Robotic Perception of Underwater Plastic Bottles for Augmented Telepresence

Aaron Smiles, Changjae Oh, and Ildar Farkhatdinov

Queen Mary University of London, UK
a.l.smiles@qmul.ac.uk

Abstract. Not enough robots are used in the fight against ocean plastics. Remotely-operated vehicles (ROVs), controlled via telepresence, can collect ocean waste, but underwater depth perception is challenging, especially for transparent objects like plastic bottles on 2D screens, where augmented visuals could assist. This study investigates stereo vision technology for underwater depth estimation, focusing on transparent objects. We recalibrated a stereo camera for underwater applications, testing depth accuracy on air- and water-filled plastic bottles under controlled lighting. Results show that recalibration improves depth perception for opaque objects but remains limited for transparent materials, particularly water-filled bottles. Our approach and performance analysis highlight the challenges of accurate depth mapping for submerged plastics, underscoring a need for advanced methods to achieve reliable augmented visuals.

Keywords: stereo vision, underwater computer vision, depth perception, augmented telepresence, transparent objects

1 Introduction

Ocean plastic pollution has become a critical environmental concern, with an estimated 8 million metric tons of plastic entering the oceans annually. While much attention has been given to surface pollution, a significant portion (99%[1]) of this plastic debris sinks below the surface or accumulates on the seafloor [2]. Plastic bottles are a big contributor, with around 8 million tons of them entering the ocean each year. The presence of this submerged plastic poses challenges for detection and removal, necessitating advanced technological solutions, and robots, or autonomous underwater vehicles (AUVs) and Remotely-operated vehicles (ROVs), will need to be involved. The SeaClear project[2] deploys collaborative AUVs for cleanup, while the Rozalia Project[3] deploys a small ROV that is controlled by a human operator with a system that includes a small 2D

[1] TOPIOS: Tracking Of Plastic In Our Seas, 2017-2022, http://topios.org/
[2] SEACLEAR: Search, Identification and Collection of Marine Litter with Autonomous Robots, https://seaclear-project.eu/
[3] Rozalia Project, https://www.rozaliaproject.org/innovation-technology

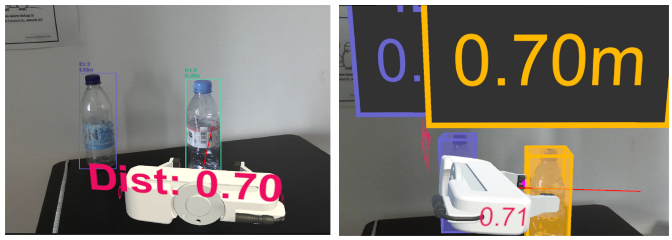

Fig. 1. Demonstration of augmented telepresence system in air. Left: 2D custom detector in Unity showing 2D bounding boxes of detected bottles and AR end effector (recorded from 2D screen). Right: 3D custom detector in Unity showing 3D bounding boxes and AR end effector (recorded from VR HMD screen).

monitor. While [3] make the case for AUVs, we argue that ROVs are overlooked and human intervention will be required in a number of situations.

Operating these vehicles in complex underwater environments requires precise control and situational awareness. Augmented telepresence, which combines real-time video feeds with computer-generated visual or auditory cues, has shown promise in enhancing operator performance and decision-making in remote operations [1]. We have developed a mixed reality system for simulated plastic waste collection with bottle detection (YOLOv7-tiny), which has been tested in air (demos: github.com/ATSEA), but underwater scene depth is required.

This research builds upon our previous work [12]. While our initial study focused on manually performing stereo vision calibration with two webcams and measuring object distances to assess the feasibility of applying these techniques to augmented telepresence, our current work takes a significant step forward. We now employ the Stereolabs ZED Mini stereo camera (ZEDM), allowing more advanced vision techniques within the Unity virtual environment. The study presents distance estimations in an experimental underwater tank for four objects, including transparent plastic bottles recorded in two states (air- and water-filled), and a transparent bottle in five lighting scenes.

Models have been deployed for detecting sub-surface marine litter. The authors [3] focus on litter floating in the water column, whereas, [9] focus on seafloor plastics, which include some transparent plastic materials; however, focus on identification and not depth. There are various methods for the detection of transparent objects in air ([5]), for example, the SuperCaustics [8] transparent object simulator was developed to generate synthetic transparent object datasets to train models to detect physical transparent objects [10], but underwater scenes are not addressed. The authors [13], note that the detection of low-contrast and translucent objects is difficult due to variable lighting and debris in the water that creates noise. They propose various methods to overcome such issues for automating video annotation, but this was not tested in a real-time scenario. Although the authors [7], use a fast detection method for detecting transparent ocean organisms, both papers do not obtain depth information of such arti-

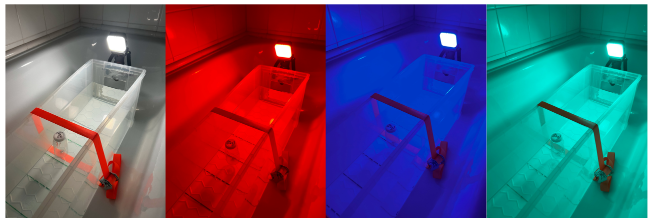

Fig. 2. Experimental underwater setup consisting of a plastic storage box as a tank, a perspex window, a 3D printed mount to attach the stereo camera to the window, and an LED light for controlling directional light and different lighting scenes, as shown: white, red, blue, and green. Yellow is omitted as the yellow hue would not photograph.

facts. SeaSplat [15] uses 3D Gaussian splatting to reconstruct underwater scenes photorealistically, whilst obtaining depth [14], markedly faster than NeRFs [6], which can be applied to ROV ops, but transparent objects are unresearched.

To our knowledge, no research exists on fast distance estimation of transparent objects or materials in water, especially when the object is filled with and surrounded by water. Providing insights for depth mapping of underwater transparent plastics distinguishes our work from other depth perception research typically focused on air-based applications. By addressing these challenges, this research aims to enhance the capabilities of ROVs in locating and removing submerged plastic debris, ultimately contributing to global efforts in ocean plastic cleanup and advancing augmented telepresence for underwater operations.

2 Implementation

Since the ZEDM is not waterproof, an experimental tank was constructed to perform underwater tests. The setup was built using a plastic storage box with dimensions H36, W27, D120 cm (enough depth to capture a range of measurements). A window was cut out of the translucent plastic and replaced with clear perspex. This was sealed with aquarium sealant. A custom mount was 3D printed to secure the ZEDM to the tank in the position of the window. The ZEDM required recalibration for use underwater due to the different refractive properties of water compared to air [4][11]. Recalibration of the ZEDM adapted the method in the Stereolabs documentation[4]. A root mean squared reprojection error of 0.22 px was achieved. To validate depth estimation, we took a series of distance measurements on objects of different materials and opacities (Fig. 3). The tank was marked with measurements at 100 mm increments to use as our ground truth (GT), and distance readings were taken between 200 and 1000 mm. The ZED stereo camera was set to Neural Depth mode, which uses a specially designed convolutional neural network (CNN)[5]. The tank was filled with clear

[4] https://www.stereolabs.com/docs/opencv/calibration
[5] https://www.stereolabs.com/en-gb/blog/neural-depth-sensing

Fig. 3. The objects used for estimating distance using the stereo vision sub-system. From left to right: Textured clear bottle, smooth clear bottle, white bottle, stone.

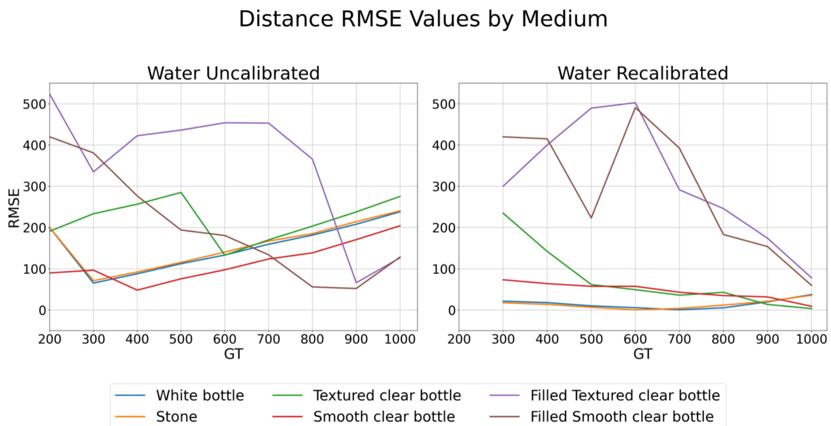

Fig. 4. Results of depth estimation in water: uncalibrated (left), recalibrated (right).

tap water and the room was naturally lit with sunlight and overhead spotlights, creating a uniform lighting in the water. The second experiments used an LED light mounted to the front of the tank to control directional and variable lighting.

3 Results

Depth Estimation by Medium: An experiment was conducted to evaluate the accuracy of distance estimation to objects in the underwater environment with the stereo vision system. The camera had obvious problems with transparent bottles, particularly empty bottles in air and water-filled bottles in water. This was due to the clear bottles containing the same fluids as the surrounding environment and becoming almost invisible to the camera. What was found was, at those points, the stereo system either returned NaN or an estimate consistent with the next visible point behind the bottle, for example, the back of the tank.

The results of recalibration show an improvement in the accuracy of opaque objects, a marginal improvement with translucent objects (air-filled clear containers submerged in water), and no improvement with transparent objects

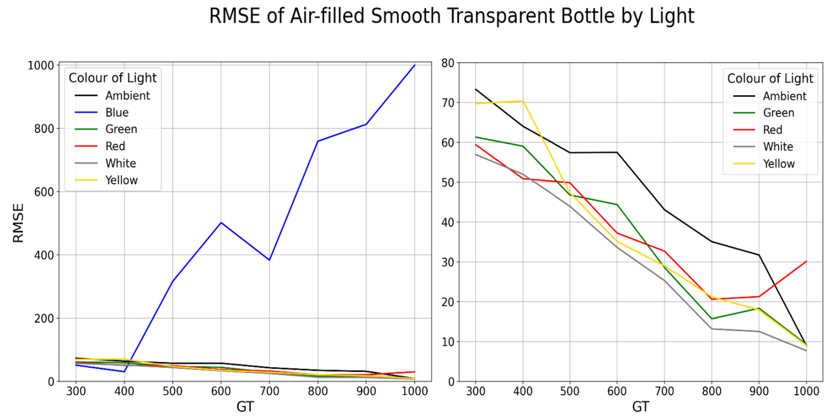

Fig. 5. Depth results on the smooth air-filled transparent plastic bottle by different colours of directional light, inc. Ambient (non-directional) light for comparison.

(water-filled clear containers submerged in water) 4. The reason the accuracy decreases for the filled transparent objects after recalibration is due to the estimation of the rear wall of the tank becoming more accurate as the detection cannot pick up the bottle and "sees" through it, providing an estimate for the distance to the next surface. After recalibration the accuracy of the filled transparent bottles appears to improve towards 1000 mm (Fig. 4); however, this is due to the error reducing as the bottles come closer to the back of the tank.

Depth Estimation by Light: Due to blue light's propensity to penetrate water more than any other wavelength of light, it penetrated the translucent walls of the tank, making it difficult for the algorithm to generate an accurate depth map, likely requiring a new recalibration to perform well. Conversely, since red light penetrates water the least, it lost accuracy beyond 800 mm. The other light colours produced a similar small improvement, but white directional light improved the accuracy most over ambient (non-directional) white light on the air-filled bottle by reducing the root mean squared error (RMSE) from 50.306 to 35.427. All results of directional light on the water-filled bottles were inconclusive, due to the algorithm returning mostly NaN results.

4 Conclusion

This research underscores the potential of recalibrated stereo vision systems, particularly the ZEDM, for underwater applications. Achieving an RMSE of 18 mm for opaque objects, this study validates the enhanced accuracy brought by recalibration but also exposes significant challenges in detecting transparent objects submerged in water. The limitations observed, especially with water-filled transparent bottles, suggest that current stereo vision systems are not yet optimised for depth estimation with transparent materials under these conditions.

Future work will explore integrating advanced transparent object detection algorithms and 3D reconstruction methods to overcome these challenges. Advancing techniques for underwater depth perception contributes towards more effective ROV-based solutions for marine conservation and plastic waste cleanup.

References

1. Brantner, G., Khatib, O.: Controlling ocean one: Human–robot collaboration for deep-sea manipulation. Journal of Field Robotics **38**(1), 28–51 (2021)
2. Chiba, S., Saito, H., Fletcher, R., Yogi, T., Kayo, M., Miyagi, S., Ogido, M., Fujikura, K.: Human footprint in the abyss: 30 year records of deep-sea plastic debris. Marine Policy **96**, 204–212 (2018)
3. Fulton, M., Hong, J., Islam, M.J., Sattar, J.: Robotic detection of marine litter using deep visual detection models. In: 2019 ICRA. IEEE
4. Hagen, O., Aly, A., Jones, R., Varga, M., Bazazian, D.: Beyond the surface: a scoping review of vision-based underwater experience technologies and user studies. Intelligent Marine Technology and Systems **2**(1) (2024)
5. Jiang, J., Cao, G., Deng, J., Do, T.T., Luo, S.: Robotic perception of transparent objects: A review. IEEE Transactions on AI **5**(6), 2547–2567 (2024)
6. Levy, D., Peleg, A., Pearl, N., Rosenbaum, D., Akkaynak, D., Korman, S., Treibitz, T.: Seathru-nerf: Neural radiance fields in scattering media (2023)
7. Luo, C., Wu, J., Sun, S., Ren, P.: Transcodnet: Underwater transparently camouflaged object detection via rgb and event frames collaboration. IEEE Robotics and Automation Letters **9**(2), 1444–1451 (2024)
8. Mousavi, M., Estrada, R.: Supercaustics: Real-time, open-source simulation of transparent objects for deep learning applications. 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA) (2021)
9. Politikos, D.V., Fakiris, E., Davvetas, A., Klampanos, I.A., Papatheodorou, G.: Automatic detection of seafloor marine litter using towed camera images and deep learning. Marine Pollution Bulletin **164**, 111974 (2021)
10. Sajjan, S., Moore, M., Pan, M., Nagaraja, G., Lee, J., Zeng, A., Song, S.: Clear grasp: 3d shape estimation of transparent objects for manipulation. 2020 IEEE International Conference on Robotics and Automation (ICRA) (2020)
11. Shortis, M.: Calibration techniques for accurate measurements by underwater camera systems. Sensors (Basel) **15**(12), 30810–26 (2015)
12. Smiles, A.L., Chavanakunakorn, K.D., Omarali, B., Oh, C., Farkhatdinov, I.: Implementation of a stereo vision system for a mixed reality robot teleoperation simulator. In: Towards Autonomous Robotic Systems. pp. 494–502 (2023)
13. Walther, D., Edgington, D., Koch, C.: Detection and tracking of objects in underwater video. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. (2004)
14. Xu, H., Peng, S., Wang, F., Blum, H., Barath, D., Geiger, A., Pollefeys, M.: Depthsplat: Connecting gaussian splatting and depth (2024)
15. Yang, D., Leonard, J.J., Girdhar, Y.: Seasplat: Representing underwater scenes with 3d gaussian splatting and a physically grounded image formation model (2024)