Real-time Vision-only Perception for Robotic Coral Reef Monitoring and Management

Matthew Dunbabin[†‡], Feras Dayoub[†], Riki Lamont[†], Steven Martin[†]

A particular challenge for Autonomous Underwater Vehicles (AUVs) operating in complex coral reef environments and conducting low-altitude reef-based monitoring and management tasks (approximately 0.5 – 1.2 m from the coral) is position estimation and obstacle detection where traditional acoustic based sensors (e.g. sonars, Doppler Velocity Logs) become less reliable due to minimum ranging distances and multipathing. Real-time stereo and monocular vision systems offer a viable means for navigation, obstacle avoidance and in-situ automated management tasks. However, such approaches are limited by their ability to robustly detect features/objects, particularly in reduced visibility conditions, false detections from dynamic objects such as fish or strong reflections/refractions from sun-glint, and motion blur. This paper provides an overview of some real-time vision-based perception approaches developed for, and applied to, small-scale AUVs for improving underwater navigation and obstacle avoidance in coral reef environments. We also present applications of our latest work for AUV-based coral reef management tasks including control of crown-of-thorns starfish and coral larvae reseeding over damaged reefs.

I. Introduction

Underwater robotic systems, for example Autonomous Underwater Vehicles (AUVs), are transforming our ability to monitor and management complex marine environments [1]. In many applications, optical imagery is a fundamental data requirement for monitoring, object classification and manipulation. Whilst traditional underwater sensors rely on sonar-based technologies, these become prohibitively expensive and less reliable for use at large-scales and in complex coral reef environments.

Underwater systems that use only visual perception for navigation, object detection and task execution offer a potentially lower-cost and expanded capability solution for certain operational regimes. This is particularly so for complex coral reef environments where survey and management tasks are often performed at very low altitudes (e.g. 0.5 - 1.2m above the seafloor), and in highly cluttered 3D environments. However, the use of vision-only AUVs in these environments has many challenges such as image consistency, lighting, dynamic objects (e.g. fish and kelp), unstructured terrain, and complex hydrodynamic forcing (wave action). As such there has been limited examples of vision-only AUVs deployed operationally. This paper provides a high-level overview of some visual perception approaches developed for, and applied to, vision-only AUVs to perform coral reef monitoring and management tasks.

[†] Queensland University of Technology (QUT), 2 George Street, Brisbane, QLD 4000, Australia. m.dunbabin@qut.edu.au.

[‡] The Australian Research Council Centre of Excellence for Robotic Vision (ACRV).

Fig. 1: The RangerBot AUV is a vision-only robot with two stereo camera pairs for performing navigation, obstacle avoidance and science/management tasks. (Inset) the downward facing stereo camera pair and LEDs.

II. Vision-only AUVs

Although many AUVs around the world employ cameras for mapping and object detection tasks, there is a paucity of platforms that rely solely on real-time vision to perform navigation and higher-level tasks. This is even more so true for platforms which have the ability to operate in close proximity to complex coral reef environments. The limited examples within the literature are the Starbug AUV developed by the CSIRO [2]–[4], and the Aqua AUV developed by McGill University [5]. Recently, driven by the availability of small and relatively low-cost and low-power Graphic Processing Units (GPUs), the use of vision to perform certain tasks onboard AUVs is increasing. The following section presents a new AUV platform that leverages this technology.

A. The RangerBot AUV

Figure 1 shows the RangerBot AUV, a small-scale vision-only AUV specifically designed for operating in complex coral reef environments with the flexibility to perform real-time perception-based tasks for management and restoration activities. The RangerBot AUV is built around two stereo camera pairs which provide all navigation, obstacle avoidance and science/management task information. The downward stereo pair has a camera baseline of 75 mm, with the forward stereo camera pair having a baseline of 120 mm. All image processing and mission execution software runs on-board the AUV using an NVIDIA Jetson TX2 module as the primary computation capability.

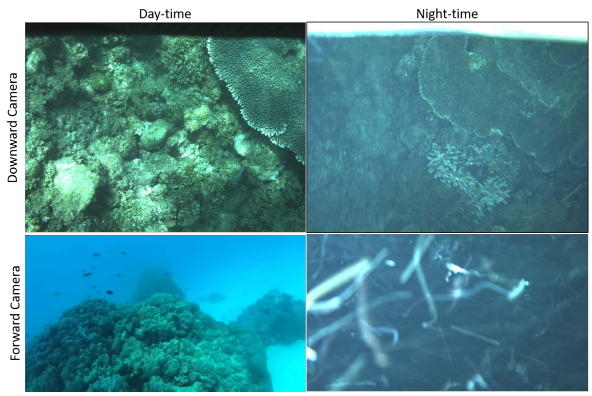

Fig. 2: Examples of images collected by the RangerBot AUV at 7.5 Hz to highlight the challenges of real-time vision processing in coral reef environments. The night-time images can be degraded by motion blur and from marine organisms being attracted to the forward lights making obstacle detection difficult. (Top row): Images from a downward facing camera, (Lower row) images from a forward facing camera.

The AUV is designed for deployment by a single person using small support vessels and shore-based operations. As such, it weighs only 16 kg with a length of 0.75 m and width of 0.44 m, and has removable and hot-swappable batteries to increase its utility in the field. Its unique thruster configuration allows full six Degree-of-Freedom control, including hover capabilities which is essential for low-altitude manoeuvring in complex coral reef environments. Operationally, the AUV has an endurance of over 6 hours (without swapping batteries) and a depth rating of 100 m. However, typically it is used from the shore to a depth of 30m for shallow coral reef monitoring activities. Other features include LED lights and a custom payload attachment point (see Section IV for examples). The RangerBot AUV has been used to help develop and evaluate the approaches presented in this paper.

III. Real-time Visual Perception for Complex Underwater Environments

Complex underwater environments, such as shallow coral reefs, exhibit many phenomena that makes robust real-time image processing challenging. These include variable lighting from shadows and wave ripple where often strong natural sunlight can overpower any strobe lighting. Additionally, night-time operations can be compromised by too little light, or by marine organisms that are attracted to the lights and clutter the image. Also when conducting low-altitude survey work, motion blur from low-lighting and camera hardware limitations can degrade image quality. Figure 2 provides examples of different image quality obtained by the RangerBot AUV to illustrate some of the challenges particularly for night-time operations.

Underwater visibility enhancement has promise to improve image quality for a range of navigation and detection tasks and has been researched for many years. However, the real-time implementation on AUVs appears limited, particularly in coral reef environments. Previous work [6] has developed an approach which exploits stereo imagery to simultaneously enhance and color correct an underwater image. Recent implementations of the approach has achieved near real-time (0.5Hz) processing on low-power GPU's. Newer enhancement approaches that use machine learning techniques, such as Generative Adversarial Networks [7], that can exploit GPU technology have promise for use on vision-only AUVs.

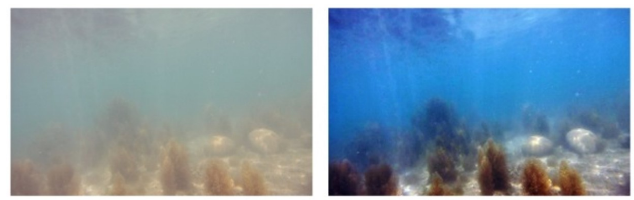

Fig. 3: Example of underwater image visibility enhancement using the method described in [6] with application to low-visibility cluttered environments. (Left) original, (Right) enhanced.

A. *Visual odometry*

The ability for an AUV to know or estimate its position whilst underwater is vital for successful mission execution. Traditionally, position estimation (whilst underwater) is provided by some form of acoustic localization (e.g. Ultra Short Baseline, Long Baseline) with this information transmitted to the AUV via an acoustic modem or tether. In shallow coral reef environments, due to multipathing, non-line-of-sight operations and logistical constraints, these approaches become less reliable with the AUV being more reliant on its own on-board sensors to estimate position.

Position estimation has been explored and successfully deployed particularly using underwater Simultaneous Localization and Mapping (SLAM) [8], [9], [10]. However, in many situations, such as AUV-based reef management applications, these approaches become less suitable due to no opportunities for loop-closure. One approach for shallow water position estimation is to use visual odometry combined with operational methods to limit odometry drift [11]. This was explored and evaluated in early work using a vision-only AUV [12]–[14] which showed navigation performance errors of $<8\%$ of distance travelled in representative coral reef environments. In recent years, more advanced open source visual odometry methods (e.g. LIBVISO2 [15]) have shown promise. Whilst considered optimized for forward looking cameras, these techniques can be modified to achieve relatively low error odometry estimates for downward looking cameras in coral reef environments. Figure 4 shows an example of a real-time estimated visual odometry track implemented on-board an AUV across a complex trajectory. In these trials, error rates of less than 2% distance travelled were achieved over coral, rubble and coarse sand substrates.

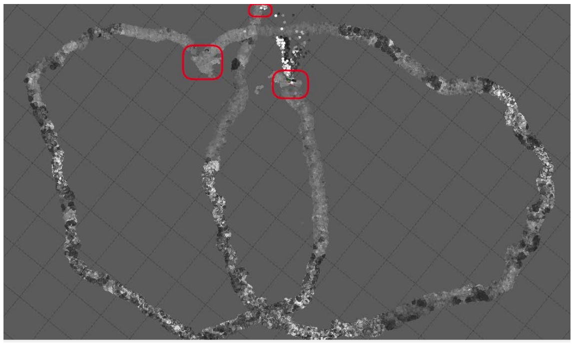

Fig. 4: Example of visual odometry during a complex multi-segment shallow water mission with partial sand cover. The AUV starts from near the boat (marked with red circles) and performs two loops returning to the boat after each loop.

B. Semantic monocular obstacle detection

Operating AUVs in previously unvisited and cluttered environments (e.g. reefs) typically requires some form of real-time and reactive obstacle detection and avoidance strategy. Sonars are typically used for mapping the immediate surrounds, however, in close proximity to structures their resolution and reliability degrade.

Vision-based underwater obstacle detection has been explored in recent years and typically employ producing range (forward looking) estimates of the scene producing sparse (e.g. SURF [16]) and dense maps (e.g. [17] and LIBELAS [18]) which can be produced in real-time on-board an AUV. The effective range of these obstacle maps is visibility dependent, with both sparse and dense maps often challenged by dynamic objects in the scene (e.g. fish, kelp) and large gaps which limit feature density.

In some recent work, Arain et al. [19] proposed the use of semantic image segmentation to enhance image-wide obstacle detection. In this approach, feature-based stereo matching is combined with learning-based segmentation motivated by [20] to produce more robust obstacle maps for AUVs operating in coral reef environments. Two methods of image segmentation are considered; The first is a binary classification which segments the image into obstacle and non-obstacle regions. The semantically labelled results can be used directly from monocular image streams for conservative obstacle avoidance. However, by combining them with sparse features (e.g. SURF) obtained from stereo imagery, the segmented image can be 'draped' over the sparse features to create a more complete obstacle map. Figue 5 shows an example of the binary labelled image with estimated 3D obstacle map.

The second approach is a multiclass classification where during training, the scene is partitioned into subjective regions (near, mid, far and free-space) based on sparse ranging data. This approach has the advantage that a monocular image stream can be analyzed to produce a semi-continuous estimate of the obstacles that the AUV may need to plan through. Examples of outputs from the multiclass obstacle segmentation are show in Figure 6.

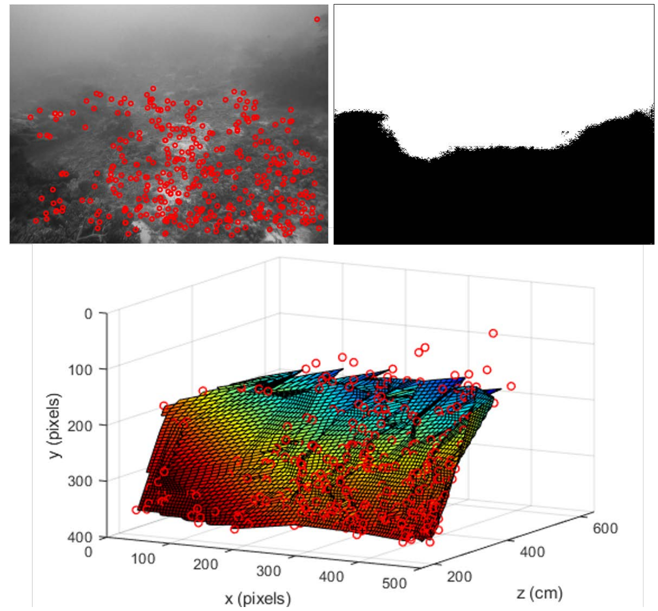

Fig. 5: Example of a 3D obstacle map by combining the sparse feature map (top left) with the semantically labelled obstacle image (top right). The lower figure shows an isometric view of the resulting 3D obstacle map (relative to the image) which an AUV could use for reactive visual servoing and local path planning (Extracted from [19]).

IV. Visual Perception-to-Action for Coral Reef Management and Restoration

Whilst most underwater vision systems focus on navigation and obstacle avoidance for monitoring only tasks, real-time vision systems have to the potential to perform higher level management actions based on what is observed as the AUV traverses the environment. We term this *Perception-to-Action*, and requires an ability to close-the-loop between the on-board vision system and a tool or instrument without human intervention. The following sections provide an overview of two novel applications of perception-to-action using vision-only AUVs for management of coral reefs.

A. Automated Crown-of-Thorns Starfish (COTS) Detection and Population Control

Crown-of-Thorns Starfish (*Acanthaster planci*) are a significant threat to the Great Barrier Reef [21]. These starfish literally eat the coral and in recent years have proliferated with accelerated outbreaks occurring. Controlling their numbers is primarily performed using manual injection of a biological agent into the starfish by a diver using a hand-held supply gun and needle [22]. In order to help upscale Crown-of-Thorns Starfish (COTS) control efforts, we proposed the use of a vision-based robotic system capable of automated COTS detection and real-time injection of the starfish.

COTS are very cryptic and can be difficult to detect within coral reefs. Early work on vision-based detection [23] was feature-based with limited performance and non-real time processing times. In recent work [24], we developed the

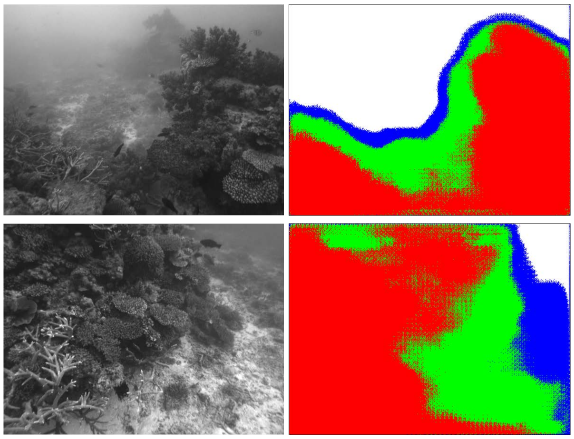

Fig. 6: Examples of monocular obstacle segmentation using a multiclass classifier; (Left) Original image, (Right) the predicted obstacle range based on training images (i.e. red is near, green is mid-field, blue is far and free-space is white). Results by Bilal Arain, QUT.

first real-time COTS detection and tracking system for use on a moving AUV in a dynamic and visually degraded environment. This was a novel Random Forest Classifier (RFC) trained from underwater footage which was embedded within a particle filter detector and tracker. The predicted class probability of the RFC was used as an observation probability to weight the particles, with sparse optic flow estimation used for the prediction step of the filter.

In order to train the classifier, cameras were attached to the hand-held injection guns used by divers during population control campaigns. Due to the variability in the way COTS are encountered in the environment, the training images were separated into three classes (easy, medium and hard) with examples shown in Figure 7. Training was performed against only the *easy* and *medium* classes due to the risk of the robotic injection system getting entangled within coral during deployment when trying to inject *hard* to observe COTS cases.

Fig. 7: Example of COTS degree-of-detection classification used for training; (Left) Easy, (Middle) Medium, and (Right) Hard.

Figure 8 shows the number of false positive detection from two large image sequences using a detection threshold (see [24] for details) of 50% against the score of the best particle. These results show for a detection threshold of 50% there were approximately 2% false positive detections. However, in practice this would not be acceptable due to the system potentially trying inject corals rather than starfish causing damage to the robot. A threshold of >80% on the best particle score reduces the false positive rate to 0.3% which when combined with hierarchical operational constraints (e.g. constant altitude image collection and depth) provides robust and acceptable detection accuracy. This classifier achieved a processing rate of 8Hz on-board the RangerBot AUV when running in conjunction with the vision-based navigation system and evaluated on the Great Barrier Reef.

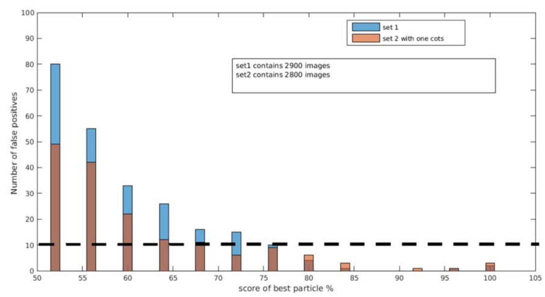

Fig. 8: Results showing the number of false positive COTS detections with a RFC detection threshold of 50% from two field image sequences against the score of the best particle.

A limitation of the approach described in [24] was that only one COTS could be tracked in an image in real-time. The approach was updated by incorporating a detector based on GoogLeNet deep network to allow detection and tracking of multiple COTS in the same image. Figure 9 shows an example output from the classifier with multiple target detections and using non-maxima suppression to avoid over labelling.

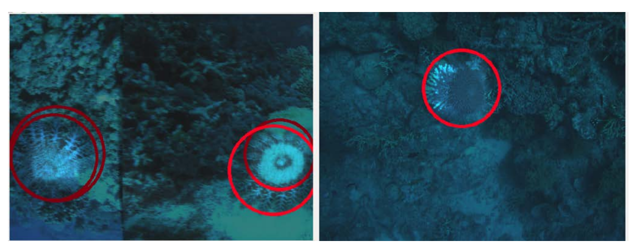

Fig. 9: Example of real-time multi-COTS detector output. (Left) Detector results before a non-maxima suppression stage where multiple over labelling detections are merged together. (Right) Improved overall COTS segmentation.

This multi-COTS detector was evaluated using the RangerBot AUV which achieves an image processing rate of approximately 10 Hz (running in addition to the vision-based navigation system). The perception-to-action task was closed with the addition of a novel injection system that can physically inject the detected COTS with Bile Salts [22] as shown in Figure 10. Field results on the Great Barrier Reef demonstrated real-time detection and control of starfish for both the *easy* and *median* classes shown in Figure 7.

B. Automated Coral Larvae Reseeding for Reef Restoration

Coral cover across the world's reefs has significantly declined due factors including coral bleaching, cyclones and

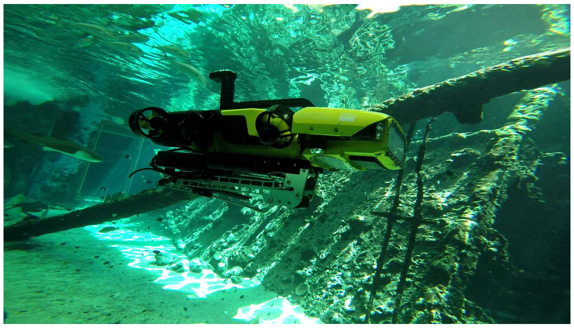

Fig. 10: The RangerBot AUV with the self-contained COTS injection system attached to the underside of the vehicle during controlled trials (shown in the retracted state). The injection arm extends out and down to autonomously deliver a dose of bile salts into the detected COTS.

destructive fishing practices (e.g. explosives). On various regions of damaged reefs, active restoration approaches are being explored to increase coral cover and return biodiversity. One such approach is coral reseeding [25] where coral larvae are captured during their mass coral spawning [26], then reared in special cages at sea for 4–7 days until they are ready to settle and then redistributed onto the damaged sections of reef.

Until recently, coral reseeding required manual redistribution of the larvae, limiting the achievable restoration area to 100-200 m^2. To increase the daily restoration area to a desired hectare to km^2 scale, we proposed the use of a robotic system to automatically distribute the coral larvae over the damaged reefs using real-time visual perception-to-action to strategically place larvae on the most appropriate substrates. As an initial trial, a custom larval delivery system was developed for a RangerBot AUV which could automatically release larvae. In order to train what is considered 'appropriate' substrate for larval settlement, the AUV was programmed to conduct a vision-based grid survey across a site with a human expert remotely triggering larvae release when over suitable locations. Figure 11 shows a RangerBot AUV during coral reseeding trials on the Great Barrier Reef.

Fig. 11: A RangerBot AUV with the first prototype coral larvae delivery system during reseeding trials on the Great Barrier Reef in December 2018.

The recorded stereo images along with expert "labelling", and derived metrics (e.g. rugosity, slope and depth) were initially combined in an SVM to predict when larvae should be released. The Great Barrier Reef trials (Figure 11) covered approximately 200 m^2. In April 2019, an upgraded larvae delivery system was demonstrated using the RangerBot AUV on severely degraded reefs in the Philippines. In these trials, a single vision-based AUV distributed over 2.5 million larvae across three hectares in 6 hours. Figure 12, shows a RangerBot AUV during automated reseeding activities in the Philippines. Current work is now investigating DCNN approaches to improve automated site selection performance.

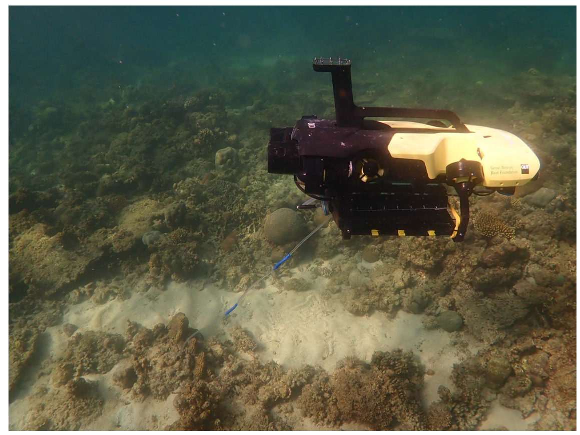

Fig. 12: A RangerBot AUV during vision-based coral larval reseeding activities in the Philippines, April 2019. The larvae delivery system is attached to the underside of the AUV.

V. Conclusions

This paper has provided an overview of some real-time vision-based perception approaches developed for, and applied to, AUVs to improve underwater navigation and obstacle avoidance in coral reef environments. The approaches include methods for image enhancement, obstacle detection and visual odometry and have been successfully evaluated in shallow, yet complex reef environments using vision-only 6DOF AUVs. In addition to the image processing approaches, examples of real-world applications of vision-based perception-to-action are presented for complex coral reef management and restoration tasks that include the robotic control of crown-of-thorns starfish and coral larvae reseeding over damaged reefs.

Acknowledgments

The work presented in this paper has been partly funded by the Great Barrier Reef Foundation though the 2016 Google Impact Challenge and the Tiffany & Co. Foundation, and the Australian Institute of Marine Science (AIMS). The authors would like to thank Peter Harrison from the Southern Cross University for his project leadership and collaborations in the coral larvae reseeding program in the Philippines through the Australian Centre for International Agricultural Research grant scheme. We would also like to thank Bilal Arain from QUT, and Paul Rigby from AIMS for their research and collaboration in learning-based underwater obstacle avoidance.

References

[1] M. Dunbabin and L. Marques, "Robots for environmental monitoring: Significant advancements and applications," *IEEE Robotics Automation Magazine*, vol. 19, no. 1, pp. 24–39, March 2012.

[2] M. Dunbabin, P. Corke, and G. Buskey, "Low-cost vision-based AUV guidance system for reef navigation," in *Proceedings of the 2004 IEEE International Conference on Robotics & Automation*, April 2004, pp. 7–12.

[3] A. Marouchos, B. Muir, R. Babcock, and M. Dunbabin, "A shallow water auv for benthic and water column observations," in *OCEANS 2015 - Genova*, May 2015, pp. 1–7.

[4] M. Dunbabin, J. Roberts, K. Usher, G. Winstanley, and P. Corke, "A hybrid AUV design for shallow water reef navigation," in *Proceedings of the 2005 International Conference on Robotics and Automation*, Barcelona, Apr. 2005, pp. 2117–2122.

[5] G. Dudek, P. Giguere, C. Prahacs, S. Saunderson, J. Sattar, L. Torres-Mendez, M. Jenkin, A. German, A. Hogue, A. Ripsman, J. Zacher, E. Milios, H. Liu, P. Zhang, M. Buehler, and C. Georgiades, "Aqua: An amphibious autonomous robot," *Computer*, vol. 40, no. 1, pp. 46–53, Jan 2007.

[6] M. Roser, M. Dunbabin, and A. Geiger, "Simultaneous underwater visibility assessment, enhancement and improved stereo," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, May 2014, pp. 3840–3847.

[7] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," *arXiv preprint arXiv:1801.04011*, 2018.

[8] M. Meireles, R. Loureno, A. Dias, J. M. Almeida, H. Silva, and A. Martins, "Real time visual slam for underwater robotic inspection," in *2014 Oceans - St. John's*, Sep. 2014, pp. 1–5.

[9] S. Rahman, A. Q. Li, and I. Rekleitis, "Sonar visual inertial slam of underwater structures," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018, pp. 1–7.

[10] S. B. Williams, O. R. Pizarro, M. V. Jakuba, C. R. Johnson, N. S. Barrett, R. C. Babcock, G. A. Kendrick, P. D. Steinberg, A. J. Heyward, P. J. Doherty *et al.*, "Monitoring of benthic reference sites: using an autonomous underwater vehicle," vol. 19, no. 1, pp. 73–84, 2012.

[11] M. Dunbabin and S. Allen, "Large-scale habitat mapping using vision-based AUVs: Experiences, challenges & vehicle design," in *Proc. OCEANS 2007 Europe*, Aberdeen, UK, 2007, pp. 1–6.

[12] M. Dunbabin, K. Usher, and P. Corke, "Visual motion estimation for an autonomous underwater reef monitoring robot," in *Proc. International Conference on Field & Service Robotics*, Port Douglas, 2005, pp. 57–68.

[13] S. S. da Costa Botelho, P. Drews, G. L. Oliveira, and M. d. S. Figueiredo, "Visual odometry and mapping for underwater autonomous vehicles," in *2009 6th Latin American Robotics Symposium (LARS 2009)*, Oct 2009, pp. 1–6.

[14] M. Nawaf, D. Merad, J.-P. Royer, J.-M. Bo, M. Saccone, M. Ben Ellefi, and P. Drap, "Fast visual odometry for a low-cost underwater embedded stereo system ," *Sensors*, vol. 18, p. 2313, 07 2018.

[15] A. Geiger, J. Ziegler, and C. Stiller, "Stereoscan: Dense 3d reconstruction in real-time," in *Intelligent Vehicles Symposium (IV)*, 2011.

[16] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (surf)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346 – 359, 2008.

[17] P. Drews, E. Hernández, A. Elfes, E. R. Nascimento, and M. Campos, "Real-time monocular obstacle avoidance using underwater dark channel prior," in *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016, pp. 4672–4677.

[18] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Computer Vision–ACCV 2010*. Springer, 2010, pp. 25–38.

[19] B. Arain, C. McCool, P. Rigby, D. Cagara, and M. Dunbabin, "Improving underwater obstacle detection using semantic image segmentation," in *2019 IEEE International Conference on Robotics and Automation (ICRA)*, May 2019.

[20] S. McMahon, N. Sunderhauf, B. Upcroft, and M. Milford, "Multimodal trip hazard affordance detection on construction sites," *IEEE Robotics and Automation Letters*, no. 1, 2018.

[21] G. De'ath, K. E. Fabricius, H. Sweatman, and M. Puotinen, "The 27-year decline of coral cover on the great barrier reef and its causes," *Proceedings of the National Academy of Sciences*, vol. 109, no. 44, pp. 17 995–17 999, 2012. [Online]. Available: http://www.pnas.org/content/109/44/17995.abstract

[22] Great Barrier Reef Marine Park Authority, "Crown-of-thorns starfish control guidelines," 2014.

[23] R. Clement, M. Dunbabin, and G. Wyeth, "Towards robust image detection of crown-of-thorns starfish for autonomous population monitoring," in *Proc. Australasian Conference on Robotics and Automation*, Sydney, December 2005.

[24] F. Dayoub, M. Dunbabin, and P. Corke, "Robotic detection and tracking of crown-of-thorns starfish," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 1921–1928.

[25] D. Dela Cruz and P. L. Harrison, "Enhanced larval supply and recruitment can replenish reef corals on degraded reefs," *Scientific Reports*, vol. 7, 12 2017.

[26] P. L. Harrison, R. C. Babcock, G. D. Bull, J. K. Oliver, C. C. Wallace, and B. L. Willis, "Mass spawning in tropical reef corals," *Science*, vol. 223, no. 4641, pp. 1186–1189, 1984.