On Understanding the Challenges in Vision-Based Shipwreck Mapping

A. Quattrini Li A. Coskun S. M. Doherty S. Ghasemlou A. S. Jagtap
M. Modasshir S. Rahman A. Singh M. Xanthidis J. M. O'Kane I. Rekleitis
Computer Science and Engineering Department, University of South Carolina

1 Abstract

State estimation is one of the most significant challenges in robot autonomy, especially, the dual problems of tracking the pose of the robot and of mapping the environment as the robot moves. In the last decade, the wide availability of camera sensors, coupled with progress in computer vision, has given rise to a variety of vision-based techniques for these problems, known as *visual odometry* or *visual SLAM*. Scaramuzza and Fraundorfer [1,2] presented a comprehensive overview this work, from the fundamentals of Visual Odometry to recent research challenges and applications.

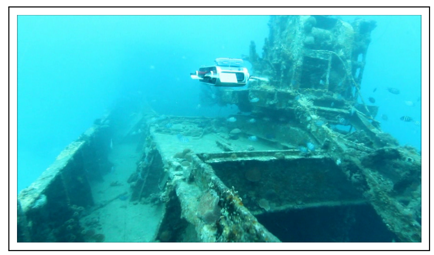

Fig. 1: Aqua robot at Pamir shipwreck, Barbados.

As a result, a lot of research papers and open source packages have been published, supported by impressive demonstrations. However, applying any of these packages on a new dataset has been proven extremely challenging, because of two main factors: software engineering challenges, such as lack of documentation, compilation, and algorithmic limitations, such as, special initialization motions, number of and sensitivity to parameters. In addition, most packages are usually developed and tested with urban settings in mind, indoor or outdoor.

In this paper, we discuss the current challenges that open-source vision-based state estimation packages face in the marine domain, and specifically in shipwreck mapping. We also present good practices for replicable results, by performing a comparison between some open source packages on underwater datasets.

The historical shipwrecks of the east coast of USA and in the Caribbean sea are an important part of history; however, they are deteriorating due to warm, salt water, human interference, and extreme weather (frequent tropical storms). Constructing accurate models of these sites will be extremely valuable not only for the historical study of the shipwrecks, but also for monitoring subsequent deterioration. Currently, limited mapping efforts are performed by divers that need to take measurements manually using a grid and measuring tape, or using handheld sensors [3]—a tedious, slow, and sometimes dangerous task. Automating such a task with underwater robots equipped with cameras—e.g., Aqua [4]—would be extremely valuable. Several datasets—see Fig. 2—have been collected over the years inside and outside shipwrecks off the coast of Barbados—i.e., Pamir, Elion, and Bajan Queen shipwrecks—by deploying an Aqua robot, an AUV equipped with an IMU and front and back cameras, or by employing a GoPro 3D Dual Hero System with two GoPro Hero3+ cameras operated by a scuba diver. Being collected at different times, the datasets vary in terms of conditions, including lighting, loss of contrast, motion type, covering a broad spectrum of situations in which the robot might find itself. These datasets are stored using the "rosbag" format[1], so that they can easily exported and processed. The datasets and detailed instructions on the usage of each package will be made available online so they could be tested and evaluated. Note that, although the ground truth is not available, the datasets were annotated with the rough estimate of the trajectory when they were collected.

Six of the most promising open source packages are tested on the datasets presented above to highlight some of the challenges in shipwreck mapping. The packages have been selected in order to cover different types of state-of-

[1] `http://www.ros.org/wiki/rosbag`

Fig. 2: Characteristic images from the evaluated datasets. AUV outside and inside wreck, Manual underwater outside and inside wreck.

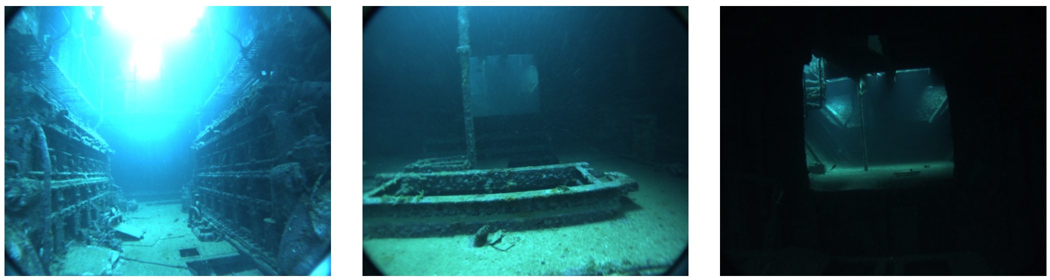

Fig. 3: Sample of images captured by Aqua robot inside Bajan Queen shipwreck in an interval of 30 seconds.

the-art techniques—i.e., feature-based methods: PTAM [5], ORB-SLAM [6]; semi-direct method: SVO [7]; direct methods: LSD-SLAM [8]; neural-based methods: RatSLAM [9]; and global optimization: g^2o [10]. We run several tests reasonably tuning the specific parameters following all available suggestions from the packages' authors.

In preliminary experiments, ORB-SLAM has shown the most promising results, being able to track features over time, while LSD-SLAM, being a method based on optical flow, is more affected by illumination changes. RatSLAM utilizes a learning process for adjusting how neurons are triggered, thus it needs the robot to visit the same place multiple times to improve the trajectory. SVO and PTAM work reasonably well in a small area resulting in a correct partial trajectory in the tested datasets.

While shipwrecks are quite rich in texture allowing the methods to reasonably track the camera, mapping them presents a different set of challenges compared to other scenarios, e.g., coral reef monitoring. Illumination can greatly vary in the scene, while the robot is moving, resulting in most cases in a loss of localization; see Fig. 3. Thus, it would be interesting to apply some image restoration techniques, e.g., [11, 12].

Additionally, shipwrecks are biodiversity hotspots, attracting plenty of marine life. Fig. 4 shows a sample run of ORB-SLAM and LSD-SLAM on a GoPro dataset, where the diver held the camera facing downwards and hovering over Pamir shipwreck. Being a direct method, LSD-SLAM is more affected by the presence of fishes, compared to ORB-SLAM which is a feature-based method. Indeed, the trajectory from ORB-SLAM more accurately represents the perceived camera motion. However, the reconstruction given by a feature-based method is sparser. Note that features could be detected over fishes and tracked for some frames, leading to some errors.

This preliminary comparison highlights also some good practices for replicable and measurable research in the field. The availability of the code allows other researchers to test and possibly adapt the methods. However, there are many solutions from other research groups that presented great performance without releasing the code—e.g., [13], thus making it hard to evaluate and use. Another important aspect is that all of the methods have plenty parameters that need to be tuned, e.g., the number of tracked features and the number of RANSAC iterations. Since finding the optimal tuning is not easy, especially if the experimenter is not the developer, care should be put in specifying the recommended values for these parameters and the effects induced by their variation. Also, some packages require initial translation motion to initialize the SLAM algorithm—e.g., PTAM requires a lateral translation—thus it is fundamental to keep that in mind, when collecting data.

Ongoing work includes a quantitative evaluation of the packages, the study of the effects of changing parameters, and the investigation of more open-source packages on the same datasets.

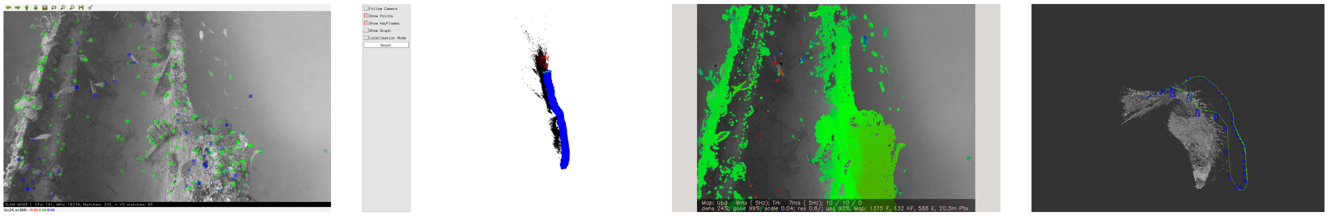

Fig. 4: Sample run of ORB-SLAM (first and second figures) and LSD-SLAM (third and fourth figure) on a footage collected with GoPro.

Acknowledgements: The authors would like to thank the generous support of a Google Faculty Research Award and of the National Science Foundation grant (NSF 1513203).

References

[1] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," *IEEE Robotics & Automation Magazine*, vol. 18, no. 4, pp. 80–92, 2011.

[2] F. Fraundorfer and D. Scaramuzza, "Visual odometry: Part II: Matching, robustness, optimization, and applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 78–90, 2012.

[3] J. Henderson, O. Pizarro, M. Johnson-Roberson, and I. Mahon, "Mapping submerged archaeological sites using stereo-vision photogrammetry," *International Journal of Nautical Archaeology*, vol. 42, no. 2, pp. 243–256, 2013.

[4] J. Sattar, G. Dudek, O. Chiu, I. Rekleitis, P. Giguere, A. Mills, N. Plamondon, C. Prahacs, Y. Girdhar, M. Nahon, and J.-P. Lobos, "Enabling autonomous capabilities in underwater robotics," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2008, pp. 3628–3634.

[5] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *IEEE and ACM Int. Symp. on Mixed and Augmented Reality*, 2007, pp. 225–234.

[6] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System," *IEEE Transactions on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[7] C. Forster, M. Pizzoli, and D. Scaramuzza, "SVO: Fast semi-direct monocular visual odometry," in *IEEE Int. Conf. on Robotics and Automation*, 2014, pp. 15–22.

[8] J. Engel, T. Schöps, and D. Cremers, "LSD-SLAM: Large-Scale Direct Monocular SLAM," in *European Conference on Computer Vision (ECCV)*, ser. Lecture Notes in Computer Science, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Springer, 2014, vol. 8690, pp. 834–849.

[9] D. Ball, S. Heath, J. Wiles, G. Wyeth, P. Corke, and M. Milford, "OpenRatSLAM: an open source brain-based SLAM system," *Autonomous Robots*, vol. 34, no. 3, pp. 149–176, 2013.

[10] R. Kummerle, G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard, "g^2o: A general framework for graph optimization," in *IEEE Int. Conf. on Robotics and Automation*, 2011, pp. 3607–3613.

[11] I. Vasilescu, C. Detweiler, and D. Rus, "Color-accurate underwater imaging using perceptual adaptive illumination," *Autonomous Robots*, vol. 31, no. 2, pp. 285–296, 2011.

[12] G. Bianco, M. Muzzupappa, F. Bruno, R. Garcia, and L. Neumann, "A new color correction method for underwater imaging," in *ISPRS/CIPA Workshop on Underwater 3D Recording and Modeling*, 2015.

[13] J. Hesch, D. Kottas, S. Bowman, and S. Roumeliotis, "Consistency Analysis and Improvement of Vision-aided Inertial Navigation," *IEEE Transactions on Robotics*, vol. 30, no. 1, pp. 158–176, 2014.