# Gesture Control and Situation Awareness in Underwater Human-Robot Teams

Antonella Wilby<sup>1</sup>

Abstract-Humans working with robots have been studied extensively, but the problem of human-robot collaboration in underwater settings has not been given much attention. Humans working underwater are exposed to a variety of risks and stressors including complex tasks, inefficiencies arising from time limitations, and task overloading. Robots can help improve the safety and efficiency of underwater operations by performing hazardous tasks in place of a human, assisting humans with complex tasks, or performing tasks in parallel with human team members to increase efficiency. However, introducing an autonomous robot into a team of humans can introduce new types of risk arising from robot performance, human-robot communication, and other factors. In this abstract, we present a case study in human-robot teaming consisting of two SCUBA divers and a robot and discuss observed issues in the scenario arising from the robot's communication abilities and situation awareness. We also discuss initial work on a gesture-based communication system for an underwater robot, as well as an omnidirectional camera system for diver tracking. This initial work outlines operational considerations for humanrobot teaming in dynamic underwater environments and builds a framework for future field experiments in this area.

### I. INTRODUCTION

In many scientific domains, humans are tasked with performing underwater work such as equipment installation, surveys, inspection, or sample collection. Underwater work has inherent risks and inefficiencies due to limited air supply and bottom times, and frequently the particular tasks at hand introduce additional risk, for example when collecting samples of hazardous biological specimens. Collaborative robots working as part of underwater operational teams have the potential to increase the safety and efficiency of underwater work, whether by collaboratively performing hazardous parts of tasks, or by offloading menial tasks to the robot to increase efficiency. However, introducing a robot into a team of humans adds a layer of complexity and thus new types of risk, arising primarily from the robot's task performance, limited communication abilities and potential for poor situation awareness. Task performance is of greatest interest, but before a robot's task performance can be safely studied as a fully-autonomous member of an underwater human-robot team, the robot must first have sufficient communication abilities and appropriate sensors for situation awareness.

The underwater environment is communication-restricted, meaning that many of the typical modes of communication employed either in human communication (*e.g.* speaking)



Fig. 1. Underwater human-robot teaming scenario.

or robot communication (*e.g.* wi-fi) are unavailable as a communication mode. Additionally, the reduced field of view created by the air-water interface of optical imaging systems results in worse situation awareness than on land. In SCUBA diving, situation awareness is a key skill for diver safety, and many of the issues that affect situation awareness in divers, such as attention tunneling and data overload [2] also affect underwater robots. Robust communication techniques and ensuring robots have good situation awareness are key to safely introducing robots to an underwater team to facilitate further research in robot task performance.

We present a case study of a human-robot team consisting of two SCUBA divers and a robot working in a coastal environment to install an array of cameras. From observations made in this case study, we discuss key factors related to both situation awareness and communication, and present initial work addressing these factors with an omnidirectional camera for situation awareness and a gesture-based communication system.

## A. Related Work

Due to the complexity of working in ocean environments, there are few examples in the literature of true underwater human-robot collaboration. One example is the OceanOne robot [4] which is a hybrid between a humanoid robot and a Remotely Operated Vehicle (ROV), which can communicate with divers underwater by making gestures with its limbs and hands, and is controlled by a topside human operator. Another example is a diver-controlled soft-robotic fish [3], where a diver operates a wireless acoustic controller to drive a robotic fish underwater.

## II. UNDERWATER HUMAN-ROBOT TEAMING

We conducted a case study of underwater human-robot teaming in the ocean off the Scripps Institution of Oceanog-

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1650112.

<sup>&</sup>lt;sup>1</sup>The author is with the Computer Science and Engineering Dept., University of California, San Diego, La Jolla, CA 92093, USA awilby @eng.ucsd.edu



Fig. 2. Omnidirectional camera for increased situation awareness in humanrobot teaming scenarios. In this scenario, Openpose [1] is run on the uncalibrated omnidirectional image to track a freediver below the robot.

raphy pier. In this scenario, two SCUBA divers were installing an array of cameras on the seabed while a teleoperated underwater robot followed the divers in the environment (Fig. 1). The videos from the camera array were later used to observe relative positions between the humans and the robot. The robot had two downward-facing cameras (typically used for visual SLAM) arranged in a stereo pair with a 17.75 in baseline, and a forward-facing camera on a tilt servo, with all cameras imaging through dome ports to minimize field of view (FOV) loss due to refraction. The stereo cameras yielded an approximately 100° horizontal by 86° vertical downward-facing FOV. The forward-facing camera had a horizontal FOV of 110° and a static vertical FOV of 83°, and was mounted on a tilt servo with a vertical rotation of approximately  $-45^{\circ}$  to  $+45^{\circ}$ . This setup yields less than 30% coverage of the entire sphere (from the combined solid angle of the two fields of view  $\Omega < 4 sr$ ) at any given time.

#### A. Situation Awareness

In a highly dynamic environment like the ocean, where humans and robots can move unconstrained in three dimensions as well as are subject to uncontrolled movement from environmental dynamics (*e.g.* swell), the limited field of view from existing cameras (such as stereo cameras used for visual SLAM) is insufficient for situation awareness, and thus operational safety. A limited FOV is analogous to attention tunneling, where a diver becomes hyperfocused on something to the neglect of all other important stimuli. Empirically, this results in the robot having limited to no awareness of its human team members unless it is directly focused on tracking them.

While the FOV constraint could be approached with a variety of search strategies, we posit that it is more effective and reliable to use a dedicated sensor for situation awareness. This will ensure the robot has sufficient information for global awareness at all times. We outfitted the robot with a Kodak SP360 omnidirectional camera in an underwater housing, with an underwater FOV of approximately 180°.



Fig. 3. Four hand gestures detected using OpenPose. From left: "ok", "ascend", "descend", "stop".

TABLE I GESTURE CLASSIFICATION ACCURACY

Class	"ok"	"ascend"	"descend"	"stop"
Classification Accuracy	97%	93%	84%	97%

This increases the sphere coverage to over 50% ( $\Omega \approx 6.4 sr$ ). We tested this camera by tracking a freediver using Openpose [1] on the uncalibrated omnidirectional image (Fig. 2). While this approach has not yet been rigorously field tested, we plan to use this setup in future human-robot field experiments.

# B. Human-Robot Communication

In addition to good situation awareness, a second aspect of safe underwater human-robot field experiments is an effective human-robot communication system. Since humans cannot speak underwater and robot communication is difficult, we developed a gesture-based communication system using the commonly understood language of SCUBA signals to enable human-robot communication. We used OpenPose to detect hands in the image, which are represented with 21 keypoints, then trained an SVM classifier with a polynomial kernel using around 1000 samples per class (Table I). The classification is very fast, but OpenPose was unable to detect hands in downsampled video. The classification is run on video from the robot's forward-facing camera at 1080p resolution, which runs at 8-10fps on an NVIDIA 1070 GPU. Currently, this gesture control system only understands four commands-"ok", "ascend", "descend", and "stop" (Fig. 3)-but we plan to extend the system to understand more sophisticated gestures and gestural instructions.

## **III. DISCUSSION AND FUTURE WORK**

We discussed a case study in underwater human-robot teaming and resulting observations for sensors and communication methods to enable safe human-robot experimentation. We plan to use our framework for situation awareness and human-robot communication for future work in task-based underwater human-robot collaboration.

## REFERENCES

- Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In CVPR, 2017.
- [2] M. R. Endsley. *Designing for situation awareness: An approach to user-centered design.* CRC press, 2016.
- [3] R. K. Katzschmann, J. DelPreto, R. MacCurdy, and D. Rus. Exploration of underwater life with an acoustically controlled soft robotic fish. *Science Robotics*, 3(16):eaar3449, 2018.
- [4] O. Khatib, X. Yeh, G. Brantner, B. Soe, B. Kim, S. Ganguly, H. Stuart, S. Wang, M. Cutkosky, A. Edsinger, et al. Ocean one: A robotic avatar for oceanic discovery. *IEEE Robotics & Automation Magazine*, 23(4):20–29, 2016.