Automatic tag detection in multimedia survey condition

Michele Russo[1], Luca Martelli[1], Roberta Ravanelli[2], Agnese Pini[3]

[1] Dept. of History, Representation and Restoration, Sapienza University of Rome, Italy - m.russo@uniroma1.it, luca.martelli@uniroma1.it
[2] Dept. of Geography, University of Liège - roberta.ravanelli@uliege.be
[3] Dept. of Civil, Constructional and Environmental Engineering, Sapienza University of Rome, Italy - agnese.pini@uniroma1.it

Technical Commission II

Keywords: Multimedia Photogrammetry, Stereo-matching, Metrology, Encoded Targets, Computer Vision, UCH.

Abstract

Recently, aquatic ecosystems have captured the interest of the international scientific community, spurring the development of new instruments and methodologies to explore, monitor, and preserve these environmental systems. In this context, the detection of shapes in shallow waters (either in submerged or semi-submerged scenarios) constitutes an intriguing research topic. Our investigation is framed within the multimedia photogrammetry domain, which aims to retrieve geometric information of static objects submerged or semi-submerged in a liquid (usually water), with one or more cameras placed outside the liquid itself. Performing a photogrammetric survey of objects in shallow waters with above-water cameras remains an open research field in the domain of Optics, presenting several theoretical problems and technical bottlenecks, starting with the study of refraction behaviour. Our goal is to develop an automatically applicable methodology to estimate a priori (and correct a posteriori) the effects of refraction on camera behaviour when capturing images of submerged or semi-submerged objects under certain conditions. We therefore tested the behaviour of the cameras in a controlled environment, through different depth levels and water motion conditions, and then elaborated a mathematical model of the optical distortion phenomenon encountered. Feature extraction presents many bottlenecks, mainly due to the particular optical conditions defined by multimedia acquisition and the eventual perturbations present if the liquid is in a condition of turbulence. This paper focuses on such technical problems, presenting part of the qualitative and quantitative results obtained at this stage of our research and the approach used to resolve some detection limitations.

1. INTRODUCTION

The research presented in this paper focuses on surveying geometric shapes in water using optical cameras out of water. In detail, the experimentation explores multimedia photogrammetry, a 3D acquisition methodology devoted to capturing geometric information of static objects immersed in a liquid or in semi-submerged conditions, with one or more cameras outside the liquid itself. This field is connected to the macro-domain of applied metrology for the analysis and knowledge of the aquatic world, creating the basis for a new acquisition approach based on passive sensing techniques for shape detection in shallow waters.

The project presents a robust multi-disciplinary connotation, because it involves multiple domains of research: from mathematical analysis to the development of computer scripts in Python language, from the study of projective geometry to instrument optics, and from metrology to hydraulics. Integrating these fields allows us to deal with innovative technological and instrumental solutions capable of solving and optimising a complex problem by acquiring data inside water.

As delved into the State of the Art, information acquisition in water, especially shallow waters, is a fascinating domain that is the focus of numerous projects at national and international scales. In fact, the aquatic world represents the newest and most enormous frontier of study in multiple application domains.

In order to study it, reliable surveying instruments capable of acquiring information within water are necessary. Several active and passive instruments can acquire information when immersed within the liquid, thus working in a coherent media. However, there is still a great deal of unexplored space on sensing techniques that work on different media, such as water and air. Using optical instruments out of water requires a priori knowledge of the light rays' behaviour within the water. Besides, it is essential to discover how this behaviour can be corrected by surveying shapes inside the liquid and introducing automatic processes to extract metrical information.

Therefore, our research[1] aims to suggest a first answer to this need. It follows up on analysing and refining the first experimental results obtained in a controlled environment, introducing an innovative 3D passive technique for multimedia aquatic surveying in shallow waters. The project aims to model the geometric distortion due to the refraction effect and define the correct geometry in multimedia photogrammetric applications.

2. STATE OF THE ART

2.1 Underwater projects

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 emerging national and international projects demonstrate growing interest. Among them, we would remind 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

[1] The authors participated equally in the experimental phase. In writing the article, M.R. was responsible for the Introduction and Conclusions, A.P. for paragraph 2 (State of the Art), R.R. for paragraph 3 (Methodology), and L.M. for paragraph 4 (Experiments and Results).

areas with depths ranging from 0 to 50 metres. 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 "i-MareCulture" (https://imareculture.eu) is aimed at opening the digitalisation of Underwater Cultural Heritage (UCH) to the public, "NAUTILOS" (https://www.nautilos-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.

2.2 Underwater research

Research on surveys in the water domain started long ago 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. 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). Some research highlights the problem of detecting shapes in shallow waters, whose conditions limit the application of submerged instruments. The application of cameras out of the water demonstrates an implicit reduction in the instrumental working distance, but it can expand 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 (Mandlburger, 2018). Besides, recent research has shown how the introduction of Machine Learning (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). This issue is cogent today, given the importance of monitoring seafloor change or carrying out morphological analysis.

3. METHODOLOGY

The experiment presented in this paper presents a follow-up of the first experimentation phase (Russo et al., 2024) which was carried out in the Hydraulics Laboratory of the Sapienza University of Rome, using a tank measuring 80×80×50 cm with 1 cm thick transparent glass walls. 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 customised metal profile and position the cameras in a predefined set-up, rotating the cameras with a precise angle (15°). The glass walls and the tank's base were covered with 0.5 mm thick white PolyVinyl Chloride (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 up to 45 cm depth. Some additional white cardboard has been added to reduce water reflection due to uncontrolled lighting conditions (Fig. 1).

Figure 1. Experimental set-up.

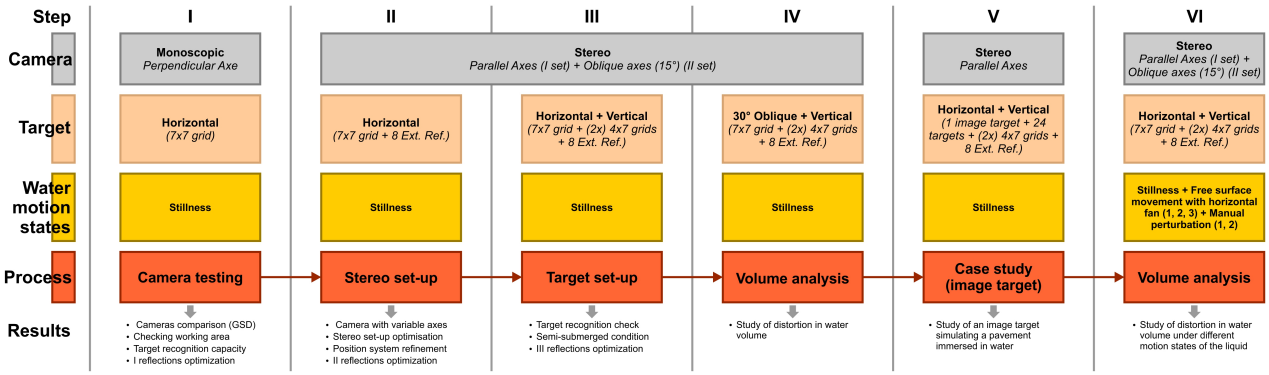

Figure 2. The experiments under consideration within the entire research project.

Beyond the first step (Russo et al., 2024), devoted to testing the camera's behaviour in different acquisition conditions, the experimental phase described here focussed on evaluating the optical camera's capacity to recognise automatic tags in different water conditions using a stereo-photogrammetry set-up (Fig. 2). The target distribution of the first experimental phase was preserved on the bottom and two sides of the box.

3.1 Instruments

In the testing phase, two compact DSC-HX60 cameras (Sony) were used in stereo configuration, dividing the acquisition sessions into two groups: 1) those with optical axes of the cameras oriented perpendicular to the bottom of the tank; 2) those with optical axes of the cameras oriented at angles of 15° to the zenithal axis. Table 1 presents the main camera parameters. The cameras were calibrated before their applications.
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).

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.

AprilTag markers are characterised 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 experimentation, 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). For both cameras in every dataset obtained in each experiment, we measured the optical distortions, caused by the water and found in the images captured at 350 Dots Per Inch (DPI), by comparing the changes in position of the marker vertices detected in every image with respect to a reference image captured with the tank empty.
Our approach stems from the Barrel Distortion model. We indeed modelled the effects of the refraction on the displacements Δ of the markers using a radial distortion law where the modulo of the displacements increases with the distance d from a central distortion point, i.e. the distortion centre C (x_C, y_C) (Fig. 3). The marker displacements indeed encode the optical distortion caused by the refraction of the optical rays due to the presence of water with increasing levels (Fig. 3). For marker displacement Δ, we mean the Euclidean distance $\underline{PP'}$ between the distorted position $P'(x_{P'}, y_{P'})$ (expressed in terms of image coordinates, Fig. 3) 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 undistorted position $P(x_P, y_P)$ of the same marker at the first filling step of the tank (no water: water level equal to 0 cm, Fig. 3). The equation of the proposed model is thus the following: $\Delta = K \cdot d^X$, where the displacements represent the observations and the four unknown parameters are the coordinates (x_C, y_C) of the distortion centre, and the two parameters K and X.

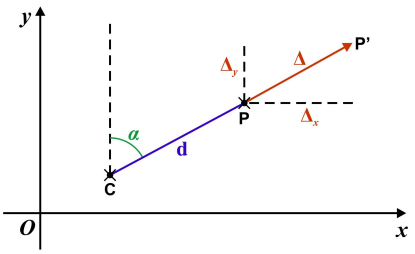

Figure 3. Distortion scheme and displacements.

The adopted model is not linear, since: 1) the coordinates of the centre of distortion are within the distance d, which is given by Pythagoras' Theorem (Fig. 3); 2) X is an exponent. We thus applied the natural logarithm to both the sides of the equation $(ln(\Delta)=ln(K)+X \cdot ln(d))$, in order to obtain a simplified form for the linearization of the model, needed to use the standard non-linear Least Squares approach – which employs the Gauss-Newton iteration scheme – for the estimation of its parameters.

3.2 Experimental set-up

In the acquisition planning, we decided to test cameras with: 1) parallel-axes configurations and a 30 cm baseline (Fig. 4a); 2) converging-axes configurations arranged at 15° to the zenithal axis and a 60 cm baseline (Fig. 4b). A mobile application was used to control the two cameras remotely, avoiding movements during the acquisition phase and trying to acquire the frames simultaneously. For the targets inside the tank, three grids of markers 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. Two 7×4 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 7×7-grid forex on the bottom was first glued onto an 8 mm plexiglass panel with the same water-resistant adhesive. Then, the forex was fixed to an additional 3-mm-thick aluminium plate to increase the overall weight of the plate and prevent buoyancy and undesired movements.

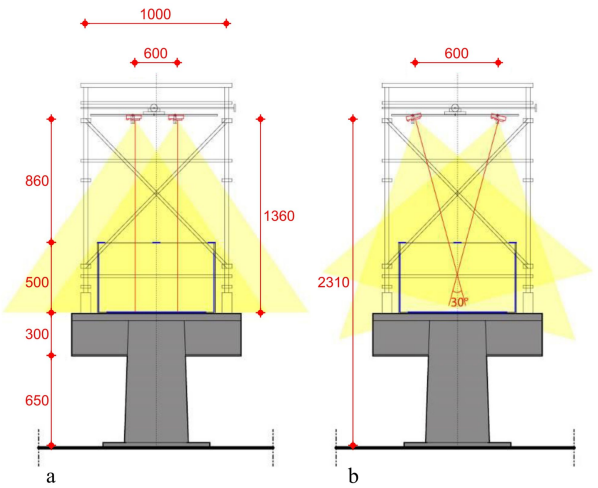

Figure 4. Acquisition schemes with: a) parallel-axes cameras; b) converging-axes cameras.

All three grids used a 75-mm tag dimension. Eight smaller tags (40 mm) were positioned on the tank's border to define the survey's invariant reference system. All the tags used have a unique I.D.

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

Table 2. Camera set-up

Once the whole system was set up, we organised the acquisition phase. In order to reduce temporary external effects, experiments prior to Test VI envisaged capturing three images with each camera for every survey state and selecting the best pair from these triplets. However, as water motion states are influenced by temporary external factors, for Test VI we analysed all three images of each triplet. In order to define the Ground Truth condition, we started capturing images without water in the tank. Then, we filled the tank with water at a 20 cm depth, capturing stereo images again with water in a still state. Next, three pairs of images were collected for each water motion condition, imposing an increasing intensity of movement on the liquid. Specifically, the first three levels of water movement in the tank were generated by a fan placed at the tank's border at three different intensity levels. The introduction of this forcing created ripples on the surface, intervening little along the water column. In contrast, adding to the fan-generated forcing a direct stir (caused by manually moving a body within the water) generated a chaotic perturbation increasing as a function of the force applied.

Thanks to the grids' known geometry, each camera's 3D position and orientation relative to the grids can be estimated. For each acquisition, we automatically tracked the position of the four corners of every tag, monitoring the displacement of the markers caused by the refraction of the optical rays due to the presence of water with increasing movement. The obtained marker displacement is represented as a quiver arrow in Fig. 5e, Fig. 6e, Fig. 7e, Fig. 8e, Fig. 9e, Fig. 10e.

3.3 Data analysis and processing

Once the capture phase was completed, we organised the files containing the images by subdividing them by their respective test iteration number, source camera (left or right), and dataset (arrangement with parallel-axes or converging-axes cameras). Each pair of images was then catalogued on a spreadsheet in .csv format where, in addition to the aforementioned information, other data concerning the survey conditions were recorded: water depth, approximate time of acquisition, and – in the case of Test VI – also the state of motion of the water. For Test VI, we schematically distinguished 11 motion conditions: 1) still water; 2) fan at slow speed; 3) fan at medium speed; 4) fan at maximum speed; 5) fan at maximum speed and additional mild stir; 6) fan at maximum speed and additional strong stir; 7) water undergoing stabilisation after 6 minutes since the strongest manual stirring; 8) water undergoing stabilisation after 9 minutes since the strongest manual stirring; 9) water undergoing stabilisation after 10 minutes since the strongest manual stirring; 10) water undergoing stabilisation after 11 minutes since the strongest manual stirring; 11) water undergoing stabilisation after 12 minutes since the strongest manual stirring. Once these metadata were entered, the automated analysis was performed with a script in Python language, harnessing the Open Source Computer Vision Library (OpenCV). First of all, the program read the intrinsic orientation parameters of each camera, previously calibrated, then it digitally reconstructed the Ground Truths associated with the marker grids, computing the 3D coordinates of the markers based on the known geometry of the grids. To increase performance of the marker recognition algorithm against shadows, reflections, and water movement blur, the script converted every image from RGB colours to grayscale and individually applied some thresholding techniques.

From our experiments so far, it appears that the combined application of binary and Otsu's thresholding methods (Otsu, 1979) determines the best probability of automated detection.

The individual AprilTag marker is detected by the dedicated library when its sequence of 36 bits (black or white square elements), its inner black bit frame, and its outer white bit frame are recognizable in the image under consideration (AprilRobotics. AprilTag 3). With these conditions satisfied, the script locates the four vertices of the marker and schematises the position of its centre as the intersection between the diagonals of a parallelogram having these same vertices. It then stores the image coordinates of these five significant points of the single AprilTag in an array associated with the analysed image.

With the marker detection concluded, the script calculated the extrinsic parameters of each camera independently through the Point-n-Perspective (PnP) approach (Marchand et al., 2016). For each tank filling stage, we divided the sets of image coordinates of the significant marker points between Ground Control Points (GCP) and Check Points (CP). This allowed us to check the mean reprojection error by using the estimated external orientation parameters, reprojecting the known 3D coordinates of the markers under examination into the image reference frame of the considered image. By taking advantage of the independent external orientation of the cameras, we carried out an independent evaluation of the baseline – which we calculated as the Euclidean distance between the two estimated positions of the camera centres – and a comparison with the values imposed in our experimental setup. The estimated values of the baselines and angles between the two camera axes are then recorded in a spreadsheet in .csv format, together with their corresponding expected values.

To begin the modelling phase of each sequence of n images captured in a given dataset with a given camera, the script compared the array of image coordinates of the significant marker points detected in a reference image (which had to be captured without any water in the tank) with the coordinates of the corresponding markers in all the other $n-1$ images. For each of these $n-1$ comparisons, the script calculated the field of geometric distortions, represented by the marker displacements, and graphically represented it as a ruled surface in the Cartesian Space (Fig. 5c, Fig. 6c, Fig. 7c, Fig. 8c, Fig. 9c, Fig. 10c). Without any powerful perturbation, the field of optical distortions induced by water refraction tends to take the form of an elliptic pseudo-paraboloid, which is a function of image coordinates x, y, and observed Δ modulo values. Regarding the estimate of the four unknown parameters of the model, we firstly applied a regularisation of the image coordinates of the markers to avoid numerical instability phenomena. The script then computes the approximate values of the coordinates of the centre of distortion, assuming it to coincide with the point of minimum of the displacement vector field, i.e. the point where the observed Δ modulo assumes the smallest value in the entire image. In the 3D graph of the optical distortion field, the program represented such a point as a dark circle with a cross (Fig. 5c, Fig. 6c, Fig. 7c, Fig. 8c, Fig. 9c, Fig. 10c).

To compute the approximate values of the parameters K and X, we applied the RANdom SAmple Consensus (RANSAC) (Fischler et Bolles, 1981) algorithm to robustly estimate the parameters of the line that fits the observations in the ln-ln domain $(ln(\Delta)=ln(K)+Xln(d))$. Then we applied iteratively the non-linear Ordinary Least Squares (OLS) according to the Gauss-Newton iteration scheme until satisfactory values of the coefficient of determination R^2 and the scale (sigma zero) σ_0 are found. At the end of the iterative process, the optimal values of the 4 parameters of the model (x_C, y_C, K, X) are found and the script plots two line fitting ln-ln plots (Fig. 5d, Fig. 6d, Fig. 7d, Fig. 8d, Fig. 9d, Fig. 10d), i.e., the RANSAC line used to find

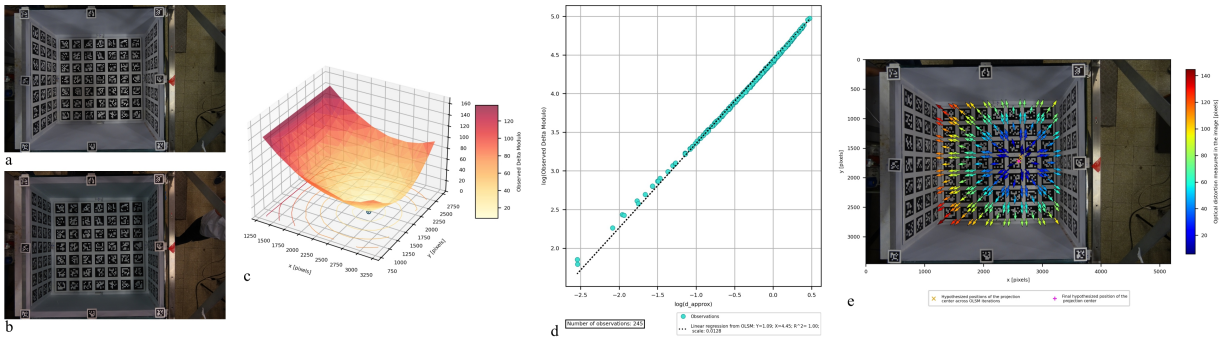

Figure 5. From the parallel-axes dataset of Test III (left camera), comparison between images DSC02940.JPG (a) – no water – and DSC02964.JPG (b) – water depth: 40 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of *log($d_{approximate}$)*, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

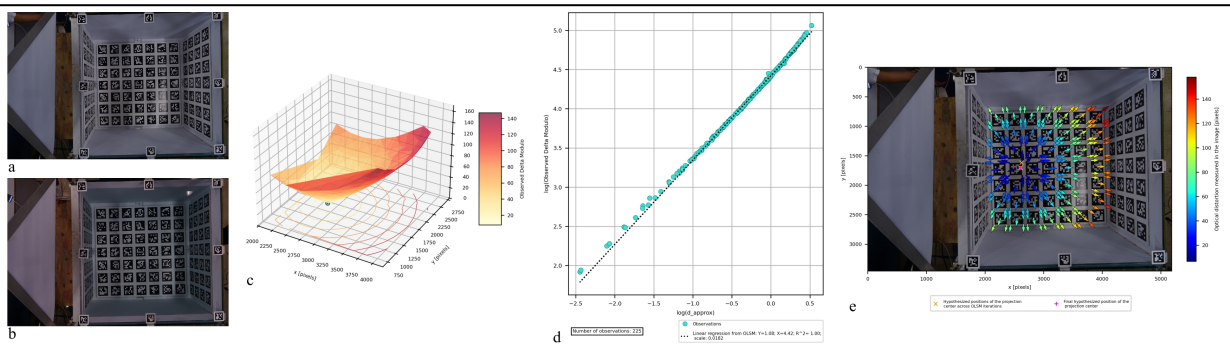

Figure 6. From the parallel-axes dataset of Test III (right camera), comparison between images DSC01379.JPG (a) – no water – and DSC01403.JPG (b) – water depth: 40 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of *log($d_{approximate}$)*, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

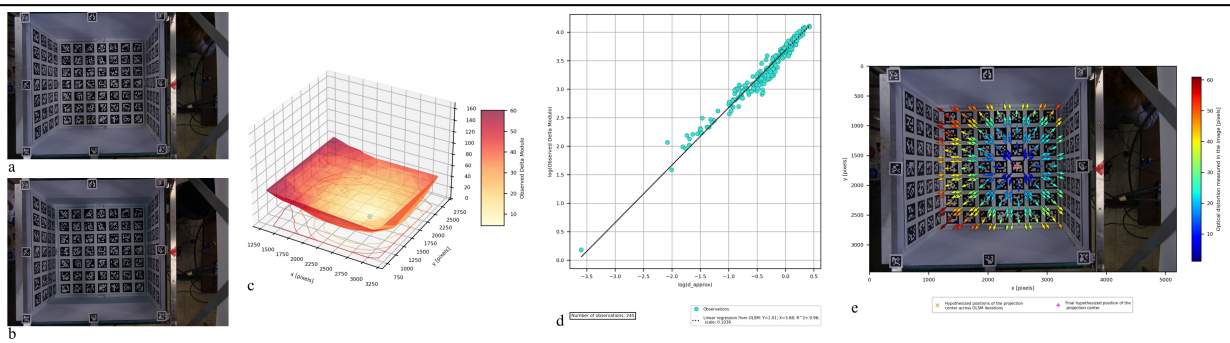

Figure 7. From the parallel-axes dataset of Test VI (left camera), comparison between images DSC06932.JPG (a) – no water – and DSC06949.JPG (b) – water depth: 20 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of *log($d_{approximate}$)*, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

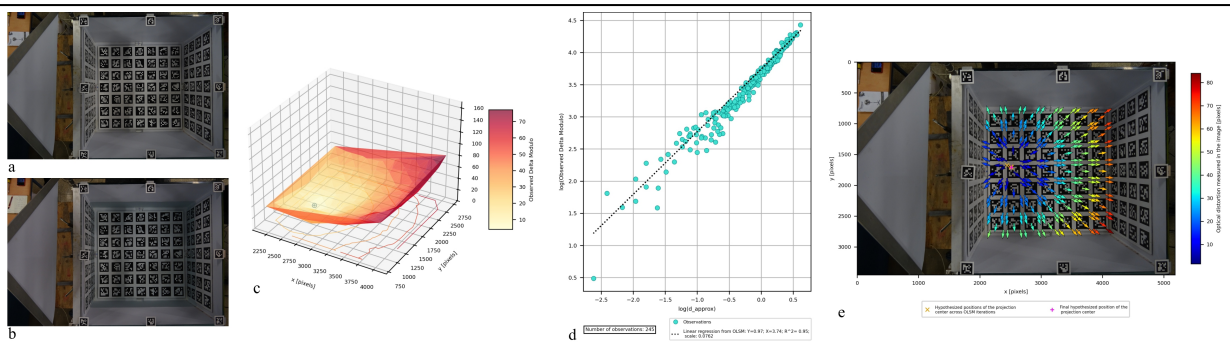

Figure 8. From the parallel-axes dataset of Test VI (right camera), comparison between images DSC04359.JPG (a) – no water – and DSC04376.JPG (b) – water depth: 20 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of *log($d_{approximate}$)*, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

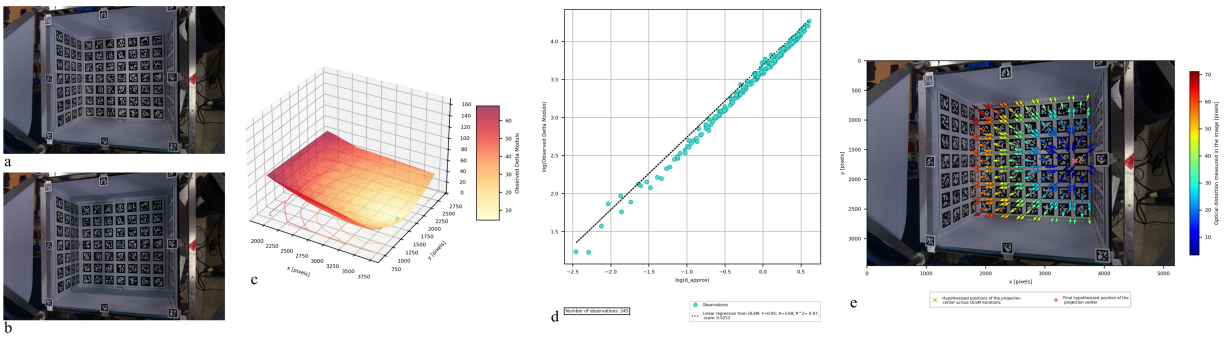

Figure 9. From the converging-axes dataset of Test VI (left camera), comparison between images DSC06998.JPG (a) – no water – and DSC06989.JPG (b) – water depth: 20 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of $log(d_{approximate})$, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

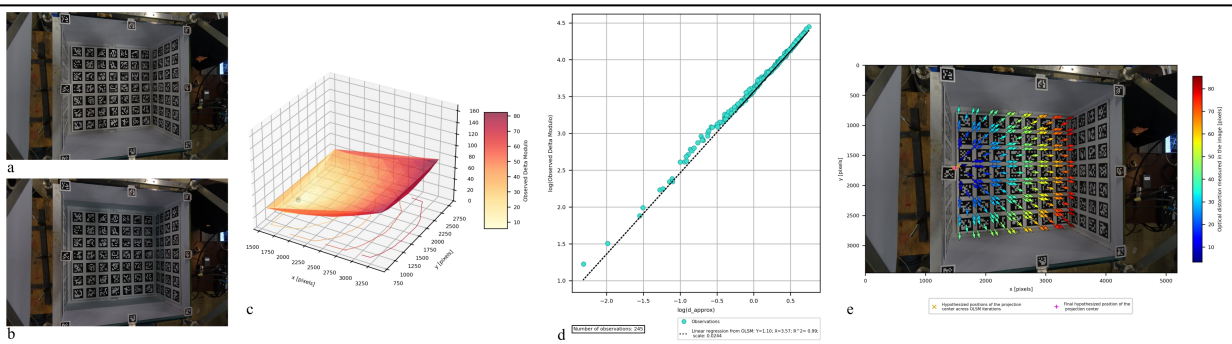

Figure 10. From the converging-axes dataset of Test VI (right camera), comparison between images DSC04426.JPG (a) – no water – and DSC04417.JPG (b) – water depth: 20 cm –; c) distortion field graph; d) plot of *log(Δ modulo)* as a function of $log(d_{approximate})$, obtained by the last iteration of the OLS Method; e) quiver plot of distortions.

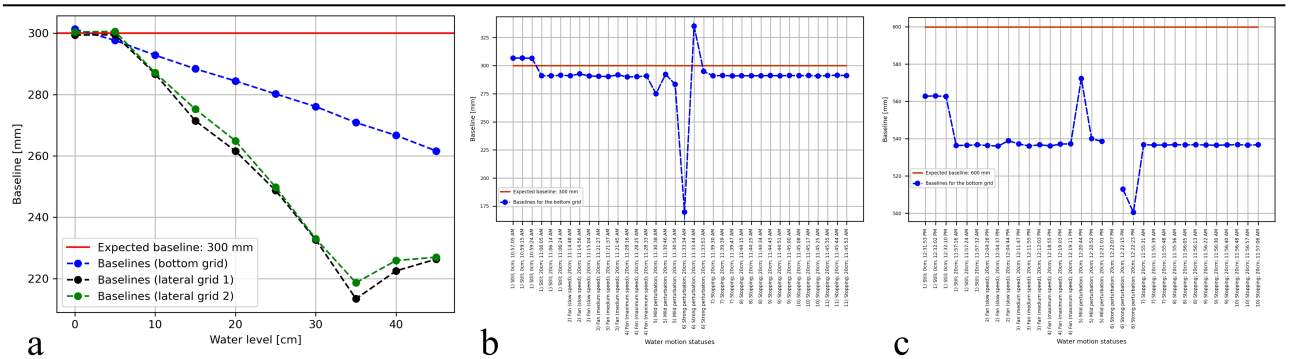

Figure 11. a) plots of baseline values as functions of water levels in the parallel-axes dataset of Test III; b) plots of baseline values as functions of water motion conditions in the parallel-axes dataset of Test VI; c) plots of baseline values as functions of water motion conditions in the converging-axes dataset of Test VI.

the approximate values of the *K* and *X* parameters, and the final optimal line whose parameters were estimated at the end of the Least Squares iteration scheme. Moreover, the script plots a quiver plot, where the optical distortions, i.e. the marker displacements occurring at all significant points of the recognised markers, are represented as vectors (Fig. 5e, Fig. 6e, Fig. 7e, Fig. 8e, Fig. 9e, Fig. 10e).
Finally, the script plots the trends of the estimated values of the baseline, the angle with respect to the zenith axis, and the four model parameters (x_C, y_C, *K*, *X*) for each image which has been compared to the reference image. In Test III these behaviour graphs included a plot of the baseline variation between various levels of tank filling at intervals of 5 cm, up to a maximum of 45 cm. For each of the three grids (bottom-side, left-side, and right-side), an analysis was conducted which allowed a different trend to be traced (Fig 11a). In Test VI, focussed on the bottom-side grid, the behaviour plots concerned the baseline variation between various states of water motion, for both the parallel-axes and the converging-axes datasets (Fig 11b, Fig11c).

4. EXPERIMENTS AND RESULTS

Two datasets were acquired for each of the case studies proposed below: one with the cameras having their optical axes oriented perpendicular to the bottom of the tank and a baseline of 30 cm; the other with the cameras having their optical axes oriented at 15° angles to the zenithal axis and a baseline of 60 cm. The ln-ln diagrams generated by processing with the RANSAC algorithm and the iterative scheme of non-linear OLS (Fig. 5d, Fig. 6d, Fig. 7d, Fig. 8d, Fig. 9d, Fig. 10d) show the model's ability to describe the physical phenomenon of refraction-induced optical distortions, as long as turbulence in water is not excessive. In fact, if the liquid is in a static state (as in Case Study 1) or in a condition of slight motion or stabilisation (as in some of the situations analysed in Case Studies 2 and 3), the observations are disposed with high R^2 values around the line representing the model and low dispersion.

4.1 Case Study 1: different water filling levels (Test III – parallel-axes dataset)

The first case study presented here is taken from Test III, which was focussed on analysing the behaviour of optical distortions through various filling levels of the tank at intervals of 5 cm, up to a maximum of 45 cm. In this case, the optical axes of the cameras were parallel to each other, perpendicular to the bottom of the tank, and separated by a distance of 30 cm. The vector field of optical deformations tended to take the form of a regular elliptical paraboloid. The higher the water level, the greater the concavity of the pseudo-parabolas that could be extracted from the vertical sections of the vector field (Fig. 5c, Fig. 6c). As seen with the comparisons between photos acquired in the absence of water, the influence of non-submerged targets on the behaviour of optical distortions can also be seen in the behaviour diagrams of the lateral marker grids (Fig. 11a).

4.2 Case Study 2: different water motion conditions (Test VI – parallel-axes dataset)

The second case study under consideration is taken from Test VI, which focussed on observing the behaviour of optical distortions through various possible states of water motion. As in the previous case, the cameras had optical axes parallel to each other, perpendicular to the bottom of the tank, and separated by a distance of 30 cm. Compared to Case Study 1, these lower parameters were due to a drastic reduction in the number of markers recognised by the script in images captured during manual stirring of the water (therefore, fewer observations could be used in mathematical modelling). As can be deduced by observing the contour lines projected from the optical deformation field diagrams on the *xy* plane (Fig. 7c, Fig. 8c), the greater the perturbations, the more irregular the vector field becomes. The automated detection and modelling procedure had no difficulty with movements caused by the fan – at any speed – nor with stabilisation of the water once the cause of the disturbance had been removed (Fig. 11b).

4.3 Case Study 3: different water motion conditions (Test VI – converging-axes dataset)

The third case study examined is also taken from Test VI and again concerns different conditions of water motion. Unlike the first two cases, the optical axes of the cameras were converging and oriented at 15° angles to the zenithal axis, while the baseline was 60 cm. As in Case Study 2, the presence of strong perturbations lowered the number of observations that could be considered as valid by the script and thus useful for mathematical modelling purposes. Again, the vector field of optical deformations tended to lose its regularity depending on the intensity of the perturbation exerted (Fig 9c, Fig. 10c). In this case, the manual perturbation had reached such a strength that the value of the baseline in one of the observed states of motion was incalculable. Analogously to Case Study 2, during detection and modelling operations, the effects of the air movement generated by the fan or the stabilisation of water after stirring did not hinder the performance of the script (Fig. 11c).

5. CONCLUSIONS

When sensing submerged or semi-submerged bodies, even in a controlled environment, mathematical modelling of refraction effects on the cameras incurs into several technical issues related to survey conditions (e.g. water depth level and motion state). We proposed a first approach, which addresses some technical bottlenecks encountered during our experiments. We recorded the intrinsic parameters of a pair of cameras and set up grids of markers in a tank, in order to reconstruct the 3D Ground Truth of these grids and detect the significant points of these targets. Having identified a number of markers in every image, the script then calculated independently the external parameters of each camera through the PnP approach. The program was then able to reproject the 3D Ground Truth points onto every image and calculate the markers' reprojection error. We modelled the effects of the refraction on the displacements Δ of the markers with a non-linear radial distortion model, where the modulo of the distortion increases with the distance *d* from the distortion centre *C*. The model was thus linearized in order to estimate the optimal values of its four parameters (the coordinates x_C, y_C of the distortion centre, and the two parameters *K* and *X)* with the standard non-linear Least Squares approach. We thus obtained a first assessment of the optical distortion field and of its point of minimum (centre *C*). Depending on the type of test, the script finally plotted behavioural diagrams describing the baseline as a function of the water depth level or the liquid motion state. The procedure we developed proved to be effective in the presence of a water level between 5 and 45 cm depth. It also operated for water in stillness, mild motion, or stabilisation conditions. Of course, if there is no water in the tank and no marker is submerged, the automated process of analysis and modelling cannot operate, because the field of optical distortions is absent. Turbulent states of water motion also hinder the modelling of the optical

distortion phenomenon, as they reduce the number of markers detected by the script. Both of these issues respectively suggest possible future developments: on the one hand, the contribution of detected, non-submerged markers to the effects schematised by our mathematical model should be investigated further; on the other hand, it is certainly desirable to improve our approach with regard to images captured in the case of particularly turbulent water.

CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

ACKNOWLEDGEMENTS

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

REFERENCES

AprilRobotics. AprilTag 3. https://github.com/AprilRobotics/apriltag.

Čejka, J.; Bruno F.; Skarlatos D.; Liarokapis F. 2019. Detecting Square Markers in Underwater Environments, *Remote Sensing* 11, no. 4: 459, doi: doi.org/10.3390/rs11040459.

Fischler M. A.; Bolles R. C. 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24, 6 (June 1981), pp. 381–395. doi: doi.org/10.1145/358669.358692

Fitzgibbon A. W. 2001. Simultaneous linear estimation of multiple view geometry and lens distortion, *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, 2001, pp. I-125–I-132, doi: doi.org/10.1109/CVPR.2001.990465.

Gribbon, K.T.; Johnston, C.T.; Bailey, D.G. 2003. A Real-time FPGA Implementation of a Barrel Distortion Correction Algorithm with Bilinear Interpolation, in *Image and Vision Computing NZ*, Palmerston North, November 2003, pp. 408–413.

Krogius, M.; Haggenmiller, 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.

Maas, H.G. 2015. On the Accuracy Potential in Underwater/Multimedia Photogrammetry. *Sensors*, 15, 18140–18152.

Mandlburger, G. 2018. A case study on through-water dense image matching. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 2018, pp. 659–666.

Marchand E.; Uchiyama H.; Spindler F. 2016. Pose Estimation for Augmented Reality: A Hands-On Survey, in *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 12, pp. 2633-–2651, 1 Dec. 2016, doi: doi.org/10.1109/TVCG.2015.2513408.

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, pp. 231–248.

Menna, F.; Nocerino, E.; Drap, P.; Remondino, F.; Murtiyoso, A.; Grussenmeyer, P.; Börline, N. 2018. Improving underwater accuracy by empirical weighting of image observations. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-2, 2018, pp. 699–705.

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, pp. 3400–3407.

Otsu, N. 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics, 9, no. 1, pp. 62-66, doi: 10.1109/TSMC.1979.4310076.

Pupil-labs. pupil-apriltags: Python bindings for the apriltags3 library. https://github.com/pupil-labs/apriltags.

Russo, M.; Ravanelli, R.; Pini, A.; Flenghi, G.; Martelli, L.; Monti, P.; Crespi, M. 2024. Multimedia photogrammetry 2.0. A first step for underwater cultural heritage application. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W4-2024, pp. 373–380.

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.

Song, Y.; Nakath, D.; She, M. et al. 2022. Optical Imaging and Image Restoration Techniques for Deep Ocean Mapping: A Comprehensive Survey. *PFG* 90, 243–267.

Wang, J. & Olson, E. 2016. AprilTag 2: Efficient and robust fiducial detection. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4193–4198. IEEE.