UUV. We designed a coordinate transformation method and a GUI to estimate the 3D position of floating debris on the water surface using images from a monocular camera mounted on a UAV. To simplify the control of UUV teleoperation, we developed a controller that generates a robust collection trajectory using teaching playback along with feedback.

To evaluate the efficacy of our system, we conducted both simulation and real-robot experiments. In the simulation experiments, 10 subjects collected multiple pieces of floating debris under different conditions. As a result, it was confirmed that UAV-UUV cooperation enables more efficient collection than UUV alone. Furthermore, from the results of the NASA-TLX and SUS questionnaires, it was confirmed that our GUI reduced the workload and realized the ease of use. In addition, from the comparison between the homographic-converted image [4] and the original image, it was confirmed that the usability improved when the original image is provided to the user. By contrast, it was

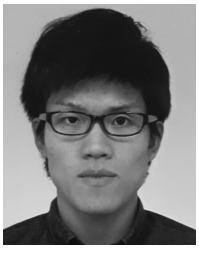
indicated that the homographic-converted image may have assisted subjects in fully comprehending the situation, such as the positional relationship between debris and UUV, and slightly improved the collection efficiency. We also tested the proposed system using real robots in an experimental pool, and we confirmed the feasibility of a semi-automatic collection of floating debris with a successful collection rate of 88.23%.

In future work, we plan to improve the successful collection rate by modifying the feedback controller. As our system currently uses simple PID control, when the deviation is small, the actuator's dead zone can make convergence to the target difficult. This can lead to possible collection failure or reduced usability; hence, a method that considers handling of the actuator's dead zone is required [26]. Furthermore, we will attempt practical collection of marine debris in an actual marine environment instead of the experimental pool using our system. Currently, we are targeting situations that are remotely controlled from a ship, where the onboard computer does not require much computing power. When our system is fully automated by implementing automatic object detection, such as that proposed by Lili [5], all processing can be made executable onboard and wirelessly by installing a computer suitable for machine learning, such as an NVIDIA Jetson.

REFERENCES

[1] *The Ocean Cleanup*. Accessed: Sep. 18, 2019. [Online]. Available: https://theoceancleanup.com/
[2] Y. R. Petillot, G. Antonelli, G. Casalino, and F. Ferreira, "Underwater robots: From remotely operated vehicles to intervention-autonomous underwater vehicles," *IEEE Robot. Autom. Mag.*, vol. 26, no. 2, pp. 94–101, Jun. 2019.
[3] S. Kawamura, "Underwater robot development for manipulation task and their uses in Biwa Lake," *IFAC-PapersOnLine*, vol. 48, no. 2, pp. 14–19, 2015.
[4] X. Xiao, J. Dufek, T. Woodbury, and R. Murphy, "UAV assisted USV visual navigation for marine mass casualty incident response," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 6105–6110.
[5] L. Zhang, Y. Zhang, Z. Zhang, J. Shen, and H. Wang, "Real-time water surface object detection based on improved faster R-CNN," *Sensors*, vol. 19, no. 16, p. 3523, Aug. 2019.
[6] J. Brooke, "SUS: A 'quick and dirty'usability," in *Usability Evaluation in Industry*. New York, NY, USA: Taylor & Francis, 1996, p. 189.
[7] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Adv. Psychol.*, vol. 52, pp. 139–183, Apr. 1988.
[8] *Clear Blue Sea, FRED (Floating Robot for Eliminating Debris*. Accessed: Aug. 15, 2020. [Online]. Available: https://www.clearbluesea.org/meet-fred/
[9] *Urban Rivers, Trashbot*. Accessed: Aug. 15, 2020. [Online]. Available: https://www.urbanriv.org/trashbot
[10] S. Hood, K. Benson, P. Hamod, D. Madison, J. M. O'Kane, and I. Rekleitis, "Bird's eye view: Cooperative exploration by UGV and UAV," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2017, pp. 247–255.
[11] J. Li, G. Deng, C. Luo, Q. Lin, Q. Yan, and Z. Ming, "A hybrid path planning method in unmanned air/ground vehicle (UAV/UGV) cooperative systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9585–9596, Dec. 2016.
[12] G. Christie, A. Shoemaker, K. Kochersberger, P. Tokekar, L. McLean, and A. Leonessa, "Radiation search operations using scene understanding with autonomous UAV and UGV," *J. Field Robot.*, vol. 34, no. 8, pp. 1450–1468, Dec. 2017.
[13] T. Miki, P. Khrapchenkov, and K. Hori, "UAV/UGV autonomous cooperation: UAV assists UGV to climb a cliff by attaching a tether," in *Proc. Int. Conf. Robot. Automat. (ICRA)*, May 2019, pp. 8041–8047.
[14] K. A. Ghamry, M. A. Kamel, and Y. Zhang, "Cooperative forest monitoring and fire detection using a team of UAVs-UGVs," in *Proc. Int. Conf. Unmanned Aircr. Syst. (ICUAS)*, Jun. 2016, pp. 1206–1211.
[15] K. Lima, E. R. B. Marques, J. Pinto, and J. B. Sousa, "Dolphin: A task orchestration language for autonomous vehicle networks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2018, pp. 603–610.
[16] O. Garcia-Garin, T. Monleón-Getino, P. López-Brosa, A. Borrell, A. Aguilar, R. Borja-Robalino, L. Cardona, and M. Vighi, "Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a Web application in r," *Environ. Pollut.*, vol. 273, Mar. 2021, Art. no. 116490.
[17] K. Kylili, I. Kyriakides, A. Artusi, and C. Hadjistassou, "Identifying floating plastic marine debris using a deep learning approach," *Environ. Sci. Pollut. Res.*, vol. 26, no. 17, pp. 17091–17099, Jun. 2019.
[18] M. Geraeds, T. van Emmerik, R. de Vries, and M. S. B. A. Razak, "Riverine plastic litter monitoring using unmanned aerial vehicles (UAVs)," *Remote Sens.*, vol. 11, no. 17, p. 2045, Aug. 2019.
[19] D. Miller, F. Dayoub, M. Milford, and N. Sunderhauf, "Evaluating merging strategies for sampling-based uncertainty techniques in object detection," in *Proc. Int. Conf. Robot. Automat. (ICRA)*, May 2019, pp. 2348–2354.
[20] M. J. Islam, M. Fulton, and J. Sattar, "Toward a generic diver-following algorithm: Balancing robustness and efficiency in deep visual detection," *IEEE Robot. Autom. Lett.*, vol. 4, no. 1, pp. 113–120, Jan. 2019.
[21] M. Fulton, J. Hong, M. J. Islam, and J. Sattar, "Robotic detection of marine litter using deep visual detection models," in *Proc. Int. Conf. Robot. Automat. (ICRA)*, May 2019, pp. 5752–5758.
[22] M. J. Islam, M. Ho, and J. Sattar, "Dynamic reconfiguration of mission parameters in underwater human-robot collaboration," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 6212–6219.
[23] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2018, pp. 7159–7165.
[24] J. Dufek and R. Murphy, "Visual pose estimation of rescue unmanned surface vehicle from unmanned aerial system," *Frontiers Robot. AI*, vol. 6, p. 42, May 2019.
[25] S. Haga and N. Mizukami, "Japanese version of NASA task load index," *Jpn. J. Ergonom.*, vol. 32, no. 2, pp. 71–79, 1996.
[26] T. Yang, N. Sun, and Y. Fang, "Adaptive fuzzy control for a class of MIMO underactuated systems with plant uncertainties and actuator deadzones: Design and experiments," *IEEE Trans. Cybern.*, early access, Feb. 2, 2021, doi: 10.1109/TCYB.2021.3050475.

NAOKI SHIRAKURA (Graduate Student Member, IEEE) received the M.E. degree from the Nara Institute of Science and Technology, Japan, in 2018, where he is currently pursuing the Ph.D. degree with the Robotics Laboratory, Division of Information Science. His research interests include navigation of unmanned aerial vehicles and unmanned underwater vehicles.



TAKUYA KIYOKAWA (Member, IEEE) received the M.E. degree from the Nara Institute of Science and Technology, Japan, in 2018, where he is currently pursuing the Ph.D. degree with the Robotics Laboratory, Division of Information Science. His current research interests include robot manipulation and robot vision for agile reconfigurable robotic systems toward agile manufacturing.

HIKARU KUMAMOTO received the B.E. degree from the Nara College, National Institute of Technology, Japan, in 2019. He is currently pursuing the master's degree with the Robotics Laboratory, Division of Information Science, Nara Institute of Science and Technology, Nara, Japan. He joined the Nara Institute of Science and Technology, in 2019. His research interests include robot manipulation and mobile robot.

JUN TAKAMATSU (Member, IEEE) received the Ph.D. degree in computer science from The University of Tokyo, Japan, in 2004. From 2004 to 2008, he was with the Institute of Industrial Science, The University of Tokyo. He was a Visiting Researcher with Microsoft Research Asia, in 2007. He joined the Nara Institute of Science and Technology, Japan, as an Associate Professor, in 2008. His research interests include robotics, including learning-from observation, task and motion planning, feasible motion analysis, 3D shape modeling and analysis, and physics-based vision.

TSUKASA OGASAWARA (Member, IEEE) received the B.S. and Ph.D. degrees from The University of Tokyo, Japan, in 1983. From 1983 to 1998, he was with the Electrotechnical Laboratory, Ministry of International Trade and Industry, Japan. From 1993 to 1994, he was a Humboldt Research Fellow with the Institute for Real-Time Computer Systems and Robotics, University of Karlsruhe, Germany. He joined the Nara Institute of Science and Technology, Nara, Japan, in 1998, where he is currently a Professor with the Division of Information Science. He is also the Vice President of the Nara Institute of Science and Technology and the Dean of the Graduate School of Science and Technology. His research interests include human–robot interaction, dexterous manipulation, human modeling, and bio-inspired robotics.