Reduced Backscatter for 3D Scene Reconstruction for Turbid Underwater Environments

Philipp Wu[1], Justin Koch[2], Torkom Pailevanian[2], Michael Garrett[2], Dan Levine[2], Christopher Yahnker[2], Matthew Gildner[2], Jeremy Nash[2], and Renaud Detry[2]

***Abstract*— [1] [2] Turbidity provides one of many challenges to underwater robotic vision. This paper studies and presents a method for extracting high quality scene geometry in short range underwater applications that is robust to high turbidity conditions. The hardware of the vision system used only requires two cameras acting as a stereo pair and a projector and can be assembled with off the shelf components. Accuracy of the reconstruction in turbid conditions is improved through partial illumination of the scene by projecting individual stripes across the environment. Using individual stripes allows for the amount of projected light to decrease, reducing the amount of backscattered light. In addition, the stripe of light provides a distinct feature in the image which additionally improves the stereo reconstruction. Our experiments show that the images taken with individual stripes allow for easier stereo matching, resulting in reconstructions with greater coverage and smaller RMS error in highly turbid scenarios (NTU 3-5).**

Fig. 1. *The vision system. The two cameras with dome ports are on the left and middle. The projector is on the right side of the image.*

I. Introduction

Computer vision, a well studied problem in air, is a critical sensing subsystem that enables robots to autonomously operate in the environment. However, underwater environments provide various challenges to robotic vision such as distortion, backscatter of light and turbid waters, which makes perception and 3D reconstruction difficult. Underwater robots have seen much use in applications such as ship inspection, documentation of ocean fauna, and maintenance of subsea infrastructure [3]. In these applications, robots are largely controlled by human teleoperators through a remote camera feed and requires significant monetary investment, time and skill from the user as well as an entire dedicated crew [4]. Such systems could see drastic improvements in efficiency and effectiveness though automation, which requires having quality 3D reconstructions. Robotic systems like the Aquasimian have been created to enable semi-autonomous execution of underwater maintenance tasks [5]. Such systems are able to autonomously perform simple tasks, such as turning a valve and inserting a stab, by localizing objects of interest through their perception systems.

Robots in underwater environments can face a wide range of turbidities that can degrade the effectiveness of 3D reconstructions algorithms. Methods to mitigate the deterioration of robotic perception in turbid underwater situations are less well studied in the literature. However, it is important for robots to be robust to turbidity as conditions can greatly very while operating in the field. Sand and dirt that is present in the environment could greatly reduce the visibility of a robot. Turbidity conditions can also be highly varying due to natural ocean phenomena, such as the turbidity currents found in the Mendocino Canyon of California [6]. Intense turbidity conditions not only make it difficult for human operators, but can significantly reduce the quality of camera images due to the increase of absorbed and reflected light from the suspended particulates.

Vision systems for underwater applications are commonly equipped with a stereo camera pair and projector. The system we used to test our methods is show in Figure 1 and was specifically designed to help maintain subsea infrastructure, allowing greater autonomy in a field that commonly requires much human supervision. The specific hardware used is the same as that described by Detry et. al., namely, two Point-Grey gigabit-ethernet Flea cameras and a Texas Instruments DLP4710 projector of 600 lumens [7]. The cameras are covered by domeports which act as a diverging lens to help mitigate distortion [8] [9].

In this paper we present a method to generate high quality reconstructions in turbid underwater environments using a standard vision suit. Experimental results conducted on an example task panel are provided, demonstrating the effectiveness of the method. We utilize the projector to partially illuminate the scene, resulting in stereo 3D reconstructions with higher accuracy and coverage than previous methods at high turbidities (NTU 3-5).

Affiliations:
[1]Electrical Engineering & Computer Science, University of California, Berkeley phil80301@berkeley.edu
[2] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA detry@jpl.nasa.gov

II. Related Work

3D reconstruction methods utilizing a vision suit composed of a stereo camera pair and projector have been well explored for underwater applications. The general system performance makes such a system ideal for short range 3D reconstruction in maintenance and documentation tasks, which are more precise than many non optical techniques such as radio and ultrasonic based sensors [10] [11]. This section presents a brief overview of existing work in this area.

A. Stereo

Stereo vision is a well studied problem in computer vision and is ideal in underwater environments. These algorithms utilize a known relative positioning of the camera pair to localize the distance of pixels though triangulation of matched features between the two camera images. Stereo vision systems for underwater applications can be easily assembled with commodity components and provide accurate short range reconstructions [12]. Many underwater stereo 3D reconstruction methods exist, but suffer from scenes with uniform texture, turbid conditions, and spatial distortion [13] [14] [15].

B. Structured Light

Structured light algorithms rely on projecting a known pattern of light onto the scene and take advantage of differences in pixel intensities to match features between a stereo camera pair. Gray codes are a common sequence of binary patterns which creates a code to localize pixels in the camera image [16]. The intensity of each pixel during every pattern is recorded, and after decoding, can be used as features for stereo 3D reconstruction. Bruno et. al. tested structured light methods underwater to generate point clouds under a variety of turbidity conditions and showed that quality drastically decreased in higher turbidity conditions [17]. Other structured light methods circumnavigate the need for two cameras by calibrating the projector relative to the camera, and treating the projector as another camera [18]. Structured light algorithms require multiple image captures with the camera in the same location in order to accurately decode the final result. A variety of single shot structured light methods have been studied and used in air, including continuous color patterns, stripe indexing methods, and grid indexing [19]. However, underwater conditions strongly degrade the quality of these other methods due to the absorbed wavelengths of light underwater.

Other methods utilize the projection and scanning of a sheet laser or project a set of stripes across the scene to provide reconstructions [20] [21]. Narasimhan et. al. has proposed a method to use a projector to scan a light stripe across the scene for underwater applications, and shows effectiveness in scene reconstruction even with the camera projector pair outside the liquid medium [22]. This method of scanning through a scene has shown robustness in a range of moderate turbidity conditions. We further discuss idea of using individual stripes as well as the advantages of using stripes for turbid conditions below.

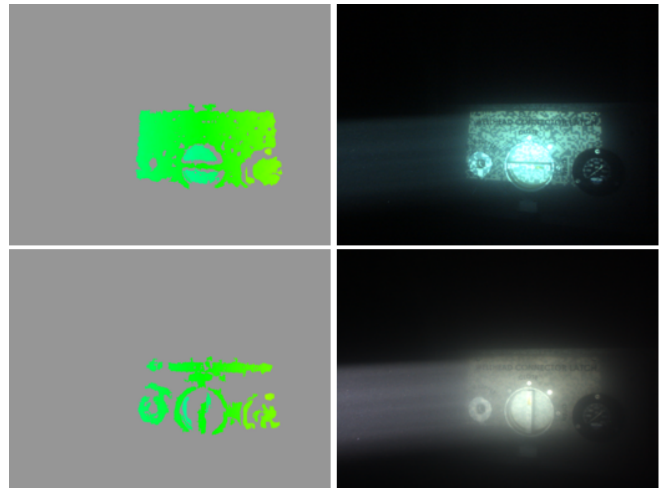

Fig. 2. *An example of degrading reconstruction quality with increased turbidity. The left column shows the disparity map, and the right column shows the image received from the right most camera. The top row demonstrates performance in natural lighting conditions, while the bottom row demonstrates the degradation in quality after increasing turbidity to 3.6 NTU.*

C. Other methods

A large variety of other methods have been used for increased quality of stereo reconstruction. One approach is to enhance the quality of the raw images taken from the cameras then passing those improved images for standard stereo reconstruction [14]. Similar methods include denoising and applying a series of filters and adjustments to the image, but generally require fine tuning, and long computation time to produce reconstructions [23]. The use of confocal imaging is another approach which utilizes an array of projectors and cameras to improve viability in turbid situations [24].

One challenge stereo methods face is monotonic textures, which is abundant in underwater settings. The lack of distinct features degrades the performance of standard stereo methods as it is difficult to find matching features. A projector can remedy the problem by creating an artificial texture on the scene. Detry et. al. used such a method to by projecting random binary blocked pattern into the scene to generate artificial texture for stereopsis, which also has the advantage of only requiring a single image capture, making it ideal for mobile robotic platforms [7]. 3D reconstruction was tested using block matching and semi-global block matching [25]. While this method was shown to be robust in moderate turbidity conditions (NTU 1 - 2.5), We tested this method in even higher turbidity conditions and found a drastic decrease in coverage. See Figure 2 for the degradation of this method in higher turbidity. This paper examines methods to improve upon this method, especially in higher turbidity situations.

III. Stereo with Image Stripes

One of the main issues for stereo systems in highly turbid scenarios is the large blur and backscatter of light due to reflection of particulates in the water. This phenomena is seen in Figure 2, where the underwater conditions result in a drastic reduction in quality of the image. Methods of partial illumination have seen success in creating reconstructions underwater and show promise in increasing quality under turbid situations [22]. An advantage of partial illumination is the reduced volume of illuminated water which is beneficial in reducing backscatter in turbid situations. We sought to overcome the increased backscatter from the assisted stereo random binary pattern method by instead projecting individual portions of the pattern onto the scene, in the form of light stripes. The partial illumination of the scene results in a reduction of illuminated liquid volume, reducing backscattered light. We aim to control the projected pattern to produce higher quality reconstructions in turbid situations.

We tested the use of image stripes in two variations, a multi capture where we split the original binary block pattern of [7] into 10 stripes, and another where we project a sequence of pure white stripes across the scene, similar to that done by Narasimhan et. al. [22]. By only projecting a stripe, the total volume of water lit by the project is greatly reduced, which will reduce the amount of backscattered light. See Figure 4 for the underwater reconstruction at a turbidity of about 3.62 NTU. Here we see the various strengths of the stripe methods compared to the baseline method. The resulting disparity map is more dense, resulting in higher coverage of the scene and better accuracy.

In order to combine the set of images from individual captures where a single stripe was illuminated, we had to stitch the images together. This was done through taking the disparity at each pixel location, and averaging the existing disparities if any existed at all. This method smooths the edges between stripes and creates a full point cloud utilizing the data collected from each of the stripes. Figure 4 visually shows a promising result of the greater improvement in 3D reconstruction in these water turbidites.

IV. Experiments

We tested the stripped random binary stereo method and single stripped methods as well as the baseline assisted stereo projection method underwater in a range of turbidity conditions on a example task panel seen in Figure 3. The experimental procedure is as follows:

- The robot frame including the vision system and task panel is placed in the water tank.
- Muddy water is mixed into tank, creating a highly turbid environment. A homogeneous environment was ensured by circulating the water with a pump.
- The data collection script is started which collects a set of images once every 11 minutes. Data is collected for shutter speeds of both 0.04 seconds and 0.16 seconds. At each time interval the following data is collected.
  - Turbidity data

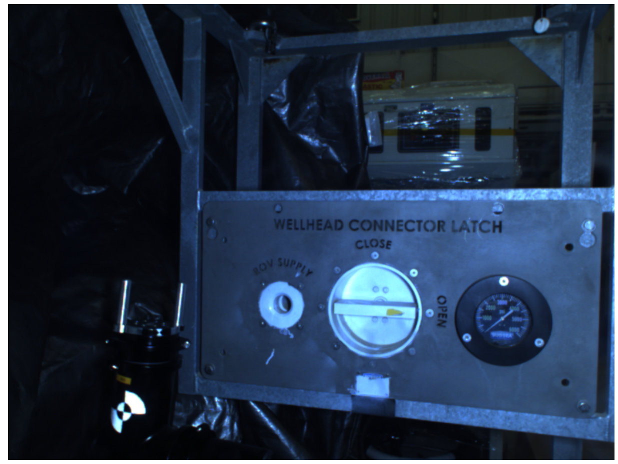

Fig. 3. *This is the example task panel used in testing the quality of 3D reconstruction methods.*

  - Random blocked pattern images
  - Striped random blocked pattern images
  - Single stripe images

As the dirt settles, turbidity decreases resulting in data collected in a wide range of turbidity values. Turbidity readings were provided with a spectrometer by Scan Messtechnik GmbH, which was located in close proximity to the stereo camera pair.

Block matching is used for stereo 3D reconstruction. Fixed shutter speeds of 0.04s and 0.16s were both used to collect image data at each time interval. A custom gain controller was used that automatically adjusted the camera gain until 0.001 present of the pixels in the scene were fully saturated. This was done to provide the best possible visibility of the task panel.

For the individual white stripes, we projected stripes that were 16 projection pixels wide, with each subsequent stripe being 16 pixels away from the previous. This was done to be more efficient in the image collection and by requiring less stripes. This still achieves a complete reconstruction as the stripes will blur across the medium, illuminating more area than is directly projected on. In addition, the nature of the block matching algorithm finds matching features surrounding the images stripe. 57 stripe images are required to cover the entire scene with the white stripes. For the patterned binary stripes, we used 10 total images.

V. Results

We were able to improve 3D underwater scene reconstruction by projecting individual stripes against the scene. Not only does this decrease backscatter, resulting in a cleaner image, but the stereo algorithm has an easier time identifying features and making correspondences. The results of the experiment can be seen in Figure 5 where the produced reconstruction was compared against the ground truth point cloud of the task panel. Each method was tested at shutter speeds of 0.04 seconds and 0.16 seconds. It is observed

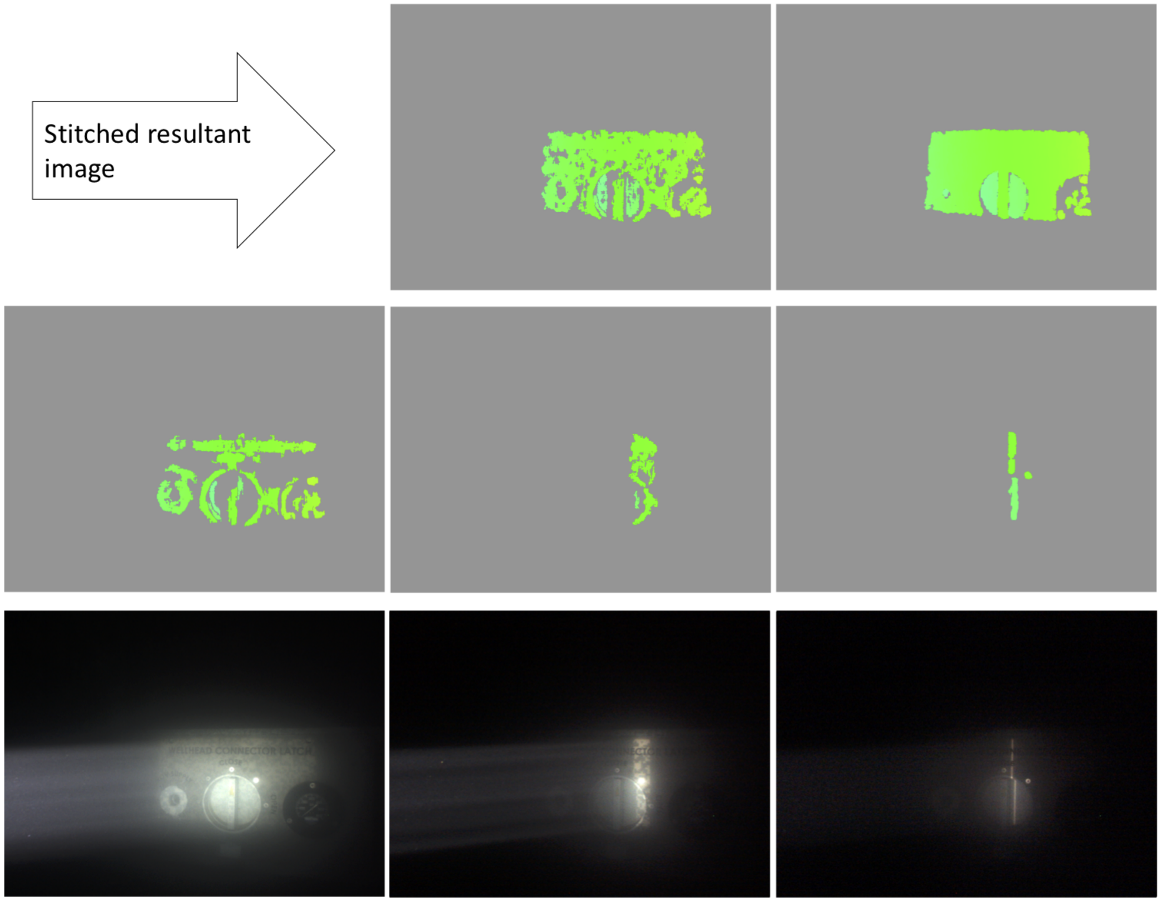

Fig. 4. *Underwater images at a turbidity of NTU 3.62 comparing the reconstructions of the stripe method. The left most image is with a random binary pattern being projected onto the scene. The middle column is using 10 stripes of the blocked pattern. The right most image uses stripes of 16 pixels wide. The top row is the stitched disparity from all of the stripes. The middle row is the disparity map from the image set of the row below.*

that the stripe methods have better performance than the original single shot assisted stereo method. This is seen mostly in the Mean Squared Error. For the shutter speed of 0.04 seconds, mean square error is lower for the stripe methods in a turbidity range of about 5 FTU to 6 FTU, and for a shutter speed of 0.16 seconds, the range of increase performance is about 5.5 to 6.5 FTU. Note that as turbidity decreases, the errors converge, as the stripe methods in these does not improve the reconstruction itself, it merely makes it easier for the system to see through turbid waters. We also see a small improvement in coverage of the scene which is defined in the work of Detry et. al., an improvement from 15-20% within the same range of improved accuracy [7].

Projecting individual stripes created better 3D reconstructions than the original binary pattern single shot method. See Figure 4, for the resultant disparity map comparisons. The striping method allows a more dense and accurate disparity map over the scene. The first major observation is that the full binary image compared to the striped binary image uses the exact same projected image, with the exception of the striped image having a smaller stripe of it. Even so, the reduction in projection volume results in a crisper and more distinguishable image, as the black and white components can be distinguished more easily. A single stripe provides the advantage of providing a unique feature in the image for stereo block matching, as the single stripe in the scene is easier to differentiate. This results in better stereo matching, as is seen though the more closely packed disparity map on the projected area. This is seen most clearly in Figure 4. Both the middle image in the first column and second column are taken at the same turbidity, but the reconstruction for the partially illuminated case renders a cleaner reconstruction.

VI. Discussion

The purpose of this work is to explore methodologies to effectively increase the quality of 3D reconstructions in turbid settings and test the limits of standard systems in extreme conditions. Our experiments show that using single stripe methods helps reduce backscatter of light and provides better stereo images due to overall reduction of light from the projector. Our results also show that individual stripes provide more accurate images, and have higher coverage.

This could provide beneficial use cases in underwater scenarios in extreme turbidity, allowing the robot to navigate and complete tasks in critical situations.

The methods may be limited due to gain saturation and projector power. The projector used is an off the shelf projector capable of projecting up to 600 lumens, and so a more luminous projector could have provided better results, and greater viability in higher turbidites. In many of the existing images at high turbidites, it was very difficult to make out the individual stripes despite saturating the camera gain, a problem that can be overcome with a stronger projector, or adjusting other parameters such as shutter speed.

A cost benefit analysis on the needs of an application must be considered before using a stripe method. Depending on the turbidity and stability of the vision system with respect to the environment, a multi shot method like the image stripes could results in lower performance reconstructions, or reconstructions that are negligibly better. Acquisition time of the total image is another consideration. In comparing acquisition time, the single stripe methods are much longer, especially at lower shutter speeds. A byproduct of this is the need for a longer shutter speed because of the light intensity of the project. The limitations on hardware result in a dimly lit image. In addition, the stripped methods is a multi shot approach, which requires the camera system to stationary when reconstructing the entire scene.

This method show promising results and can be helpful in creating 3D reconstructions in high turbid scenarios. However, this method will have only negligible benefits with a downside of higher acquisition time in lower turbid scenarios. A user could create an adaptive method to adjust the method based on turbidity readings to overcome this. Additionally the stripe method can be used on a specific region on interest after acquiring a more general lower quality reconstruction. This would allow a compromise between speed and accuracy.

VII. Future Work

Further exploration in the various patterns to project onto the scene, as well as how to improve the acquisition time to performance ratio needs additional study. More effective patterns can be used to more efficiently cover the entire scene to allow for faster acquisition. Often times only a portion of the scene is of interest. Utilizing partial illumination to search for the area of interest is a natural extension of this work. A major disadvantage of this method is the need for multiple captures. This can be greatly relieved by integrating partial illumination techniques with SLAM algorithms to allow for continuous reconstruction of the scene. This requires scene understanding of in order to best partially illuminate the scene to provide the best reconstruction.

Acknowledgments

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Copyright 2019 California Institute

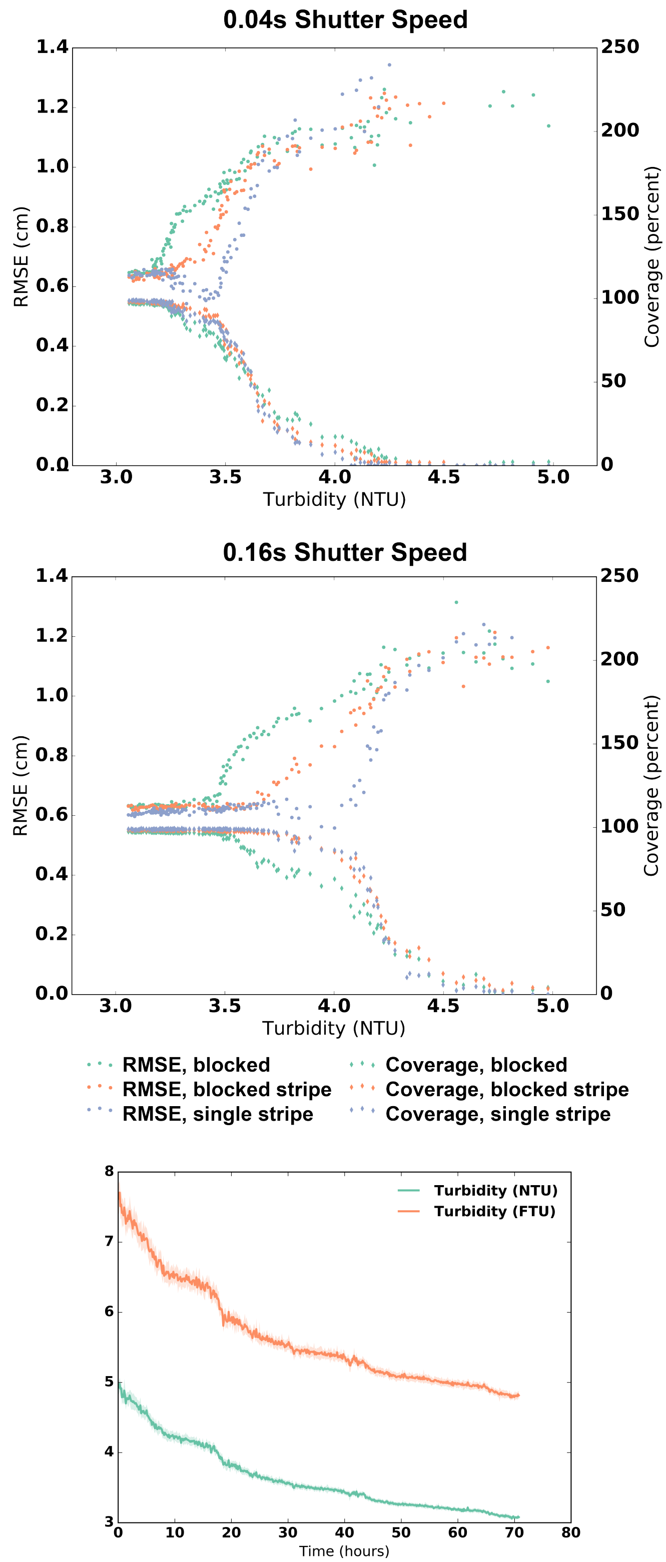

Fig. 5. *Plots from the underwater test. The top two images show RMS and Coverage as a function of turbidity for shutter speeds of 0.04 and 0.16 seconds. The bottom image shows the turbidity measurements throughout the 70 hours of testing.*

of Technology. U.S Government sponsorship acknowledged. The authors would like to thank Hanumant Singh for his enthusiastic insights on underwater structured light.

REFERENCES

[1] D. P. Kingma, D. J. Rezende, S. Mohamed, and M. Welling, "Semi-supervised learning with deep generative models," *CoRR*, vol. abs/1406.5298, 2014. [Online]. Available: http://arxiv.org/abs/1406.5298

[2] M. Livne and D. J. Fleet, "Tzk: Flow-based conditional generative model," *CoRR*, vol. abs/1902.01893, 2019. [Online]. Available: http://arxiv.org/abs/1902.01893

[3] A. Bodenmann, B. Thornton, R. Nakajima, H. Yamamoto, and T. Ura, "Wide area 3d seafloor reconstruction and its application to sea fauna density mapping," 09 2013.

[4] C. Mai, S. Pedersen, L. Hansen, K. L. Jepsen, and Z. Yang, "Subsea infrastructure inspection: A review study," in *2016 IEEE International Conference on Underwater System Technology: Theory and Applications (USYS)*, Dec 2016, pp. 71–76.

[5] J. Koch, T. Pailevanian, M. Garrett, C. Yahnker, R. Detry, D. Levine, and M. Gildner, "Development of a robotic limb for underwater mobile manipulation," in *MTS/IEEE OCEANS*, 2018.

[6] E. J. Sumner and C. K. Paull, "Swept away by a turbidity current in mendocino submarine canyon, california," *Geophysical Research Letters*, vol. 41, no. 21, pp. 7611–7618, 2014.

[7] R. Detry, J. Koch, T. Pailevanian, M. Garrett, D. Levine, C. Yahnker, and M. Gildner, "Turbid-water subsea infrastructure 3d reconstruction with assisted stereo," in *MTS/IEEE OCEANS*, 2018.

[8] F. Jenkins and H. White, *Fundamentals of Optics*. Tata McGraw-Hill Education, 1937.

[9] A. Jordt, "Underwater 3d reconstruction based on physical models for refraction and underwater light propagation," 2014.

[10] M. L. R. A. M. M. R. Markus-Christian Amann, Thierry M. Bosch, "Laser ranging: a critical review of unusual techniques for distance measurement," *Optical Engineering*, vol. 40, pp. 40 – 40 – 10, 2001.

[11] M. Massot-Campos and G. Oliver, "Optical sensors and methods for underwater 3d reconstruction," in *Sensors*, 2015.

[22] S. G. Narasimhan, S. K. Nayar, B. Sun, and S. J. Koppal, "Structured light in scattering media," in *Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 - Volume 01*, ser. ICCV '05. Washington, DC, USA: IEEE Computer Society, 2005, pp. 420–427.

[12] O. Fabio, K. Fabjan, L. R. Dario, A. Jacopo, and C. Stefano, "Performance evaluation of a low-cost stereo vision system for underwater object detection," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3388 – 3394, 2014, 19th IFAC World Congress.

[13] S. B. W. Matthew Johnson-Roberson, Oscar Pizarro and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *J. Field Robotics*, vol. 27, pp. 21–51, 01 2010.

[14] 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.

[15] A. Jordt-Sedlazeck, K. Kser, and R. Koch, "3d reconstruction based on underwater video from rov kiel 6000 considering underwater imaging conditions," 05 2009.

[16] K. Herakleous and C. Poullis, "3dunderworld-sls: An open-source structured-light scanning system for rapid geometry acquisition," *CoRR*, 2014.

[17] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. Razionale, "Experimentation of structured light and stereo vision for underwater 3d reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 4, pp. 508 – 518, 2011.

[18] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *2012 Second International Conference on 3D Imaging, Modeling, Processing, Visualization Transmission*, Oct 2012, pp. 464–471.

[19] J. Geng, "Structured-light 3d surface imaging: atutorial," *Adv. Opt. Photon.*, vol. 3, no. 2, pp. 128–160, Jun 2011.

[20] M. Massot-Campos and G. Oliver, "One-shot underwater 3d reconstruction," 09 2014.

[21] C. Roman, G. Inglis, and J. Rutter, "Application of structured light imaging for high resolution mapping of underwater archaeological sites," in *OCEANS'10 IEEE SYDNEY*, 2010, pp. 1–9.

[23] C. J. Prabhakar and P. U. P. Kumar, "3d surface reconstruction of underwater objects," *CoRR*, 2012.

[24] H. S. Marc Levoy, "Improving underwater vision using confocal imaging," 2009.

[25] H. Hirschmller, "Stereo processing by semi-global matching and mutual information," *in IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 328–341, 02 2008.