# FROG: A PORTABLE UNDERWATER MOBILE MAPPING SYSTEM

F. Menna [1], R. Battisti[1], E. Nocerino [2,3], F. Remondino [1]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Web: http://3dom.fbk.eu – Email: <fmenna> <rbattisti><remondino>@fbk.eu
[2] Dipartimento di Scienze Umanistiche e Sociali, University of Sassari, Sassari, Italy; email: enocerino@uniss.it
[3] NBFC, National Biodiversity Future Center, Palermo 90133, Italy

**Commission I**

**KEY WORDS:** vSLAM, Real-Time, Low-Cost, Raspberry, Mobile Mapping, GNSS-denied environments, Indoor, Underwater

**ABSTRACT:**
Browsing the scientific and professional literature it appears that the concept of mobile mapping underwater is not as common as in 'terrestrial' applications. Nevertheless, exploring and mapping the ocean's depths is a priority challenge for humankind, today more than ever. Radio waves, such as the GNSS or UWB signal, have a very limited transmission underwater, resulting in the absence of an underwater global positioning system. Consequently, the main sounding methods (i.e., depth measuring systems) are based on the fusion of inertial and acoustic sensors, which allow for systematic mapping of vast seafloor areas. However, photogrammetric surveying methods are preferred when high resolution and reliable colour information are essential aspects in the project economy. This class of approaches include visual odometry and visual SLAM (vSLAM), which represent a valid tool for navigation and positioning in GNSS-denied environments, such as underwater. In this paper, we present a portable underwater mobile mapping system, named FROG, which implements a vSLAM based solution to guide the survey according to photogrammetric principles. FROG is built upon the Guided Photogrammetry - GuPho concept and, thanks to its modular design, can be used by a diver or installed on a micro ROV and controlled remotely from a support vessel. In the paper, FROG characteristics will be detailed, and its potentialities demonstrated in real case applications at sea and in lakes.

## 1. INTRODUCTION

### 1.1 Underwater MMTs

Mobile mapping technology (MMT) has been recognised as a standard surveying approach for a broad variety of 'terrestrial' applications in urban, natural, and built environments in the last 30 years. Schwarz and El- Sheimy (2004) coincide the birth of the idea of mobile mapping, i.e., mapping form moving vehicles, with early applications of photogrammetry. They emphasize how the advantages in digital imaging and direct georeferencing had fostered the development and popularization of mobile mapping systems, leading to the development of several systems which can be mounted on different terrestrial and aerial platforms or even carried by human operators.

Because of the constraints due to the physical properties of the water medium, the development of mobile mapping systems underwater did not go hand-in-hand with the development of analogous systems above the water. And this is even more critical when it comes to portable mobile mapping systems. Indeed, in underwater applications the term 'mobile mapping' is not so common, although seabed mapping or bathymetry has been a major challenge since ever. Originally, the term bathymetry indicated the ocean's depth with respect to the water surface, while currently is commonly adopted to indicate the measurement and study of the topography of water bodies, including the ocean, rivers, streams, and lakes[1]. Mankind has always been interested in investigating the depths of the oceans (i) to ensure safe navigation and, more recently in history, (ii) to understand the fundamentals of geological and oceanographic processes, (iii) for the exploration and location of mineral and energy resources, (iv) to study and monitoring climate change, (v) for an effective implementation of marine conservation and protection strategies (Smith Menandro & Cardoso Bastos, 2020; Wölfl et al., 2019). In the last years, underwater mapping is gaining more and more attention with the launching of important world-wide programs such as the United Nations Decade of Ocean Science for Sustainable Development, where underwater 3D mapping is set among the highest priorities, and the initiative promoted by GEBCO, General Bathymetric Chart of the Oceans, and the Nippon Foundation to facilitate the complete mapping of the global ocean floor by the year 2030.

Exploration and mapping of remote ocean depths require the employment of autonomous surface vehicles (ASV) and autonomous underwater vehicles (AUV), which rely on automatic navigation, localization and mapping algorithms, such as SLAM (simultaneous localisation and mapping). Differently from terrestrial or aerial SLAM approaches that largely depend on optical sensors, underwater SLAM systems mainly use acoustic based instruments, i.e., sonars (Jiang and al., 2019) most often integrated with inertial measurements (Jørgensen, 2016). Underwater, the GNSS signal is not transmitted, resulting in the lack of a global positioning system. Thus, the fusion between inertial and acoustic measurements is very common underwater providing the six degrees of freedom (6DoF) and allowing systematic mapping of vast underwater areas.

In applications where high resolution and colour information are key, such as underwater archaeology (Menna et al., 2018), marine biology and industrial inspections (Chemisky et al., 2021), optical surveying methods based on photogrammetry have become a standard practice. In particular, visual odometry and visual SLAM systems have shown great potential as a valid tool for underwater navigation and positioning in a GNSS denied environment in different applications ranging from archaeology to subsea metrology (Menna et al., 2019; Nawaf et al., 2018; Drap et al. 2015).Visual SLAM has also been used as positioning technique, providing the 6DoF necessary for mobile mapping systems based on structured light for subsea inspection and monitoring from mining environments to cultural heritage (Bleier et al 2019a,b; Bräuer-Burchardt et al., 2023).

[1] https://oceanservice.noaa.gov/facts/bathymetry.html

a) b)

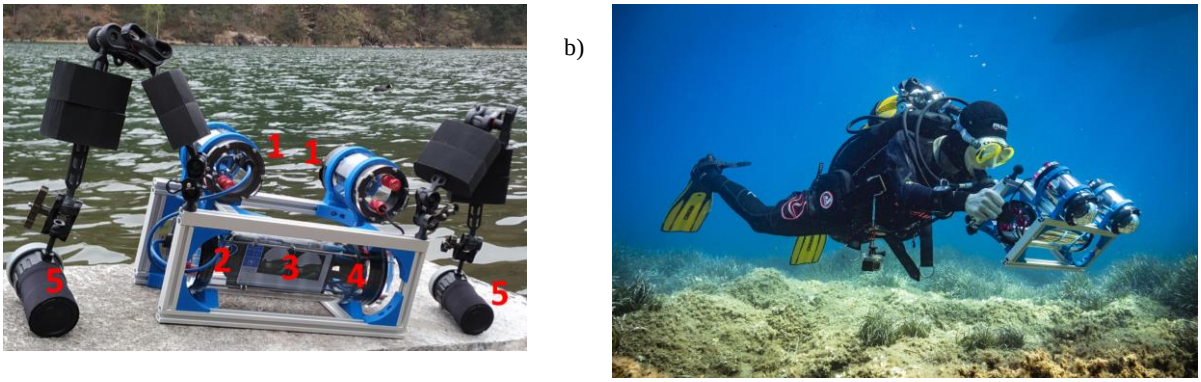


**Figure 1**. FROG prototype above the water on a lake shore (a) and in shallow water during a test dive in the Mediterranean Sea (b) used for vSLAM-based mobile mapping applications. The main hardware components are shown: 1) cameras; 2) computing unit; 3) visualisation unit; 4) battery; 5) underwater lights.

To the best of our knowledge, none of the available solutions based on visual odometry and vSLAM provide, at low-cost, embedded real-time processing, guidance and control of the system, such as automatic image acquisition, distance to the object, maximum allowed speed to avoid motion blur, exposure control. In previous works (Menna et al., 2019; Nawaf et al., 2018; Drap et al. 2015), the images, captured by a stereo-camera, were transferred to a computer for real-time processing to support navigation while performing the photogrammetric survey. These systems need powerful computational resources, often with GPU capabilities and on a remote computer, either wirelessly over the Internet or via a cable.

In this paper we present a novel low-cost portable underwater system called FROG (Figure 1) which is built upon the *Guided Photogrammetry - GuPho* concept presented in Menna et al. (2022). Preliminary results on the algorithmic aspects related to the real-time feedback and image acquisition control of the developed vSLAM-based mobile mapping system were presented in Torresani et al. (2021). In Di Stefano et al. (2021) a first metric comparison with other portable mobile systems was presented in the case of underground heritage documentation. In Menna et al. (2022), a broader overview of the motivations driving the development behind such a system and its lightweight architecture were provided and different surveying scenarios enabled by its modularity were presented.

In this paper, we focus on the underwater version of GuPho, researching into the application aspects of the system in complex and challenging scenarios such as in turbid water (lakes) and in narrow submersed environments.

## 2. FROG: A VSLAM BASED UNDERWATER PHOTOGRAMMETRIC SYSTEM

### 2.1 Real-time guidance for diver-operated and micro ROVs underwater photogrammetric surveys

In many underwater photogrammetry projects, especially in shallow water, the image acquisition is still carried out by divers using open circuit SCUBA or rebreather apparatuses, maybe assisted by a DPV (diver propulsion vehicle). Applications may range from ecological (Nocerino et al., 2020), to archaeology (Nawaf et al., 2021, Drap et al., 2015) and geological studies (Caruana et al., 2022). In these circumstances (such as in shallow water), underwater photogrammetry surveys carried out by divers have several advantages that make them preferred to remotely operated vehicles. Among the several factors, there are costs, ease of implementation, dexterity of humans, better perception of the environment, especially when surveying fragile scenery such as cultural heritage or due to the risk of entanglement of the ROV umbilical, such as for delicate benthonic species like corals, or in underwater caves. Nevertheless, image acquisition by divers is still completely carried out manually, most of the time using regular digital cameras enclosed in waterproof pressure housings. In these cases, the survey is based solely on the training and expertise of the photogrammetrist with no feedback provided in real-time with respect to the image acquisition parameters both from a geometrical (e.g. distance to the object, overlap, speed, motion blur) and radiometric (e.g. exposure) standpoints. The lack of a real-time guidance can lead to undesired loss of image coverage in the photogrammetric survey, an event that in most of the cases is avoided by planning a hyper redundant camera network with larger overlaps and sidelaps. Consequently, the survey is carried out with the diver that needs to take many more photographs than necessary, staying longer underwater, and thus resulting in an unnecessary, far-from-efficient, and more risky survey. Moreover, the environmental conditions underwater are very often unpredictable, requiring slight changes in the camera network due for example to greater than expected turbidity of the water, current or swell. In these conditions, the survey needs to be adjusted on site by the photogrammetrist who needs to be able to reconsider all the parameters such as, for example, distance to the object, swimming speed, interval shooting, shutter speed to avoid motion blur. Moreover, even in ideal conditions, when surveying vast areas that need more than a simple transect (i.e. a photogrammetric block), it is very difficult for a human to localise oneself and navigate in the environment to carry out parallel strips only with visual aids, trying to figure out the correct trajectory to follow.

Moved from the above considerations, we designed a low-cost, lightweight, and portable underwater mobile mapping prototype system called FROG (Figure 1). FROG is engineered on top of a vSLAM stereo-vision system, named GuPho developed by the authors (Torresani et al., 2021) to provide real-time guidance to the underwater surveyor during the image capturing phase, ensuring a more reliable, and effective photogrammetric data acquisition and processing. Given its modular design, FROG can be installed on a micro ROV, and controlled remotely from the support vessel.

### 2.2 Hardware

FROG uses the same hardware of GuPho (Torresani et al., 2021), enclosed in three waterproof housings made of polycarbonate and is depth rated to 100m. A cylindrical housing contains the main electronic components, i.e., the computing unit, the power bank and the visualisation device. Two additional waterproof housings, equipped with dome ports, contain the cameras. The

housings are connected to a rigid frame made of aluminium (Figure 1).

FROG is composed of three modular components, i.e. imaging, computing and visualization. Each of these can be customised or replaced in order to meet specific performance requirements of the application of interest. For example, the imaging module consists of a stereo camera whose configuration such as the baseline, the convergence angle of the lenses, the type of lenses (rectilinear or fisheye), can be easily changed. Depending on whether FROG is used handheld by a diver or installed as skid of a remotely controlled vehicle (ROV), the real-time 3D reconstruction, along with the visual feedback used to help the surveyor, can be displayed either on a mobile phone, placed inside the cylindrical enclosure, or on the remote computer on the support vessel, respectively. For the diver version, two buttons have been added to allow simple controls by hand (e.g., as start and stop) that normally are activated through the touchscreen of the mobile phone for the terrestrial version. Similarly, the camera model and computing unit can be changed for higher resolution applications thanks to the modular software architecture that is based on virtual machine containers.

Regarding the imaging unit, the version presented in this contribution features a pair of 1.3 MP global shutter cameras mounting fisheye lenses with parallel optical axes in normal configuration and baseline of about 250mm. During the experimental study, fisheye lenses have proved high flexibility, especially in challenging scenarios, such as in turbid water in lakes, or in narrow passages like those encountered in underwater canyons and caves (Nocerino and Menna, 2020). This confirms the benefits of fisheye lenses for very close-range environmental surveys as previously shown for terrestrial cases (Perfetti et al., 2017). In these cases, a wide field of view is often the unique solution for an efficient photogrammetric survey. Indeed, due to the required close working distance (from 3m up to less than 50 cm from the surveyed scene in very turbid water), a field of view greater than 100 degrees may be preferrable to minimise the number of strips in the photogrammetric block. These constraints guided the choice of the dome ports against flat ports, despite their higher costs because the field of view of flat ports is limited to about a maximum of 97 degrees due to total internal reflection. Moreover, images provided by a dome port are of superior quality, as images formed behind a flat port suffer from chromatic aberrations and astigmatism. Dome ports allow to minimise the refractive effects of water and keep the optical characteristics of the lens almost unchanged when immersed in water, such as focal length, distortions, field of view.

The computing unit in FROG is a Raspberry Pi4 microcomputer while a mobile phone is used for 3D visualization. The current version of the system uses software synchronisation, although hardware synchronisation is possible. The current synchronisation error is less than 1 ms as measured with a self-developed led stopwatch. Considering a swimming/cruise speed of a diver/ROV ranging from 0.1 to 0.5 m/s, the errors introduced by the asynchronous camera triggering can be considered not significant.

The entire system is powered by a USB power bank of less than 100Wh that ensures more than two hours of continuous operation. Given the low power battery requirements of FROG, it can be carried as hand luggage on a plane. The FROG weight about 5kg above the water and, with lights, is neutrally buoyant once immersed.

### 2.3 Software

FROG uses the OpenVSLAM framework (Sumikura et al., 2019), a project based on ORB-SLAM2 (Mur-Artal and Tardós, 2017). OpenVSLAM provides important features supporting high flexibility and modularity of the system such as: i) different camera models for rectilinear and fisheye lenses, and ii) real-time visualisation on a web-based viewer, that relieves the computing unit from 3D visualisation, thus delegating it to the visualisation unit (i.e. the mobile phone or the remote computer connected via LAN). During surveying operations, the computing unit sends synchronisation triggers to the two cameras and waits for the images to be downloaded. The stereo images are then input into the vSLAM algorithm to estimate in real-time the pose of the system and a sparse three-dimensional reconstruction of the observed scene.

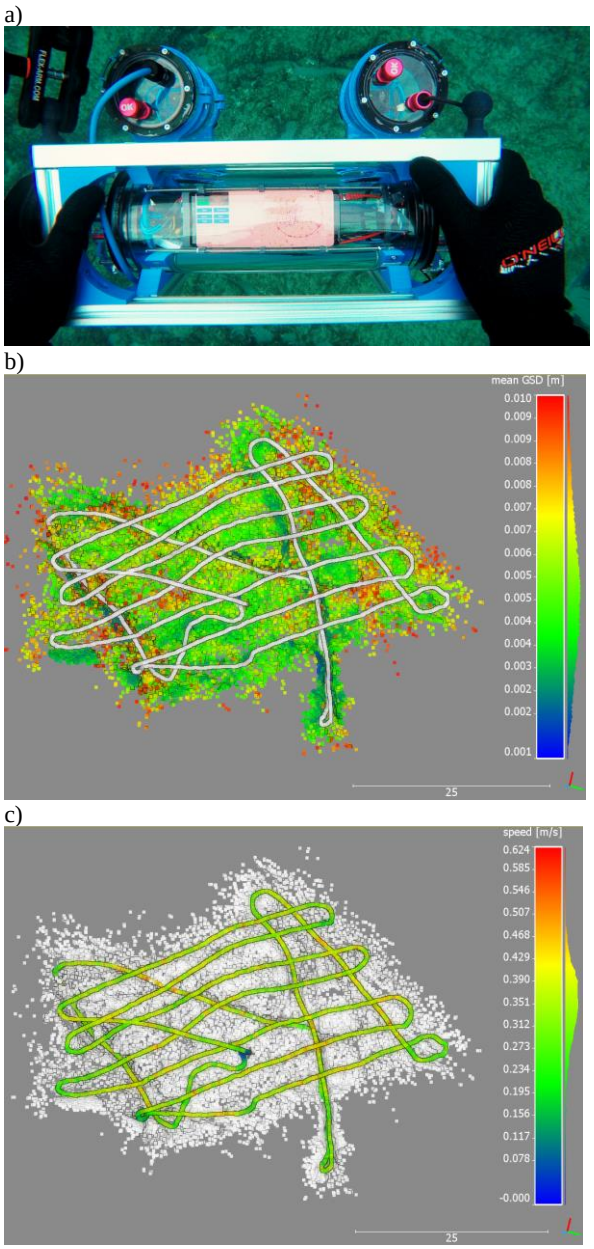


**Figure 2.** FROG in action with its waterproof housing and outputs of the vSLAM process in real-time useful for quality check. Trajectory of the underwater survey for the Argentiera shallow water case study (section 4) as seen by the diver while being recorded in real-time (a) and corresponding sparse point cloud coloured according to the mean GSD (b) or according to the swimming speed (c) for motion blur analysis.

The estimation of camera poses and the 3D sparse reconstruction performed in real-time are used to help the image acquisition, introducing visual feedbacks shown on the display of the

visualisation unit (Figure 2). As reported in Menna et al. (2022), different features are implemented to support the image acquisition in real-time, for example: (i) camera-to-object distance guidance to guarantee the planned ground sample distance (GSD), (ii) an efficient and automatic image acquisition to meet the planned overlap, (iii) current speed warnings to avoid motion blur, (iv) automatic image exposure computed only on the 3D surveyed scene and not on the entire image.
With these real-time feedbacks, the surveyor can more easily follow regular paths underwater with the graphical support of the trajectory and the sparse point clouds of the scenery. Additionally, traffic light style real-time warnings are provided to indicate whether the surveying parameters are being met (e.g. GSD, motion blur).
The real-time feedback is beneficial not only for a greater reliability of the survey, with minimised risks of gaps in the image acquisition, but also for improving its efficiency and safety as the maximum allowed speed, which does not introduce systematic motion blur effects, is checked throughout the image acquisition, helping the surveyor to keep it as constant as possible.
Once the survey is completed, the vSLAM outputs, namely the stored images, the camera exterior orientations, the image observations and the sparse point cloud of tie points can be utilised for further photogrammetric processing, such as dense cloud densification and meshing. In the experiments reported in this paper, Agisoft Metashape (https://www.agisoft.com/) via python API scripting was used, although we have successfully tested OpenMVS (https://github.com/cdcseacave/openMVS).
Despite different processing strategies can be envisaged, such as computing the dense point cloud directly from the exterior orientation parameters estimated from the vSLAM, in Menna et al. (2022) we showed that performing the bundle adjustment (BA) can provide significant accuracy improvements with minimum costs in terms of computational time. Moreover, even when re-orienting the entire set of collected images, exploiting the known calibrated baseline constraints, and the approximate image orientations from the vSLAM, provided a time saving of a factor at least 5.

## 3. FROG CALIBRATION

### 3.1 Optical alignment and focusing

A spherical dome port has several benefits in underwater photogrammetry (Menna et al., 2016). When well centred on the entrance pupil (EP) of the objective lens, the refractive effects of water are practically negligible. Nevertheless, the dome acts as a negative lens that creates a virtual image in front of the dome at about 3 to 4 times its radius from the EP of the lens. Also, the lens, centred within the dome, must be able to focus at such short distance. As main implication, especially for larger sensors (e.g., full frame), the limited depth of field available at a focusing distance of only few centimetres does not allow the system to be used both above and below the water with the same focus settings, and consequently, calibration.
However, FROG uses ½ inches sensors that in combination with fisheye lenses provide a very large DoF, making the images always acceptably in focus both under and above the water.
The spherical dome surface is centred with respect to the EP position (for paraxial rays) of the two cameras using the optical alignment procedure described in Menna et al. (2016). In this procedure, the centre of the dome is assumed to be lying on the optical axis of the lens but not coincident with entrance pupil EP. This assumption is most of the time true within 1-2 mm due to the easier centring of the lens barrels with the dome port circular flange. The virtual image of the EP, as seen along the optical axis is collimated by a camera with a macro-objective lens serving as optical collimator; then, using a linear stage, the optical collimator is shifted along the optical axis until the surface of the dome, at its apex, is in focus. The amount of shift is measured with a calliper and checked against the radius of the spherical dome. If the offset does not coincide with the radius of the dome the camera and its lens are shifted accordingly to minimize the difference. In alternative the shift can be determined analytically and thus adjusted, after computing the decentring offsets as proposed in Rofallski et al. (2022). The procedure is iterative but usually only requires a single underwater calibration of the system.
As described in Menna et al. (2020, 2016) and She et al. (2022), a non-centred dome port introduces a change of magnification when the camera is immersed in water. Conversely, an image captured from a camera with the lens properly centred on the dome does not show magnification changes. This property is often exploited in underwater photography for showing seamless split views of both underwater and coastal landscapes. For photogrammetrists, capturing a partially immersed checkerboard (Figure 3) provides a straightforward practical method to understand whether the lens is properly centred. Figure 4 shows a split view of a partially immersed checkerboard for a frame captured by one of the cameras of FROG prototype showing no visible refraction effects, except for the very borders of the image format due to apparent movement of the EP for peripheral rays (field of view larger than 120 degrees). As it is shown in section 4.1, this design makes it possible to use the device both above and under the water with the same calibration when relaxed accuracy requirements exist (e.g. 1:100 relative accuracy).

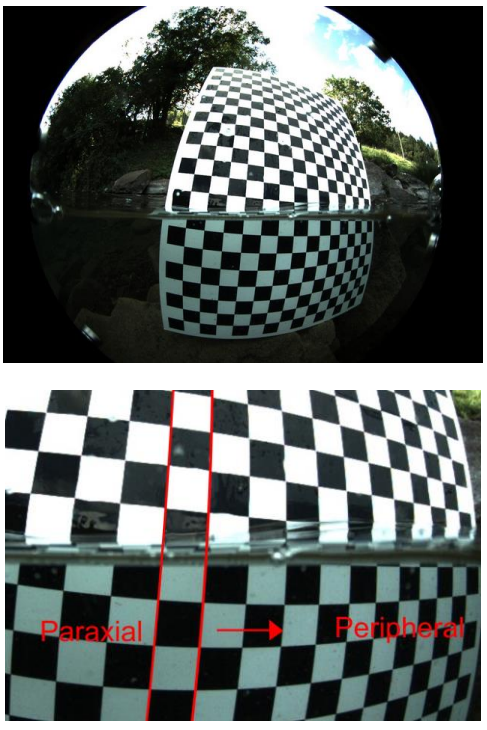


**Figure 3**. Split views of a semisubmersed fisheye image with both the dome and checkerboard partially immersed in water to show the absence of significant refractive effects (for paraxial rays) and thus a proper centering procedure of the dome.

### 3.2 Stereo system self-calibration

For the vSLAM algorithm to properly work guiding the diver throughout the photogrammetric survey, the stereo configuration must be calibrated and the calibration parameters pre-loaded into

FROG. The parameters required include the interior orientation parameters of each camera and the asymmetrical relative orientation (baseline and attitude) of the right camera with respect to the left, considered the 'master' camera.
The most rigorous solution would require FROG to be calibrated in the prevailing working conditions, i.e., at the planned operative depth, with the same visibility, illumination and current. However, in the present version, the calibration parameters cannot be updated on-site, but they need to be preloaded and, once estimated and entered into the system, they are kept fixed during the vSLAM process. Although introducing some inaccuracies into the process, the most practical solution entails to calibrate FROG 'in air' and use the calibration result in the vSLAM process to guide the diver underwater in performing the surveying. For increasing the reliability of the process, it is advisable to acquire additional calibration dataset also underwater, as the parameters estimated a posteriori can then be used to optimise the orientation step and generate the dense models. FROG is calibrated using a portable testfield, or targets and scale bars prior to entering the water and, for extra safety, at the operative depth of the survey (Figure 4). The calibration images are collected following standard self-calibration protocols, which should involve a highly redundant camera network comprising rolled and convergent images of the three-dimensional testfield. The stereo system calibration parameters, which include interior as well exterior relative orientation parameters, are estimated combining the observations from the two cameras in a unique bundle block adjustment with baseline constraints. In our formulation, the baseline distance between the two cameras is an unknown and may vary from one era to another within a set tolerance. The mean relative orientation parameters are estimated from the bundle following the procedure described in Nocerino and Menna (2020).

**Figure 4**. Example of temporary test field in an alpine lake using 1 meter long scalebars, color and resolution checkers, coded targets used for geometric and radiometric calibration of the sytem.

## 4. FROG IN ACTION

To show the potentiality of FROG as portable underwater mobile mapping system for 3D surveying, two experiments are hereafter reported:

1) a high-resolution (1 cm) bathymetric map and orthophoto of an area of about 1600 $m^2$ in very shallow water in Sardinia (Italy);
2) monitoring an underwater artwork for preventive restoration in the Garda lake (Italy).

The first area, the Argentiera mining park near Sassari, Italy is part of the Sardinian UNESCO sites (https://www.castelmeteo.it/argentiera/libro/libro.htm). The mining area has been used since the roman times and is now an open-air museum. FROG was tested to evaluate its potential in documenting the peculiar geological features of the shallow water submersed parts, rich of narrow passages, overhang environments, small caverns, and long canyons.
The second test site is the "Cristo Silente" (silent Christ) statue in Riva del Garda, near Trento, Italy, standing vertically at a depth of about 16m (base) up to 12 m (head). The statue was made by artist Germano Alberti and placed underwater in 1970. It is a very popular SCUBA dive location, used by several thousand divers every year both for training and leisure. The statue is made of metal and is regularly subject to restoration activities. FROG was used to test its potential in 3D modelling and monitoring of the statue.

### 4.1 Argentiera mining park

**4.1.1 System calibration**: FROG was calibrated both above and under the water to verify, analytically, the centring of the domes as well as showing the effects of using a dry calibration underwater. For the assessment, we used more than 10 scalebars characterised by an average and maximum length of 25 cm and 1 m ca respectively, and an estimated reference length accuracy better than 0.05 mm. Overall, in both calibrations (above and under the water), the root mean square (RMS) of the length measurement error (LME) were below 0.2 mm, corresponding to a relative LME of 1:5000.
Table 1 reports the interior as well as exterior orientation parameters of the two fisheye underwater cameras when calibrated above and under the water, respectively. The slight change in the calibration parameters implies residual refractive effects that prevent using the calibration estimated above the water for underwater surveys when high accuracy is required. Nevertheless, the minor differences enable FROG to work underwater even using the dry calibration, with some compromises on the achievable accuracy. This approximation introduces systematic errors in the real-time reconstruction, yet it proved to be enough for real-time navigation and guidance. To estimate the accuracy degradation when using a dry calibration in a close-range underwater survey, we used the dry calibration parameters in Table 1 in the underwater calibration image datasets. These were kept fixed letting the bundle adjustment (BA) to solve only for camera poses, tie points and coded target 3D coordinates. Overall, the RMS of LME worsened to 5 mm with an average length error of 1 cm on the 1 m long scalebars (relative error 1:100).

**4.1.2 Shallow water bathymetry**: A quadrilateral area of approximately 45x35$m^2$ with depths ranging from 1.5 to 9 meters was surveyed using FROG set with a target GSD of 1 cm that corresponds to a maximum camera to object distance of about 3.5 meters from the object (for paraxial rays). The area was surveyed with seven parallel strips and 5 cross strips optimised to follow underwater narrow passages and canyons present in the area (Figure 2), resulting in about 2500 images in total. With an average speed kept during the survey of 35 cm/s, the entire trajectory (about 530 m long) was completed in 25 minutes.
To evaluate the accuracy of the survey, two one-meter long scalebars were located at two opposite corners of the quadrilateral at about 50 m distance from each other. Moreover, a total of five targets were placed on the seabed, four at the corners plus one at the centre. For these targets, a floatable GNSS receiver and high resolution (2 mm) depth meters, both developed by the authors (Menna et al., 2021) were used to measure the planimetric coordinates and the depths.

| | | Interior orientation [pixel] | | | | | | Asymmetric relative orientation [mm/degrees] | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **f** | **PPAx** | **PPAy** | **k1** | **k2** | **k3** | **Dx** | **Dy** | **Dz** | **ω** | **φ** | **κ** |
| **DRY** | **L** | 385.07 ± 0.02 | 9.02 ± 0.009 | 19.16 ± 0.008 | -8e-03 ± 9e-05 | -2e-03 ± 7.4e-05 | -1e-03 ± 1.9e-05 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **R** | 383.597 ± 0.02 | 16.8444 ± 0.009 | 28.1947 ± 0.008 | -7e-03 ± 9e-05 | -2e-03 ± 7.3e-05 | -1e-03 ± 1.9e-05 | 245.61 ±3.3e-02 | 8.252 ±6.5e-03 | 6.338 ±1.2e-02 | -0.077 ±0.001 | -0.895 ±0.002 | 2.8212 ± 0.0003 |
| **UW** | **L** | 389.99 ±0.05 | 9.02 ±0.02 | 18.83 ±0.02 | -4e-03 ±0.0002 | 3.7e-04 ±2e-04 | -2e-03 ±4.5e-05 | 0 | 0 | 0 | 0 | 0 | 0 |
| | **R** | 389.633 ±0.04 | 18.38 ±0.019 | 29.0418 ±0.017 | -0.004 ±2e-04 | -0.0007 ±2e-04 | -0.0019 ±4.2e-05 | 244.20 ±9.2e-02 | 7.732 ±1.5e-02 | 6.623 ±2.8e-02 | -0.169 ±0.002 | -0.799 ±0.003 | 2.716 ±0.0007 |

Table 1. Interior and asymmetric relative orientation parameters of the two FROG fisheye cameras in their waterproof housing with dome ports calibrated both above and under the water. The observed slight change in the calibration parameters implies residual refractive effects.

While the GNSS receiver coordinates are only used for approximate georeferencing (the device was let float over the target using a finger spool cable for an estimated accuracy of not better than 5 m), the depths, each averaged over a period of more than 1 minute (more than 150 observations) can be considered accurate to about 1-2 cm, also considering the very favourable weather conditions during the survey (overall RMS of water surface elevation of less than 2 cm). The approximate camera poses estimated by FROG in real-time were then input in Agisoft Metashape for a refined BA. The scaling was provided by the relative orientation of the two cameras, kept fixed along with the interior orientations during the adjustment. The RMS of LME on the two scalebars resulted in about 1.7 mm, a highly satisfying result considering the GSD of about 10 mm achieved with only 1.3MP cameras. The quality of the result was evaluated also on the coordinates of the five targets by computing a 3D Helmert transformation without scale factor and weighting the planimetric coordinates according to their expected accuracy. However, stating the low accuracy of the GNSS measurements, only the RMS on the Z coordinates is considered, resulting equal to 1.4 cm. This value shows that, despite the large area and the high resolution of the survey, residual systematic errors were well compensated. Indeed, for elongated strips and large aerial like photogrammetric blocks, systematic errors may affect the overall accuracy of 3D coordinates resulting in deformed blocks, especially in Z (Nocerino et al., 2014).

For the same test area, with the aim to show the accuracy degradation due to the use of a dry camera calibration, we swapped the calibration sets as done in section 4.1.1. The resulting LME on the 1 m scalebars resulted in average larger by 3.9 cm (+4% relative error) while the RMS on the Z coordinates on the 5 targets raised to 68 cm. Comparing the planimetric distances of the two diagonals of the quadrilateral between the two BA solutions, an average systematically positive difference of 1.5 m over 30 m was observed, corresponding to a relative length error of 5%.

The dataset processed with the underwater calibration was then further elaborated to produce classical photogrammetric products such as a DEM, orthophoto and 3D mesh model (Figure 5).

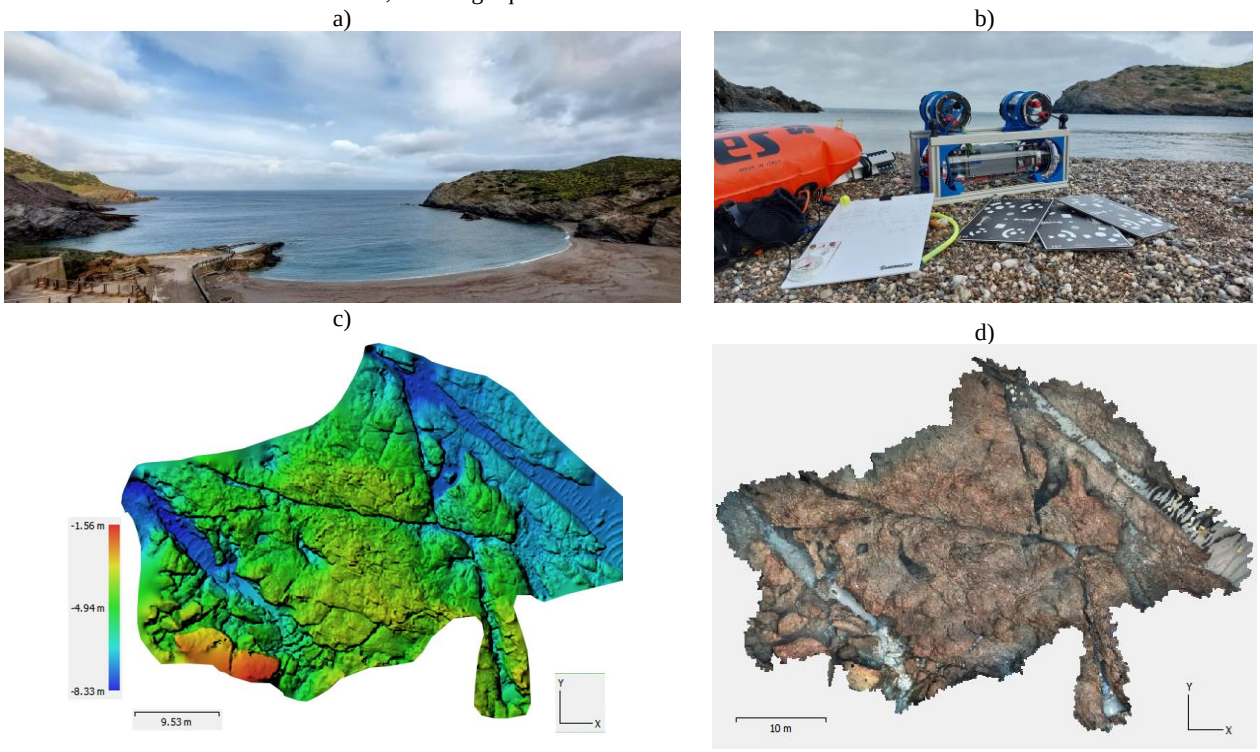


Figure 5. The bay of the Argentiera mining park (a) where FROG (b) was tested for producing a high-resolution bathymetric map and orthophoto of an area of about 1600 m$^2$ in very shallow water. The resulted DEM (c) and orthophotomosaic (d).

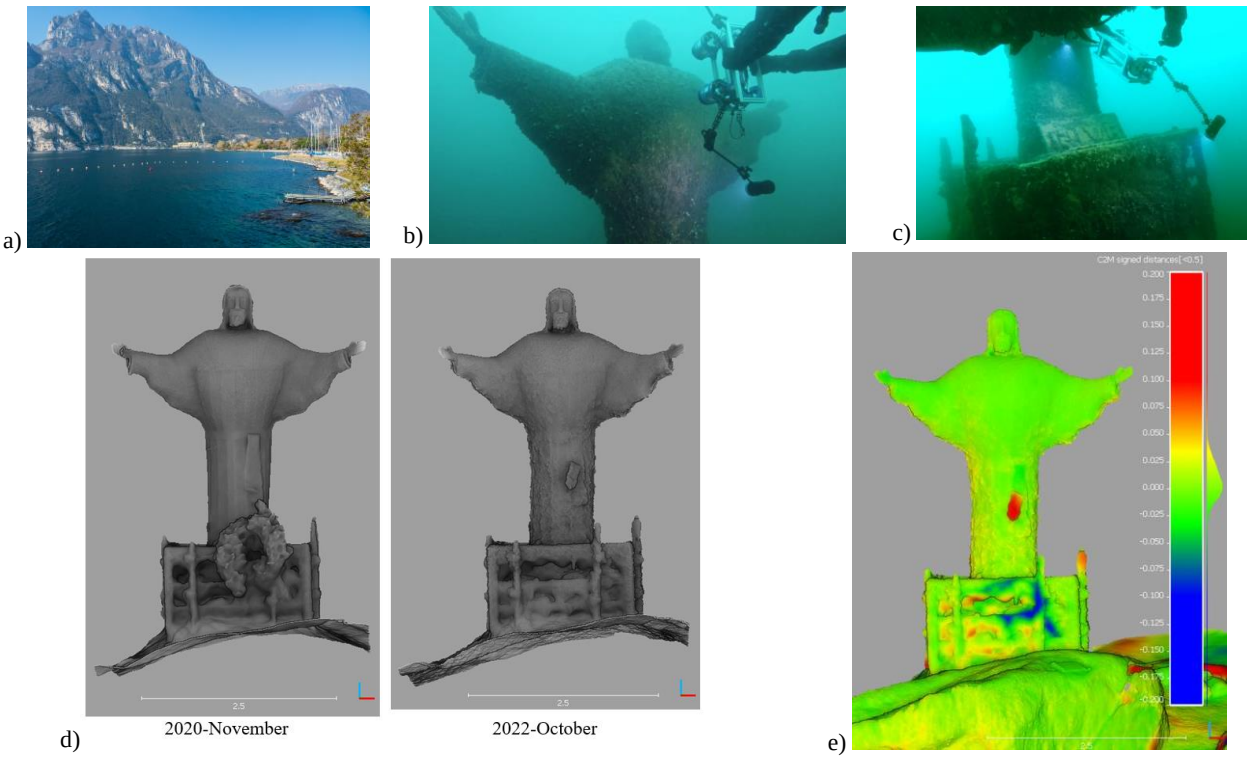


**Figure 6.** The Garda lake (a) where FROG was used for the 3D monitoring of Cristo Silente statue. A diver operating the system at about 16 m depth (b, c). 3D models of the statue previously surveyed in 2020 with standard photogrammetry and in 2022 with FROG (d). Mesh-to-mesh distance showing the damage of the belt of the tunic of Christ in red (e).

### 4.2 Cristo Silente

The survey of the statue was carried out within a test dive in October 2022 aimed also at evaluating the performances of FROG during long operations. The system continuously tracked the bottom of Garda Lake up to a depth of about 32m in a loop more than 300m long for 74 minutes. The Cristo Silente statue was surveyed at closer distance providing a GSD of better than 5 mm and for the statue plus the surrounding promontory, 2700 images were automatically collected by FROG. The two scale bars were repeatedly placed along the trajectory for checking the LME, which resulted in agreement with what was also shown in section 4.1.2 (GSD ca 1 cm, LME less than 5 mm). The post-survey processing was then carried out in Agisoft Metashape where a 3D mesh model was computed. This mesh model was then compared to a former 3D model of Cristo Silente surveyed by underwater photogrammetry by the authors two years before. The comparison highlighted a damage of the belt of the tunic of Christ for a more than 7 cm mesh-to-mesh distance, most probably due to corrosion and weakening of the soldered metal plates, of which the statue is made (Figure 6).

## 5. CONCLUSIONS AND FUTURE WORK

The FROG prototype presented in this paper shows the potential of vSLAM techniques combined with state-of-the-art photogrammetric procedures for high resolution 3D optical mapping of the underwater environment. The system, developed as a handheld underwater portable mapping device is built with low-cost hardware, yet the accuracy results achieved in the two experiments in a lake and at sea are more than encouraging. The RMS of LME on the scale bars was always below the planned GSD, a result often difficult to achieve even in regular terrestrial photogrammetric surveys. Yet, more importantly, without a real-time guidance it would have not been possible to entirely cover the surveyed scene without gaps if not recording a hyper redundant camera network. The combination of vSLAM for real-time guidance, BA techniques, and multi-view stereo reconstruction proved to be a cost effective as well as an efficient solution for high resolution underwater 3D measurements. We expect that this approach will be more and more common in the future of photogrammetry, not only underwater. In the future we will integrate pressure and inertial sensors in FROG to improve the reliability of the system in more challenging conditions such as in very turbid water, over moving seagrass or in presence of light caustics. To this aim we will also explore artificial intelligence solutions. Furthermore, we will consider extending the modularity by controlling an optional higher resolution camera.

## ACKNOWLEDGEMENTS

This work is partly supported by EIT RawMaterials GmbH under Framework Partnership Agreement No. 19018 (AMICOS – Autonomous Monitoring and Control System for Mining Plants) and by the project “Ross Sea Benthic Monitoring Program: new non-destructive and machine-learning approaches for the analysis of benthos patterns” (“RosS-BMP”,_PNRA18_00263).
The authors acknowledge the support of NBFC to Univ of Sassari, funded by the Italian Ministry of University and Research, PNRR, Missione 4 Componente 2, “Dalla Ricerca all’impresa”, Investimento 1.4, Project CN00000033.
The authors are thankful to Fabio Mosna, Sara Messina, Andrea Marconi, Carlo Longin, Andrea Montagner and Emiliano Trentinaglia from Rane Nere Sub Trento for the support during diving operations.

## REFERENCES

Bleier, M., Van der Lucht, J. and Nüchter, A., 2019a. Scout3D–an underwater laser scanning system for mobile mapping. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, pp.13-18.

Bleier, M., Almeida, C., Ferreira, A., Pereira, R., Matias, B., Almeida, J., Pidgeon, J., van der Lucht, J., Schilling, K., Martins, A. and Silva, E., 2019b. 3D Underwater Mine Modelling in the¡ VAMOS! Project. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42, pp.39-44.

Bräuer-Burchardt, C., Munkelt, C., Bleier, M., Heinze, M., Gebhart, I., Kühmstedt, P. and Notni, G., 2023. Underwater 3D Scanning System for Cultural Heritage Documentation. *Remote Sensing*, 15(7), p.1864.

Caruana, J., Wood, J., Nocerino, E., Menna, F., Micallef, A. and Gambin, T., 2022. Reconstruction of the collapse of the 'Azure Window'natural arch via photogrammetry. *Geomorphology*, 408, p.108250.

Chemisky, B., Menna, F., Nocerino, E. and Drap, P., 2021. Underwater survey for oil and gas industry: A review of close range optical methods. *Remote Sensing*, 13(14), p.2789.

Di Stefano, F., Torresani, A., Farella, E.M., Pierdicca, R., Menna, F., Remondino, F., 2021. 3D Surveying of Underground Built Heritage: Opportunities and Challenges of Mobile Technologies. *Sustainability*, Vol.13, 13289

Drap, P., Merad, D., Hijazi, B., Gaoua, L., Nawaf, M.M., Saccone, M., Chemisky, B., Seinturier, J., Sourisseau, J.C., Gambin, T. and Castro, F., 2015. Underwater photogrammetry and object modeling: a case study of Xlendi Wreck in Malta. *Sensors*, 15(12), pp.30351-30384.

Jiang, M., Song, S., Li, Y., Jin, W., Liu, J. and Feng, X., 2019. A survey of underwater acoustic SLAM system. Proc. 12th *ICIRA*, Part II 12, pp. 159-170.

Jørgensen, M.J., 2016. Enhanced Subsea Acoustically Aided Inertial Navigation. PhD thesis, Technical University of Denmark, ISSN: 0909-3192, 164 pages.

Menna, F., Torresani, A., Battisti, R., Nocerino, E. and Remondino, F., 2022. A modular and low-cost portable vSLAM system for real-time 3D mapping: from indoor and outdoor spaces to underwater Environments. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 48, pp.153-162.

Menna, F., Nocerino, E., Chemisky, B., Remondino, F. and Drap, P., 2021. Accurate scaling and levelling in underwater photogrammetry with a pressure sensor. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, pp.667-672.

Menna, F., Nocerino, E., Ural, S. and Gruen, A., 2020. Mitigating image residuals systematic patterns in underwater photogrammetry. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, pp.977-984.

Menna, F., Agrafiotis, P. and 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., Fassi, F. and Remondino, F., 2016. Geometric and optic characterization of a hemispherical dome port for underwater photogrammetry. *Sensors*, 16(1), p.48.

Mur-Artal, R. and Tardós, J.D., 2017. Orb-slam2: An open-source slam system for monocular, stereo, and RGB-D cameras. *IEEE transactions on robotics*, 33(5), pp.1255-1262.

Nawaf, M., Drap, P., Ben-Ellefi, M., Nocerino, E., Chemisky, B., Chassaing, T., Colpani, A., Noumossie, V., Hyttinen, K., Wood, J. and Gambin, T., 2021. Using virtual or augmented reality for the time-based study of complex underwater archaeological excavations. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 8(M-1-2021), pp.117-124.

Nawaf, M.M., Merad, D., Royer, J.P., Boï, J.M., Saccone, M., Ben Ellefi, M. and Drap, P., 2018. Fast visual odometry for a low-cost underwater embedded stereo system. Sensors, 18(7), p.2313. Nocerino, E. and Menna, F., 2020. Photogrammetry: linking the world across the water surface. *Journal of Marine Science and Engineering*, 8(2), p.128.

Nocerino, E., Menna, F., Gruen, A., Troyer, M., Capra, A., Castagnetti, C., Rossi, P., Brooks, A.J., Schmitt, R.J. and Holbrook, S.J., 2020. Coral reef monitoring by scuba divers using underwater photogrammetry and geodetic surveying. *Remote Sensing*, 12(18), p.3036.

Nocerino, E., Menna, F. and Remondino, F., 2014. Accuracy of typical photogrammetric networks in cultural heritage 3D modeling projects. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 45, pp. 465-472.

Perfetti, L., Polari, C. and Fassi, F., 2017. Fisheye photogrammetry: tests and methodologies for the survey of narrow spaces. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42(W3), pp.573-580.

Rofallski, R., Menna, F., Nocerino, E. and Luhmann, T., 2022. An efficient solution to ray tracing problems for hemispherical refractive interfaces. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2, pp.333-342.

Schwarz, K.P. and El-Sheimy, N., 2004. Mobile mapping systems–state of the art and future trends. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 35(Part B), p.10.

She, M., Nakath, D., Song, Y. and Köser, K., 2022. Refractive geometry for underwater domes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 183, pp.525-540.

Smith Menandro, P. and Cardoso Bastos, A., 2020. Seabed mapping: A brief history from meaningful words. *Geosciences*, 10(7), p.273.

Torresani, A., Menna, F., Battisti, R., Remondino, F., 2021. A V-SLAM Guided and Portable System for Photogrammetric Applications. *Remote Sensing*, Vol.13(12), 2351.

Wölfl, A.C., Snaith, H., Amirebrahimi, S., Devey, C.W., Dorschel, B., Ferrini, V., Huvenne, V.A., Jakobsson, M., Jencks, J., Johnston, G. and Lamarche, G., 2019. Seafloor mapping–the challenge of a truly global ocean bathymetry. *Frontiers in Marine Science*, p.283.