ICRA 2022
IEEE International Conference on Robotics and Automation

MINNESOTA ROBOTICS INSTITUTE
UNIVERSITY OF MINNESOTA

Using Monocular Vision and Human Body Priors for Autonomous Diver Approach

Michael Fulton, Jungseok Hong, Junaed Sattar
{fulto081, jungseok, junaed}@umn.edu

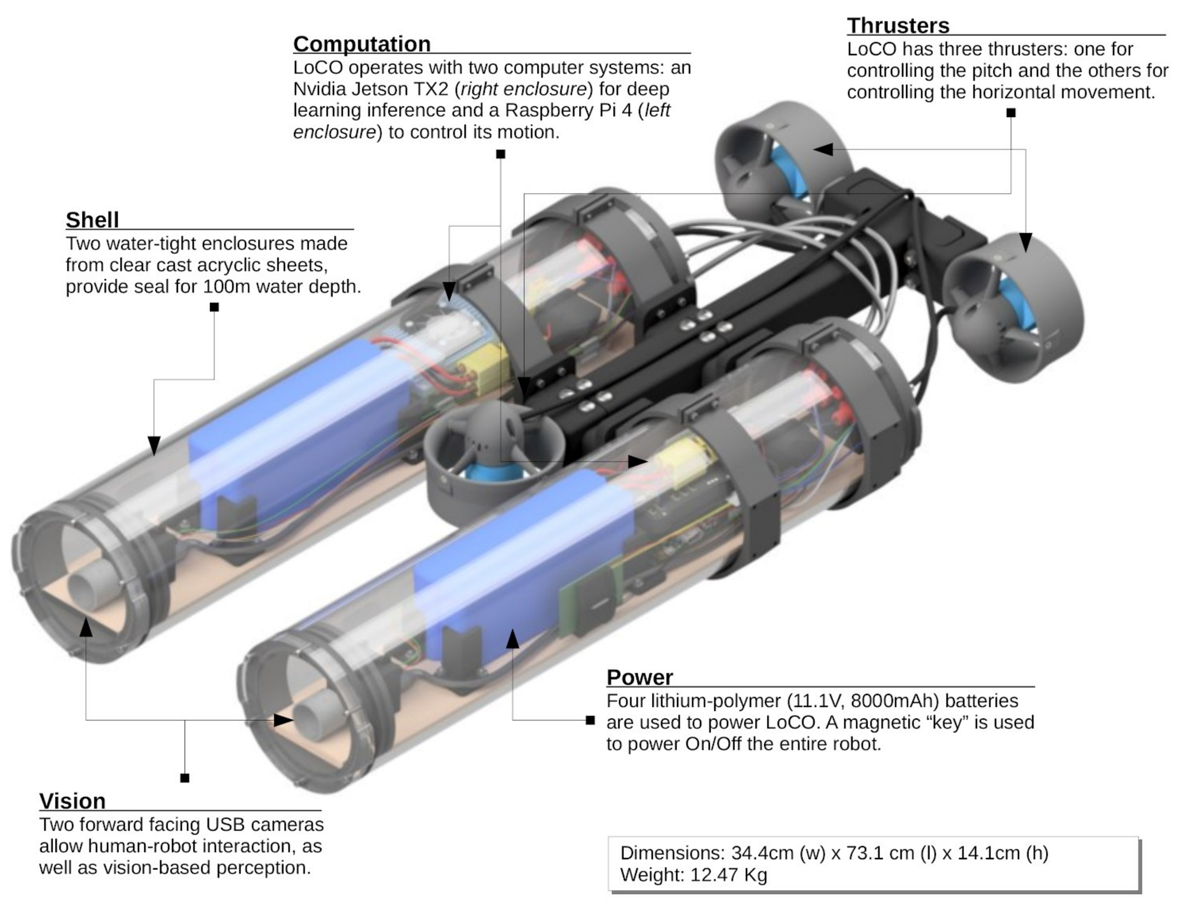
SCAN ME
irvlab.cs.umn.edu

Can Monocular Vision Be Used for Diver Approach?

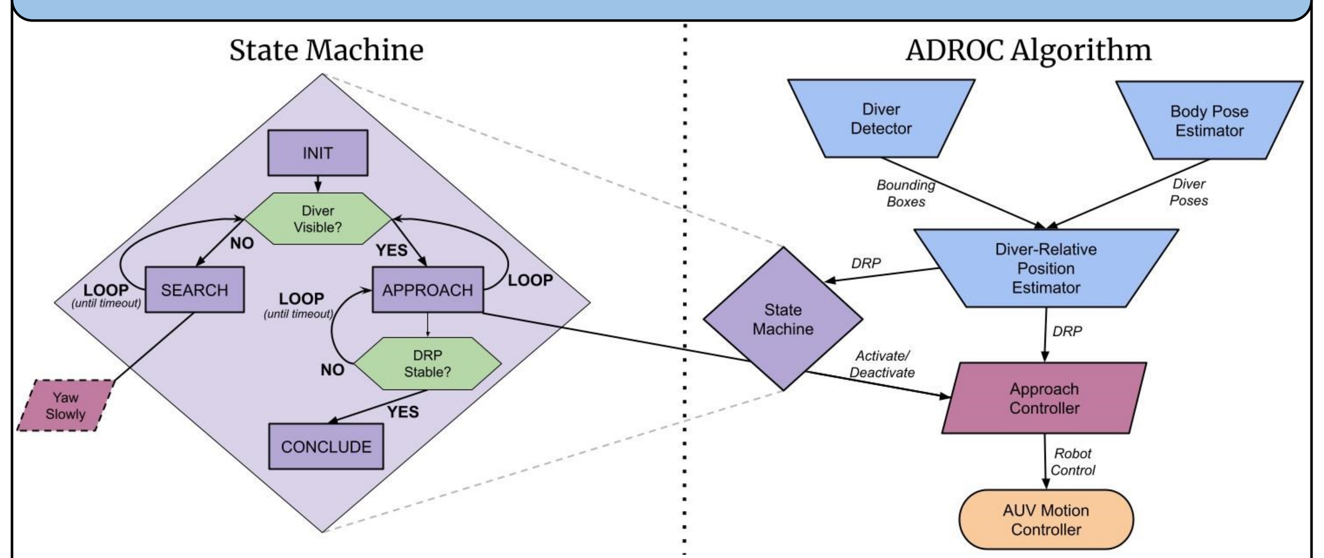
Autonomous Underwater Vehicles (AUVs) which work collaboratively with divers must be able to navigate around their human partners. While research has addressed diver-relative navigation more directly, the task of diver approach has never been addressed for AUVs.

Approaching a diver requires estimating the relative position of the target, including the distance to them. This can be achieved with sensors such as sonar, stereoscopic cameras, etc. To keep the requirements for running this algorithm as minimal as possible, we have developed a monocular vision-based diver approach algorithm utilizing a novel diver distance estimator which uses human body priors. Our algorithm, **A**utonomous **D**iver-**R**elative **O**perator **C**onfiguration, can be run on any AUV with just one camera, an embedded GPU, and sufficient motion control capabilities (at least 3DOF).

LoCO AUV: Designed for Interaction



Flexible, Robust, and Modular Design

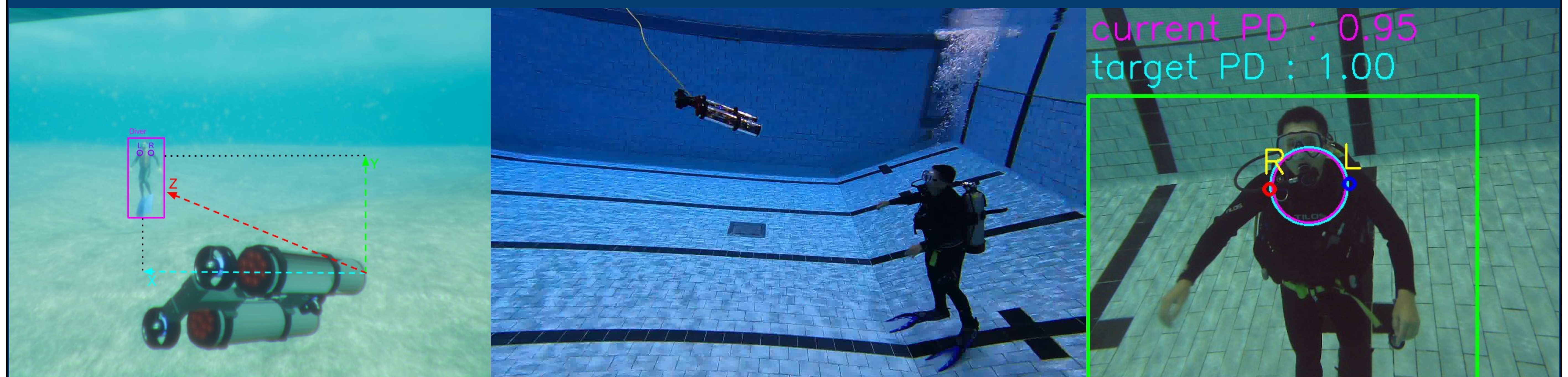


ADORC is a modular algorithm comprised of the following parts:

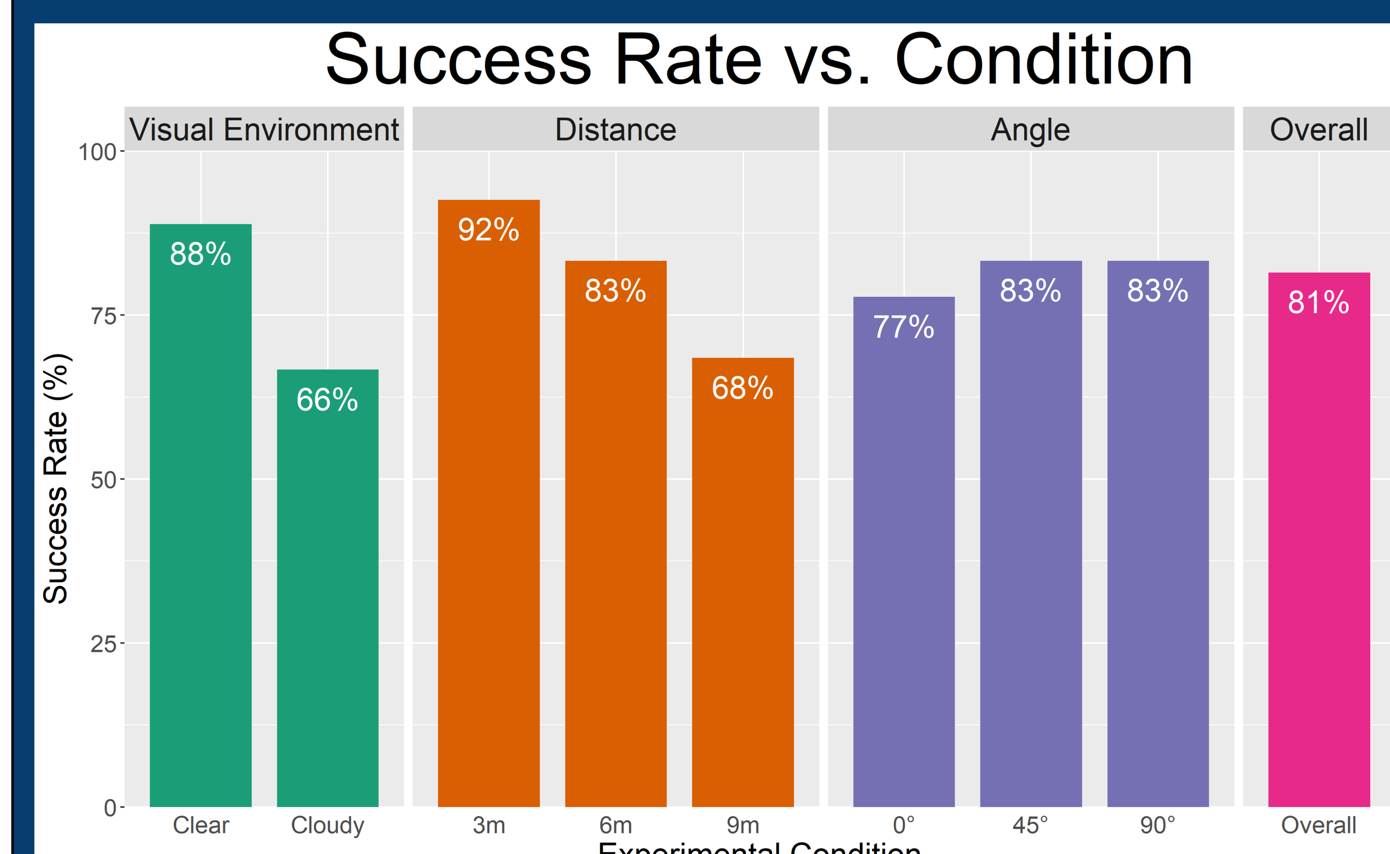
- Two perception modules: a **diver detector** (YOLOv4-Tiny) and a **human body pose estimator** (trt_pose)
- The **diver relative position estimator**, which combines information from the perception modules into a diver-relative position (DRP) estimate. The DRP is comprised of an (x,y) center point and *pseudodistance*, which is the ratio of current diver distance to ideal diver distance.
- An **approach controller** which uses 3 PID controllers to control vehicle yaw, pitch, and surge.
- A **state machine** which controls the robot approach by switching from states of searching and approaching based on the visibility of a diver. The state machine decides when the approach is over by entering a CONCLUDE state once the DRP is stable in an acceptable range .

The core components of ADROC are the state machine, diver-relative position estimator, and approach controller. The perception modules could be replaced with different methods, and the entire algorithm can easily be ported to new AUVs since it is implemented using the Robot Operating System (ROS),

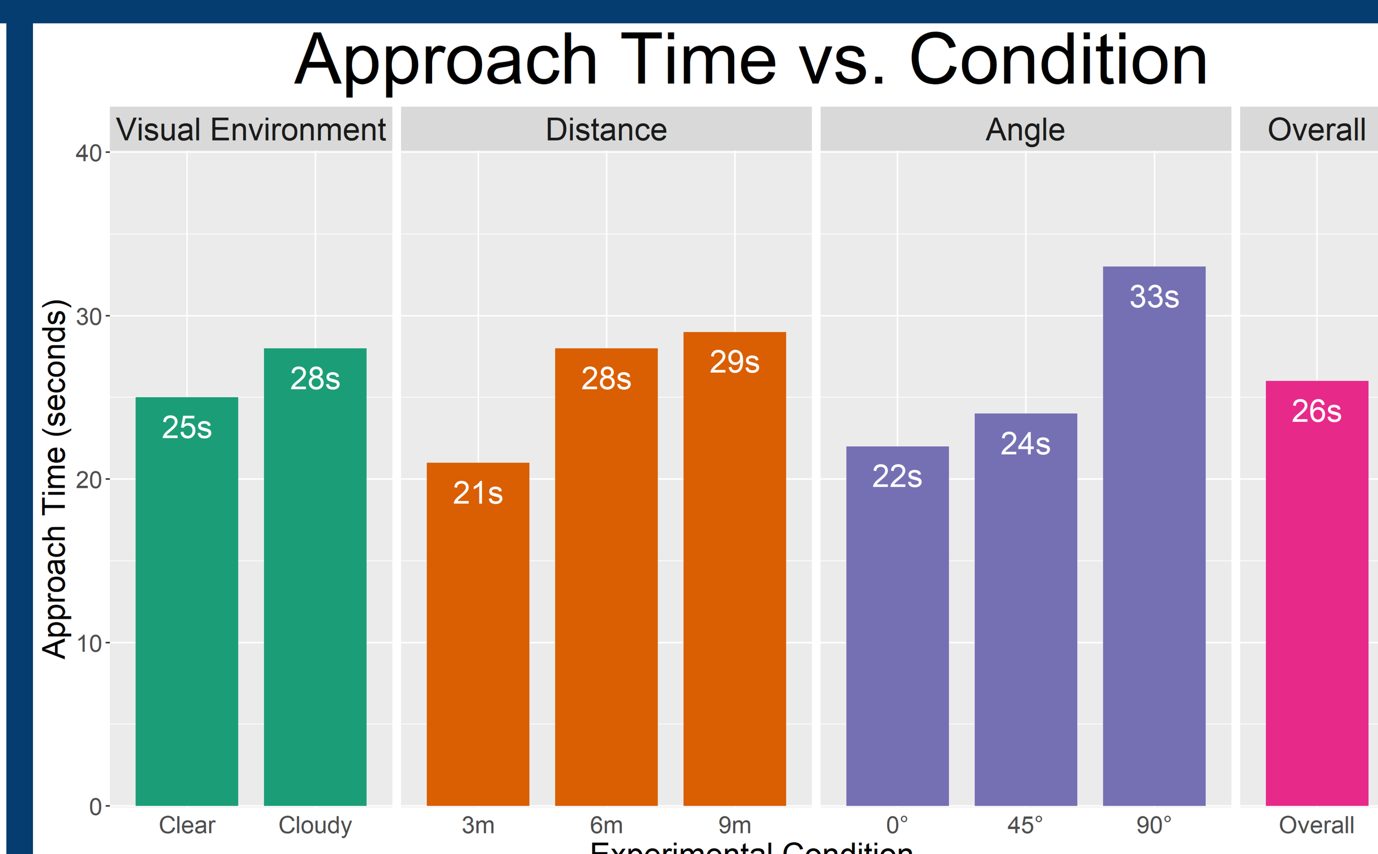
Autonomous Diver-Relative Configuration (ADROC) is a novel monocular vision-based diver approach algorithm.



ADROC allows an AUV to approach a diver using a single camera with an overall success rate of up to 81%.



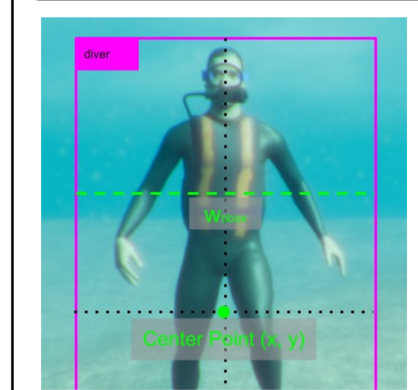
Success rate is the ratio of successful approaches to attempted approaches.



Approach time averaged for only successful approaches.

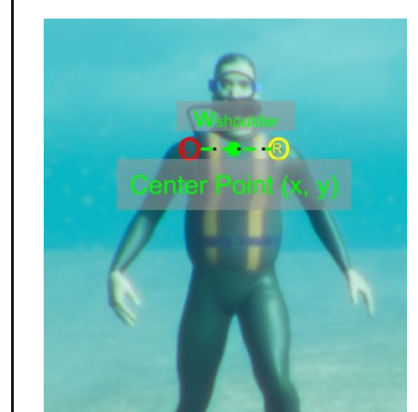
Success rate is negatively correlated with initial distance, but initial angle has less of an effect on success. Visual conditions have the greatest effect on success.

Pseudodistance Enables Monocular Distance Estimation



$$x = x_{\min} + \frac{w_{bbox}}{2}$$
$$y = y_{\min} + \frac{h_{bbox}}{2}$$
$$pseudodistance = \frac{w_{bbox}}{w_{image}} \times \frac{1}{target\_ratio}$$
$$target\_ratio = 0.255$$

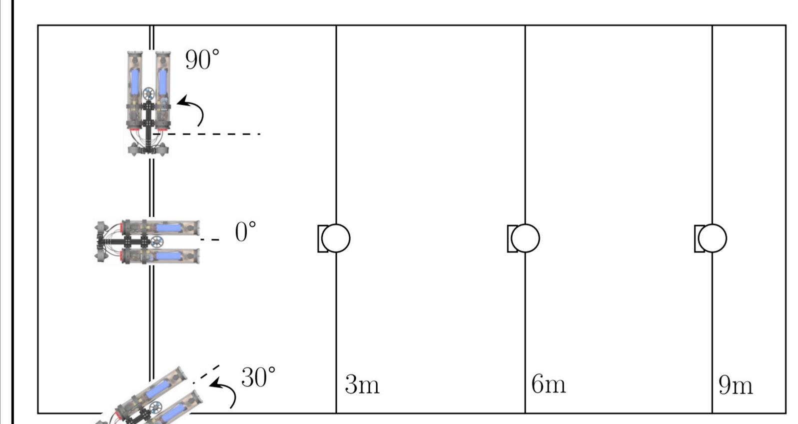
Estimating DRP (*x, y, pd*) from bounding box input. The bounding box DRP is only used if no pose information is available, because bounding box width can vary in unpredictable ways.



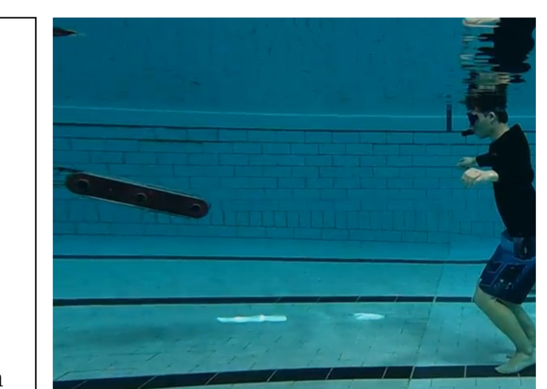
$$x = (x_l + x_r)\,/\,2$$
$$y = (y_l + x_r)\,/\,2$$
$$pseudodistance = \frac{\sqrt{(x_l - x_r)^2 (y_l - y_r)^2}}{w_{image}} \times \frac{1}{target\_ratio}$$
$$target\_ratio = 0.17$$

Estimating DRP (*x, y, pd*) from body pose. Pose-based DRP is used whenever available, as it is based more directly on body priors.

AUVs Can Approach Divers Reliably Using ADROC

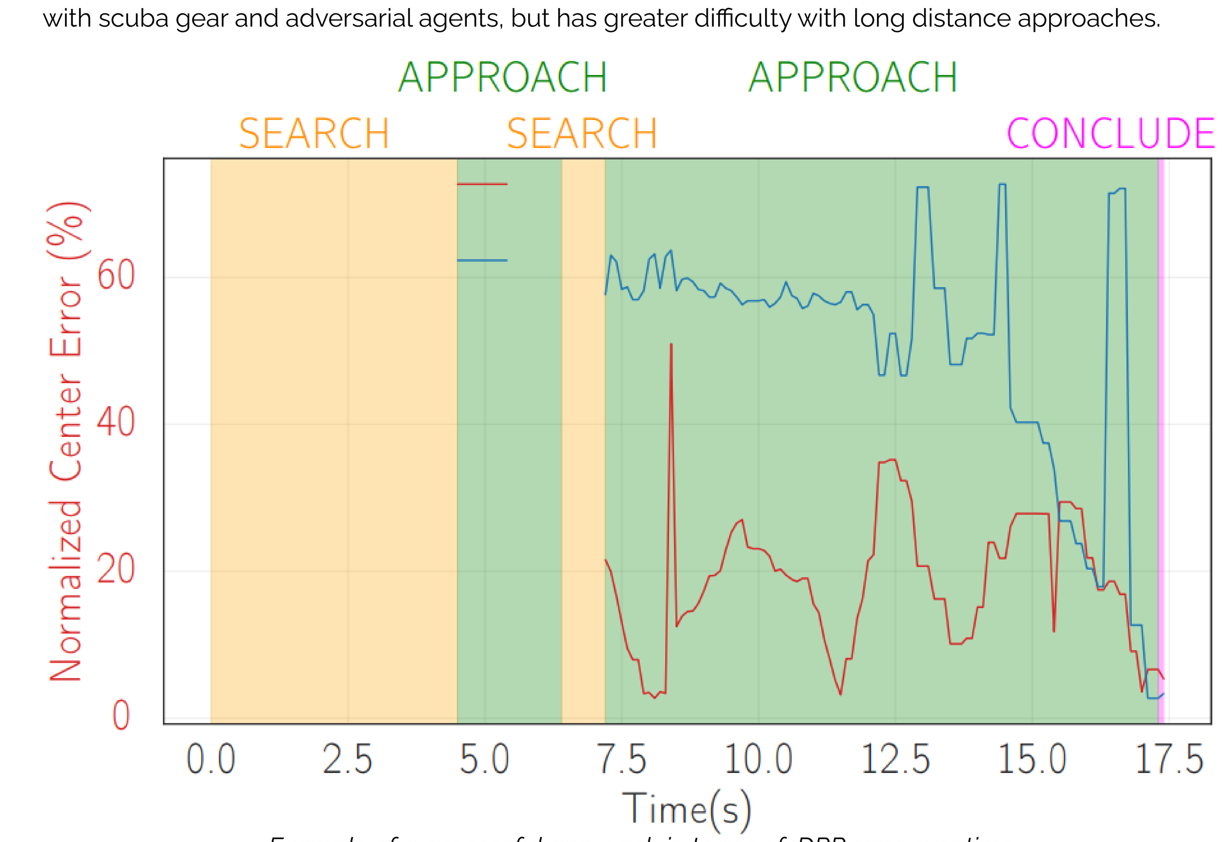


Experimental setup with initial distances and angels.



ADROC on the Aqua AUV (not used for evaluation).

- ADROC testing was conducted by running 162 approaches over the course of two experimental sessions. One session was conducted in clear water, the other was conducted in cloudier water.
- Nine divers were used as the approach target. Three initial distance and angle conditions were used, with two approaches attempted from each initial pose condition (nine poses, 18 attempts per participant).
- An approach is considered a success if ADROC reaches the CONCLUDE state when it is in front of the diver, within arm's reach. Failing to find the diver, oscillating between search and approach, and early CONCLUDE states are all considered failed approaches.
- ADROC has also been evaluated from advanced distances (15m), with adversarial agents (pushing the AUV away), and with divers in scuba gear (results are not in analysis). ADROC can operate successfully with scuba gear and adversarial agents, but has greater difficulty with long distance approaches.



Example of a successful approach in terms of DRP error over time.