CUREE: A Curious Underwater Robot for Ecosystem Exploration

Yogesh Girdhar[1], Nathan McGuire[1], Levi Cai[2], Stewart Jamieson[2], Seth McCammon[1], Brian Claus[1], John E. San Soucie[2], Jessica E. Todd[2], and T. Aran Mooney[1]

Abstract— The current approach to exploring and monitoring complex underwater ecosystems, such as coral reefs, is to conduct surveys using diver-held or static cameras, or deploying sensor buoys. These approaches often fail to capture the full variation and complexity of interactions between different reef organisms and their habitat. The CUREE platform presented in this paper provides a unique set of capabilities in the form of robot behaviors and perception algorithms to enable scientists to explore different aspects of an ecosystem. Examples of these capabilities include low-altitude visual surveys, soundscape surveys, habitat characterization, and animal following. We demonstrate these capabilities by describing two field deployments on coral reefs in the US Virgin Islands. In the first deployment, we show that CUREE can identify the preferred habitat type of snapping shrimp in a reef through a combination of a visual survey, habitat characterization, and a soundscape survey. In the second deployment, we demonstrate CUREE's ability to follow arbitrary animals by separately following a barracuda and stingray for several minutes each in midwater and benthic environments, respectively.

I. INTRODUCTION

Robots have successfully been deployed for extensive adaptive monitoring of underwater ecological phenomena, such as toxic cyanobacteria [1], phytoplankton [2], [3], tracking animals in the mid-water column [4], [5], and benthic surveys of coral reef-like environments [6] [7], [8]. However, their use has either been limited to mapping tasks over the seafloor, or informative path planning over simple scalar observations using mission specific sensors in the mid water column. Coral reef ecosystems consist of complex seafloor geometry, a large number of species, and the manifold of interactions between organisms and their habitats. To monitor reefs, scientists still primarily rely on divers to conduct survey missions. Autonomous Underwater Vehicles (AUVs) offer an unparalleled ability to measure phenomena of scientific interest distributed across space and time. However, to fully observe these ecosystems, robots must develop an understanding of different components of the ecosystem and use these components to perform more targeted hypothesis-driven data collection. This paper presents the Curious

[1]Y. Girdhar, N. McGuire, S. McCammon, B. Claus, and T. A. Mooney are with the Woods Hole Oceanographic Institution (WHOI) {yogi, nmcguire, smccammon, bclaus, amooney}@whoi.edu
[2]L. Cai, S. Jamieson, J. E. San Soucie, and J. E. Todd are with the MIT-WHOI Joint Program in Oceanography/Applied Ocean Science and Engineering {lcai, sjamieson, jsansoucie, jtodd}@whoi.edu

Fig. 1: Illustration of CUREE conducting a reef survey. CUREE uses vision and passive acoustics to collect information about its environment. CuSA assists CUREE by providing a high-bandwidth communications link to scientists and improved localization of CUREE with GPS and USBL.

Underwater Robot for Ecosystem Exploration (CUREE), a robotic system that provides ecologically relevant behaviors and perception subsystems that can be used as building blocks to rapidly construct missions to build understanding of different aspects of the ecosystem.

Images are perhaps the best approach for capturing species-level information about organisms without destructive sampling [9]. However, existing approaches to computer vision underwater are still very noisy. Even in the best case scenario, vision can only be used to make observations within a small range (<5 m) of the robot location. Passive acoustic observations made by hydrophones enable the robot to make observations over much longer ranges. However, these acoustic observations do not contain as much species level information, since not every animal makes measurable sounds. Recent efforts have shown the use of sound source localization (SSL) [10] techniques by robots for the purpose of target acquisition and tracking. We hypothesize that integrating visual and acoustic sensors would be an ideal combination for a robot tasked with monitoring a complex ecosystem. Such a robot would primarily use images for species level observation, identify different habitat types, and for collision avoidance close to the seafloor, and use passive acoustics observation to approximate biodiversity and animal presence over longer spatial scales.

The novel physical design of the robot was guided by identifying requirements for transportability and deployment

to remote parts of the world, heat dissipation and computation associated with running state-of-the-art perception and planning algorithms, need for in-water behaviors such as drifting, and enabling continuous rapid research and development efforts through a modular design.

CUREE can be optionally accompanied by CUREE's Surface Assistant (CuSA), a surface vehicle which autonomously follows CUREE while acting as a radio communication relay and providing CUREE with its global position information (Fig. 1).

II. MONITORING ECOSYSTEMS: KEY CAPABILITIES

The biological components of a coral reef ecosystem are made up of stationary plants, corals, and algae attached to the seafloor, as well as fish, turtles, other mobile animals, and microscopic organisms which move about in the water column. The geophysical component is defined in terms of seafloor substrate types, depth, and the physical and chemical properties of the water. CUREE aims to perform missions that target observations of different components of the ecosystem with its cameras and hydrophones. CUREE's core capabilities are summarized in the subsections below.

A. Low Altitude Benthic Surveys

Conducting benthic surveys is perhaps the most common mode of operation for most AUVs. However, unlike standard bathymetric surveys with sonar devices, such as Doppler Velocity Loggers (DVLs) or echosounders, which can be performed while keeping the vehicle at high altitudes, the visual benthic surveys used to assess coral and fish biodiversity require AUVs to operate close to the sea floor to obtain high-resolution imagery. Acoustic sensors work well when there is a hard, mostly flat bottom. However, they often fail in the presence of soft or narrow objects such as soft corals and plants that are commonly encountered by AUVs operating at <1.5m altitude in coral reefs.

To enable low-altitude observations we rely on the vehicle's DVL to provide a robust estimate of the distance between CUREE and the seafloor, up to 1.5m. By fusing the altitude estimate with local position information from CUREE's DVL and IMU, and with global positioning from its USBL, CUREE can localize itself in complex reef-like environments.

Similarly, CUREE can localize the high-resolution imagery as well. We can then process this imagery offline to produce 3D representations of the reef [11], which is represented by the distribution $P(\text{img}|x)$. An example of such a 3D reef survey is shown in Fig. 2.

However, we anticipate future experiments where CUREE would need to operate even closer to the seafloor (e.g. to take a water sample near a specific coral head) or near narrow objects, such as pillar coral. Doing so in soft or narrow terrains may make the DVL estimate unreliable. As an alternative CUREE can produce an altitude estimate created from stereo-vision to provide a high-resolution alternative to the DVL estimate.

Fig. 2: 3D reconstruction of Booby Rock Reef, St. John, USVI, produced from a CUREE visual survey. The AprilTag [12] shown in the cutout is 20cm × 20cm, while the complete survey is approximately 20m × 20m.

B. Unsupervised Substrate Characterization

Ecologists often characterize reef habitats in terms of different substrate types such as coral, rubble, algae, and sand, and how they are distributed in space. The task of discovering and mapping different habitat types can be understood as learning the factorization of the spatial distribution of the visual seafloor observations $P(\text{img}|x)$ in terms of an appearance model of each habitat distribution $P(\text{img}|\text{habitat})$ and the spatial distribution of different habitats $P(\text{habitat}|x)$:

$$P(\text{img}|x) \quad = \quad \sum_{\text{habitats}} P(\text{img}|\text{habitat})P(\text{habitat}|x).$$

Giving AUVs the ability to discover habitat types and provide habitat labels to locations in real-time can enable observations of organisms to be understood in the context of their habitat, identifying anomalous migrations.

However, factors such as bad visibility, presence of different underlying species or corals or algae, overlapping categories, altitude, and the varying health condition of the reef can make it hard to categorically label the observed seafloor using a supervised neural network classifier. Moreover, moving beyond coral reefs, there exist many different types of substrate types for which there is insufficient prior data to train a convolution neural net (CNN).

Instead of using a pre-trained supervised classifier to detect specific habitat types, CUREE uses an unsupervised approach to automatically discover different habitat types in realtime, and use them to produce a continuously updating habitat map $P(\text{habitat}|x)$. The approach uses a Hierarchical Dirichlet Process (HDP) based topic model [13], which does inference in realtime using a Gibbs sampler. The unsupervised learning and realtime inference properties of this approach enables CUREE to be deployed in arbitrary environments with no requirements on prior knowledge.

C. Soundscape Surveys

Marine soundscapes are composed of biological, geological, physical, and anthropogenic sound sources. While visual observations can provide dense and specific information about the robot's immediate surroundings, they suffer from

Fig. 3: CUREE's primary sensing capabilities come from forward and downward-looking stereo cameras and a four-channel hydrophone array (Photo credit: Austin Greene).

Fig. 4: Sample Acoustic Survey conducted at Joel's Shoal Reef, US Virgin Islands. During drifting periods, there is no noise from the thrusters, allowing CUREE to detect sounds from marine animals, such as snapping shrimp.

Previous attempts to capture the spatial variation of soundscapes have been through the deployment of fixed hydrophone arrays [16], drifting buoys equipped with hydrophones that passively drift across a region of interest [17], or hydrophone-equipped gliders [18], [19], [20]. However, these approaches are limited in their effectiveness in capturing spatial variation. Static arrays can only observe in one location and can only capture the direction of arrival of a sound, not the exact source location. Since they can move through an environment, drifters and gliders can better capture the spatial variation of soundscapes. However, drifters' ability to cover an environment is dictated entirely by the ocean currents, and gliders cannot operate in the shallow waters of many coral reefs. Due to its size and maneuverability, CUREE is capable of operating in these shallow, highly dynamic environments. However CUREE's thruster noise saturates its hydrophone's dynamic range making it difficult to collect useful acoustic data while moving. CUREE addresses this by interleaving periodic drifting behavior, free from thruster noise, during a regular benthic survey (Fig. 4). As a result, CUREE can make soundscape observations at arbitrary locations along a mission, enabling adaptive soundscape-guided missions.

D. Animal Tracking

CUREE provides a novel capability to follow arbitrary animals in the wild. Unlike tracking animals in the mid-water column, where traditional computer vision techniques like blob detection are sufficient to lock on to slow-moving targets (e.g. jellyfish or larvaceans) [21], tracking animals on a coral reef requires the ability to separate fast-moving fish from complex backgrounds. CUREE addresses this problem through the use of a semi-supervised visual tracker [22][23]. Given a single bounding box annotation of an animal in the robot's view, CUREE visually locks on and follows, enabling it to observe the spatiotemporal distribution of the animal

$$P(x, t | \text{animal})$$

over long periods of time.

These types of observations (long continuous tracks of organisms) are invaluable to marine biologists, since they provide a unique perspective on the animal's behavior that may not be replicable in a laboratory setting. Stationary camera traps will fail to capture this distribution, since they can only observe $P(\text{species} | x)$. The primary other way to obtain these observations is to use human divers, an approach which scales poorly. This visual tracker provides the robot with the relative position of the animal, which then is used by the tracking controller to move the robot to keep the target animal in the center of the robot's view.

CUREE's animal tracking behavior, when coupled with other instances of the same or other behaviors, enables us to answer many interesting ecologically relevant questions. For example, correlating the spatiotemporal distribution of the animal with the habitat distribution gives us information

several drawbacks. Underwater cameras have a limited range and can only provide information about objects within the cameras field of view. In contrast, hydrophones provide an omnidirectional sensor with much greater range. Multiple hydrophones in an array enable Sound Source Localization (SSL) by comparing the time difference of arrival of sound at the different array elements, or through more complex signal analysis techniques, including beamforming [14]. CUREE's hydrophone payload consists of four synchronized hydrophones positioned away from the robot's body at the end of aluminum arms as can be seen in Fig. 3.

Soundscapes observed over a coral reef show variability in both space and time due to diurnal animal activity, tides, and storms [15]. The observed distribution of acoustic features (spectrogram) can be represented by the distribution

$$P(\text{sound} | x, t).$$

about the animal's preferred habitats

$$P(\text{habitat}|\text{animal}) = \sum_x P(\text{habitat}|x)P(x|\text{animal});$$

and correlating with observations of other animal tracks, assuming independent tracks of individual animals in the same region, gives us information about inter-species interactions

$$P(\text{animal}_1, \text{animal}_2) \approx \sum_x P(\text{animal}_1|x)P(\text{animal}_2|x).$$

III. Field Demonstrations on Coral Reefs

CUREE has been field-tested in dozens of deployments on reefs around St. John in the U.S. Virgin Islands. In this section, we describe two deployments that demonstrate CUREE's effectiveness in evaluating ecosystem-related hypotheses and monitoring dynamic phenomena. The first of these is an Audio-Visual survey of a coral reef conducted at Joel's Shoal. This survey combines the unsupervised substrate characterization and soundscape survey to learn the preferred habitat of snapping shrimp. The second deployment was conducted at Tektite Reef, where CUREE demonstrated autonomous animal tracking by following a barracuda and a stingray, for several minutes, after CUREE was provided with an initial bounding box for each.

A. Audio-Visual Benthic Survey

In this experiment, we deployed CUREE at Joel's Shoal reef in USVI, and conducted an audiovisual survey with 9370 images, interleaved with 50 drifting soundscape observations. The goal of this deployment was to collect a dataset to test whether it is possible to predict the presence of snapping shrimp using purely visual information. Snapping shrimp are known for their continuous high-frequency snaps, and their presence is mainly observed acoustically as they hide in rock crevices and burrows [24]. To simplify, we ignore acoustic multipath effects, though these can be considered in future experiments with the multi-hydrophone configuration.

Given the observed soundscape, we first built an acoustic snapping shrimp snap detector using a method similar to [25], where we identify transient spikes in acoustic activity within the shrimp band (2 kHz to 24 kHz) above a threshold of 0.1σ over the mean. The resultant counts of shrimp snaps form the distribution $P(\text{shrimp}|\text{sound})$. Given this shrimp detector and the soundscape survey, we then estimate the snap distribution over time as

$$P(\text{shrimp}|t) = \sum_{\text{sound}} P(\text{shrimp}|\text{sound})P(\text{sound}|t).$$

This distribution is visualized as the black dashed line in the upper plot of Fig. 5 and the black line in Fig. 6.

Alternatively, we assume a generative model for snap distribution which only assumes knowledge of a timeseries of visual observations:

$$P(\text{shrimp}|\text{img}, t)$$
$$= \sum_{\text{habitat}} P(\text{shrimp}|\text{habitat})P(\text{habitat}|\text{img}, t),$$

where $P(\text{habitat}|img, t)$ is computed offline using HDP-ROST [13]. $P(\text{habitat}|\text{img}, t)$ can be seen in the middle plot in Fig. 5, where each color in the time series corresponds to the relative prevalence of that habitat type in the imagery collected at time t. From a qualitative examination of the images associated with each topic, we were able to determine that Topic 3, shown in blue, corresponded with observations of coral-covered regions. $P(\text{shrimp}|\text{habitat})$ is modeled by a linear least squares regression mapping habitat distribution to $P(\text{shrimp}|\text{sound})$. The temporal snap distribution $P(\text{shrimp}|t)$ is plotted as the blue dashed line in Fig. 6.

We find that the snap density as predicted by only visual observations correlates positively with the snap density computed using the soundscape data. The only topic for which there was a positive coefficient in the linear regression is Topic 3, the topic associated with rocky coral habitat. While more experiments are necessary, this supports the hypothesis that snapping shrimp prefer the dense corals around Joel's Shoal as their habitat, where there are many more rock crevices and tiny spaces to hide in.

B. Tracking Organisms

In this experiment we evaluated CUREE's ability to autonomously track moving objects in a reef environment using vision. Vision can be especially useful in shallow (¡100m) reef environments since, during daytime, there is typically ample light and low turbidity, while sonar is generally ineffective due to severe multipathing in reef environments. For the tracker we used a state-of-the-art semi-supervised tracker, SiamMask [26], which we deployed on CUREE without requiring any additional training for the marine domain. During the experiment, SiamMask was running at 15 fps on 360p images, fully on-board the vehicle. CUREE's tracking controller operated in along 3 axes: yaw, surge, and heave to try to keep the bounding box produced by SiamMask centered within the image frame. The heave and yaw actions are straightforward to compute from the bounding box position; the control input is proportional to the distance between the center of the bounding box and the center of the image. However, since SiamMask is a monocular tracker, we cannot directly compute the surge feedback from the distance to the tracked target without performing registration to an object in a stereo depth image. Instead, we opted for a simpler heuristic where the tracking controller attempts to keep the ratio of the width of the SiamMask bounding box to the width of the image constant. In doing so, as the object grows larger in the image CUREE will move backwards away from it, and move closer if the object appears to get smaller.

This tracker was deployed in two experiments near Tektite Reef in USVI. In the first experiment, CUREE tracked a Barracuda as it swam in the water column above the reef, and in the second CUREE tracked a stingray swimming near the bottom of the reef. In the Barracuda tracking experiment, CUREE was able to maintain its track for five minutes. During the track, the barracuda swam through schools of small fish, as can be seen in Fig. 7 (upper-right, lower-

Fig. 5: CUREE is capable of conducting simultaneous visual and acoustic surveys over complex seafloor environments such as a coral reefs. The approach mixes low-altitude terrain and waypoint following behavior with drifting periods to capture soundscapes. In this survey of Joel's Shoal in the USVI, CUREE drifts for 10 seconds at every waypoint to enable soundscape observation at that location. Top: Spectrogram of audio observations at each waypoint and average observed snapping shrimp snap rate during the drifting window. Middle: Automatically computed visual topic labels representing different habitat types. Bottom: Examples of imagery captured by the cameras at different points of time. We see that images 2, 3, and 5 have similar topic distribution and correspond to a coral-covered habitats, whereas images 1 and 4 have similar topic distributions (but distinct from 2,3,5) and show sandy habitats. We see that Topic 3 (blue), which indicates the presence of dense corals in the images, correlates positively with high rates of snapping shrimp snaps.

Fig. 6: Plot shows normalized snapping shrimp snap rate as computed for soundscape observations (black), and the least squared linear fit of visual topic distribution to the normalized snap rate observations (blue dashed). We see that there is a strong correlation between the observed shrimp snaps and the predicted ones.

left). Another significant challenge for the tracker was the considerable change in appearance and apparent size of the barracuda depending on whether it was being viewed from the side (Fig. 7 top-left), or from behind (Fig. 7 bottom-right). This created an unforeseen emergent behavior within the CUREE tracking controller, where CUREE followed the barracuda much more closely while the fish was swimming directly away from CUREE. However, despite being closely followed, the barracuda didn't seem perturbed by CUREE's presence, instead choosing to swim slowly through the water.

To further test the limits of the tracker, we selected a benthic target that had a more complex background. We were able to track a stingray, which can be seen in Fig. 7, for several minutes. The stingray was followed by a symbiotic fish, which occasionally caused the tracker to follow the fish instead, though CUREE was able to eventually continue tracking the stingray when the fish returned to the stingray.

This situation highlights the inherent challenges in this type of tracking. For further experiments and analysis refer to [27].

In future work, the stereo-cameras on-board the AUV can be used to extract relative position, and hence an estimate of telemetry of the targets can be computed by merging those measurements with vehicle position estimates.

IV. HARDWARE DESIGN AND CONTROL

CUREE is designed to be easily transportable and deployable by a single person anywhere in the world. CUREE and its essential gear can fit in a pair of $27 \times 16 \times 9$ inch Pelican cases, so it fits in the baggage allotment of most commercial airlines. However, the dry weight of 55 lbs means that transporting CUREE this way will likely incur overweight baggage fees. CUREE can operate for multiple hours using one or two 14.8V 10Ah batteries from BlueRobotics, which can be carried on most commercial flights.

Aspects of CUREE's hardware design, specifically motor configuration and default camera configuration, were inspired by the Rangerbot [28]. However, CUREE is designed to be significantly more flexible, to support a larger range of scientific and engineering research, with modular sensor payloads and options for the head module that can be tailored to mission-specific camera configurations. CUREE utilizes a series of SubConn and BlueTrail connectors to enable this modularity. Fig. 8 shows an expanded view of CUREE with its different housings and sensors.

High maneuverability is necessary for low-altitude missions and collision avoidance in reef like environments and to track marine animals. Whereas stable forward motion is important for performing higher altitude survey tasks. To achieve both of these requirements, we use six Blue Robotics T200 thrusters in a configuration inspired by the

Fig. 7: Still frames taken from autonomous tracking experiments. (TOP) CUREE follows a barracuda for five minutes, through significant changes in target appearance, and as the barracuda was partially occluded while swimming through schools of other fish. (BOTTOM) CUREE tracks a stingray for three minutes in a complex benthic scenario, though a symbiotic fish occasionally distracts the tracker.

Fig. 8: Exploded view of CUREE. The head module contains the twin stereo pairs (survey cameras and tracking cameras) and can be swapped between a head with the tracking camera mounted forward facing and downward at $45°$. While the Sensor Payload housing currently holds the hydrophone recorder, it is connected to the main housing by ethernet and a power cable, and can therefore be easily swapped for other sensors (e.g. a water sample collector).

Rangerbot [28]. Four thrusters are mounted in a quadrotor-like configuration but are then tilted 45-degrees inward to also allow for sideways motion. Two thrusters on the rear of the vehicle point directly aft, and control forward motion and yaw. This thruster configuration minimizes the amount of internal space taken up by the thrusters, keeps the vehicle streamlined in the forward direction, and provides full six degress of freedom. Polyurethane foam can be attached to

attain neutral buoyancy, improve stability, and level trim in cases where surveying is more important than tracking.

CUREE's head module contains a downward-looking stereo camera, for substrate characterization or altitude estimation, and a forward-looking stereo camera suitable for animal tracking and collision avoidance, different head modules provide different angles, numbers, and types of cameras. In addition, the head module contains a NVIDIA Jetson Xavier AGX computer that runs all the neural networks and inference algorithms needed to implement the perception and path planning pipelines. CUREE also dissipates heat from the high-powered computers directly into the water by ensuring they are connected to the top of the aluminum head.

The main robot housing is a 6-inch cylindrical acrylic tube which contains the primary battery, networking equipment, and the primary control computer: a Raspberry Pi 4. A Water-linked DVL and an IMU observe local changes in CUREE's position, orientation, altitude, and velocity, enabling dead reckoning positioning. When deployed in combination with the USBL-equipped CuSA, CUREE also gets global positioning updates which are fused with dead reckoning based state estimates using an Extended Kalman Filter (EKF). Altimetry or terrain relative navigation is obtained by fusing the vehicle's echosounders, DVLs, and stereo imaging based depth estimates. Both the control computer in the main housing and the Jetson in the robot's head run ROS, and all micro-controller code is written using Arduino, providing a fully open-source environment for development of science missions and implementation of new algorithms on CUREE.

CUREE's surface assistant (CuSA) has a GPS receiver and a USBL modem, allowing CUREE to obtain a global position reference while underwater. CuSA also has a radio antenna, which allows high-bandwith information to be transmitted to scientists at the vehicles' base station. However, all of the computation and power necessary to operate CUREE is onboard the vehicle, allowing it to operate independently from CuSA if desired.

V. SUMMARY AND FUTURE WORK

CUREE is a robot designed to explore underwater ecosystems, observe complex interactions between the organisms that live there and their habitats, and use these observations in real-time to adapt its behavior as an intelligent partner for marine science. As a compact system designed to be deployed and operated by teams as small as a single person, CUREE can be taken anywhere in the world in checked luggage on commercial airlines and deployed without a need for significant supporting infrastructure. In experiments in the U.S. Virgin Islands, we demonstrated how CUREE can be used to study coral reefs, by combining audio and visual observations of a coral reef to infer the preferred habitat of snapping shrimp, or by tracking a barracuda as it hunts above a reef. As we continue the development of CUREE we plan to incorporate more active decision-making into its missions, enabling it to make informed decisions about which habitat types it should observe to gain the most information about the species that live there, or when to switch tracking targets.

REFERENCES

[1] G. Hitz, E. Galceran, M. E. Garneau, F. Pomerleau, and R. Siegwart, "Adaptive continuous-space informative path planning for online environmental monitoring," *Journal of Field Robotics*, vol. 34, no. 8, pp. 1427–1449, 12 2017. [Online]. Available: http://doi.wiley.com/10.1002/rob.21722

[2] T. O. Fossum, G. M. Fragoso, E. J. Davies, J. E. Ullgren, R. Mendes, G. Johnsen, I. Ellingsen, J. Eidsvik, M. Ludvigsen, and K. Rajan, "Toward adaptive robotic sampling of phytoplankton in the coastal ocean," *Science Robotics*, vol. 4, no. 27, p. eaav3041, 2 2019. [Online]. Available: http://robotics.sciencemag.org/lookup/doi/10.1126/scirobotics.aav3041

[3] J. Das, F. Py, J. B. Harvey, J. P. Ryan, A. Gellene, R. Graham, D. A. Caron, K. Rajan, and G. S. Sukhatme, "Data-driven robotic sampling for marine ecosystem monitoring," *International Journal of Robotics Research*, 2015.

[4] K. Benoit-Bird, T. Patrick Welch, C. Waluk, J. Barth, I. Wangen, P. McGill, C. Okuda, G. Hollinger, M. Sato, and S. McCammon, "Equipping an underwater glider with a new echosounder to explore ocean ecosystems," *Limnology and Oceanography: Methods*, vol. 16, no. 11, pp. 734–749, 2018.

[5] D. R. Yoerger, A. F. Govindarajan, J. C. Howland, J. K. Llopiz, P. H. Wiebe, M. Curran, J. Fujii, D. Gomez-Ibanez, K. Katija, B. H. Robison, B. W. Hobson, M. Risi, and S. M. Rock, "A hybrid underwater robot for multidisciplinary investigation of the ocean twilight zone," *Science Robotics*, vol. 6, no. 55, jun 2021. [Online]. Available: https://www.science.org/doi/10.1126/scirobotics.abe1901

[6] T. Manderson, J. C. Gamboa, S. Wapnick, J. F. Tremblay, F. Shkurti, D. Meger, and G. Dudek, "Vision-Based Goal-Conditioned Policies for Underwater Navigation in the Presence of Obstacles," in *Robotics: Science and Systems*, 2020.

[7] R. Armstrong, H. Singh, S. Rivero, and F. Gilbes, "Monitoring coral reefs in optically-deep waters," *Proceedings of the 11th Int Coral Reef Symposium*, no. 17, pp. 7–11, 2008. [Online]. Available: http://www.gers.uprm.edu/pdfs/armstrong_etal_08.pdf

[8] H. Singh, C. Roman, O. Pizarro, R. Eustice, and A. Can, "Towards High-resolution Imaging from Underwater Vehicles," *The International Journal of Robotics Research*, vol. 26, no. 1, pp. 55–74, 1 2007. [Online]. Available: http://ijr.sagepub.com/cgi/doi/10.1177/0278364907074473

[9] B. L. Richards, O. Beijbom, M. Campbell, M. Clarke, G. Cutter, M. Dawkins, Edginton D, D. Hart, M. Hill, A. Hoogs, D. Kriegman, E. Moreland, T. Oliver, W. Michaels, M. Piacentino, A. Rollo, C. Thompson, F. Wallace, I. Williams, and K. Williams, "Automated Analysis of Underwater Imagery: Accomplishments, Products, and Vision A Report on the NOAA Fisheries Strategic Initiative on Automated Image Create new project "Kahekili Herbivore Fisheries Management Area" View project Automated Image Analysis," *NOAA Technical Memorandum NOAA-TM-NMFS-PIFSC-83*, p. 59, 2019. [Online]. Available: https://doi.org/10.25923/0cwf-4714

[10] C. Rascon and I. Meza, "Localization of sound sources in robotics: A review," *Robotics and Autonomous Systems*, vol. 96, pp. 184–210, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.robot.2017.07.011

[11] R. Ferrari, W. F. Figueira, M. S. Pratchett, T. Boube, A. Adam, T. Kobelkowsky-Vidrio, S. S. Doo, T. B. Atwood, and M. Byrne, "3D photogrammetry quantifies growth and external erosion of individual coral colonies and skeletons," *Scientific Reports*, vol. 7, no. 1, p. 16737, dec 2017. [Online]. Available: http://www.nature.com/articles/s41598-017-16408-z

[15] A. Lillis and T. A. Mooney, "Snapping shrimp sound production patterns on caribbean coral reefs: relationships with celestial cycles and environmental variables," *Coral Reefs*, vol. 37, no. 2, pp. 597–607, 2018.

[12] E. Olson, "AprilTag: A robust and flexible visual fiducial system," in *2011 IEEE International Conference on Robotics and Automation*. Shanghai, China: IEEE, May 2011, pp. 3400–3407. [Online]. Available: http://ieeexplore.ieee.org/document/5979561/

[13] Y. Girdhar, L. Cai, S. Jamieson, N. McGuire, G. Flaspohler, S. Suman, and B. Claus, "Streaming Scene Maps for Co-Robotic Exploration in Bandwidth Limited Environments," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 5 2019, pp. 7940–7946. [Online]. Available: http://arxiv.org/abs/1903.03214https://ieeexplore.ieee.org/document/8794132/

[14] R. J. Urick, "Principles of underwater sound-2," 1975.

[16] S. E. Freeman, M. J. Buckingham, L. A. Freeman, M. O. Lammers, and G. L. D'Spain, "Cross-correlation, triangulation, and curved-wavefront focusing of coral reef sound using a bi-linear hydrophone array," *The Journal of the Acoustical Society of America*, vol. 137, no. 1, pp. 30–41, 2015.

[17] A. Lillis, F. Caruso, T. A. Mooney, J. Llopiz, D. Bohnenstiehl, and D. B. Eggleston, "Drifting hydrophones as an ecologically meaningful approach to underwater soundscape measurement in coastal benthic habitats," *Journal of Ecoacoustics*, vol. 2, no. 10.22261, 2018.

[18] M. F. Baumgartner and D. M. Fratantoni, "Diel periodicity in both sei whale vocalization rates and the vertical migration of their copepod prey observed from ocean gliders," *Limnology and Oceanography*, vol. 53, no. 5part2, pp. 2197–2209, 2008.

[19] A. Dassatti, M. Van der Schaar, P. Guerrini, S. Zaugg, L. Houégnigan, A. Maguer, and M. Andre, "On-board underwater glider real-time acoustic environment sensing," in *OCEANS 2011 IEEE-Spain*. IEEE, 2011, pp. 1–8.

[20] C. C. Wall, C. Lembke, and D. A. Mann, "Shelf-scale mapping of sound production by fishes in the eastern gulf of mexico, using autonomous glider technology," *Marine Ecology Progress Series*, vol. 449, pp. 55–64, 2012.

[21] D. R. Yoerger, A. F. Govindarajan, J. C. Howland, J. K. Llopiz, P. H. Wiebe, M. Curran, J. Fujii, D. Gomez-Ibanez, K. Katija, B. H. Robison, *et al.*, "A hybrid underwater robot for multidisciplinary investigation of the ocean twilight zone," *Science Robotics*, vol. 6, no. 55, p. eabe1901, 2021.

[22] L. Cai, R. Hanlon, and Y. Girdhar, "[poster] Evaluating semi-supervised, in-situ visual tracking methods for marine aninmals," in *Ocean Sciences Meeting (OSM)*, 2022.

[23] W. Hu, Q. Wang, L. Zhang, L. Bertinetto, and P. H. S. Torr, "SiamMask: A Framework for Fast Online Object Tracking and Segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, jul 2022. [Online]. Available: http://arxiv.org/abs/2207.02088

[24] J. Duffy, "Crustacean social evolution," in *Encyclopedia of Animal Behavior*, M. D. Breed and J. Moore, Eds. Oxford: Academic Press, 2010, pp. 421–429. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780080453378003508

[25] C. A. Radford, A. G. Jeffs, C. T. Tindle, and J. C. Montgomery, "Temporal patterns in ambient noise of biological origin from a shallow water temperate reef," *Oecologia*, vol. 156, no. 4, pp. 921–929, 2008.

[26] Q. Wang, L. Zhang, L. Bertinetto, W. Hu, and P. H. Torr, "Fast online object tracking and segmentation: A unifying approach," in *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, 2019, pp. 1328–1338.

[27] "Semi-supervised visual tracking of marine animals using autonomous underwater vehicles," *International Journal of Computer Vision*, 2023.

[28] L. Zeldovich, "The starfish terminator," *Mechanical Engineering*, vol. 140, no. 10, pp. 36–41, 2018.