Underwater Cave Mapping using Stereo Vision

Nick Weidner, Sharmin Rahman, Alberto Quattrini Li, and Ioannis Rekleitis

Abstract— **This paper presents a systematic approach for the 3-D mapping of underwater caves. Exploration of underwater caves is very important for furthering our understanding of hydrogeology, managing efficiently water resources, and advancing our knowledge in marine archaeology. Underwater cave exploration by human divers however, is a tedious, labor intensive, extremely dangerous operation, and requires highly skilled people. As such, it is an excellent fit for robotic technology, which has never before been addressed. In addition to the underwater vision constraints, cave mapping presents extra challenges in the form of lack of natural illumination and harsh contrasts, resulting in failure for most of the state-of-the-art visual based state estimation packages. A new approach employing a stereo camera and a video-light is presented. Our approach utilizes the intersection of the cone of the video-light with the cave boundaries: walls, floor, and ceiling, resulting in the construction of a wire frame outline of the cave. Successive frames are combined using a state of the art visual odometry algorithm while simultaneously inferring scale through the stereo reconstruction. Results from experiments at a cave, part of the Sistema Camilo, Quintana Roo, Mexico, validate our approach. The cave wall reconstruction presented provides an immersive experience in 3-D.**

I. INTRODUCTION

The importance of underwater cave mapping spans several fields. First, it is crucial in monitoring and tracking groundwater flows in karstic aquifers. According to Ford and Williams [1] 25% of the world's population relies on karst water resources. Our work is motivated by the Woodville Karst Plain (WKP) which is a geomorphic region that extends from Central Leon County around the "Big Bend" of Florida [2]. Due to the significance of WKP, the Woodville Karst Plain Project (WKPP) has explored more than 34 miles of cave systems in Florida since 1987 [3], proving the cave system to be the longest in USA [4]. This region is an important source of drinking water and is also a sensitive and vulnerable ecosystem. There is much to learn from studying the dynamics of the water flowing through these caves. Volumetric modeling of these caves will give researchers a better perspective about their size, structure, and connectivity. These models have even greater importance than simply enhancing the mapping. Understanding the volume of the conduits and how that volume increases and decreases over space is a critical component to characterizing the volume of flow through the conduit system. Current measurements are limited to point-flow velocities of the cave metering system and a cross-sectional volume at that particular point. This paper presents a first step towards robotic mapping of an underwater cave. Fig. 1 shows an underwater cave environment. The proposed approach results in 3-D reconstructions which will give researchers the above described capabilities. Furthermore, volumetric models, will be incredibly helpful for those involved with environmental and agricultural studies throughout the area, and once perfected this technology could help map other subterranean water systems, as well as any 3-D environment that is difficult to map. The Woodville Karst Plain area is sensitive to seawater intrusions which threaten the agriculture and the availability of drinking water; for more details see the recent work by Zexuan et al. [5]. Second, detailed 3-D representations of underwater caves will provide insights to the hydrogeological processes that formed the caves. Finally, because several cave systems contain historical records dating to the prehistoric times, producing accurate maps will be valuable to underwater archaeologists.

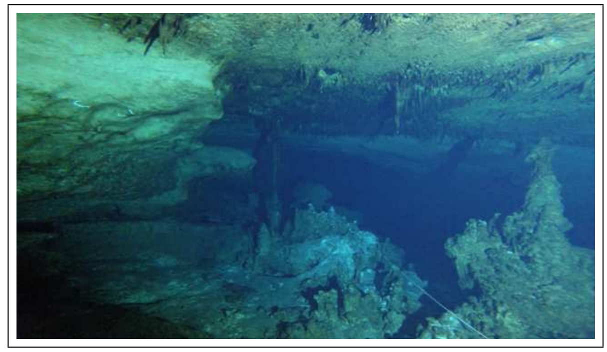


Fig. 1. Typical scene from an underwater cave.

Operations in underwater caves can be grouped under three categories: motion inside the known part of the cave; exploration of new territory; and surveying of newly explored areas. Most transportation in the explored part of caves is performed using diver propulsion vehicles (DPVs). All explored areas are marked by permanently attached cave line, which provides a direct route to the exit; see Fig. 2 where a diver is inspecting the line. When divers explore uncharted territory, they proceed without the DPVs, laying out line and tying it to protrusions on the floor, walls, or ceiling. The third phase, surveying, consists of two divers measuring distances, using a cave-line with knots every 3 m between attachment points. Simultaneously, the divers also measure the water depth at each attachment point, as well as the azimuth of the line leading to the next attachment point. All the information is recorded on a slate or waterproof paper. Estimates of the height and width of the passage can also be recorded, if time permits. The above described process is

*This work was supported by NSF (CRI:II-New 1513203).

The authors are with the Computer Science and Engineering Department, University of South Carolina, Columbia, SC, USA [weidnern,srahman]@email.sc.edu, [albertoq,yiannisr]@cse.sc.edu

Fig. 2. A cave diver attaching a branch line to the main line of the cave.

error-prone and time consuming, and at greater depths results in significant decompression times, where total dive time can reach between 15 to 28 hours per dive. This paper presents a first step of utilizing robotic technology to assist in cave exploration via the use of a stereo camera and a video-light. In many cases, during DPV rides, the divers attach cameras to their DPV and/or to themselves in order to document the exploration. Consequently, introducing a stereo camera, with a GoPro form-factor, does not complicate the standard operating procedures and does not increase the cognitive load of the divers.

The presented approach utilizes the presence of the artificial lighting to produce a rough model of the traversed area. In particular, the video-light cone is used to identify the walls of the cave from a single stereo pair. Furthermore, motion between consecutive stereo pairs is estimated and the 3-D reconstruction is utilized to produce an approximate volumetric map of the cave.

The next section discusses related work. Section III illustrates the challenges present in the underwater cave domain and presents an overview of the proposed approach. Experimental results from an underwater cave, part of the Sistema Camilo, Quintana Roo, Mexico, are presented in Section IV. The paper finishes with a discussion of lessons learned and an outline of future work.

II. Related Work

The majority of underwater mapping up to now consists of fly-overs with downward pointing sensors mapping the floor surface. The resulting representation consists of 2.5 dimensional mesh-maps or image mosaics with minimal structure in the third dimension. In addition to underwater caves, several other underwater environments exhibit prominent three dimensional structure. Shipwrecks, are significant historical sites. Producing accurate photorealistic 3-D models of these wrecks will assist in historical studies and also monitor their deterioration over time. Finally, underwater infrastructure inspection [6] is another dangerous and tedious task that is required to be performed at regular intervals. Such infrastructure includes bridges, hydroelectric dams [7], water supply systems [8], and oil rigs. For more information please refer to the Massot-Campos and Oliver-Codina survey [9] for an overview of 3-D sensing underwater.

Most of the underwater navigation algorithms [10]–[13] are based on acoustic sensors such as Doppler Velocity Log (DVL), ultra-short baseline (USBL) and sonar. Gary et. al. [14] presented a 3D model of a cenote using LIDAR and sonar data collected by DEPTHX (DEep Phreatic THermal eXplorer) vehicle having DVL, IMU and depth sensor for underwater navigation. Corke et. al. [15] compared acoustic and visual methods for underwater localization. However, collecting data using DVL, sonar, and USBL while diving is expensive and sometimes not suitable in cave environments.

Using stereo vision underwater has been proposed by several groups, however, most of the work has focused on open areas with natural lighting, or artificial light that completely illuminates the field of view. Small area dense reconstruction of a lit area was proposed by Brandou et al. [16]. Mahon et. al. [17] proposed a SLAM algorithm based on the viewpoint augmented navigation (VAN) using stereo vision and DVL in underwater environment. A framework proposed by Leone et al. [18] operated over mainly flat surfaces. Several research groups have investigated the mapping and/or inspection of a ship's hull using different techniques [19]–[22], the most famous shipwreck visual survey being that of the Titanic [23]. Error analysis was performed recently by Sedlazeck and Koch [24]. The problem of varying illumination was addressed by Nalpantidis et al. [25] for above-ground scenes in stereo reconstruction. More recently, Servos, Smart and Waslander [26] presented a stereo SLAM algorithm with refraction correction in order to address the transitions between water, plastic, and air that exist in the underwater domain.

III. 3-D Reconstruction using stereo video

A. Challenges

As can be seen in Fig. 3, the complete absence of natural illumination in combination with the presence of several sources of artificial illumination, such as: each diver's primary light and also one or more video-lights, results in huge lighting variations in the scene. In particular each diver's primary light generates a tightly focused beam which is constantly moving with the motion of the diver. In Fig. 4a, there are three divers present: one holding the video light, his tanks visible at the bottom of the image; one traveling with the camera, not visible; and a third one whose DPV is visible at the top of the image. The primary light of the third diver can be seen as a blue beam pointing downwards, starting at the left of the DPV.

The lighting variations make the success of traditional visual odometry [27] algorithms near impossible. The main assumption of Brightness Constancy Constraint underlying most visual odometry algorithms is violated by the constantly moving light-sources. Table I presents tests of five open source packages of vision based SLAM on underwater cave vision datasets; as expected most of them failed on the longer sequence and the rest were not able to extract the scale of the environment. It is worth noting that several of the packages are expecting specific motions in order to initialize [28]. Complete results are not presented due to space constraints; interested readers should refer to the work of Quattrini Li et al. [29] for a detailed analysis of more packages

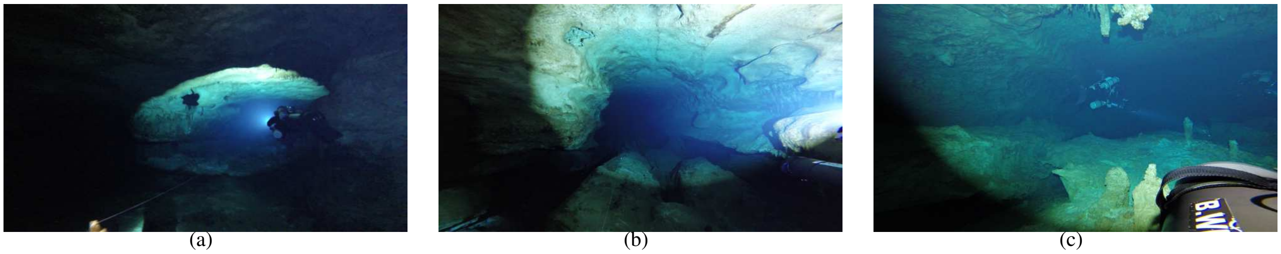

(a) (b) (c)

Fig. 3. Left camera images of an underwater cave with different illuminations. Illumination in the cave is provided by the lights individual divers have and also from a strong video-light. (a) Diver in front holds a strong video-light; see how the cone of light outlines the boundaries of the cave. (b) Diver with video-light follows behind the camera. (c) The diver with the camera also holds the light.

and a variety of datasets. The selection of the algorithms presented here was motivated of testing a variety of methods; feature based [30], [31], semi-direct [32], direct [33], and global [34]. The main challenge these algorithms face is the constant change of the field of view and the dramatic lighting variations resulting from occlusions and from light absorption over distance. Among the most successful was the ORB-SLAM [30] with its latest incarnation as ORB-SLAM 2, still in beta version, working with stereo images. While some of these packages, produced an acceptable trajectory, their shape reconstruction from the detected 3-D features was plagued by noise.

B. Wireframe reconstruction

Using light variations to infer shape has been used extensively in the past [35]–[37]. The 3-D reconstruction consists of several steps. First the images have to be rectified; a process achieved through a process called camera calibration.

Camera Calibration: While calibrating a camera is a well studied problem, the camera used (stereo Dual GoPro Hero[1]) presented us with a major challenge. By default the camera is utilizing a *SuperView* mode which stretches the wide angle image even further. Above water traditional Camera Calibrations packages such as the ones in MATLAB [2] and OpenCV[3] were unable to calibrate the camera. For underwater footage, the refraction of light through the water, the port, and finally the lens, resulted in partial elimination of the artificially introduced distortion of the *SuperView* mode. The image though was still distorted enough that OpenCV was unable to perform satisfactory calibration; several pixel calibration error. MATLAB in contrast, by selecting images through the complete viewing sphere, and rejecting images from the areas where images were already used, avoided overfitting and produced Camera Calibration with an error of 0.8 pixels. Utilizing the MATLAB produced internal parameters, the left (see Fig. 4(a)) and right images are rectified in order to remove the strong distortions from the wide angle lens of the GoPro camera; see Fig. 4(b) for the rectified image.

Contour Tracking: Adaptive thresholding is used in order to identify the areas with different illumination; see Fig. 4(c) for the thresholded image where the cone of the video-light meets the cave walls. Selecting the right value for thresholding the image required some domain knowledge, and currently was perform per video sequence, by a human. Current experiments consider adjusting the threshold based on keeping a balance between the amount of light and dark areas, but that work is outside the scope of this paper.

[1]http://gopro.com/
[2]http://www.mathworks.com/products/matlab/
[3]http://opencv.org/

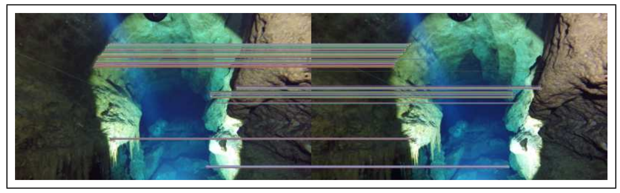


Fig. 5. Select features matched at the boundary between left and right image of a stereo pair.

During the next step, edge detection marks the boundaries between light and dark areas; see Fig. 4(d) for the boundaries of Fig. 4(b). The OpenCV Canny edge detector [38] is used to identify the edges marking the lighter area boundaries. As can be seen, the edge map is very noisy and thus not suitable for estimating the walls of the cave. A filter is applied to the contour list, eliminating short contours. More specifically, for every contour, its bounding box is calculated and then only the highest fifth percentile is kept. While this method can eliminate elongated contours, experiments with the actual underwater cave video footage proved to not affect the main boundaries. The filtered contours can be seen in Fig. 4(e). Figure 4(f) superimposes the filtered contours on the rectified image; the areas where the cone of light meets the cave walls are clearly identifiable. In addition, the area with acceptable lighting is extracted for use at the motion estimation. The edge map of the boundaries is used then as input to a stereo reconstruction algorithm.

Sparse Stereo Reconstruction: The 3-D structure of the cave boundaries is estimated for each stereo pair. For every point on the contour of the left image a SURF feature descriptor [39] is calculated using the left rectified image. Consequently, the same descriptor is matched on the right rectified image. Outlier rejection is facilitated by searching only locations at the same row and to the right of the left-image feature's coordinates. As the camera calibration error is 0.8 pixels, it justifies the assumption above. Previous work on feature quality [40]–[42] for underwater images indicated SURF [39] to be the most appropriate feature

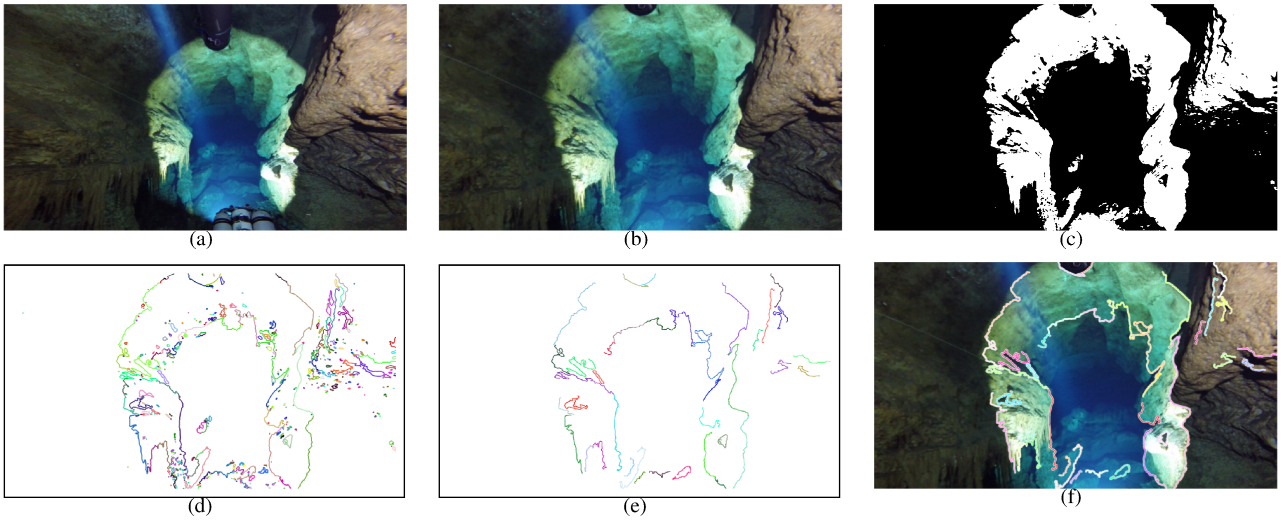


Fig. 4. (a) Left camera image of an underwater cave. (b) The rectified image. (c) The rectified image thresholded based on light intensity. (d) An edge map of the boundaries of the thresholded image. (e) The boundaries filtered to eliminate small contours. (f) The longer contours superimposed on the rectified image.

TABLE I

PERFORMANCE OF DIFFERENT OPEN SOURCE VISION BASED SLAM PACKAGES ON UNDERWATER DATA; FOR A DETAILED ANALYSIS PLEASE REFER TO QUATTRINI LI ET AL. [29]

	[30] ORB-SLAM	[31] PTAM	[32] SVO	[33] LSD-SLAM	[34] Colmap
10 sec	noisy	no initialization	partial trajectories/no scale	loss of track	partial trajectories/no scale
448 sec	noisy	no initialization	partial trajectories/no scale	loss of track	partial trajectories/no scale

descriptor. Furthermore, the OpenCV Canny edge detector groups the edges in a list of continuous contours, as such consecutive points belonging to the same contour can be filtered for consistency. Figure 5 presents select feature matches corresponding to the contours between the left and right image of a stereo pair.

Figure 6 presents a comparison of the performance of dense stereo reconstruction using OpenCV's semi-global block matching (SGBM) stereo algorithm [43] and the contour calculation. The standard output of dense stereo algorithms is a depth map, a normalized image where depth is quantified between 0 and 255; as such the values are discretized; see Figs. 6(a),6(d) for the 3-D reconstruction using the SGBM stereo algorithm on Fig. 4(b). The noise is quite noticeable, Figs. 6(b),6(e) present the same reconstruction using only the lighted areas. Finally, Figs. 6(c), 6(f) present only the contours of high intensity variation extracted from Fig. 4(b) projected in 3-D using SURF feature matching between left and right image. The noise is largely reduced, and the cave boundaries are clearly identifiable. While the first row of Fig. 6 presents a frontal view, and the error is not noticeable, the second row, presents a side view and the outliers are obvious.

C. Visual Odometry

Brute force application of VO algorithms [44], [45] is quite challenging in the underwater cave domain, due to the lighting variations and the sharp contrasts existing in the image, as discussed earlier. However, by thresholding the image, the areas of adequate illumination in the left and right camera feed can be used to apply one of the latest VO algorithms, ORB-SLAM 2[4], a variant of ORB-SLAM [30], [46] for stereo vision. Figure 7 presents tracked features in the areas with higher illumination. It is worth noting that during some segments of the video the third diver swimming below the camera exhaled sending a cloud of bubbles in the field of view, however, the VO algorithm was robust enough to handle these dynamic features. This event highlighted one of the challenges of underwater vision.

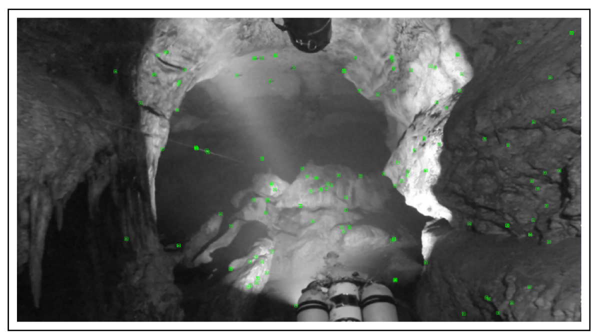


Fig. 7. ORB features tracked by ORB-SLAM 2.

Figure 8 presents the trajectory of the stereo camera and the 3-D position of stable features as extracted from ORB-SLAM 2 from a trajectory of seven minutes, twenty eight seconds. While there was no ground truth, observing the video one gets a qualitative verification for the estimated

[4]https://github.com/raulmur/ORB_SLAM2

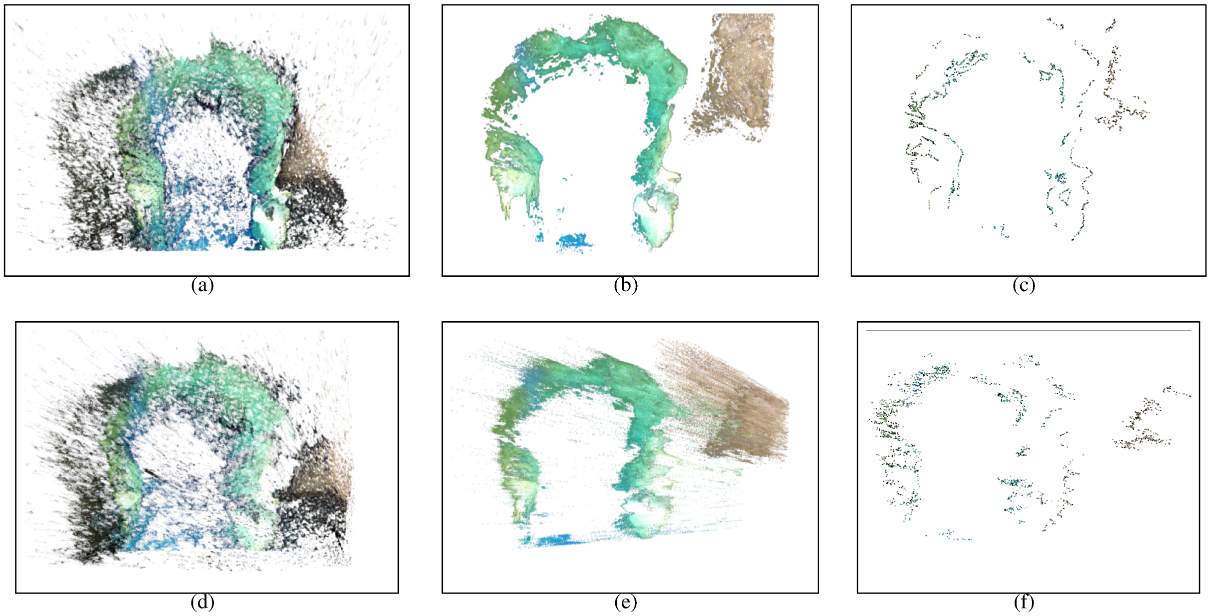


Fig. 6. Three different reconstructions from two different angles are presented. (a-c) Present a frontal view; (d-f) present a side view. (a), (d) Disparity map of the Fig. 4(b) using the the OpenCV's semi-global block matching (SGBM) stereo algorithm. (b), (e) Applying the SGBM algorithm only to the lighted part. (c), (f) The contour in 3-D using feature matches; see Fig. 5. It is worth noting elimination of outliers makes the contours much more distinct.

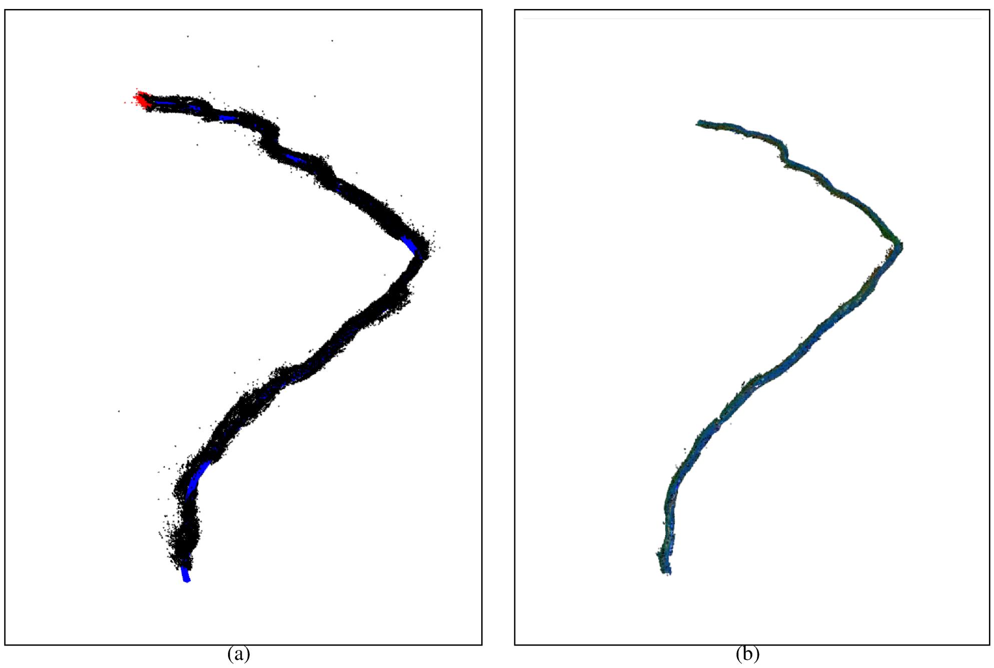


Fig. 8. (a) The trajectory calculated by ORB-SLAM 2 of a 7 min 28 sec traversal and the 3-D points estimated from ORB features. (b) The wireframe reconstructed from the proposed stereo algorithm. Please note, the reduced number of outliers compared to (a).

trajectory. The estimated trajectory is then used as an input to produce a volumetric map by transforming the boundaries calculated above through space using the estimated pose of the stereo camera at each instant. It is clear that the contour based reconstruction; see Fig. 8(a), has eliminated several outliers which were present in the ORB-SLAM reconstruction; see Fig. 8(b). The next section presents results from an actual cave.

IV. Experimental Results

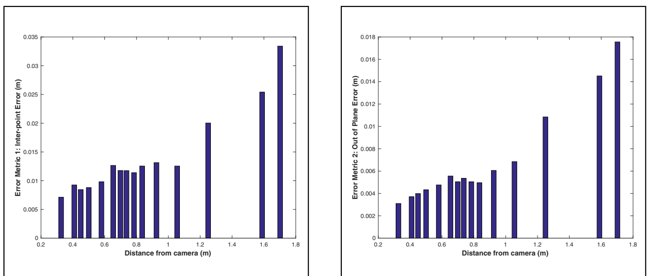
A. Experimental setup

In January 2015, the authors requested from a cave exploration team in Mexico to acquire sample footage using a Dual Hero stereo camera from GoPro during a dive at an already explored cave. The selected cave is part of the Sistema Camilo, the 11th longest submerged cave system in the world, located at Quintana Roo, Yucatan peninsula, Mexico.

The camera was mounted on a DPV and the video-light was carried in different configurations in order to demonstrate alternative lighting schemes.

B. Camera Calibration

As mentioned above, the stereo camera used utilizes a recording mode termed superview, which stretches the image in order to produce more aesthetically pleasing videos. Post-processing all the calibration footage collected, error analysis showed, as expected, the error to slightly increase with distance; see Fig. 9.



(a) (b)

Fig. 9. (a) Average error of the inter-point distance of the target; (b) Average error of the reconstructed points from the best plane fitting 3-D points of the checkerboard. The results were from 4,000 images of the calibration target presented to the stereo camera underwater.

C. Stereo reconstruction

Figure 8(b) presents the 3-D reconstruction of a long video of 7 min 28 sec. The structure corresponds with the cave morphology, however it is difficult to discern in the still image. The accompanying video presents a fly-through the cave. Figure 10 presents the 3-D reconstruction of a cave segment from a short ten seconds traversal. The left and right walls are clearly identifiable, while the floor and ceiling are occluded from the two divers that swam in the field of view.

V. CONCLUSIONS

This paper presented first ever reconstruction results from an underwater cave using a novel approach utilizing the artificial lighting of the scene as a tool to map the boundaries. The proposed technique was applied on real stereo video footage from a cave in Mexico, where an exploration team collected visual data using a light in different configurations. Central to our approach was the strategy of minimum interference with the standard procedures of the dive team. As cave diving is considered one of the most extreme activities, increasing the cognitive load, or hampering the functionality of the teams equipment was out of the question.

We are currently working on developing a stereo camera/light configuration that will produce the best reconstruction results without interfering with the operations of the divers. It is worth noting that in the presented experiments the video-light and the camera were carried by different divers thus constantly changing their relative pose. Selecting appropriate lighting and fixing it to the camera has proven to be a challenging task. We will continue to experimentally test different VO algorithms on the stereo footage and adapt the most promising ones to operate inside the segmented part of the image with adequate illumination, ensuring accurate pose estimation for consecutive stereo image pairs. Future work will consider the characterization of shadows, other divers, and dark areas due to light absorption which will be eliminated from the shape calculations.

The final result of the discussed work will be an algorithmic solution producing a volumetric map of the cave and an estimate of the camera's trajectory. The proposed approach will advance the state of the art for Visual SLAM [47], [48] in extreme conditions.

ACKNOWLEDGEMENT

The authors are indebted to the cave divers: John Rose, Steve Cox, Mark Garland, and Blake Wilson for collecting the video data. Part of this research was supported through a Magellan scholarship from the University of South Carolina. We acknowledge support for this work from a Google Faculty Research Award and the National Science Foundation grants (NSF 1513203, 1637876).

REFERENCES

[1] D. C. Ford and P. W. Williams, *Karst Geomorphology and Hydrology*. Chapman & Hall, 1994.

[2] E. Lane, "The Spring Creek Submarine Springs Group, Wakulla County, Florida," Florida Geological Survey, Tallahasee, Fl, Tech. Rep. Special Publication 47, 2001.

[3] C. McKinlay, "Woodville Karst Plain Project (WKPP)," URL:http://www.wkpp.org, Apr. 2015.

[4] B. Gulden, "World longest underwater caves," URL:http://www.caverbob.com/uwcaves.htm, Apr. 2015.

[5] Z. Xu, S. W. Bassett, B. Hu, and S. B. Dyer, "Long distance seawater intrusion through a karst conduit network in the Woodville Karst Plain, Florida," *Scientific Reports*, vol. 6, pp. 1–10, Aug 2016.

[6] D. Ribas, P. Ridao, and J. Domingo, "Underwater SLAM in Man-Made Structured," *J. of Field Robotics*, vol. 25, pp. 898–921, 2008.

[7] P. Ridao, M. Carreras, D. Ribas, and R. Garcia, "Visual inspection of hydroelectric dams using an autonomous underwater vehicle," *J. of Field Robotics*, vol. 27, no. 6, p. 759 778, 2010.

[8] C. White, D. Hiranandani, C. S. Olstad, K. Buhagiar, T. Gambin, and C. M. Clark, "The Malta cistern mapping project: Underwater robot mapping and localization within ancient tunnel systems," *J. of Field Robotics*, vol. 27, no. 4, p. 399 411, 2010.

[9] M. Massot-Campos and G. Oliver-Codina, "Optical sensors and methods for underwater 3d reconstruction," *Sensors*, vol. 15, no. 12, pp. 31 525–31 557, 2015.

[10] C.-M. Lee, P.-M. Lee, S.-W. Hong, S.-M. Kim, and W. Seong, "Underwater navigation system based on inertial sensor and doppler velocity log using indirect feedback kalman filter," *Int. J. of Offshore and Polar Engineering*, vol. 15, no. 02, 2005.

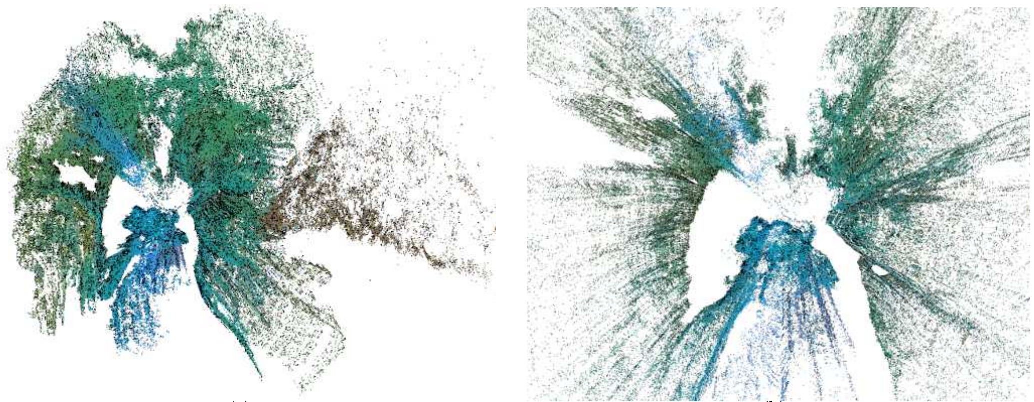
[11] J. Snyder, "Doppler Velocity Log (DVL) navigation for observation-class ROVs," in *MTS/IEEE SEATTLE OCEANS*, 2010, pp. 1–9.

[12] H. Johannsson, M. Kaess, B. Englot, F. Hover, and J. Leonard, "Imaging sonar-aided navigation for autonomous underwater harbor surveillance," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2010, pp. 4396–4403.

[13] P. Rigby, O. Pizarro, and S. B. Williams, "Towards geo-referenced AUV navigation through fusion of USBL and DVL measurements," in *MTS/IEEE OCEANS*, 2006, pp. 1–6.

[14] M. Gary, N. Fairfield, W. C. Stone, D. Wettergreen, G. Kantor, and J. M. Sharp Jr, "3d mapping and characterization of sistema zacatón from depthx (deep phreatic thermal explorer)," in *Proc. of KARST08: 11th Sinkhole Conference ASCE*, 2008.

[15] P. Corke, C. Detweiler, M. Dunbabin, M. Hamilton, D. Rus, and I. Vasilescu, "Experiments with underwater robot localization and tracking," in *Proc. IEEE Int. Conf. on Robotics and Automation*, 2007, pp. 4556–4561.



(a) (b)

Fig. 10. (a) The 3-D walls of the cave extracted from a short ten second traversal. (b) A second view more inside the reconstruction.

[16] V. Brandou, A. G. Allais, M. Perrier, E. Malis, P. Rives, J. Sarrazin, and P. M. Sarradin, "3-D Reconstruction of Natural Underwater Scenes Using the Stereovision System IRIS," in *MTS/IEEE OCEANS 2007 - Europe*, June 2007, pp. 1–6.

[17] I. Mahon, S. B. Williams, O. Pizarro, and M. Johnson-Roberson, "Efficient view-based SLAM using visual loop closures," *IEEE Trans. on Robotics*, vol. 24, no. 5, pp. 1002–1014, 2008.

[18] A. Leone, G. Diraco, and C. Distante, *Stereo Vision*. InTech, ISBN: 978-953-7619-22-0, 2008, ch. 11. A Stereo Vision Framework for 3-D Underwater Mosaicking, pp. 173–196.

[19] A. Hogue, A. German, and M. Jenkin, "Underwater environment reconstruction using stereo and inertial data," in *IEEE Int. Conf. on Systems, Man and Cybernetics, ISIC*, Oct 2007, pp. 2372–2377.

[20] F. S. Hover, J. Vaganay, M. Elkins, S. Willcox, V. Polidoro, J. Morash, R. Damus, and S. Desset, "A Vehicle System for Autonomous Relative Survey of In-Water Ships," *Marine Technology Society J.*, vol. 41, no. 2, pp. 44–55, June 2007.

[21] B. Englot and F. S. Hover, "Three-dimensional coverage planning for an underwater inspection robot," *The Int. J. of Robotics Research*, vol. 32, no. 9-10, pp. 1048–1073, Sept. 2013.

[22] A. Kim and R. M. Eustice, "Real-time visual SLAM for autonomous underwater hull inspection using visual saliency," *IEEE Trans. on Robotics*, vol. 29, no. 3, pp. 719–733, 2013.

[23] R. Eustice, H. Singh, J. Leonard, and M. Walter, "Visually Mapping the RMS Titanic: Conservative Covariance Estimates for SLAM Information Filters," *Int. J. of Robotics Research*, pp. 1223–1242, 2006.

[24] A. Sedlazeck and R. Koch, "Perspective and non-perspective camera models in underwater imaging overview and error analysis," in *Outdoor and Large-Scale Real-World Scene Analysis*. Springer Berlin Heidelberg, 2012, vol. 7474, pp. 212–242.

[25] L. Nalpantidis and A. Gasteratos, "Stereo vision for robotic applications in the presence of non-ideal lighting conditions," *Image and Vision Computing*, vol. 28, no. 6, pp. 940 – 951, 2010.

[26] J. Servos, M. Smart, and S. L. Waslander, "Underwater stereo SLAM with refraction correction," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Nov 2013, pp. 3350–3355.

[27] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, vol. 1, 2004.

[28] G. Klein and D. Murray, "Parallel tracking and mapping for small AR workspaces," in *Proc. Sixth IEEE and ACM Int. Symposium on Mixed and Augmented Reality*, 2007, pp. 225–234.

[29] A. Q. Li, A. Coskun, S. M. Doherty, S. Ghasemlou, A. S. Jagtap, M. Modasshir, S. Rahman, A. Singh, M. Xanthidis, J. M. O'Kane, and I. Rekleitis, "Experimental Comparison of open source Vision based State Estimation Algorithms," in *Int. Symposium of Experimental Robotics (ISER)*, Tokyo, Japan, 2016.

[30] R. Mur-Artal and J. D. Tardos, "ORB-SLAM: Tracking and Mapping Recognizable Features," in *Robotics: Science and Systems (RSS) Workshop on Multi VIew Geometry in RObotics (MVIGRO)*, 2014.

[31] G. Klein and D. Murray, "Parallel tracking and mapping for small ar workspaces," in *ACM Int. Symposium on Mixed and Augmented Reality, 2007. ISMAR 2007. 6th and*, Nov 2007, pp. 225–234.

[32] 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.

[33] J. Engel, T. Schops, and D. Cremers, "LSD-SLAM: Large-Scale Direct Monocular SLAM," in *Eur. Conf. on Computer Vision*, 2014, pp. 834–849.

[34] J. L. Schönberger and J.-M. Frahm, "Structure-from-Motion Revisited," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 2016.

[35] J. R. Kender and E. M. Smith, "Shape from darkness; deriving surface information from dynamic shadows," in *Proc. AAAI National Conf. on Artificial Intelligence (AAAI)*, 1986, pp. 664–669.

[36] M. Daum and G. Dudek, "On 3-D Surface Reconstruction Using Shape from Shadows," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 1998, pp. 461–468.

[37] K. N. Kutulakos and S. M. Seitz, "A theory of shape by space carving," *Int. J. of Computer Vision*, vol. 38, no. 3, pp. 199–218, 2000.

[38] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, pp. 679–698, 1986.

[39] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded up robust features," in *European Conf. on Computer Vision*, 2006, pp. 404–417.

[40] J. Shi and C. Tomasi, "Good features to track," in *IEEE Conf. on Computer Vision and Pattern Recognition*, 1994, pp. 593 – 600.

[41] F. Shkurti, I. Rekleitis, and G. Dudek, "Feature tracking evaluation for pose estimation in underwater environments," in *Conf. on Computer and Robot Vision*, 2011, pp. 160–167.

[42] M. Xanthidis, A. Q. Li, and I. Rekleitis, "Shallow coral reef surveying by inexpensive drifters," in *MTS/IEEE Oceans*, 2016, pp. 1–9.

[43] H. Hirschmüller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 328–341, 2008.

[44] S. S. da Costa Botelho, P. Drews, G. L. Oliveira, and M. da Silva Figueiredo, "Visual odometry and mapping for underwater autonomous vehicles," in *Latin American Robotics Symposium*, 2009, pp. 1–6.

[45] P. Corke, C. Detweiler, M. Dunbabin, M. Hamilton, D. Rus, and I. Vasilescu, "Experiments with underwater robot localization and tracking," in *Proc. of IEEE Int. Conf. on Robotics and Automation*, 2007, pp. 4556–4561.

[46] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardós, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Trans. on Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[47] J. Salvi, Y. Petillot, S. Thomas, and J. Aulinas, "Visual SLAM for underwater vehicles using video velocity log and natural landmarks," in *OCEANS 2008*, sept. 2008, pp. 1 –6.

[48] S. Wirth, P. L. Negre Carrasco, and G. O. Codina, "Visual odometry for autonomous underwater vehicles," in *MTS/IEEE OCEANS-Bergen*, 2013, pp. 1–6.