

Underwater Laser-based Structured Light System for One-Shot 3D Reconstruction

Miquel Massot-Campos, Gabriel Oliver-Codina

Departament de Matemàtiques i Informàtica, Universitat de les Illes Balears

Cra. Valldemossa, km. 7.5 07122 Palma de Mallorca (Spain)

Email: {miquel.massot,goliver}@uib.es

Abstract—A Laser-based Structured Light System (LbSLS) has been designed to perform underwater close-range 3D reconstructions even with high turbidity conditions and outperform conventional systems. The system uses a camera and a 532 nm green laser projector. The optical technique used is based on the projection of a pattern obtained placing a Diffractive Optical Element (DOE) in front of the laser beam. In the experiments described in this manuscript, the DOE used diffracts the laser beam in 25 parallel lines providing enough information in a single camera frame to perform a 3D reconstruction.

I. INTRODUCTION

In recent years, 3D imaging sensors have become the novelty in several fields as in human-machine interaction, maps and films. To generate this data, a 3D sensor device has to be used to generate and construct the scene. In engineering, this workflow is known as 3D reconstruction, and nowadays is seen as a new useful tool in Computer Science. It is widely used, from topological maps as Google Maps, to industrial engineering for the verification of assembly parts.

Underwater sensors with high 3D reconstruction resolution capabilities are still under research and development. These sensors are useful, for example, to increase the underwater robots manipulation autonomy or enabling them with robust 3D mapping capacity. The sensor will require some features and protection due to the hazardous environment and the ability to work under high turbidity conditions.

Underwater environments are frequently covered by mud and flora. Thus, conventional stereo cameras needing texture to detect and characterize singular points, are useless in these conditions in order to recover 3D information from the object to be manipulated and its surrounding area [1], [2]. In addition, the performance of traditional optical imaging systems are limited underwater by absorption and scattering. These two terms depend on the turbidity of the water the light is propagating in and the light source used. There are two main limitations: a system can be power limited when, increasing the illumination power, only the backscatter increases; or contrast limited, when the propagation signal is too weak to be detected by the sensor.

An external light can be used to help in the lightning but if there was no texture, a light may not solve the problem. An active system, where the light is used cleverly to project a texture or pattern and the camera is used to recover the deformed pattern can overcome a passive system where there is nothing else but the environment [3]. In terms of illumination, laser light presents two major benefits compared to conventional lightning or projectors: a decreased absorption coefficient if



Figure 1. Laser and camera system. The camera is housed in the black cylinder and the laser is in the white one. Both cylinders have a flat acrylic port.

its wavelength is chosen wisely and a decrease in scattering, thanks to its small illumination volume [4].

Laser as a light source has been widely used as stripe scanning system [5], where the movement of the slit can be done with a manipulator or moving a vehicle while scanning. However, depending on the accuracy of the pose of the vehicle, the reconstruction may be not detailed enough. For this reason, a one shot reconstruction system would benefit from a detailed reconstruction even in movement.

The system presented in this paper is capable of providing a non-dense point cloud in one camera shot, using a projection pattern and a triangulation method between the sensed pattern in the camera frame and the calibration knowledge.

This article is structured as follows: the description of the sensor is given in section II, the experimental setup in section III, the results in section IV and the conclusions in section V.

II. DESIGN AND IMPLEMENTATION

The system presented in this paper is formed by a housed 1920 × 1440 color camera with a 12 mm optic and a housed 532 nm green laser (figure 1). A Diffractive Optical Element (DOE) has been placed in front of the laser to modify the beam shape to a set of 25 parallel lines. These lines are projected on the scene and recovered by the camera.

The system has been design to be modular, reconfigurable and easy to mount in an underwater robot to provide 3D information of the environment.

A 532 nm laser has been chosen as light source, as color is extremely important in underwater environments due to

absorption and scattering. Those coefficients vary depending on the wavelength of the light source [6]. In order to transmit the maximum light they have to remain low. Blue-green color spectra present a good compromise between absorption and scattering.

The camera and the laser housings are made out of acetal plastic and are fixed to a deployable structure to keep its relative position fixed throughout the experiments.

A. Reconstruction pipeline

The process beginning with the acquisition of images to the final point cloud is split in five steps:

1) *Acquisition*: This step handles the image acquisition, where illumination and color changes have to be taken into account. The exposure of the camera is set so that the only saturated pixels in the image lie in the area where the strongest laser spot is projected, making for an easy detection of the central laser beam and a better accuracy in a peak detector algorithm for those laser points different to the main beam.

Due to the DOE, the lines have a Gaussian profile and the central line has a brighter dot where the laser would project its beam if the DOE was not mounted. This dot or central laser beam is used to decode more easily a line segment.

2) *Segmentation*: Once the acquisition process has finished correctly setting the camera parameters, the segmentation process removes the background illumination by subtracting the red channel to the green channel. Therefore, the green lines that are only present in the green channel do not get affected by the subtraction but the background illumination.

The red channel has been chosen instead of the blue channel because part of the sensitive spectra of the blue channel in the camera lies close to the wavelength of the laser and not only the background would be removed but also part of the laser lines.

Finally, a binary image is computed by thresholding the previous result, and the line centers are found for each row of the image. For each center, the neighboring values at the original image are the input for a peak detector algorithm, using the center of mass method [7]. The centers that do not reach a certain intensity value are also discarded. Other authors have used the first, or even the second derivative to compute those peaks [8].

3) *Decoding*: The output image of the segmentation stage is scanned row by row for rows where all the lines are indexable. A row is defined indexable if 25 crossings are found, meaning that there is a one to one correspondence between the crossings and the 25 different planes.

These indexable rows are used as seeds for a flood fill algorithm that scans the lines vertically, following the segmented line if and only if there is at least one common pixel vertex.

However, when one or more crossings are missing, the correspondence between the planes and the crossings is not straightforward and has to be solved. To solve these cases, two constraints are used. The first constraint fills in the gaps left by the flood filling algorithm. For example, if there is an unclassified line segment that has a left neighbor labeled as

index $n - 1$ and a right neighbor labeled as $n + 1$ for all the pixels in the line segment, then the unclassified line must be labeled as n .

The second constraint to check is that the leftmost line has to be completed by the closest unindexed right line, if there is any, and if the distance of the gap is similar to the distance of the unindexed line segment.

The focus of this work, however, is not on the decoding step of the reconstruction pipeline but on the ability to outperform a conventional sensor under high turbidity conditions.

4) *Triangulation*: With the labeling and the calibration, each 3D point $p(t)$ can be computed by triangulating its corresponding laser plane π_n^L to the line formed by joining the segmented pixel to the camera focal point, which depends on the scale factor t .

$$\pi_n : Ax + By + Cz + D = 0 \quad (1)$$

$$p(t) = \left(\frac{u - c_x}{f_x} t, \frac{v - c_y}{f_y} t, t \right) \quad (2)$$

$$t = \frac{-D}{A \frac{u - c_x}{f_x} + B \frac{v - c_y}{f_y} + C} \quad (3)$$

where (f_x, f_y) is the camera focal length in x and y axes. (c_x, c_y) is the central pixel in the image. (u, v) is the detected laser peak pixel in the image.

Replacing 3 in 2, the 3D coordinates of the point are obtained.

B. Calibration

In terms of calibration, the camera has been considered as a pinhole camera with an unknown projection matrix, and the laser as a pinhole projector (inverse of a camera) with a set of 25 planes whose equations are unknown.

The calibration of this system has been done in two stages: first, with the laser projector switched off, the camera intrinsic parameters have been obtained underwater using a known checkerboard pattern. The detected squares of the pattern in the image and the known real size of the checkerboard pattern are the input 2D and 3D points for a Perspective'n'Point problem, which has been solved using a Levenberg-Marquardt optimization. The solver finds such a pose that minimizes reprojection error, that is the sum of squared distances between the observed projections and the projected 3D points in the image.

Then, the unknown plane equations are solved by a plane fitting method: The calibration pattern is placed in the field of view of the camera and is detected, as shown in figure 2. As the camera is already calibrated, the plane equation of this pattern can be obtained. Then, the laser pattern is projected onto this plane and it is detected and decoded. The observed projections of the detected laser points are raytraced to the calibration plane and their corresponding 3D points are obtained. More 3D points at different depth are obtained by redoing the same work for different poses of the calibration plane respect to the system.

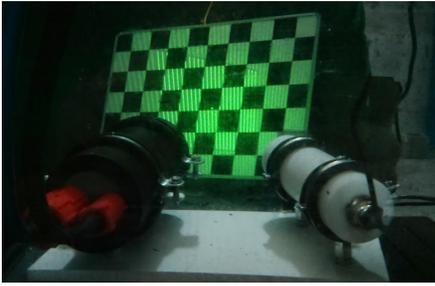


Figure 2. Laser and camera calibration in the water tank. The laser pattern is projected onto the calibration pattern. Both patterns are recovered by the camera.

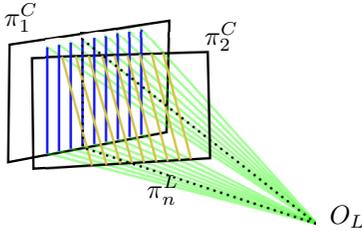


Figure 3. Diagram of the plane fitting calibration method. Two calibration planes π_1^C and π_2^C are shown with the detected laser lines on top. All points belonging to the lines with the same labeling are then fitted to a plane.

All points belonging to the same labeled line are used to fit a 3D plane and calibrate the system. At this stage, the labeling is controlled by a human operator. In figure 3, two calibration planes π_1^C and π_2^C are shown, with the laser line detections in blue and yellow. A laser plane π_n^L can be fitted to minimize the distance to the 3D points. With more calibration planes and detections, better accuracies can be achieved.

This calibration stage does not fix the condition that all the planes must cross at a line passing the laser focal point. This condition may be true in air, where there is no optical port in front of the laser, but when the laser is housed, small distortions are introduced which are taken into account by calibrating each plane separately, without considering the fixed geometry of the DOE.

III. EXPERIMENTAL SETUP

The experimental setup for this conference article focus on the ability of the camera to recover the laser pattern in different turbidity conditions. The system has been deployed in a 125 l, $1.2 \times 0.35 \times 0.35$ m water tank and the same scene has been reconstructed in nine different turbidity conditions. In figure 4 the setup is depicted. The scene consists of a textureless white bottle at the back place approximately at 0.7 m from the camera, some stacked tiles and a brown jar on top, approximately at 0,5 m from the camera.

A conventional camera or stereo system may find enough features in the tiles or the jar to perform any kind of feature matching, but not on a textureless object such as the bottle. Furthermore, with the addition of turbidity, the number of detectable features decreases making even more difficult to extract keypoints.

The camera and the laser have been fixed together at an angle so that the projecting pattern can be seen from the camera

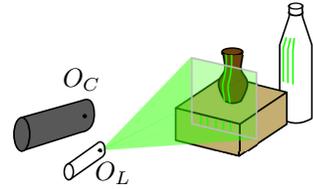


Figure 4. Experimental setup. The camera and the laser are deployed in a water tank, pointing to a jar and a bottle. O_C is the coordinate origin of the camera, and O_L is the origin of the laser. The transform between these two coordinates frames needs to be calibrated to obtain a valid 3D reconstruction.

Table I. NUMBER OF 3D POINTS DETECTED BY THE SYSTEM AT DIFFERENT TURBIDITY LEVELS.

Experiment	Turbidity (ml)	Turbidity (%)	3D points
1	0	0	13,202
2	5	1/250	12,661
3	10	1/125	13,516
4	15	3/250	13,315
5	20	2/125	13,062
6	25	1/50	13,013
7	30	3/125	9,882
8	35	7/250	5,600
9	45	9/250	77

from 0.5 m to 1 m due to the water tank dimensions. Once fixed and deployed, the calibration has been performed.

Turbidity has been obtained by pouring small quantities of whole milk into the water tank and then the mix has been stirred. Nine different milk concentrations, starting from 5 ml up to 45 ml in steps of 5 ml have been measured.

IV. EXPERIMENTAL RESULTS

Using the setup depicted in figure 4, and nine different turbidity levels, the number of 3D points reported by the system are shown in table I. Note that the number of points do not change until a high turbidity value is reached. Then the number of detected points falls until there is almost no detection at all. That is happening because the sensor is not able to discern the laser from the background, due to the scattered light.

In figure 5 experiments 1, 4, 7 and 9 are shown together with the detected points and the triangulated 3D points.

The recovered 3D information is not affected by turbidity. The geometry of the scene is clear and the laser line segments remain the same throughout the low to medium turbidity experiments. Although some points are missing, most of them are missed due to a very steep angle between the scene surface and the projection, causing the line segments to be very thin when projected in the camera image. Besides, the surfaces whose orientation is similar to the heading of the camera do not get affected by pollution until the image loses contrast.

Note how in the higher turbidity scenario mostly only the central laser beam is detected. Images of the jar and the bottle have poor contrast and colors in the presence of scattering, as expected.

V. CONCLUSION AND FUTURE WORK

The designed system has been tested in a variable turbidity scenario and one shot 3D reconstructions that reproduce the

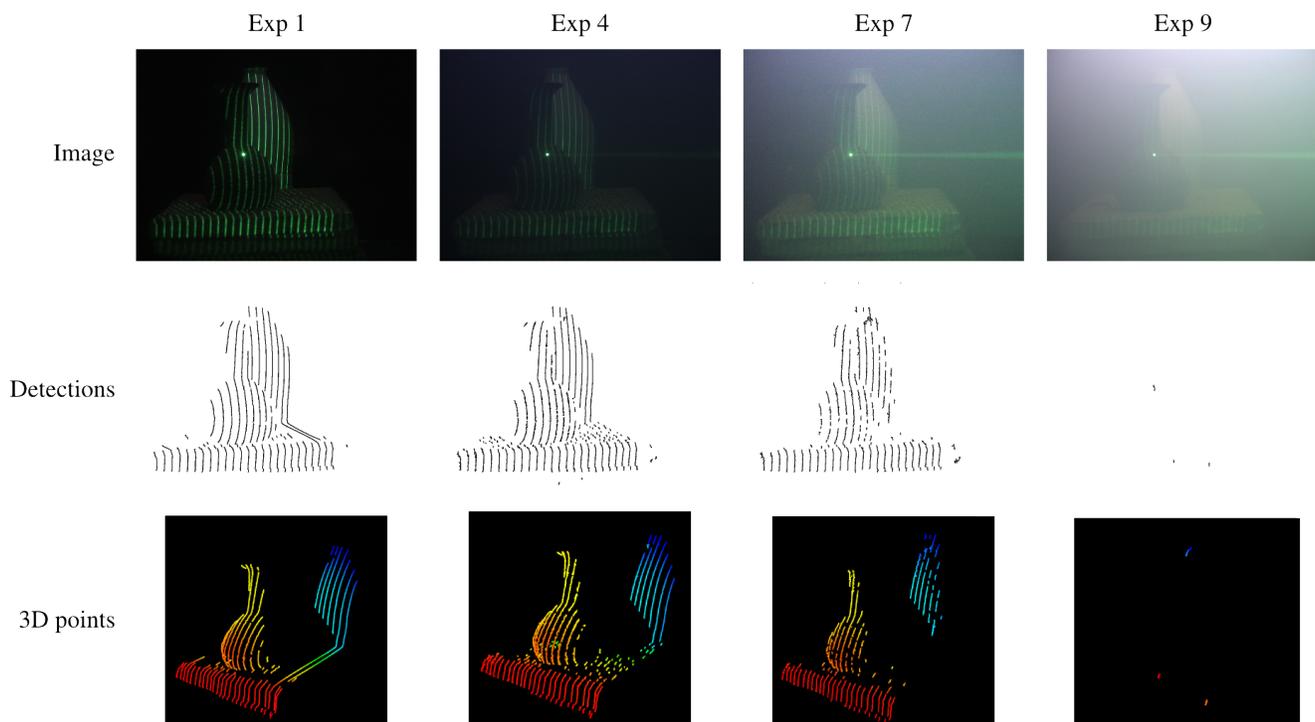


Figure 5. Laser images, detections and 3D points for different milk concentrations. The 3D point cloud has been rotated and it is presented in a isometric view, similar to the one in figure 4. For better quality, please refer to the digital version.

geometry of the scene have been obtained for the different configurations. The output 3D data can be used to find objects or to match them to a known object database. The modular system can be mounted on top of an underwater robot, as well as a standalone deployable sensor.

As seen in the results, the system has proved to be mostly invariant from low to mid turbidity levels. At high turbidity, the number of detected points decrease until the camera cannot discern the objects from the scattered light. The correct detection of the laser depends on the contrast between the lines and the background, thus it may not work when ambient light is high on in a high scattering medium. This scenario could be solved using pulsed lasers and range gated cameras with a complex timing and sensing pipeline, and more expensive hardware. Even so, the system has been designed to operate under tenths of meters of water, where darkness is guaranteed and there is not as much scattering as in shore or in a bay.

Future work includes the development of a better automatic decoder that correctly solves the correspondence between the line segments and the laser planes, as well as new experimental data comparing this system to a stereo rig or a multibeam sonar in terms of resolution, range and cost.

Other DOEs are also in the scope of this study. More lines or different patterns could be used, even from different lasers to provide a wider field of view.

ACKNOWLEDGMENTS

This work has been partially supported by grant BES-2012-054352(FPI) and contract DPI2011-27977-C03-02 by the Spanish Ministry of Economy and Competitiveness, FEDER Funding and by Govern Balear (Ref 71/2011).

REFERENCES

- [1] C. Beall, B. J. Lawrence, V. Ila, and F. Dellaert, "3D reconstruction of underwater structures," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 04481111. IEEE, Oct. 2010, pp. 4418–4423.
- [2] O. Pizarro, R. M. Eustice, and H. Singh, "Large Area 3-D Reconstructions From Underwater Optical Surveys," *IEEE Journal of Oceanic Engineering*, vol. 34, pp. 150–169, 2009.
- [3] 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, Jul. 2011.
- [4] E. P. n. González, F. S.-T. Díaz-Pache, L. P. Mosquera, and J. P. Agudo, "Bidimensional measurement of an underwater sediment surface using a 3D-Scanner," *Optics & Laser Technology*, vol. 39, no. 3, pp. 481–489, Apr. 2007.
- [5] M. Prats, J. J. Fernandez, and P. J. Sanz, "An approach for semi-autonomous recovery of unknown objects in underwater environments," in *2012 13th International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*. IEEE, May 2012, pp. 1452–1457.
- [6] J. S. Jaffe, K. Moore, J. McLean, and M. Strand, "Underwater Optical Imaging: Status and Prospects," *Oceanography*, vol. 14, no. 3, pp. 64–75, 2001.
- [7] R. Fisher and D. Naidu, "A comparison of algorithms for subpixel peak detection," *Image Technology, Advances in Image Processing, Multimedia and Machine Vision*, 1996.
- [8] J. Forest Collado, J. Salvi, and E. Cabruja, "Laser stripe peak detector for 3D scanners. A FIR filter approach," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR'04)*, 2004, pp. 2–5.