

Human Diver-Inspired Visual Navigation: Towards Coverage Path Planning of Shipwrecks

AUTHORS

Nare Karapetyan

James V. Johnson

Ioannis Rekleitis

Computer Science and Engineering
Department, University of
South Carolina

Introduction

Mapping and inspection of underwater structures is an important task found in environmental studies, marine archaeology, resource acquisition, and infrastructure maintenance, often performed by underwater vehicles. In order to inspect a structure, an underwater vehicle has to move over most (ideally all) of the region of interest, a task often termed as coverage. The coverage problem in an underwater environment inherits all the challenges faced by planning for an aerial robot in three dimensions related to the increased dimensionality of the planning space. The coverage path planning (CPP) problem's complexity exponentially increases when moving from two to three dimensions. In addition, the underwater environment presents novel challenges from both the coverage and the navigation perspective. Underwater vision presents unique challenges such as hazing, color attenuation (Roznere & Quattrini Li, 2019; Skaff et al., 2008), and lack of good features (Quattrini Li et al., 2016b; Shkurti et al., 2011). Dynamics of the water

ABSTRACT

This work proposes vision-only navigation strategies for an autonomous underwater robot. This approach is a step towards solving the coverage path planning problem in a 3-D environment for surveying underwater structures. Given the challenging conditions of the underwater domain, it is very complicated to obtain accurate state estimates reliably. Consequently, it is a great challenge to extend known path planning or coverage techniques developed for aerial or ground robot controls. In this work, we are investigating a navigation strategy utilizing only vision to assist in covering a complex underwater structure. We propose to use a navigation strategy akin to what a human diver will execute when circumnavigating around a region of interest, in particular when collecting data from a shipwreck. The focus of this article is a step towards enabling the autonomous operation of lightweight robots near underwater wrecks in order to collect data for creating photo-realistic maps and volumetric 3-D models while at the same time avoiding collisions. The proposed method uses convolutional neural networks to learn the control commands based on the visual input. We have demonstrated the feasibility of using a system based only on vision to learn specific strategies of navigation with 80% accuracy on the prediction of control command changes. Experimental results and a detailed overview of the proposed method are discussed.

Keywords: underwater navigation, path planning, coverage, autonomous system

(Moulton et al., 2018) and visibility constraints contribute to instability, drifting, and error in localization of an autonomous underwater vehicle (AUV). For details on the challenges of underwater sensing, please refer to the comparison studies in Joshi et al. (2019) and Quattrini Li et al. (2016a).

The focus of this work is visual mapping of a shipwreck. Historical shipwrecks tell an important part of history and at the same time have a special allure for most humans, as exemplified by the plethora of movies and artworks of the *Titanic*, see, for example, the work of Eustice et al. (2006) for the visual mapping of the

Titanic. Shipwrecks are also one of the top scuba diving attractions all over the world. The historical shipwrecks are deteriorating due to warm, salt water; human interference; and, in some places, extreme weather (frequent tropical storms). Reconstructing accurate models of these sites will be of high importance not only for the historical study of the shipwrecks but also for monitoring subsequent deterioration (National Park Service, U.S. Department of the Interior, 2020; National Oceanic and Atmospheric Administration [NOAA] National Marine Sanctuaries, 2020); see Figure 1(a) for the different floors of a shipwreck

FIGURE 1

(a) The different levels of the Stavronikita shipwreck, Barbados, after a partial collapse. (b) Diver collecting visual data for training at the Pamir shipwreck, Barbados.



exposed after a partial collapse. Currently, limited mapping efforts are performed by divers who take measurements manually using a grid and measuring tape or using handheld sensors (Henderson et al., 2013)—a slow and sometimes dangerous task; see Figure 1(b) for a diver collecting data manually to be used in training the navigation model. While acoustic sensing sound navigation and ranging (SONAR) is common, the resulting maps do not contain the details vision can provide. Vision has been utilized successfully to map underwater structures (Hogue et al., 2007) or even underwater caves (Rahman et al., 2019).

This method presents a methodology for training a learning system to guide an underwater vehicle covering a shipwreck using only vision. The trained system's aim is to perform coverage in a completely unknown space by learning a set of motion strategies based on the previous visual observations of different environments.

In order to perform coverage or reconstruction of 3-D structures traditionally, reliable state estimation is required (Bircher et al., 2018). However, even with ideal state estimation, performing a coverage in an unknown underwater environment exposes new

and very different challenges. The aim of this work is to answer the question: is it possible to achieve autonomous behavior that will provide meaningful coverage decisions relying only on vision? In contrast to the works presented in the literature on 3-D coverage (Galceran & Carreras, 2013), this work is solely based on vision and does not rely on state estimation nor requires a map of the environment for navigation.

The main contribution of this article is a deep learning framework for learning the motions for navigating around a shipwreck. The proposed system is based on data collected by human operators who are trained to guide an autonomous system in similar behavior around different shipwrecks using only vision. Results demonstrate the accuracy of the learned system in validation datasets.

The article is structured as follows. First, we will discuss the related work and highlight challenges that each approach is facing. In the following section, the formal problem definition along with the proposed method are outlined. In the next section, the experimental results from real and simulated data are reported. The last section concludes with discussions on lessons learned and how this work will progress.

Related Work

The 3-D coverage problem is quite challenging as there are exponentially many solutions to choose from. Some approaches have been proposed and shown to have feasible solutions both with single and multiple robots using octomaps (Dornhege & Kleiner, 2013; Dornhege et al., 2016). In addition, some coverage solutions in three dimensions have been presented for ensuring the even distribution of spray paint in the automotive industry (Atkar et al., 2005). Another work has addressed the coverage of an unknown environment using a frontier-based approach (Dornhege et al., 2013).

Cheng et al. (2008) addressed the problem by representing an area through well-defined 2.5-D features, thus reducing the complexity of structure dependency of the 3-D coverage. In this work, the coverage problem is specifically designed for an aerial vehicle with a conical field of view that can rotate around a fixed point with 3 degrees of freedom, thus somehow limiting its application to other domains.

Similar to Palomeras et al. (2019), Bircher et al. (2018) combine the problem of covering an unknown environment with the given structure by sampling random next-best views around the target structure. It is an extension of their previous work (Bircher et al., 2016) that builds a tree of next-best views and selects the best branch using the size of unexplored area covered as the metric. As most next-best view-based approaches, these works are also subject to the sensor data quality, system localization, and vehicle dynamics.

The coverage problem has been studied also for underwater environments and has valuable environmental and archaeological importance. A

number of works presented seabed and underwater CPP methods with autonomous surface vehicles and AUVs (Galceran et al., 2014; Hollinger et al., 2013). Behavior-based control of an underwater vehicle for coral-reef inspection was proposed by Rosenblatt et al. (2002). The behavior selection is implemented using both fuzzy logic and utility fusion. The behaviors ensure collision avoidance, proper distance from the reef, and rope-following or target-following actions, but a similar approach has not been developed for surveying more complicated environments.

To ensure complete coverage without overlaps, Galceran and Carreras (2012) suggest segmenting the environment based on similar depths. Each of these segments then is considered as an individual planning problem. The proposed algorithm extends cellular decomposition (Choset, 2000) performing 2.5-D coverage by traveling on constant depth from the surface. This work has been extended to also take into account the state estimation uncertainty and perform replanning as needed (Galceran et al., 2015). The above works rely on knowing the approximate depth to split the area, and it only views the environment from above.

When the environment is unknown, Vidal et al. (2017) proposes a next-best view approach, but it is constrained by the certainty of the state estimation. To overcome the complexity of 3-D exploration, simplified formulation of the problem is considered, such as 2-D mapping of an underwater structure (Vidal et al., 2019). In this work, authors use a view planner and frontier-based strategies. The environment is represented as a quadtree occupancy map, and it is also used to generate viewpoints for the exploration.

It is worth noting that quadtree representations are computationally becoming more expensive when representing large-scale complex structures and require accurate localization.

The 3-D coverage also has been shown to have a wide range of applications for mapping historical artifacts and structures in underwater environments (Gracias et al., 2013). Most of the presented works either assume reliable localization or some type of prior information about the environment. Nevertheless, when operating in an underwater environment, reliable state estimation is very challenging. In contrast, a wide range of vision-based navigation methods have been used in literature for aerial robots and ground vehicles for either goal-oriented path planning or coverage. Such examples are work by Smolyanskiy et al. (2017), where they collect real and simulation data to train a drone to navigate over a trail.

Motivated by similar works, Manderson et al. (2018) have built a 3-D navigation deep learning framework for underwater systems by proposing a vision-based navigation in an unknown environment for coral reef exploration. More recent work (Manderson et al., 2020) has incorporated path planning in conjunction with obstacle avoidance and a bias towards areas with corals. Their approach is trained for navigating over the coral reef and not navigating around an underwater structure.

More generally, in computer vision literature, methods grouped under the terms shape-from-motion and next-best-view have been utilized for the visual mapping of an object or a structure. By definition, shape-from-motion techniques always estimate the location of the camera together

with the shape of the mapped structure, thus requiring accurate state estimation of the mapping system—a difficult task as shown in Joshi et al. (2019). Furthermore, all next-best-view techniques require the knowledge of where the camera is and an ability to navigate to the next viewpoint. In contrast, the proposed system navigates around based only on the visual input and does not require a representation of the vehicle's location.

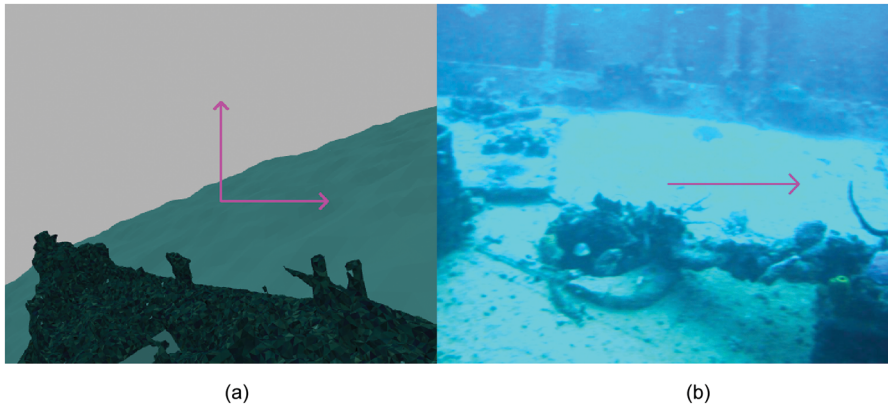
Proposed Approach

In general terms, the CPP problem is concerned with finding a trajectory that moves an end effector/sensor of the robot over each available point of the environment (area or volume) while avoiding obstacles. For the purposes of this work, the robot is autonomous, operates without a tether, and has 6 degrees of freedom from which controllable are the three orientations: yaw (rotation around the z-axis), pitch (rotation around the y-axis), and roll (rotation around the x-axis) in a robot-centric frame of reference, up and down (depth change) and forward motions. There is no prior information about the environment, and it is deployed in a 3-D-bounded area of interest $\varepsilon \subset \mathbb{R}^3$. The robot follows a path π that will result in the acquisition of a sequence of camera frames $V = \{f_1, f_2, \dots, f_n\}$, where f_i is the i^{th} image frame. The objective of the CPP problem is to ensure that π is obstacle free and that the 3-D reconstruction resulting from V is covering the entire surface of the ε object.

The objective of this work is to build a system that will use only the current frame V to guide the robot and build a path π , in such a way that it will imitate the behaviors of

FIGURE 2

The labeled data, showing the desired change in orientation along the yaw and pitch angles. (a) The gazebo simulation, presenting changes in both the yaw and pitch angles, and (b) the underwater video data, where there is a change only along the yaw angle.



a human diver when surveying a shipwreck. As such, the main strategies, which will be the building blocks for CPP, considered within this work are

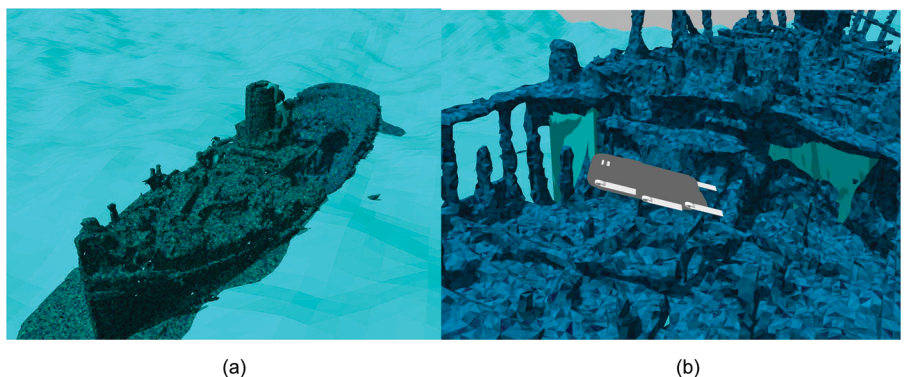
1. keep the shipwreck in the field of view,
2. follow the shipwreck's side,
3. turn around the bow and stern of the ship, and
4. circumnavigate the mast.

To create a vision-based navigation system, we need to have a large dataset of different shipwrecks. We begin with 3-D meshes of shipwrecks—the data consist of Gazebo (Koenig & Howard, 2004) models of shipwrecks provided by the NOAA (<https://nauticalcharts.noaa.gov/data/wrecks-and-obstructions.html>). In addition, we have generated test data from the coverage of the Stavronikita shipwreck in Barbados where we have collected different sets of images utilizing an underwater Aqua2 robot (Dudek et al., 2005), a stereo sensor (Rahman et al., 2018), and a GoPro camera. The Aqua2 vehicle is capable of autonomous operations (Sattar et al., 2008) up to a depth of 30 m.

A diver was asked to label data based on the action that they would take if they were to perform coverage around a shipwreck. The possible values that the diver selects are the directions in a 2-D image—the label windows are illustrated in Figure 2. The labeling performed is based on a diver's strategy of data collection, consistent with the abovementioned four behavioral strategies. Following the same strategies, simulation data have been collected by driving an Aqua2 robot in the Gazebo simulator and recording the yaw and pitch changes corresponding to the current camera view (Figure 3).

FIGURE 3

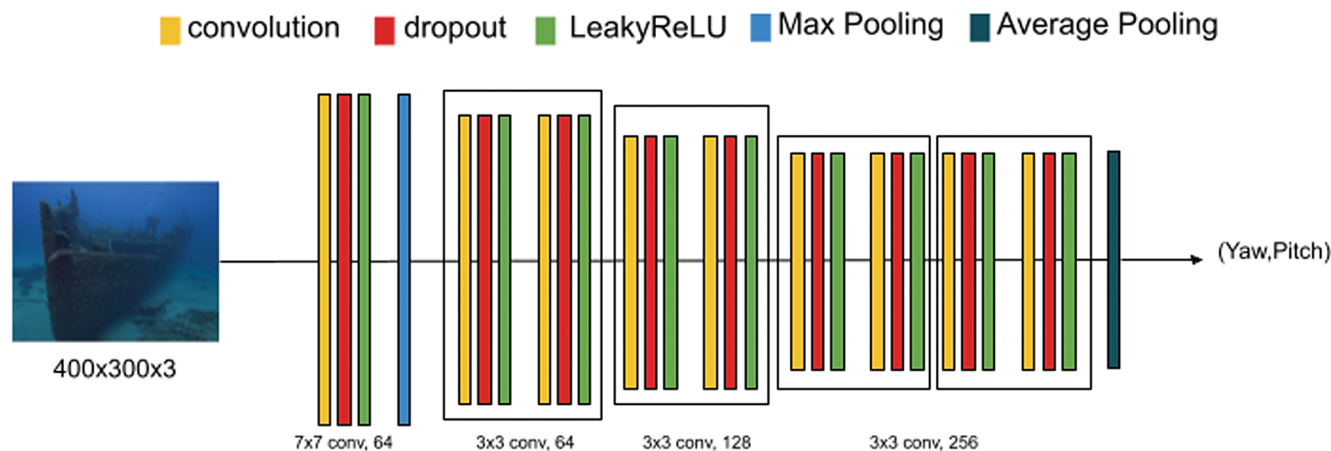
(a) Gazebo model of a shipwreck used for training. (b) The Aqua2 robot navigating over a shipwreck in Gazebo.



The labeled data are fed to an 18-layer residual network with similar architecture to the one proposed by Smolyanskiy et al. (2017) and adopted later by Manderson et al. (2018) (see Figure 4). The network is a variation of standard ResNet-18 (He et al., 2016). The first convolution layer has 7×7 , and the rest are 3×3 —all are downsized using strides of 2. After each convolutional layer, a dropout layer is added with a 0.2 rate. To avoid the dead neuron problem, we use the LeakyReLU activation function. We used the momentum gradient descent optimizer with a learning rate of 0.001 and a momentum of 0.1. Similar to Smolyanskiy et al. (2017), we also treat this as a classification problem rather than regression, to ensure we have a consistent data labeling process. As such, similar to Manderson et al. (2018), the change in yaw and pitch have seven possible values for each. Thus, our network operates on each frame of the incoming video, classifying each image into one of possible 49 classes. Each of these classes consists of a different yaw/pitch command composed of two integers in the $[-3, 3]$ range. The classification indicates the chosen action for

FIGURE 4

The overview of neural network architecture.



the robot. After training, the network is given images as input producing classifications that are directly used as input commands to the Aqua2 robot.

In training, the network used a batch size of 32 and a dropout rate of 0.2 over 2,048 epochs, a stochastic gradient descent method optimized network with a categorical cross-entropy loss function. These variables were tuned until the expected behavior could be observed. We ended up with a system with 10 layers that did not result in data overfitting and achieved increased accuracy in both the testing and validation data. The output of the system are direction

commands that are then converted to the yaw, pitch, and roll commands to control the Aqua2 robot.

Experimental Results

The experiments were performed using the Aqua2 simulator in Gazebo. It emulates the real dynamics of an underwater environment and allows control of the robot. The Aqua2 robot used in simulation uses the motion from six flippers, each independently actuated by an electric motor, to swim. It has 6 degrees of freedom, of which five are controllable. The robot's primary sensing modality is vision. It is equipped with

three iDS USB 3.0 UEye cameras: two facing forward and one in the back. The front-facing cameras are used for navigation and data collection.

The training has been performed separately on simulation data and real-world data, as well as on combined data samples. Both simulation and real-world datasets consisted of over 25,000 labeled images. The simulation dataset was collected by driving Aqua2 in Gazebo using the keyboard, which resulted in ready labeled data, whereas the data from the Stavronikita shipwreck were collected with handheld cameras and later were labeled by a human with the above-mentioned approach.

TABLE 1

Comparison of accuracy of training using simulation and real-world data, and the convergence speed (accuracy values are approximations).

Training Dataset	Simulation Based	Real-World Collected	Simulation and Real-World
Validation accuracy on simulation	82%	12%	34%*
Validation accuracy on real world	6%	80%	34%*
Test accuracy	78%	80%	47%
Convergence on epochs	500	1,000	1,000
Size of dataset	29,180	26,817	61,984

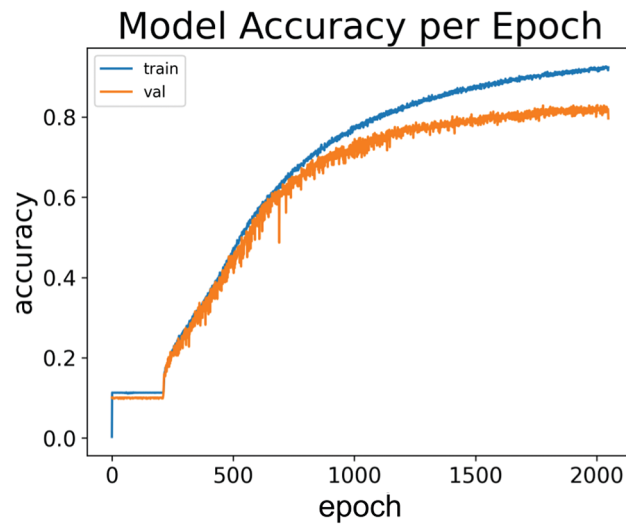
*Note that the validation is given for combined data, hence reported as the same.

The focus of this paper is on answering the question if a deep neural network can learn the motion controls of an underwater vehicle using human-driven commands informed solely by the visual input. It is worth noting that, during training with a simulator, the human user only sees the view from the robot's camera, while in the footage taken by a human diver, the diver has higher awareness as they swim around the structure. In order to answer the above question, we evaluate how well the learned system is able to predict the motions similarly to the human annotations. Thus, within this work, we are concerned with performance of the classifier and how different datasets affect the prediction accuracy. The training data, model test, and validation test data were selected with a proportion of 6:2:2.

The model accuracy results using different datasets are depicted in Table 1. We performed separate training using only simulation data and then only real-world data collected from shipwrecks, and finally, we performed training on combined data. Test accuracy is reported on the same type of data, whereas additionally, we performed validation tests using models trained in a real environment on simulation and vice versa. We also report the number of epochs after which the model is converging, for example, when more iterations of training do not result in an improvement of the model. When the Convolutional Neural Network (CNN) is trained on separate datasets, the validation accuracy converges close to 80% (see Figure 5). In contrast, when the training happens on the combined dataset, the validation accuracy drops to 34% (see Table 1). This discrepancy between

FIGURE 5

The accuracy per epoch plot of the proposed method trained only on real data.



simulation and real-world datasets is indicative of the different visual appearances between the simulator and the real world. In future work, the simulation data will be enhanced by using generative adversarial networks (GANs) (Goodfellow et al., 2020) to resemble the real-world imagery, similar to the approach used for training vision-based cooperative localization (Joshi et al., 2020). It is worth noting that, in Joshi et al. (2020), a separate

GAN was necessary for pool and ocean images.

Nevertheless, the quantitative results on separate datasets highlight the feasibility of the method. The resulting control commands have been used to navigate the Aqua2 robot in Gazebo simulation over different shipwreck models. From the earlier mentioned target behaviors, the Aqua2 in simulation performed the “keeping the shipwreck in the field of view”

FIGURE 6

Portion of a trajectory of robot in simulation produced by a prediction-based controller.



and “following the side of shipwreck” strategies (see Figure 6), which are critical for the majority of the exploration. These experiments demonstrate the capabilities of the proposed system trained only on one real shipwreck dataset using vision as the only sensor.

Conclusions

This work proposed a new approach tackling underwater navigation around a complex structure based solely on visual input. The method is based on the diver’s expertise in performing navigation collecting data around a structure. Reported preliminary results of 80% accuracy achieved from training on separate sets of simulated and real underwater data showed that this approach is feasible and capable of performing some of the expected behaviors. Thus, this is a step towards building a vision-based navigation system for coverage of shipwreck structures.

Key challenges of this work are related to the complexity of understanding the underwater scene from limited amounts of data. To address this, more restrictions and more systematic data labeling must be performed. Given the extremely challenging circumstances of data collection of underwater shipwrecks, the introduction of GAN-generated artificial training data is necessary for enabling smooth transitions between real and simulated environments.

When working in the underwater domain, we are limited by the technical constraints of the autonomous platform more than when operating with surface, ground, or aerial vehicles. The constraints of underwater vehicles include but are not limited to the battery life, computational power, and cost. In order to be able

to successfully execute the online learning-based method proposed above, for 3-D coverage, our main experimental platform, the Aqua2 robot, must be upgraded to include a Jetson TX2 Module for computations. New experiments and more data have to be collected with Aqua2 to illustrate the robustness of the proposed system. An alternative to AUV data collection is manual collection by divers, using a sensor suite (Rahman et al., 2018), where the human guides the exploration. In addition, a 3-D reconstruction of the underwater structure should be generated using the proposed method and state-of-the-art 3-D coverage methods employed to highlight the qualitative differences.

Acknowledgments

The authors would like to acknowledge the generous support of the National Science Foundation grants (NSF 2024741 and 1943205).

Corresponding Author:

Nare Karapetyan
Computer Science and
Engineering Department,
University of South Carolina
550 Assembly Street, Columbia,
SC 29201
Email: nare@email.sc.edu

References

Atkar, P.N., Greenfield, A., Conner, D.C., Choset, H., & Rizzi, A.A. 2005. Uniform coverage of automotive surface patches. *Int J Robot Res.* 24(11):883–98. <https://doi.org/10.1177/0278364905059058>.

Bircher, A., Kamel, M., Alexis, K., Oleynikova, H., & Siegwart, R. 2016. Receding horizon “next-best-view” planner for 3D exploration. In: *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1462–68.

Stockholm, Sweden: IEEE. <https://doi.org/10.1109/ICRA.2016.7487281>.

Bircher, A., Kamel, M., Alexis, K., Oleynikova, H., & Siegwart, R. 2018. Receding horizon path planning for 3D exploration and surface inspection. *Auton Robot.* 42:291–306. <https://doi.org/10.1007/s10514-016-9610-0>.

Cheng, P., Keller, J., & Kumar, V. 2008. Time-optimal UAV trajectory planning for 3D urban structure coverage. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2750–7. Nice, France: IEEE. <https://doi.org/10.1109/IROS.2008.4650988>.

Choset, H. 2000. Coverage of known spaces: The boustrophedon cellular decomposition. *Auton Robot.* 9:247–53. <https://doi.org/10.1023/A:1008958800904>.

Dornhege, C., & Kleiner, A. 2013. A frontier-void-based approach for autonomous exploration in 3D. *Adv Robotics.* 27(6):459–68. <https://doi.org/10.1080/01691864.2013.763720>.

Dornhege, C., Kleiner, A., Hertle, A., & Kolling, A. 2016. Multi-robot coverage search in three dimensions. *J Field Robot.* 4(33): 537–58. <https://doi.org/10.1002/rob.21573>.

Dornhege, C., Kleiner, A., & Kolling, A. 2013. Coverage search in 3D. In: *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pp. 1–8. Linköping, Sweden: IEEE. <https://doi.org/10.1109/SSRR.2013.6719340>.

Dudek, G., Jenkin, M., Prahacs, C., Hogue, A., Sattar, J., Giguere, P., ... Rekleitis, I. 2005. A visually guided swimming robot. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1749–54. Edmonton, AB, Canada: IEEE. <https://doi.org/10.1109/IROS.2005.1545231>.

Eustice, R.M., Singh, H., Leonard, J.J., & Walter, M.R. 2006. Visually mapping the RMS Titanic: Conservative covariance estimates for SLAM information filters. *Int J Robot Res.* 25(12):1223–42. <https://doi.org/10.1177/0278364906072512>.

Galceran, E., Campos, R., Palomeras, N., Carreras, M., & Ridao, P. 2014. Coverage

- path planning with real-time replanning for inspection of 3D underwater structures. In: IEEE International Conference on Robotics and Automation (ICRA), pp. 6586–91. Hong Kong, China: IEEE. <https://doi.org/10.1002/rob.21554>.
- Galceran, E., Campos, R., Palomeras, N., Ribas, D., Carreras, M., & Ridao, P.** 2015. Coverage path planning with real-time replanning and surface reconstruction for inspection of three-dimensional underwater structures using autonomous underwater vehicles. *J Field Robot.* 32(7):952–83. <https://doi.org/10.1002/rob.21554>.
- Galceran, E., & Carreras, M.** 2012. Efficient seabed coverage path planning for ASVs and AUVs. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 88–93. Vilamoura, Algarve, Portugal: IEEE. <https://doi.org/10.1109/IROS.2012.6385553>.
- Galceran, E., & Carreras, M.** 2013. A survey on coverage path planning for robotics. *Robot Auton Syst.* 61(12):1258–76. <https://doi.org/10.1016/j.robot.2013.09.004>.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y.** 2020. Generative adversarial networks. *Commun ACM.* 63(11):139–44. <https://doi.org/10.1145/3422622>.
- Gracias, N., Ridao, P., Garcia, R., Escartín, J., l'Hour, M., Cibeccchini, F., ... Mallios, A.** 2013. Mapping the moon: Using a lightweight AUV to survey the site of the 17th century ship 'La Lune'. In: MTS/IEEE OCEANS-Bergen, pp. 1–8. Bergen, Norway: IEEE. <https://doi.org/10.1109/OCEANS-Bergen.2013.6608142>.
- He, K., Zhang, X., Ren, S., & Sun, J.** 2016. Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–8. Las Vegas, NV: IEEE. <https://doi.org/10.1002/2014EF000239>.
- Henderson, J., Pizarro, O., Johnson-Roberson, M., & Mahon, I.** 2013. Mapping submerged archaeological sites using stereo-vision photogrammetry. *Int J Naut Archaeol.* 42(2): 243–56. <https://doi.org/10.1111/1095-9270.12016>.
- Hogue, A., German, A., & Jenkin, M.** 2007. Underwater environment reconstruction using stereo and inertial data. In: IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 2372–7. Montreal, ON, Canada: IEEE. <https://doi.org/10.1109/ICSMC.2007.4413666>.
- Hollinger, G.A., Englot, B., Hover, F.S., Mitra, U., & Sukhatme, G.S.** 2013. Active planning for underwater inspection and the benefit of adaptivity. *Int J Robot Res.* 32(1):3–18. <https://doi.org/10.1177/0278364912467485>.
- Joshi, B., Modasshir, M., Manderson, T., Damron, H., Xanthidis, M., Li, A.Q., Rekleitis, I., & Dudek, G.** 2020. Deep url: Deep pose estimation framework for underwater relative localization. In: International Conference on Intelligent Robots and Systems (IROS), pp. 1777–84. Las Vegas, NV: IEEE. <https://doi.org/10.1109/IROS45743.2020.9341201>.
- Joshi, B., Rahman, S., Kalaitzakis, M., Cain, B., Johnson, J., Xanthidis, M., ... Rekleitis, I.** 2019. Experimental comparison of open source visual-inertial-based state estimation algorithms in the underwater domain. In: International Conference on Intelligent Robots and Systems (IROS), pp. 7221–7. Macau, China: IEEE. <https://doi.org/10.1109/IROS40897.2019.8968049>.
- Koenig, N., & Howard, A.** 2004. Design and use paradigms for gazebo, an open-source multi-robot simulator. In: International Conference on Intelligent Robots and Systems (IROS), Vol. 3, pp. 2149–54. Sendai, Japan: IEEE. <https://doi.org/10.1109/IROS.2004.1389727>.
- Manderson, T., Higuera, J.C.G., Cheng, R., & Dudek, G.** 2018. Vision-based autonomous underwater swimming in dense coral for combined collision avoidance and target selection. In: International Conference on Intelligent Robots and Systems (IROS), pp. 1885–91. Madrid, Spain: IEEE. <https://doi.org/10.1109/IROS.2018.8594410>.
- Manderson, T., Higuera, J.C.G., Wapnick, S., Tremblay, J.F., Shkurti, F., Meger, D., & Dudek, G.** 2020. Vision-based goal-conditioned policies for underwater navigation in the presence of obstacles. *Robot Sci Syst XVI.* <http://roboticsproceedings.org/rss16/p048.pdf>.
- Moulton, J., Karapetyan, N., Quattrini Li, A., & Rekleitis, I.** 2018. External force field modeling for autonomous surface vehicles. In: International Symposium on Experimental Robotics (ISER), pp. 328–38. Buenos Aires, Argentina: Springer. https://doi.org/10.1007/978-3-030-33950-0_29.
- National Park Service, U.S. Department of the Interior.** 2020. Abandoned shipwreck act guidelines. <https://www.nps.gov/archeology/submerged/document.htm> (accessed 25 February 2020).
- NOAA National Marine Sanctuaries.** 2020. The world's underwater cultural heritage. <https://sanctuaries.noaa.gov/science/monitoring/mimnms.html> (accessed 25 February 2020).
- Palomeras, N., Hurtós, N., Vidal, E., & Carreras, M.** 2019. Autonomous exploration of complex underwater environments using a probabilistic next-best-view planner. *IEEE Robot Autom Lett.* 4(2):1619–25. <https://doi.org/10.1109/LRA.2019.2896759>.
- Quattrini Li, A., Coskun, A., Doherty, S.M., Ghasemlou, S., Jagtap, A.S., Modasshir, M., ... Rekleitis, I.** 2016a. Experimental comparison of open source vision based state estimation algorithms. In: International Symposium of Experimental Robotics (ISER), pp. 775–86. Tokyo, Japan: Springer. https://doi.org/10.1007/978-3-319-50115-4_67.
- Quattrini Li, A., Coskun, A., Doherty, S.M., Ghasemlou, S., Jagtap, A.S., Modasshir, M., ... Rekleitis, I.** 2016b. Vision-based shipwreck mapping: On evaluating features quality and open source state estimation packages. In: MTS/IEEE OCEANS-Monterrey, pp. 1–10. Monterey, CA: IEEE. <https://doi.org/10.1109/OCEANS.2016.7761095>.
- Rahman, S., Karapetyan, N., Quattrini Li, A., & Rekleitis, I.** 2018. A modular sensor suite for underwater reconstruction. In: MTS/IEEE

- OCEANS-Charleston, pp. 1–6. Charleston, SC: IEEE. <https://doi.org/10.1109/OCEANS.2018.8604819>.
- Rahman**, S., Quattrini Li, A., & Rekleitis, I. 2019. Contour based reconstruction of underwater structures using sonar, visual, inertial, and depth sensor. In: International Conference on Intelligent Robots and Systems (IROS), pp. 8048–53. Macau, China: IEEE. <https://doi.org/10.1109/IROS40897.2019.8967697>.
- Rosenblatt**, J., Williams, S., & Durrant-Whyte, H. 2002. A behavior-based architecture for autonomous underwater exploration. *Inform Sciences*. 145(1-2):69–87. [https://doi.org/10.1016/S0020-0255\(02\)00224-4](https://doi.org/10.1016/S0020-0255(02)00224-4).
- Roznere**, M., & Quattrini Li, A. 2019. Real-time model-based image color correction for underwater robots. In: International Conference on Intelligent Robots and Systems (IROS), pp. 7191–6. Macau, China: IEEE. <https://doi.org/10.1109/IROS40897.2019.8967557>.
- Sattar**, J., Dudek, G., Chiu, O., Rekleitis, I., Giguere, P., Mills, A., ... Lobos, J. P. 2008. Enabling autonomous capabilities in underwater robotics. In: International Conference on Intelligent Robots and Systems (IROS), pp. 3628–34. Nice, France: IEEE. <https://doi.org/10.1109/IROS.2008.4651158>.
- Shkurti**, F., Rekleitis, I., & Dudek, G. 2011. Feature tracking evaluation for pose estimation in underwater environments. In: Canadian Conference on Computer and Robot Vision (CRV), pp. 160–7. St. John, NF, Canada: IEEE. <https://doi.org/10.1109/CRV.2011.28>.
- Skaff**, S., Clark, J., & Rekleitis, I. 2008. Estimating surface reflectance spectra for underwater color vision. In: British Machine Vision Conference (BMVC), pp. 1015–24. Leeds, UK: DBLP. <https://doi.org/10.5244/C.22.101>.
- Smolyanskiy**, N., Kamenev, A., Smith, J., & Birchfield, S. 2017. Toward low-flying autonomous MAV trail navigation using deep neural networks for environmental awareness. In: International Conference on Intelligent Robots and Systems (IROS), pp. 4241–7. Vancouver, BC, Canada: IEEE. <https://doi.org/10.1109/IROS.2017.8206285>.
- Vidal**, E., Hernández, J.D., Istenič, K., & Carreras, M. 2017. Online view planning for inspecting unexplored underwater structures. *IEEE Robot Autom Lett*. 2(3):1436–43. <https://doi.org/10.1109/LRA.2017.2671415>.
- Vidal**, E., Palomeras, N., Istenič, K., Hernández, J.D., & Carreras, M. 2019. Two-dimensional frontier-based viewpoint generation for exploring and mapping underwater environments. *Sensors*. 19(6):1460. <https://doi.org/10.3390/s19061460>.