Towards real-time underwater photogrammetry for subsea metrology applications

Conference Paper · June 2019

DOI: 10.1109/OCEANSE.2019.8867285

CITATIONS
0

READS
65

8 authors, including:

Fabio Menna
Fondazione Bruno Kessler
87 PUBLICATIONS **1,520** CITATIONS
SEE PROFILE

Erica Nocerino
Aix-Marseille Université
92 PUBLICATIONS **1,169** CITATIONS
SEE PROFILE

Mohamad Motasem Nawaf
Aix-Marseille Université
24 PUBLICATIONS **82** CITATIONS
SEE PROFILE

Julien Seinturier
COMEX SA
53 PUBLICATIONS **341** CITATIONS
SEE PROFILE

Towards real-time underwater photogrammetry for subsea metrology applications

Fabio Menna
Innovation Department
COMEX SA
36 bd de l'Océan - CS 80143 - 13275 Marseille, France
f.menna@comex.fr

Erica Nocerino
LIS UMR 7020
Aix-Marseille Université, CNRS, ENSAM, Université De Toulon
Bâtiment Polytech, Avenue Escadrille Normandie-Niemen, 13397, Marseille, France
erica.nocerino@univ-amu.fr

Mohammad Motasem Nawaf
LIS UMR 7020
Aix-Marseille Université, CNRS, ENSAM, Université De Toulon
Bâtiment Polytech, Avenue Escadrille Normandie-Niemen, 13397, Marseille, France
mohamad-motasem.nawaf@univ-amu.fr

Julien Seinturier
Innovation Department
COMEX SA
36 bd de l'Océan - CS 80143 - 13275 Marseille, France
j.seinturier@comex.fr

Alessandro Torresani
3D Optical Metrology (3DOM) unit
Bruno Kessler Foundation (FBK)
Via Sommarive 18, 38123 Povo-Trento, Italy
atorresani@fbk.eu

Pierre Drap
LIS UMR 7020
Aix-Marseille Université, CNRS, ENSAM, Université De Toulon
Bâtiment Polytech, Avenue Escadrille Normandie-Niemen, 13397, Marseille, France
pierre.drap@univ-amu.fr

Fabio Remondino
3D Optical Metrology (3DOM) unit
Bruno Kessler Foundation (FBK)
Via Sommarive 18, 38123 Povo-Trento, Italy
remondino@fbk.eu

Bertrand Chemisky
Innovation Department
COMEX SA
36 bd de l'Océan - CS 80143 - 13275 Marseille, France
b.chemisky@comex.fr

***Abstract*— High accuracy underwater inspections are getting more and more important in the underwater industry where time and cost optimization represent nowadays the main innovation drivers. The subsea industry is undergoing a digital transformation process and for this reason, methods that can provide real-time accurate 3D digital measurements are increasingly demanded. This paper provides a short review of the main techniques currently used in subsea metrology to then present an experimental study carried out to evaluate the accuracy potential of three vision-based techniques well-known in photogrammetry, namely visual odometry with and without windowed bundle adjustment, and keyframe based simultaneous localization and mapping (SLAM). The accuracy evaluation is done using an ORUS 3D® subsea photogrammetry system using a certified 3D underwater reference test-field available at COMEX facilities, whose spatial coordinates are known with sub-millimetre accuracy. A critical assessment of results is presented against currently set tolerances for the subsea metrology industry.**

Keywords*— *Underwater photogrammetry, SLAM, visual odometry, subsea metrology, 3D monitoring, accuracy evaluation

I. INTRODUCTION

Over the last decade, technologies for underwater mapping, 3D inspections, and subsea metrology have gained a leading role within the oil and gas industry [1]. The growing global demand for energy, accompanied by a general reduction of hydrocarbon resources have pushed towards the development of fields in deeper waters and harsher environments. Following the complexity of deep-sea operations, and a general improved efficiency demand, concepts like the Industry 4.0 and Smart maintenance have developed also for the subsea production scenario [2].

Such concepts have brought to a new production paradigm that sees the integration of artificial intelligence (AI), smart sensors, big data monitoring, predictive analysis and dynamic simulations. Within this complex framework, real-time 3D digital inspections and monitoring are fundamental for the development and maintenance of underwater production infrastructures [1]. With the aim to develop cost-effective technologies and safer offshore operations, a recent European research project has demonstrated the potential for underwater 3D surveys to be carried out even remotely, in presence of communication latencies [3].

Amongst the different applications of underwater 3D surveying technologies, subsea metrology surveys aim to provide accurate and traceable 3D measurements, typically positions and orientations, of structures like pipeline interconnections linking hydrocarbons reservoirs to processing and storage facilities (Figure 1). Indeed, many parts such as interconnecting pipes needed for subsea field installations are manufactured and finalized at sea on the surface vessel, on site. Because of the high cost rates per day of operations at sea, over the last years great efforts have been made to improve the accuracy of real-time methods, for example by developing data fusion techniques that integrate inertial and acoustic methods [1]. The current state-of-the-art techniques in subsea metrology make use of expensive acoustic localization systems consisting of a network of Long Baseline (LBL) beacons deployed at sea bed and on the vehicle to be tracked, typically remotely operated vehicles (ROV) or autonomous underwater vehicles (AUV).

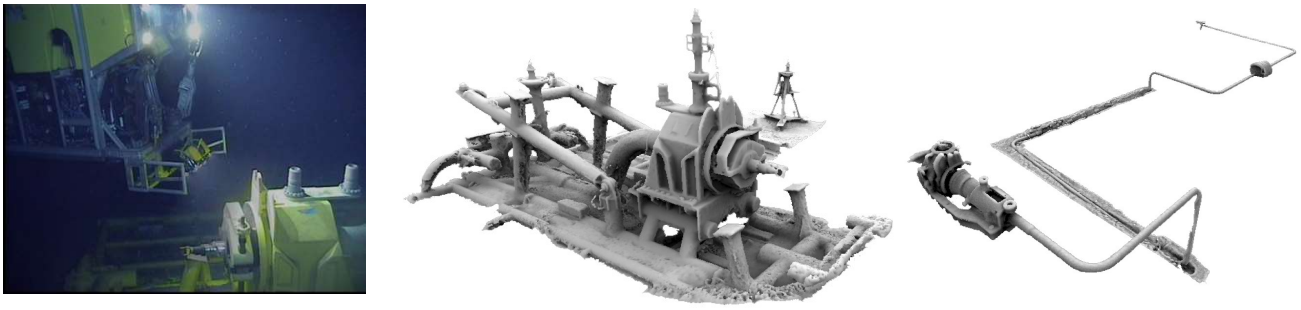

Fig. 1. Example of photogrammetry based subsea metrology operations using an ORUS3D® 3kv system mounted on a work class ROV (left) and derived 3D models of a FLET structure (center) and a spool (right).

Such localization systems integrate acoustic range measurements, depth sensing and inertial navigation techniques into a multilateration computation and require specialized team, with a time-consuming installation and set-up which increase the time and subsequently costs of subsea metrology operations.

At research level, automatic localization, navigation, and mapping algorithms, such as visual odometry and Simultaneous Localization And Mapping (SLAM), are receiving more and more attention. Tanks to their successful application, flexibility, scalability, and low cost, these methods are becoming very attractive also for the subsea industry. However, in this specific field, tolerances are very strict and validation methods are mandatory.

This paper provides a first attempt to evaluate vision-based localization and mapping algorithms within the specific context of subsea metrology measurements using a highly accurate underwater environment ground truth (pool) and a certified photogrammetry based imaging system (ORUS3D®).

In this study the authors first present an introductory review of subsea metrology survey methods to then investigate the use of vision based real time techniques, namely SLAM and Visual Odometry, as a navigation tool to assist and control a Remotely Operated Vehicle (ROV) while performing inspection and monitoring tasks underwater. In particular, the paper focuses on methods of investigation able to assess the accuracy of both trajectory and 3D tie points used in the image orientation process and evaluate whether the preliminary real time reconstruction meets the tolerances defined in typical subsea metrology surveys.

II. Review Of Subsea Metrology Survey Methods

The International bureau of weights and measures (BIPM) defined metrology as the '*science of measurement, embracing both experimental and theoretical determinations at any level of uncertainty in any field of science and technology.*'[1] Its main aims are the qualification, verification and validation of measured data according to accepted standards, which require traceability, i.e. the '*property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty*' [4]. In other words, traceability implies the identification of references or standards. These allow for the assessment of measurement uncertainty and comparison of different measurement results, which become traceable to the same reference. In industry, metrology ensures that parts and components are of the required quality and accuracy, i.e. manufactured according to standards developed at different levels, from international and national basis to industry specific or even customized for internal purposes.

When it comes to the underwater environment, metrology is commonly referred as the process of acquiring accurate and traceable dimensional measurements of subsea structures [1]. Subsea structures are pipeline interconnections (hubs or flanges), joining subsea assets, widely used in the offshore, marine and underwater engineering companies (Figure 1, centre and right). The main measurements required in subsea metrology are [1, 5]:

- horizontal position and depth of the hubs;
- hub-to-hub slant and horizontal distances (also called baseline);
- hub-to-hub relative heading and attitude;
- spool azimuth (i.e., the bearing of the spool from the hub) and angle of approach (difference between the spool azimuth and hub headings);
- seabed profile along the structure route.

Typical subsea metrology tolerances are defined by [1]:

$$Relative\ distance_{X,Y,Z} = [50 \div 150]\ mm \quad (1)$$

$$Relative\ angle_{roll,pitch,heading} = [0.5^o \div 1.0^o] \quad (2)$$

They take into account the permissible hub misalignment, which may be due to several factors, such as stress analysis, fabrication tolerance and possible deformation resulting from deployment operations.

Today, the most commonly used techniques for subsea metrology are: long baseline (LBL) acoustics, diver taut wire, digital taut wire, photogrammetry, Inertial Navigation Systems (INSs), Simultaneous Localization And Mapping (SLAM)

[1] https://www.bipm.org/en/worldwide-metrology/

techniques and laser scanning. In the followings a short description for each technique is provided after [1, 6]:

A. LBL acoustic systems

Absorption of sound waves in water is weak. Thanks to this property, acoustic systems are so widely used for underwater positioning in marine (offshore) mining that they are often compared to Global Navigation Satellite Systems (GNSS) on land [7]. LBL systems derive their name from the distance between seabed transponders or beacons, distances usually ranging from 50 to 2000 m. The accuracy of the system depends on the frequency applied, which may vary from 0.02 m to 0.5 m. For underwater positioning, LBL typically consists of at least three transponders, installed on the seafloor in accurately known positions, and one transceiver, on a submersible or a surface vessel. The measurement principle is based on trilateration: the time the signal travels from the transponders to the receiver and backwards is measured and converted to the travelling distance, knowing the velocity of sound. For subsea metrology, the transponders are placed on the structures using specific metrology interface tooling and further spread around the structures to form a robust geometrical network. Pressure/depth sensors and gyros are used to measure the structures depth and attitude, respectively. An accurate determination of the speed of sound at the measurement site is an essential component of the accuracy potential of this method.

B. Diver taut wire

It provides the direct distance between the flanges based on tape measurements. Jig plates with protractor markings are mounted directly above the flanges and wire ropes are tensioned to measure the distance, orientation and attitude between the subsea structures.

C. Digital taut wire

Primarily developed for ROV operations, it also entails the use of digital inclinometers and pressure sensors. Algorithms can be implemented to correct for catenary effects caused by, for example, currents or gravity.

D. Photogrammetry

Photogrammetry is a well-known high-accuracy surveying technique based on measurements from photographs; interested readers may refer to specific textbooks on this topic such as [8]. In the context of subsea metrology, photogrammetry heavily relied on dimensional control, requiring surveyed scale bars or known coded targets on the subsea structures. Today the trend is to fully automatize the acquisition and measurement process and to use contactless systems.

E. INSs

The principle involved in inertial navigation system (INS) metrology is to use three orthogonal accelerometers measuring linear acceleration in the X, Y and Z plane, combined with three orthogonal gyroscopes measuring angular velocity likewise.

F. Acoustically aided INS SLAM

In the subsea metrology context, SLAM usually refers to the integration of INS with acoustic methods, mainly LBL which is employed as range-based sensor [5, 9]. INS integrated SLAM is considered an augmentation to conventional LBL method, allowing for a reduction in the number of beacons required to guarantee the desired accuracy. In section III, automatic real time vision-based SLAM approaches are presented in details, being the main focus of this work

G. Sonar

Subsea sonar scanners, a development of the more traditional multi-beam echo-sounders, work similar to a terrestrial laser scanner. Mounted on a tripod or an underwater vehicle, they measure the XYZ position of individual points, relative to the sonar scanner position.

H. Laser scanning

Underwater lasers are orders of magnitude more powerful than their counterpart on land, because light absorption is much higher in water than in air. They require output powers generally in the megawatt range, which make them expensive and difficult to build. They are used both in static (mounted on tripod) and dynamic (attached to a moving vehicle) mode.

III. VISION-BASED SLAM

SLAM methods (Table I) are receiving more and more attention for subsea measurements because of their significantly lower cost with respect to acoustic positioning methods [10-13]. Moreover, open source frameworks are becoming publicly available, making the integration of such technology much easier. Born for real time robot navigation [14], their use can be seen as a stand-alone alternative solution for limited areas or as an additional technology, integrated with current state-of-the-art navigation technology (INS and acoustic systems).

A. SLAM, SfM or VO?

SLAM and structure from motion (SfM) aim both at estimating the pose of the agent (the robot for the SLAM and a camera for the SfM) and at reconstruction (or mapping) the environment (or the 'structure') at the same time. The two main classes of algorithms were originally developed in two different communities, respectively robotic and computer vision [15]. According to [16], before his seminal work on SLAM with a single camera [17], the mobile robotics community had almost completely abandoned pure vision-based navigation approaches and the computer vision community had been almost completely disinterested from real-time and robotics applications.

In a SLAM-based approach, data coming from different sensors, or modules, are fused together to estimate the system position and attitude (the state vector) and build the map of the environment. Crucial for SLAM approaches is the identification of the so called 'loop closure', i.e. the detection of a previously mapped place and consequent re-localization of the system within the already measured environment. The 'loop closure' reduces the drift accumulated in the SLAM solution over time [18].

When cameras are the only sensors used, SLAM is based on visual information only, and is therefore called visual SLAM (V-SLAM or vSLAM, [20-21]).

At the early stage of development, the main difference between visual SLAM and SfM was that the first was mainly thought for real-time (or on-line) computation while the second was traditionally performed off-line, meaning that all the data

and measurements (i.e. images) are provided and (post-) processed together. [15] considered MonoSLAM [17] the first approach to bring the general SLAM problem from the robotic community into pure vision.

However, it should be noticed that also SLAM can be formulated as full or off-line problem, when the whole trajectory and the map is estimated providing all the sensor data and measurements. On the contrary, online SLAM updates incrementally both the agent pose and map with the most recent estimates from the sensors. The two approaches differ in the estimation techniques implemented [21]: filter-based approaches (such as the Kalman filter) are most suitable for iterative, real-time implementation; optimization-based methods (bundle adjustment, BA, or graph-based SLAM) are usually adopted for solving the full SLAM approach. The feature measurements are integrated by estimating the probability distribution in filter-based approaches or through optimization in BA [15].

Visual odometry (VO) consists in estimating the motion of a single camera or stereo systems from visual input (images or video frames) alone [22]. The main differences between SLAM and VO are explained in [23] and are here summarised. While SLAM and visual SLAM aim to obtain a global and consistent estimate of trajectory and map, VO is mainly devoted to recover the path incrementally, potentially optimizing only over the last n poses (also called windowed BA). VO can be implemented as step for a complete SLAM algorithm, where also loop closure and possibly a global optimization step are performed. Visual SLAM is potentially more accurate of VO, because more constraints are enforced on the mobile path; however, this does not ensure higher robustness, since outliers not detected in the loop closure can critically affect the map correctness.

Table I summarizes the main hardware and algorithmic components of vision-based SLAM algorithms (after [24]). For a recent review of visual-inertial SLAM approaches, interested readers are addressed to more specific papers on the topic, such as the work by [25].

B. Underwater SLAM

Under the water, SLAM techniques have been used to fuse inertial and acoustic positioning systems in particular in subsea metrology industry [1], or for autonomous underwater robot navigation and localization using imaging sonar and visual sensors. First methods were focused on using acoustic images [26-30] to move then to visual based methods [10-13]. The specific challenges to face in implementing SLAM for real underwater environments are critically presented in [31], who proposed a bio-inspired algorithm based on the localization approach performed by neural structures found in mammals. A more comprehensive review of different techniques used for localization and mapping under the water can be found in [14,31,32].

C. Underwater SLAM benchmarking

Although the use of real time techniques for underwater navigation and mapping is not recent, few studies have been presented to evaluate the accuracy of such methods in a real environment using a certified ground truth benchmark surveyed with high accuracy techniques. Besides the well-known adverse environmental issues (visibility, light attenuation, suspended particles, darkness, etc.), which may negatively impact the performances of well-established methods for in-air applications, critical is the set-up of a ground truth for the evaluation of proposed methods.

Studies and benchmark datasets for the evaluation of different SLAM techniques exist above the water [33-34]; using publicly available datasets, issues have been reported in the literature such as trajectory or scale drifts reporting also a metric accuracy evaluation [35-37].

Under the water, remarkable studies have been published [24, 38] with the aim of evaluating state-of-the-art open-source SLAM approaches. The assessment is carried out on datasets acquired in real subaquatic environments; however, a metric and reliable ground truth is missing.

To overcome this issue, [12] created simulated underwater scenarios with the Underwater Simulator – UWSim [39]. This is a very interesting approach, which, however, can hardly reproduce the full complexity of the real environment.

Summarizing, under the water, comparative studies have been focused on either (i) simulated trajectories [12], or on comparing real-time results with (ii) reference trajectory computed using visual, inertial, depth and sonar range data [24], or (iii) photogrammetrically derived trajectories obtained offline, in post processing, using for example BINGO software [40], Agisoft Photoscan [41], Colmap SfM application [13].

TABLE I. TAXONOMY OF MAIN VISION-BASED SLAM APPROACHES

IMAGE(VISUAL)-BASED LOCALIZATION AND MAPPING METHODS							
Number of cameras	*Additional sensors*	*Indirect vs direct*	*Sparse vs dense*	*Loosely vs tightly coupled optimization*	*With or without loop closure*	*Online vs offline*	*Filter vs optimization-based approach*
Mono, stereo or more	IMU, acoustic, laser, etc.	Feature based with geometric error optimization vs actual pixel intensities with photometric error optimization	Without or with geometry prior (connectedness of image region, smoothness) to estimate dense or semi-dense geometry	Measurements from the different sensors are pre-processed before the optimization or directly used without pre-processing and jointly optimized (correlation between all the measurements can be taken into account)	Recognition of already visited places (relocalization)	Real time vs post-processing (with all the data and measurements available)	Kalman filters vs BA or graph-based methods

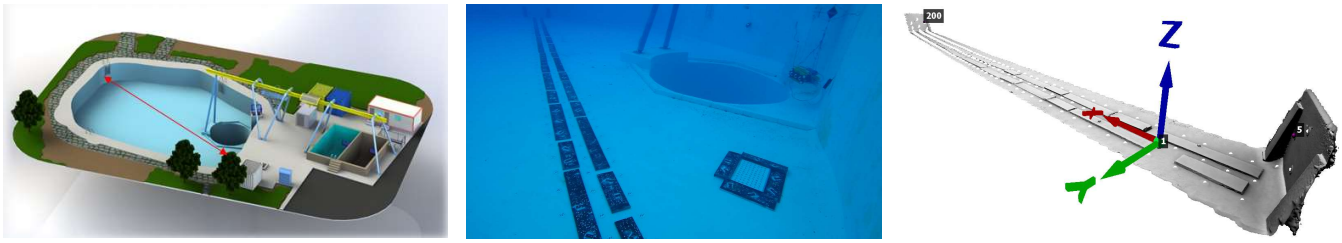

Fig. 2. COMEX test pool (left), an underwater view of the COMEX 3D photogrammetry reference test-field (center) and its 3D mesh representation with refernce axes superimposition (right).

IV. Experiment Design

With the aim to metrically assess the accuracy potential of real time vision-based techniques in a subsea metrology context, an experiment was carried out at COMEX test facilities in Marseilles, France. A comparison is performed between three vision-based techniques characterized by an increasing algorithmic complexity, namely:

1) A simple visual odometry algorithm relying on forward stereo-triangulation and image resection;

2) Visual odometry with local windowed BA;

3) Keyframe based SLAM with loop cloure (ORB-SLAM2).

A description of utilized instruments and methods is hereafter provided.

B. COMEX 3D underwater photogrammetry reference test-field

In 2018, a high-accuracy 3D reference test-field was set-up in the COMEX pool as part of quality control procedures of the ORUS 3D® system, and research and development tasks of COMEX Innovation department, such as evaluating the accuracy performances of vision-based algorithms against an accurate certified ground truth. Photogrammetric targets made of a material that is resistant to UV and chlorinated water were installed along a linear section of the pool. The realized reference consists of a horizontal transect joining two opposite submerged vertical walls of the pool as visible in Figure 2.

Table II reports the main dimensional characteristics of the 3D underwater reference test-field.

TABLE II. COMEX 3D U/W TEST-FIELD MAIN SPECS

Length (X)	**Width (Y)**	**Depth Variation (Z)**
30 m	1.2 m	1.4 m
Number of photogrammetric targets	**Estimated accuracy of reference coordinates σ_{XYZ}**	
200	<0.5 mm	

Metrology techniques using a laser tracker were performed after emptying the pool to determine with high accuracy the 3D targets' coordinates. The laser tracker Spherically Mounted Retroreflector (SMR) was aligned to be in tangency with the photogrammetric circular target in four points, then the centre of the photogrammetric target was determined through a best fit circle. The multi-station measurements were then adjusted through least square procedures. The reference coordinate system is set with the X axis pointing along the main direction of the test-field, Z vertical and Y according to the right-hand rule convention.

C. Photogrammetric sensor: ORUS3D® 3kv system

The ORUS 3D® (Figure 3) is an underwater photogrammetry measurement system born for real-time contactless deep-sea surveys, today certified by Bureau Veritas - Marine & Offshore division to achieve sub millimetre accuracy up to 5 m distance and sub centimetre accuracy on 30 m long transects. The system was developed upon the outcomes of the ROV3D project [40].

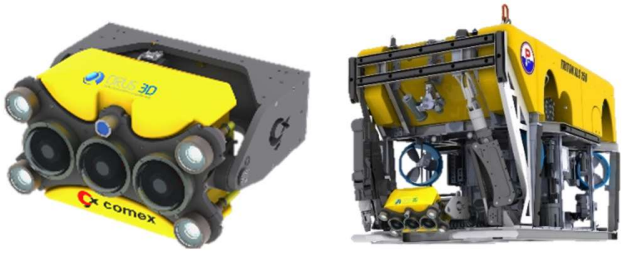

Fig. 3. ORUS 3D® trifocal sensor (left) and an example of its installation in a work class ROV (right).

Today ORUS 3D® is manufactured in 4 different models, corresponding to their maximum operating depth, respectively 300m, 1000m, 3000m, 6000m. Its design, materials and calibration procedures are optimized to guarantee high reliability and accuracy, taking into account the refractive effects of water, thermal and pressure influences.

The ORUS 3D® is composed of three main parts:

1) an Embedded Processing Unit (EPU) managing the synchronization and real-time raw data processing of connected sensors;

2) a cluster of sensors including, among the others, three industrial global shutter cameras (one high resolution-HR and two low resolution-LR), 4 LED based strobes, an attitude/ heading reference system, underwater altimeter (acoustic range finder), temperature and pressure sensors.

3) Surface Control Unit Computer (SCUC) connected to the EPU via umbilical of the ROV managing the remote control of the system parameters and displaying the remote real-time

visual-inertial odometry performed on-board the EPU with real time 3D measurement capabilities.

The synoptic schema for the main optical sensors is shown in Figure 4.

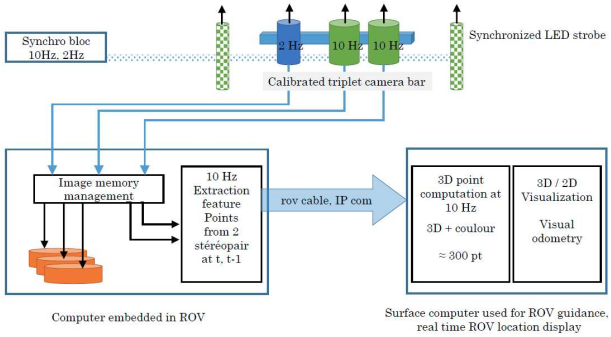

Fig. 4. Synoptic schema of the ORUS 3D® trifocal sensor [40].

The version tested in this study is the 3Kv, rated at a depth of 3000m with camera pressure housings made of titanium and dome ports made of optical glass. It can be installed on skids starting from Mid-Size ROVs class. The current photogrammetric processing relies on ad-hoc developed procedures of calibration and image triangulation with and without inertial sensor integration in the BA. From the camera calibration process, images are rectified to provide distortion free images based on single view point pinhole camera model. This pre-processing step allows a more flexible and easier cross platform and software use of the images.

D. Image dataset acquisition and reference trajectory computation

For the experiments presented in this paper, the image dataset consists of a session of 10 HZ LR stereo-rectified full HD images, subset of a full trifocal sensor acquisition carried out with an ORUS 3D® 3Kv over the COMEX test-field. Because of the featureless surface texture of the pool floor, some metallic plates with contrast random pattern were installed at the pool floor (Figure 5) to provide a minimum number of features, necessary to carry out photogrammetric orientation and successive measurements. The image acquisition was carried out with the ORUS 3D® hanged from a floating support maneuvered by a diver. The images were acquired at night to test the system in lighting conditions similar to those available at typical operative depths.

TABLE III. IMAGE DATASET CHARACTERISTICS

Number of stereo pairs	4615
Camera to object average distance	1.2 m
Nominal Baseline	165 mm
Baseline to distance nominal ratio	1/7
Ground Sample Distance – GSD	1 mm
Sidelap (left to right)	92%
Frame rate	10 Hz
Average speed (diver operated)	0.10 m/s
Overlap (along the tack)	>99%

The image sequence used in this experimentation included approximately two thirds to record the vertical walls of the pool facing each other, and another third for the rectilinear part.

Table III summarizes the main characteristics of the image dataset acquired for this experimentation.

A reference trajectory was generated by orienting the images in Agisoft Metashape software[2], using the full dataset (including HR cameras), then a bundle adjustment was run by constraining the solution using the inertial sensor and coded target ground control point coordinates as soft constraints (COMEX ORUS 3D® software).

A dense point cloud through dense image matching techniques and a mesh were generated. A cloud to mesh distance check was performed to verify the consistency between the reconstructed mesh and the reference coded target coordinates. The root mean square (RMS) of distances resulted below 1 mm. Left and right camera exterior orientations (trajectories and angles), coded targets and mesh were used as reference for comparing the real-time visual odometry and SLAM, processed offline in this experiment.

Fig. 5. Example of image pair from the image dataset used in the presented experiments.

V. DESCRIPTION OF TESTED ALGORYTHMS

A. Visual odometry based on stereo-triangulation and image resection

A simple visual odometry algorithm was implemented in C++ with the main purpose to provide a reference point from which to evaluate improvements provided by more sophisticated methods, such as windowed BA and keyframe based SLAM. Indeed, it is known that methods without map optimization are prone to drift because of systematic error accumulation. The visual odometry approach relies on feature extraction, description, and matching followed by stereo triangulation and then RANSAC based image resection. To optimize the computing resources, a constant velocity dynamic model is used to process successive image pairs only if their distance from the previous one is above a threshold. A threshold of 10 cm was set in this experiment.

[2] https://www.agisoft.com/

B. *Visual odometry with windowed bundle adjustment*

The visual odometry method tested in this study is a mono variant of the method originally developed in [40] and further improved in [41] with the addition of a windowed BA. The method is composed of two steps, first, a relative pose estimation is performed on each new image following multiple view geometry fundamentals, second, a SfM and BA adjustment approach is applied to a set of images defined by a sliding window that selects the last n images. The implementation was tested using the python scripting API of Agisoft Metashape software. A visual odometry procedure was simulated so that images are processed in sequence as if they are acquired in real-time. As for the method described in the previous subsection, a constant velocity dynamic model is used and only successive image pairs whose distance is larger than 10 cm from the previous one is processed. This solution guarantees that 3D points are triangulated with an acceptable baseline to distance ratio within the local window. In this study a window size 4 is reported for a down sampling factor of 2 of the images (half width and height). Calibration parameters were kept fixed in the BA procedure.

C. *ORB-SLAM2 mono and stereo test*

The monocular and the stereo pipelines of ORB-SLAM2[3] have been used to obtain the estimated device trajectory and the mapping of the environment. ORB-SLAM2 was developed on top of ORB-SLAM[4] adding the support for stereo and RGB configurations. The system requires as input a visual vocabulary, used to speed up feature matching and loop detections, and a configuration file containing the camera calibration and the parameters for the ORB extraction. During the tests, the default parameters were used. The input images were down sampled to a factor of 2 of original size (half width and height). As output the system provides the estimated structure of the environment (sparse point cloud of 3D tie points) and the estimated poses of the keyframes. The keyframes is a subset of the input images on which the mapping and the BA optimisation is performed; this is done to ensure the real-time behaviour of the system.

VI. Evaluation Of Results

A. *Real time trajectory and tie points accuracy evaluation*

In order to compare the trajectories from the three tested algorithms with the reference one, a datum transformation was carried out using a best fitting similarity transform (according to least-squares principle) computed using commonly available camera positions along the trajectories of the tested methods and the reference one. The transformation was then applied to calculate the new camera positions and angles in the reference coordinate system and consequently translation and angular errors as difference against the reference values. The procedure is a common practice in surveying disciplines and corresponds to the absolute trajectory error (ATE) presented in [33].

Table IV summarizes the positional and angular RMS errors (RMSEs) between the reference 3D trajectory and the tested algorithms. A drift analysis is provided by Figure 6 that shows the spatial error vectors from the reference trajectory to the one obtained with the tested algorithms. Also, using the same transformation, 3D coordinates of tie points were brought in the reference coordinate system and Euclidean distances from the mesh were calculated (Figure 7). This comparison provides the empirical accuracy of the preliminary sparse reconstruction. The scale factor was always considered, except for the visual odometry method based on stereo-triangulation and image resection, and for the stereo version of ORBSLAM2.

VII. Discussion, Conclusions, And Future Work

This paper reported about an experimental investigation carried out to evaluate the accuracy potential of real time algorithms when used for subsea metrology purposes. An ORUS 3D® (Figure 3) subsea photogrammetry system certified by Bureau Veritas - Marine & Offshore division was used to acquire a single photogrammetric strip on the underwater 3D reference set up at COMEX facilities in Marseille, France. For the first time a sub-millimetre accuracy 3D reference test field was used for the metrological performance analysis.

Three different algorithms for real-time visual-based localization and mapping, characterized by increasing complexity, were tested using standard parameters provided by their developers and documented in the related publications. The experiments provided a preliminary assessment of the achievable potential accuracy in a single strip-based nadir image acquisition, the simplest to implement and most effective in terms of reduced data acquisition and management but known to be the most prone to systematic error accumulation. Such errors, visible as lateral and vertical drifts in Figure 6, mainly depend on whether the algorithm implement or not a global trajectory and map optimization. The tested algorithms showed very promising results with trajectories differing only for few centimetres from the reference one. As expected, the simple visual odometry implementation displays strong vertical drift that makes this technique effective for real time navigation but not well suited for real time subsea metrology applications.

By looking at the Euclidean difference maps of 3D tie points against the reference mesh, and from the angular error table, it is worth noticing that, at the moment, distance tolerance of 10 cm and angular tolerance of 1 degree may be potentially met only for transects below 30 m (under the assumption of a correct external scaling, except the stereo VO and ORBSLAM2 stereo that already provides scaled measurements). A repeatability test (Figure 8) was carried out by running 5 times each algorithm showing a low repeatability of ORBSLAM2 (in particular the stereo version). This behaviour is in agreement with the results reported in [24].

[3] https://github.com/raulmur/ORB_SLAM2

[4] https://github.com/raulmur/ORB_SLAM

TABLE IV. POSITIONAL AND ANGULAR ROOT MEAN SQUARE ERRORS (RMSE) BETWEEN THE REFERENCE 3D TRAJECTORY AND THE TESTED ALGORITHMS

	VO (stereo triangulation and resection) no scale needed	**VO (windowed BA, window size = 4)**	**ORBSLAM2-MONO (LRL)**	**ORBSLAM2 STEREO no scale needed**
RMSE X|Y|Z [mm]	27 | 05 | 103	19 | 5 | 23	7 | 15 | 15	17 | 10 | 47
RMSE EULER ANGLES ω|φ|κ [deg]	1.89 | 1.33 | 1.24	0.44 | 0.40 | 0.31	0.81 | 0.34 | 0.53	1.14 | 1.09 | 1.18

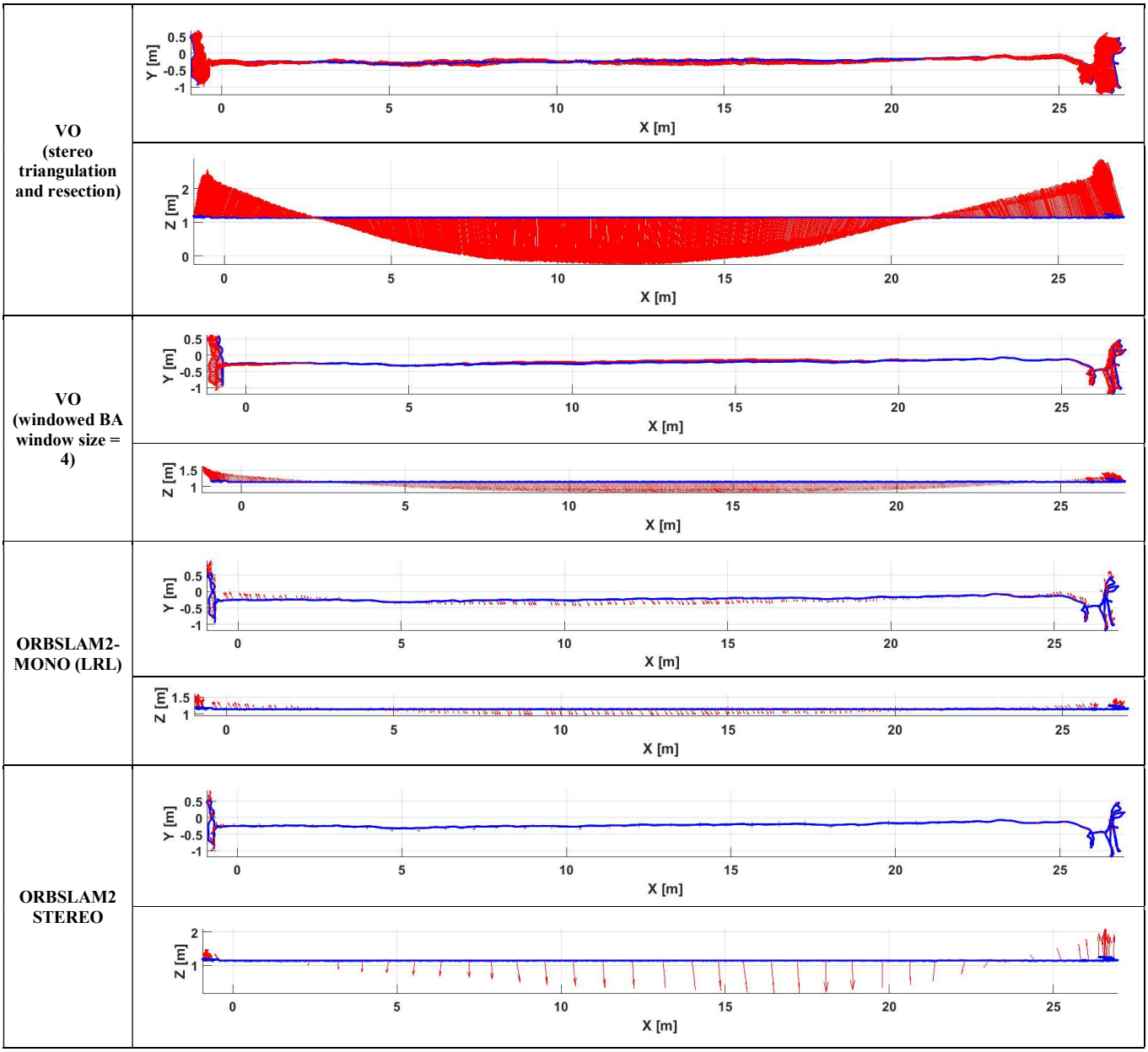

Fig. 6. Trajectory drift analysis for the tested algorythms: The vectors indicate the difference between the reference trajectory and the tested methods. Vector scale: 10. for each algorithm top (OXY) and elevation (OXZ) views are shown

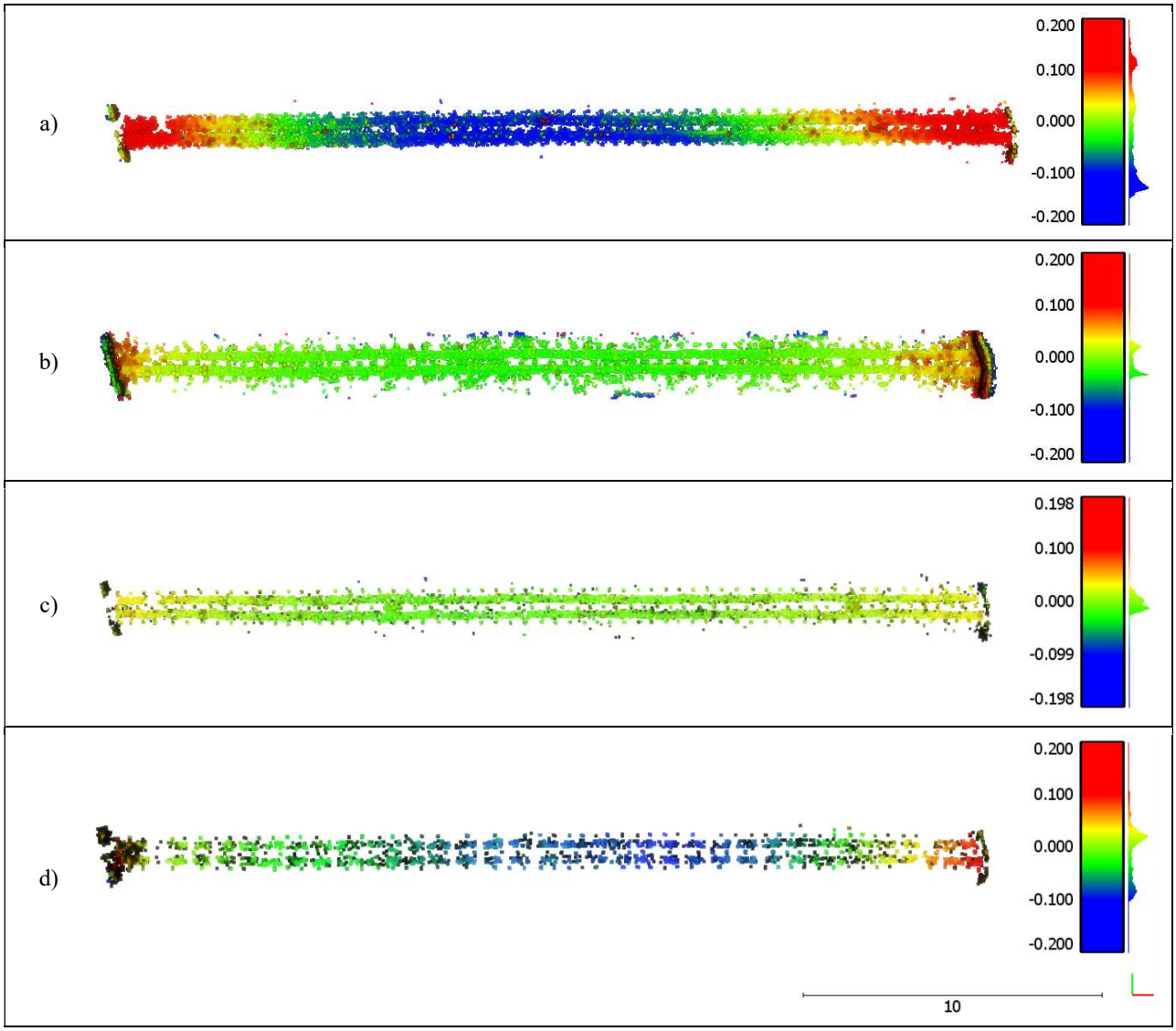

Fig. 7. 3D tie points colour coded according to their euclidean distance from the reference mesh for the global best fit test: a) VO (stereo triangulation and resection), b) VO (windowed BA window size = 4), c) ORBSLAM2 mono and d) ORBSLAM2 stereo. Values are in meters.

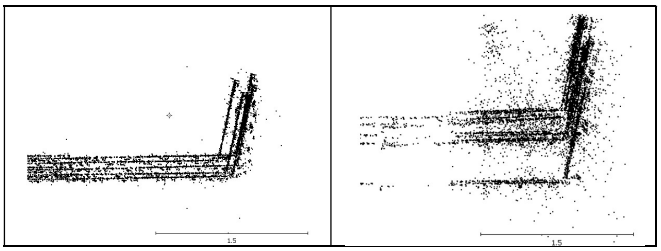

Fig. 8. Repeatability analysis for the ORBSLAM2 mono (left) and stereo (right) algorithms. A section view (plane OXZ) of one end of the pool shows that the 3D tie points from 5 different runs of the algorythm do not overlap well onto each other.

A further investigation is required to assess the sensitivity of results on the algorithm parameters. A future work will also investigate the accuracy of tested algorithms against the ROV speed and the integration of inertial sensors as well as the effect of different camera networks on the final accuracy.

REFERENCES

[1] IMCA, International Marine Contractors Association, “Guidance on subsea metrology,” IMCA S 019, 2017. https://www.imca-int.com/publications/318/guidance-on-subsea-metrology/

[2] A. Marhaug and P. Schjølberg, “Smart Maintenance-Industry 4.0 and Smart Maintenance: from Manufacturing to Subsea Production Systems,” in 6th International Workshop of Advanced Manufacturing and Automation. Atlantis Press, 2016, November.

[3] J. Gancet, D. Urbina, P. Letier, M. Ilzokvitz, P. Weiss, F. Gauch, G. Antonelli, G. Indiveri, G. Casalino, A. Birk, and M.F. Pfingsthorn, "DexROV: Dexterous undersea inspection and maintenance in presence

of communication latencies," *IFAC-PapersOnLine* 48, no. 2, 2015: pp. 218-223.

[4] VIM3 2.41, Joint Committee for Guides in Metrology, "International vocabulary of metrology – basic and general concepts and associated terms," (VIM), BIPM, 2017 (https://jcgm.bipm.org/vim/en/2.41.html).

[5] M.J. Jørgensen, N.K. Poulsen, and M.B. Larsen, "Enhanced Subsea Acoustically Aided Inertial Navigation," DTU Compute, 2015.

[6] A. Bakare, "Subsea Field Development: A Critical Review of Metrology Methods and Achievable Accuracies in Spool Tie-In Operations,". Master Thesis in Subsea Engineering at the University of Aberdeen, 2013.

[7] A. Tomczak, "Modern methods of underwater positioning applied in subsea mining," Górnictwo i Geoinżynieria, 35, 2011, pp.381-394.

[8] T. Luhmann, S Robson, S. Kyle, and J. Boehm, "Close-range photogrammetry and 3D imaging," 2013, Walter de Gruyter.

[9] P. Newman and J. Leonard "Pure range-only sub-sea SLAM," in *IEEE International Conference on Robotics and Automation* (Cat. No. 03CH37422), Vol. 2, 2003, pp. 1921-1926. IEEE.

[10] R.M. Eustice, O. Pizarro, and H. Singh, "Visually augmented navigation for autonomous underwater vehicles," IEEE Journal of Oceanic Engineering, 33(2), 2008, pp.103-122.

[11] A. Kim and R. Eustice, "Pose-graph visual SLAM with geometric model selection for autonomous underwater ship hull inspection," In IEEE/RSJ International Conference on Intelligent Robots and Systems, 2009, pp. 1559-1565. IEEE.

[12] A.C. Duarte, G.B. Zaffari, R.T.S. da Rosa, L.M. Longaray, P. Drews, and S.S Botelho, "Towards comparison of underwater SLAM methods: An open dataset collection," In *OCEANS* 2016 MTS/IEEE Monterey (pp. 1-5). IEEE.

[13] M. Ferrera, J. Moras, P. Trouvé-Peloux, and V. Creuze, "Real-Time Monocular Visual Odometry for Turbid and Dynamic Underwater Environments," Sensors, 19(3), 2019, p.687.

[14] F. Hidalgo and T. Bräunl, "Review of underwater SLAM techniques," in Automation, Robotics and Applications (ICARA), 2015, 6th International Conference on (pp. 306-311). IEEE.

[15] M.R.U. Saputra, A. Markham, and N. Trigoni, "Visual SLAM and structure from motion in dynamic environments: A survey," ACM Computing Surveys (CSUR), 51(2), 2018, p.37.

[16] A. J. Davison. "15 Years of Visual SLAM," Presentation at The Future of Real-Time SLAM: 18th December 2015 (ICCV Workshop), http://wp.doc.ic.ac.uk/thefutureofslam/wp-content/uploads/sites/93/2015/12/slides_ajd.pdf

[17] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera," In Iccv, vol. 3, 2003, pp. 1403-1410.

[18] P. Newman and Ho, "SLAM-loop closing with visually salient features," In Proceedings of the 2005 IEEE International Conference on Robotics and Automation (pp. 635-642). IEEE.

[19] J. Fuentes-Pacheco, J. Ruiz-Ascencio, and J. M. Rendón-Mancha, "Visual simultaneous localization and mapping: a survey," Artificial Intelligence Review 43.1 (2015): 55-81.

[20] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual SLAM algorithms: A survey from 2010 to 2016," IPSJ Transactions on Computer Vision and Applications 9.1, 2017, 16.

[21] G. Bresson, Z. Alsayed, L. Yu, and S. Glaser, "Simultaneous localization and mapping: A survey of current trends in autonomous driving," IEEE Transactions on Intelligent Vehicles, 2(3), 2017,pp.194-220.

[22] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. (Vol. 1, pp. I-I). IEEE.

[23] D. Scaramuzza and F. Fraundorfer, "Visual odometry [tutorial]," IEEE robotics & automation magazine, 18(4), 2011, pp.80-92.

[24] B. Joshi, S. Rahman, M. Kalaitzakis, B. Cain, J. Johnson, M. Xanthidis, , N. Karapetyan, A. Hernandez, A.Q. Li, N. Vitzilaios, and I. Rekleitis, "Experimental Comparison of Open Source Visual-Inertial-Based State Estimation Algorithms in the Underwater Domain," 2019, arXiv preprint arXiv:1904.02215.

[25] C. Chen, H. Zhu, M. Li, and S. You, "A Review of Visual-Inertial Simultaneous Localization and Mapping from Filtering-Based and Optimization-Based Perspectives," Robotics, 7(3), 2018, p.45.

[26] A. Fusiello, R. Giannitrapani, V. Isaia, and V. Murino, "Virtual environment modeling by integrated optical and acoustic sensing. In Second International Conference on 3-D Digital Imaging and Modeling," (Cat. No. PR00062), 1999, pp. 437-446. IEEE.

[27] U. Castellani, A. Fusiello, V. Murino, L. Papaleo, E. Puppo, and M. Pittore, "A complete system for on-line 3D modelling from acoustic images," Signal Processing: Image Communication, 20(9-10), 2005, pp.832-852.

[28] C.N. Roman, "Self consistent bathymetric mapping from robotic vehicles in the deep ocean," Doctoral dissertation, Massachusetts Institute of Technology, 2005.

[29] C.M. Clark, C.S. Olstad, K. Buhagiar, and T. Gambin, " Archaeology via underwater robots: Mapping and localization within maltese cistern systems," in 2008 10th International Conference on Control, Automation, Robotics and Vision, 2008, pp. 662-667. IEEE.

[30] D. Ribas, P. Ridao, J.D. Tardós, and J. Neira, "Underwater SLAM in a marina environment," In 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, 2007, pp. 1455-1460. IEEE

[31] F. Guth, L. Silveira, S. Botelho, P. Drews, and P. Ballester, "Underwater SLAM: Challenges, state of the art, algorithms and a new biologically-inspired approach," in 5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics, 2014, pp. 981-986. IEEE.

[32] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," IEEE Journal of Oceanic Engineering, 39(1), 2014, pp.131-149.

[33] J. Sturm, N. Engelhard, F. Endres, , W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012, pp. 573-580. IEEE.

[34] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," in The International Journal of Robotics Research, 32(11), 2013, pp.1231-1237.

[35] R. Mur-Artal and J.D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," IEEE Transactions on Robotics, 33(5), 2017, pp.1255-1262.

[36] R. Mur-Artal, J.M.M. Montiel, and J.D. Tardos, "ORB-SLAM: a versatile and accurate monocular SLAM system," IEEE Transactions on Robotics, 31(5), 2015, pp.1147-1163.

[37] X. Yuan, J.F. Martínez-Ortega, J.A.S. Fernández, and M. Eckert, "Aekf-slam: a new algorithm for robotic underwater navigation," *Sensors*, 17(5), 2017, p.1174.

[38] A.Q. Li, A. Coskun, S.M. Doherty, S. Ghasemlou, A.S. Jagtap, M. Modasshir, S. Rahman, A. Singh, M. Xanthidis, J.M. O'Kane, and I. Rekleitis, "Experimental comparison of open source vision-based state estimation algorithms," in International Symposium on Experimental Robotics, 2016, pp. 775-786. Springer, Cham.

[39] M. Prats, J. Perez, J.J. Fernández, and P.J. Sanz, "An open source tool for simulation and supervision of underwater intervention missions," in IEEE/RSJ international conference on Intelligent Robots and Systems, 2012, pp. 2577-2582. IEEE.

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

[41] M. Nawaf, D. Merad, J.P. Royer, J.M. Boï, M. Saccone, M. Ben Ellefi, and P. Drap, "Fast Visual Odometry for a Low-Cost Underwater Embedded Stereo System," Sensors, 18(7), 2018, p.2313.