

MULTIMEDIA PHOTOGRAMMETRY 2.0. A FIRST STEP FOR UNDERWATER CULTURAL HERITAGE APPLICATION

Michele Russo ¹, Roberta Ravanelli ², Agnese Pini ², Giulia Flenghi ¹,
Luca Martelli ¹, Paolo Monti ², Mattia Crespi ²

¹ History, Representation and Restoration Dept., Sapienza University of Rome
(m.russo; giulia.flenghi; luca.martelli)@uniroma1.it

² Civil, Constructional and Environmental Engineering Dept., Sapienza University of Rome
(roberta.ravanelli; agnese.pini; paolo.monti; mattia.crespi)@uniroma1.it

Commission II

KEY WORDS: Multimedia Photogrammetry, Stereo-matching, Camera orientation, Metrology, Encoded Targets, Controlled environment, Image processing, Underwater Cultural Heritage

ABSTRACT:

The research is framed in multimedia photogrammetry, a specific domain aimed at acquiring geometric information about static objects immersed or semi-submerged in a liquid through one or more cameras external to the liquid. If the liquid is water, this field belongs to the broader field of applied metrology for analysing and understanding the aquatic world. Specifically, the various passive sensing techniques for acquiring Underwater Cultural Heritage (UCH) in shallow water are central to understanding the research domain's underlying issues. Our research is framed in the domain, implementing the automatic analysis to estimate a priori (and correcting a posteriori) the camera's behaviour under certain conditions for acquiring submerged or semi-submerged objects. The first analytical results are framed in a two-year project, which aims to define a behaviour model in a controlled environment with encoded targets and stereo-photogrammetry, automatically extracting the camera orientation parameters under different water height conditions. A planar reproduction of a CH artefact, which simulates an immersed architectonic floor, has been applied to validate the process in a first case study, testing the system's capacity to extract the correct coordinates of the image. At the end of this first experimental phase, the aim is to define a model for the behaviour of water deformation. It will make it possible to predict and correct the water refraction by calculating the correct coordinates within the liquid. In the future, this model will be tested under different and incrementally more complex acquisition conditions. The global project's primary goal is to arrive at the application of this model in an uncontrolled environment for the survey of UCH in shallow water.

1. INTRODUCTION

This research is framed by surveying geometric shapes in water using cameras. The research explores the topic of multimedia photogrammetry, a surveying methodology devoted to acquiring geometric information of static objects in liquid or semi-submerged with one or more cameras outside the liquid. This field is part of the macro-domain of applied metrology for the analysis and knowledge of the aquatic world, introducing an innovative acquisition approach for a broader application of passive sensing techniques for shape acquisition in shallow waters.

Knowledge of the aquatic ecosystem defines a vast field on which primary interests have recently emerged nationally and internationally. First, we report the conferences "ACQAE, Il futuro è nell'oceano" (<http://aquae.cnr.it>) and the "MaGIC2 project" (<https://www.protezionecivile.gov.it/en>), which delved into areas with depths ranging from 0 to 50 meters. At the European scale, some research activities and projects are related to the topic, such as the "Our Ocean Conference" (<https://ourocean2019.no/>), which introduces an ocean monitoring system. At the research level, the project "MareCulture" (<https://imareculture.eu>) is aimed at opening the digitalization of Underwater Cultural Heritage (UCH) to the public, "NAUTILOS" (<https://www.nautilus-h2020.eu>), which aims to obtain a wide range of seafloor data with dense spatial resolution, "TECTONIC" (<https://www.tectonicproject.eu>), aimed at safeguarding UCH. The projects present a strong

multi-disciplinary connotation to deal with innovative technological and instrumental solutions capable of solving and optimizing the acquisition of information inside water. This last point highlights a limitation in applying such methodologies in so-called shallow waters, whose conditions make the use of submerged instrumentation extremely difficult or impossible.

The use of optical instruments out of the water requires a priori knowing the light rays' behaviour within the water. Besides, it is essential to discover how this behaviour can be corrected to survey shapes in the liquid. This issue is cogent today, given the importance of monitoring seafloor change or carrying out morphological analysis.

Therefore, the research¹ aims to provide an initial answer to this need through experimental response in a controlled environment, introducing an innovative 3D passive technique for multimedia aquatic surveying in shallow water. The aim of the whole project is to model the geometric distortion due to refraction effect, defining the correct geometry in the multimedia photogrammetric application.

¹ The authors participated equally in the experimental phase. In writing the article, P.M. M.C. were responsible for the Introduction, G.F. for the paragraph 2 (State of the Art), L.M. for the paragraphs 3.1 (Instruments), A.P. for the paragraph 3.2 (The experimental set-up), R.R. for the paragraphs 3.3 (Camera orientation in multimedia condition) and 4 (Case Study), MR for the paragraph 3 (Methodology) and 5 (Conclusions).

2. STATE OF THE ART

The aquatic ecosystem is an extensive field of research. The vastness and the crucial role of the aquatic area have led to the discovery of new instruments and methodologies to explore the contents and monitor the ecosystem to preserve it. The research on the surveys in the water started with monitoring oil platforms and large vessels in the 1950s (Leatherdale and Turner, 1983). In the last decade, an ISPRS Commission (Underwater Data Acquisition and Processing - ISPRS WG II/7) has been devoted to the topic, promoting many conferences and activities mainly related to integrating acquisition instruments and testing data processing pipelines related to the water domain. Besides, many articles have been proposed to analyse the in-water acquisition with active and passive sensors (Menna et al., 2018), showing different applications, as in the volume 3D Recording and Interpretation for Maritime Archaeology (Springer) or the Special Issue Underwater 3D Recording & Modelling (MDPI). A state-of-the-art can be found in (Song et al., 2022). Finally, an increasing number of conferences and projects at national and international scales demonstrates the attention, framing the scenarios for the near future in the domain.

A research bottleneck is the detection of shapes in shallow waters, whose conditions limit the application of submerged instruments. Using cameras out of the water demonstrates an implicit reduction in the instrumental working distance. However, it expands both the typology of optical instruments no longer adapted to aquatic conditions and the application scale to satellite (Lyzenga, 1978; Stumpf et al., 2003), airborne/drone (Agrafiotis et al., 2019), or terrestrial levels. Several studies have explored the geometry of projective rays passing through different gas or liquid materials (Maas, 2015; Menna et al., 2018), finding a field of application in seabed analysis (Mandlbürger, 2018). Besides, recent research has shown how the introduction of ML algorithms can improve the extraction of geometric features (Mohamed et al., 2020). The topic of combined above-water photogrammetry remains an open field of research, containing several challenges and extraordinary potential to unveil new UCH, starting with the study of refraction behaviour (Skarlatos and Agrafiotis, 2018).

3. METHODOLOGY

The experiment presented in this paper presents the first attempt of a long research path, which has been carried out in the Hydraulics Laboratory of the Sapienza University of Rome

using a tank measuring 80x80x50 cm with 1 cm thick transparent glass walls (Fig. 1). The tank is framed by a metal structure equipped with a millimetric movement system positioned on the top of the structure. Two pulleys allow the precise moving of a central plate. This latter has been used to fix a customized metal profile and position the cameras in a predefined set-up. The glass walls and the tank's base were covered with 0.5 mm thick white PVC sheets fixed to the glass with silicone sealant. A fixed taper has been positioned vertically in a corner to check the height of the water in real-time. It was decided to fill the water in steps of 5 cm, with eight different acquisition sets available (up to 45 cm depth) to have a sufficient framework for the geometric distortion in the box. At last, many attempts have been made to test the external lighting condition of the laboratory, which is not controlled, by checking the water reflections and light spots in the tank.



Figure 1. Experimental set-up.

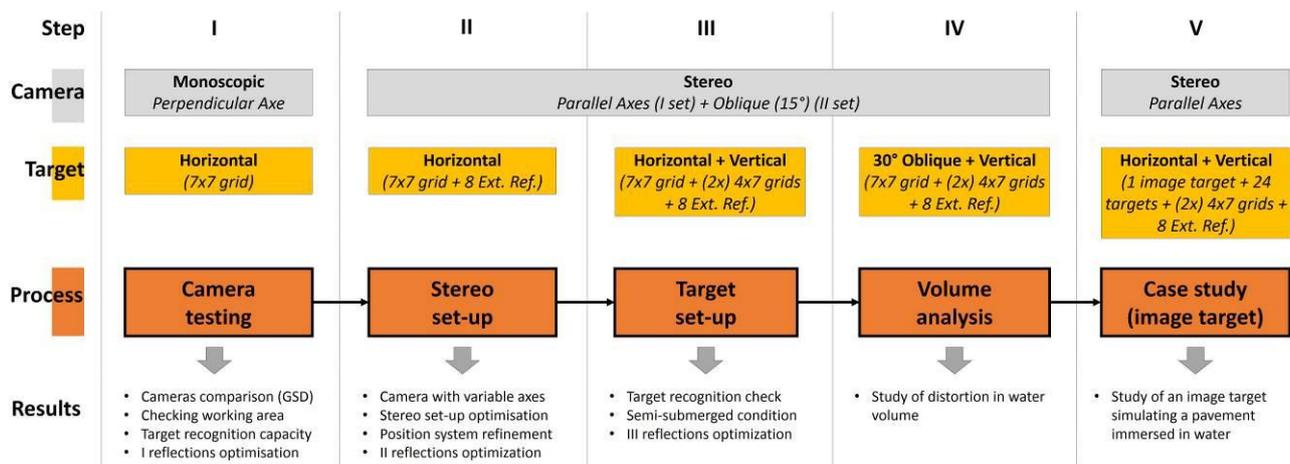


Figure 2. Experimental schema.

The introduction of white cardboard panels shaped and fixed to the structure made it possible to define the set-up that minimizes reflections without excessively reducing the scene's illumination (Fig. 1).

Several steps to refine the experimentation were planned, facing different variables. The first passage regards the choice of the camera (camera testing) in the experiment. Therefore, we compared cameras in monoscopic acquisition and camera axes perpendicular to the water, evaluating Ground Sampling Distance (GSD), working area/volume and target recognition (Fig. 2). The second step regarded the camera position (camera set-up). In this delicate phase, the best camera configurations for stereo-photogrammetry (baseline, axis rotation) and their positioning systems concerning the metal structure were planned, trying to simulate the results a priori (Fig. 5).

A third essential step regarded the target positioning (target set-up) respect to the acquisition system. The target distribution was refined during the experimentation to fit some sides of the box. We arrived at the definition of best target positions (horizontal, vertical) and conditions (submerged, semi-submerged or out of the liquid) at varying water depths (Fig. 6), with an external reference system defined by a sequence of targets (see section 3.2).

These different conditions allowed for evaluating target recognition in several acquisition conditions. The introduction of the oblique target in the fourth step allowed us to complete the analysis of the distortion variation in the whole tank (volume analysis). In the end, a printed ortho-image of a floor was used in the last step (Fig 4), simulating a pavement immersed in water and introducing an actual situation (case study, Section 4).

3.1 Instruments

Once the acquisition set and the experiment pipeline had been established, the focus was on the research's two main aspects: the cameras and the targets. In the first monoscopic step devoted to target recognition (Fig. 2), two cameras were used: a compact DSC-HX60 camera (Sony) and a OnePlus 9 smartphone. Their application helped compare and verify the working area, the image quality, the GSD (Fig. 3) and the capacity of the target detection algorithm to recognize the encoded targets in a submerged scenario. After this step, we decided to use the DSC-HX60 for a better result in the laboratory condition and the possibility of using two identical cameras, suggesting an ideal stereo-camera system. In Table 1 the main camera parameters are presented. Besides, both the cameras were calibrated before their applications.

Lens	Sony G
Sensor	CMOS Exmor R - 7,76 mm (1/2,3")
CMOS Dimension	5184x3456 px (6.03 x 4.62 mm)
Working Distance	1360 mm (to the base of the tank)
GSD	0.4 mm

Table 1. Camera parameters and working set-up

The targets applied belong to the family of AprilTag (Olson, 2011; Wang and Olson, 2016; Krogius et al., 2019), a visual fiducial system commonly used in robotics and specifically designed to be efficiently detected with automatic algorithms. Their uses range from camera calibration and ground truthing to object detection and tracking (Wang and Olson, 2016).

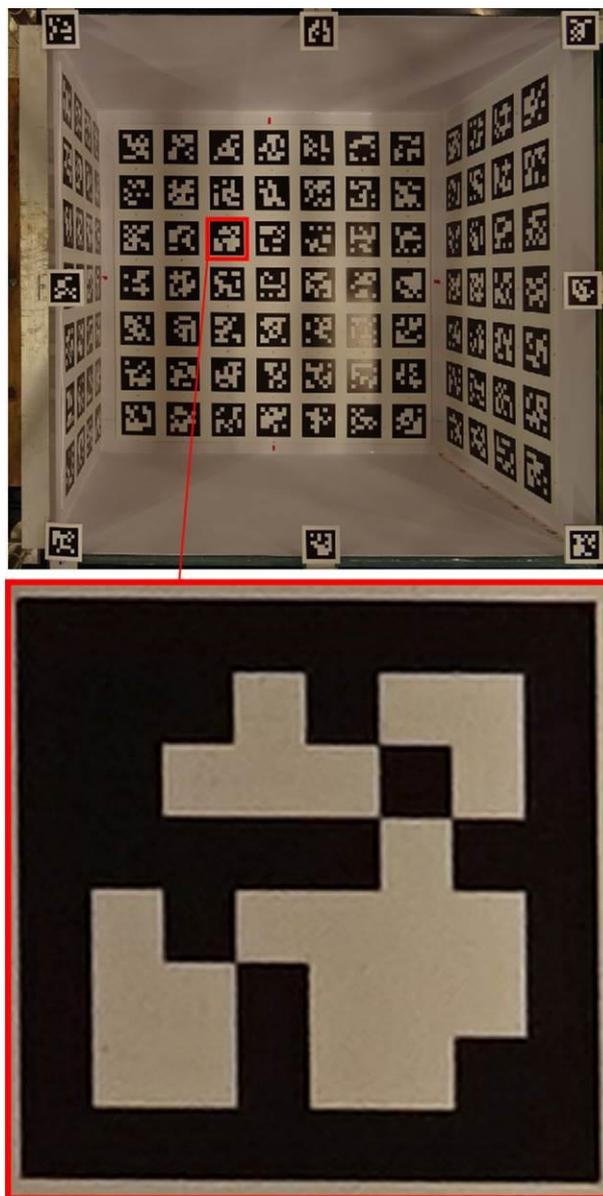


Figure 3. On the top the image acquired by the camera, on the bottom single target detected.

AprilTag markers are characterized by a standard layout. It is based on a square border surrounding a unique pattern of data bits (Krogius et al., 2019) and integrated by a fast detection algorithm. In this work (Fig. 4), we used the tag36h11 family and the most recent version of the AprilTag detection algorithm, AprilTag 3 (AprilRobotics. AprilTag 3), implemented in the pupil-apriltags open-source Python library (Pupil-labs. pupil-apriltags). This version includes a faster detector, improved detection rate on small tags, flexible tag layouts, and pose estimation capabilities (AprilRobotics. AprilTag 3).

The orthophoto of a pavement was finally chosen as the target image (Fig. 4). The 49.5 x 49.5 cm image at 300 dpi (0.08 mm/px) reproduces an area photogrammetrically surveyed by Remote Pilot Aircraft System (RPAS). Considering that the image represents the working area at a scale of 1:10, the experiment reproduces the boundary conditions of a survey more or less 16 meters above ground level with RPAS.

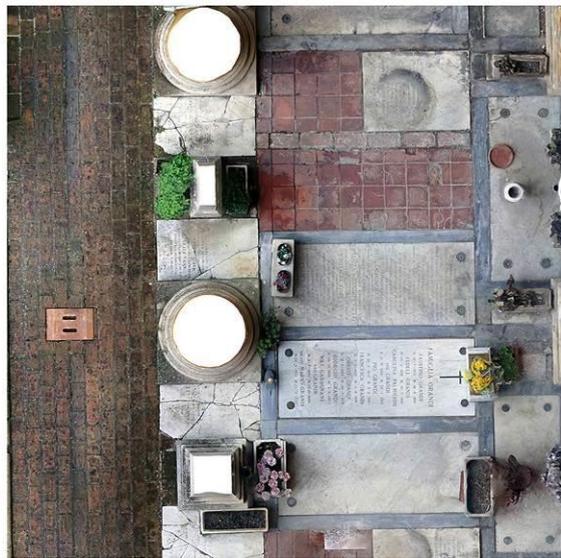
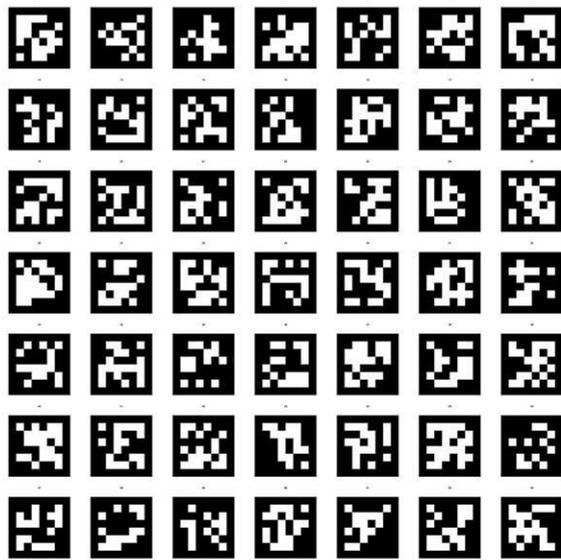


Figure 4. Visual fiducial grid (top) and image target (bottom).

3.2 Experimental set-up

In the acquisition planning, we decided to test two different survey configurations (Fig. 5): cameras with parallel and convergent axes to verify the different behavior of the light rays in the water. A mobile application was used to control the two cameras remotely, avoiding movements during the acquisition phase and trying to acquire the frames simultaneously. We also acquired three frames for each water height, choosing the best pair of images concerning water movement.

To ensure that the acquisition area was consistent with the available experimental space and allowed most of the targets in the scene to be acquired, a configuration with a 30-cm baseline with parallel-axis and a 60-cm baseline with oblique camera axes by 15 degrees was chosen. A metal plate was handcrafted to support these two configurations, customizing the gripping points for the cameras with the chance to rotate them according to specific angles. After testing different grip set-ups and verifying image quality, it was decided to lock both cameras with the following configuration:

Focal Length	Shutter Speed	ISO	Diaphragm
4 mm	1/10	80	3.5

Table 2. Camera set-up

We started acquiring a couple of images without water, and then we proceeded to catch a sequence of stereo-images every 5 cm of water up to 45 cm. At this stage, we also verified the impact of ambient light (laboratory neon lights) on the targets, covering the metal structure to reduce the effect of reflections.

Finally, as for the targets inside the tank, three grids of targets of known size and structure were printed on different forex panels of 3 mm, covering almost entirely the bottom of the tank and the two sides visible by the tilt-axis camera. Two 7x4 grids were glued with water-resistant one-component assembly adhesive onto 4 mm plexiglass panels and then fixed on the two lateral sides with the same adhesive.

The 7x7-grid forex on the bottom was first glued onto an 8 mm plexiglass panel with the same water-resistant adhesive. Then, it was fixed to an additional 3-mm-thick aluminium plate to increase the overall weight of the plate and prevent buoyancy and unwanted movement.

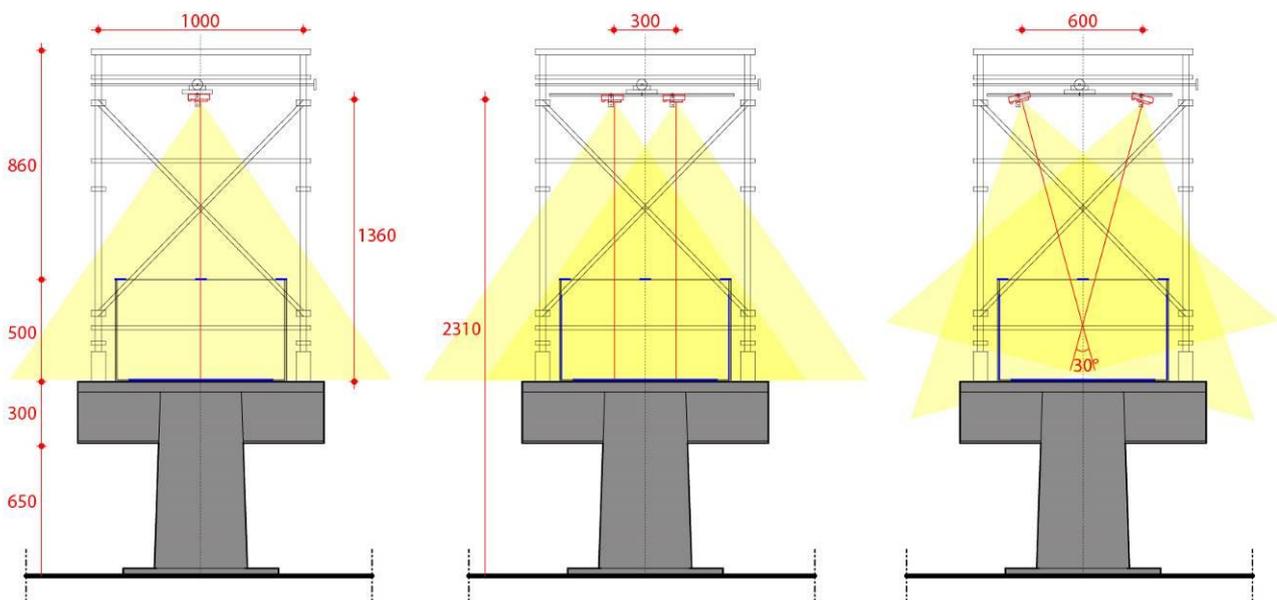


Figure 5. The three acquisition schemes applied in the photogrammetric experimentation.

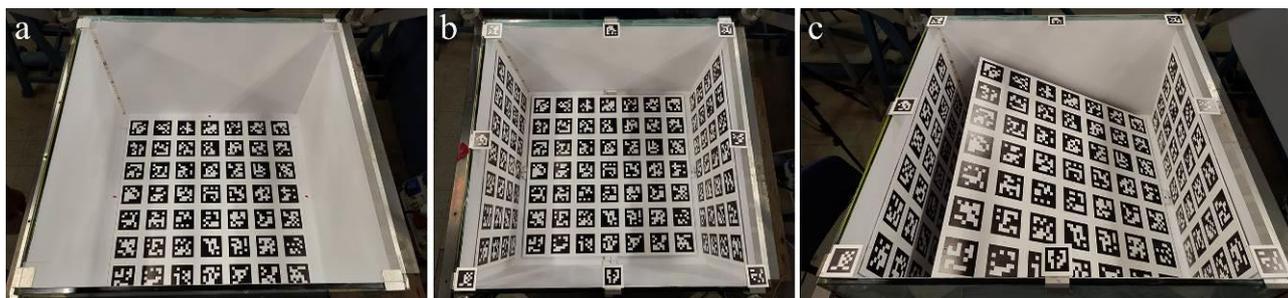


Figure 6. Target set-up: a) single horizontal target; b) vertical-horizontal target; c) vertical-oblique target scenario.

The bottom grid was positioned in two ways: backed horizontally on the bottom with some side shims to prevent movement and inclined at 30° with triangular metal profiles (Fig. 6). In the latter configuration, the inclination was calculated to cover the water depth of 45 cm completely. A tag dimension of 75 mm was used for all three grids. In addition, eight more petite tags (40 mm) were positioned on the border of the tank to define the invariant reference system of the survey. All the tags used have a unique ID.

The image target was printed on 5 mm thick plexiglass. It was then fixed to the forex plate, bearing the background targets with removable adhesive.

3.3 Camera orientation in multimedia condition

Thanks to the known geometry of the grids, the 3D position and orientation of each camera relative to the grids can be estimated. It can provide us with practical observations to set up a model able to handle the refraction effects on the optical path in function of the different water levels and camera set-up, preparing the immersion of image target (Section 4) or future

three-dimensional geometric shapes to verify their dimension and reliability concerning their known geometry.

For each experiment, characterized by a different set-up of the cameras and of the bottom grid (Section 3.2), we automatically tracked the position of the four corners of every tag (Fig. 3) at each filling step of the tank. It allowed for monitoring the displacement of the markers caused by the refraction of the optical rays due to the presence of water with increasing level (Fig. 7).

For marker displacement, represented by the quivers in Fig. 7, we mean the Euclidean distance between the position (expressed in terms of image coordinates) of the i th marker at the k th filling step of the tank (water level equal to $5 \cdot k$ cm with $k \geq 1$) and the position of the same marker at the first filling step of the tank (no water: water level equal to 0 cm). Fig. 7 shows how the pattern of marker displacements varies depending on the considered grid and, thus, implicitly, on the water level. Indeed, suppose we consider a certain fixed water level in the tank.

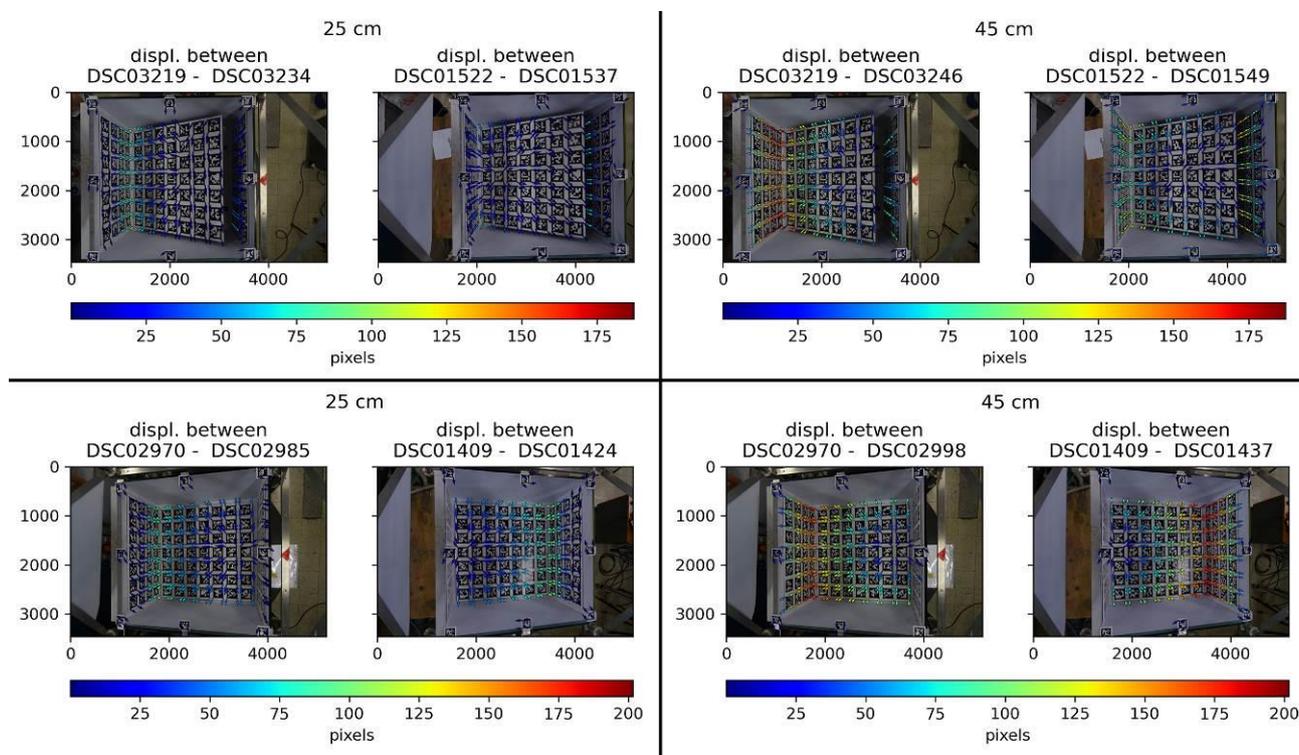


Figure 7. Map of target vertex displacements with 25 and 45 cm of water extracted from images acquired with parallel-axis camera set-up and inclined base targets (first row) and converging-axis camera set-up with horizontal base targets (second row).

In that case, all the markers of the bottom grid (in the planar configuration) are submerged at the same water level, and the displacements increase radially with the distance from a central distortion point. Moreover, the central distortion point remains fixed in all the filling steps of the tank while the magnitude of the displacements increases with the water level. Let us consider a certain fixed water level in the tank in the lateral grids oriented vertically. The markers are situated at varying depths based on their respective grid rows, and the displacements increase with the water level. In the case of the inclined bottom grid, both effects are present when considering a fixed water level, while the central distortion point changes its position with the increasing water level.

It is important to note that in the case of semi-submerged markers, i.e., targets for which the water level partially covers the markers, the target detection algorithm did not work, even after applying standard thresholding algorithms to limit the radiometric differences between the two parts of the targets. A possible hypothesis is that the geometric distortion of just one part of the target could create problems in the marker detection algorithm. The correct distortion modelling could thus improve the target detection in these cases, but further investigations are needed.

For each experiment, the external orientation parameters (position and orientation in the ground reference system) of each of the two cameras were independently estimated at each filling step of the tank to monitor the effects of the marker displacements and, thus, of the water distortion, at increasing

water levels in the ground reference frame. The intrinsic orientation parameters of the two cameras were previously estimated through the "Lens calibration" tool of Agisoft Metashape using the same acquisition parameters employed during the experiments. To estimate the external orientation parameters of the camera, we employed the Perspective-n-Point (PnP) approach (Marchand et al., 2015). The method estimates the pose of a calibrated camera using a given set of n 3D points in the ground reference system and the corresponding 2D projections of the points in the image (image coordinates) – by leveraging the known grid geometry and the detected image coordinates of the markers.

The external orientation was carried out independently considering the three grids – keeping fixed the previously found intrinsic orientation parameters – thus having three different estimations of each camera's pose for each filling step of the tank. For each filling step of the tank, the set of marker coordinates was split in half between Ground Control Points (GCP) and Check Points (CP). It allowed to check the value of the mean reprojection error obtained using the estimated external orientation parameters to project the known 3D coordinates of the considered markers in the ground reference frame in the considered image. The independent external orientation of the cameras made it possible to have an independent evaluation of the baseline (estimated as the Euclidean distance between the two estimated camera centre positions) and to be compared with the imposed values (Section 3.2), as shown in Fig. 8.

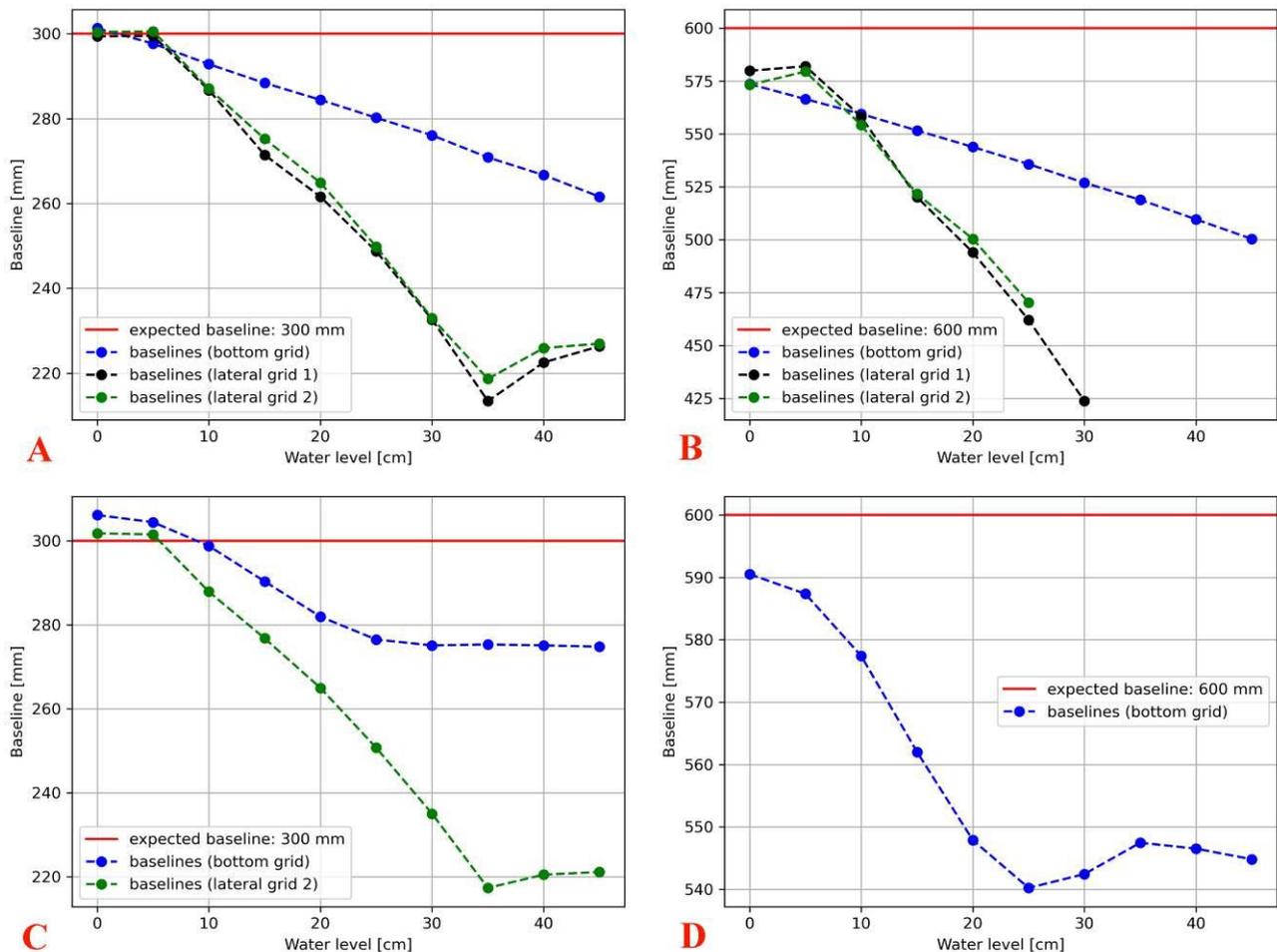


Figure 8. Extracted graphs with baseline values at varying water depth: A) parallel-axis cameras and planar target; B) converging-axis cameras and planar target; C) parallel-axis cameras and inclined target; D) converging-axis cameras and inclined target.

From Fig. 8A, relative to the experiments with the parallel camera axis (I set) and the planar bottom grid, it is possible to notice how the different patterns of the displacements in function of the type of the grid (bottom or vertical) influence the estimation of the external parameters and thus the estimation of the baseline. In estimating the baseline from the bottom grid markers, the baseline decreases linearly with the water level. In the case of the estimation of the baseline from the lateral grid markers, the estimated baseline decreases linearly with the water level until the water level of 35 cm, but with a different slope, (the baseline is estimated from markers at different water levels, from 0 to 35 cm). Until this level, there are both wholly submerged markers outside the water. From the water level of 40 cm, all the markers of the lateral grids are entirely submerged, and the variations in the estimated baselines are minor.

From Fig. 8B, relative to the experiments with the camera oblique axis (II set) and the planar bottom grid, the situation is similar to Fig. 8A. The difference is that, due to the oblique axes of the cameras, one of the lateral grids for each camera in the two set of images is very inclined and so the marker detection, in this case, is less reliable, especially when the water level is high. From the water level of 30 cm, no marker of one of two lateral grids is detected (Fig. 7, second row).

In Fig. 8C, relative to the experiments with the parallel camera axis (I set) and the inclined bottom grid, it was possible to retrieve the baseline from just one of the two lateral grids since the inclined grid projects a shadow on one of the two lateral sides of the bottom and this makes the marker detection less reliable. In this case, the baseline retrieved from the one lateral grid follows the same trend as Fig. 8A. In the case of the baseline retrieved from the inclined bottom grid, the decrease of the estimated baseline with the increase of the water level is smoother. It does not follow a linear trend since, due to the inclined grid; there are markers submerged and markers outside the water until the water level is 30 cm. All the markers are submerged from 30 cm and up but at different water levels due to the inclined grid, and the estimated baseline remains nearly constant.

In Fig. 8D, relative to the experiments with the camera oblique axes (II set) and the inclined bottom grid, only the markers of the bottom grid are reliably detected due to the shadows and occlusions of the inclined grid and the inclined views of the cameras. In this case, the trend of the estimated baselines is similar to Fig. 8C. All the baselines estimated should be equal to the expected values when the water level is 0 cm; this is generally true, with some residual errors due to the not perfect implementation of the desired setup or in the retrieval of the marker coordinates or in the estimation of the external orientation parameters. A refinement of the camera external orientation is thus under investigation.

4. CASE STUDY

For the case study with the mosaic target simulating an immersed architectural floor, the acquisition scenario was the same as the experiment with parallel camera axes (I set) and the planar bottom grid. In this case, the mosaic target was positioned on top of the bottom grid itself (Fig. 9): the mosaic target was specifically designed to leave the external markers of the bottom grid in view to have an easily detectable reference in the water. After a first acquisition with just the bottom grid in the tank without water, we collected the images with the mosaic target at an increasing water depth (from 0 cm to 45 cm).

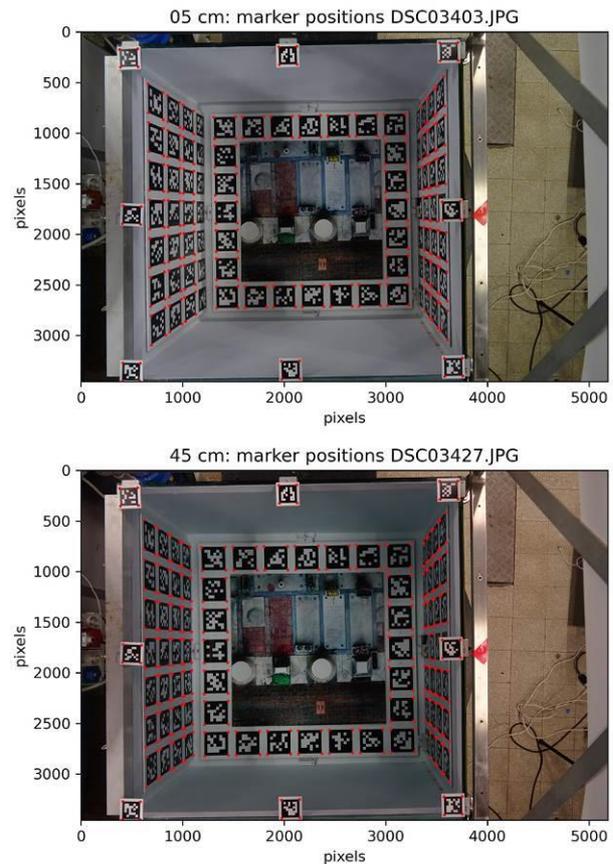


Figure 9. Comparison between the marker position and image distortion with 5 cm and 45 cm of water

The external orientation parameters of each of the two cameras were independently estimated at each water filling step of the tank. We used the same approach described in Section 3.3. The results in Fig. 10 show that the baseline trends are very similar to the ones reported in Fig. 8A: the acquisition set-up is thus stable and replicable. The water distortion, observable through the marker displacements, follows the same pattern of the analogous experiments. So, once modelled, it can be corrected. A first water distortion model is currently under development to prefigure the displacement pattern in function of the camera-viewing angle and the water depth, based on a radial distortion.

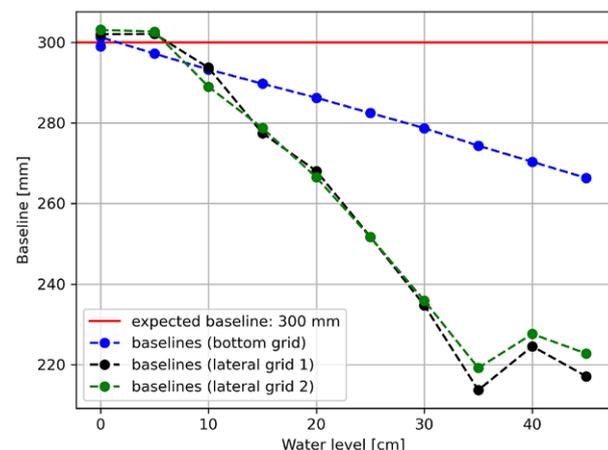


Figure 10. Extracted graph with baseline values at varying water depth of the target in the image target scenario.

5. CONCLUSIONS

This research aimed to provide an initial answer to collect geometric information about static objects immersed or semi-submerged in a liquid through two cameras external to the liquid. We designed and refined the experimental set-up to capture the data required to model the refraction effects on the optical path in function of different water levels and camera set-ups. Planar grids of encoded targets were used to monitor the displacements of the image coordinates distorted by the progressively increasing levels of water. A planar mosaic target was used to simulate an immersed architectural floor to verify the data collected in the experiments. The proposed experimental set-up allowed the observation of an evident radial pattern at a fixed water level whose magnitude increases with the water level. Moreover, the effects of the water refraction on the estimation of the external orientation parameters of the cameras were evaluated. The mosaic case study noted the same displacement pattern, providing data to define a first water distortion model currently under development. Future research stages will foresee the separate analysis of submerged markers (at different degrees of submersion) from non-submerged ones; the testing of the model to correct the image coordinates in the case study; the evaluation of errors in the target detection and its improvement when the targets are partially submerged.

ACKNOWLEDGEMENTS

The authors would like to thank the Sapienza University for the research funding that supported the experimental path.

REFERENCES

- Agrafiotis, P., Skarlatos, D., Georgopoulos, A., Karantzalos, K., 2019: DepthLearn: Learning to Correct the Refraction on Point Clouds Derived from Aerial Imagery for Accurate Dense Shallow Water Bathymetry Based on SVMs-Fusion with LiDAR Point Clouds. *Remote Sens.* 11, 2225.
- AprilRobotics. AprilTag 3. <https://github.com> (May, 2023).
- Krogus, M., Haggemiller, A., Olson, E., 2019: Flexible layouts for fiducial tags. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE.
- Leatherdale, J.D., Turner, D.J., 1983: Underwater photogrammetry in the North Sea. *The Photogrammetric Record* 11, 151-167.
- Lennert, M., GRASS Development Team, 2017. Addon i.segment.stats. Geographic Resources Analysis Support System (GRASS) Software, Version 7.2, Open Source Geospatial Foundation. grass.osgeo.org (1 June 2017).
- Lyzenga, D. R., 1978: Passive remote sensing techniques for mapping water depth and bottom. *Applied optics* 17(3), 379-383.
- Maas, A., Rottensteiner, F., Heipke, C., 2017: Classification under label noise using outdated maps. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, IV-1/W1, 215-222.
- Maas, H. G., 2015: On the Accuracy Potential in Underwater/Multimedia Photogrammetry. *Sensors* 15, 18140-18152.
- Mandlbürger, G., 2018: A case study on through-water dense image matching. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 659-666.
- Marchand, E., Uchiyama, H., Spindler, F., 2015. Pose estimation for augmented reality: a hands-on survey. *IEEE transactions on visualization and computer graphics*, 22(12), 2633-2651.
- Menna, F., Nocerino, E., Drap, P., Remondino, F., Murtiyoso, A., Grussenmeyer, P., Börlin, N., 2018: Improving underwater accuracy by empirical weighting of image observations. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 699-705.
- Menna, F., Agrafiotis, P., Georgopoulos, A., 2018: State of the art and applications in archaeological underwater 3D recording and mapping. *Journal of Cultural Heritage* 33, 231-248.
- Michalis, P., Dowman, I., 2008: A Generic Model for Along-Track Stereo Sensors Using Rigorous Orbit Mechanics. *Photogrammetric Engineering & Remote Sensing* 74(3), 303-309.
- Mohamed, H., Nadaoka, K., Nakamura, T., 2020: Towards Benthic Habitat 3D Mapping Using Machine Learning Algorithms and Structures from Motion Photogrammetry. *Remote Sens.* 12, 127.
- Olson, E., 2011: AprilTag: A robust and flexible visual fiducial system. *IEEE International Conference on Robotics and Automation*, Shanghai, 3400-3407.
- Pupil-labs. pupil-apriltags: Python bindings for the apriltags3 library. <https://github.com/pupil-labs/apriltags>.
- Skarlatos, D., Agrafiotis, P., 2018: A Novel Iterative Water Refraction Correction Algorithm for Use in Structure from Motion Photogrammetric Pipeline. *J. Mar. Sci. Eng.* 6, 77.
- Smith, J., 1987a: Close range photogrammetry for analyzing distressed trees. *Photogrammetria* 42(1), 47-56.
- Smith, J., 1987b. Economic printing of color orthophotos. Report KRL-01234, Kennedy Research Laboratories, Arlington, VA, USA.
- Smith, J., 2000: Remote sensing to predict volcano outbursts. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XXVII-B1, 456-469.
- Song, Y., Nakath, D., She, M., Köser, K., 2022: Optical Imaging and Image Restoration Techniques for Deep Ocean Mapping: A Comprehensive Survey. *PFG* 90, 243-267.
- Stumpf, R. P., Holderied, K., Sinclair, M., 2003: Determination of water depth with high-resolution satellite imagery over variable bottom types. *Limnology and Oceanography* 48 (1part2), 547-556.
- Wang, J., Olson, E., 2016: AprilTag 2: Efficient and robust fiducial detection. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4193-4198.