Towards Multi-Robot Shipwreck Mapping

Marios Xanthidis[a], Bharat Joshi[a], Nare Karapetyan[a], Monika Roznere[b], Weihan Wang[c], James Johnson[a], Alberto Quattrini Li[b], Jesse Casana[b], Philippos Mordohai[c], Srihari Nelakuditi[a], Ioannis Rekleitis[a]

***Abstract*— This paper introduces on-going work about a novel methodology for cooperative mapping of an underwater structure by a team of robots, focusing on accurate photo-realistic mapping of shipwrecks. Submerged vessels present a history capsule and they appear all over the world; as such it is important to capture their state through visual sensors. The work in literature addresses the problem with a single expensive robot or robots with similar capabilities that loosely cooperate with each other. The proposed methodology utilizes vision as the primary sensor. Two types of robots, termed distal and proximal observers, having distinct roles, operate around the structure. The first type keeps a distance from the wreck providing a "bird's"-eye-view of the wreck in sync with the pose of the vehicles of the other type. The second type operates near the wreck mapping in detail the exterior of the vessel. Preliminary results illustrate the potential of the proposed strategy.**

I. Introduction

This paper initiates the discussion on underwater structure mapping at different scales utilizing a team of cooperative robots. Underwater structure modeling is crucial for operating in different natural and man-made environments. These environments are diverse and include shipwrecks, oil-rigs and hydroelectric dams, submerged historical sites, and cave systems. Operating in the underwater domain is dangerous, tedious, labor intensive, and physically exhausting for humans; fortunately, underwater robots can enable such operations. However, the underwater domain poses unique challenges, including absence of localization systems (e.g., GPS) and communication infrastructure (e.g., WiFi), challenging visibility conditions, and external forces (currents). This paper proposes a novel methodology for cooperative mapping of an underwater structure by a team of robots. In particular, this paper introduces and discusses the following research questions with some preliminary results on this on-going work:

- **RQ1:** How to robustly achieve cooperative localization in the presence of occlusions?
- **RQ2:** How to fuse the different sources of information on-board in real time for reconstruction?
- **RQ3:** How should the co-robots cooperate for the mapping task?

[a]University of South Carolina, Columbia, SC, USA, 29208, {mariosx, bjshi,nare,jvj1}@email.sc.edu, {yiannisr,srihari}@cse.sc.edu.
[b]Dartmouth College, Hanover, NH, USA, 03755 {monika.roznere.gr, alberto.quattrini.li, jesse.j.casana}@dartmouth.edu
[c]Stevens Institute of Technology, Hoboken, NJ, USA, 07030, {wwang103,pmordoha}@stevens.edu
This research has been supported in part by the National Science Foundation under grants 1943205, 2024741, 2024541 and 2024653, 1919647

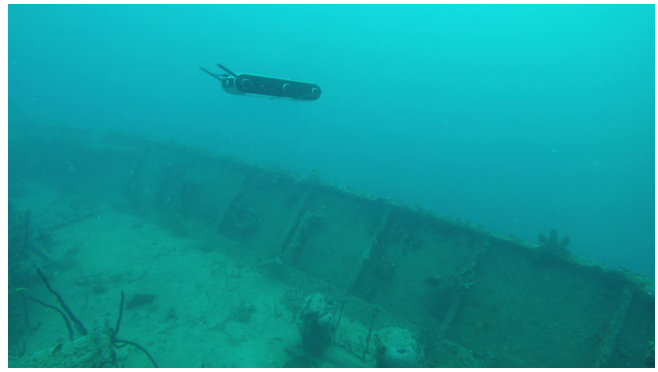


Fig. 1: Aqua2 AUV navigating over the Stavronikita shipwreck, Barbados.

- **RQ4:** How to efficiently and robustly use limited resource communication channels to share information between a team of robots and between robots and operator?

Utilizing an Autonomous Underwater Vehicle (AUV) to map an underwater structure presents the following dilemma. On one hand, if the AUV is close enough to map the observed details accurately, it is missing the big picture, *i.e.*, in what direction is the most unknown part, and also, incremental SLAM generates drift. On the other hand, if the AUV is far enough to sense a large portion of the structure, details are missing due to the underwater sensing conditions. The proposed approach utilizes AUVs in two distinct roles: *proximal observers* are AUVs which operate near the underwater structure (see Fig. 1); *distal observers* are AUVs which operate at a distance, keeping a large portion of the structure together with the proximal observer(s) in their field of view; see Fig. 3a. Central to the proposed approach is the ability of the distal observers to detect and localize the proximal observers via a cooperative localization framework [1].

II. Related Work

Wreck mapping has been studied using a variety of techniques all around the world. Photogrammetry of manually obtained images resulted in mosaics in Demesticha *et al.* [3], or from an ROV, see Nornes *et al.* [4]. While the Arrows EU project provides an overview of robotic technology use [5]. Menna *et al.* [6] provide a comprehensive review of techniques used. Mapping projects extend from Italy [7], Spain [8], Canada [9], Qatar [10], up to the arctic [11]. With the most famous wreck explorations of the Titanic [12] and the Antikythera [13] shipwrecks.

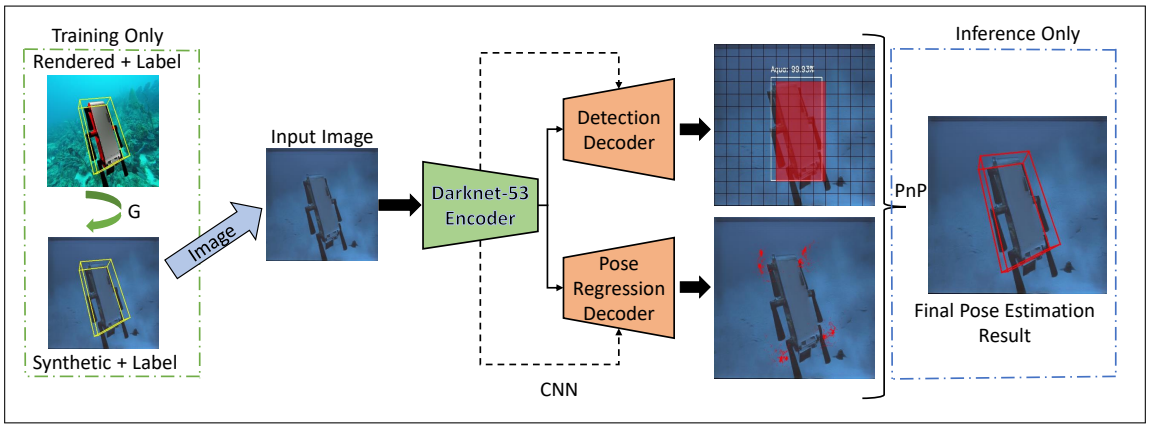


Fig. 2: The Deep Underwater Relative localization framework from Joshi *et al.* [2].

Kurazume *et al.* [14], [15] first introduced the idea of localizing two robots based on mutual observations, termed it as cooperative positioning system, although in the literature it is most often termed as Cooperative Localization (CL) [1]. Dieudonné *et al.* [16] proved that CL with an arbitrary set of sensors is NP-hard, while Roumeliotis and Rekleitis [17] provided a theoretical analysis of factors that affect error growth in a CL system. Further studies have examined the performance [18], the sensing modalities [19], and consistency [20], proposing decentralized solutions [21] and integration with inertial sensors [22].

Active sensing methods for dense 3-D modeling have been reported in the literature, but none of them benefit from two or more tightly collaborating robots, such as the ones proposed here. Several approaches require enumerating and simulating sensing from discrete pose hypotheses in 6-D [23]–[27] incurring high computational cost; others are limited to 2-D slices of constant height or depth [28], [29] or require coarse initial models [30], [31]; while most operate on occupancy grids drastically reducing the resolution of the reconstructed surfaces. Exploration strategies [32], [33] that guide a robot towards frontier voxels without requiring sampling in pose space are closely related to our work, but they are limited to single robot, rely on a prior map, and do not model uncertainty. Multirobot 3-D reconstruction methods have been presented [34], but robots are assigned areas to map independently without tight cooperation.

III. The Proposed System

A. Overview

The main idea of our approach is to have a team of co-robots collaborating with a human operator. There are two types of robots: proximal observers, which will operate close to the structure in order to produce an accurate map, and distal observers, which will be at distance maintaining the global picture of the structure and the pose of the proximal observer. Currently Cooperative Localization (see III-B); Motion strategies for navigating near obstacles (see III-C); Underwater State Estimation [35]; and photorealistic reconstruction (see III-E) have been explored, while, the remaining components of the proposed approach are still under investigation.

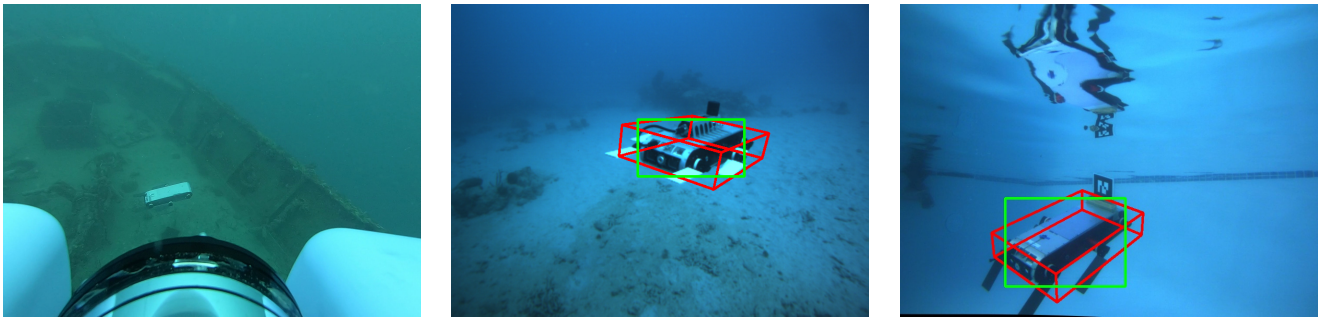
B. Cooperative Localization

First, we need to address the problem of cooperative localization: its solution provides the relative pose between Distal and Proximal Observers. We recently proposed a deep learning based framework for detecting the relative pose between the two robots [2] from a single image.

To overcome the challenge of obtaining training data with accurate 6D poses, we utilize the Unreal Engine 4 to generate a rendered dataset by projecting the 3D model of the robot swimming over underwater images. However, these rendered images differ from real underwater images due to color loss and poor visibility quality in the underwater domain. Thus, we employ CycleGAN [36], an image-to-image translation network, to bridge the gap between rendered and real images, producing a synthetic dataset containing images that are closer in appearance to real underwater images. In effect, a CNN is trained on the synthetic dataset and then tested on real underwater images, as shown in Fig. 2.

We modified YOLOv3 [37] by adding a pose regression decoder, whereas the object detection decoder and backbone encoder remain unchanged. The pose regression decoder predicts projected keypoint locations of 8 corners corresponding to the 3D the model of the robot in an image, along with their confidence scores. An image is divided into grids, and each image patch corresponding to grid votes for the object detection box. Instead of using all the grid cells for 2D keypoint prediction, we select grid cells that fall inside the object bounding box, thus focusing on regions that belong to the robot. From these candidate keypoint predictions, the most dependable 2D keypoint candidates for each 3D keypoint are selected to yield a set of 2D-to-3D correspondences. More specifically, we select the 12 most confident keypoint predictions for each corner of the robot's 3D model. These selected 2D keypoints are used in the RANSAC-based PnP [38] algorithm to obtain a robust 6D pose estimate.

The proposed framework has been tested in different environments – pool, ocean – and different cameras including an Aqua2 robot and several GoPro cameras, demonstrating



(a) (b) (c)

Fig. 3: (a) BlueROV observing an Aqua2 robot navigating over the Stavronikita shipwreck, Barbados (b) Aqua2 AUV detected while navigating over coral reef, Barbados (c) Aqua2 AUV detect in pool experiments.

its robustness with respect to variations in underwater environment, camera intrinsics, and color calibration, Fig. 3. Moreover, we demonstrate better accuracy in terms of translation and orientation error on pool dataset compared to the state-of-the-art methods [39].

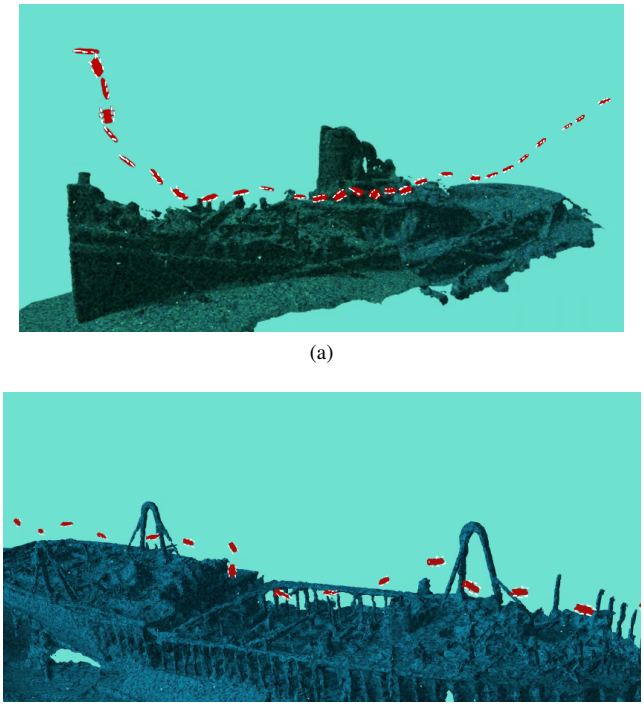
Our pose estimation framework might fail to produce a correct pose estimate sometimes when the bounding box detection is not accurate. Hence, relative pose of the robot from the observer can be determined accurately for each frame apart from failure cases. We are currently fusing other sensor data, e.g., IMU and pressure-based depth sensor to detect discrepancies in estimated pose across multiple frames and further improve the pose estimate.

C. Proximal Observer: Active Exploration

The main goal of the proximal observer is to operate safely in proximity to the target structure and collect close observations. In the absence of prior information, the proximal observers will greedily explore the underwater structure. We plan to investigate two different directions:

1) Learning-based Exploration: A Deep Learning framework navigating the underwater robot to collect data in proximity to the structure is explored. Recently, we proposed a CNN framework [41] which is trained based on the way human divers collect data. In this framework, human operators annotated the training dataset, by marking the best direction of motion and orientation to guide the autonomous underwater robot for shipwreck coverage. Although due to the nature of the technique some limited unpredictable behavior is expected, we speculate that such techniques will avoid any issues that might arise from state estimation uncertainty.

2) Optimization-based Exploration: A path-optimization-based technique is considered to navigate the robot near the structure maximizing visibility and information gain. Past work introduced AquaNav [42], a robust underwater navigation framework utilizing a path optimization planner that allows extensions in the form of cost functions. We will investigate the addition of novel cost functions maximizing visibility of target structures, determined on-line by a CNN similarly to [43]. Such a technique would require sufficient localization capabilities that could utilize a robust SLAM method such as SVIN2 [35], or can be provided locally by dead reckoning. We expect that although robust localization is needed, unlike the first approach, such a framework could ensure a predictable and safe performance with completeness, if paired with a high-level global planner. Preliminary results in simulation already show strong potential for this approach, as can be observed in Fig. 4.



(a)

(b)

Fig. 4: Example of the enhanced AquaNav [40] pipeline navigating the Aqua2 safely in proximity to two initially unknown simulated shipwrecks increasing visibility; a desired behavior for the proximal observers. The Aqua2 robot is highlighted with red.

D. Distal Observer: Active Positioning

The main objective of the distal observer is to track the proximal observers while simultaneously keeping in frame from a larger distance a bigger portion of the target structure. To achieve such behavior, the distal observers have to be

capable to visually recognize and localize the proximal observers, predict their motions, visually recognize relevant segments of the target structure, and plan motions that satisfy the desired tracking objectives.

1) Proximal Observers Tracking: Prior work has already successfully addressed this issue for the Aqua2 underwater robot [2]. The method introduced a DNN technique trained on realistic simulated data that enabled an underwater robot using cameras to extract the relative position and orientation of another visible robot.

2) Structure Segment Recognition: Past research has already produced a CNN-based recognition method to corals [43]. We plan to investigate similar ways to recognise human-made underwater structures, such as shipwrecks, by employing DNNs trained on annotated datasets.

3) Motion Prediction: The distal observer has to produce motions that successfully track the proximal observer, thus the future positions of the tracked robot should be provided or extracted. This problem could be resolved either by utilizing a motion predictor combined with fast replanning, or, when allowed by the underwater communication conditions, the plans of the proximal observers could be shared directly to inform motion planning.

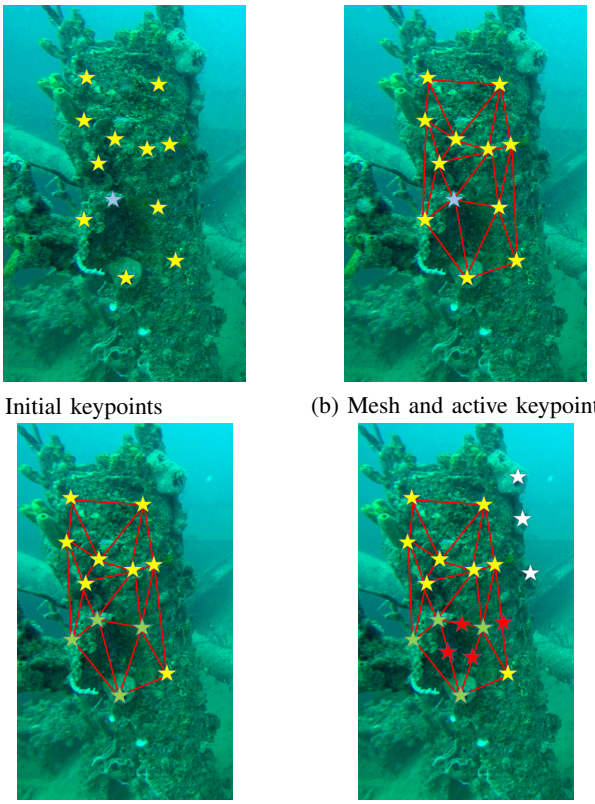
4) Navigation and Motion Planning: Given the map, areas of interest of the target underwater structure represented as segments, and a predicted trajectory of the proximal observer, the distal observer has to produce a sequence of motions that satisfy our objectives. Having already utilized AquaNav [42] – a powerful autonomous underwater navigation package for safe operation in proximity – similarly to the planning problem formulation of the proximal observers, we will focus on enhancing the optimization process with novel cost functions encouraging the visibility of the future positions of the tracked robot and the target structure at a desired distance.

E. Photo-realistic reconstruction

Our goal in this project is to obtain photorealistic 3D models useful to the robots and their operators. Even though 3D point clouds may be sufficient for obstacle avoidance, they are not well suited for tasks such as visibility estimation or detection of geometric features such as openings and locations of high curvature. They are also a poor form of visualization for users to appreciate and understand underwater structures. We therefore aim to represent surface models using triangular meshes.

We propose a representation comprising a set of *3-D keypoints* detected on the left (reference) image of the first stereo pair and reconstructed in 3-D via triangulation after they are matched on the right image. The keypoints are connected on the images to form triangles which are then lifted to 3-D after their vertices have been reconstructed. The proposed processing steps are as follows – see also Fig. 5, which shows a sketch of the approach.

1) Extract keypoints from the first image of the object and reconstruct them in 3-D. Harris corners or any other feature detection mechanism can be used here.



(a) Initial keypoints (b) Mesh and active keypoint

(c) Explained keypoints (d) Generation of new keypoints

Fig. 5: Illustration of surface reconstruction. (a) Keypoints are detected in the reference image of the first stereo pair and reconstructed in 3-D. (b) An initial mesh is generated by applying Delaunay triangulation on the detected keypoints in the image. The keypoint whose expected error is closest, but higher, than the specified tolerance (see text) is chosen as the active keypoint (in blue). The next best viewpoint is computed based on the active keypoint. (c) After acquiring additional images, the expected error of some keypoints falls below the tolerance (in green). (d) New keypoints are generated in previously unseen parts of the object (in white) and after triangle subdivision (in red).

2) Apply 2D Delaunay triangulation on the keypoints on the image to obtain a mesh.
3) Select as *active keypoint* the one with the smallest expected reconstruction error still above the tolerance.
4) Compute the Next-Best View (NBV) for the active keypoint using the approach of Freundlich et al. [44], [45].
5) Update position estimates and error covariance matrices for all visible keypoints based on new images and label as "explained" those with acceptable expected error.
6) Detect new keypoints in newly observed parts of the scene.
7) Insert as new keypoints the vertices that were generated due to triangle subdivision.
8) Attach new keypoints to the mesh and return to Step 3 unless all keypoints are explained.

As the robot approaches a surface, triangles project onto larger areas in the image plane and new keypoints may be detected in their interior. When this occurs, the planarity of

these triangles will be assessed and non-planar triangles will be subdivided to allow a tighter approximation of the surface. A classifier will be trained to make these decisions considering image appearance and reprojection errors computed by warping one image to another via the mesh. The output of this module will be a 3-D mesh comprising compact triangles, since their projections satisfy the Delaunay condition on some views, which will be texture-mapped by blending the input images with weights emphasizing fronto-parallel images observing the triangles at high resolution.

F. Underwater Information Sharing

The underwater domain is a very harsh environment for communication, which is primarily based on acoustic devices characterized by very low data rate (tens of kilobits), relatively high packet loss, and distance-dependent performance [46]–[48]. This requires decisions over time on what information to share between robots and human operators.

We propose a hybrid representation comprising, on one hand, a set of 3-D keypoints and triangles, and, on the other hand, an octree-based occupancy grid [49]. The keypoints and triangles capture the details of the surface with the highest precision that has been achieved so far, while the octree naturally allows the hierarchical visualization and transmission of the current state of the estimated model.

Initially, a coarse level of the octree with a few bits per voxel, indicating whether it is free, occupied or unknown, will be transmitted. Information about voxels at finer levels will be transmitted subsequently, if their parent voxels have been subdivided. Additional information to give the operator a better view of the current surface estimate and the progress of the modeling process, such as 3-D reconstruction uncertainty within each voxel, will also be transmitted progressively. Finally, as the octree reaches a steady state and voxels at the finest resolution have been assigned stable labels, the list of keypoints and triangles will be transmitted. Note that each voxel is marked with the timestamp of the last transmission and last update, so that they can be possibly added back to the transmission list: a metric that identifies the change will determine the importance of retransmission, that can be used by the optimization method proposed in the following section.

We propose to optimize data transmission by performing a joint cross-layer optimization of the data representation at the application layer, and the operations of the communication network.

IV. Discussion

Our research on underwater structure mapping and inspection requires advanced state estimation, active sensing, and communications and will enable more robust situational awareness, autonomy, and robot coordination. The main contributions include the use of co-robots in tight cooperation that increases robustness and efficiency of the system inspecting an underwater structure by assigning different roles to the robots; an explicit cross-layer optimization for communication that improves the communication channel utilization under very limited bandwidth; a sensor data fusion approach that can be run in real time on-board the robot. These contributions will be deployed and demonstrated on inexpensive vehicles, such as the BlueROV2. The software developed to enable these operations will be open-sourced to facilitate research in the challenging field of marine robotics.

Taking a wider view, we aspire that our work will reduce the cost and risk of having divers perform exploration and mapping tasks, possibly under adverse conditions. We explicitly plan to rely on low-cost platforms to reduce the cost of underwater robotic operations. Reducing the cost and risk would enable broader mapping efforts leading to the discovery of historically important, valuable or otherwise interesting artifacts. Human divers will be engaged only after the robots have found promising evidence. Besides government entities and non-profit organizations, such as museums, the proposed research will benefit archaeologists who can monitor whether a recovery is needed for preservation as mandated by the UNESCO Convention on the Protection of the Underwater Cultural Heritage [50].

References

[1] I. Rekleitis, G. Dudek, and E. E. Milios, "On multiagent exploration," in *Vision Interface*, Jun. 1998, pp. 455–461.

[2] B. Joshi, M. Modasshir, T. Manderson, H. Damron, M. Xanthidis, A. Q. Li, I. Rekleitis, and G. Dudek, "Deepurl: Deep pose estimation framework for underwater relative localization," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 1777–1784.

[3] S. Demesticha, D. Skarlatos, and A. Neophytou, "The 4th-century bc shipwreck at mazotos, cyprus: New techniques and methodologies in the 3d mapping of shipwreck excavations," *Journal of Field Archaeology*, vol. 39, no. 2, pp. 134–150, 2014.

[4] S. M. Nornes, M. Ludvigsen, Ø. Ødegard, and A. J. SØrensen, "Underwater photogrammetric mapping of an intact standing steel wreck with rov," *IFAC-PapersOnLine*, vol. 48, no. 2, pp. 206–211, 2015.

[5] B. Allotta, R. Costanzi, A. Ridolfi, C. Colombo, F. Bellavia, M. Fanfani, F. Pazzaglia, O. Salvetti, D. Moroni, M. A. Pascali, *et al.*, "The arrows project: Adapting and developing robotics technologies for underwater archaeology," *IFAC-PapersOnLine*, vol. 48, no. 2, pp. 194–199, 2015.

[6] F. Menna, P. Agrafiotis, and A. Georgopoulos, "State of the art and applications in archaeological underwater 3d recording and mapping," *Journal of Cultural Heritage*, vol. 33, pp. 231–248, 2018.

[7] C. Balletti, C. Beltrame, E. Costa, F. Guerra, and P. Vernier, "Underwater photogrammetry and 3d reconstruction of marble cargos shipwreck.," *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*, 2015.

[8] N. Palomeras, N. Hurtós, M. Carreras, and P. Ridao, "Autonomous mapping of underwater 3-d structures: From view planning to execution," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1965–1971, 2018.

[9] D. F. Coleman, "Underwater archaeology in thunder bay national marine sanctuary, lake huron—preliminary results from a shipwreck mapping survey," *Marine Technology Society Journal*, vol. 36, no. 3, pp. 33–44, 2002.

[10] K. Bain, R. Cuttler, E. Kieran, and S. Al Naimi, "Wreck and ruin—mapping the shipwreck resource of qatar," in *Proceedings of the Seminar for Arabian Studies*, JSTOR, 2014, pp. 35–41.

[11] A. A. Mogstad, Ø. Ødegård, S. M. Nornes, M. Ludvigsen, G. Johnsen, A. J. Sørensen, and J. Berge, "Mapping the historical shipwreck figaro in the high arctic using underwater sensor-carrying robots," *Remote Sensing*, vol. 12, no. 6, p. 997, 2020.

[12] R. M. Eustice, H. Singh, J. J. Leonard, and M. R. Walter, "Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters," *The Int. Journal of Robotics Research*, vol. 25, no. 12, pp. 1223–1242, 2006.

[13] S. B. Williams, O. Pizarro, and B. Foley, "Return to antikythera: Multi-session slam based auv mapping of a first century bc wreck site," in *Field and Service Robotics*, Springer, 2016, pp. 45–59.

[14] R. Kurazume, S. Nagata, and S. Hirose, "Cooperative positioning with multiple robots," in *Proc. ICRA*, vol. 2, 1994, pp. 1250–1257.

[15] R. Kurazume and S. Hirose, "An experimental study of a cooperative positioning system," *Auton. Robot.*, vol. 8, pp. 43–52, 2000.

[16] Y. Dieudonné, O. Labbani-Igbida, and F. Petit, "Deterministic robot-network localization is hard," *IEEE Trans. Robot.*, vol. 26, no. 2, pp. 331–339, 2010.

[17] S. I. Roumeliotis and I. Rekleitis, "Propagation of uncertainty in cooperative multirobot localization: Analysis and experimental results," *Autonomous Robots*, vol. 17, no. 1, pp. 41–54, Jul. 2004.

[18] A. Mourikis and S. Roumeliotis, "Performance analysis of multirobot cooperative localization," *IEEE Trans. Robot.*, vol. 22, no. 4, pp. 666–681, 2006.

[19] X. Zhou and S. Roumeliotis, "Determining 3D relative transformations for any combination of range and bearing measurements," *IEEE Trans. Robot.*, vol. 29, no. 2, pp. 458–474, 2012.

[20] G. P. Huang, N. Trawny, A. I. Mourikis, and S. I. Roumeliotis, "On the consistency of multi-robot cooperative localization," in *Proc. RSS*, 2009, pp. 65–72.

[21] K. Y. K. Leung, T. D. Barfoot, and H. H. T. Liu, "Decentralized localization of sparsely-communicating robot networks: A centralized-equivalent approach," *IEEE Trans. Robot.*, vol. 26, pp. 62–77, 2010.

[22] A. Martinelli and A. Renzaglia, "Cooperative visual-inertial sensor fusion: Fundamental equations," in *Proc. Int. Symp. Multi-Robot and Multi-Agent Systems (MRS)*, 2017.

[23] S. Wenhardt, B. Deutsch, E. Angelopoulou, and H. Niemann, "Active Visual Object Reconstruction using D-, E-, and T-Optimal Next Best Views," in *Proc. CVPR*, 2007.

[24] E. Dunn, J. van den Berg, and J.-M. Frahm, "Developing visual sensing strategies through next best view planning," in *Proc. IROS*, 2009, pp. 4001–4008.

[25] C. Potthast and G. S. Sukhatme, "A probabilistic framework for next best view estimation in a cluttered environment," *Journal of Visual Communication and Image Representation*, vol. 25, no. 1, pp. 148–164, 2014.

[26] A. Kim and R. M. Eustice, "Active visual SLAM for robotic area coverage: Theory and experiment," *The International Journal of Robotics Research*, vol. 34, no. 4-5, pp. 457–475, 2015.

[27] J. Daudelin and M. Campbell, "An adaptable, probabilistic, next-best view algorithm for reconstruction of unknown 3-d objects," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1540–1547, 2017.

[28] F. Fraundorfer, L. Heng, D. Honegger, G. H. Lee, L. Meier, P. Tanskanen, and M. Pollefeys, "Vision-based autonomous mapping and exploration using a quadrotor mav," in *Proc. IROS*, IEEE, 2012, pp. 4557–4564.

[29] E. Vidal, J. D. Hernández, K. Istenič, and M. Carreras, "Online view planning for inspecting unexplored underwater structures," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1436–1443, 2017.

[30] B. Hepp, M. Nießner, and O. Hilliges, "Plan3d: Viewpoint and trajectory optimization for aerial multi-view stereo reconstruction," *ACM Transactions on Graphics (TOG)*, vol. 38, no. 1, p. 4, 2018.

[31] N. Smith, N. Moehrle, M. Goesele, and W. Heidrich, "Aerial path planning for urban scene reconstruction: A continuous optimization method and benchmark," in *SIGGRAPH Asia*, 2018, p. 183.

[32] R. Shade and P. Newman, "Choosing where to go: Complete 3d exploration with stereo," in *Proc. ICRA*, 2011, pp. 2806–2811.

[33] G. A. Hollinger, B. Englot, F. S. Hover, U. Mitra, and G. S. Sukhatme, "Active planning for underwater inspection and the benefit of adaptivity," *The International Journal of Robotics Research*, vol. 32, no. 1, pp. 3–18, 2013.

[34] S. Golodetz, T. Cavallari, N. A. Lord, V. A. Prisacariu, D. W. Murray, and P. H. Torr, "Collaborative large-scale dense 3d reconstruction with online inter-agent pose optimisation," *IEEE transactions on Visualization and Computer Graphics*, vol. 24, no. 11, pp. 2895–2905, 2018.

[35] S. Rahman, A. Q. Li, and I. Rekleitis, "Svin2: An underwater slam system using sonar, visual, inertial, and depth sensor," in *2019 IEEE/RSJ International*

Conference on Intelligent Robots and Systems (IROS), IEEE, 2019, pp. 1861–1868.

[36] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *International Conference on Computer Vision*, 2017.

[37] J. Redmon and A. Farhadi, *YOLOv3: An incremental improvement*, 2018.

[38] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An accurate O(n) solution to the PnP problem," *Int. Journal of Comp. Vision*, 2009.

[39] K. Koreitem, J. Li, I. Karp, T. Manderson, F. Shkurti, and G. Dudek, "Synthetically trained 3d visual tracker of underwater vehicles," in *MTS/IEEE OCEANS*, 2018.

[40] M. Xanthidis, N. Karapetyan, H. Damron, S. Rahman, J. Johnson, A. O'Connell, J. O'Kane, and I. Rekleitis, "Navigation in the presence of obstacles for an agile autonomous underwater vehicle," in *IEEE International Conference on Robotics and Automation*, 2020, pp. 892–899.

[41] N. Karapetyan, J. V. Johnson, and I. Rekleitis, "Coverage path planning for mapping of underwater structures," in *Global Oceans 2020: Singapore–US Gulf Coast*, IEEE, 2020, pp. 1–6.

[42] M. Xanthidis, N. Karapetyan, H. Damron, S. Rahman, J. Johnson, A. O'Connell, J. M. O'Kane, and I. Rekleitis, "Navigation in the presence of obstacles for an agile autonomous underwater vehicle," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 892–899.

[43] M. Modasshir and I. Rekleitis, "Enhancing coral reef monitoring utilizing a deep semi-supervised learning approach," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 2020, pp. 1874–1880.

[44] C. Freundlich, P. Mordohai, and M. Zavlanos, "A hybrid control approach to the next-best-view problem using stereo vision," in *ICRA*, 2013.

[45] C. Freundlich, Y. Zhang, A. Z. Zhu, P. Mordohai, and M. M. Zavlanos, "Controlling a robotic stereo camera under image quantization noise," *IJRR*, 2017.

[46] D. Pompili and I. F. Akyildiz, "Overview of networking protocols for underwater wireless communications," *IEEE Communications Magazine*, vol. 47, no. 1, pp. 97–102, 2009.

[47] A. Song, M. Stojanovic, and M. Chitre, "Editorial underwater acoustic communications: Where we stand and what is next?" *IEEE Journal of Oceanic Engineering*, vol. 44, no. 1, pp. 1–6, 2019.

[48] N. Nowsheen, C. Benson, and M. Frater, "A high data-rate, software-defined underwater acoustic modem," in *Proc. MTS/IEEE OCEANS – Seattle, WA*, 2010, pp. 1–5.

[49] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Autonomous Robots*, 2013.

[50] *The world's underwater cultural heritage*, http://www.unesco.org/new/en/culture/themes/underwater-cultural-heritage/underwater-cultural-heritage/, Accessed 01/20/2020 2020.