Comparative Accuracy Assessment of two Low-Cost Devices for Underwater Structure-from-Motion 3D Reconstruction

Lukas Quirin[1], Gunnar Lelle-Neumann[1], Ferdinand Maiwald[1]

[1] Chair of Optical 3D-Metrology, Geodetic Institute, TUD Dresden University of Technology, 01069 Dresden, Germany
(lukas.quirin, gunnar.lelle-neumann, ferdinand.maiwald) @tu-dresden.de

Keywords: underwater photogrammetry, structure-from-motion, low-cost, 3D reconstruction

Abstract

The accurate three-dimensional documentation of underwater environments is critical for infrastructure integrity assessments and archaeological studies, yet it often relies on costly equipment. Advances in low-cost Structure-from-Motion techniques provide accessible alternatives to conventional methods, enabling detailed reconstructions using consumer-grade devices. This study evaluates and compares the 3D reconstruction accuracy of two low-cost imaging systems, GoPro HERO13 and Huawei P30, in underwater and dry conditions, using three test objects with varying surface characteristics. High-precision reference models were generated via an ATOS Triple Scan system to quantify absolute accuracy. Results highlight significant performance differences between the cameras in dry conditions, with the GoPro showing improved reliability in underwater conditions, albeit facing challenges with lighting and low contrast environments. However, the Huawei P30 offered superior image resolution but suffered from usability issues in aquatic settings. Overall, dry condition scans yielded higher-quality point clouds, while underwater reconstructions were affected by environmental factors such as reflections, contrast, and lighting limitations. These findings provide practical insights into optimizing workflows for low-budget underwater photogrammetry and suggest potential improvements, including enhanced lighting setups and alternative imaging devices for future studies.

1. Introduction

Accurate three-dimensional (3D) documentation of underwater environments is essential for evaluating the structural integrity of submerged infrastructure such as dams, pipelines or offshore platforms, as well as for repair operations or monitoring sites affected by potential pollution hazards including underwater chemical or ammunition residues. Automatic 3D surveying plays a key role in fulfilling these tasks remotely with a spectrum of uncrewed systems, such as remotely operated (underwater) vehicles (ROV), autonomous underwater vehicles (AUV) or robots (Grömer et al. (2024); Kahmen and Luhmann (2022); Diamanti et al. (2025)).

Conventional underwater surveying methods, including high resolution imaging sonars and laser-based techniques, often require expensive instrumentation. Advances in photogrammetry and Structure-from-Motion (SfM) techniques enable detailed 3D reconstructions from standard imagery. Assuming good visibility conditions, this creates opportunities for low-cost underwater mapping (Menna et al. (2018)). Affordable camera systems such as action cameras or smartphone cameras could significantly reduce costs, but their applicability especially in turbid water conditions requires further investigation.

This study presents a comparative accuracy assessment of two imaging devices for underwater SfM-based 3D reconstruction, giving practical workflow recommendations for low-budget underwater inspection and survey tasks. Aiming to quantify performance differences and identify influencing factors such as optical distortion, image quality and lighting, this study is inspired by recent work by Grömer et al. (2024), Lochhead and Hedley (2022) and Tanduo et al. (2025).

For a comprehensive accuracy evaluation, a good reference model, which ideally has the exact measures of the test object, is crucial, but often difficult to accomplish in practice. The presented work aims to enhance prior research by providing absolute accuracy of the tested methods with high precision reference scans of test objects using a GOM ATOS Triple Scan system (Carl Zeiss GOM Metrology GmbH). The underwater and above-water reconstructions can then be compared to the highly accurate reference model.

2. Related Work

Recent advances in underwater photogrammetry have explored a range of imaging devices and reconstruction techniques. The following section reviews relevant studies with regard to low-cost camera systems, underwater SfM workflows and approaches to quantifying reconstruction accuracy.

In Grömer et al. (2024) dam inspections are undertaken on the upstream side of a dam. The authors present a method for underwater dam inspection using a GoPro Hero 10 mounted on an ROV. At the Pack reservoir in Austria, approximately $1\,000\,m^2$ of a gravity dam was surveyed and compared against a terrestrial laser scan (TLS), achieving a mean deviation of $0.04\,m$ ($\pm 0.18\,m$). The authors highlight the importance of GCP weighting in the bundle adjustment and recommend disabling electronic image stabilisation, as it alters the camera's interior orientation.

Lochhead and Hedley (2022) show a suitable workflow for building 3D SfM models, comparing image data captured in dry and wet conditions. The authors systematically evaluated the impact of underwater conditions on SfM accuracy using an ornamental aquarium coral. Photographs were captured both on dryland and in a saltwater tank using a Sony Alpha a6500 on a custom-built imaging platform with controlled LED lighting. The 180-photo wet-lab test achieved a mean measurement error of $0.26\,\%$, comparable to the dryland result ($0.28\,\%$). However, dryland tests consistently produced denser point clouds, and reducing the number of photographs below 90 led to a notable decline in model quality.

In Tanduo et al. (2025) the authors compare the accuracy and effectiveness of a conventional SfM photogrammetry workflow (Agisoft Metashape) for underwater 3D reconstruction with the workflows of two novel neural rendering techniques: Neural Radiance Fields (NeRF), SeaThru-NeRF and 3D Gaussian Splatting (3DGS). Using a data set of 323 images from a submerged Roman archaeological site, the authors evaluated geometric accuracy via M3C2 cloud-to-cloud analysis in CloudCompare. SeaThru-NeRF achieved the highest geometric accuracy (mean deviation of $0.003\,m \pm 0.089\,m$) via M3C2 cloud-to-cloud analysis. Both neural methods reduced processing time significantly, though at the cost of lower point cloud density.

Perfetti et al. (2026) compare three multi-camera configurations for SfM-based 3D mapping of confined spaces: a low-cost GoPro rig, a custom-built multi-camera prototype and an industrialised version integrating visual SLAM. Evaluated against a TLS reference via M3C2 distances, the GoPro rig performed worst with deviations exceeding $30\,cm$, attributed to rolling shutter effects and lens distortion. The study demonstrates the limitations of off-the-shelf action cameras in challenging environments.

Kanki et al. (2026) demonstrate an applied use of underwater photogrammetry for ecological mapping. Using three GoPro HERO8 cameras, the authors reconstructed more than $2\,500\,m^2$ of seafloor topography off Himeshima Island, Japan, at an average resolution of $7.9\,cm$. While the study illustrates the practical potential of consumer-grade action cameras for large-scale surveying, no quantitative accuracy assessment against a high-precision reference was performed.

The reviewed studies vary considerably in scale, equipment and evaluation methodology. Several rely on TLS references at centimetre-level accuracy (Grömer et al. (2024); Perfetti et al. (2026)), manual measurements (Lochhead and Hedley (2022)) or relative comparisons between methods (Tanduo et al. (2025)), while others omit a quantitative accuracy assessment entirely (Kanki et al. (2026)). The present study addresses this gap by providing absolute accuracy benchmarks at the sub-millimetre level through high-precision ATOS reference scans, using exclusively consumer-grade devices operated handheld. Furthermore, the evaluation is conducted directly on point cloud level rather than on derived 3D surface models.

3. Materials, Equipment and Experimental Design

The following sections describe the test objects, imaging devices and reference scanner, as well as the experimental procedure for the dry and underwater scans.

3.1 Test Objects

Three objects of comparable size were selected for this study: A coral stone with a diameter of nearly $10cm$, a ceramic replica of the Dresdner Frauenkirche (a prominent local church) with a diameter of $8cm$ and a ball valve (size DN15, $1/2''$) commonly used in water supply systems (fig. 1). Comparable to real-world uniform underwater objects, the coral stone and the church replica have both a similar matte white color. However, the two objects differ significantly in their degree of surface roughness and structural details. Besides its smoother surface, the church replica is nearly rotational symmetric, which can be a challenge for photogrammetric scans (Remondino and El-Hakim (2006)). The coral stone on the other hand shows many complex, asymmetric features, including undercuts and holes. In contrast to the other two objects, the ball valve represents a typical component found in technical infrastructure, featuring optical surface characteristics that pose particular challenges for photogrammetric reconstruction: threaded recesses and a predominantly low-textured, reflective surface (Remondino and El-Hakim (2006); Luhmann et al. (2020)).

Figure 1. Test objects used for photogrammetric reconstruction: coral stone, church replica, and ball valve

Initially, an octagonal test object was considered for this study, as referenced in the preliminary project outline. However, this object was replaced by a ball valve. This decision was motivated by two considerations: First, rotational symmetry as a challenging geometric property was already represented by the ceramic replica, rendering the octagon partially redundant in terms of the investigated surface characteristics. Second, the ball valve offers a more industrially relevant test case while also being partly rotational symmetric.

3.2 Equipment

ATOS Triple Scan system The aim of this study is to obtain accurate 3D point clouds of underwater objects based on acquired image data by using low-cost devices. For the reference point clouds, all objects are scanned with the ATOS Triple Scan system, a high precision structured-light 3D scanner, manufactured by Carl Zeiss GOM Metrology GmbH, Zeiss Group. The scans were conducted in the laboratory of the Chair of Photogrammetry, TUD Dresden University of Technology. Based on fringe projection technology, the “Triple Scan” concept uses three measurement angles from a single sensor position. The sensor consists of one projector and two cameras. The scanner’s accuracy is roughly estimated to be 1:50 000 ($20\mu m$) in in-plane direction and 1:25000 ($40\mu m$) in out-of-plane direction for a measurement volume of $1m^3$ (GOM (2010)). These values are influenced by multiple factors, like the measurement environment or the object’s surface properties. The reference point clouds are also used for scaling the image data point clouds. The scanner in action is shown in figure 2.

Cameras To compare measurement accuracy, noise level and 3D point cloud quality respectively, experiments will be conducted using two different low-cost camera systems both underwater and in dryland conditions for each object. The cameras used for the scans are the GoPro HERO13 and the smartphone camera of the Huawei P30, a Leica Triple camera (see table 1). For the experiments, only the main wide camera is used for the smartphone scans. For the underwater scans, the GoPro camera and the smartphone both will be encapsulated in waterproof cases, commonly used for diving. The GoPro Hero13 Black is a compact, waterproof action camera manufactured by GoPro, Inc. The camera supports both photo and video modes, allowing the extraction of individual video frames as an alternative to dedicated photo series for SfM-based reconstruction. It is equipped with a $1/1.9$” CMOS sensor with a resolution of 27.13 megapixels (MP) and captures video footage at up to 12.9 MP resolution at 60 $frames\ per\ second$ (fps). Without additional housing, the camera is rated waterproof to a depth of $10m$, however, it was decided to use the camera in a waterproof case. The built-in fixed wide-angle lens, with a minimal focal length equivalent of approximately $12mm$ ($35mm$ equivalent), provides a large field of view, enabling the capture of the test objects from relatively close distances. However, the wide-angle lens introduces considerable radial distortion, which must be accounted for during the photogrammetric processing pipeline. The phone camera used in this study offers a higher native image resolution and a narrower field of view with less radial distortion, compared to the GoPro Hero13. In the context of this paper, the GoPro Hero13 camera will be referred to as the “GoPro camera” and the Huawei/Leica phone camera as the “phone camera.”

Camera	Mode	Resolution	Image Rate
GoPro	photo	27.13 MP	2 fps
	video	5.48 MP	120 fps
	video	12.9 MP	60 fps
phone	photo	9.98 MP	single shots
	video	2.07 MP	60 fps
	video	8.29 MP	30 fps

Table 1. Camera configurations used for the experiments

Software Several software packages were employed throughout the different stages of this study. The ATOS Triple Scan system was operated using GOM Inspect, a dedicated software for structured-light 3D scanning developed by GOM GmbH. Following the acquisition of GOM by the Zeiss Group (Carl Zeiss GOM Metrology GmbH), GOM Inspect has been succeeded by ZEISS Inspect, which was used for post-processing and editing of the reference point cloud data. For the photogrammetric reconstruction workflow, Agisoft Metashape Professional Version 2.2.2 (Agisoft LLC) was used. The subsequent accuracy assessment, comprising the comparison of the photogrammetric point clouds against the ATOS reference scans, was carried out using CloudCompare Version 2.13.2 (Kharkiv), an open-source software for 3D point cloud and mesh processing. A detailed description of the processing steps performed in each software is provided in the following sections.

3.3 Experimental Design

The experimental design comprises three stages: the acquisition of reference scans, the image-based object scans under dry and underwater conditions, and the subsequent post-processing and accuracy assessment.

3.3.1 Object Scans

Reference Scan Prior to scanning with the ATOS Triple Scan system, self-adhesive circular reference point markers were attached to each object. These markers serve as common tie points between overlapping scan positions, enabling the registration of multiple partial scans into a unified coordinate system. It is recommended to avoid placing the markers in symmetric patterns, as this may lead to ambiguities during the automatic registration process.

After preparation, each object was placed on a small support structure, a matchbox for instance, positioned on a box in front of the scanner. The use of a small support object was intended to visually decouple the test object from the surrounding surface area. In standby mode, the scanner’s projector illuminates the measurement area with blue light. Within this illuminated area, all surfaces visible to the cameras are captured, including the underlying box’s surface (fig. 2, left). This excess data had to be removed manually during post-processing in GOM inspect or ZEISS inspect respectively. The alignment of the object within the scanner’s field of view was facilitated by two red laser beams, one emitted per camera, which indicate the focal intersection of both sensor lines of sight. The scanner’s position could also be adjusted. Once the object and scanner were correctly positioned relative to each other, the laser beams were switched off and the scanning process was initiated.

Throughout the scanning process, the scanner’s projector projects a fringe pattern of alternating bright and dark stripes onto the measurement area (fig. 2, right). By shifting these patterns across the object surface, the two cameras are able to derive three-dimensional spatial information from each individual scan. A complete reference scan for each object consisted of approximately 50 to 120 partial scans, depending on the geometric complexity of the object. Between each partial scan, the object’s orientation was slightly changed, typically by only a few degrees. This incremental rotation ensures that the system identifies a sufficient number of reference point markers from one partial scan to the next, thereby enabling a reliable registration. This cycle of repositioning and scanning was repeated until the entire object surface was captured.

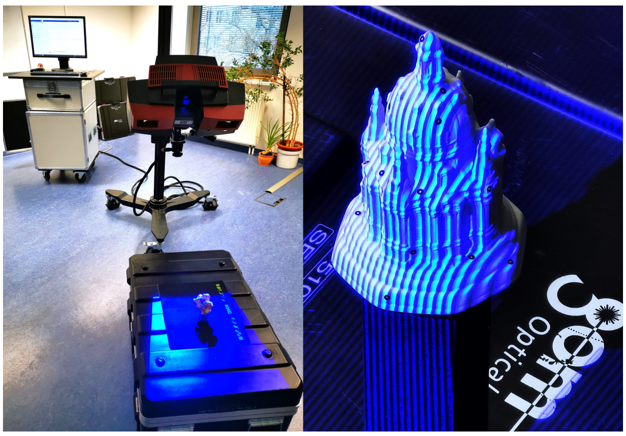

Figure 2. ATOS Triple Scan (left) and church replica with the projected fringe pattern (right) during the scan process

After removing the excess data in the post-processing, the reference point clouds (fig. 3) were exported and saved for further analysis in CloudCompare.

Figure 3. Reference point clouds of the ball valve (left), coral stone (center) and the church replica (right).

Object Scans in Dry Conditions The experiments conducted in dry conditions serve as a best-case scenario for the low-cost reconstructions, as they indicate the achievable point cloud quality without the effects of the underwater environment. For each of the three test objects, image data was acquired using both the GoPro camera and the phone camera. As recommended by Lochhead and Hedley (2022), the image capture followed a spiral and meridian-like pattern around each object, ensuring multi-view coverage from varying angles and heights. For each camera, image data was obtained both as a photo series and as video footage, resulting in four independent data sets per object under dry conditions. For the GoPro camera, photos were captured using its internal time-lapse function at an image rate of 2 fps, yielding approximately 200 to 300 images per object. As the phone camera does not offer a comparable function, individual photos had to be taken manually, resulting in approximately 100 images per object. Each video recording lasted approximately two to three minutes. From the video footage, recorded at frame rates of 120 fps (GoPro) and 60 fps (phone) respectively, approximately 1 000 to 1 500 individual frames were initially extracted. This deliberately high number of extracted frames was necessary to account for the phone camera’s autofocus behaviour, which intermittently attempted to focus on surrounding objects during camera movement. During the subsequent alignment preparation in Agisoft Metashape, low-quality frames were filtered out, reducing the usable data sets to several hundred images. The GoPro camera, in contrast, operates with a fixed focus. For the video footage of the ball valve, the frame rates were reduced to 60 fps (GoPro) and 30 fps (phone), resulting in higher image resolutions (see table 1).

The dry condition scans were conducted in the Hydraulic Engineering Hall of the Hubert Engels Laboratory at the Institute of Hydraulic Engineering and Technical Hydromechanics (IWD), TUD Dresden University of Technology, where the underwater experiments were also carried out. This ensured comparable illumination conditions across all experiments. An exception are the scans of the ball valve, which were performed in the workshop of the Chair of Optical 3D-Metrology at TUD, where a vise was used to position the valve to achieve optical access from all directions. Additionally, for the church replica, an additional scan was conducted under near-ideal conditions using the phone camera, with a spotlight illumination and a felt pad as a plain, low reflection background surface. With the exception of the ball valve, a stainless steel ruler was placed next to each object to provide defined reference points for scaling during post-processing in Agisoft Metashape. For the ball valve, scaling was based on its known nominal dimensions, as it is a standardised industrial component.

Underwater Scans The underwater image acquisition was conducted in a water tank of approximately $1m^3$ volume, located in the Hydraulic Engineering Hall (fig.4, left). To faciliate optical access, the objects were placed on a platform inside the tank. The initial platform configuration, a structure mounted on the wall of the tank, was found to be impractical for achieving a 360 optical access to the objects (fig.4, center). Furthermore, this arrangement positioned the objects at a greater distance from the overhead light source mounted above the tank. Only the underwater scans of the church replica taken with the GoPro are used with this configuration. For all subsequent underwater scans, the objects were placed on a freestanding table structure within the tank, which provided an improved optical accessibility and better illumination conditions (fig 4, right). This table structure could also be modified, regarding scale bar integration and surface characteristics of the object support. The first two underwater experiments were conducted without additional spotlight illumination, which proved to be insufficient for a successful reconstruction (see section 5).

The underwater experiments were designed to replicate the acquisition configurations used under dry conditions for all three test objects. However, operating the phone camera inside the waterproof housing prevented taking manual single-shot photos. Consequently, all underwater image data for the phone camera was only acquired as video footage. For the underwater scans, filming the video footage took one to three minutes longer than for the dry scans. This is primarily attributed to reflections on the water surface, obstructing the operator’s ability to visually monitor the camera’s position and orientation relative to the object. An overview of all underwater scan configurations is provided in table 2.

Figure 4. Water tank, used for the underwater images (left). Inside, the objects placed on a support structure (center, right).

4. Result Assessment

The image data sets, obtained in the experiments (see section 3.3) were then processed with the photogrammetry software Agisoft Metashape, in order to create the 3D point clouds. The point clouds were then loaded into CloudCompare, and evaluated for their accuracy.

4.1 Post Processing in Agisoft Metashape

The photogrammetric reconstruction was carried out in Agisoft Metashape Professional (see section 3.2). The processing workflow comprised the following steps:

Image Import and Frame Extraction For the video-based data sets, individual frames were extracted within Metashape, with the extraction interval determined by the frame rate and the duration of the recording, targeting approximately 1 000 to 1 500 frames per data set. The first and last seconds of each video were excluded from extraction to avoid frames captured during placement or removal of the camera. Photo series were imported directly, after a quick manual review.

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-979-2026 |

Image Quality Assessment Prior to alignment, the estimated image quality was assessed using Metashape's built-in image quality estimation function. This parameter is calculated based on the sharpness level of the most focused part of each image Agisoft LLC. All images with a quality value below 0.5 were excluded from further processing, as poorly focused images can affect the alignment results.

Image Alignment The image alignment was performed using the highest accuracy setting with sequential preselection, which takes into account the image's sequence number, assuming that consecutive images are overlapping sufficiently. The key point limit was set to 80 000 and the tie point limit to 20 000 per image. During alignment, both interior orientation parameters (focal length, principal point coordinates and lens distortion coefficients) and exterior orientation parameters (three translation components and three Euler rotation angles) were estimated automatically. The initial interior orientation was derived from the Exchangeable Image File Format (EXIF) metadata of the input images. Where necessary, reference points were created manually to support the alignment process. It should be noted that Metashape treats these points as fixed references, requiring careful placement.

Tie Point Filtering Following alignment, the quality of the resulting tie point cloud was assessed. Points observed in fewer than three images were removed. Additionally, points exhibiting high reprojection errors were identified and excluded from alignment aswell. In the model view, the spatial distribution of the tie points was visually inspected: a diffuse, cloud-like distribution of tie points beyond the object boundary indicates a poor alignment result. An example for such an alignment is shown in fig. 5, center. The use of spotlight illumination proved beneficial during this stage, as areas outside the light beam appeared significantly darker in the images. These areas were then disregarded by the software during tie point detection, effectively reducing the number of non-object-related tie points.

Dense Point Cloud Generation After tie point filtering, a dense point cloud was generated using depth maps at high quality with the point confidence calculation enabled. This parameter records the number of depth maps contributing to each point, which can later be used for filtering. Following point cloud generation, all points with a confidence value below 2 were removed to obtain a cleaner result. No 3D surface model was generated, as the accuracy assessment was conducted only on a cloud-to-cloud basis. Data sets for which the alignment failed or which produced strongly distorted or doubled features in the point cloud (see fig. 5, bottom) were excluded from further analysis. In the point cloud, so called markers can be placed. These markers can be used for scaling the model, by defining distances between these markers. In this study, the known width of the stainless steel ruler, placed next to the objects, was used for scaling.

4.2 Point Cloud Comparison in CloudCompare

The accuracy assessment was carried out in CloudCompare (see section 3.2). Both the ATOS reference point cloud and the photogrammetric point cloud of each object were imported into CloudCompare. It was found that the scaling between Agisoft Metashape and CloudCompare did not transfer reliably. While partly the different default unit systems (metres in Agisoft Metashape and millimetres in CloudCompare) can be responsible, the observed scaling inconsistencies appeared to be non-systematic. Consequently, each photogrammetric

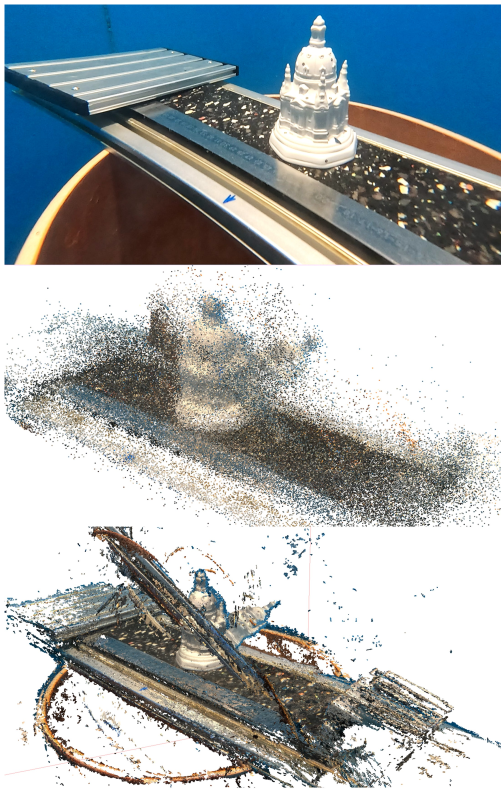

Figure 5. Example of a failed alignment: Object (top), resulting dense point cloud (center) and derived 3D model (bottom)

point cloud was manually scaled to match the corresponding ATOS reference point cloud within CloudCompare. Following the scaling, the two point clouds were coarsely aligned by manually selecting corresponding point pairs. Then, the *Iterative Closest Point (ICP)* registration tool was applied for fine alignment. This algorithm iteratively minimises the distance between two point clouds by computing optimal translation and rotation parameters until convergence is reached. The accuracy of each reconstruction was then evaluated using CloudCompare's *cloud-to-cloud distance* computation, which calculates the nearest-neighbour distance from each point in the photogrammetric point cloud to the reference point cloud. From this analysis, the mean distance and standard deviation were derived as quantitative accuracy measures. Additionally, color-coded distance maps were generated to visualise the spatial distribution of deviations across the object surface.

5. Analysis and Results

5.1 Feasibility of generating the point clouds

Under dry conditions, all experimental configurations successfully produced usable point clouds. For the cloud-to-cloud distance analysis, only the best-performing data set per object and camera — either the photo series or the video footage — was selected based on the visual quality of the resulting point cloud.

For the underwater experiments, the success rate was considerably lower. Out of 15 underwater scan configurations, 9 yielded a point cloud (table 2). The ball valve produced a point cloud in all 4 configurations. The coral stone yielded point clouds in 2 out of 3 configurations. The church replica proved to be the most difficult object, with only 3 out of 8 configurations producing a point cloud. Notably, the two experiments conducted without spotlight illumination both failed entirely. Among the successful reconstructions, those marked with (*) in table 2 were not considered for further analysis in Cloudcompare, because other data sets with the same or very similar (in case of the church replica) experimental configurations showed better results. The configuration result marked with (**) produced a point cloud that was doubled, resulting in the church replica model having two tips and two intertwined main bodies. Besides that, the point cloud did not show too many artefacts. Because of its doubled structure, it was not considered for further analysis. The successful church replica phone camera scan was achieved by combining the higher quality image frames of three shorter video footages that did not capture the entire church.

Object	Camera	Spotlight	Image Source	Underlying Surface	Point Cloud
valve	GoPro	yes	video	high contrast	yes
	GoPro	yes	video	high contrast	yes*
	phone	yes	video	high contrast	yes
	phone	yes	video	high contrast	yes*
coral	GoPro	yes	video	high contrast	yes
	GoPro	yes	video	high contrast	no
	phone	yes	video	high contrast	yes
replica	GoPro	yes	photo	aluminium, tank floor	yes
	GoPro	yes	video	aluminium, tank floor	no**
	GoPro	yes	video	shiny	yes*
	GoPro	yes	video	high contrast	no
	GoPro	no	video	aluminium, tank floor	no
	GoPro	no	video	aluminium, tank floor	no
	phone	yes	video (combined)	high contrast	yes
	phone	yes	video	high contrast	no

Table 2. Overview of underwater scan configurations and results

5.2 Point Cloud Quality

The accuracy of the photogrammetric point clouds was evaluated using two metrics: First, the cloud-to-cloud distance to the ATOS reference point cloud, expressed as mean distance and standard deviation, and also shown as color-coded distance maps (fig. 6, 7, 8), and second, the proportion of points falling within defined distance thresholds ($< 2\,mm$, $< 1\,mm$ and $< 0.5\,mm$) as a measure of overall point cloud cleanliness.

Cloud-to-Cloud Distance Table 3 summarises the mean cloud-to-cloud distances and standard deviations for all successfully reconstructed image data sets. Under dry conditions, the phone camera consistently achieved lower mean distances than the GoPro across all three objects. The lowest mean distance was obtained for the coral stone with the phone camera ($0.10\,mm \pm 0.09\,mm$), while the highest was recorded for the ball valve with the GoPro ($1.21\,mm \pm 1.02\,mm$). For the phone camera, an additional scan of the church replica with a plain background and spotlight illumination further reduced the mean distance from $0.27\,mm$ to $0.22\,mm$. In the underwater experiments, the GoPro yielded mean distances comparable to or lower than those obtained under dry conditions for the coral stone ($0.46\,mm$ vs. $0.45\,mm$) and the church replica ($0.51\,mm$ vs. $0.65\,mm$). For the ball valve, the underwater result ($0.80\,mm$) also improved compared to the dry scan ($1.21\,mm$). The phone camera's underwater results remained comparable to the dry conditions for the ball valve ($0.64\,mm$ vs. $0.62\,mm$), while the coral stone showed an increase from $0.10\,mm$ to $0.46\,mm$. The church replica exhibited the largest degradation underwater, with the mean distance increasing from $0.27\,mm$ to $1.04\,mm$ ($\pm 1.13\,mm$), based on the combination of three underwater data sets for the phone camera.

Object	Data Source	Mean [mm]	Std. Dev. [mm]
GoPro			
valve	video frames, dry conditions	1.21	1.02
	photos, underwater	0.80	0.67
coral	photos, dry conditions	0.45	0.53
	video frames, underwater	0.46	0.52
replica	photos, dry conditions	0.65	0.49
	photos, underwater	0.51	0.44
Phone Camera			
valve	video frames, dry conditions	0.62	0.53
	video frames, dry conditions (structured ground)	0.64	0.53
	video frames, underwater	0.64	0.48
coral	photos, dry conditions	0.10	0.09
	video frames, underwater	0.46	0.35
replica	photos, dry conditions	0.27	0.16
	photos, dry conditions, plain background	0.22	0.13
	video frames, underwater (3 data sets)	1.04	1.13

Table 3. cloud-to-cloud distance results for both the GoPro and the phone camera

Point Cloud Density and Distribution Table 4 provides the total point count and the proportion of points within three distance thresholds for each data set. Across all objects and conditions, the phone camera produced significantly denser point clouds than the GoPro, typically by a factor of 5 to 10. Under dry conditions, the phone camera achieved the highest proportions of points within $0.5\,mm$: $99.12\,\%$ for the coral stone and $98.77\,\%$ for the church replica scan in the plain background. The GoPro reached $69.20\,\%$ and $47.74\,\%$ for the same objects respectively. For the ball valve, both cameras remained below $55\,\%$. In the underwater experiments, the phone camera's proportions within $0.5\,mm$ decreased notably compared to dry conditions, from $99.12\,\%$ to $61.95\,\%$ for the coral and from $92.87\,\%$ to $36.61\,\%$ for the church replica. The GoPro's underwater values, on the other hand, remained comparable to or slightly above the dry condition results.

The spatial distribution of the cloud-to-cloud distances is visualized in color-coded distance maps for all three objects (fig. 6, 7, 8). In figure (6) the color-coded distance maps are shown for the ball valve point cloud, with (a) showing the results for the GoPro camera scan under dry conditions, (b) the phone camera scan under dry conditions, (c) the GoPro scan and (d) the phone scan both for the underwater image data. The results in figure 6 match with the results in 4, the higher amount of points for the phone camera scans are represented by more intense colors. Both cameras' point clouds show lower accuracy in areas that typically show reflections in at least some of the pictures, such as edges or the ball of the valve. None of the scans was able to reproduce the thread correctly.

Figure 7 presents the distance maps for the four point clouds of the coral, (a) for the GoPro and (b) for the phone camera, both in dry conditions. The other two graphs represent the underwater scans, (c) the GoPro and (d) the phone camera. The

Object	Camera	Data Source	Points (total)	< 2mm [mm]	< 2mm [%]	< 1mm [mm]	< 1mm [%]	< 0.5mm [mm]	< 0.5mm [%]
valve	ATOS	reference scan	697 843	–	–	–	–	–	–
	GoPro	video frames, dry conditions	56 518	46 923	83.02	30 687	54.30	14 227	25.17
	GoPro	photos, underwater	69 441	65 670	94.57	51 753	74.53	26 380	37.99
	phone	video frames, dry conditions	406 474	396 014	97.43	349 739	86.04	221 704	54.54
	phone	video frames, underwater	508 605	495 646	97.45	430 444	84.63	252 430	49.63
coral	ATOS	reference scan	6 804 196	–	–	–	–	–	–
	GoPro	photos, dry conditions	87 830	86 531	98.52	79 197	90.17	60 777	69.20
	GoPro	video frames, underwater	223 136	217 245	97.36	203 517	91.21	154 779	69.37
	phone	photos, dry conditions	792 847	792 832	100.00	792 706	99.98	785 862	99.12
	phone	video frames, underwater	1 201 358	1 200 481	99.93	1 104 251	91.92	744 233	61.95
replica	ATOS	reference scan	238 150	–	–	–	–	–	–
	GoPro	photos, dry conditions	227 997	224 054	98.27	183 369	80.43	108 848	47.74
	GoPro	photos, underwater	208 352	205 685	98.72	186 075	89.31	130 565	62.67
	phone	photos, dry conditions	184 413	184 382	99.98	183 351	99.42	171 267	92.87
	phone	photos, dry conditions, plain background	618 251	618 231	100.00	618 066	99.97	610 650	98.77
	phone	video frames, underwater (3 data sets)	841 279	744 357	88.48	544 489	64.72	307 990	36.61

Table 4. Point distribution analysis of cloud-to-cloud distances

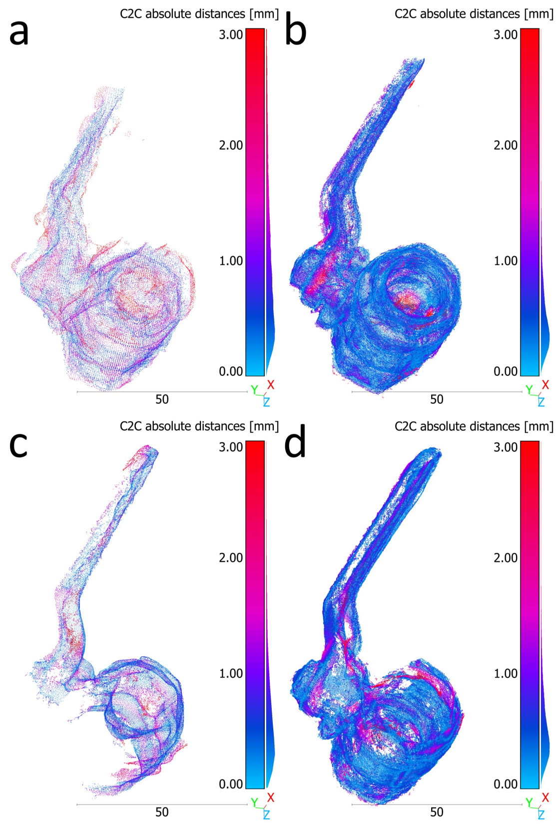

Figure 6. Cloud-to-cloud-distance between the photogrammetric scans of the ball valve and the reference point cloud

distribution of the point distances to the reference scan matches the results shown in table 4. In point cloud (b) the target marker stickers are visible as areas with higher distance values. The underwater scans show areas with higher point distances around the edges. Overall the coral scans seem to adequately reproduce the structure.

The point clouds distances for the church replica are shown in

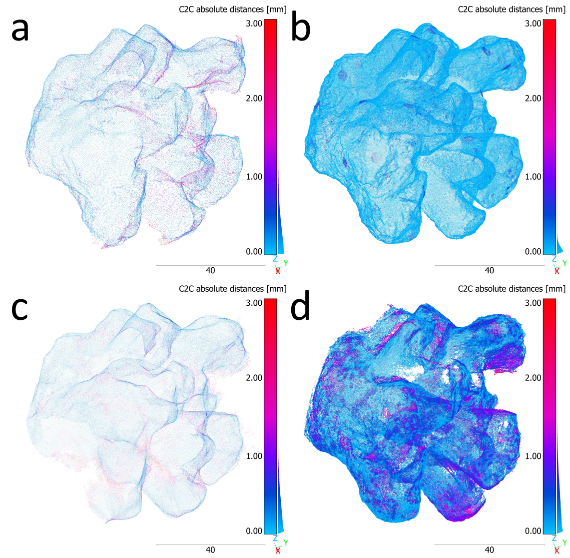

Figure 7. Cloud-to-cloud-distance between the photogrammetric scans of the coral stone and the reference point cloud

figure 8, (a), (b) and (c) represent the scans in dry conditions and (d) and (e) the underwater scans. The image data of (b) and (d) are captured with the GoPro, the remaining three with the phone camera. Clearly visible is, that both point clouds reconstructed from the phone camera data in dry conditions are matching very well to the reference point cloud. The underwater phone scan shows strong shifts in the point cloud and only a partial representation of the object. The improvement for the GoPro scans from the dry condition data to the underwater data (see tables 3 and 4) is visible in the point distance graphs as well: (d) shows less purple and pink areas than (b).

6. Discussion

6.1 Point Cloud Generation

The failure of the two experiments conducted without spotlight illumination confirms that dedicated lighting is essential for underwater image acquisition in the given tank setup. The ambient lighting alone did not provide sufficient contrast for reliable feature detection. The church replica’s low success rate underwater (3 out of 8 configurations) can be attributed to its smooth, nearly rotationally symmetric and low-textured surface. These properties, which already pose a challenge for photogrammetric reconstruction under dry conditions (Remondino and El-Hakim (2006); Luhmann et al. (2020)), combined with the adverse effects of the water medium, severely limited the number of detectable features. For the successful church replica underwater scan with the phone camera the image frames of three shorter video footages had to be combined, since the main scan did not provide enough sharp image frames for a successful reconstruction. In contrast, the coral stone scan benefited from the stone’s rich surface texture and its asymmetric features, which provided a sufficient number of detectable features even underwater. The ball valve, despite its reflective and low-textured metallic surface, produced point clouds in all underwater configurations. The reason could be the high-contrast background surface used during these experiments, which supported the alignment process with additional tie points in the object’s surroundings. This observation suggests that even the underwater photogrammet-

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-979-2026 |

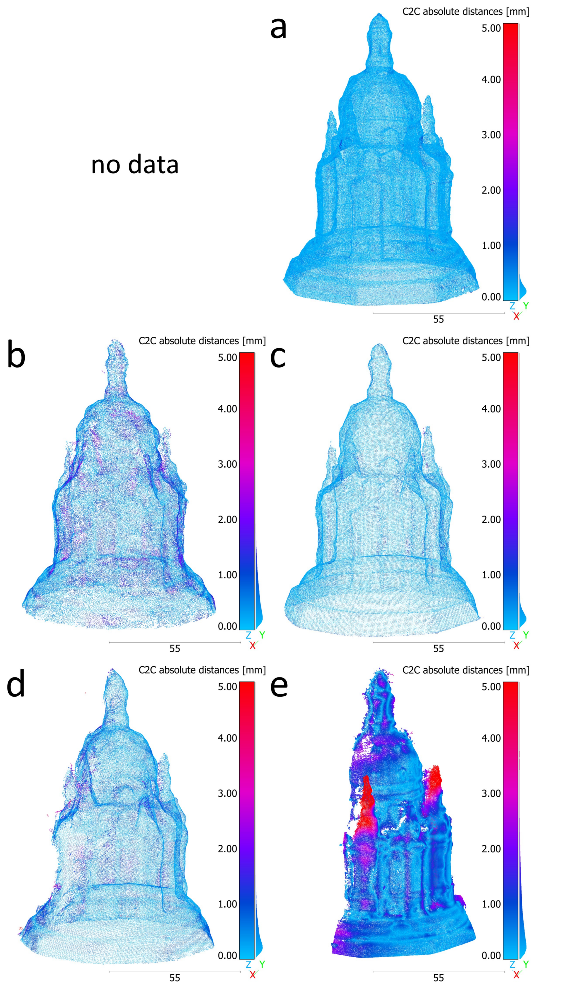

Figure 8. Cloud-to-cloud-distance between the photogrammetric scans of the church replica and the reference point cloud

ric reconstruction of objects with challenging surface properties can be successful, if the surrounding conditions are supportive.

6.2 Point Cloud Quality

The phone camera consistently outperformed the GoPro under dry conditions, both in terms of mean cloud-to-cloud distance and point cloud density (tables 3 and 4). This is primarily attributed to its higher native image resolution and narrower field of view with less radial distortion. To the underwater environment the two cameras exhibited notably different responses. The GoPro's mean distances remained comparable to or improved over the dry condition results for all three objects. This robustness can partly be explained by its fixed-focus lens, which maintained consistent image sharpness regardless of the medium and may partly be explained by the experimental setup: the underwater scans consistently used spotlight illumination. The phone camera, on the other hand, showed a clear degradation underwater. The severely worse results for the church replica (from $0.27\,mm$ to $1.04\,mm$) must be interpreted carefully, since this point cloud is the result of the combined alignment of three different video footages (the position of the object was not changed though). And as figure 8 shows it could not be reconstructed entirely. The counterintuitive improvement of the GoPro's results underwater is also remarkable, because the underwater image data showed a much stronger distortion than the data for dry conditions. It indicates that the alignment and automatic calibration in Agisoft Metashape are quite robust. The results confirm that surface texture and geometry are strong factors for reconstruction quality. The coral stone achieved the best results across nearly all configurations. As expected, the church replica suffered from its near-rotational symmetry and smooth surface, limiting the number of matching points. The ball valve's reflective metallic surface posed challenges for both cameras, especially underwater, as it can be seen in the edge areas and the valve ball in figure 6. Due to its surface characteristics, the ball valve was an especially challenging object, even for the ATOS reference scan. The influence of surface reflection can also be seen in figure 7 (b), where an otherwise clean point cloud is disturbed by the comparably more reflective target marker stickers.

The threshold analysis (table 4) reveals that a higher point density does not necessarily indicate a better reconstruction. The phone camera's underwater scan of the church replica produced 841 279 points (more than any other replica configuration) yet only $36.61\,\%$ fell within $0.5\,mm$, the lowest proportion across all data sets. This indicates that a substantial share of the additional points represents noise rather than accurate surface geometry. The GoPro, despite producing significantly fewer points, maintained stable or slightly improved proportions within $0.5\,mm$ underwater, consistent with the robustness of its fixed-focus lens. Notably, the phone camera's $< 2\,mm$ values remained high even underwater (e.g. $99.93\,\%$ for the coral), while the $< 0.5\,mm$ proportions dropped substantially. This suggests that the underwater environment primarily affects fine-scale accuracy rather than introducing large-scale geometric distortions.

Overall, the results show the expected challenges in SfM based reconstruction. Illumination and geometry and surface texture are important factors to consider. Especially the GoPro's performance improved with better lighting.

7. Conclusion and Future Work

This study assessed the potential of low-cost imaging devices, including the GoPro HERO13 camera and the phone camera of the Huawei P30, for underwater SfM-based 3D reconstruction. The experiments revealed that while the phone camera outperformed the GoPro in dry conditions, the GoPro showed improved results underwater. Underwater, the GoPro could outperform itself and yielded similar adequate results as the smartphone camera underwater. Environmental factors, such as lighting, contrast and object texture, were identified as critical to reconstruction quality, with reflective and low-texture surfaces posing significant challenges. Despite inherent limitations, such as radial distortion in the GoPro and operational constraints with the smartphone housing, the results suggest that low-cost photogrammetry can provide practical solutions for underwater surveys, particularly when paired with careful planning and optimized workflows.

Future work will explore the integration of additional sensors to enhance data acquisition, particularly in low-visibility or highly reflective environments. This could include the development of hybrid approaches combining photogrammetry with complementary technologies, such as acoustic or optical sensors, to address current limitations in underwater reconstructions. Additionally, more robust experimental setups, improved image acquisition techniques, and the use of multiple synchronized cameras are planned to further refine reconstruction accuracy and scalability for larger and more complex underwater environments. These advancements aim to expand the applicability of low-cost underwater 3D reconstruction methods to industrial and scientific contexts.

Acknowledgments

We want to thank the staff of the Hubert Engels Laboratory at the Institute of Hydraulic Engineering and Technical Hydromechanics for providing room and equipment for the experiments. Further, we want to thank Carl Zeiss GOM Metrology GmbH for provision of the ATOS Triple Scan and the staff of the Chair of Photogrammetry for providing room and a brief introduction to how to operate the scanner.

References

Agisoft LLC, 2026a. Agisoft Metashape Professional – version 2.2.2. https://www.agisoft.com/. Accessed: 2026-04-01.

Agisoft LLC, 2026b. Agisoft metashape user manual. https://www.agisoft.com/downloads/user-manuals/. Accessed: 2026-04-01.

Carl Zeiss GOM Metrology GmbH, 2026a. Atos triple scan - technical specifications. https://www.zeiss.com/metrology/en/systems/optical-3d/3d-scanning/atos/atos-compact-scan.html. Accessed: 2026-04-05.

Carl Zeiss GOM Metrology GmbH, 2026b. ZEISS Inspect - 3D inspection software. https://www.zeiss.com/metrology/en/software/zeiss-inspect.html. Accessed: 2026-04-01.

CloudCompare, 2026. CloudCompare - 3D point cloud and mesh processing software, version 2.13.2 (kharkiv). https://www.cloudcompare.org/. Accessed: 2026-04-02.

Diamanti, E., Ødegård, Ø., Mentogiannis, V., Koutsouflakis, G., 2025. Underwater Drones as a Low-Cost, yet Powerful Tool for Underwater Archaeological Mapping: Case Studies from the Mediterranean. *Journal of Computer Applications in Archaeology*, 8(1): 10–24.

GOM, 2010. ATOS Triple Scan. data sheet brochure. GOM GmbH.

GOM GmbH, 2018. GOM Inspect - 3D inspection software. Now succeeded by ZEISS Inspect. Accessed: 2026-04-01.

Grömer, M., Nocerino, E., Calantropio, A., Menna, F., Dreier, A., Winiwarter, L., Mandlburger, G., 2024. High-detail and Low-Cost Underwater Inspection of Large-Scale Hydropower Dams. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-2-2024.

Kahmen, O., Luhmann, T., 2022. Monocular Photogrammetric System for 3D Reconstruction of Welds in Turbid Water. *PFG(2022)*, 90:19–35.

Kanki, T., Sano, W., Sannoh, M., Kan, H., 2026. Seafloor vegetation map of man-made boulders reef by underwater photogrammetry: Suggestions for site selections in macroalgal bed creations. *PLoS ONE*, 21(3), e0341865.

Lochhead, I., Hedley, N., 2022. Evaluating the 3D Integrity of Underwater Structure from Motion Worklflows. *The Photogrammetric Record*, 37(177): 35-60.

Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2020. *Close-Range Photogrammetry and 3D Imaging*. De Gruyter.

Menna, F., Agrafiotis, P., Georgopoulos, A., 2018. State of the art and applications in archaeological underwater 3D recording and mapping. *Journal of Cultural Heritage*, Volume 33, Pages 231-248.

Perfetti, L., El-Alailyi, A., Fassi, F., Remondino, F., Sgrenzaroli, M., Vassena, G. P. M., 2026. Ant3D: Evolution of a Multi-Camera Fisheye System for 3D Mapping in Confined Spaces. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W12-2026, 375–382.

Remondino, F., El-Hakim, S., 2006. Image-Based 3D Modelling: A Review. *The Photogrammetric Record*.

Tanduo, B., Matrone, F., Murtiyoso, A., 2025. Initial Experiments on the Use of Radiance Fields for Underwater 3D Reconstruction. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-9-2025.