

A PORTABLE OPTO-ACOUSTIC SURVEY SOLUTION FOR MAPPING OF UNDERWATER TARGETS

B. Chemisky^{a*}, E. Nocerino^a, F. Menna^b, M.M. Nawaf^a, P. Drap^a

^a Aix Marseille University, CNRS, ENSAM, Université De Toulon, LIS UMR 7020, 13397 Marseille, France - (erica.nocerino, pierre.drap, motasem.nawaf)@univ-amu.fr, bertrand.chemisky@etu.univ-amu.fr

^b 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy – fmenna@fbk.eu

Commission II, WG II/9

KEY WORDS: Underwater photogrammetry, Remote Operated Vehicle, forward looking sonar, stereo-camera, SLAM, calibration.

ABSTRACT:

During underwater investigations, whatever the mission objective and the type of vehicle, obstacles detection and avoidance are essential tasks. They can either represent a target of interest that is the object of the mission or, on the contrary, represent obstacles that can hinder or affect the navigation of the vehicle. The underwater optical cameras that are usually fitted to underwater vehicles only offer a narrow field of view. The absorption of electromagnetic waves in the first few meters and the diffusion of light by the particles limit the use of these sensors to only a few meters range. The use of acoustic sensors, such as the forward looking sonar (FLS), is then necessary to enlarge the volume in which a target can be detected during the progression of the vehicle. Traditionally, sonars featured mechanical rotating parts, but lately bidirectional forward looking sonar, which directly produces a 2D image of the area, are becoming more and more common. Although these sonars can operate at frequency higher than 1MHz, their spatial resolution remains much lower if compared to current optical sensors and can be insufficient to identify and characterize a target. The combination of these two sensors in an operational scenario is essential to take advantage of each technology. In this paper we describe a low cost, multi-sensor, underwater survey solution for the identification, tracking, and 3D mapping of targets. After a description of the architecture of the opto-acoustics data acquisition and processing platform, we will focus on the calibration of the rigid transformation between the two sensors.

1. INTRODUCTION

In underwater surveying we continuously deal with the antagonism between the range and resolution of the sensors used. In an underwater environment, the performance of optical sensors is greatly affected by the level of turbidity and their measurement capacity are therefore highly dependent on a combination of spatial, temporal and climatic factors that require a high degree of flexibility. Acoustic sensors, on the other hand, depending on the frequency used, can reach ranges of several kilometres. The lower the frequency, the greater the range. At the same time the greater the range, the lower the achievable resolution.

Search, mapping identification and reconnaissance of underwater targets are often carried out by multiple vectors and at different times. This means that the position of the target, in a continuously changing environment without stable references, can be highly uncertain and requires an extensive search when the target needs to be revisited (Mari et al., 2017). As optical sensors offer only a short range and coverage, the use of acoustic sensors during the search phase may be the only advisable solution, all the more so when visibility is poor. We thus propose the integration of a high-frequency forward looking sonar coupled with a stereo optical sensor to optimize target identification and mapping. The targets, once identified by the "long range" acoustic sensor, can be tracked to help the pilot of the remotely operated vehicle to get close to the object up to an acceptable distance for a high-resolution optical survey. This is a common procedure in underwater reconnaissance and mapping where a coarse-to-fine resolution approach is adopted car-

rying out several subsequent surveys (e.g. side scan sonar, then multibeam, then punctual inspection by an ROV). In this paper we propose a low-cost, lightweight and compact solution, based on off-the-shelf components, for easy deployment from a light, semi-rigid boat and, in the longer term, a USV (Unmanned Surface Vehicle). The remotely operated platform described in this article is intended to be a platform for experimentation that goes beyond the operational scenario retained in this study. The software architecture is broken down into modules, multi-environments (C++, Java, Python, etc.), open-source libraries such as openCV¹ and commercial software such as Agisoft Metashape². This choice is based on the wish for great modularity in order to meet different needs of educational as well as scientific projects.

A prior calibration of the relative orientation between the optical and the acoustic sensors allows the expression of the position of a target detected in the sonar image, in the photogrammetric reference frame and vice versa. The paper is structured as follows. First, an overview of related works on optical and acoustic data sensor calibration and fusion is presented. Then, the developed platform architecture from a hardware and software point of view is described. We will then detail the contribution of photogrammetric processing in our calibration approach for a method applicable in operational conditions and provide some experimental results. The paper concludes with a discussion regarding the operational deployment and future works on the system architecture.

¹ <https://opencv.org/>

² <https://www.agisoft.com/>

* Corresponding author

2. RELATED WORKS

ROVs used in inspection operations can range from small light vehicles called Observation Class ROVs, often limited to visual inspections, to Work Class ROVs, which can carry a multitude of larger and heavier sensors (IMCA, International Marine Contractors Association, 2014). Small ROVs have the great advantage of being easily transportable onboard cars and can be deployed from small boats or the shore even by a single person in the case of micro ROVs. Their main drawback is the limited range and stability when equipped with several additional sensors. For more information about inspection ROVs, the reader may want to refer to (Capocci et al., 2017, Ledezma et al., 2015). (Poore et al., 2016, Aristizábal et al., 2016, Aguirre-Castro et al., 2019) present the development of light remotely operated underwater vehicles, dedicated to education and research. Their work focuses on the development of the vehicle itself and the remote control architecture, which is made open-source by (Aristizábal et al., 2016). (Clark et al., 2008) describes the integration of multi-sensors on a micro ROV for mapping and localization within maltese cistern systems. The very small VideoRay Pro III ROV³, equipped with a Tritech SeaSprite scanning sonar, successfully completed the surveys by combining video recordings and fixed station sonar scans. An interesting review on micro ROVs, their architecture and battery systems is presented by (Capocci et al., 2017). Our approach is more dedicated to the extension of carrying capacities of multimodal sensors on a flexible vehicle, by extending its power supply and data transmission capabilities (see section 4). Some related works have been already done on a BlueROV2 to integrate an Oculus multibeam sonar as described in the presentation of Stevens Institute of technology about autonomous Subsea Pipeline Inspection project⁴.

Another study described in (Tang et al., 2020) use an Oculus M750d sonar on a portable ROV to extract altitude information of the object from their acoustic shadow and echo detected in a single sonar image.

These studies are limited to the integration of only one additional sensor, mainly because of a datalink limitation. From the perspective of an open platform for research and education needs we wanted to access as many sensors as possible. The study presented here shows the simultaneous integration of a stereo camera sensor with an embedded Inertial Motion Unit (IMU) and an obstacle avoidance sonar. For more than a decade, the mechanical scanning sonar commonly used on underwater vehicles for obstacle detection or avoidance has been progressively replaced by new high-frequency generation sonars. These new forward looking sonars are based on a transducer array and on-board signal processing modules allowing the production of 2D images at each recurrence.

A summary of the fusion methods between optical and acoustic sensors is described in (Ferreira et al., 2016). Data fusion is achieved by the geometric relationship between the two sensors allowing the data to be expressed in a common reference frame. This approach is called mapping-oriented method, according to the classification proposed in (Nicosevici et al., 2004). Some methods are based on a "direct" computation of the rigid transformation matrix by aligning the respective 3D point clouds of each sensor of a same scene (Lagudi et al., 2016, Bruno et al., 2015). For example, (Drap et al., 2014) show the fusion of opto-acoustic data between a 3D scanner and 3 synchronised

cameras. Referencing of the data in a common coordinate system is then carried out in post-processing by exploiting the point clouds obtained by each type of sensor. A big difference exists between 3D and 2D forward looking sonars. In the latter case, in a conceptually similar way to a photograph, one dimension is lost (see section 3.2.1). This is due to the loss of elevation angle information due to the projection geometry (Aykin and Negahdaripour, 2013). Few works have been carried out for data fusion between a 2D forward looking sonar and an optical sensor. In (Negahdaripour et al., 2007, Negahdaripour et al., 2009) the authors propose a new approach of opto-acoustic stereo reconstruction using the epipolar geometry of a stereo system composed by an optical camera and a 2D sonar. A more recent example of opto-acoustic fusion is presented in (Rahman et al., 2018). The system operated by divers is equipped with a stereo camera, an IMU and a mechanical scanning sonar to run visual inertial odometry (VIO) of underwater structures. They propose a new approach to fuse range data from sonar into the traditional VIO framework. Their method is based on the selection of a visual patch around each sonar point, and introduce extra constraints in the pose graph using the distance of the sonar point to the patch. In (Liu et al., 2020), they develop a scale-adaptive matching algorithm for underwater acoustic and optical images based on a Gaussian scale-space and correlation filter. As our objective is to propose a multi-sensor platform for the combination of acoustic and optical data, the determination of the relative orientation between these sensors is essential. In (Negahdaripour et al., 2009) the relative positions of the camera and sonar were estimated through an optimization algorithm that minimizes the distances between 3D reconstructions of optical and acoustic matching projections. A planar grid is used as a target with prominent opto-acoustic features. The planar grid constraint with more than five points allow to compute the relative pose parameters in solving a nonlinear optimization problem based on a suitable error measure. As detailed in the section 5, our calibration approach of the opto-acoustic system is direct and sequential. The interior orientation parameters of the two sensors are calibrated separately. A photogrammetric survey is carried out obtaining synchronised acoustic and image data. 3D coordinates of targets recognisable in both optical and acoustic systems are estimated within the photogrammetric survey in a global reference frame. Their correspondences in the sonar image and in the camera simultaneous position allow to optimally estimate the relative orientation between the two systems.

3. VEHICLE AND SENSORS DESCRIPTION

Within the framework of our study, we aimed to integrate a multitude of sensors on a lightweight and portable vehicle that could be easily handled by one person during the launching and recovery phases. With the proposed solution mainly aimed to research and educational projects that require easy access to the control functionalities with open-source software, flexibility, expandability, and a moderate cost, we opted for the BlueROV2 vehicle from the company BlueRobotics⁵. The developed system is based on this versatile portable ROV, carrier which we equipped with a stereo camera and an acoustic sensor (looking forward sonar).

3.1 ZED2 passive stereo-camera system

The decision to use an optical stereo system is evidently due to the possibility of directly scaling the survey, a crucial prerequisite when mapping unknown underwater environments.

³ <https://www.videoray.com/>

⁴ www.stevens.edu/sites/stevens.edu/files/files/MS/BlueROV_slides.pdf

⁵ <https://bluerobotics.com/>

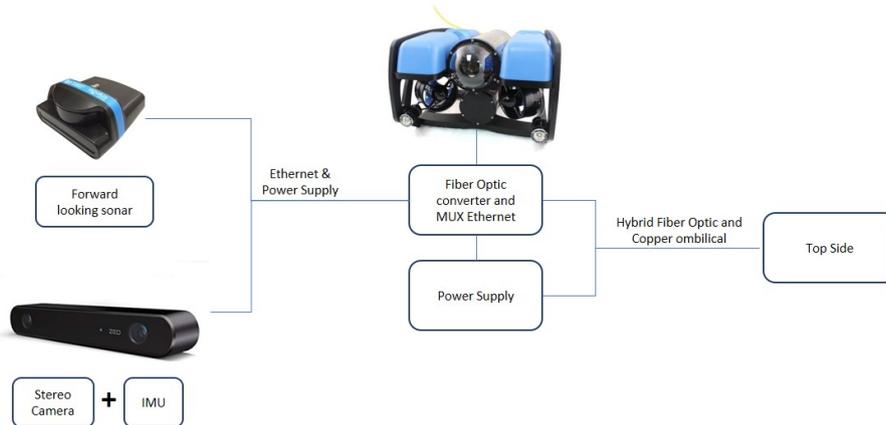


Figure 1. BlueRov2 Architecture

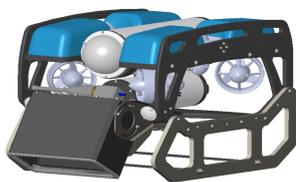


Figure 2. ZED2 waterproof housing installed on the BlueROV2

The optical sensor chosen is the StereoLab's ZED2 stereo camera⁶. This off-the-shelf device combines two synchronised high resolution cameras and an Inertial Motion Unit (IMU). The provided Software Development Kit (SDK) integrates a Simultaneous Localization and Mapping (SLAM) module that used in our study. Sensor specifications are provided in Table 1.

Sensor type	1/3" 4MP CMOS		
Sensor size	2688 x 1520 (pixel)		
Image Resolution	2208 x 1242 (pixel)	@15fps	Cropping mode
	1920 x 1080 (pixel)	@15/ 30fps	Cropping mode
	1280 x 720 (pixel)	@15/ 30/ 60fps	2x2 mode binning
	672 x 376 (pixel)	@15/ 30/ 60/ 100fps	4x4 mode binning
Shutter	Electronic synchronized rolling shutter		
Sensitivity	1900 mV/Lux-sec		
Nominal baseline	120 mm (4.7")		
Nominal focal Length	2.12mm (0.08")		
Aperture	f/1.8		

Table 1. ZED2 specifications

The ZED2 sensor is not designed to be immersed and its USB3 communication protocol greatly limits its operating distance (max 15m with an extender). We have therefore designed a parallelepipedal housing waterproof up to 60 meters equipped with a flat port (Figure 2). Although hemispherical portholes would have been preferable (Nocerino et al., 2016), the flat porthole was chosen for its design simplicity and lower cost. Due to bandwidth limitations on the Ethernet link between the vehicle and the surface control unit, we had to degrade the resolution to

⁶ <https://www.stereolabs.com/zed-2/>

672 x 376 pixels. At this resolution, a frame rate of 5 fps can be ensured for a proper SLAM process.

3.2 Oculus M750D forward looking sonar

We employ the BluePrintSubsea Oculus M750D forward looking sonar⁷. The dual-frequency sensor (750kHz/1.2MHz) is composed of 512 beams that allow a horizontal aperture of 130deg/70deg with 1deg/0.6deg angular resolution respectively in Low and High frequency. Technical specifications are provided in table 2.

	Low Frequency	High Frequency
Range max (m)	120	40
Range min (m)	0.1	0.1
Range resolution* (mm)	4	2.5
Horizontal aperture (°)	130	70
Vertical aperture (°)	20	12
Number of Beams (max)	512	512
Angular Resolution (°)	1	0.6
Beam Separation (°)	0.25	0.16

Table 2. Oculus Sonar M750d specification (* indicates parameter is dependent on range)

3.2.1 Forward looking sonar principle. The principle of high frequency 2D forward looking sonar is based on the emission of an acoustic signal in a beam as wide as the maximum angular aperture allowed by the sonar. The signal is reflected by the bottom and objects into this emission beam. The backscattered signal is then sampled by a transducer array and after a signal processing phase known as "beam-forming", forms a number n of beams containing backscatter intensity along a path of length r . n is a parameter depending on the sonar design and more specifically on the number of transducers and r is a parameter generally adjustable by the user, which determines the reception time window. Thus, the mapped volume is divided into a number n of beams whose angular resolution varies according to the frequency and size of the antenna formed during beam-forming. For each beam, the amplitude of the received

⁷ <https://www.blueprintsubsea.com/pages/product.php?PN=BP01032/>

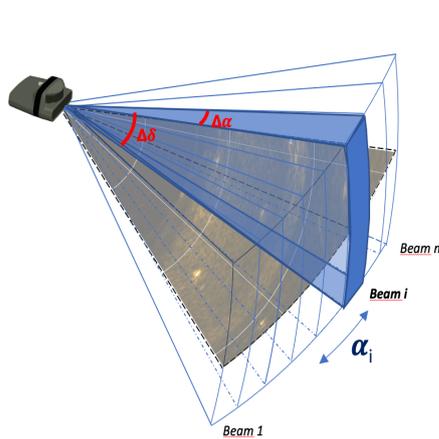


Figure 3. Forward Looking Sonar: beam forming and 2D image projection ($\Delta\alpha$ and $\Delta\delta$ are receptive azimuth and elevation resolution and α_i is the azimuth angle of beam i)

signal as a function of propagation time is recorded. The 2D image is produced from the amplitude of the time samples of the signal and the azimuth angle α of each beam with respect to the sonar axis. For the same distance from the sonar it is therefore not possible to discriminate between two targets that have a difference in azimuth or elevation less than the resolution of each beam. While the azimuth resolution is generally small, less than 1° , the elevation resolution often exceeds 10° due to the small size of the antenna in its vertical axis. When represented in the sonar image, a point P loses its elevation information and thus its vertical dimension with respect to the sonar frame. A 3D point P_S can be expressed in rectangular or spherical coordinates in the sonar reference frame as shown in equation 1 and Figure 4.

$$P_s = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_s = \rho \begin{bmatrix} \cos \delta \sin \alpha \\ \cos \delta \cos \alpha \\ \sin \delta \end{bmatrix} \quad (1)$$

where $X, Y, Z = 3D$ rectangular coordinates
 $\rho =$ range
 $\delta =$ elevation angle
 $\alpha =$ azimuth angle

4. OPTO-ACOUSTIC SURVEY SOLUTION ARCHITECTURE

4.1 Vehicle and sensors description and global architecture

In its standard configuration, the BlueROV2 is energy self-sufficient, powered by an on-board battery. Data communication to the surface is limited due to an Ethernet VDSL protocol and a four twisted pairs tether, that do not allow the transmission of video streams from additional cameras and sonars. We have reviewed the architecture with a power source from the surface (400VDC) and data communication by high speed Ethernet via an optical fibre. Thus, all the data produced by the additional sensors are transmitted over this Ethernet link. The ROV is controlled by the BlueRobotics software module which communicates with a Raspberry Pi3 board embedded in

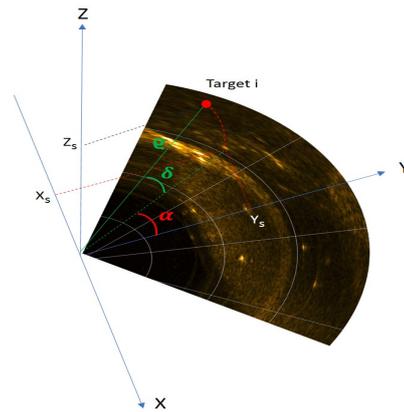


Figure 4. Forward Looking Sonar: target rectangular coordinates (X_S, Y_S, Z_S) and polar coordinates (ρ, α, δ)

the vehicle. In parallel, the Zed2 stereo camera and the Oculus M750D forward looking sonar are controlled through their respective SDK. The control and data acquisition of the ZED2 sensor are supervised by a software module developed in Python 3.7. Regarding the sonar sensor data, a TCP/IP server has been added into its SDK in order to export the data flow to a Python 3.7 module. The data message is detailed in Figure 5.

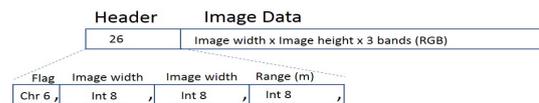


Figure 5. Structure of the message transmitted by the Oculus SDK to Python module

Note that the two sensors are not synchronised at a hardware level, which means that the image capture by the ZED2 sensor is not synchronized with the emission-reception of the Oculus sonar. The delay, which can be variable between the acquisition of these two sensors is therefore unknown. Only a time stamp on the data reception event, at the level of the respective SDKs is provided to ensure a temporal referencing. At this stage of the study, we will overcome this temporal variable by collecting data at successive fixed stations of the vehicle in order to ensure the calibration of the relative orientation between camera and sonar.

4.2 Management module and graphical user interface

The management module developed in Python is the main thread launched by the operator. Based on the Qt package⁸, the user interface is composed of a table to read and send parameters to the sensors and processing modules. A 3D viewer based on the Open3d package⁹ is also used to display point clouds and camera poses. The calibration file for the stereo camera is set at this stage. The underwater intrinsic calibration of each camera (left and right) and the external orientation of the stereo sensor are compute in a previous phase.

4.3 Data acquisition and recording

The data stream generated by the ZED2 stereo camera includes both left and right images and all associated information (IMU

⁸ <https://pypi.org/project/PyQt5/>

⁹ <http://www.open3d.org/>

data, temperature, timestamp) at a frame rate defined in the initialization step. The images are stored in jpeg format with the associated information saved as EXIF metadata.

4.4 SLAM module

Simultaneous Localization And Mapping (SLAM) is an algorithmic approach developed in robotics that aims to simultaneously estimate the position of the vector in its environment and produce the map in real time (Davison, 2003).

Existing SLAM methods differ in terms of number and type of integrated sensors and computational and estimation methods. The sensor development kit (SDK) of the ZED2 offers a SLAM solution that integrates data from its stereo-camera and the inertial measurements provided by the integrated IMU sensor.

In our current development, we adopt the ZED2 proprietary SLAM module, although the whole system has been implemented to easily integrate alternative solutions, including open-source ones, such as ORBSLAM2¹⁰ (Mur-Artal et al., 2015). The SLAM solution computed on-site in real time is stored and made available as initial approximation to run an off-line full photogrammetric process. It also forms the basis of the opto-acoustic calibration procedure developed.

4.5 Incremental SfM

In comparison to SLAM, the SfM approach comes from the field of computer vision (Saputra et al., 2018) and is traditionally performed off-line.

In our case, following the approach presents in (Nawaf et al., 2018), we implemented a sliding window Bundle Adjustment process based on Agisoft Metashape¹¹ photogrammetry software. We exploit the Metashape python API for executing in parallel with only slight delay the two tasks of image acquisition and sequential photogrammetric processing. The system is designed for the integration of more advanced display methods such as the one proposed in (Nocerino et al., 2020).

5. OPTO-ACOUSTIC CALIBRATION PROCEDURE

The calibration procedure we propose requires the identification of common features between the optical and acoustic sensors and consists of the following steps.

5.1 Deployment of opto-acoustic targets

It is rather unusual to find natural features easily recognisable in optical and acoustic images at the same time. Therefore, the calibration procedure includes a deployment phase of targets specifically designed to be well measurable by both the optical and acoustic systems. Five is the number of targets envisaged, thus ensuring a minimum of redundancy even when estimating the relative orientation between the two sensors from a single ROV position (see section 5.4).

5.2 ZED2 underwater calibration

The ZED2 stereo camera is provided with a factory calibration file containing the interior orientation parameters for each camera and the rigid transformation between them (relative orientation). The parameters provided refer to the use of the sensor in 'air'. It is also possible to refine the parameters using a self-calibration procedure included in the SDK.

¹⁰ <https://github.com/raulmur/ORBSLAM2/>

¹¹ <https://www.agisoft.com/>

For our applications, the first step consists in the calibration in water of the optical stereo system within its underwater housing with flat port (section 3.1). Based on previous experience and practical evidence, we opt for calibration based on the classic photogrammetric model carried out under practical working conditions, a method that allows the effects of refraction to be absorbed into the estimated orientation parameters. A preliminary calibration is carried out in laboratory (pool) under controlled conditions. The estimated interior and relative orientation parameters are then used as approximate values in a self-calibrating SLAM processing in operative scenarios.

5.3 Opto-acoustic survey and estimate of targets 3D coordinates

The area where the opto-acoustic targets have been positioned is surveyed with the ROV empowered with the developed survey architecture (section 4). Real-time acquisition is verified thanks to the SLAM module and opto-acoustic visualisation tools implemented. Once the survey has been completed, the SLAM solution is fed into a full photogrammetric workflow to produce a complete model of the surveyed area, in which the coordinates of the opto-acoustic targets are also determined.

5.4 Estimating opto-acoustic relative orientation

The final step consists in estimating a seven-parameter Helmert transformation, which provides the opto-acoustic system calibration (relative orientation) as shown in equation 2

$$\begin{bmatrix} X_{T_i} \\ Y_{T_i} \\ Z_{T_i} \end{bmatrix}^{Pho_j} = \lambda R_S^{Pho} \begin{bmatrix} \rho_{i_j} \delta_{i_j} \sin \alpha_{i_j} \\ \rho_{i_j} \cos \delta_{i_j} \cos \alpha_{i_j} \\ \rho_{i_j} \sin \delta_{i_j} \end{bmatrix}^{S_j} + \begin{bmatrix} X_{OS} \\ Y_{OS} \\ Z_{OS} \end{bmatrix}^{Pho} \quad (2)$$

where $i = 1, \dots, n$ target identification number
 $j = 1, \dots, m$ ROV position (observation in time or frame)
 $Pho_j = j$ -th optical sensor position
 $S_j = j$ -th sonar position
 $X_{T_i}, Y_{T_i}, Z_{T_i} = 3D$ coordinates of the i -th target in the j -th Pho position
 $\rho_{i_j}, \delta_{i_j}, \alpha_{i_j} =$ range, elevation and azimuth angles of the i -th target in the j -th S sonar image
 $\lambda =$ scale factor between the sonar and the optical sensor
 $R_S^{Pho} =$ rotation matrix from sonar to optical sensor
 $X_{OS}, Y_{OS}, Z_{OS} =$ translation vector from sonar to optical sensor

The unknowns are the three coordinates (X_{OS}, Y_{OS}, Z_{OS}) of the acoustic sensor centre with respect to the optical sensor (in our case, the ZED2 left camera), the three rotations (ω, ϕ, κ) and a scaling factor (λ) between the two reference systems. Due to the loss of elevation information in the sonar image, the elevation angles (δ_{i_j}) of each target are also unknown.

The observations consist of the distance (ρ_{i_j}) and azimuth angle (α_{i_j}) of the targets measured in the sonar projection plane at every acquisition time. The 3D coordinates of the targets in the corresponding optical sensor position reference frame are matched to them. Each target in a sonar image therefore provides three equations, but increases the number of unknowns by one.

6. POOL EXPERIMENT

Before proceeding to sea trials, we tested the developed multi-sensor platform, the software architecture and implemented calibration procedure under controlled conditions, such as those of a pool.

6.1 Description of the experimental conditions

The first tests took place in a pool installed in our premises. The tank measures 4 meters by 2 meters with a depth of 1.20m. The bottom is covered by a rigid plate measuring 2.20m by 1.5m and featuring a random pattern to support automatic approaches of image orientation. On the plate, 360 circular coded targets are homogeneously distributed whose coordinates were photogrammetrically measured without the water.

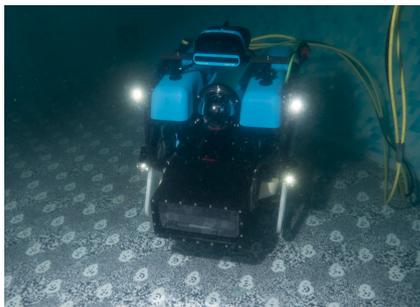


Figure 6. BlueROV2 equipped with Stereolab ZED2 and Oculus M750d sensors

On the photogrammetric grid, five opto-acoustic targets are distributed. For these initial tests, we used aluminium cylinders of 90 mm diameter and 20 mm height, i.e. of material, size and shape such to be recognisable both in optical and sonar images.



Figure 7. Opto-acoustic targets deployment

The BlueROV2 is deployed in the pool to first adjust its balance and then to orientate the sensors in order to have images including the five opto-acoustic targets.

Two acquisition sessions were then carried out (Table 3). The first for the calibration of the ZED2 stereo optical sensor underwater and the second for the estimation of the optico-acoustic relative orientation. Note that the optico-acoustic targets were deployed only for the second session of acquisition

In these experiments, one laptop is dedicated to the control/command of the ROV and a second one to data acquisition. The pilot can thus guide the vehicle by combining the piloting camera (BlueROV2) with the real time return of the ZED2 stereo images, the M750d sonar image and finally the navigation

Sensor	Dataset1		Dataset2	
	ZED2	M750d	Zed2	M750d
Location	Pool			
Description	Zed2 calibration		opto-acoustic calibration	
Image Resolution (pixel)	672 x 376	1920 x 1080	672 x 376	1920 x 1080
Frame Rate (fps)	2	≈15	2	≈15
Nb of pictures processed	768	NA	916	NA
Sonar range (m)	NA	2	NA	2
Sonar aperture (°)	NA	130	NA	130

Table 3. Datasets description

aid provided by the SLAM processing. Both SLAM process and incremental SfM can then be running for trajectory control during the opto-acoustic survey (figure 8).

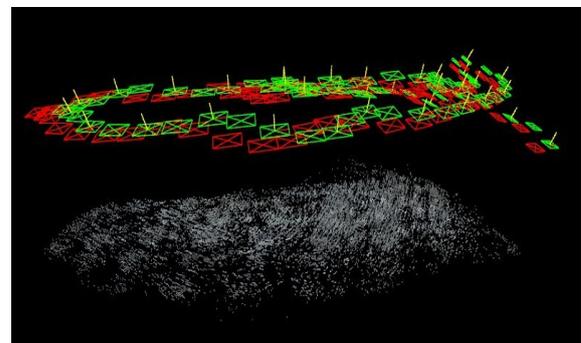


Figure 8. RealTime SLAM process

6.2 Results

The first step is to calibrate the ZED2 sensor (calibration parameters of each camera and of the stereo system) using Dataset1 (Table 3). With the zed calibration parameters it is possible to process Dataset2 (Table 3) from which the object coordinates of the opto-acoustic targets can be estimated. At the moment the procedure requires the manual selection of sonar frames in which the five targets are visible, whose correspondence with the ZED2 images is achieved through their respective times.

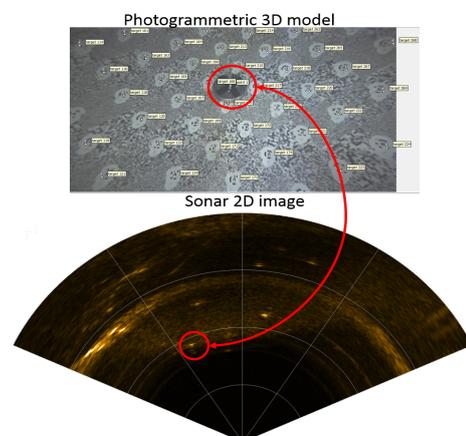


Figure 9. Target identification and correspondence between ZED2 and sonar images

We measured the five opto-acoustic targets in four sonar views, for a total of 20 observations to solve the system in equation (2). The estimated values with their standard deviations (Std) are shown in the Table 4.

	Value	Std
X_{OS} (cm)	4.3	0.8
Y_{OS} (cm)	16.4	3.0
Z_{OS} (cm)	30.7	1.1
ω ($^{\circ}$)	-89.70	0.02
ϕ ($^{\circ}$)	0.01	0.01
κ ($^{\circ}$)	-0.02	0.01
λ	1.00	0.01

Table 4. Sonar to Optical sensors relative orientation parameters

7. CONCLUSIONS

The main advantage of the opto-acoustic calibration method described in this paper lies in its applicability in an operational environment. Contrary to the approaches described in section 2, here the deployment of targets on the bottom is quite flexible, with the only constraint being that targets must be measurable in both the optical and acoustic systems. Although the minimum number of targets is four, given the high uncertainties involved (software synchronisation, low sensor resolution), redundancy is advisable. This can be accomplished not only by increasing the number of targets, but also by increasing the number of sonar images where they are marked.

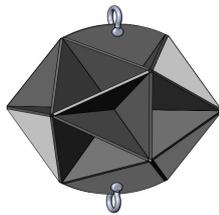


Figure 10. Sonar target Multi-triplane design

The experimental phase described in this paper was performed in a controlled environment where the opto-acoustic targets are easily identifiable in both the 3D photogrammetric model and the sonar images. The signature intensity of the targets in the acoustic image is however low and can become problematic in a real environment. In operational conditions, both the topography, artifacts or objects present in the area can be a source of echoes making it difficult to discriminate these signatures from calibration targets. In 1975, (Wallace et al., 1975) presents the results of an experimental investigation of several passive sonar targets used for calibration, marking, and tracking. Although it is concluded that, of the targets studied, only the sphere can be used as calibration standards, we will favor a multi-triplane target design because of its higher target strength. Future works will delve into the accuracy assessment of the relative orientation computation method here presented. Further experiments will then be carried out in a natural environment using the targets shown in figure 10 to assess the solution in an operational setting.

ACKNOWLEDGEMENTS

This work has been partially supported by The Direction Generale de l'Armement (DGA) of France within the framework of the CAPRICA project.

REFERENCES

- Aguirre-Castro, O. A., Inzunza-González, E., García-Guerrero, E. E., Tlelo-Cuautle, E., López-Bonilla, O. R., Olguín-Tiznado, J. E., Cárdenas-Valdez, J. R., 2019. Design and Construction of an ROV for Underwater Exploration. *Sensors*, 19(24), 5387.
- Aristizábal, L., Rúa Pérez, S., Gaviria, C., Osorio, S., Zuluaga, C., Posada, N., Vasquez, R., 2016. Design of an open source-based control platform for an underwater remotely operated vehicle. *DYNA*, 83, 198-205.
- Aykin, M. D., Negahdaripour, S., 2013. Forward-look 2-d sonar image formation and 3-d reconstruction. *2013 OCEANS-San Diego*, IEEE, 1–10.
- Bruno, F., Lagudi, A., Ritacco, G., Muzzupappa, M., Guida, R., 2015. Opto-acoustic data fusion for supporting the guidance of remotely operated underwater vehicles (ROVs). *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W5, 47–53. <http://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XL-5-W5/47/2015/>.
- Capocci, R., Dooly, G., Omerdić, E., Coleman, J., Newe, T., Toal, D., 2017. Inspection-Class Remotely Operated Vehicles—A Review. *Journal of Marine Science and Engineering*, 5(1). <https://www.mdpi.com/2077-1312/5/1/13>.
- Clark, C. M., Olstad, C. S., Buhagiar, K., Gambin, T., 2008. Archaeology via underwater robots: Mapping and localization within maltese cistern systems. *2008 10th International Conference on Control, Automation, Robotics and Vision*, 662–667.
- Davison, 2003. Real-time simultaneous localisation and mapping with a single camera. *Proceedings Ninth IEEE International Conference on Computer Vision*, 1403–1410 vol.2.
- Drap, P., Merad, D., Boï, J.-M., Mahiddine, A., Peloso, D., Chemisky, B., Seguin, E., Alcalá, F., Bianchimani, O., 2014. Underwater Multimodal Survey: Merging Optical and Acoustic Data. O. Musard, L. Le Dû-Blayo, P. Francour, J.-P. Beurier, E. Feunteun, L. Talassinis (eds), *Underwater Seascapes*, Springer International Publishing, Cham, 221–238.
- Ferreira, F., Machado, D., Ferri, G., Dugelay, S., Potter, J., 2016. Underwater optical and acoustic imaging: A time for fusion? a brief overview of the state-of-the-art. *OCEANS 2016 MTS/IEEE Monterey*, 1–6.
- IMCA, International Marine Contractors Association, 2014. Code of Practice for The Safe & Efficient Operation of Remotely Operated Vehicles. Technical Report IMCA R 004 Rev. 3, IMCA, International Marine Contractors Association.
- Lagudi, A., Bianco, G., Muzzupappa, M., Bruno, F., 2016. An alignment method for the integration of underwater 3D data captured by a stereovision system and an acoustic camera. *Sensors*, 16(4), 536.
- Ledezma, F. D., Amer, A., Abdellatif, F., Outa, A., Trigui, H., Patel, S., Binyahib, R., 2015. A Market Survey of Offshore Underwater Robotic Inspection Technologies for the Oil and Gas Industry. *SPE Saudi Arabia Section Annual Technical Symposium and Exhibition*, Society of Petroleum Engineers, Al-Khobar, Saudi Arabia.

Liu, J., Li, B., Guan, W., Gong, S., Liu, J., Cui, J., 2020. A Scale-Adaptive Matching Algorithm for Underwater Acoustic and Optical Images. *Sensors*, 20(15). <https://www.mdpi.com/1424-8220/20/15/4226>.

Mari, Z., Nawaf, M. M., Royer, J.-P., Seinturier, J., Chemisky, B., Drap, P., 2017. Target reacquisition and identification using real-time visual odometry and 3D, reconstruction of underwater scenes. *Undersea Defence Technology Bremen*, 8.

Mur-Artal, R., Montiel, J. M. M., Tardós, J. D., 2015. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics*, 31(5), 1147-1163.

Nawaf, M. M., Merad, D., Royer, J.-P., Boi, J.-M., Saccone, M., Ben Ellefi, M., Drap, P., 2018. Fast Visual Odometry for a Low-Cost Underwater Embedded Stereo System †. *Sensors*, 18(7). <https://www.mdpi.com/1424-8220/18/7/2313>.

Negahdaripour, S., Sekkati, H., Pirsivash, H., 2007. Opti-acoustic stereo imaging, system calibration and 3-d reconstruction. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8.

Negahdaripour, S., Sekkati, H., Pirsivash, H., 2009. Opti-Acoustic Stereo Imaging: On System Calibration and 3-D Target Reconstruction. *IEEE Transactions on Image Processing*, 18(6), 1203–1214. <http://ieeexplore.ieee.org/document/4815402/>.

Nicosevici, T., Garcia, R., Carreras, M., Villanueva, M., 2004. A review of sensor fusion techniques for underwater vehicle navigation. *Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No.04CH37600)*, 3, 1600–1605 Vol.3.

Nocerino, E., Menna, F., Chemisky, B., Drap, P., 2020. 3D sequential image mosaicing for underwater navigation and mapping. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 43, 991–998.

Nocerino, E., Menna, F., Fassi, F., Remondino, F., 2016. Underwater calibration of dome port pressure housings. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*.

Poore, K., Kitts, C., Wheat, G., Kirkwood, W., 2016. A small scale roV for shallow-water science operations. *OCEANS 2016 MTS/IEEE Monterey*, 1–6.

Rahman, S., Li, A. Q., Rekleitis, I., 2018. Sonar visual inertial slam of underwater structures. *2018 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 5190–5196.

Saputra, M. R. U., Markham, A., Trigoni, A., 2018. Visual SLAM and Structure from Motion in Dynamic Environments. *ACM Computing Surveys (CSUR)*, 51, 1 - 36.

Tang, Z., Lu, J., Wang, Z., Ma, G., 2020. Three dimensional height information reconstruction based on mobile active sonar detection. *Applied Acoustics*, 169, 107459.

Wallace, R. H., Hillery, H. V., Barnard, G. R., Marks, B. M., McKinney, C. M., 1975. Experimental investigation of several passive sonar targets. *The Journal of the Acoustical Society of America*, 57(4), 862–869. <http://asa.scitation.org/doi/10.1121/1.380526>.