

Alma Mater Studiorum - Università di Bologna

DOTTORATO DI RICERCA IN
INGEGNERIA CIVILE, CHIMICA, AMBIENTALE E DEI MATERIALI

Ciclo 35

Settore Concorsuale: 08/A4 - GEOMATICA

Settore Scientifico Disciplinare: ICAR/06 - TOPOGRAFIA E CARTOGRAFIA

FIRST STEPS IN UNDERWATER PHOTOGRAMMETRY WITH THE SUPPORT OF
AN UNMANNED UNDERWATER VEHICLE (UUV)

Presentata da: Mattia Carello

Coordinatore Dottorato

Alessandro Tugnoli

Supervisore

Luca Vittuari

Co-supervisore

Alessandro Lambertini

Esame finale anno 2023

Abstract

The evolution of modern and increasingly sensitive image sensors, the increasingly compact design of the cameras, and the recent emergence of low-cost cameras allowed the Underwater Photogrammetry to become an infallible and irreplaceable technique used to estimate the structure (3D coordinates of scene points) of the seabed with high accuracy. On the other hand, the innumerable processes related to ocean ecosystem are still poorly understood as well as vast areas of the seabed remain largely unexplored. In order to fill this gap, the European Union has initiated and is initiating numerous research programmes to combine sustainable development goals with scientific needs. Within this context, the main topic of this work is the Underwater Photogrammetry from a geomatic point of view and all the issues associated with its implementation, in particular with the support of Unmanned Underwater Vehicle (UUV). Questions such as: how does the technique work, what is needed to deal with a proper survey, what tools are available to apply this technique, and how to resolve uncertainties in measurement will be the subject of this thesis. The study conducted can be divided into two major parts: one devoted to a series of ad-hoc surveys and tests, thus a practical part, another supported by the bibliographical research, which thus represents the theoretical one. However the main contributions are related to the experimental section, in which two practical case studies are carried out in order to improve the quality of the underwater survey of some calibration platforms. The results obtained from these two experiments showed that, for sure the refractive effects due to water and underwater housing can be compensated by the distortion coefficients in the camera model, but if the aim is to achieve high accuracy then a model that takes into account the configuration of the underwater housing, based on ray tracing, must also be coupled. The major contributions that this work brought are: an overview of the practical issues when performing surveys exploiting an UUV prototype, a method to reach a reliable accuracy in the 3D reconstructions without the use of an underwater local geodetic network, a guide for who addresses underwater photogrammetry topics for the first time, and the use of open-source environments such as OpenCV, Octave and the Pinax algorithm implementation.

Acknowledgements

I would like to express my sympathy and admiration to:

- My supervisor who allowed me to complete this work with great clarity and freedom
- Erica Nocerino and Fabio Menna, who are luminaries in the field of Underwater Photogrammetry and they helped me in the calibration platform design
- My colleagues, in particular Enrica, Anna and Giulia for their patience
- My family and friends that support me during these three years
- The Interregional European funding for two projects, Sushidrop and Techera, which gave me access to Blucy's equipment

The sea, once it casts its spell, holds one in its nets of wonder forever.

Jacques Yves Cousteau

Contents

Abstract	i
Acknowledgements	iii
1 Introduction	1
1.1 Outline	4
2 Terrestrial survey of calibration platforms	5
2.1 Introduction	5
2.2 Platforms Design	6
2.2.1 Calibration grid design	7
2.2.2 Calibration frame design	9
2.3 Geomatic techniques for calibration platforms survey	11
2.3.1 Total station and photogrammetry	13
2.3.2 Structured-Light projection Scanning	15
2.3.3 Terrestrial Laser Scanner	16
2.4 Comparison techniques	18

2.4.1	Calibration grid	20
2.4.2	Calibration frame	20
3	Underwater Photogrammetry	26
3.1	Introduction	26
3.2	Water physics	29
3.3	Geometry	35
3.3.1	Dome port	38
3.3.2	Flat port	40
3.3.3	Calibration algorithms	43
3.4	Colorimetry	50
4	Unmanned Underwater Vehicles (UUVs)	57
4.1	Introduction	57
4.2	Overview	60
4.3	The case of Blucy	76
4.3.1	Positioning instruments	82
4.3.2	Surveying instruments	91
5	Structure from Motion technique	97
5.1	Introduction	97
5.1.1	Projective Geometry	98
5.1.2	Camera Models	100

5.2 Theory	105
5.2.1 Epipolar Geometry and Triangulation	106
5.2.2 Depth and Depth Maps	108
5.2.3 Bundle Adjustment	109
5.2.4 Camera Pose Estimation	111
5.3 Challenges in the underwater environment	113
6 Underwater experience	117
6.1 Introduction	117
6.2 Underwater surveys context	119
6.2.1 Calibration grid	119
6.2.2 Calibration frame	120
6.3 Experiment setup	125
6.3.1 Calibration grid	126
6.3.2 Calibration frame	136
6.4 Analysis of the results	140
6.5 Several applications of color correction algorithms	153
7 Conclusion	156
Bibliography	157

List of Tables

2.1 Average Standard Deviation (1 Sigma) taken from STAR*NET Adjustment of Terrestrial Survey and Total Error Factor.	14
2.2 Error in meters of the control points in the photogrammetric survey.	15
2.3 Qualitative features of the two measurement techniques.	18
4.1 Outline of the main internal sensors usually equipped on a AUV belonging to the proprioceptive navigation category.	68
4.2 Outline of the main external sensors usually equipped on a AUV belonging to the geophysical navigation category.	69
4.3 Web-based resources on UUV. Accessed: 2022-07-18.	77
4.4 Blucy's general technical characteristics	80
4.5 Instruments for positioning of Blucy	84
4.6 iXBlue Phins Compact C3 performances.	86
4.7 Trittech PA-500 characteristics.	87
4.8 GNSS Aided Inertial Navigation System main specifications.	88
4.9 Instruments for surveying of Blucy	92
4.10 Sonic 2020 multi beam system specifications.	94

4.11 MiniCT and MiniSVS technical specifications. 94

6.1 Residuals of the errors associated with the control points in the case of uncorrected images(left) and corrected images(right). In the first column there's the target Id. There is target Id 16, however the target number is 15, because target number 7 was affected by a misprint. 134

6.2 Parameters workflow used in Agisoft Metashape. 142

6.3 Standard deviations values used as input of the Monte Carlo method expressed in meters. They are obtained as mean values between those given by the underwater experiment and those found in literature. 151

6.4 Properties of the underwater calibration frame experiment. It is reported the mean value (μ) for each spatial coordinates, as well as the standard deviation (σ), the maximum value and minimum one. UW stands for underwater. P stands for Pinax. 152

6.5 Properties of the Monte Carlo simulations. It is reported the average of the mean values (μ) for each spatial coordinates, as well as the average of the standard deviations (σ), the average of maximum values and minimum ones. UW stands for underwater. P stands for Pinax. 153

List of Figures

1.1	UUV prototype, called Blucy, components and subsystems in the context of the research project called SUSHIDROP.	3
2.1	The newly assembled handmade calibration grid following the information on the sheet placed at its center. Attention should be paid to how the joints between the various metal elements were made and how the targets were attached.	7
2.2	The drawing of the handmade calibration frame subsequently printed on Dibond™ material without the targets at altitude.	9
2.3	The calibration frame after printed on Dibond™. It is possible to notice that there are also two plates placed at different heights with respect to the calibration frame.	10
2.4	Available geomatics techniques, sensors, and platforms for 3D recording purposes, according to the scene' dimensions and complexity. A color legend is provided to show the average accuracy of the techniques.	11
2.5	Geodetic network established for the survey of the calibration frame. T1 and T2 stationing points are used to survey the same 3D points, while T3 is used only as a foothold.	12
2.6	Schematic view of the geomatic network provided by STAR*NET.	13
2.7	Textured point cloud provided from the standard SfM pipeline in Agisoft Metashape.	15

2.8 An operator while conducting the calibration grid survey with an empty tank. It can be seen that the tank is below ground level, which is why this instrument was chosen to carry out the survey. It is also possible to identify the use of a laptop connected to the instrument that can display the data acquired by the instrument in real time. 19

2.9 Matrices of relative distance differences expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles. 24

2.10 Frequency histograms of the difference of the coordinates between Total Station and both the techniques used for the survey. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value. For clarity, along the axis the order of magnitude is millimeters. 25

3.1 Different technological approaches suitable at different spatial scales, depths, and 3D-complexity of the investigated area for underwater surveying. Source: [1]. . . 26

3.2 A schematic view of the effect of the sea’s selective filter applied to the visible wavelengths of the light. 30

3.3 Effect of the artificial light’s configuration on underwater pictures. It is easy to note how a proper light configuration improves the quality of the images. Correct configuration at the bottom of the figure. 32

3.4 Behaviour of the ray of light meeting different surface's shapes. a) the case of a flat port: when the rays travel from a slow medium to a fast one, they are bent towards the normal direction and the image's symmetry is lost due to the change in direction b) the case of a dome port: the rays orthogonally meet the surface preserving the image's symmetry even if the two mediums are different. 36

3.5 The example of the broken pencil. A pencil is partially submerged into a glass of water to demonstrate the change in direction due to the Snell's law. 37

3.6 Schematic representation of refractive effects in the image formation for a dome port. Red line indicates the light path in air or the collinearity air line. Blue line indicates the light path in water according to the Snell's law. P is the submerged point projected on the sensor at the distance \bar{r}_i . $\bar{\Delta r}$ is function of the distance \bar{Z} and it represents the quantity projected on the sensor that differs for the in-air case and the in-water one. c is the principal distance. O is the entrance pupil. D is the center of the dome. \overline{DS} is the dome port radius. When $O \neq D$ we deal with decentered dome. The thickness of the glass flat port is not considered in this model. 38

3.7 Graphical demonstration of the principle of virtual image's formation when using a dome port. 40

3.8 Schematic representation of refractive effects in the image formation for a flat port. Red line indicates the light path in air or the collinearity air line. Blue line indicates the light path in water according to Snell's law. P is the submerged point projected on the sensor at the distance \bar{r}_i . $\bar{\Delta r}$ is function of the distance \bar{Z} and it represents the quantity projected on the sensor that differs from the in air case to the in water one. c is the principal distance. O is the entrance pupil. $\overline{OS'}$ is the flat port distance from the entrance pupil. It is not considered in this model the thickness of the glass flat port. 41

3.9 Ray tracing in a multi-media configuration:water, glass and air. Oblique dotted lines indicate direction when no refraction effect occurs. Thickness of layers is only indicative. 43

3.10 Schematic visualization of a Flat camera system. \mathbf{n} is the normal vector to the glass interface. The blue line represents the physically accurate light ray; the green line is the apparent ray traced back to the camera’s optical axis. The focus distance x for each light ray is computed as $x = \tan\delta(d_0\tan\alpha + d_1\tan\beta)$. Figure extracted from [2]. 48

3.11 The three components of underwater optical imaging: direct component (straight line), forward component (dashed line) and backward scatter component (dash-dot line). 53

3.12 Jerlov water types. Each type is indicated with a colour and those types that are coastal are distinguished by 'C' letter. Gray lines represent 280 randomly chosen observations from the seaBASS database, which lay within the Jerlov’s water types space. On the right is showed color table represents the Jerlov’s water types. 54

4.1 An ancient DPV system next to a recent one. 57

4.2 The submarine REMORA 2000 of COMEX during the Grand Ribaud excavation, 2000. Photo Frederic Bassemayousse. 58

4.3 Three examples of UUV that work at different depth and with different aims. . . 59

4.4 Implementing the latest technology is one of the pillars of the GEBCO Seabed 2030 project. The concept of an unmanned mapping barge, monitoring by satellite communication and equipped with an ultra-narrow beam deep-water multi-beam (left), is just one of many ways technology could be used to generate new data. The sub-meter of detail sometimes needed to investigate small scale processes at the seabed is today only possible to achieve in the deep ocean using UUVs equipped with high-resolution high frequency multibeam systems. 60

4.5 Two navigation modes: (a) ROV and (b) AUV. 66

4.6 Outline of underwater navigation classification. It may happens that some methods are combined to provide increased performance, which is the case of multi-sensory information fusion. 70

4.7 (a) USBL and (b) LBL acoustic navigation systems. 71

4.8 The three types of real-time maps produced during SLAM activities. 73

4.9 Main Blucy components: (a) Wi-Fi communication. (b) USBL transponder. (c) Radio communication. (d) Hook. (e) AHRS: GNSS and INS. (f) Thruster: vertical. (g) Main buoyancy foam. (h) Lateral buoyancy foam panel. (i) HDPE structure. (j) Thruster: lateral. (k) Fiber optical cable. (l) Thruster: longitudinal. (m) Altimeter. (n) Frontal LED lights. (o) PilotCAM. (p) Adjustable LED lights. (q) MBES sonar head. (r) MiniSVS. (s) MiniCT. (t) BottomCAM. (u) MBES computer. (v) Bottom LED lights. (w) Navigation, Guidance and Control (NGC) computer. (x) UUV 24V Battery. (y) FOG. (z) DVL. As described in [\[3\]](#). 79

4.10 Blucy’s functional architecture. Main modules are indicated with the corresponding arrow: remote station (located on the vessel or on the shore), surface communication and navigation modules (communication with the remote station via Wi-fi or radio modem and navigation based on GNSS), propulsion modules (8 thrusters), underwater communication modules, navigation payload modules (acoustic altimeter, fiber-optic-gyro, doppler velocity log, pressure sensor and the USBL modem), powering module (a double lithium battery 24V 50 Ah), sensor modules (conductivity e temperature sensor, multi-beam), navigation guidance and control module (the main canister equipped with an huge number of connectors that allows to communicate with the other parts of the vehicle, it is provided with: i)a CPU for managing the system and the navigation, driving and control system ii)a CPU for mission control management iii)the servo-controllers of the motors iv)the multiplexers for data/video transmission on the optical fiber cable v)conversion (DC/DC) and distribution components, including relays for enabling/disabling, of the energy to the other modules), and auxiliary systems (cameras and LED lights). In addition it is highlighted the tether cable made of optical fiber. 83

4.11 USBL positioning system, in ROV and AUV mode. On top of the figure it is possible to note the connections with the instruments placed on the vessel. Data coming from transceiver and vessel instruments are processed within the navigation computer to provide a precise positioning. 85

4.12 Setup for USBL positioning on surface vessel: (a) GNSS antennas and INS platform. (b) Structure for USBL. 86

4.13 MQTT communication scheme. 89

4.14 Transect performed by UUV partially above and partially under water surface. . 90

4.15 Blucy prototype and fiber optical cable. 91

4.16 Main surveying instruments: (a) BottomCAM and its network data transmission router before insertion into the waterproof case. (b) BottomCAM 15 mm frontal glass. (c) MBES sonar head (top) and final BottomCAM canister (bottom) installed in UUV.	94
4.17 Dense 3D point cloud of seabed processed from BottomCAM high-resolution imagery with SFM techniques.	96
5.1 Incremental Structure-from-Motion pipeline. Source: [4].	98
5.2 Central projection. Source: [5].	99
5.3 Perspective transformations. Source: [5].	100
5.4 Simplified model of the camera's image acquisition process. A ray from the scene enters the camera through the pinhole or center of projection, crossing the optical axis. After that, it is recorded in the image at the point of intersection with the image plane. Source: [6].	101
5.5 Simplified camera intrinsics. Source: [5].	103
5.6 Different types of radial lens distortion compared with the case of absence of distortion. Source: [7].	104
5.7 Epipolar geometry between two views. (a) The two cameras are indicated by their centres C and C' and image planes. The camera centres, point X and its images x and x' lie in a common plane π . (b) An image point x back-projects to a ray in 3-space defined by the first camera centre C and x . This ray is imaged as a line l' in the second view. Source: [5].	106
5.8 Misaligned image pairs. The measured points do not satisfy the epipolar constraint. Since the rays do not intersect, x' does not lie on l' , as well as x does not lie on l . Source: [5].	107

5.9	Depth map of an observed object. Notice the single channel color scale which gives an idea of the depth. Source: [5].	108
5.10	Bundles of light rays for multiple cameras observing the same scene. Source: [5].	110
5.11	GSD relationship illustration.	113
5.12	Illustration of the relationship between motion speed and shutter time. Inspired by [8].	115
5.13	The photographic exposure triangle. Source: [8].	116
6.1	European Interreg research projects.	117
6.2	Two snapshots taken during the experiment conducted in the water tank of the Department of Hydraulics of the Faculty of Engineering, University of Bologna. Notice the presence of a traverse, which was a fairly significant obstacle during the survey. In (a) there is a zoom of the camera system raised through the use of some ropes tied to an appropriate hook integral with an overhead crane. In (b) a view of the entire system where it is possible to observe the connecting cables with the battery and the control panel.	121
6.3	Map of the area in front of Miramare marine protected area in the Gulf of Trieste. Red bullets indicate the three points of interest of the survey mission. The calibration frame survey was conducted in the vicinity of 'C1 Buoy', in front of the Miramare castle.	121
6.4	View of MPA Miramare, where you can see the vessel used during the mission docked at 'C1 buoy' while Miramare Castle appears in the background.	122
6.5	Vessel used to carry out the survey mission in Miramare MPA moored in the harbor. Also the presence of mechanical arm useful for loading and unloading operations of Blucy can be seen.	123

6.6 Two photos taken during the survey mission conducted in the water of AMP Miramare. In (a) it is showed Blucy during its activity on the surface of water. Pay attention to the fiber optic cable that connects Blucy with the control panel and the fact that there are floats integral to the cable that prevent drift effects. Also present is a buoy to signal the presence of Blucy. In (b) an underwater photo taken by a diver while Blucy takes photos of calibration frame. Due to poor visibility of the water, the rectangular-shaped frame appears highly disturbed but still visible. In addition, it is possible to observe the artificial light system needed to illuminate at great depths, where sunlight is faintly present. 124

6.7 Scheme of the steps carried out to validate the calibration and correction flow. The flow is split in three parts,each of which is interconnected. Pay attention to different colors, they are associated to different elements. 127

6.8 Camera system, leaning on top of a stand, during the in air calibration. A chessboard was placed on a rigid plane and the size of the alternating squares was given as input for the calibration algorithm. 128

6.9 Determination of the refractive index of the underwater hosing flat glass port. In (a) is showed the experiment setup and a schematic view of the experiment's mechanism. The deflection that the laser beam undergoes due to the abrupt change in refractive index is $\Delta x = d\{tan(\theta) - tan[asin(\frac{sin(\theta)}{N_w})]\}$ where d is the thickness of the glass port. In (b) it is report the experimental data fit, where $p0$ and $p1$ are respectively d and N_w , which is the unknown parameter. Its computed value is $N_w = 1.60 \pm 0.04$. Thanks to Sara Rabaglia which makes available the experiment tools. 132

6.10 Tree bash command called to display the standard structure of a working directory. For the sake of simplicity, data set consists of only three images. The correction map is split in X and Y spatial coordinates and they are represented by .txt files. 134

6.11 Matrices of relative distance differences among SLS target coordinates and underwater photogrammetric extracted target coordinates expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles. . . . 135

6.12 Frequency histograms of the difference of the center target coordinates between those obtained by means of SLS and those estimated with photogrammetric techniques. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value. 137

6.13 Image shows control and check point locations and error estimates. Z error is represented by ellipse color. X,Y errors are represented by ellipse shape. Estimated locations are marked with a dot for control points or T-crossing for check points. 138

6.14 First pictures obtained by means of Blucy. In (a) there is the snapshot of the calibration frame. In (b) the same thing except for the corrected distortions that are strongly removed as you can see from the perfect rectangular shape of the platform. It has to be noticed the two resolution bar place on two adjacent sides of the platform. They are designed starting from the smallest GSD considered, please refer to 2.2.2 section for more details. During the survey they are used to stabilize the focus of the camera. 140

6.15	Matrices of relative distance differences among TS target coordinates and underwater photogrammetric extracted target coordinates expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles. It has to be noticed that targets at altitude are those with a target id greater than 125, to which correspond higher measurement errors.	141
6.16	Data processing for underwater images. GCPs selected as ground truth are those measured with TS.	141
6.17	Textured point cloud provided from the standard SfM pipeline in Agisoft Metashape that represents the calibration frame placed underwater during the Techera surveying mission.	143
6.18	Camera positions. An average flying altitude of 2.33 m is computed.	143
6.19	GCP locations and error estimates. Z error is represented by ellipse color. X and Y errors are represented by ellipse shape. Estimated GCP locations are marked with a dot or crossing. Case of underwater images corrected with Pinax.	144
6.20	Cloud-to-cloud distance computed with ICP algorithm. We use as reference the terrestrial point cloud computed in the terrestrial surveys obtained by means of the standard photogrammetric technique. Values associated to color scale are reported in metres.	146
6.21	Frequency histograms of the cloud to cloud distance with the associated Gauss curve. Mean and standard deviation provided as subtitles.	146

6.22 Frequency histograms of the difference of the center target coordinates between those obtained by means of SLS and those estimated with underwater photogrammetric techniques. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value. 149

6.23 Computation of convergence coefficients as the number of simulations increases. (a) is related to a single case extracted randomly within the thousand distributions. (b) is related to all cases of y coordinate without using Pinax algorithm. (c) is related to all cases of z coordinate using Pinax algorithm. 152

6.24 Application of ACE, CLAHE, LAB, and SP color correction algorithm obtained with standard parameters. 154

Chapter 1

Introduction

The hostility of the marine environment, regardless of the huge improvements in the recent years, remains a major challenge to overcome and often an extremely costly mission. The water medium with its characteristics makes the environment difficult to manage either in terms of operability and in terms of potential risks. Some of these dangers are removed using remotely piloted devices, also called Unmanned Underwater Vehicles (UUVs), which in turn can be divided between Autonomous Underwater Vehicles (AUVs) and Remotely Underwater Vehicles (ROVs). Thanks to the role played by the UUV devices, together with the efforts in the past and current researches of the underwater scientific community, this challenge is going through a very florid phase, in which also the public interest becomes a fundamental driver for the development and growth in this field. In 1968, the whole world was looking up: the moon landing was one year away and it seemed that the future would be in the space. Today, however, our attention is turned rather downwards: the changing climate of our planet. In response to this problem, the European Union has outlined approaches and rules to lessen its impact globally and prevent the Earth's temperature from rising above 2°C, the insurmountable limit sanctioned by the Intergovernmental Panel on Climate Change IPCC, or rather limit warming to 1.5°C relative to the pre-industrial period (<https://www.ipcc.ch/sr15/>). Through two strategies, the European Union is trying to address the environmental needs of the marine ecosystem without forgetting the role played by fishing and resources exploitation. Blue

economy and blue growth represent the two main concepts on which the adopted policies are based, which are also aimed at enhancing the value of the underwater heritage. Recently (2014) some researchers (ITCG-international Thwaites glacier collaboration) have shown that one of the largest glaciers on our planet (Thwaites) is melting, causing an irreversible rise in the oceans [9]. In this context, the use of UUVs has made it possible and still makes it possible to understand how much and how fast this melting will happen. As well as in enhancing the heritage hidden beneath our waters and increasing common interest in its exploration, fostering the process of education and respect toward the heritage itself. Therefore, the underwater exploration provides answers to our questions, not only related to climate change and cultural heritage, but also related to biology, oceanography, and biosphere physics in general. Further, hidden under our eyes, there are countless remnants of our past waiting to be rediscovered and catalogued to be made available for use by the entire community (as testified by the Convention on the Protection of the Underwater Cultural Heritage-United Nations Educational, Scientific and Cultural Organization-UNESCO 2001), including those who could not directly access them (virtual reality), such as people with physical disabilities. Nevertheless, the ocean ecosystem and its innumerable processes are still poorly understood, as vast areas of the seabed remain largely unexplored. For this reason any exploration campaign, experiment, experience, survey conducted underwater acquire importance as it enhances this field of study, which has yet to evolve and reach the peak of its expression. Within this context, the underwater photogrammetry is placed. More precisely, it allowed the birth of optical-based 3D reconstruction methods useful for restoring an accurate 3D representation of smaller areas of interest with high resolution [10]. The evolution of modern and increasingly sensitive image sensors, the increasingly compact design of the cameras, and the recent emergence of low-cost cameras allowed the Underwater Photogrammetry to become the main technique used to estimate the structure (3D coordinates of scene points) with high accuracy. Therefore, the main topic of this work is the Underwater Photogrammetry from a geomatic point of view and all the issues associated with its implementation. Questions such as: how does the technique work, what is needed to deal with a proper survey, what tools are available to apply this technique, and how to resolve uncertainties in measurement will be the subject of this discussion. The study conducted can

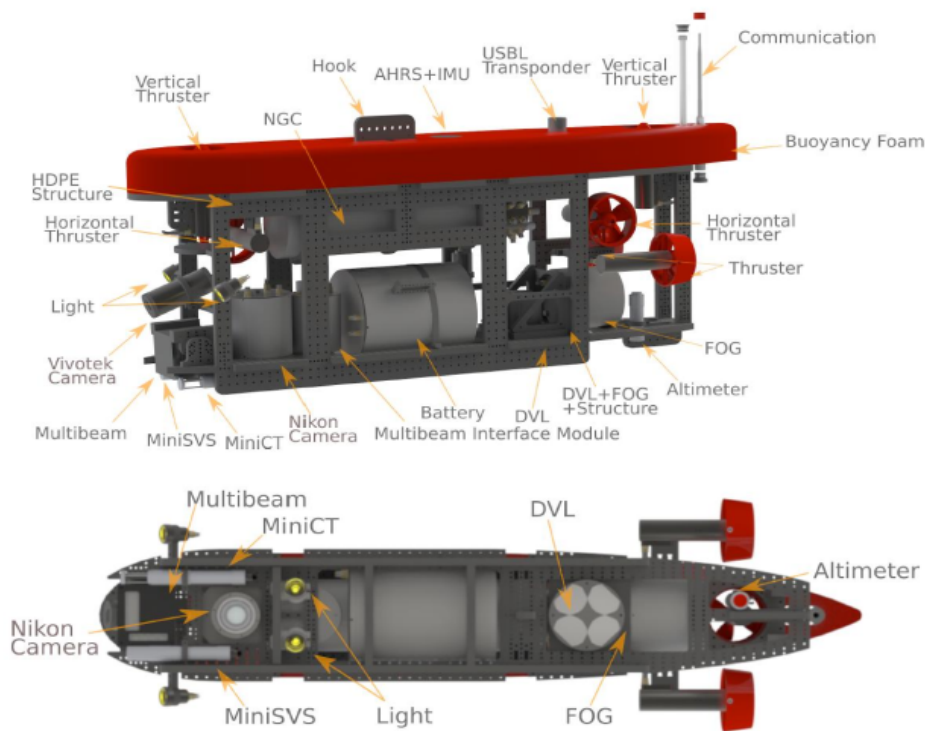


Figure 1.1: UUV prototype, called Blucy, components and subsystems in the context of the research project called SUSHIDROP.

be divided into two major parts: one devoted to a series of ad-hoc surveys and tests, thus a practical part, another supported by the bibliographical research, which thus represents the theoretical one. For the sake of brevity, the theoretical part will neglect some basic concepts to leave space to fundamental factors which most affect the application of the technique. For what concerns the practical part, the backbone is the use of free of charge and open source tools, in a context of low budget activities and home made test easily reproducible. These practical experiences were carried out thanks to the participation in two European research projects which enabled the complex procedure of obtaining underwater data to be implemented. SUsustainable fiSHeries wIth DROnes data Processing (SUSHIDROP) is the main actor in this context; the project concerned the building of a prototype of an hybrid UUV (Figure 1.1) from scratch, while maintaining low manufacturing costs and at the same time the same performance as a high-precision commercial device. TechERA, on the other hand, began as an European project to capitalise on and enhance the previous project, enabling underwater drone campaigns in the Adriatic Sea.

1.1 Outline

This document is structured into the following chapters:

- Chapter 2: the instruments through which underwater photogrammetry studies will be carried out are introduced. In particular, the focus is on how to carry out the survey with these instruments, also comparing different available geomatic techniques.
- Chapter 3: the subject of underwater photogrammetry is introduced at a general level, starting with a small historical excursus of the technique. The challenges this technique has to overcome due to the presence of water are noted. In addition, the geometric problems to be faced and in particular the issue of camera calibration are introduced, as well as a brief discussion of color correction methods.
- Chapter 4: it provides an overview of the use of UUVs in science and then focuses on Blucy (Figure [1.1](#)) and its technical specifications
- Chapter 5: the general theory of the 3D reconstruction starting from 2D images, that is Structure from Motion (SfM), is introduced. Some basic concepts of SfM are faced, as well as its main challenges in the underwater field.
- Chapter 5: it describes the set-up of the experiments conducted to analyze the results. It contains a summary of the results obtained in the context of this thesis. Finally, the application of certain color correction methods is provided.
- Chapter 6: the last chapter presents the conclusions and some guidelines for future work.

Chapter 2

Terrestrial survey of calibration platforms

2.1 Introduction

The first element that has been carried out in order to have a control on the camera calibration process is that to build some calibration platforms, therefore the first topic that will be faced is the realisation and surveying of calibration platforms. This operation was done in steps. For logistical reasons, it started with a first calibration platform carried out in a simpler manner with coded target which had stability issues during the displacement of the calibration platform, consider, for example, when the calibration platform has to be moved from one place to another or when it has to pass through narrow places. After, a second calibration platform was built, by looking for to correct the previous errors. Therefore, in the description it will follow the same temporal steps, from the calibration grid to the calibration frame. This chapter will deal with the application of several geomatic techniques used for the survey of some calibration platforms. Among these techniques there exist some already consolidated and therefore considered by the scientific community as reliable techniques, while others are fruit of the recent technological innovations that need to be taken into consideration since they usually provide optimal results in short time. In this regard, a summary description of the techniques chosen for the survey

will be provided, as well as it will show what the design criteria of the calibration platforms were. Finally, a comparison of the techniques used in terms of accuracy and operability will be carry out so as to provide a guide that can support operators who will be in the same survey condition. Focus has to be paid to survey's place. As the title of the chapter also says, here we consider the survey of the calibration platforms when they are above the sea level, therefore the presence of the water is neglected and you can use a large number of techniques since there are not many limitations here, which you would have underwater. Any object with encoded or appropriately design targets can be considered as a calibration platform; here we consider two object examples used for calibration and removal of refractive effects of an underwater camera equipped on an Underwater Unmanned Vehicle (UUV).

2.2 Platforms Design

As already mentioned, a calibration platform is an object which coded targets are firmly attached. The midpoint of these targets is a point of the 3D space that has to be surveyed by appropriate available geomatic techniques. The key point in the calibration platform design is the rigidity. The more is rigid the platform, the more reliable the platform is. A second aspect that has to be taken into account is its portability. Indeed, it needs to imagine that these platforms will have to be moved to the place where the survey is to be carried out, so they will have to be strong enough and rigid to withstand small bumps, which could change the spatial arrangement of the target centres, but at the same time they have to be small enough to be moved inside a car or a van for example. As far as the dimension of calibration platforms a trade-off between portability and visibility has to be found. To solve the portability problem, one could imagine designing a demountable platform, but this choice is incorrect since in the disassembly-assembly stages the error in target placement naturally increases. However this choice could be correct if a calibration platform is used only for one time. It was decide to design two calibration platforms; actually the design process followed an evolution from a calibration grid to the construction of a calibration frame. In this fashion, starting from a preliminary idea of the calibration platform, that is the calibration grid, it was possible to address the main

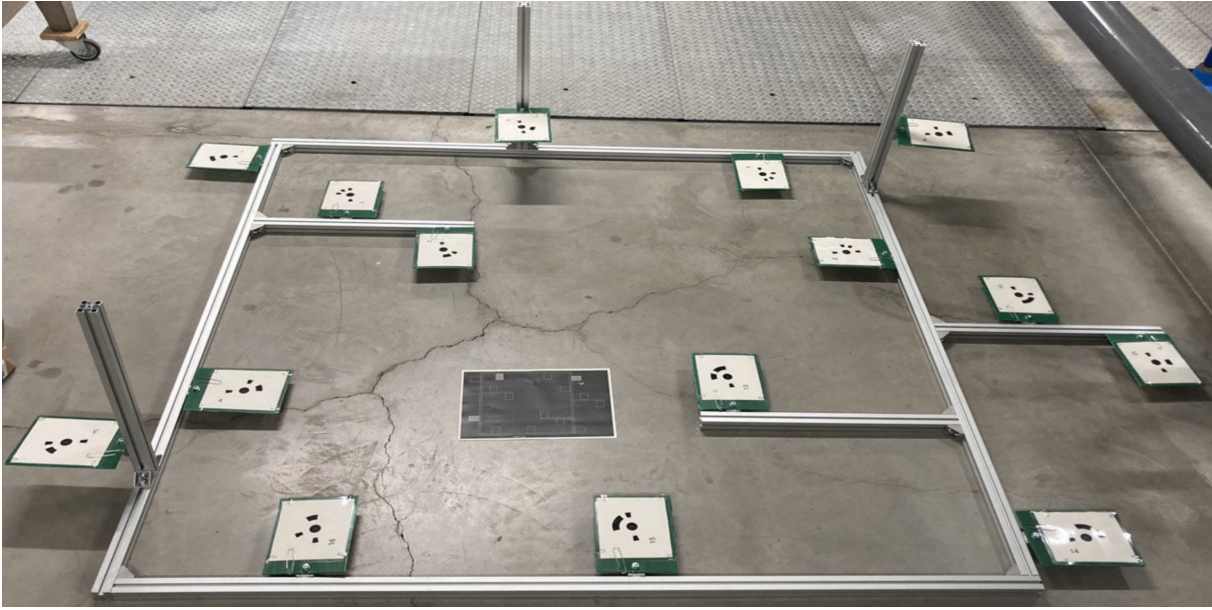


Figure 2.1: The newly assembled handmade calibration grid following the information on the sheet placed at its center. Attention should be paid to how the joints between the various metal elements were made and how the targets were attached.

issues associated with surveying a calibration platform so that the same errors would not be repeated in the final version of the calibration frame. So in this context, the calibration grid can be considered as the primordial project of the calibration platform while the calibration frame the final one. In addition, an important role is played by the cost in manufacturing the calibration platform, obviously when the cost of investment increases the reliability of the platform increases at the expense of the cost-effectiveness of the platform. For this reason, some hand-made solutions are found using products that come from the mass market.

2.2.1 Calibration grid design

The calibration grid is showed in Figure [2.1](#). It was realised in square aluminum bars with a cross section of 3 cm. The main square profile consists of 4 aluminum bars that are 1.5 m long, while the elements that protrude from the profile (6) are exactly 0.5 m long; therefore its area is 2.25 m^2 . There are 15 circular 20 bit coded targets, three of which are mounted at different heights from the plane of the calibration grid in order to provide centre coordinates of the targets that belong to different planes from the main one; in this fashion extracting the target centres becomes more accurate since the addition of new 3D planes improve the robustness of

the computation; for further details check this paper [11] where it was proven that a stratified distribution of the targets is essential to optimize the vertical accuracy. Circular coded targets are currently the standard type of targets applied in photogrammetric calibration [12]; they are centripetal encoding where unique sections of a disk surrounding a central circular target are used to encode the target. The target's radius is set to $1 \text{ cm} \approx 2 \times 10 \times GSD$, where GSD is the Ground Sample Distance that is computed as

$$GSD = \frac{\mu \times d}{f}$$

μ is the pixel size in meters, while d is the distance between object and camera in meters and finally f is the focal length of the camera in meters, in this case, multiply by a factor equals to 1.333 (water refractive index with respect to air) to account for refractive effects introduced by water. Two works exist that examine this effect more closely: [13] and [14]; both conclude that refraction is compensated for multiplying the focal length measured in air by the index of refraction for water $1.333f_a = f_w$. In this case it was considered $d \approx 3 \text{ m}$, $\mu = 5.956 \text{ } \mu\text{m}$ and $f = 24 \text{ mm}$. The targets themselves are printed in simple rectangular paper sheets of the dimension of $14 \times 12 \text{ cm}^2$, which were subsequently plasticized to avoid being damaged by the presence of water. Since in the water the stability of the plasticized target is compromised, a thick sheet of rigid plastic (of green color in Figure 2.1) is used as a table top and several zinc plated paper clips are employed to anchor the target to the plastic support. In addition, a hole is drilled on the plastic holder to allow its attachment to the calibration grid. The connections between the metal rods and those between them and the plastic supports are made with metal brackets equipped with bolts that make for a very rigid bond even underwater. As mentioned before, this is a preliminary version of the calibration platform therefore, after careful analysis of the results obtained during tests, the following critical issues were considered:

- lack of a color correction device
- increase the number of targets
- choose a water-resistant material with which to carry out the targets

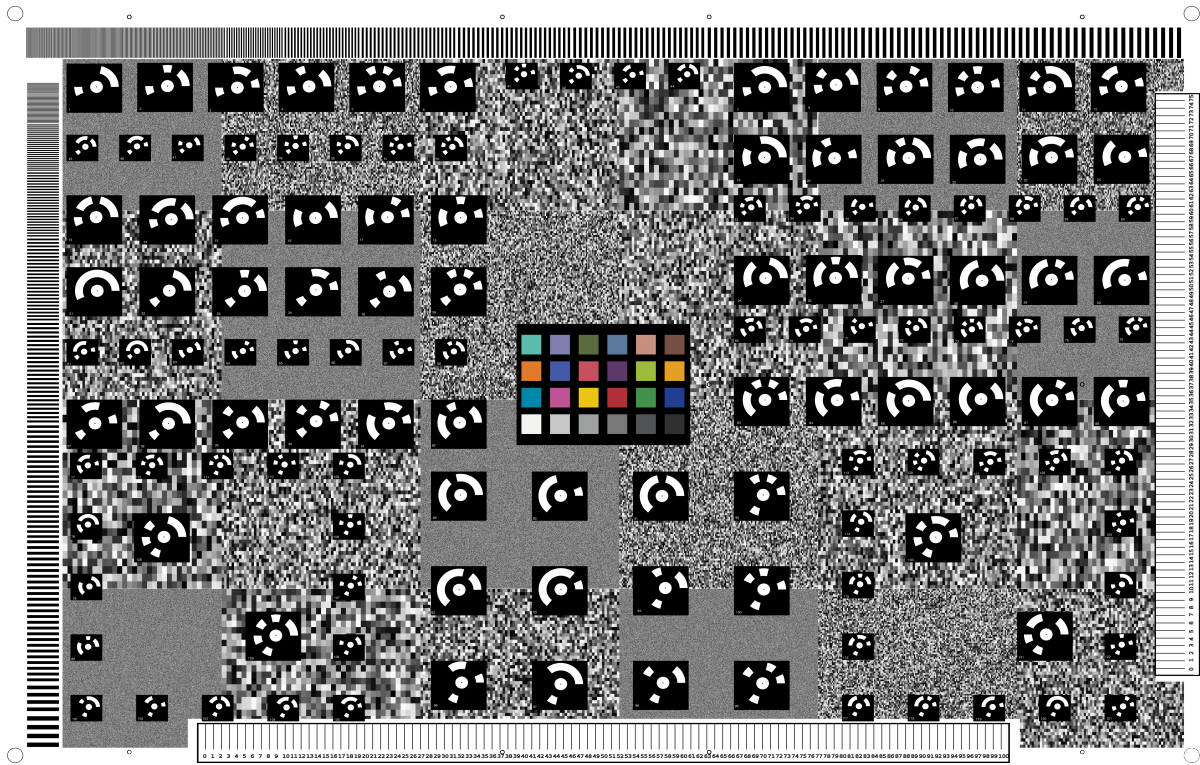


Figure 2.2: The drawing of the handmade calibration frame subsequently printed on Di-bond™ material without the targets at altitude.

- lack of resolution patterns

In an improvement perspective the previous critical issues will be resolved with the next version of the calibration platform.

2.2.2 Calibration frame design

The calibration frame is showed in Figure [2.2](#). It has dimension of $1.5 \times 1.0 \text{ m}^2$ and it is split in two part of $0.75 \times 1.0 \text{ m}^2$ for the sake of portability. It owns 131 inverted (white on black) coded targets placed on the rectangular profile, and 20 targets split in 4 plates placed at different heights with respect the rectangular profile. In order to identify the coded targets dimension two distances were assumed, a maximum and a minimum, from which snapshots of the object will be taken, that is $d_{min} = 1.5 \text{ m}$ and $d_{max} = 3.0 \text{ m}$. Using the same data of the calibration grid, except for the refractive index of water, that in this case is related to salt water, the two distances correspond to two GSDs: $GSD_{min} = 0.3 \text{ mm}/pix$ and $GSD_{max} = 0.6 \text{ mm}/pix$.

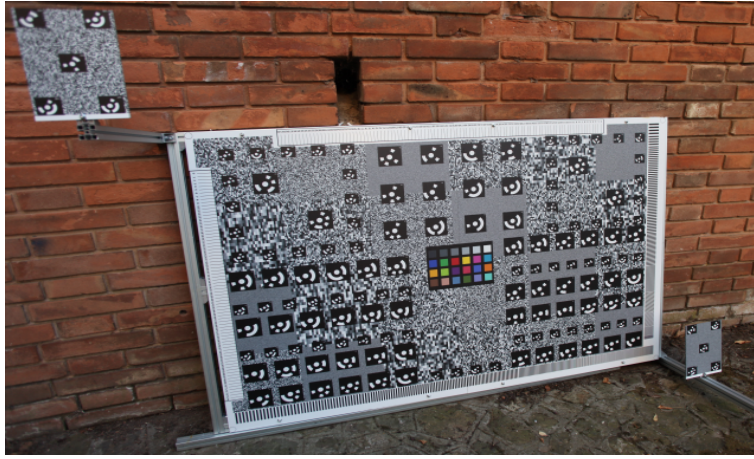


Figure 2.3: The calibration frame after printed on Dibond™. It is possible to notice that there are also two plates placed at different heights with respect to the calibration frame.

Considering that the dimension of a target should be ten times bigger than the respective GSD in order to have good level of visibility, it obtains a minimum dimension of the targets equals to 3 mm and a maximum one equals to 6 mm. Therefore it was decided to carry out two classes of target, one of dimension of 4 mm and another of the dimension of 8 mm. The two target classes are clearly visible in Figure [2.2](#), the small targets correspond to the class with the smallest values and vice versa. The targets were spread around the total area randomly but uniformly. The material used to build the handmade calibration frame is Dibond™ ($\sim 100 \text{ €/m}^2$), an alloy formed by two thin aluminum sheets wrapping a layer of polyethylene going to form a 3mm-thick rigid panel. It is used for outdoor signage since it is very weather resistant and allows the same print quality to be achieved as would be on paper. With the aim of providing a metric reference, two rulers of 1 m and 0.75 m respectively were placed along two sides of the rectangular profile. In addition, two resolution bands were placed on the remaining sides. A resolution band alternates rows of black and white spaced progressively starting from the minimum GSD. As the height from which a photo of the object is taken changes, the resolution band appears more or less detailed; the closer you get to the object, the more the alternation of bands is visible. In this way it is possible to determine the focus of the object according to the distance at which the photos are taken. Finally the coded targets are placed on several random patterns, similar to the effect when the old analog TVs lost the signal, which consists of a grid of 4x6 square of grey scale random patterns with different pixel size (special thanks to Danilo Menna for providing the code to make the random patterns). This choice was made in order

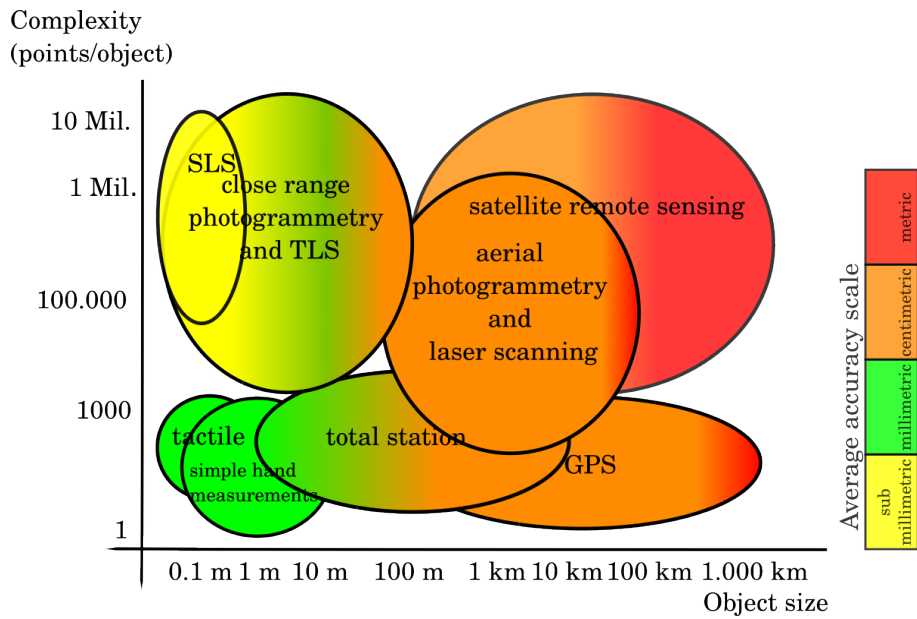


Figure 2.4: Available geomatic techniques, sensors, and platforms for 3D recording purposes, according to the scene' dimensions and complexity. A color legend is provided to show the average accuracy of the techniques.

to improve the features extraction phase of the Structure from Motion pipeline. Four holes are placed at the corners of the rectangular profile for the purpose of supporting maneuvering the calibration frame into position with the help of ropes. Finally the two Dibond™ panels are attached together by means of the aluminum bars used in the calibration grid implementation, which allows to fix rigidly the panels to the frame. The final implementation of the calibration frame is showed in Figure 2.3, where only two of the four plates placed at different heights are mounted.

2.3 Geomatic techniques for calibration platforms survey

This section will describe the geomatic techniques that were chosen to survey the calibration platforms among the available techniques depicted in the Figure 2.4, where it is possible to see the applied techniques based on the scene size and scene complexity. It is important to clarify the fact that among the available geomatic techniques, the choice was constrained by the instruments present in the LARIG (Laboratorio di Rilievo e Geomatica) Survey and

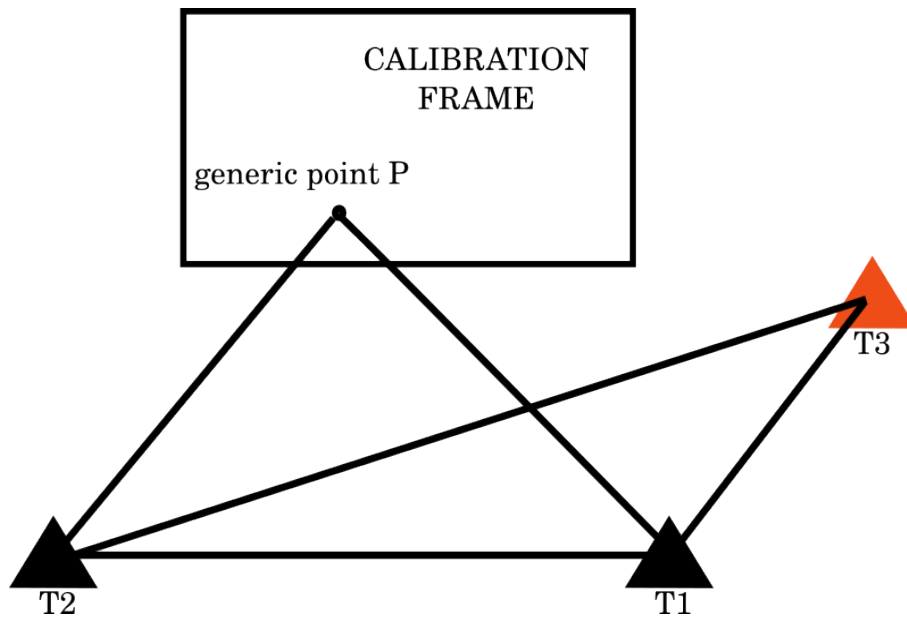


Figure 2.5: Geodetic network established for the survey of the calibration frame. T1 and T2 stationing points are used to survey the same 3D points, while T3 is used only as a foothold.

Geomatic laboratory of the DICAM (Dipartimento di Ingegneria Civile, Chimica, Ambientale e dei Materiali) Department of the University of Bologna. Since calibration platforms can undergo changes in the position of target centres relative to their structure over time due to their own movement, it is necessary to survey the calibration platform before each survey campaign. This is an operation that can take different amounts of time and accuracy depending on the technique you have chosen to use. Considering the two platforms that have to be surveyed, among the several instruments present in the laboratory, three are chosen to carry out the survey. The first is an hybrid technique that employs a Total Station (TS) and a photogrammetric approach. The second is the measure of the calibration platform by means of a Structured Laser Scanner (SLS) and the third is carry out through a Terrestrial Laser Scanner (TLS). In order to carry out a comparison, one of the techniques has to be taken as a ground truth and it was chosen the first as a ground truth since it is an already established technique, very accurate, used as a support for other techniques, that in this case is the photogrammetric technique.

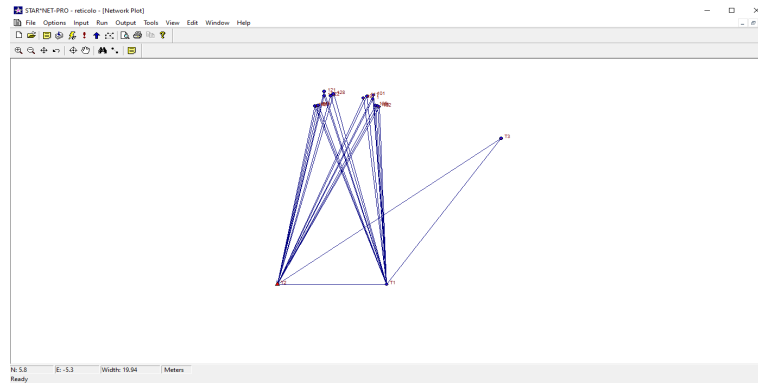


Figure 2.6: Schematic view of the geomatic network provided by STAR*NET.

2.3.1 Total station and photogrammetry

This hybrid technique is used only for the survey of the calibration frame. As you can see from the Figure [2.5](#), it has been created a geodetic network of triangular shape, where two vertices has been used to survey the 3D coordinates of the centres of several (exactly 15) coded targets, while the last remaining has been used as a foothold for the network. The same points were then collimated from two different station points in order to achieve maximum accuracy. Horizontal, vertical angles, and slope distances were observed from two survey stations that were positioned to allow favorable intersecting angles, while simultaneously allowing line of sight to all targets. A Leica Geosystem TS30 total station was used for all observations; this instrument has a 0.5 seconds of arc angular precision with a range measurement precision of $\pm 2 \text{ mm} + 2 \text{ ppm}$ to reflective tape and $\pm 0.6 \text{ mm} + 1 \text{ ppm}$ using precise circular prisms among stations (Leica 2000). In order to achieve higher accuracy, the survey scheme depicted in the Figure [2.5](#) was adopted, which by exploiting redundancy manages to achieve sub-millimeter accuracy. Angles were observed from each station, and distance observations were also made to each point allowing the target coordinates to be calculated using two methods: polar observation and a combination of polar and intersection. All the processing was performed using the MicroSurvey STAR*NET least squares adjustment package; the schematic view of the measurements in STAR*NET is reported in Figure [2.6](#). Table [2.1](#) shows the average accuracy of the targets and the total error factor for the two methods calculated using the STAR*NET adjustment package. The average standard deviation of the coordinates achieved using the combination method provided better results since the polar measurements were of lower precision. In addition, the total error factor,

computed as the square root of the ratio between the total sum of the standard residuals and the number of redundancies, shows the same result. It should be roughly within a range of 0.5 and 1.5; in general an error factor may be large for several reasons. There may be one or more large errors in the input data, there may be a systematic error, or it may have assigned standard errors that are unrealistically small. Undoubtedly the greater redundancy due to the presence of the distance observations contributes to greater reliability of the results, which results in smaller mean square deviations. Evidently, the measurements came out well even though they were made without the use of prisms, but considering the relative weights between angles and distances, due to the very short lengths, in fact the most influential and precise measurements in the calculation are the angular ones consequently the less precise distances do not jeopardize the final results and indeed they provide a beneficial small contribution to the final result. Therefore, it can be concluded that the combination approach produces the best result: a standard deviation related to the X and Z coordinates below a tenth of a millimeter and a lower total error factor than the polar approach of about 0.124. In light of these results, therefore, only the coordinates of the target centres obtained by the combination method will be used. As previous mentioned, the TS30 was used as a support for the photogrammetric survey since it needs a scale factor to estimate the target centres. A detailed discussion of the photogrammetric surveying technique can be found in Chapter 5, in this section we will only mention it. Several images of the calibration frame were taken from different points of view, promoting a high overlap among them and using a Canon EOS 5D Mark II with a focal length of 20 mm. After, the standard SfM pipeline has been carried out in Agisoft Metashape, where the target centre coordinates surveyed by the TS30 are used as control points. The accuracy level achieved by this hybrid technique has lead to the result showed in Table 2.2 where it is possible to see a metric error of the order of magnitude of a tenth of a millimeter. As a final

Method	Standard deviation (mm)			Total Error Factor
	X	Y	Z	
Polar	0.10	0.26	0.04	1.187
Combination	0.07	0.15	0.03	1.063

Table 2.1: Average Standard Deviation (1 Sigma) taken from STAR*NET Adjustment of Terrestrial Survey and Total Error Factor.

	Error(m)	X error (m)	Y error (m)	Z error (m)
Control points	0.0009	0.0006	0.0004	0.0005

Table 2.2: Error in meters of the control points in the photogrammetric survey.



Figure 2.7: Textured point cloud provided from the standard SfM pipeline in Agisoft Metashape.

result, Agisoft Metashape provides textured point cloud, which in this case is showed in Figure [2.7](#), where it is possible to see the calibration frame leaning on a wall near the engineering faculty of the University of Bologna.

2.3.2 Structured-Light projection Scanning

This technique is used for the survey of both the calibration platforms. 3D scanning technologies based on structured-light projection are increasingly being used in a wide variety of fields, from medicine to cultural heritage. These types of techniques allow for extremely high accuracy (up to the micrometre level), even though the costs of the instruments are among the highest in the field of 3D scanners. The working principle of these scanners is based on the projection of a light pattern (or code), in the EM spectrum portion of visible light (380 - 740 nm), consisting of stripes or points. The pattern is projected on the surface of the object and it is distorted by the interaction with its geometry. The distorted pattern is detected by 1 or more CCD sensors in the scanner at different positions with respect to the pattern projector. As a result, the

acquisition of images from multiple viewpoints allows the identification of the 3D coordinates of the pattern points and, consequently, of the object geometry, thanks to the triangulation principle. Accuracy in geometric reconstruction therefore depends on the size and arrangement of the stripes or points forming the pattern: the smaller the distance between the points, the greater the scanner's ability to detect the smallest surface feature. This type of process is much faster than structure-from-motion photogrammetry (please refer to chapter 5), because the number coordinates of the points detected on the object are much larger. In accordance with the area covered by the light pattern, structured light instruments will have a different range: the greater the area the pattern covers and detects, the greater the size of the object to be scanned. The operator must then position at a certain distance and move slowly (depending also on the number of frames per second acquired), trying to cover the entire object surface. The small coverage area of these scanners is one of the limitations of this technique; in fact, it is not always possible to make a single scan for a fairly large area: if the scanner cannot "connect" a portion of the object with the adjacent one, the two portions will be aligned in two different reference systems. In this situation, the scans will have to be aligned during post-processing in dedicated software. Unlike laser scanning (described in detail in the next section), structured light projection scanning first of all is based on different physical and theoretical principles; moreover, the non-coherent type of radiation emitted by SLS, unlike lasers, also allows to avoid speckle effect in the signal. In addition, the range of the instruments and thus their intended use is different, since SLS is intended for small to medium size objects, while TLS can cover much larger areas, up to hundreds of meters.

2.3.3 Terrestrial Laser Scanner

This technique is used only for the survey of the calibration frame. 3D scanning techniques based on laser emission are widely used in several fields, especially when the area to be scanned ranges from few to hundreds of meters. This type of instrument exploits LASER (Light Amplification by Stimulated Emission of Radiation), consisting in electromagnetic energy usually emitted in IR range. The peculiarity of LASER is that the radiation is not diffused (with

photons propagating in all the directions at different wavelength, as for SLS). The radiation beam is instead monochromatic (only emitted at a single wavelength) and spatially coherent; namely, all the electromagnetic energy photons emitted are in phase, allowing for highly precise focusing of the light beam. Among the most used scanning devices based on LASER (especially for architectural applications), there are the Time-of-Flight terrestrial laser scanners. These devices are constituted by a source emitting a LASER beam and a solid-state image sensor as a receiver. These are included in the scanner and are placed at a fixed distance. The emitted LASER beam, deflected by a mirror, hits the object surface, whose image is acquired by a camera composed by lens and sensor. An electronic clock, much more sensitive than an ordinary quartz one, registers the time employed by the laser pulse to reach the detector. In order to establish an absolute very accurate time resolution (3×10^{-12}) only atomic clocks can be used. With this time measurement, it is possible to calculate the distance (range) between the scanner centre and each point of the surveyed surface reflecting the signal. Together with this range values, other important data are acquired to obtain a 3D point clouds of the surveyed object: the angles in the horizontal and vertical planes passing through the centre of the instrument and changing at every measurement due to the rotation of both the mirror (in the vertical direction up to 270°) and the scanner head (in the horizontal direction up to 360°). According to the range values and the angles measurements, the X, Y and Z coordinates for each points are obtained and a point cloud is generated. The coordinates of the points refer to a relative reference system, in which the scanner is the origin and the Z (the elevation) is the height above the ground. Moreover, also RGB data can be acquired to enable texture processing of the surveyed area, through digital cameras integrated in the scanner capturing colour values that are coupled with the XYZ coordinates values for each point of the cloud. In TLS surveys, the instrument is chosen depending on the range for each specific application, varying from the medium (60-80 m) to the long-range (within 1-2 km). The instrumental range affects the accuracy of the final survey, since the point density (the number of point at each distance step) varies according to the distance. This range-dependant accuracy for TLS may vary from 2 to 6 mm. In terms of how TLS surveys are conducted, the operator must position the scanner at the correct distance to the object or scenario to be surveyed in accordance with

TLS	SLS
Advantages	Advantages
Celerity Low post-processing effort Automatic measurements Accuracy: 1 mm + 10 ppm	Costs High resolution Small objects Accuracy: 0.1 mm
Drawbacks	Drawbacks
Cost Big objects	Slowness High post-processing effort Non-automatic measurements

Table 2.3: Qualitative features of the two measurement techniques.

the instrumental range. For large areas, it is very likely that multiple scans will be required; in this case, the instrument has to be moved, ensuring to maintain an overlap area between scans of at least 30-40%. The scans acquired will be in different reference systems as these are instrument-centred. Some scanners allow pre-alignment of scans through systems based on positioning and tracking systems coming from the world of robotics. If pre-alignment does not occur, the scans will have to be aligned in post-processing either through ICP algorithm (Iterative Closest Point, in which each point in the cloud is approximated to its homologous in another cloud), or through a rotation and translation of the cloud to be aligned to another cloud, whose reference system will be considered as the "master." These strategies are often a source of alignment errors that can be propagated as transformations occur on each individual point of the cloud. The error arising from the alignment of multiple clouds, will affect the global error of the geometric reconstruction obtained from the survey. The orientation of the point clouds obtained from TLS is always relative to the instrument centre, in other words the reference system is local. If the point clouds have to be oriented in a global reference system, it is necessary to use targets in the cloud to be used as control points to be associated with 3D coordinates obtained with GNSS-based positioning systems.

2.4 Comparison techniques

Technical comparisons between SLS and TLS measurement techniques will be made in the following sections where the results obtained will be shown quantitatively. In this one, on the



Figure 2.8: An operator while conducting the calibration grid survey with an empty tank. It can be seen that the tank is below ground level, which is why this instrument was chosen to carry out the survey. It is also possible to identify the use of a laptop connected to the instrument that can display the data acquired by the instrument in real time.

other hand, the two measurement techniques will be compared in qualitative terms, i.e., time taken to make the measurements, resources used and associated costs. The two instruments are by construction very different from each other, and therefore they are also operationally different. The SLS is mostly used for small objects (<1 m) and needs experienced operator for the data acquisition phase. The post-processing phase also requires a great deal of experience, since it is generally necessary to divide the same survey into several parts, which will then be aligned by means of alignment between point clouds. TLS is used to detect large objects, usually inside enclosed spaces. It does not need an operator during the data acquisition phase; indeed, it is enough to set the survey parameters properly to obtain obtained results. However, the post-processing phase requires much attention especially during data filtering. Table [2.3](#) attempts to summarize the characteristics of each tool used during the survey so as to highlight what advantages and disadvantages are associated with each technique used.

2.4.1 Calibration grid

For the terrestrial survey of the calibration grid only the SLS was used, exactly the ARTEC EVA SLS which has a 3D nominal accuracy of 0.1 mm. This choice was driven by the fact that the use of a Total Station would have been inconvenient since the water tank where the calibration grid was located was about 3m underground, please see Figure [2.8](#), and there was a limited space to manoeuvre around it. Unlike the calibration frame, the survey of the calibration grid was carried out after taking the underwater photos. In order to limit movement relative to the target centres as much as possible, the calibration grid was placed at the bottom of the tank at the beginning of the experiment and only then it was filled with water. When the photo-capture phase was finished, the tank was slowly emptied. When it was completely empty, an operator safely surveyed the calibration grid where it had previously been placed in order to limit as much as possible any movement of it and thus reduce the risk of minor impacts that could have changed the position of the center of the targets. The tank bottom turned out to be an irregular enough surface and it does not require the use of random patterns during the survey phase. With this type of instrument, the survey was carried out within a time interval of 20 to 30 minutes and the accuracy level achieved is that related to the precision of the instrument, as mentioned before around a tenth of a millimeter. The instrument, during the survey, was held at an average distance from the object of approximately 30-40 cm and then the data was imported and processed through software provided by the instrument's parent company that was able to provide the coordinates of the centres of the targets. These coordinates will later be used to make a comparison with those estimated by photogrammetric techniques and those with a correction due to water-induced refraction effects.

2.4.2 Calibration frame

For the terrestrial survey of the calibration frame both techniques, SLS and TLS, were used, exactly the ARTEC EVA SLS and the Leica Geosystems RTC360 3D Laser Scanner. For what concerns the SLS survey, an operator stands at a distance of approximately 20-30 cm from the object, as required by the manufacturer, to perform the calibration frame survey.

This required several acquisition phases, corresponding to different point clouds that were subsequently aligned. In order to also detect targets placed at different heights with respect to the plane of the calibration frame, as the instrument is very sensitive to abrupt changes along the vertical, objects were placed, which were subsequently removed during the cloud alignment phase, in order to alleviate the sudden change in altitude. As for the TLS survey, the instrument was simply placed in front of the calibration frame on its tripod. The instrument was started up and in a few minutes finished the survey without the need for any operator. In order to find out which of the two surveying techniques achieved the best results, it was necessary to compare the two techniques based on the difference between the coordinates obtained with these two techniques and those obtained from photogrammetry, used as a ground truth. To obtain these differences, it was necessary to report the coordinates of the centres of the targets obtained with SLS and TLS in the reference system of the photogrammetric technique defined by means of the TS. By means of a rigid rototranslation it was possible, since the scale factor for the photogrammetric case it is induced by the TS, while for the other two cases it is induced by the instruments themselves. At least three homologous points are needed to make a rototranslation, therefore five points have been chosen localized at the edges and center of the calibration frame so that they were evenly distributed. Assuming that we want to bring back the homologous points defined in reference system A to reference system B, it is necessary to solve the following equation

$$RA + t = B$$

If the data is noisy thus it is necessary to minimize the least square error

$$err = \sum_{i=1}^N ||RA^i + t - B^i||^2$$

Finding the optimal rigid transformation matrix can be broken down into the following steps:

1. Find the centroids of both data set
2. Bring both data set to the origin then find the optimal rotation R
3. Find the translation t

The centroids are just the average of the points and can be calculated as $centroid_A = \frac{1}{N} \sum_{i=1}^N A^i$ and $centroid_B = \frac{1}{N} \sum_{i=1}^N B^i$ where A^i and B^i are 3x1 vectors eg. $[x y z]^T$ and $N = 5$. For finding the optimal rotation a Singular Value Decomposition (SVD) is carry out. It provides a matrix called E decomposed into three other matrices, such that:

$$[U, S, V] = SVD(E) \quad E = USV^T$$

E is square matrix so U, S and V will have the same size of E . To find the optimal rotation it first re-centre both data set so that both centroids are at the origin. This removes the translation component, leaving on the rotation to deal with. The next step involves accumulating a matrix, called H , and using SVD to find the rotation as follows:

$$H = (A - centroid_A)(B - centroid_B)^T \quad [U, S, V] = SVD(H) \quad R = VU^T$$

H is the covariance matrix. $A - centroid_A$ and $B - centroid_B$ are operation that subtracts each column in A/B by $centroid_{A/B}$. Now that it has solved for R it can solve for t . Placing the centroids into the original equation it gets:

$$R \times A + t = B \quad R \times centroid_A + t = centroid_B \quad t = centroid_b - R \times centroid_A$$

The procedure seen so far takes its cue from Kabsch's famous algorithm, a method for calculating the optimal rotation matrix that minimizes the RMSD (root mean squared deviation) between two paired sets of points. To perform this transformation, a routine was created on Octave that yielded the following results:

$$R_{SLS} = \begin{bmatrix} 0.9988 & 0.0129 & -0.0463 \\ -0.0476 & 0.1361 & -0.9895 \\ -0.0065 & 0.9906 & 0.1366 \end{bmatrix} \quad R_{TLS} = \begin{bmatrix} -0.9973 & -0.0737 & 4.875 \times 10^{-5} \\ 0.0737 & -0.9973 & -3.0765 \times 10^{-4} \\ 7.1310 \times 10^{-5} & -3.0321 \times 10^{-4} & 1.0000 \end{bmatrix}$$

$$t_{SLS} = \begin{bmatrix} -129.3326 \\ -203.0421 \\ 97.1698 \end{bmatrix} \quad t_{TLS} = \begin{bmatrix} 0.1450 \\ 7.9599 \\ 1.5898 \end{bmatrix}$$

Once the rotation matrix and translation vector were identified for both instruments used, the coordinates of all target centres were reported in the same TS reference system. To get a global idea of how the differences between the coordinates are distributed for both measurement techniques, matrices of the relative distances between the targets were created. A matrix of relative distances between targets is defined as a symmetrical square matrix in which the leading diagonal is zero since each row contains the 3D distance between a certain target and all others. To show the results for these matrices, two heatmap graphs were constructed that associate a scalar with each cell in the matrices, which is then represented through an appropriate color. In constructing these graphs, the modulus of coordinate differences between TS and the two measurement techniques was calculated. The results are shown in Figure 2.9, where it can be seen that the results attributable to SLS possess a more uniform distribution, in the sense that they have very low values (blue color) for almost all targets. Results related to TLS, on the other hand, show a more random pattern as there are many cells that have medium to high values (yellow color). In general, it can be assumed, based on the value obtained from the average of the measurements, that the SLS technique works better here than the TLS solution. In both graphs it is possible to observe a yellow area related to the targets with higher numbers. They are the targets that are located at altitude, and a higher measurement error is associated with them. The heatmap graphs show a global assessment and they suggest that SLS technique performs better than TLS one. In order to witness this global assessment even better, frequency histograms of the coordinate differences in the three coordinate axes X, Y and Z were created and the results obtained were analyzed. The Figure 2.10 summarizes the frequency histograms obtained for both measurement techniques in addition to indicating other useful information to improve the reliability of the evaluation of the results. The description of the results can be conducted by looking in columns the Figure 2.10. For the first column the mean of values associated to SLS is lower than that associated to TLS by two-tenths of a millimeter, while the value associated to standard deviation of the difference of the coordinates associated to SLS

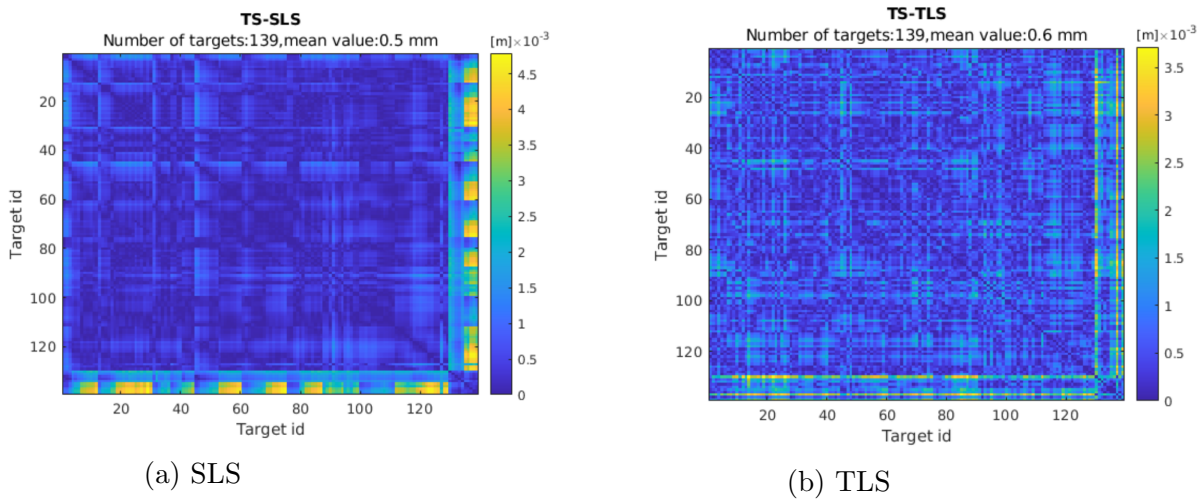
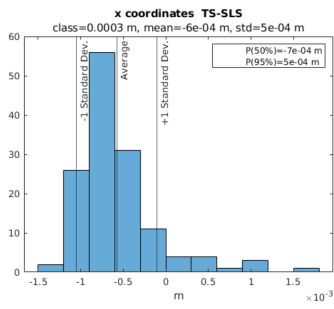
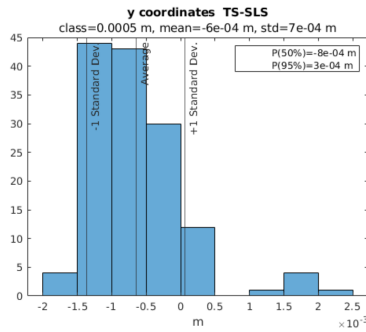


Figure 2.9: Matrices of relative distance differences expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles.

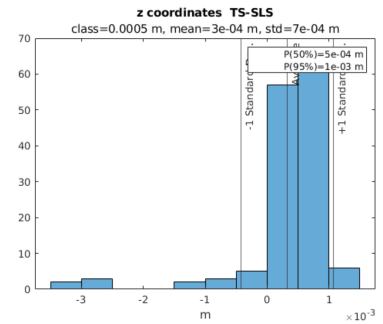
is lower than that associated to TLS by exactly one-tenths of a millimeter. For the second column the mean of values associated to SLS is lower than that associated to TLS by exactly two-tenths of a millimeter, while the value associated to standard deviation of the difference of the coordinates associated to SLS is equals to that associated to TLS. For the last column the mean of values associated to SLS is lower than that associated to TLS by exactly two-tenths of a millimeter, while the value associated to standard deviation of the difference of the coordinates associated to SLS is lower than that associated to TLS by exactly two-tenths of a millimeter. It can therefore be assumed, in light of the results just shown, that the SLS measurement technique is more accurate than the TLS measurement technique, and that the gain in accuracy that is achieved with the SLS measurement instrument is about two-tenths of a millimeter.



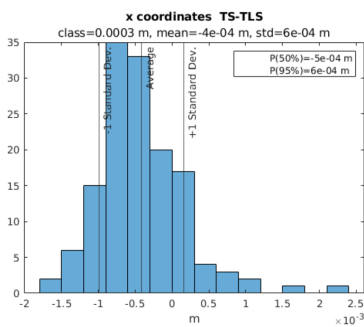
(a) Frequency histogram of the difference of the x coordinates between Total Station and SLS.



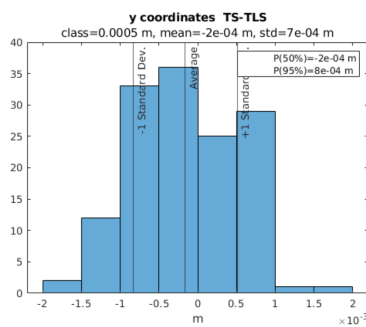
(b) Frequency histogram of the difference of the y coordinates between Total Station and SLS.



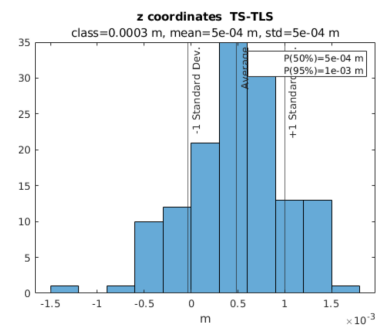
(c) Frequency histogram of the difference of the z coordinates between Total Station and SLS.



(d) Frequency histogram of the difference of the x coordinates between Total Station and TLS.



(e) Frequency histogram of the difference of the y coordinates between Total Station and TLS.



(f) Frequency histogram of the difference of the z coordinates between Total Station and TLS.

Figure 2.10: Frequency histograms of the difference of the coordinates between Total Station and both the techniques used for the survey. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value. For clarity, along the axis the order of magnitude is millimeters.

Chapter 3

Underwater Photogrammetry

3.1 Introduction

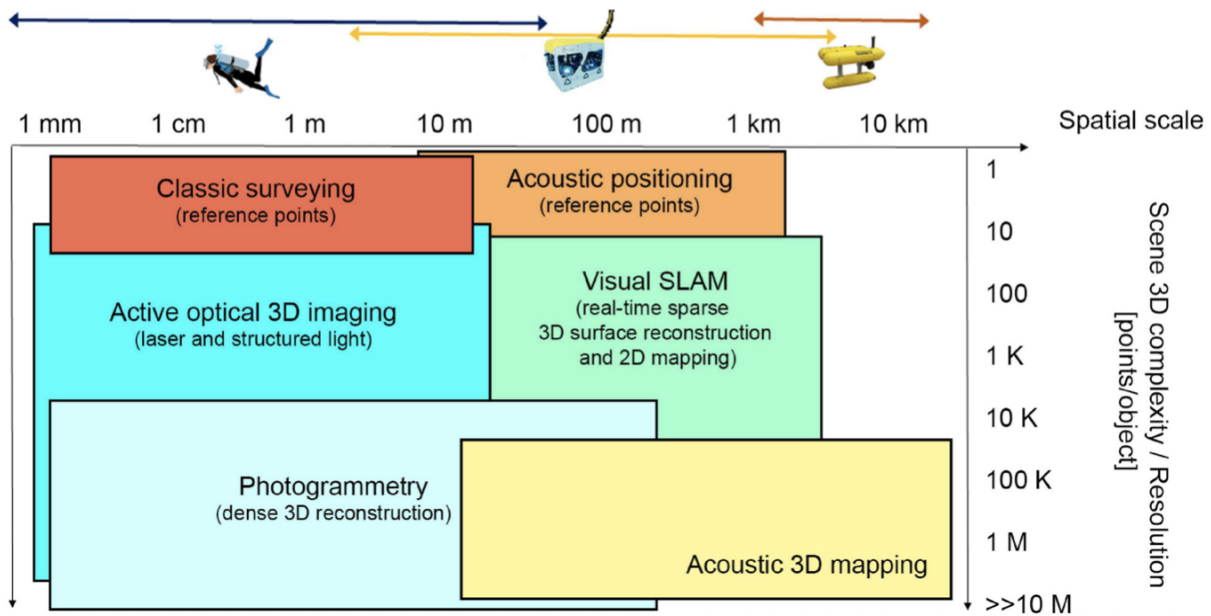


Figure 3.1: Different technological approaches suitable at different spatial scales, depths, and 3D-complexity of the investigated area for underwater surveying. Source: [1].

Underwater photogrammetry saw its born thanks to the efforts in two connected fields: underwater photography and the development of underwater vehicles. These sectors play an important role, both for the economy and the research, as basements for the principles and applications of the underwater photogrammetry. The main techniques which can be employed for

underwater surveys are shown in Figure 3.1, according to the related spatial scales, depths and 3D complexity of the study area. Close-range photogrammetry is placed among the other techniques within a spatial range from about 1 mm to 100m and a complexity/resolution window between 10.000 and 10.000.000. As the spatial scale increases, different underwater platforms should be employed for the survey, as shown in the upper side of this Figure: scuba divers until the meter-level, and ROV/AUV underwater unmanned vehicles for higher scales. The techniques used for underwater survey are roughly the same of the terrestrial field, except for the global navigation satellite system (GNSS) and laser scanning, which are replaced by acoustic techniques, due to the difficulty of electromagnetic waves to spread in the water. Reasonably, the nature of these techniques is exactly the same for the terrestrial and underwater environments, while their practical application should be reconsidered due to the presence of water, that makes the underwater environment particularly complex for the surveying activities. In order to understand the evolution of the underwater photogrammetry, it is necessary to provide a development's framework of the underwater photography and vehicles over time. In the following list the main findings and developments that allowed and supported the spread of underwater photogrammetry are shown:

- 1853: William Bauer took the first underwater photos through the portholes of a submarine. He was one of the first people realizing that taking underwater photographs would need severe modifications from standard photographic equipment.
- 1856: William Thompson took the first underwater image of the seabed in the Weymouth bay, England, with an exposure time of 10 min.
- 1864: A torpedo device, by Luppis-Whitehead Automobile in Austria, was employed as the first underwater navigating vehicle.
- 1893: The french marine biologist Louis Boutan and his collaborators developed the first underwater camera system with flash.
- 1915: From still underwater photography to underwater movie: "Thirty Leagues Under the Sea" by Carl Gregory.

- 1926: Dr. William Longley and the National Geographic photographer Charles Martin took the first underwater colour photograph.
- 1943: the Engineer Émile Gagnan and the Naval Lieutenant Jacques-Yves Cousteau invented the aqualung, an underwater breathing apparatus, improving on the Le prier and Commeinhes "SCUBA".
- 1952: The first tethered Remotely Operated Vehicle (ROV), named POODLE, was developed by Dimitri Rebikoff.
- 1953: Hans Hass and Franke & Heidecke the Rolleimarin developed the first commercial underwater camera housing. Development of an underwater vehicle for divers: the "Pegasus".
- 1957: Cousteau and de Wouters developed the CALYPSO-PHOT 35mm underwater camera, leading to the commercialization of Nikon's series cameras. The SPURV (Self-Propelled Undersea Research Vehicle), developed by the University of Washington APL, saw its first application.
- 1970s: Giddings pushed the use of dome ports.
- 1986: The Titanic shipwreck was found and 2000 images were taken with the deep submergence vehicle (DSV) Alvin.
- 2000s: Low-cost underwater digital cameras arise on the market with affordable prices making the general public able to take underwater photographs.
- 2012: James Cameron reached the bottom of the Mariana Trench with the submersible vessel Deepsea Challenger, diving to a depth of 10908 m.

Alongside the reached goals in the fields of underwater vehicles and photography, the main achievements in the underwater photogrammetry over time should be analysed in order to appreciate the efforts that led to the currently available technology.

- 1990: Recovering Shape and Motion from Undersea Images (Negahdaripour et al., 1990).

- 1990s: Design and Development of OTTER (Ocean Technology Testbed for Engineering Research) with visionbased capabilities (Wan et al., 1996).
- 1995: Stereo-video measurements of subtidal organisms (Harvey & Shortis, 1995).
- 1997: A two-phase model for quantitative analysis of underwater stereo images (Li, et al., 1997).
- 2006: Photogrammetry for virtual exploration of underwater archeological sites (Drap et al. 2007).
- 2010: SfM based approach to map archaeological sites (Skarlatos et al., 2010).
- 2010s: Refractive camera calibration for underwater applications (Telem & Filin, 2010; Agrawal et al., 2012; Jordt-Sedlazeck & Koch, 2012).
- 2013: A photogrammetric approach to survey floating and semisubmerged objects (Menna et al., 2013; Nocerino & Menna, 2020).

Taking the described timeline into account, the following sections will deal with the issue of Underwater Photogrammetry. Firstly, the measurement instruments' behaviour inside the water will be analysed, also discussing about the optical effects introduced by the used RGB sensors and possible methods to import the proper color from underwater pictures. Finally, in the context of the experiment carried out within this thesis, a camera calibration algorithm will be deeply described and applied.

3.2 Water physics

Underwater Photogrammetry is a quite flexible technique suitable for high resolution mapping. Besides, it presents some limitations especially in case of mapping large objects, as for instance several dozens of square meters. Indeed, it always provides high resolution maps but it is very time consuming compared to other techniques, such as the acoustic ones. The presence of water makes the underwater environment full of pitfalls for the common instruments and

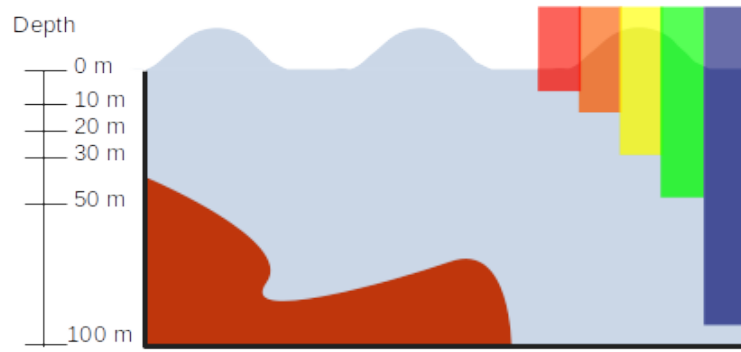


Figure 3.2: A schematic view of the effect of the sea's selective filter applied to the visible wavelengths of the light.

techniques, as those employed for the terrestrial sites. Underwater Photogrammetry is based on taking a certain number of images under the sea on selected conditions. In order to understand how performing this simple operation could be tricky in the underwater, suffice is to say that the pressure felt due to the atmospheric weight is around $1g/dm^2$ while the pressure felt underwater is one thousand times bigger, around $1kg/dm^2$; seawater is nearly 800 times denser than air, and this fact influences the underwater images' formation as the path of the optical rays is modified. For these reasons, dealing with Underwater Photogrammetry, the equipment must be protected using a proper housing able to avoid the water intrusion exerting at the same time enough force to resist the water's weight. In addition, it has been demonstrated that the seawater's density is not constant with the depth, as it is a function of temperature, salinity and pressure. In a nutshell, underwater the water pressure increases of 1 atmosphere ($1.033N/cm^2$) every 10 m, thus the internal arrangement of optical systems may be altered as the working depth varies. In this scenario, the most important characteristic is the light's absorption. The optical remote sensing, such as the one from satellites, exploiting the sun light cannot be applied. Indeed, in open ocean, only the 94% of the solar radiation goes through the water surface and thus it is transmitted [15]. In addition, it is proven that within the first 10 m of depth, the water absorbs more than 50 percent of the visible light energy. In general, the amount of light upward reflected depends on the height of the Sun (time of the day and season) and on the sea's conditions (waves, wind). The phenomenon of light's absorption is not homogeneous, but it is differently distributed according to the wavelength; see Figure 3.2. Spectrum of the solar light is composed by several wavelengths and the water acts like a selec-

tive filter. The red light is almost instantaneously obstructed on the water surface: indeed, at 6 m of depth a very small amount of red light can be detected, while the other wavelengths can go through a slower absorption process. Therefore, the great amount of light is absorbed (converted in heat) and only 1% reaches 100 m of depth. The absorption increases together with the wavelengths, meaning that longer wavelengths (red, orange) and UV are first absorbed, followed by shorter ones. In particular, wavelengths belonging to the infrared are absorbed in the first few meters while those belonging to Ultraviolet, in the case where the water is clear, are able to penetrate as far as 20 m. In addition, especially for shallow water (4-5 m of depth), another phenomenon to be taken into account is the presence of caustics. These are sunlight's reflections which should be avoided for a proper photogrammetric acquisition. Indeed, since the traditional automatic SfM's pipeline is based on contrast changes, caustics represent a source of noise because they change from one picture to another. Different methods able to remove these variable patterns exist: a straightforward approach could be to take pictures in the early afternoon hours, when the water ripple reflection phenomenon is milder. If there is no light, especially for higher depths, artificial light must be a solution to perform Underwater Photogrammetry, since it compensates for color loss and light attenuation. When we have to deal with artificial light, a good procedure is to direct the light sources as far outside the field of view as possible. Two main factors, turbidity and backscattering, play an important role when we have to use artificial light. The first is due to suspended particles that cause light scattering or diffuse reflection, while the second arises when the light from an artificial source is reflected from the particle back to the camera. The lights should be directed in such a way that the cone, created by the overlap of the light sources, does not illuminate the particles between the sensor and the object, see Figure [3.3](#). One of the characteristics related to signal transmission in the water medium is the abrupt attenuation of the electromagnetic waves. Global Navigation Satellite Systems (GNSS) are based on information's exchange through electromagnetic waves, thus it is impossible to establish a GNSS connection underwater. An alternative solution is given by the acoustic systems, that, unlike electromagnetic waves, are based on compression waves and exploit the same feature of marine animals to communicate underwater. In the underwater environment, the sound undergoes a minimal attenuation. Besides, it is considered a

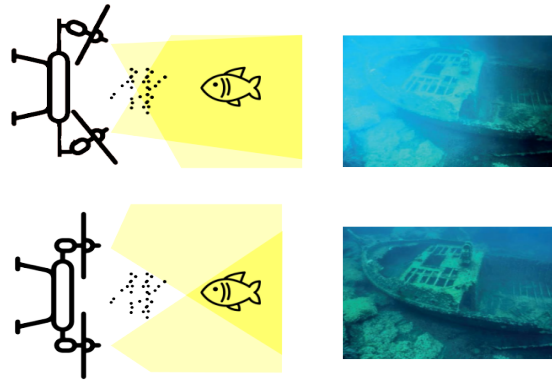


Figure 3.3: Effect of the artificial light's configuration on underwater pictures. It is easy to note how a proper light configuration improves the quality of the images. Correct configuration at the bottom of the figure.

noisy source, which has to be managed using a stable algorithm able to overcome limits related to multipath phenomena and signal's loss. In general, acoustic systems are more expensive than GNSS receivers and their practical arrangement is more complex. Therefore, the scientific community was pushed to carry out low-cost solutions to exploit the GNSS' power even in the underwater, in order to reach reliable accuracy and to provide an absolute georeferencing. Among the possible methods to obtain a GNSS positioning in the underwater environment, two are worth listing: GNSS Intelligent Buoys and GNSS poles. The former are hybrid systems consisting of a surface GNSS receiver exploiting a differential approach (DGPS) and a submerged hydrophone. The hydrophone receives the acoustic impulses periodically emitted by a synchronized pinger installed on-board the underwater platform and records their time of arrival (TOA). The buoys communicates via radio with a central station (on the vessel) where the position of the underwater target can be computed [16]. Conversely, GNSS poles are low-cost solutions that can be employed in shallow water at maximum depths of about 6 m. In these cases, poles higher than the sea level in the area of interest can be used. A GNSS receiver is installed at the pole's tip with attached a tilt sensor, while at the base a handy holder, usually pointed, allows to be threaded into the Ground Control Points (GCPs) placed on the seabed. Moreover, this second approach requires at least:

- a diver, who has to conduct the underwater activities, such as laying down the GCPs and holding the pole as stable as possible

- a swimmer, who helps the diver to hold the pole and communicates between the divers and the vessel's crew during the procedure of fixing the pole in the correct position over each GCP;
- a vessel, able to host the crew and the instruments
- a person who has to start a PPK(Post Processed Kinematic), or RTK(Real Time Kinematic), or a NRTK (Network Real-Time Kinematics) session to record the position of the GNSS receiver.

In this way it is possible to create an underwater GNSS network by knowing the height of the pole [17]. The explained configuration suffers from different sources of error, due to the fact that the pole can be deformed under the water's pressure and the tilt sensor, consisting of a tiny inertial platform placed at the bottom of the antenna able to measure the pole tilt, is not enough to correct the planimetric errors when the pole is not perfectly vertical. Therefore, in the best situations, the reachable accuracy is within 10 cm, considering poles 4-5 m long. Reasonably, when the underwater GCPs' network is located near the shoreline, it is not necessary to arrange for a vessel and a GNSS receiver. Rather, it is sufficient to mount on the pole's tip a prism which can be collimated from a TS placed on the land [18]. If it is not necessary to create products aligned to an absolute reference system, comparable precision can be obtained using a local reference system. Note that when we refer to high precision, an accuracy not overcoming the centimeter level is meant. High accuracy underwater surveys provided important answers to all the questions related to underwater archaeology, biology, engineering, oceanography and so on. In particular, for what concerns biology, the coral reef growth's monitoring is one of the main cases where high precision Underwater Photogrammetry becomes fundamental. Typical coral growth's rates are in the order of 1.1 mm/year (*Agaricia agaricites*), up to 180 mm/year (*Acropora cervicornis*) [19]. However, the coral's species that represent bio-indicators or play important roles within the feed chain have growth rates among 10-20 mm/year. For this reason, their monitoring needs the application of high accuracy underwater surveying, since even small variations have to be detected in the growth of these animals. Often, these marine animals live distributed over a large areas, where modest height

differences can be appreciated and the seabed is complex, rocky, full of cavities and inhabited by numerous species. Acquiring high accuracy underwater networks in these kinds of areas is a non-trivial task and still a topic of the current research, especially when relatively large areas need to be covered. In order to establish a reliable monitoring of the coral growth and consequently reaching a stable level of accuracy of about 1 cm, Authors of [20] suggest to exploit potentialities of both geodetic control networks and Underwater Photogrammetry at the same time. However, it is expected that the photogrammetric accuracy is at least one order of magnitude lower. It has been demonstrated that a geodetic network, capable of highlighting error's accumulation and non-linear deformation over the photogrammetric model, is strongly needed to overcome possible systematic errors. Typically, an underwater network is established considering the principles of trilateration, height measurements using diving computers, and point-to-point measurements. Distances are measured using measurement tapes as well as rigid graduated bars. The quality of a photogrammetric image processing's pipeline strongly depends on the image acquisition strategy, i.e. nadir and oblique acquisition angles combined with a double grid (cross-wise) flight path, as well as a well distributed network of GCPs. Neyer et al. try to summarize the most important aspects of high precision geodetic underwater network for photogrammetric applications, which derive from the strong physics laws that control the submerged environment. Due to their significance, they are reported in the following paragraph aiming to guide anyone wants to reach this level of accuracy.

- Stable GCP points: stability must be provided during the measurement of GCP coordinates and during the photogrammetric acquisition.
- Accurate measurements: the measurements for deriving the geodetic coordinates have to be as accurate as possible.
- Good geometry: measurement connections in both distances and height need to be optimized for any given scenario, however the topography might prohibit important connections.
- Height component: in underwater surveys the vertical axis is generally the more important, therefore it is necessary to carry out precise height estimates.

- Minimization of mechanically related effects: GCPs have to be attached as close to the topography as possible by ensuring line-of-sight visibility from measuring geodetic network. For the photogrammetric network, if they are directly attached to the topography the danger to accidentally damage a GCP is reduced and they can be used in post-processing to obtain a fine point cloud registration solely based on these targets.
- Additional signaled points: in addition to control points of known position, these points with unknown coordinates should be placed and used as check points.

3.3 Geometry

All underwater applications requires the cameras to be placed into particular waterproof housings. In the last decades, there is great availability of different possible housings, ports and materials, but usually the underwater scene is observed through a dome or a flat port. Therefore, optical effects on the resulting images, such as focusing shifts and field of view's variations, need to be considered by the experts depending on the chosen type of port. In most cases, typical ports can produce optical aberrations, such as lateral and axial chromatic ones, astigmatism and field's curvature. Thanks to the current knowledge in the photogrammetric field and the technological development, nowadays it is possible to take into account all the listed issues providing several optimal solutions to improve the final accuracy of the underwater products, such as the images. In the last few years, the consumer-grade market of Underwater Photogrammetry got very popular and underwater housings are available for a big range of digital and compact cameras, directly designed by the manufacturer, or from third parties companies. In this context, both experts and amateurs share the same tools available to document the submerged environment. Reasonably, from the scientific point of view, it is necessary to distinguish between flat and dome ports and to evaluate the possible advantages and disadvantages of each solution in a photogrammetric context. Lenses used in photography are designed to minimize global optical aberrations. However, residual aberrations are always present and their amount depends on the optical design, the glasses' quality and the cost of the lens them-

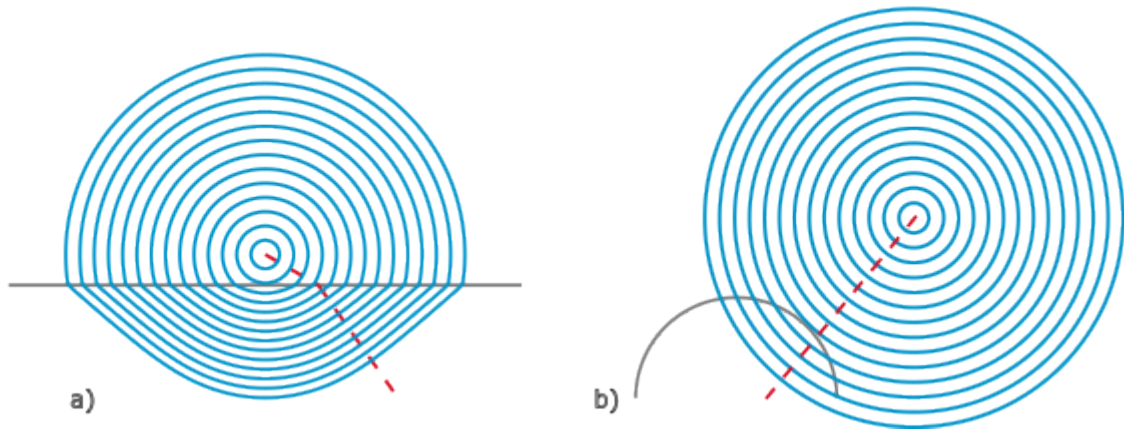


Figure 3.4: Behaviour of the ray of light meeting different surface's shapes. a) the case of a flat port: when the rays travel from a slow medium to a fast one, they are bent towards the normal direction and the image's symmetry is lost due to the change in direction b) the case of a dome port: the rays orthogonally meet the surface preserving the image's symmetry even if the two mediums are different.

selves. Moreover, even if the best available camera on the market is chosen, its use inside the underwater housings leads to some physical limits that cannot be totally removed. For this reason, a proper trade-off between the camera and port's quality has to be selected. We can assert that the shape (flat or hemispherical) of the camera port deeply affects the optics of the image's formation taking place within the camera sensor. Basing on the Snell's law, when the incident light's beam forms a null angle with respect to the normal of a surface, it continues its travel with the same direction; therefore the refraction effect is null since no direction changes occur, see Figure b) [3.4](#); and the light beam is only slowed. This happens when a light's ray gets in contact with an hemispherical port. On the other hand, when an incident light's ray from water meets a flat port, it is bent towards the surface's normal forming a refracted angle proportional to the sin of the incident angle, see Figure a) [3.4](#). An example of how the Snell's law is applied in the everyday life is the experiment of the broken pencil: a pencil partially submerged in a glass of water demonstrates the refraction's effect, that is the change of direction. It appears as a broken pencil, although it actually continues to have its original shape, see Figure [3.5](#). In general, two categories of approaches to correct the refraction effect exist: the first are based on the theory of ray tracing, also called explicit methods, while the second are based on perspective projection, also called implicit methods. They will be deeply explained in



Figure 3.5: The example of the broken pencil. A pencil is partially submerged into a glass of water to demonstrate the change in direction due to the Snell's law.

the following paragraphs under certain conditions. Before addressing this topic, it is necessary to clarify that imaging process can be affected by random errors, gross errors or systematic errors. For what concerns random and gross errors, several methods are able to recognize and detect outliers coming from observation's errors, in order to isolate or remove them from the images. Unfortunately, for what concerns the systematic errors, which are those generated on the functional or stochastic model, there is no single method to detect and remove them, since they can be identified only when they result as residuals in either the image space or at GCPs in the object space. Otherwise, these errors will deform the object space without the possibility for the user to notice their presence. However, when they arise in recognizable and repeatable residual patterns, some self-calibration models can be applied to compensate them, depending on the models' strength. In addition, when using incomplete functional models, the systematic residuals in the image plane are not only dependent on the position of the image observations within the sensors but they are also function of the 3D positions of the points in the object space. The importance of the camera's geometry in reducing the deformation in the object space was analyzed by Menna et al. in [21], where an interesting comparison between the use of flat or dome port in terms of systematic errors affecting the images can be found. They provided two solutions to solve this problem: a stochastic approach, consisting in the

radial weighting of the images applied for flat ports which improves the accuracy in the object space up to 50%, and an iterative look-up correction table that was used to reduce the evident systematic residual patterns in the case of a dome port. Through different simulations, Menna et al. investigated the depth-dependent systematic errors introduced by unmodelled refraction effects when flat and dome ports are used. Therefore, the refractive effects can introduce systematic errors which are not considered by the standard pinhole camera model due to several reasons, such as the use of a flat port, the non-perpendicularity of the flat port with the optical axis of the camera, and the non-concentricity of the entrance pupil of the lens with the center of the dome port as well as the non-sphericity of the dome port. Most of the information about the camera's geometry can be found in the following sections.

3.3.1 Dome port

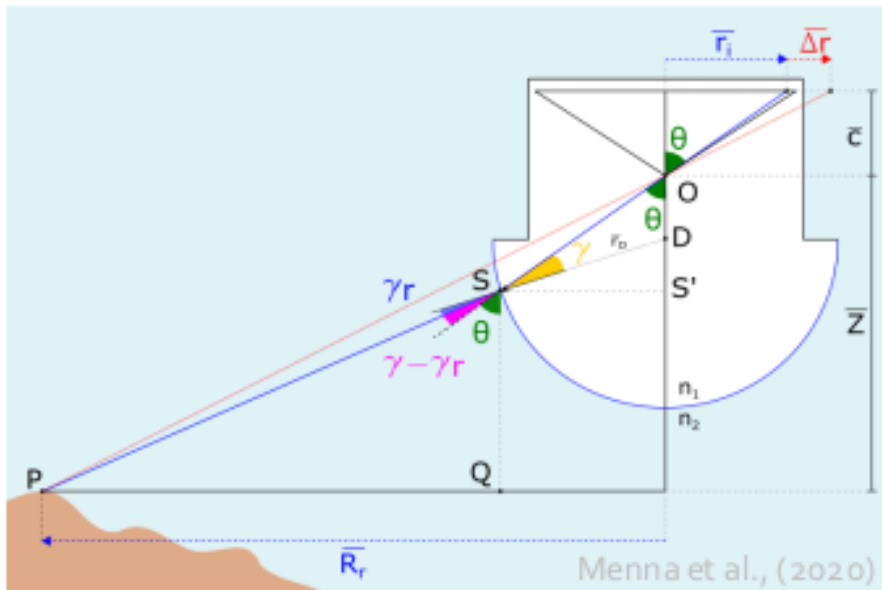


Figure 3.6: Schematic representation of refractive effects in the image formation for a dome port. Red line indicates the light path in air or the collinearity air line. Blue line indicates the light path in water according to the Snell's law. P is the submerged point projected on the sensor at the distance \bar{r}_i . Δr is function of the distance \bar{Z} and it represents the quantity projected on the sensor that differs for the in-air case and the in-water one. c is the principal distance. O is the entrance pupil. D is the center of the dome. \overline{DS} is the dome port radius. When $O \neq D$ we deal with decentered dome. The thickness of the glass flat port is not considered in this model.

Considering the situation described in Figure [3.6](#) and using the law of sines, the following

equations can be obtained:

$$\theta = \tan^{-1}\left(\frac{\overline{r}_i}{c}\right)$$

$$\gamma = \sin^{-1}\left(\frac{\overline{OD}}{\overline{SD}}\sin\theta\right)$$

$$\gamma_r = \sin^{-1}\left(\frac{n_1}{n_2}\sin\gamma\right)$$

$$\overline{SS'} = \overline{SD}\sin(\gamma + \theta)$$

$$\overline{DS'} = \overline{SD}\cos(\gamma + \theta)$$

$$\overline{PQ} = (\overline{Z} - \overline{OD} - \overline{DS'}\tan(\theta + \gamma - \gamma_r))$$

$$\overline{R_r} = \overline{SS'} + \overline{PQ}$$

The error $\overline{\Delta r}$ changes depending on the position of point P along the segment \overline{PS} : the closer the point P to the camera, the larger the $\overline{\Delta r}$ variations. In addition, an axial misalignment with the center of the dome shifted towards the object space increases the field of view of the camera and vice versa. Hemispherical dome ports are concentric lens acting as additional optical elements. They are more expensive than flat ports due to an increased construction's complexity since sphericity has to be ensured to avoid abnormal refraction effects. This kind of ports is used for wide angle resolutions cameras thanks to their resistance to high pressure. When dome ports are perfectly centered with the entrance pupil no lens distortion occurs. Using this type of ports allows preserving the field of view (FOV) with respect to the camera-lens system as well as for the focal length and the magnification effect. Concerning the FOV, its maximum value is not limited since it only depends on the properties of the camera. One of the main features in the use of dome ports concerns the effect of the appearance of a virtual image in the object space, see Figure [3.7](#). Taking an underwater picture with this configuration produces an upright smaller virtual image of the object at a distance of approximately three times the dome's curvature radius from the its surface. The camera-lens system focuses on this virtual image, so that the camera system needs to be able to focus on close range objects. For instance, a full frame photogrammetric camera usually has a minimum focus distance equal to 30 cm. Therefore, even if a point lies at an infinite distance from the dome port's surface,

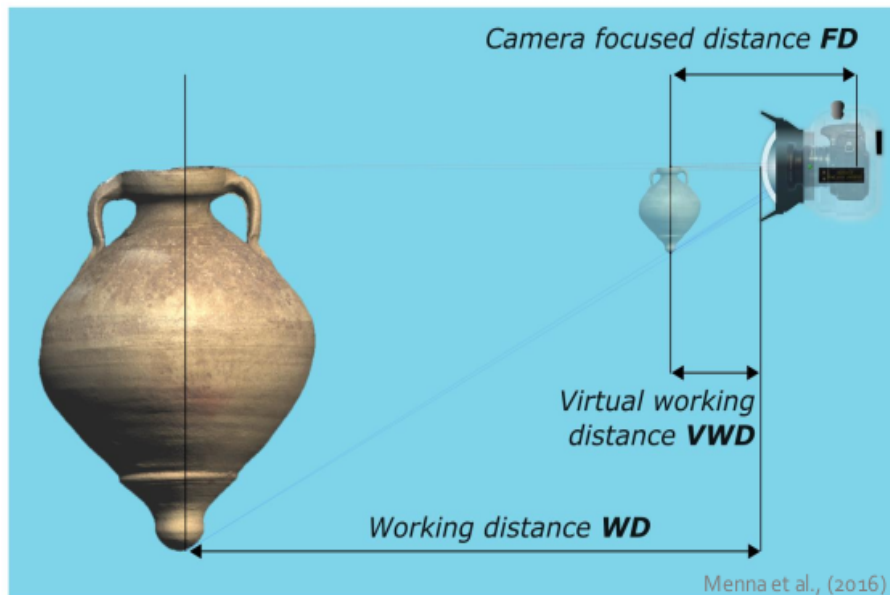


Figure 3.7: Graphical demonstration of the principle of virtual image's formation when using a dome port.

it is observed at a distance equal to three times the curvature radius of the dome port itself. For these reasons, it is necessary to consider this virtual distance when setting the suitable focus for the camera before starting the surveying measurements. In order to compute the virtual distance more precisely dome material and its geometry represents the main inputs to feed a lens software as it can be Winlens. Other effects, that are involved in using dome port, are related to an increase of the depth of field (DOF), by a factor proportional to the ratio between the refraction indices of water and air, to a spherical aberration, optical rays passing through the peripheral parts of the dome do not converge in the same focal point of the rays passing through the center, and to a field curvature that causes flat object to be projected on a paraboloidal surface, producing unsharp images.

3.3.2 Flat port

By observing Figure [3.8](#) and using the law of sines, it is possible to derive the following equations:

$$\theta = \tan^{-1}\left(\frac{\bar{r}_i}{c}\right)$$

$$\gamma = \theta$$

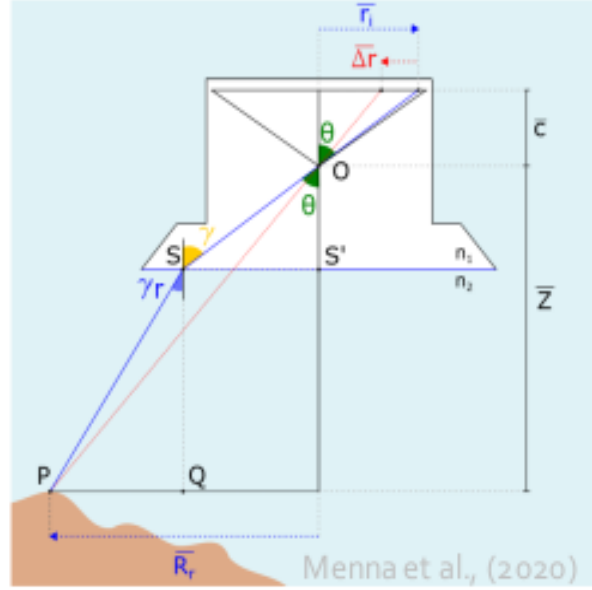


Figure 3.8: Schematic representation of refractive effects in the image formation for a flat port. Red line indicates the light path in air or the collinearity air line. Blue line indicates the light path in water according to Snell’s law. P is the submerged point projected on the sensor at the distance \bar{r}_i . $\overline{\Delta r}$ is function of the distance \bar{Z} and it represents the quantity projected on the sensor that differs from the in air case to the in water one. c is the principal distance. O is the entrance pupil. $\overline{OS'}$ is the flat port distance from the entrance pupil. It is not considered in this model the thickness of the glass flat port.

$$\gamma_r = \sin^{-1}\left(\frac{n_1}{n_2}\sin\gamma\right)$$

$$\overline{SS'} = \overline{OS'}\tan\theta$$

$$\overline{PQ} = (\bar{Z} - \overline{OS'})\tan\gamma_r$$

$$\overline{R_r} = \overline{SS'} + \overline{PQ}$$

The error $\overline{\Delta r}$ changes depending on the position of point P along the segment \overline{PS} : the closer the point P to the camera, the larger the $\overline{\Delta r}$ variations. As well as, for small $\overline{OS'}$ compared to \bar{Z} , the entrance pupil can be imagined on the flat port, therefore O and S' coincide. In this case the $\overline{\Delta r}$ is only function of the angle γ and it can be demonstrated that the error $\overline{\Delta r}$ is very well modeled by standard pinhole image formation. Flat ports consist of flat plane of optically transparent glass or polycarbonate. They are cheaper than dome ports, since they are simply material layers of a certain thickness which do not require any geometrical constraints. Usually they are used for compact digital cameras. As opposed to dome port, flat port acts as a

boundary between two different media and consequently the rays of light coming from object in water deviate from their original path when pass through port and reach camera sensor. This refraction effect causes the reduction of FOV and, conversely, the increasing of the focal length. Even if we use wide-angle lens the reduction of FOV occurs as the Snell's law demonstrates. Indeed when we use flat port the light rays pass through two interfaces: water-glass, glass-air. At the second interface, a light ray passes from a means with higher refraction index to another with lower one arising the total internal refraction effect, that is when the refraction angle is 90° . Therefore any ray of light, with incident angle higher than the critical angle, can not reach the image sensor. By applying Snell's law, the critical angle can be calculated, leading to a maximum FOV for every flat port of about 96° . By looking to Figure [3.9](#) in order to obtain the critical angle we have to set $\theta_3 = 90^\circ$. We can write for each interface an equation follows the Snell's law:

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2)$$

$$n_2 \cos(\theta_2) = n_3 \sin(\theta_3)$$

Therefore, $n_3 \sin(\theta_3) = 1$ and

$$n_2 \cos(\theta_c) = 1$$

$$\theta_c = \arccos\left(\frac{1}{n_2}\right) \approx 48^\circ$$

That is why the field of view, when using a flat glass port, can not exceed $96^\circ = 48^\circ \times 2$. In addition, the focal length and magnification effect increase by a factor equal to approximately the ratio between the refraction indices of water and air, with respect to the camera lens-system. While the effect on the observed object is that the object appears closer to the camera by a factor of about 25%, approximately the reciprocal of the refraction index of water. For that concerns lens distortion, with this type of port, a pincushion distortion is observable. While other effects are related to chromatic aberration and astigmatism. The former occurs when a lens causes different wavelengths of light to have differing focal lens resulting in a failure of a lens to focus all colors to the same point. The second it is included inside the group of optical aberrations and it is color dependent resulting in a blurring effect. If the red channel of color

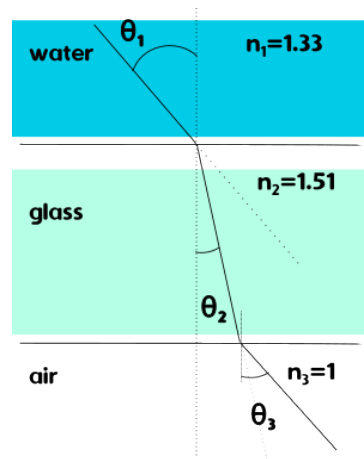


Figure 3.9: Ray tracing in a multi-media configuration: water, glass and air. Oblique dotted lines indicate direction when no refraction effect occurs. Thickness of layers is only indicative.

is not really affected by astigmatism, the blue channel is tremendously.

3.3.3 Calibration algorithms

Accurate estimates of locations and dimensions within the object space represent the main interest for an expert photogrammetrist. Within this context, calibration of camera system plays an essential role in order to ensure that measurements are reliable. Indeed, small errors in the perspective projections result in systematic errors within the measurements. In order to prevent the formation of these errors it is necessary to find a model that can recognize them. With respect to the in air case, the underwater camera calibration gains even more importance, as refractive effects and the presence of multiple media increase the complexity of the system. Underwater camera calibration is characterized by an additional uncertainty caused by:

- attenuation of the light through the housing port and water media
- small errors in the refracted light path due to modelling assumptions
- non-uniformity of the media

For these reasons, the precision and accuracy of calibration underwater are expected to be lower than calibration conducted in air. According with Shortis in [22], underwater camera calibration

can be avoided when high level of accuracy does not required and the object to be measured can be approximated to a planar surface. A scenario difficult to find in underwater surveys. In all other cases, it is a mandatory step because neglecting camera calibration parameters introduces systematic errors. It exists two kinds of calibration approaches, one that focuses on the use of correction lenses or dome ports that are able to reach a near-perfect central projection resulting in a lack of refraction effects; one that uses targets disposed on the field, which are recognizable by image analysis techniques or proper software. The latter is more expensive in terms of time and resources but it allows to reach a greater accuracy and reliability. With respect to dry camera calibration, underwater calibration has a significant advantage since the locations of the photographs are relatively unconstrained. The kernel of this approach is to capture multiple, convergent images of a portable or fixed calibration fixture in order to determine the geometrical/physical parameters of the camera calibration. The mathematical basis of the technique is described in [23]. The main parameters of the camera that rules image formation are:

- principal distance c and principal coordinates (x_p, y_p)
- parameters for the radial (k_1, k_2, k_3) distortions
- parameters for the decentring (P_1, P_2) distortions
- affinity s_x
- shear α

For that concerns radial distortions SfM software offer five polynomial terms, but usually three terms are enough, indeed five terms can be used in case of camera lenses with extreme distortion profiles. Decentring distortion is described by five terms, however in practice two terms are enough in order to obtain optimal results. There are two reason why it is necessary to carry out the calibration of a camera system:

1. determine the internal characteristics of the camera system

2. determine the pose estimation, that is the relative orientation of the cameras with respect to another or exterior orientation with respect to an external reference

As mentioned before, there are two main approach of camera system calibration:

- Absorption of refractive effects
- Geometric correction of refractive effects

In this context, it has to be considered the entire path that the ray of light has to undertake, including the camera lens, housing port and water medium. The first method is based on the concept that the camera parameters of the calibration are able to absorb the refraction effects. Considering that the camera optical axis is approximately perpendicular to a flat or dome port, the primary refraction effect due to the passage through the two interfaces (air-port and port-water) can be considered radially symmetric around the principal point. This effect can be absorbed by the three terms of the radial distortion. However, it exists others asymmetric effects due to, for instance, alignment errors between optical axis and housing port or non uniformity along the thickness of the glass port. Fortunately, these secondary effects can be absorbed by the decentring distortion and affinity parameters. The drawback of this method is that there will be always some systematic errors which do not recognize by the model, due to the fact that the principal hypothesis of the model, that is the single projection, fails because of refraction effects. These kind of errors translate in scale changes when the measurements are taken outside the range used for the calibration step. However, if the range of measurement is not so different from that of calibration, the level of systematic error is lower than the precision with which measurements can be extracted. The alternative to the absorption method is the process of geometric correction, based on the theory of ray tracing. Unlike the previous method, with geometric correction is possible to compute exactly the refraction effects and therefore remove them in precise fashion. Additional unknowns, as the distance between the camera perspective centre and the housing, and the normal of the flat port are added in the system of equations in order to find for a best solution. The type of port, refraction indices and thickness's port have to be known. For this kind of method it is required two phase of

calibration: a dry calibration, conducted in air and a wet calibration, conducted underwater. However, it exist an algorithm that are able to avoid the wet calibration that can be found in [2]. Therefore the disadvantages of this method rely on the two phase calibration, on the fact that strictly depends on the knowledge of refraction indices and finally on the issue that these types of methods can only be used with the type of port they are designed for, unlike absorption methods which are more versatile. The perspective camera model is abandoned, since it gives errors as the distance and slope of the refractive surfaces increase, to make way for a physical error-correction model that, at least at the theoretical level, should allow for a complete error compensation. If you are interested in how the perspective model fails when considering geometrical correction of refractive effects, especially for flat port, Treibitz et al. (2012) in [24] provide a complete discussion on the issue as well as define the boundaries in which the single view point (SVP) model can be applied. While Agrawal et al. (2012) in [25] provide a demonstration that describes how an axial camera model is more suitable for flat ports instead of perspective camera model, where they derive a 12th degree polynomial that must be solved to project a 3D point onto an image plane. The absorption models are also called implicit modelling of refraction while the geometric models are also called explicit modelling.

The Pinax Model

As we said before there exists an algorithm able to avoid the second traditional phase of wet calibration among the explicit modelling, and its name is Pinax by Łuczyński et al. (2017). The Pinax model can be used for accurate and efficient refraction correction of underwater cameras in flat-plane housing. The model is perfectly suited for the calibration and rectification of single camera images as well as for providing excellent initial guesses for the camera parameters allowing faster convergence in a multiview approach. The Pinax model is based on a virtual pinhole camera model, which is demonstrated that it is applicable for underwater housings where the camera is relatively close to the flat port, using the projection function of an axial camera. In order to understand the method operation it is necessary clarify the setup of the system. There are:

- physical camera that follows the standard SVP model is enclosed in a water housing with a flat glass panel through which it observes the underwater environment
- virtual camera that has the task to handle glass panel distortions introduced

The virtual camera model interprets the environment scene from the physical camera. Therefore it is used to handle the refraction induced by distortion. The parameters used in Pinax model and depicted in Figure [3.10](#) are:

- d_0 : distance from the center projection of physical camera to the glass port
- d_1 : thickness of the flat glass port
- x : distance to point of intersection of the light ray with the camera axis
- Δx : length of the focus section
- n_a, n_g, n_w : refraction indices
- \mathbf{n} : normal vector to the glass surface
- α : incident angle

While the outputs of the model are:

- d_0^* in order to build an underwater housing with flat port.
- A correction maps useful to undistort underwater images since it contains refraction corrections.

By considering the demonstration in [\[25\]](#), the physical accurate model of a flat-port underwater camera is that of an axial camera. In this kind of model, the rays of light creating the image do not intersect in a single point, as in the SVP model, instead they all intersect in one line, also called the axis of the camera. Therefore when we use the pinhole camera model we approximate this line in a point, that's why we introduce systematic errors in the measurements when this

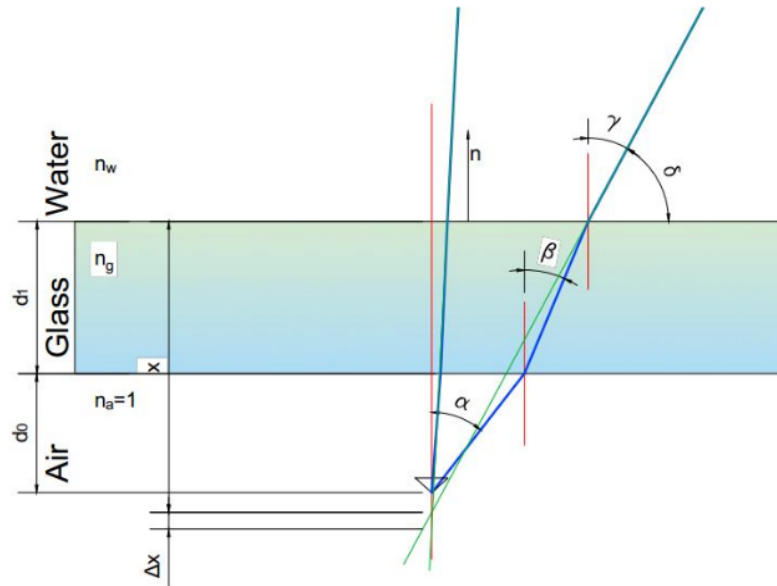


Figure 3.10: Schematic visualization of a Flat camera system. \mathbf{n} is the normal vector to the glass interface. The blue line represents the physically accurate light ray; the green line is the apparent ray traced back to the camera's optical axis. The focus distance x for each light ray is computed as $x = \tan\delta(d_0 \tan\alpha + d_1 \tan\beta)$. Figure extracted from [2].

approximation is not based on any assumptions. The quality of the approximation depends on the length Δx of this axis, the better the line is comparable to a point, the better is the approximation between an axial camera model and pinhole camera model. In other words, the pinhole camera can be seen as an axial camera where the focus section of the camera axis is infinitesimally short. The Pinax model is based on a virtual pinhole camera model while using the projection function of an axial camera model. In the work of Łuczyński et al. (2017) it is demonstrated that the best approximation of the axial camera model with a pinhole model occurs for small values of d_0 , by emphasizing how the length of the focus section is deeply dependent on this parameter instead of d_1 or α . In addition, within this context, it is possible to assert that for applications in which there is a constant altitude control of the observing UUV, the standard SVP pinhole camera performs very well, in terms of projection errors, only at a certain distance, which should be roughly the same as that with which we calibrated that camera as well as the same distance used by the UUV to surveying. Therefore it is necessary to carry out the calibration phase at a distance comparable to that used for surveying purposes. In order to be sure on the axial camera model substitution instead of a SVP pinhole camera model, it is necessary to satisfy these four assumptions, in order of importance:

1. The distance d_0 is small and near the optimal spot d_0^* where the rays traced back from the water cross in a maximum focus section Δx^* .
2. The optical axis of the camera is perpendicular to the glass surface.
3. The glass thickness and its approximate refraction index are known.
4. The water refraction index is approximately known from tables.

In general, the most stringent requirement (1) is totally fulfilled since underwater housings are designed to save space, so the physical camera inside the housing is placed as closely as possible to the flat port. Most underwater camera calibration algorithms are based on a calibration carried out on two phases, Pinax instead uses external parameters, such as refraction index, which translate the external world inside the camera system, and it allows to overcome the underwater calibration phase, in order to carry out a single calibration of the camera in a standard configuration as for that of the in-air calibration phase. It enables to save resources and money during the preparation of a survey campaign. For what concerns the water refraction index the authors demonstrated the results of the Pinax model, taking into account changes in the index, are deeply dependent on the accuracy with which insert the index within the model. The model takes the water refraction index into account for which a rough estimate of the salinity is sufficient, and for which in-situ measurements from a CTD-sensor can also be used to derive perfectly fitting correction maps. As such, this calibration model was chosen in order to become an excellent tool for calibrating the underwater UUV camera of the SushiDrop (please see chapter [4](#)) project in addition to rectifying the images produced to be fed a SfM algorithm later. Pinax is able to compute correction maps based on the input parameters, such as the water refraction index, in less than a few minutes. In order to have accurate results, it is possible to compute these correction maps as the measurements provided by the CTD sensor installed on the UUV change. You may want to go to Chapter [6](#) for further discussion of the method.

3.4 Colorimetry

In order to address the topic of underwater imaging it is mandatory to consider the physical basis of the light propagation in the water medium. The physical properties of the water cause degradation effects that in normal conditions there are not in air. Underwater images are characterized by a poor visibility due to the exponential light attenuation since the light travels water and the scene appears to be poor in contrast and hazy. The attenuation process of light is due to:

- absorption (it removes energy)
- scattering (it changes direction of light from its path)

Both these phenomena manipulate the overall performance of underwater imaging. The blurring of image features is due to the forward scattering, light is randomly deviated from its path from the object to the camera. The backward scattering leads to a poor contrast that generating a characteristic veil that overlap the image itself causing occlusion of the scene. Its contribution is given by the portion of reflected light from water towards the camera first it reaches the scene objects. But, absorption and scattering are not only due to the water medium, indeed in water there are small floating particles or dissolved organic matter, also called "marine snow" [26], which increase the absorption and scattering effects. The visibility range may be increased with artificial lights, however these sources are not only subjected to absorption and scattering, but they cause a bright illumination spot localized in the center of the image while the surrounding area is poor enlightened. Finally, since the amount of light decrease as column depth increases, colors decay based on their wavelength. As for instance, blue color travels much more than other colors since it is characterized by a shortest wavelength, therefore in the underwater images blue is dominant. Therefore, we may assume that the problems that could affect underwater images are the following:

- limited range of visibility
- low contrast

- bright artifacts artefacts
- non uniform lighting
- blurring
- noise

These are the problems that we have to face in the underwater imaging context. The elaboration of underwater images can be seen from two different point of view:

1. image restoration technique: this is an inverse problem. By knowing the image degradation model and the formation model an attempt is made to recover the appearance of the undegraded image. These methods are considered as rigorous, but they need numerous parameters such as attenuation and scattering coefficients, which are poorly known through tables and they may be tremendously variable.
2. image enhancement methods: they use qualitative subjective criteria to produce a more visually pleasing image without considering any physical models.

Light interacts with water medium through absorption and scattering, the former represents a power loss due to water depending on the refraction index, while the latter a deviation from the straight path of light also due to the small floating particles. The empirical law of Lambert-Beer assumes that the decay of the intensity of the light is related to the medium properties in which light travels and its trend is exponential. Therefore, the irradiance E at position r can be modeled as:

$$E(r) = E(0)e^{-\beta r}$$

where $E(0)$ is the amount of irradiance at the start position and β is the total attenuation coefficient of the medium. This coefficient is a measure of the light loss due to the contribution effects of absorption and scattering over a unit of length in the path inside the attenuating medium. In general, the attenuation coefficients of oceanic, coastal and bay waters are respectively $0.05m^{-1}$, $0.2m^{-1}$ and $0.33m^{-1}$. If we consider the approximation on the color channels,

β becomes β_c where $c = R, G, B$ denotes the visible colors. In this context, attenuation coefficients are defined relative to the ratios between blue attenuation coefficient and red or green one

$$\beta_{BR} = \frac{\beta_B}{\beta_R} \quad \beta_{BG} = \frac{\beta_B}{\beta_G}$$

By considering an homogeneous and isotropic medium, β can be decomposed as the sum of a and b , that are respectively the attenuation coefficient of absorption and scattering

$$E(r) = E(0)e^{-ar}e^{-br}$$

The total scattering coefficient b is the superposition of all scattering events at all angles through the volume scattering function $\phi(\theta)$ which gives the probability for a ray of light to be deviated of an angle θ from its direction of propagation

$$b = 2\pi \int_0^\pi \phi(\theta) \sin(\theta) d\theta$$

Therefore, we may assume that $a, b, \beta, \phi(\theta)$ represent the inherent properties of the medium and their knowledge should allow to theoretically predict the light propagation the medium. However, all these parameters also depend on the 3D position as well as on the time. McGlamery in [27] provides a model that takes into accounts all the previous issues and according to his model the underwater image formation model may be represented by the linear superposition of three components:

1. E_d the direct component: light reflected directly by the object that has not been scattered in the water
2. E_f the forward-scattered component: light reflected by the object that has been scattered at a small angle
3. E_b the backward-scattered component: light reflected by objects not on the target scene but that enters the camera, for example due to floating particles

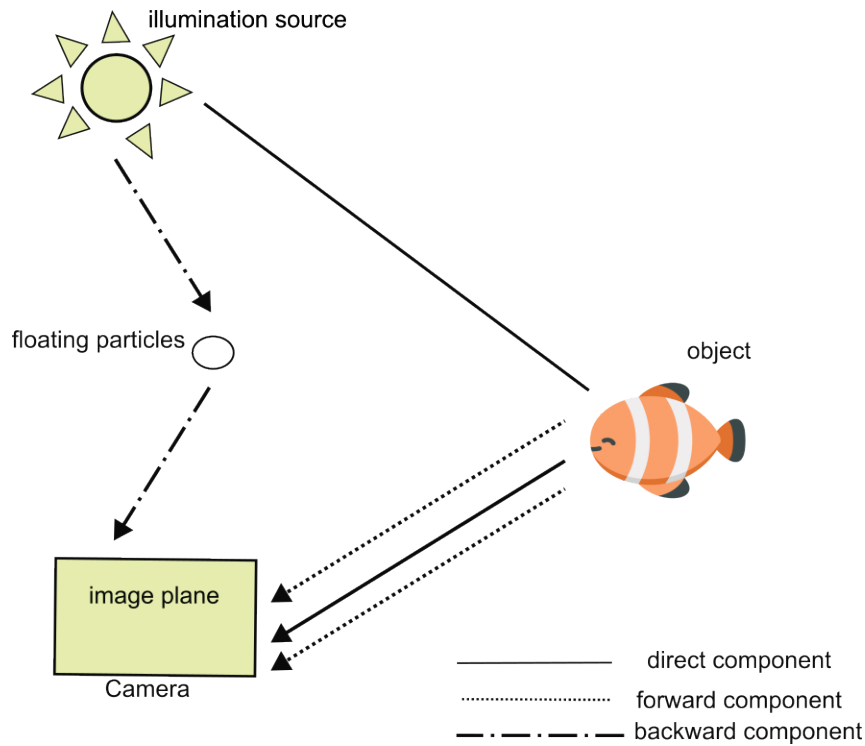


Figure 3.11: The three components of underwater optical imaging: direct component (straight line), forward component (dashed line) and backward scatter component (dash-dot line).

Therefore, the total irradiance E_T , as depicted in Figure 3.11, is

$$E_T = E_d + E_f + E_b$$

For how the terms of total irradiance are defined, see the work done by Jaffe et al. (1990) in [28], for the purpose of the colorimetry dissertation, it is sufficient to know what contributions lead to the formation of the underwater image and that the model partitions 3-dimensional space into planes parallel to the camera plane, and the radiation scattered toward the camera is computed superposing small volume elements weighted by an appropriate volume scattering function. In order to restore underwater images, as we said before, several parameters has to be known. Therefore, scientists have been engaged in carrying out measurement campaigns with the aim of determining these parameters. Between 1947 and 1948 Jerlov and colleagues took part of the Swedish Deep Sea Expedition [29] during which the first systematic assessment of the optical properties of the global ocean was conducted. After the expedition, they derive from the measurement of the downwelling irradiance at various depth ten optical classes, five oceanic and five coastal ones, known as Jerlov's water types, see Figure 3.12. From the measurements it

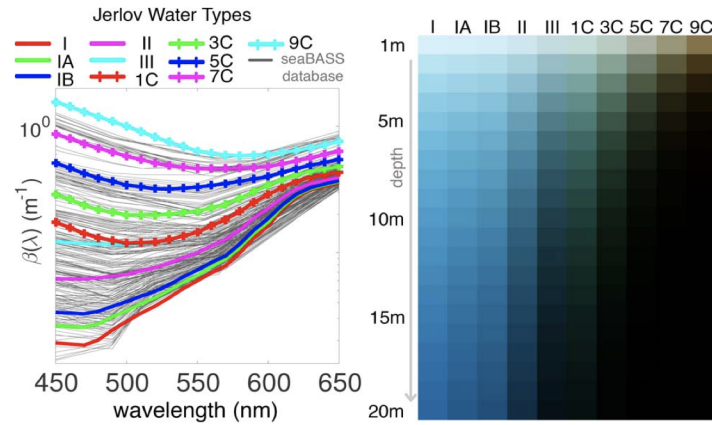


Figure 3.12: Jerlov water types. Each type is indicated with a colour and those types that are coastal are distinguished by 'C' letter. Gray lines represent 280 randomly chosen observations from the seaBASS database, which lay within the Jerlov's water types space. On the right is showed color table represents the Jerlov's water types.

was possible to determine the total attenuation coefficient $\beta(\lambda)$, which depends also on the type and density of the particles suspended in the water column. For that concerns oceanic waters, particles are almost always drifting algae (phytoplankton) with a specific spectral signature, attenuating the longer wavelengths, such as red colors, much faster than shorter ones, resulting in an overall bluish appearance. In the coastal oceans, due to the contribution of rivers or to agricultural runoff, optically active impurities cause short wavelengths to attenuate just as strongly as long ones. In order to fill up the space left free among the ten classes, a new and recent database called seaBASS was carried out [30] using novel in-situ measurements. The knowledge of wideband attenuation coefficients per color channel is essential to apply underwater image reconstruction methods. Akkaynak et al. in [31], based on the measurements of global attenuation coefficient taken by Jerlov and during seaBASS campaign, constrain the set of physically-feasible wideband attenuation coefficients in the ocean in order to feed image reconstruction methods with parameters that have a physical meaning. They calculate the space of valid wideband effective attenuation coefficients in the 3D RGB domain and find that a bound manifold in 3-space sufficiently represents the variation from the clearest to murkiest waters. In this fashion, it was possible to understand what are the potential values that the attenuation coefficients can take on, as well as it is also demonstrated that against the common single image formation model, which use the these optical water parameters to derive their corrections in order to achieve visibility enhancement and color restoration, the coefficients depend on the

imaging range and object reflectance, and they quantified the errors resulting from ignoring these dependencies, which are considered non-negligible. In general, single image reconstruction methods do not take into account of the dependency on wavelength attenuation, and therefore they avoid the need to estimate color-dependent attenuation; however it results in unconvincing results. Other methods use multiple images of the same object from several distances to recover channel dependent optical depth. But the basic estimation method for attenuation coefficients is to acquire an image of a known calibration target at known distances. It results in a major resource consumption since it requires external hardware and distance measurement. But none of these methods verified the results to be physically feasible. Indeed ill-posed computer vision problems benefit from constraining the set of solution, since it narrows the scope of the search of unknowns, as well as it increases the speed of the algorithm. Therefore the contribution of researches in this context it may considered important in order to delineate the existence space of the values of the attenuation coefficients related to underwater. Finally, it can be said that all methods using these attenuation coefficients derived from previous measurement campaigns use pre-existing values to feed into the image formation model. In contrast, there are others which do not take into account the knowledge of any parameter, such those that use information from depth maps [32], available in any photogrammetric dataset through SfM techniques, in order to apply the color correction model. In addition to this type of methods, Bekerman et al. (2020) in [33] developed a procedure able to restore image color based on single image analysis through the knowledge of water physical properties instead of depth maps. Another example within this category is the ACE model, developed by Rizzi et al. (2003) in [34], an unsupervised color equalization algorithm which is based on some adaptation mechanisms of the human vision system that translates into a particular lightness constancy and color constancy. A very promising source summarizing the qualities of several color restoration and enhancement algorithms, which also takes into account computational efficiency, was compiled by Mangeruga et al. (2018) in [35]. It provides several guidelines in the choice of the right algorithm depending on the underwater scene that we want to enhance as well as useful framework that contains several color enhancement algorithms together, so that they can be applied simultaneously on an exemplary database image and consequently choose which of the methods turns out to be more

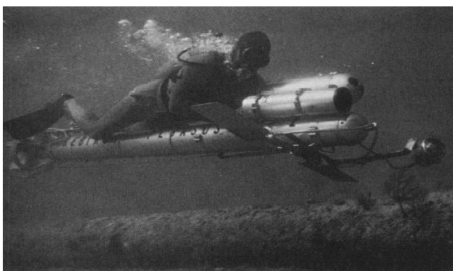
suitable. In addition, there is also a dissertation about the metrics used to rank enhancement color methods. The main gaps that still need to be filled are the lack of an official database suitable for evaluating the performance of these methods, which is able to collect all different image processing conditions in order to establish common quantitative as well as qualitative estimates. As well as, the lack of underwater imaging methods that take into account data from different sources as is the case, for example, in the marine animal vision system, which use their physical optics, photoreceptors and neurophysiological mechanisms. Last but not least, a color correction technique simple to apply but at the same time very efficient is the white balance. It can be applied in post-processing when we deal with raw images, or in processing placing a white panel in front of the camera before to start the surveying campaign, considering as stable all the optical properties of a examined area. Actually white balance is used to make the colors in a photograph appear as our eyes see them, since the color of what we see is influenced by the color of the light that strikes it, and in underwater case, it is not the white light of the sun but is actually a light with a blue hue. Modern cameras provides several option to set white balance, therefore it represents an expeditious technique to improve image quality.

Chapter 4

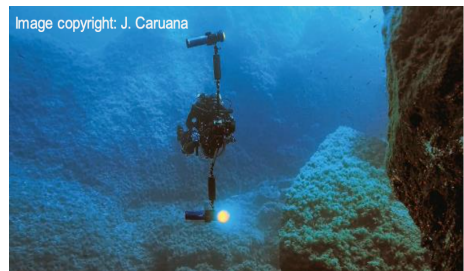
Unmanned Underwater Vehicles (UUVs)

4.1 Introduction

The evolution of Unmanned Underwater Vehicles arise from the need to have a stable and efficient platforms for underwater photogrammetry. Any means able to carry a single, stereo or multiple camera systems to capture underwater images would help the diver to explore the underwater environment, and in that sense starting in the second half of the 1960s, new vehicles appropriately designed for underwater monitoring are being built. Rebikoff (1968)



(a) The Rebikoff Photogrammetric PE-GASUS System.



(b) CCR diving with a DPV by Caruana et al. (2022).

Figure 4.1: An ancient DPV system next to a recent one.

was the first to design a diver propulsion vehicle (DPV) shallow water (up to 40m of depth)

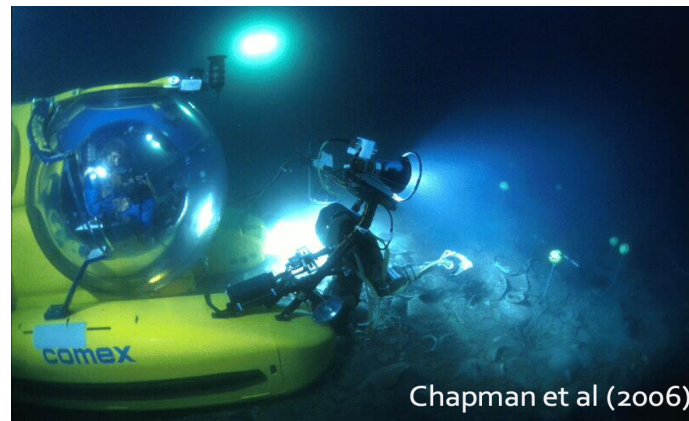
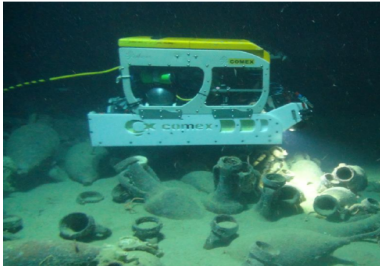


Figure 4.2: The submarine REMORA 2000 of COMEX during the Grand Ribaud excavation, 2000. Photo Frederic Bassemayousse.

open circuit scuba diving. It was immediately successful in producing the first and only really sharp and color-correct underwater photos of marine life, geology and archaeology. It was fully effective only because it was designed as a complete integrated system with each component subsystem, camera, corrected lens, strobe light, controls and housings simultaneously designed for each other to fully accomplish the required job. Nowadays, thanks to the technological development, Caruana et al. (2022) developed a Closed Circuit Rebreather (CCR) diving in shallow and deep water (up to 150m of depth) with a DPV, much lighter and more flexible. It was used as support tool to document the collapsed Azure Window, a natural arch situated in the west coast of Gozo (Maltese Archipelago). They demonstrate use of this application to document complex underwater geomorphology spread across a large area, and draw up a geomorphic assessment of the site and collapse event on the basis of this 3D-model. Alongside these methods used for shallow water, other means of exploration have been developed to reach greater depths (up to 600m of depth), as for instance, the Remora 2000 system that is a submersible system for deep water surveys developed by Chapman et al. (2006). This project improved the accessibility of underwater sites by generating 3D models for virtual and augmented reality applications (Figure 4.2). Up to this point we have talked about vehicles that are directly piloted by experienced personnel (manned vehicles), but without referring to a large category of vehicles, that are unmanned, born out of the need to decrease the risks associated with an underwater survey campaign as well as to provide tools that can perform a survey more systematically. Among these vehicles it exists those are remotely controlled by a



(a) Comex UUV-ROV developed by Drap et al. during an archaeological survey.



(b) Work class UUV-ROV developed by Menna et al. used for Oil&Gas inspections.



(c) GuPho prototype developed by Torresani et al. and used by a diver for benthic surveying.

Figure 4.3: Three examples of UUV that work at different depth and with different aims.

cable and those wireless controlled. Therefore, a series of vehicles corresponding to this category as the depth increases will be shown in the following. Rende et al. (2015) presented a towed system for shallow water mapping of *Posidonia Oceanica*, since this kind of plants are considered as a bio-indicator of the quality of coastal marine waters (up to 50m of depth). Chemisky et al. (2021) assembled a microROV opto-acoustic system for mapping underwater targets (up to 100m of depth). Drap et al. (2015) built an observation class ROV based optical system for deep archaeological application (up to 1000m of depth) (Figure (a) 4.3). Menna et al. (2019) designed a work class ROV based optical system for subsea inspection in the Oil&Gas (Figure (b) 4.3), while Kwasnitschka et al. (2016) used a torpedo like AUV based optical system for large subsea area and Sigh et al. (2006) built a seabed like AUV based optical system for high resolution photomosaics. These last vehicles are able to reach large depths (up to 6000m of depth). Last but not least, it is worth considering underwater systems based on Visual SLAM technology, they represent a wide support for divers during the surveying. Indeed, it is not at all easy for a diver to orient himself in the underwater space, and these systems provide a support to diver in the sense that they can provide precious information during the diving,

as for instance, guidance on trajectory, GSD, and so through a visual feedback provided by a screen. Torresani et al. (2021) provided a guided and portable system for photogrammetric applications, which can reach a sub-millimeter length measurement accuracy (Figure (c) [4.3](#)).

In the following sections will be addressed the UUV topic, and in particular it will show the functioning of Blucy, an UUV developed by a teams of experts within University of Bologna.

4.2 Overview

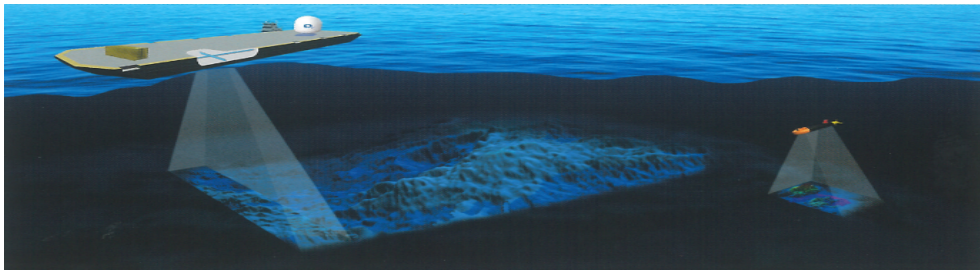


Figure 4.4: Implementing the latest technology is one of the pillars of the GEBCO Seabed 2030 project. The concept of an unmanned mapping barge, monitoring by satellite communication and equipped with an ultra-narrow beam deep-water multibeam (left), is just one of many ways technology could be used to generate new data. The sub-meter of detail sometimes needed to investigate small scale processes at the seabed is today only possible to achieve in the deep ocean using UUVs equipped with high-resolution high frequency multibeam systems.

In July 2022 it was released a new version of the General Bathymetric Chart of the Oceans (GEBCO) [\[36\]](#), which respect with the older one released in 2021, it provides 2.08% more of mapped seafloor; therefore, right now, it is estimated that only 23.4% of the seafloor has been mapped. GEBCO is a bathymetric representation of oceans and seas of the world; the project was born with the goal to create and provide the most authoritative publicly-available bathymetry of the world's oceans, under supervision of International Hydrographic Organization (IHO) and Intergovernmental Oceanographic Commission (IOC/UNESCO). The complete target is still far to reach (it was established to map the entire ocean floor by 2030), since it lacks about 70-75% of the global seafloor which is yet undetected, but the current improvements without any doubts they have been possible even with the support of Unmanned Underwater Vehicles or UUVs. Therefore to date, only a portion of the submarine ecosystems are mapped in detail, even if the oceans cover about two-thirds of the earth and about 37% of the world's

population lives within 100 km of the ocean [37]. Actually, the ocean is generally overlooked since we focus our attention on land and atmospheric issues, where surveying campaigns can be conducted without such difficulties that instead characterized marine environment.

From an International point of view to an European one, it is worth mention the recent initiative of the European Commission called Destination Earth; the project was born with the aim to simulate atmosphere, land, ice and ocean with unrivaled precision thanks to the computational effort that recent technologies can reach nowadays, as for instance exascale computational power. Oceans are within the Destination Earth project since they represent principal drivers in the climate context; therefore to hone climate forecasts a deep knowledge of them could provide a more accurate model. Indeed, the main purpose of the project is to create a Digital Twin (DT) of the Earth, where with DT it is intended an ultra-realistic, multi-scale, multi-physics simulation of a system to mirror its evolution; therefore having robust representation of ocean shape and dynamic as well as a ocean DT could provide better answers to our question related to global dynamics issues that affect our planet. Even in this case, the advent of the use of UUVs, for what concerns oceans, represents a key element to accomplish the task of create a DT of Earth.

The underwater infrastructure and oil&gas sectors need to have huge precision and operability in the marine environment, indeed activities as submarine cable laying and maintenance, marine mineral resources exploration and accurate submarine microtopography require an intensive technological development able to face marine environment hostilities. Therefore, in the last decades, the use of UUVs, moreover in these fields, it has undergone exponential growth thanks to the efforts made by researchers and to the increasingly advanced technologies, such as new materials, sensors and algorithms.

Within this context knowledge of the marine environment has and will continue to have impact on the future survival of humanity and his development in coastal areas, but even on the animals and, in general, living beings that inhabit our waters.

The main issues and difficulties closely related to the submarine environments cab be listed as follows:

- resistance to the water column pressure

- resistance to corrosion and encrustations
- poor visibility
- hermeticity
- navigation and positioning
- communication
- security and safety
- economic issues
- logistic issues

During submerged activities equipment has to be strong enough to withstand the weight of the overlying water, therefore it is necessary to use only rigid materials by looking for a trade off between weight and mechanical characteristics. Because of the electrical potential difference between salt water and metals, the latter undergo a degradation process, called corrosion, able to weaken themselves until they are practically useless; therefore the use of noble metals is preferred, such as titanium and stainless steel, but on the other hand they are characterized by a very high price with respect of standard metals. As the depth increase, sun rays have to cross over thick layers of water, therefore one can assume that the loss of visibility is proportional to the height of the water column (beyond depth of 1000m no light from the surface can be detected). In addition, in comparison with in-air case, the visibility is totally different since water medium, and all the content within it, interacts with sun rays by generating distortion effects which are able to degrade the quality of visibility. Usually equipment are built to work in standard conditions, that is when the medium in which they are used is the air, therefore it is necessary protect them from the water action. A lot of techniques exist to avoid water entrance, such in the case of electrical circuits or electronic devices, most of them are based on the use of o-ring trimmings and stainless steel containers, which are placed airtight in a hyperbaric chamber. In addition, underwater equipment has to be tested before every mission since it is possible that there is some imperfection that could cause water entrance. A body placed in the sea or ocean is

subjected to the action of the currents as well as to the water pressure, therefore it is necessary to develop some systems able to avoid drift effect and powerful enough to face water inertia. Most positioning systems are designed to work under standard condition, where there are not visibility problems and material corrosion; it makes difficult positioning in underwater condition, even because it is very cumbersome to materialize points that need to establish a reference system. In addition, the signals coming from satellites are severely weakened as soon as they touch the water surface; it makes satellite positioning below the water surface impracticable. Thus, an hybrid technique is used to solve this problem, since it can join the position obtained above water surface with that obtained under water surface; only in this way it is possible to obtain an absolute positioning. Signals, such as Wi-fi, GPS, Bluetooth and others, in underwater condition, due to power loss and consequent absorption by water, they do not work. As work for marine animals, the means of communication, in underwater conditions, it is the acoustic one; indeed, sound in water retains characteristics that make it usable for information exchange. Sure is that the content of these information packages cannot be very large and that the sound means is very noisy, but so far it is an excellent solution to solve, for instance, the problem of underwater positioning. When human being is involved in underwater mission, the risk you have to pay is huge, therefore only highly specialized personnel is approached; he must know underwater emergency procedures to limit the possible damage and he must have experience to answer in the best possible way to unforeseen incidental situation. One of the main factors that makes hostile the underwater environment is the economic one, indeed, when it organizes underwater missions a huge amount of money has to be spent to allow the entire system to work: crew of the ship, underwater personnel, navigation systems, sensors, fuel, control center, lifeboats, and so on. In general it could be say that submarine equipment is very expensive. It is not possible to rely on permanent infrastructures that are available only in small numbers and finally the particularly harsh and hostile environmental conditions make logistic element very important: it is necessary to rent a boat, unless you own it, at each mission, without being entirely certain of the weather conditions you will encounter. In short, the risk of renting a boat with unfavorable weather conditions, which do not allow the success of the mission, is high; therefore this randomness in the mission preparation introduces a further element that

has to be considered in the logistic of an underwater campaign.

The harsh and unstructured nature of the underwater environment causes significant challenges for UUVs. However, with recent advances, this field is progressing at an unprecedented rate. Thus, it seems clear that taking into account all the issues argued above, the advent of UUVs is a means to face the hostilities of the underwater environment since they are designed properly to face them.

UUVs can not only replace dangerous underwater work typically undertaken by people but also conduct comprehensive exploration as well as exploration at depths that cannot be achieved by ordinary diving technology. They can be an important tool for the study of oceans and can help lead the development of oceans in a new era, indeed they are now being used for a variety of tasks, including oceanographic surveys, demining, and bathymetric data collection in marine and riverine environments. Thus it is possible to list the main application areas in which UUV are used:

- Military, defense, and coastal security applications
- Environmental monitoring systems
- Scientific applications
- Ocean mining and oil&gas industry
- Other applications: hull inspections, commercial harbor security missions

To each type of application it maps a different kind of UUV. As for instance, for military and defense, torpedoes shaped UUVs are used since they are designed considering the computational fluid dynamics (CFD) model and therefore they have to be performing from the point of view of navigation and rapidity, as well as being able to accomplish underwater mechanical actions. For what concerns scientific and environmental application, priority is given to endurance and payload, since a lot of sensors have be equipped on the UUV; in this case UUVs can have simple shapes, such as rectangle and cubic ones, but they can also have curious shapes who try to emulate the shape of marine animals (jellyfishes, stingrays, and sharks), called biomimetic

structures, since they represent the best examples of mobility in environments such as this. This latter type of UUV are used in the study of ocean currents, ocean volcanoes, tsunami detection, deep-sea biological phenomena, and migration and changes in major ecosystems [38]. In the case of ocean mining and oil&gas industry, the main private sector that uses UUVs, they are important either endurance and payload either the possibility to work remotely at deep depths.

One of the most important factors taken into account in the manufactured material of UUVs is the density, indeed material with low density, as weight increases, promote buoyancy, and this means that more force is required to submerge the UUV. The materials often used to built UUVs are: aluminium, acrylic, PVC (polyvinyl chloride), PE (polyethene), PMMA (polymethylmethacrylate), PLA (polylactic acid), and ABS (acrylonitrile butadiene styrene). Obviously, the deeper the UUV reaches, the greater the external pressure is during the UUV's operation. The UUVs that have a torpedo shape are made of aluminium, with chambers filled with pressure air, and their frontal sections are made in plastic. With the advent of 3D printers and ease with which custom parts can be obtained, the majority of UUVs therefore consist of carbon polymers, from which those with the best characteristics in contact with water are chosen. In addition the use of 3D printers allows to reduce the fabrication cost of an UUV. The energy by which UUVs ensure all their functionalities, including motion, derive from the use of battery (lithium, lead) or fuel cells, who are able to produce electricity, but also through the use of simple motors, diesel-powered, in the form of mechanical energy. In addition, in a context where energy, and therefore durability of UUV is a main issue, as for instance in the case of glyders [39], some researchers have devoted themselves to the supply of energy from alternative sources, such as ocean thermal and salinity gradients [40].

Over the last few decades, there has been significant progress in a number of technologies critical to these vehicle systems. Even with these advances, a number of research issues remain. Autonomy, endurance, navigation, sensors/sensors processing, underwater robotic manipulation and excavation, and acoustic communications remains as the enabling technologies for these systems [41]. In particular, the communication issue still suffer from many limits such as:

- small bandwidth, which means that only short content messages can be shared: no pics,

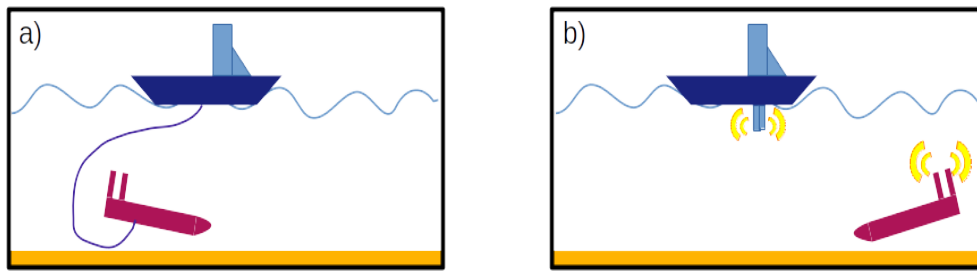


Figure 4.5: Two navigation modes: (a) ROV and (b) AUV.

no video, and so far. But, for instance, National Marine Electronics Association (NMEA) messages can be shared. They are currently the most common data format supported by positioning equipment [42]

- low data rate, which limits the number of messages that can be shared.
- high latency since the speed of sound in water is 1500 m/s, compared with the speed of light
- multipath transmissions due to the presence of free surface (up) and sea bottom (down)
- unreliability since there are frequent data loss in transmissions
- variable sound speed due to the relationship with the temperature

Notwithstanding the deficiencies referred to above, acoustic localization has become the most important and reliable technical means to solve positioning and tracking in a variety of applications and fields.

The main difference on which is based an UUV classification is the navigation system; indeed, there exists three types of piloting an UUV, see Figure 4.5:

- Remotely Operated Vessel (ROV): the ROV is a vehicle joined to an umbilical cable coming from a vessel; usually it is driven from navigation computer to which is linked, which is able to visualize that the vehicle are observing.
- Autonomous Underwater Vehicle (AUV): the AUV is an independent vehicle, equipped with an automatic navigation system. It does not need to be piloted.

- Hybrid: an UUV that can be piloted either in ROV either in AUV.

When dealing with AUV mode, because of the modest bandwidth of acoustic transmissions, small amount of real-time telemetry data is transferred, and this greatly complicates navigation (e.g., for obstacle avoidance). Real-time visual inspection from the control panel boat is not possible. Once the vehicle has been retrieved, the data can be downloaded. While with the connection via fibre-optic umbilical cable (ROV), the operator can see the acquisitions captured by the navigation camera in real time and make manoeuvre decisions. For what concerns sensors installed on the UUVs, a classification based on the function they perform can be applied:

- Internal: main function is to measure the changing parameters of the UUV during its movement. Its position error will continue to accumulate with continuous movement of the process; therefore long-term direct navigation data cannot be collected. Such sensors include inertial navigation devices, Doppler odometers, gyroscopes, speedometers, and odometers [43]. An overview of the main internal sensors is showed in Table 4.1.
- External: main function is to obtain the characteristic information from the surrounding environment. They do not affect with the unbounded position error growth, therefore they are used as support to bound the error of the internal sensors [44]. Such sensors include infrared, sonar, ultrasonic, and machine vision sensors [45]. An overview of the main external sensors is showed in Table 4.2.

In the case of AUV, there exists a further classification based on the navigation and localization techniques, which help to fix better in mind all possible kinds of sensors that can be installed on-board of an UUV, see Figure 4.6:

- Inertial/dead reckoning: when the travel speed of AUV is known, new position can be estimated by integrating consecutively in time, thanks to instruments able to detect movement: Doppler velocity log (DVL) together with inertial systems and compass [53].
- Acoustic transponders and modem: this category include all the techniques based on measuring the time of flight (TOF) of signals from acoustic beacons or modems to perform

Name	Description
Compass	It provides a globally bounded heading reference. It works by measuring the magnetic field variation in three referential axes, that are subtracted from the Earth's magnetic field. Nevertheless it is very sensitive to noise caused by others sources like the operation of other sensors, motors and others.
DVL	Doppler Velocity log uses acoustic measurements to capture bottom tracking and determine the velocity vector of an AUV moving across the seabed. It determines the AUV surge, sway and heave velocities by transmitting acoustic pulses and measuring the Doppler shifted returns from these pulses off the seabed [46]. DVLs will typically consist of 4 or more beams [47], 3 beams are needed to obtain a 3D vector.
IMU	<p>Inertial Measuring Unit uses a combination of accelerometers and gyroscope to estimate the vehicle's orientation, velocity, and gravitational forces.</p> <ul style="list-style-type: none"> • Gyroscope: Measure angular rates. For underwater applications the following two categories are mainly used: <ul style="list-style-type: none"> - FOG: In Fiber Optic Gyroscope light is passed through a series of fibre optical cables in different directions. The angular rates are determined based on the phase change of the light after passing through the fibre [48] - MEMS: In micro-electromechanical systems an oscillating mass is suspended within a spring system. Rotation of the gyroscope results in a perpendicular Coriolis force on the mass which can be used to calculate the angular rate of sensors [49] • Accelerometer: Measures the force required to accelerate a proof mass. Common design include a pendulum, MEMS and vibrating beam among others.
Pressure sensor	Underwater depth can be measured with a barometer or indirectly through the measure of others sensors.

Table 4.1: Outline of the main internal sensors usually equipped on a AUV belonging to the proprioceptive navigation category.

Type	Description
Optical	<p>Monocular and Stereo cameras: they capture images of the seabed and then match these images to navigate. Within this context, the process of determining the AUV pose by analyzing subsequent photographs it is called visual odometry. This can be achieved through SfM. In underwater environments there are limitations for the optical systems such as the reduced range of cameras, susceptibility to scattering, and inadequacy of lightning. As result, RGB color camera are more commonly installed on hovering AUVs, with respect to Stereo cameras, because they can get close to objects of interest. In addition, visual odometry and feature extraction relies on the existence of features. Therefore, optical underwater navigation methods are definitely well suited to small-scale mapping and feature-rich environments [50].</p>
Acoustic	<p>Sonar: The production of images through sonar predates the advent of AUVs; therefore they represent a robust technology. Sonar are designed to work at specific frequencies depending on the range and resolution required. Usually to small frequencies correspond small ranges but an appreciable resolution and vice versa.</p> <ul style="list-style-type: none"> • Imaging Sonar: the intensity of the acoustic returns from the seabed of this fan-shaped beam also depends on the bottom type and is recorded in a series of cross-track slices [51]. When mosaicked together along the direction of travel, these assembled slices form an image of the seabed within the swath of the beam. Hard objects coming out from the seabed send a strong return which is represented as a dark image. Shadows and soft areas, such as mud and sand, send weaker returns which are represented as lighter images • Ranging Sonar: with multibeam instead have a single transducer pointing down there are multiple ones from an array of transducers. The sound bounces off the seabed at different angles and is received by the AUV at slightly different times [52]. The received signal are then processed onboard the AUV, converted in water depths, and make as a bathymetric map. The multibeam resolution depends on its transducer quality, frequency and distance from the seabed
Magnetic	See compass in Table 4.1

Table 4.2: Outline of the main external sensors usually equipped on a AUV belonging to the geophysical navigation category.

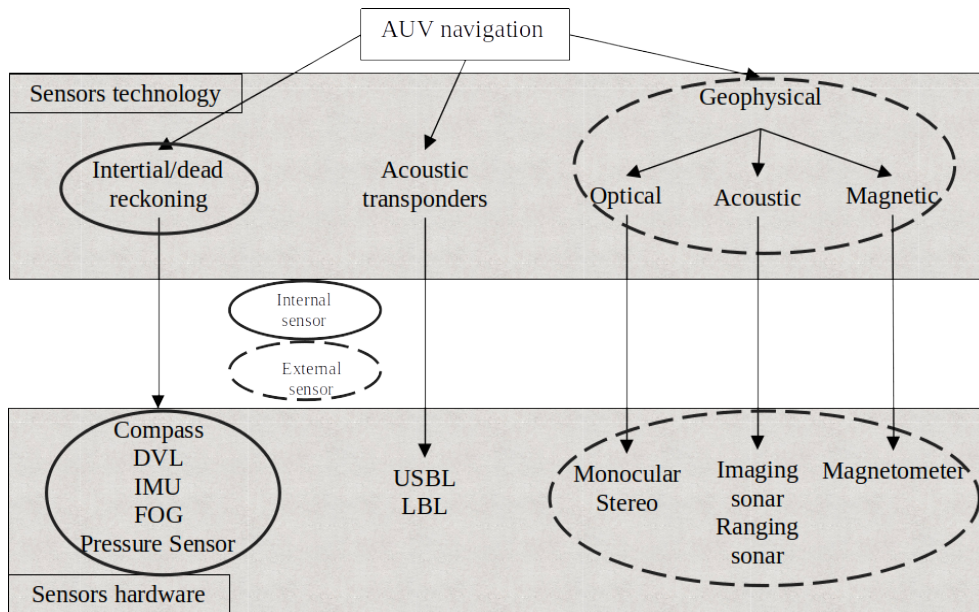


Figure 4.6: Outline of underwater navigation classification. It may happen that some methods are combined to provide increased performance, which is the case of multisensory information fusion.

navigation.

- Geophysical: info coming from external environment are used as a reference for navigation, therefore through sensors and processing which are able to identifying and classifying some environmental features.

It should be notice that there is a difference between navigation and localization. Navigational accuracy is linked to the precision with which the AUV guides itself from one point to another. Localization accuracy is the error in how well the AUV localizes itself within a map. One of the most challenge in underwater environments is the navigation and localization, and consequently to establish an absolute reference system, since there is a huge difference with respect land environment case, where satellite signal is available. The scenario that comes closest to the underwater one is that of the indoor navigation and localization [54], indeed it suffers from the same shortcoming, that is the lack of satellite signal. In underwater environments the electromagnetic waves emitted by satellites are absorbed when they come in contact with the water surface, therefore a global positioning system (GPS) receiver, placed underwater, cannot capture them [55]. The joint between underwater devices, who are used to establish a local positioning system, and a global positioning system allow to geo-reference the UUV. At the

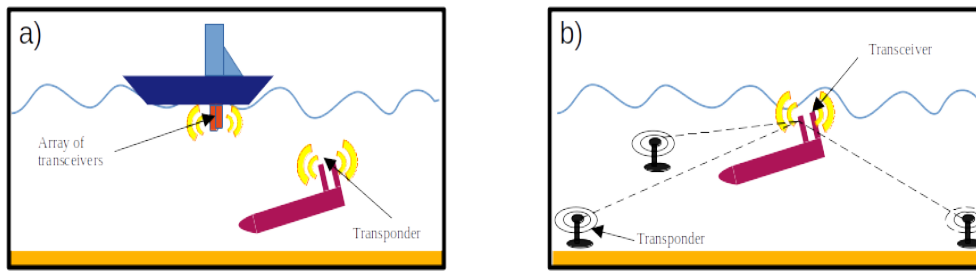


Figure 4.7: (a) USBL and (b) LBL acoustic navigation systems.

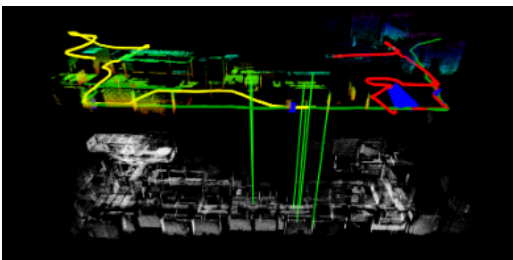
time when this joint is made, since the UUV is fixed inside an absolute reference system, all its products, such as photos, videos, and surveying objects inherit its reference system and so they can be localized within a map. Usually, these underwater devices are represented by ultrashort baseline (USBL) and long baseline (LBL) who are acoustic systems based on time of flight (TOF) of acoustic signals, see Figure 4.7; acoustic waves are appropriate for underwater propagation due to minimal attenuation. An LBL positioning system consists of two parts: a transceiver mounted on-board the UUV and an array of three or more transponders fixed to the seabed [56] in known positions. In most cases, the transponders are globally referenced before the start of the mission by a surface ship. In a normal operation, the transceiver sends out an interrogation signal, and the transponders reply in a predefined sequence. The two-way travel time of the acoustic signals is used to determine ranges. Some limitations of LBL are the cost and time associated with setting up the network and restricted area in which the UUV can move by receiving the acoustic signals. A solution to avoid the seabed installation of the transponders is the use of multi-GPS-intelligent-buoys positioning which can reduce installation costs and the need for the recovery of the transponders. These buoys are able to receive the electromagnetic satellite signals and to translate them into acoustic signals. Instead, USBL navigation allows to localize an UUV relative to a surface ship, without the need to setting up a seabed network [57]. This type of configuration works similarly to the standard LBL configuration; but the vehicle has multiple acoustic transceivers, because they must determine not only at which distance they are from the transponders but also the angle with which the replies of the signal arrive. An acoustic pulse is transmitted by the transceiver and detected by the subsea transponder, which replies with its own acoustic pulse. This return pulse is then detected by the transceiver. The time between the transmission of the initial acoustic

pulse and the reply is then measured by the USBL system and it is converted into a range. An USBL head contains an array of transducers separated by a baseline from a centimetres to decimetres. When the subsea transponder reply pulse is detected by the different transducers, a phase-differencing method within the transducer array is applied to estimate the direction of the subsea transponder, also called bearing. When range and bearing are known, the relative position of the subsea transponder is then computed. USBL positioning system offers positioning accuracy and easy operation in the short range, but error accumulates with increasing distance, generally from 1-2% of the slant distance for basic equipment and up to 0.06% for the ultimate USBL systems. The error also depends on assistance of external sensors, such as depth ones. Finally to compute the subsea position in an absolute reference, as we said before, the USBL system must be coupled to GNSS, and a gyro compass that will compensate the orientation of the surface ship at the detection instant reply.

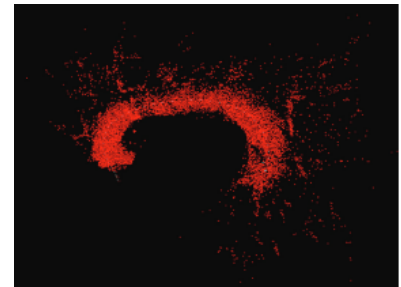
Localization and navigation is performed through multi-sensory information fusion, as described in Figure 4.6: one has to imagine that localization and navigation is based on the info coming from different sensor technologies at the same time. Indeed, with the increasing demand of UUVs and the further research conducted by different countries, localization and navigation requires more and more complexity and accuracy; thus, a single sensor or a combination of two or more sensors do not match the needs of nowadays. This gap is filled with the development of multi-sensory information fusion positioning navigation technology. It uses several sensors of the same type or multiple measurements of a single sensor to obtain multiple independent observations. Subsequently these observations are integrated and processed in such a way as to that they can improve the accuracy of position and velocity estimates of the tracking object, improve the positioning accuracy, and achieve optimal navigation. After the fusion of multi-sensory information from multiple observations processes, each observation process may undergo one or more fusion algorithms. The most commonly used fusion algorithms are those based on Bayesian network [58], Kalman filters [59], statistical decision theory [60], Dempster-Shafer evidence theory, clustering analysis, voting methods, neural network, and fuzzy set theory. These fusion algorithms are able to predict the data obtained by sensors, therefore they support the UAV decision making pipeline to avoid artifacts and to detect some outliers. In

addition, they are capable of recognizing noise and manage it, as well as, ensuring stability and robustness of the system; if one or more sensors do not work, this kind of algorithms are ready to handle this eventuality and arrive at the simplest and most practical solution avoiding data loss from working sensors.

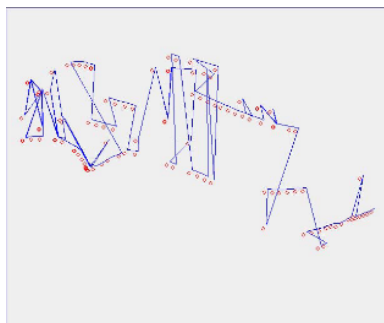
Nowadays, beside the use of fusion algorithms, another technique is developing called Simultaneous Localization and Mapping (SLAM), which requires a large amount of computation to achieve real-time positioning. It is a positioning and navigation technology that uses multi-sensory information fusion technology. An UUV moves from one point to another in a completely unknown environment: the UUV uses the acquired information coming from sensors to estimate and position itself while building an incremental map [61]. The UUV use its sensor's perception of the environment (internal sensors) to continuously update a map during movement and collect information to further improve the accuracy of the map; the rules behind which the UUV builds the map determines the type of the map, see Figure 4.8: raster, characteristic, and topology map. A raster map is a mosaic of equally sized raster that together



(a) Raster map in black and white with three trajectories (green, yellow and red) estimated in real-time in [62].



(b) Characteristic map built with the algorithm proposed in [63].



(c) Topological map generated by the underwater rover ROVFURGII in movement described in [64].

Figure 4.8: The three types of real-time maps produced during SLAM activities.

represent the environment. Each raster has a binary value to represent the state of the raster, therefore, if there is an obstacle or not. These kind of maps are built using light detection and ranging (LiDAR) sensors, who can reach very high accuracy, even in underwater environment. The drawback is that it is expensive and higher requirements for installation [62]. Characteristic map or feature map is a vision based SLAM technology which use a camera as a sensor, therefore it less expensive and lighter than the previous case. Indeed, image acquisition formation is more abundant and feature differentiation is higher. The disadvantage is that the real-time processing of image formation requires very high computational power, a problem that nowadays is facing with the support of graphics processing unit (GPU). Through image analysis, and thus through the recognition of features among different images, it is possible to recognize points, lines and polygons that describe the surrounding environment. Then these information are used to determine the relationship between the UUV and the environment to achieve the positioning of the UUV [63]. The pose of the UUV is estimated based on difference between feature positions and observed positions. For what concerns topology map, we can imagine that it consists of location nodes and connection lines: nodes represent the location of the characteristic points in the environment, while connection lines between nodes represent the path information between the characteristic locations. Each node contains information regarding the movement of one node to another one, including topological information and previously identified sections [64]. A hierarchical structure can be accommodated on top of this "behaviour graph", where nodes at one level of the hierarchy represent sets of nodes in lower levels. The meaning of nodes and lines in a topological map varies according to the application as well as the algorithm used to build them. Nevertheless, they are common elements to most of the topological map descriptions, namely, the use of the sensory input descriptions in order to identify nodes, connectivity relations among the nodes in the map, and local metrical information associated with lines in the map.

In conclusion, the combination of hardware and software has created unlimited possibilities for SLAM: a comprehensive technology and a complex system that has attracted the continuous exploration and application of researchers. Indeed, with the continuous development of science and technology, there is an increasing demand for intelligent UUVs, in particular in terms

of accuracy and speed. For these reasons the attention, in the field of automatic guidance systems, is focused on SLAM technology: map-bases multi-sensory information fusion of the main algorithms is used to build a precise positioning and navigation system for AUV in an underwater complex environment. One of the important component of study is the sensor system, since SLAM use sensors to take its decisions during movement; it plays an important role in collecting environmental information, building environment maps, positioning, navigation and path planning. Therefore further studies on sensor system has to be conducted in terms of improvement its accuracy, operability and integrability; this could allow to increase SLAM technique performance and reliability. Last but not least, it needs to note the latest research frontier related to UUVs, namely, the multiple underwater robots' coordination of positioning and navigation. Cooperative positioning and navigation of multiple UUVs means that they can share information through underwater acoustic communication and further improve their positioning accuracy through information fusion technology [65]. Some of the advantages of implement a multi-UUVs navigation system are low cost, high efficiency, strong fault tolerance and reconfiguration [66]. Indeed, in cooperative navigation there exists two paradigms of navigation: parallel and pilot. In the parallel one, each UUV has exactly the same function and structure of the other members, therefore each component of the UUV teams plays the same role of the others. Instead, in the pilot one, it can be found a master-slave relationship, in which there is a pilot UUV equipped with high-precision positioning and navigation sensor system, in contrast with to the others that are equipped with low-precision positioning and navigation equipment. The pilot UUV sends out its position through a sound signal pulse and the other following UUVs receive the position signal and compute the relative position. This allows to have less expensive cost and a redundancy in the system that increase the global measurement accuracy. In particular, the position error associated to the underwater acoustic communication delay decrease, as well as, the probability to have a fault. Actually, only the pilot paradigm of navigation and localization is implemented, since in terms of fault tolerance, the parallel approach does not solve the problem.

The global UUV market size is projected to grow from USD 2.0 billion in 2020 to USD 4.4 billion by 2025, at a CAGR (compound annual growth rate) of 16.4% from 2020 to 2025 [67].

The market is driven by various factors, such as the growing number of deep-water offshore oil&gas production activities and increasing maritime security threats. The UUV market includes major players Lockheed Martin Corporation (US), Saab AB (Sweden), Kongsberg Gruppen (Norway), Northrop Grumman Corporation (US), Oceanering International, Inc. (US). These players have spread their business across various countries includes North America, Europe, Asia Pacific, Middle East, Africa, and South America. COVID-19 has impacted their businesses as well. Industry experts believed that COVID-19 affected UUV production and services by 25-30% globally in 2020. Work class vehicles is the product type segment contributing the largest share of the remotely operated vehicles market. Work class ROVs with robotic arm manipulators are used to grasp objects, pumps, and brushes for cleaning operations as well as for underwater pipeline or deep water rig construction. The electronic instruments used in work class ROVs include underwater cameras and lights; acoustic positioning instruments; Conductivity, Temperature, and Depth (CTD) recording instruments; tracking systems; and side-scan, bottom scan, and multi-beam sonar. The electric propulsion is the fastest-growing segment of the UUV market. Electric systems contain lithium-ion batteries that are used in ROVs. The lithium-ion batteries are used in small ROVs for increased endurance and greater operating range. These batteries are commonly used in combination with other propulsion systems for improving the efficiency and reliability of ROVs.

There are a certain number of web-based resources, either public either commercial, which provide significant and a better understanding of the UUV technology. Table 4.3 lists the main online services in which a newbie interested to the topic or simply a curious person or an expert in the field can take cues for further research.

4.3 The case of Blucy

Usually project missions for surveying benthic zone are conducted using trawl nets, therefore with the removal of the individuals from their natural environment. Subsequently through a visual census analysis, that consists of the identification and counting of species, it can estimate the variety, number and even common sizes of the detected species [68]. This kind of technique

Name	Description	Link
Unmanned Underwater Vehicle Simulator	It is an UUV simulator. An open-source package containing the implementation of Gazebo plugins and ROS nodes necessary for the simulation of unmanned underwater vehicles, such as ROVs and AUVs. It is a software very useful to design and control, from scratch until the end of life, of a personal UUV. It a research prototype, originally devveloped for the EU ECSEL Project 662107 SWARMS.	https://uuvsimulator.github.io
IEEE Marine robotics	The Marine Robotics Technical Committee try to set up a major contribution to understanding large scale problems related to the spread of UUVs. Emerging marine robotic developments afford scientists advanced tools to explore and exploit the oceans at an unprecedented scale, in a sustainable manner. Priority research areas are: Actuation and Sensing Systems, Communication, Manipulation, Interaction, Guidance, Mission control systems, Localization, Multi-vehicle coordinaiton, and Persistent monitoring.	https://www.ieee-ras.org/marine-robotics
AUVAC	Autonomous Undersea Vehicle Application Center brings together academic, private sector, and government organizations in support of Autonomous Undersea Vehicles (AUVs), in order to advance AUV system technology, promote AUV interoperability, and increase AUV availability in support of the international national ocean community needs.	https://auvac.org/
Uwe R.Zimmer	A list of AUVs and developers.	http://www.transit-port.net/Lists/AUVs.html
OET	Ocean Exploration Trust (OET) and the Nautilus Exploration Program aim to explore the ocean, seeking out new discoveries in the fields of geology, biology, maritime history, archaeology, and chemistry while pushing the boundaries of education, outreach, and technological innovation. Their expeditions center on scientific exploration of the seafloor identifies priority regions and phenomena, and shares their expeditions with explorers around the world via live telepresence.	https://nautiluslive.org/
Energy Dais	Listo of top AUV Manufacturers	https://www.energydais.com/list-of-top-auv-manufacturers/

Table 4.3: Web-based resources on UUV. Accessed: 2022-07-18.

is considered invasive and requires a team of divers. There exists several areas, protected areas, where this technique can not be performed, since it may preclude the development and growth of the species that live there. A solution that can solve these limitations is the use of an UUV, indeed it does not require the intervention of divers, and through the use of an optical sensor the need to remove individuals from the seabed is abolished. The possibility of carrying out a survey, at a safe distance using a vehicle without an operator and with reduced procedures and cost compared to the past, has been particularly attractive and has allowed the development of a large number of UUVs. Hopefully, thanks to the ongoing technological and knowledge advances, UUVs could already in the next decades reach a level of reliability, performance, efficiency, and economy of use similar to that which today belongs to their terrestrial twins. Thanks to the images provided by an UUV it may be possible to census numerous benthic species, as long as the visibility conditions are good. Otherwise, it can be used techniques based on acoustic technologies, such as multibeam echosounder (MBES), which accurately map and provide an automatic unsupervised or supervised classification of benthic communities without the use of invasive techniques [69]. At one time, the recognition of identified species required a great amount of time, as it was carried out at the expense of human operators who, photo by photo, identified and counted specimens belonging to the different species under examination. Today, instead, thanks to the computer vision techniques this task is accomplished by machine learning algorithms, that are able to identify specimens automatically [70], even through high-resolution satellite images; in this fashion time related to detection is exponentially reduced. Underwater ecosystems present greater challenges and remain largely unexplored, but the need for a reliable and up-to-date information motivated the birth of the Interreg Italy-Croatia SUSHIDROP project (Sustainable fisheries with DRONES data Processing) [71]. The aim of the project is to map ecosystems for sustainable fishing and produce several digital models of the underwater environment; in order to achieve these goals a prototype of an UUV, named Blucy, has been designed and developed [72], see Figure 4.9. Blucy was built from scratch in order to combine the best possible features in terms of positioning and surveying while maintaining a low total cost, in Table 4.4 the main Blucy's features are shown; in order to carry out exploration and survey activities of submerged environments, the use of an appropriate

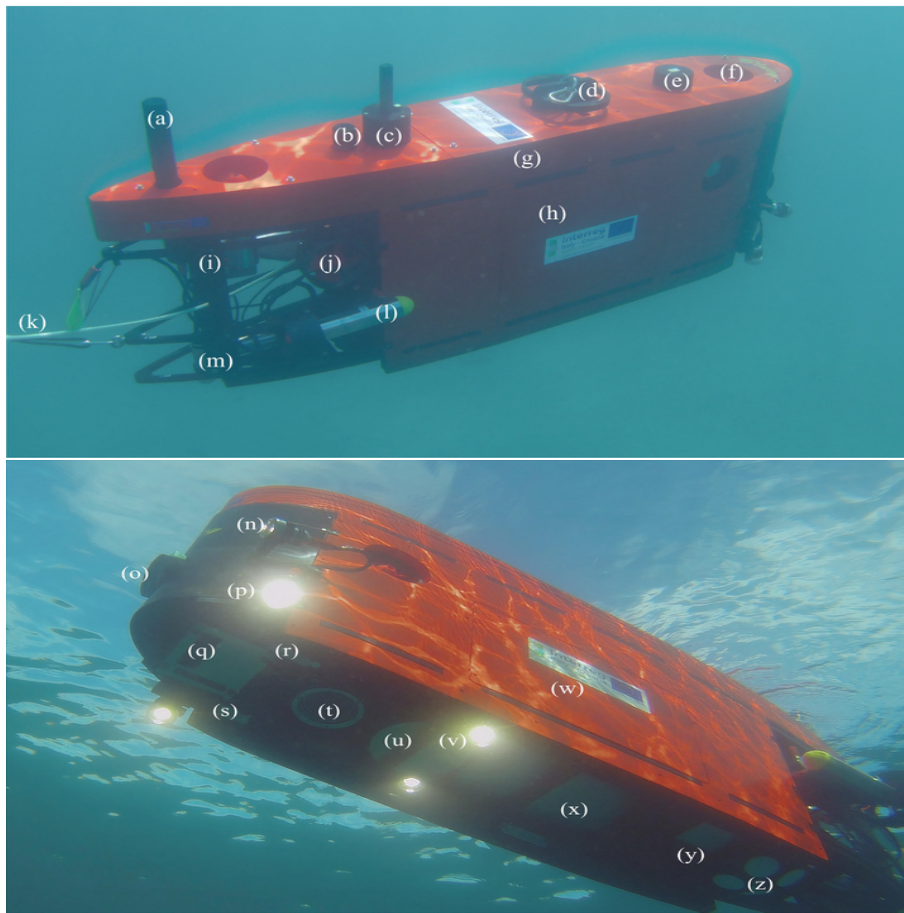


Figure 4.9: Main Blucy components: (a) Wi-Fi communication. (b) USBL transponder. (c) Radio communication. (d) Hook. (e) AHRS: GNSS and INS. (f) Thruster: vertical. (g) Main buoyancy foam. (h) Lateral buoyancy foam panel. (i) HDPE structure. (j) Thruster: lateral. (k) Fiber optical cable. (l) Thruster: longitudinal. (m) Altimeter. (n) Frontal LED lights. (o) PilotCAM. (p) Adjustable LED lights. (q) MBES sonar head. (r) MiniSVS. (s) MiniCT. (t) BottomCAM. (u) MBES computer. (v) Bottom LED lights. (w) Navigation, Guidance and Control (NGC) computer. (x) UUV 24V Battery. (y) FOG. (z) DVL. As described in [3].

subaqueous vehicle position estimation is essential. Indeed, only by accurately knowing the position of the vehicle and its sensors it is possible to ensure one of the main principle of surveying, namely the *repeatability*; when monitoring benthic areas is essential to know precisely where the survey is conducted in order to come back there in the future and analyze the differences between two different temporal instants in terms of control variables that have been designated as environmental indicators, usually available as metadata for the correct scale and georeferencing of spatial information. The possibility to compare several temporal instants, of the same area, allows to know the evolution of the submerged environment in such a way as to monitor it and in case make decisions to safeguard it. One of the most exemplary cases using this technique is the volumetric control of coral reef bommies [73]; through the use of SfM

Features	Notes and description
Main dimensions and characteristics	
Length L	2000 (mm)
Breadth B	630 (mm)
Height H	1050 (mm)
Main propulsion units	2 (+2 additional)
Transversal propulsion units	2
Vertical propulsion units	2
Maximum operating depth D	300 m
Maximum Weight W	215 kg
Design Power Consumption	500 W
Velocity range U	0,5-1.5 m/s
Standard mission endurance	12 h
Materials	HDPE, Stainless Steel, PVC Foam, ABS
Other features	
Underwater Navigation	Altimeter, Acoustic depth meter, DVL, FOG+IMU
Underwater Positioning	USBL(300m of operativity-distance in width and depth)
Cameras	2
Led lights	4
Surface Navigation	GPS
Surface Communication	Wi-fi and Radio Modem
Electronics	PC/104- based modules
Reserve Payload Ws	35 kg

Table 4.4: Blucy's general technical characteristics

technique is possible to detect coral bommies and compute the volume associated to these in order to have a quantitative idea of their growth during years. The accurate estimation of object positions under examination besides temporal instant in which they are detected are important information in order to build a time series of the states of the submerged environment and the species that live there. For what concerns synchronization of time, all sensors installed on Blucy and consequently all data produced by sensors are synchronized using UTM timestamps: fundamental for the correct alignment of the acquired data and to guarantee a high quality of the following processes. However a further difficulty in the synchronization of time is the frequency: not all sensors detect with the same frequency; some of them can detect with high speed frequency while other, since of their operation, they cannot reach high frequencies. Therefore it is necessary decide a method through which follows an absolute temporal line, maybe by interpolating or by mediating on time using a prederminated period of time, which is resulted suitable for the sensor global system. For what concerns positioning system a great deal of attention has been paid to the different stages in which the UUV will find, namely superficial and submerged one with respect to water. All information regarding the prototype was derived from the work done by Lambertini et al. [3]. In a functional context Blucy consists of the following modules or subsystems, which are showed in Figure 4.10:

- Navigation-Guidance-Control (NGC) module and communication (NGC-C) which include: i) A Central Processing Unit (CPU) for plant management and the navigation, guidance and control system; ii) a CPU for mission control management; iii) the servo controllers of the motors; iv) the multiplexer for the data/video transmission through the optic fiber cable; v) AC/DC converters and on/off.
- Navigation and communication underwater sensory system (U-NavCom) which includes: i) a Fiber-Optic Gyro; ii) a Doppler velocimeter that measures the linear velocity of the vehicle; iii) a mini-SVS that measures depth; iv) an acoustic altimeter; v) an acoustic transponder USBL/modem.
- On-board surface communication sensory module (S-Com) which includes: i) a Wi-fi/radio communication system between the vehicle and the vessel; ii) the power dis-

tribution electronics

- On-board surface navigation sensory module (S-Nav) which includes an integrated unit GPS-AHRS.
- Propulsion system (Prop-S) which includes: four helix horizontal propellers and two helix vertical propellers with brushless motors.
- Auxiliary system (Aux-S) which includes: i) a pilot camera; ii) a nadiral camera; iii) LED spotlight lighting systems (L).
- Scientific instrumentation package (Sc-P) which includes: i) mini-SVS-P; ii) multibeam sonar.
- Power modules: batteries.
- Connectors and cables necessary to connect the above modules so as to ensure their the power supply and communications.
- A polyethylene chassis in which to allocate the aforementioned modules and the necessary buoyancy reserve to make the vehicle neutral in water.
- A remote surface station (Remote-Station) that includes: i) a system of georeferencing; (ii) a USBL/modem acoustic transceiver; (iii) a communication system in optical fiber including cable and winch; iv) a wi-fi/radio communication system between the surface station and the vehicle ; v) a remote supervision system; vi) a system of power supply and power distribution.

4.3.1 Positioning instruments

The development of the Blucy prototype has followed the principle of maintaining a low cost of manufacture but at the same time providing the best performance in terms of underwater positioning. In order to reach these kind of performances a set of positioning instruments has been integrated in a customized fashion, instead of using commercial solutions since they

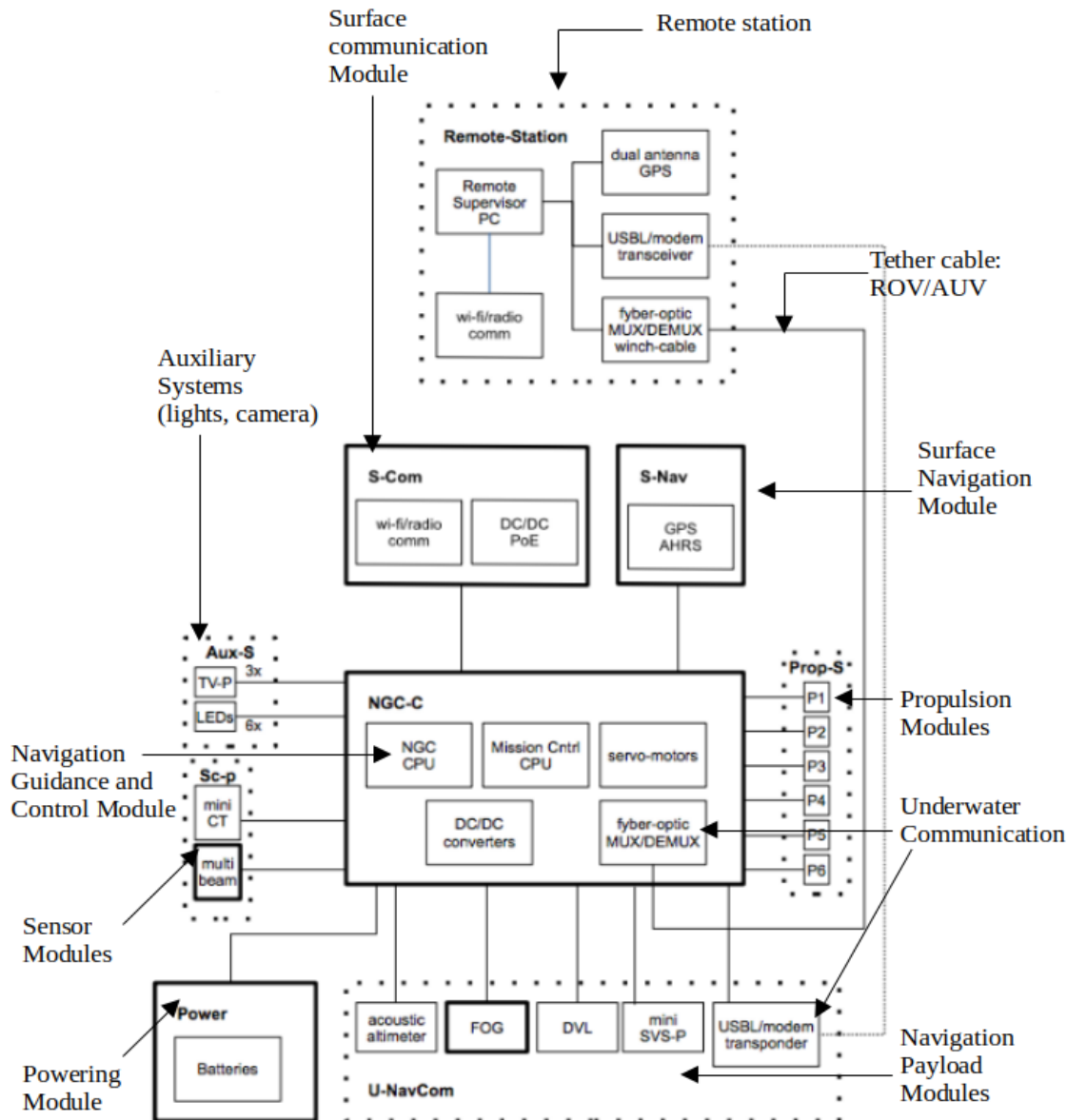


Figure 4.10: Blucy's functional architecture. Main modules are indicated with the corresponding arrow: remote station (located on the vessel or on the shore), surface communication and navigation modules (communication with the remote station via Wi-fi or radio modem and navigation based on GNSS), propulsion modules (8 thrusters), underwater communication modules, navigation payload modules (acoustic altimeter, fiber-optic-gyro, doppler velocity log, pressure sensor and the USBL modem), powering module (a double lithium battery 24V 50 Ah), sensor modules (conductivity e temperature sensor, multi-beam), navigation guidance and control module (the main canister equipped with an huge number of connectors that allows to communicate with the other parts of the vehicle, it is provided with: i)a CPU for managing the system and the navigation, driving and control system ii)a CPU for mission control management iii)the servo-controllers of the motors iv)the multiplexers for data/video transmission on the optical fiber cable v)conversion (DC/DC) and distribution components, including relays for enabling/disabling, of the energy to the other modules), and auxiliary systems (cameras and LED lights). In addition it is highlighted the tether cable made of optical fiber.

Sensor	Parameters	Model
FOG	Position, Attitude	iXblue Fiber Optic Gyroscope Compact C3
DVL	Linear Speed and Acceleration	Nortek Doppler Velocity Log DVL 500-300 m
AHRS	Latitude, Longitude	LORD 3DM-GX5-45 GNSS/INS
ALT	Altitude	Tritech PA200
USBL	Relative Position	EvoLogics S2C M 18/34

Table 4.5: Instruments for positioning of Blucy

are black-box systems. However, a commercial positioning system is more reliable and stable than customized systems. A list of the positioning instruments provided on board the UUV is showed in Table 4.5. As was mentioned in the previous section, some instruments are used when Blucy is over the water surface, such as AHRS, while others are used for dead reckoning, such as FOG, DVL, and USBL, which provide for relative positioning during submarine activities, when satellite signal is absent. The Attitude and Heading Reference System (AHRS) instrument is a GNSS receiver and an Inertial Measurement Unit (IMU) platform. The information acquired by the instruments are processed in real time by an Extended Kalman Filter (EKF) according to a Fossen model [74], which is based on wave filtering. Further information are provided by other sensors, such as MiniCT and MiniSVS, which are essential for precise positioning: temperature, conductivity, pressure and speed of sound. Indeed, these sensors improve the performance in obtaining the depth estimation. In addition, MiniSVS products are used for sound velocity correction within the algorithm for USBL. The same data are mandatory for proper operation of Multibeam Echosounder (MBES) during beam forming and beam steering. During the over water stage, the AHRS system is able to provide the position and the attitude of the UUV in a geographic reference system. The characteristics of the FOG, DVL and USBL are already mentioned in the previous sections, but a further explanation, especially for the USBL system is necessary, see Figure 4.11. In particular, for the correct relative positioning between the ship and the UUV, it is necessary measure with high precision possible the offset between the transceivers of the USBL system and the center of the GNSS on the vessel. Therefore it is crucial to implement a correct reference on the surface vessel by accurately measuring the offsets among the instruments on the boat. On the vessel are positioned two GNSS antennas which provide an absolute positioning and a heading angle, while pitch and roll are provided by another INS platform properly positioned (in such a way as to avoid heading measurement)

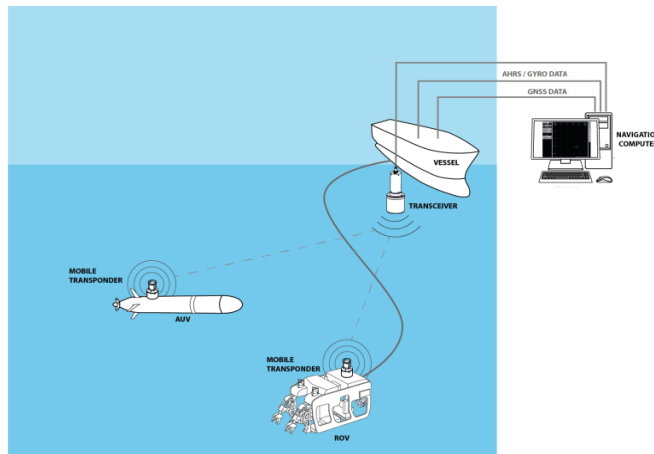


Figure 4.11: USBL positioning system, in ROV and AUV mode. On top of the figure it is possible to note the connections with the instruments placed on the vessel. Data coming from transceiver and vessel instruments are processed within the navigation computer to provide a precise positioning.

and calibrated. As mentioned before, offsets among these instruments have to accurately detected, with the aim to reduce as much as possible the position error propagation. During the implementation of the USBL system, it is possible to insert within the setting procedure, the 3D offset (X, Y, Z) and rotation (rX, rY, rZ) between the GNSS surface vessel Antenna Reference Point (ARP) and the USBL transponder ARP. By knowing the shape of the vessel, the configuration of the USBL system and relative position with respect to the vessel of the GNSS surface vessel ARP, a 3D vector is provided and it is used as preliminary setting in the implementation of the USBL system. Due to the propagation of noise in the water, it is necessary for the USBL system to be placed several meters away from the hull of the vessel; this is made possible by the intervention of a pole to which the USBL system is connected. The setup for the USBL positioning on surface vessel is showed in Figure [4.12](#). The USBL positioning and communication system provides quite precise relative positioning of the vehicle with a range of almost 500 m and it also allows to communicate with an underwater vehicle; therefore it allows a simultaneous positioning and communication. It is designed to exploit the Sweep Spread Carrier (S2C) technology, which continuously spreads the signal energy over a wide range of frequencies and adapt the signal structure so that the multipath signal components do not interfere with each other, in others words, this technology tries to mimic dolphins known to use a wide variety of acoustic signals to efficiently communicate over long distances. Indeed, Multipath propagation, fading, variable delays and limited bandwidth make underwater acous-

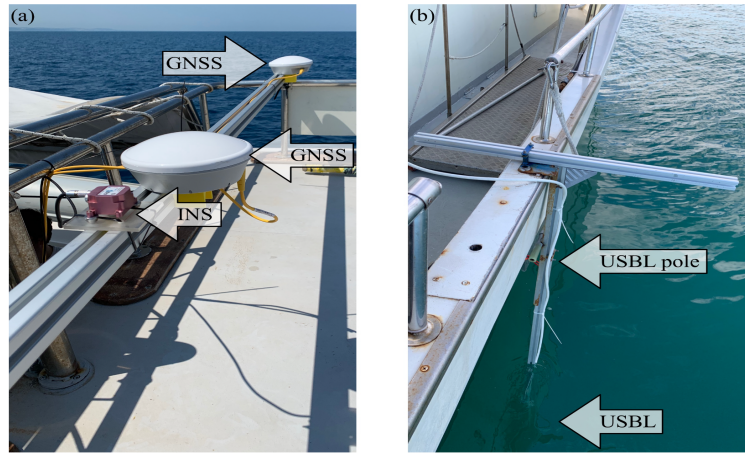


Figure 4.12: Setup for USBL positioning on surface vessel: (a) GNSS antennas and INS platform. (b) Structure for USBL.

	C3	C5	C7
Heading	0.15 deg	0.05 deg	0.01 deg
Roll pitch	0.05 deg	0.01 deg	0.01 deg
Position accuracy	0.2% DT	0.1% DT	0.05% DT
Power consumption	12 watts	20 watts	20 watts

Table 4.6: iXBlue Phins Compact C3 performances.

tic communication a challenge even in calm seas, where stratification of the water column can cause severe signal reverberations and dynamic fluctuations of the channel properties, resulting in poor transmission quality and below-average transmission speeds. It ensures an accurate 3D positioning and reliable data transmissions with up to 13.9 kbit/s. The USBL system uses an hemispherical beam pattern, optimized for medium range transmissions in vertical and salt channels. The Doppler Velocity Log (DVL) installed on Blucy combines compact design with a wide range of functionalities and it used for recording velocity. It can fly higher in the water column and closer to the seabed than similar equipment. This 500 kHz DVL is used by industry leaders in the subsea market because of its high accuracy and state-of-the-art technology. It owns a depth rating of 300 m and it can detects at a minimum altitude of 0.3 m up to a maximum altitude of 200m with a velocity resolution of 0.01 mm/s. The data recorded by DVL are integrated with those coming from the Fibre Optic Gyro (FOG) in order to reach a precise guidance of the underwater vehicle with a good accuracy for what regards the vehicle heading, see its characteristics in Table 4.6. FOG and DVL are located in the same canister, therefore they represent a single unit. The altimeter (ALT) can work in range of 50m through acoustic

Operating frequency	500 kHz
Beamwidth	6° conical
Range	0.3 to 50 m(0.1 to 10 m optional)
	Resolution
Digital	1mm
Analogue	0.025% of range

Table 4.7: Tritech PA-500 characteristics.

signal; its main features are showed in Table [4.7](#). The on-board surface navigation sensory module (S-Nav) allows to measure the position of the vehicle when it is on the surface and to have GPS positioning. In addition it used as Inertial Measurement Unit (IMU) when the vehicle is both on surface both underwater. The IMU consists of an accelerometer, a gyroscope and a magnetometer, which are high frequency sensors. The AHRS features a high-performance, integrated multi-constellation GNSS receiver utilizing the GPS, GLONASS, BeiDou, and Galileo satellite constellations, which provides GNSS data from a 1 Hz frequency up to 4 Hz. Sensor measurements are fully calibrated, temperature-compensated, and mathematically-aligned to an orthogonal coordinate system for highly accurate outputs. The auto-adaptive estimation filter algorithm produces highly accurate computed outputs under dynamic conditions. Compensation options include automatic compensation for magnetic anomalies, gyro and accelerometer noise, and noise effects. The computed outputs include pitch, roll, yaw, heading, position, velocity, and GNSS outputs- making it a complete GNSS/INS (GNSS Aided Inertial Navigation System) solution. The main characteristics of the AHRS are showed in Table [4.8](#). Through these positioning techniques it is possible to check in real time the position of Blucy during underwater navigation and to know its attitude. Positioning information is provided by cable in ROV mode and by acoustic channel in AUV mode. All the instruments listed above need to go through a calibration phase, which in the case of the USBL system, it can represent a difficult operation, especially in ROV mode, since it is required that the UUV has to make a round around the vessel with the eventuality that the cable can twist. The choice of several positioning systems, beyond the performance characteristics, it is driven by the possibility to integrate all these instruments among themselves; before to equip Blucy with an instrument, communication among instruments has to be ensured. In order to allow communication among instruments different protocols of communication are chosen, however it was used a single com-

General	
Integrated sensors	Triaxial accelerometer, triaxial gyroscope, triaxial magnetometer, pressure altimeter, temperature sensor, and GNSS receiver
Data outputs	<ul style="list-style-type: none"> • Imu outputs: acceleration, angular rate, magnetic field, ambient pressure, Delta-theta, Delta-velocity • Extended Kalman Filter (EKF) computed outputs: filter status, GNSS timestamp, LLH position, NED velocity, attitude estimates, linear and compensated acceleration, bias compensated angular rate, pressure altitude, gyroscope and accelerometer bias, scale factors and uncertainties, gravity and magnetic models • Complementary Filter (CF) computed outputs: attitude estimates stabilized, north and up vectors, GNSS correlation timestamp • Global Positioning System outputs (GPS) • GNSS computed outputs):LLH position, ECEF position and velocity, NED velocity, UTC time, GNSS time, SV. GNSS protocol access mode available
IMU Sensor Outputs	
Sampling rate	accelerometer 1 kHz, gyroscope 4 kHz, magnetometer 100 kHz
IMU data output rate	1 Hz to 500 Hz (standard mode) and 1 Hz to 1000 Hz (sensor direct mode)
Alignment error	accelerometer $\pm 0.05^\circ$, gyroscope $\pm 0.08^\circ$, magnetometer $\pm 0.05^\circ$
Computed outputs	
Position accuracy	± 2 m RMS horizontal, ± 5 m RMS vertical (typ)
Velocity accuracy	± 0.1 m/s RMS (typ)
Attitude accuracy	EKF outputs: $\pm 0.25^\circ$ RMS roll and pitch, $\pm 0.8^\circ$ RMS heading (typ)
GNSS outputs	
GNSS data output rate	1 Hz to 4 Hz
Velocity accuracy	0.1 m/s
Heading accuracy	0.5°
Horizontal position accuracy	GNSS: 2.5 m CEP, SBAS: 2.0m CEP

Table 4.8: GNSS Aided Inertial Navigation System main specifications.

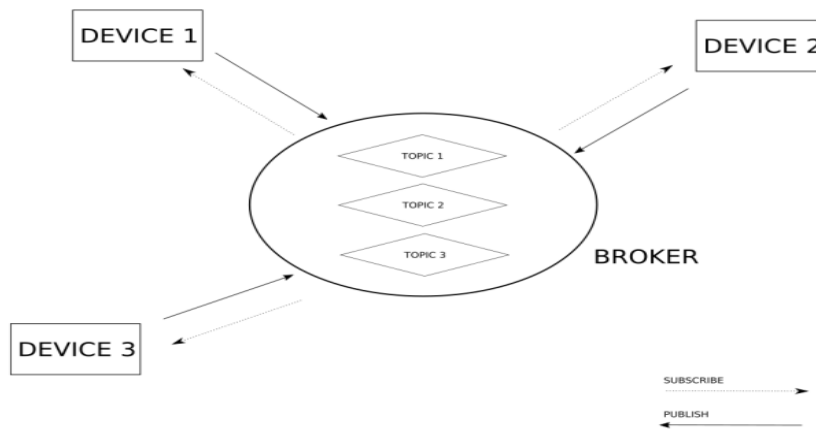


Figure 4.13: MQTT communication scheme.

munication protocol to gather all the sensor data and provide them outside; it is based on new paradigm of communication of Internet of things (IoT) that is the MQ Telemetry Transport or Message Queue Telemetry Transport (MQTT) [75]. MQTT is a standard ISO light messaging protocol of publish-subscribe type, see Figure 4.13. An MQTT network protocol is used for telemetry transmission from Blucy to the remote station. MQTT is an efficient and light weight open protocol to deliver publish/subscribe messaging between remote devices, maintained by OASIS Message Queuing Telemetry Transport Technical Committee and ISO/IEC 20922 compliant. Each device can communicate with the broker, represented by Blucy, by publishing a message but at the same time it is able to receive from him, who has the responsibility to dispatch several messages among devices. Data coming from devices are organized into topics in the broker system, in which the published messages are listed and sent to the outside, who can be the fiber optic cable or a USBL message. This communication protocol works on top of a TCP/IP protocol and it was designed for low bandwidth and low impact, in terms of exchanged data, communication as the case of Blucy. Each device has a set of input topics (used to receive commands) and a set of output topics (used to share telemetry data). In this fashion it is possible to use from different devices different topics in such a way that a single device can read the data coming from other devices without communication problems. MQTT server is initialized when the drone is turned on and is executed by the main canister CPU. The data is transmitted in real time through the fiber optic cable in ROV mode, or in AUV mode an efficient summary of only the key information for positioning and health of the

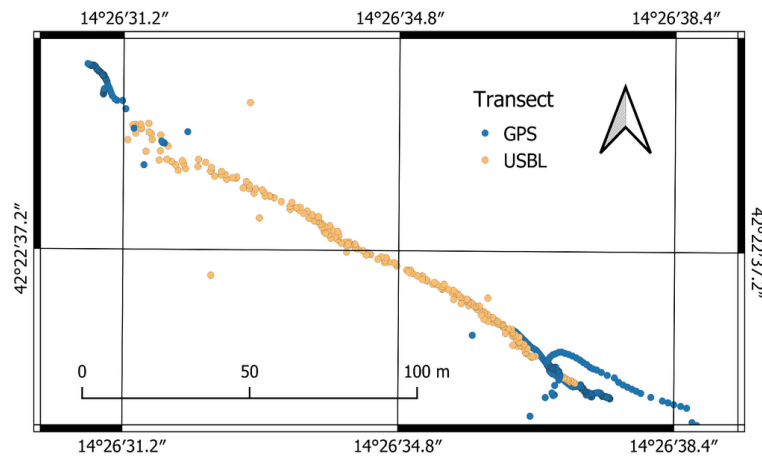


Figure 4.14: Transect performed by UUV partially above and partially under water surface.

drone is transmitted through the acoustic channel when the drone is submerged underwater. The remote station computer runs an MQTT client. All information is then displayed in real time and, depending on the mission performed, some of it is recorded and immediately sorted according to the type of sensor used for their acquisition. The point and telemetry data are saved on a fast SSD disk using a script created ad hoc. Data are recorded within a database formed by a sequential order of text files. Each mission performed produces a new database. Further, each device provide a customizable configuration file, which contains its parameters and all devices are developed in C++ language; each device is independent and it handles a specific hardware of the underwater vehicle. All devices are launched from a single shell script and the raw data coming from instruments are parsed by a C language script able to sort and making them human readable within the navigation computer. In Figure [4.14](#) it is possible to see a transect performed by the UUV, partially above sea water and partially submerged. In blue there are points that belong to the above sea water phase, while in yellow there are point that belong the underwater phase obtained by means of USBL positioning systems, but also whit the support of the dead-reckoning instruments such as FOG and DVL. The sequence starts from the southeast with position available from GPS receiver , and it stops in the north-west part of the image when the underwater vehicle re-emerges and the GNSS signal is again available. There is a substantial coincidence of positioning, however there are some outliers maybe due to the points that are in the water-air interface, where neither GNSS signal nor acoustic one is stably received.



Figure 4.15: Blucy prototype and fiber optical cable.

4.3.2 Surveying instruments

As we said before the main mission of Blucy was to monitor and survey the Adriatic benthic zone, however it was used also to provide support in mussel farm activities [3]. Blucy is an *hybrid* UUV, indeed, when dealing with submarine activities it is important have this feature. Several mission scenarios can happen during a submerged mission, therefore it needs to have a versatile system able to answer to our needs in the shortest possible time. For example, in the ROV mode, Blucy can be used when there is an irregular seabed, characterized by several obstacles or by a low visibility. This kind of mode is used when it has to observe a specific object, often difficult to reach: for close inspection of the seabed performed in hostile environments, in particularly tight spaces at moderate cruising speed, and when high position accuracy is required. Instead, when Blucy is in this condition, a long fiber optical cable, about 1000 m long, allows the connection between the navigation computer and the submarine vehicle: photos, videos and data coming from sensors are sent through cable to navigation computer. The selected cable has a nominal overall diameter of 7.80 mm, specific gravity of 0.95 kg/dm^3 , LCP fiber braid strength member, LDPE UV resistant sheath, hydrolysis UV resistant PUR outer sheath and a total breaking strength of 500 kg, see Figure 4.15. Instead, the AUV navigation mode is suitable for surveying large areas at high cruising speed in safe conditions and with a path previously planned through the use of way-points at a safe distance from the seabed and any other obstacle. The lack of connecting cable between the vessel and the navigation computer in underwater environment does not allow a heavy data transmission,

Sensor	Parameters	Model
MBES	Bathymetry, Water Column	Multibeam Echosounder R2Sonic 2020
PilotCAM	Live Stream Video	Vivotek IB8369A Network Camera
BottomCAM	High Resolution Imagery	Nikon Z6 with NIKKOR Z 24mm f/1.8 S
MiniCT	Conductivity, Temperature	Valeport MiniCT
MiniSVS	Pressure, Sound Speed	Valeport MiniSVS

Table 4.9: Instruments for surveying of Blucy

since only acoustic transmission can be performed, which consists of an exchange of small data packets. Therefore, it is necessary that when in this mode, each instrument be equipped with an appropriate solid-state drive (SSD) in which to store surveying data, which after will be downloaded through the Wi-fi antenna when the vehicle is on the water surface. A list of the surveying instruments provided on board the UUV is showed in Table 4.9, see Figure 4.16. The optical sensors equipped on Blucy allow to capture single frames and video streams to be transmitted in real time to the remote station through the fiber optic cable when the drone is on mission in ROV mode. The specific orientation of the two optical sensors allows to differentiate the type of acquisition according to a dual approach. There are two cameras:

- PilotCAM transmits a video stream with a resolution of 2560x1920 and a constant frequency of 30 fps, compressed with high efficiency h265 coding. The video stream is transmitted over a network protocol and displayed on the remote station to allow manual navigation of the UUV. The same video stream is recorded on board the remote station disk and the resulting recording can be used for processing such as automatic recognition of marine features using machine learning algorithms, as well as being an excellent tool for inspection and visual census. It is positioned in the frontal part of the underwater vehicle, which coupled to its lens has been selected to obtain an image with wide Field of View (FOV). It provides a video stream in real time when Blucy is in ROV mode. It is essential in this navigation mode since it allows precision navigation at particularly close range to the seabed. It is installed on a special custom support that allows to change inclination angle of the sensor, typically at an angle of about 45 degrees with respect to the seabed.
- BottomCAM consists of a high-resolution full frame camera capable of recording 24-

megapixel frame sequences that are transmitted in real time to the remote station through a dedicated fiber optic channel. The image transmission is managed through a subsystem dedicated to the conversion of the USB signal into a network protocol through which the images are rapidly transmitted. The same network interface allows the remote snapshot of frames at the frequency necessary to achieve the photogrammetric coverage of the seabed. It is installed at the bottom of the UUV with a nadiral viewpoint in order to be used for photogrammetric purposes in terms of three-dimensional reconstruction and the realization of the Digital Model Elevation (DEM) and orthomosaic of the seabed. According to Ground Sample Distance (GSD) and visibility conditions in the water column, several main parameters such as overlap, shutter speed and height above the seabed are then chosen. It provides high resolution images in real time when the UUV is in ROV mode and through the fiber optical cable it is possible to ensure a bidirectional communication between the sensor and the navigation computer, indeed it can take pictures directly from the navigation computer and at the same time they are sent to it for visualization. The BottomCAM is positioned inside a waterproof container where a signal conversion module suitable to remotely control the camera allows the data transmission. In addition, in order to protect from water pressure the BottomCAM, in front of the waterproof container it is placed a thick glass, which induces distortions that have to be properly removed to obtain an accurate surveying product. Regarding this camera and its technical specifications, please refer to the chapter on the optical sensor.

In addition, there is a MBES active acoustic sensor which is used to detect information and produce maps related to more distant objects, as well as to acquire high quality bathymetric data, water column information and quantitative fisheries stock assessment for habitat mapping. The MBES sensor is installed on the vehicle in a proper frame with the possibility to vary the angle of acquisition and its main features are shown in Table 4.10. Other information related to the water column are punctually acquired during the navigation by MiniCT and MiniSVS sensors, such as conductivity, temperature, pressure, and sound velocity. Their main characteristics are shown in Table 4.11. Due to the low space available on board of the underwater vehicle and the need of low energy consumption, which otherwise would weigh to

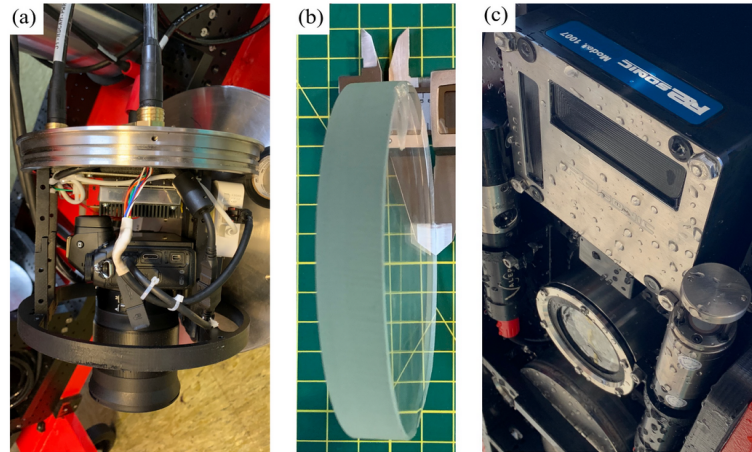


Figure 4.16: Main surveying instruments: (a) BottomCAM and its network data transmission router before insertion into the waterproof case. (b) BottomCAM 15 mm frontal glass. (c) MBES sonar head (top) and final BottomCAM canister (bottom) installed in UUV.

System Feature	Specification
Frequency	400 kHz to 200 kHz
Beamwidth - Across Track (at nadir)	2.0° @ 400 kHz/ 4.0° @ 200 kHz
Beamwidth - Along Track (at nadir)	2.0° @ 400 kHz/ 4.0° @ 200 kHz
Number of Bems	256
Swath Sector	10° to 130°
Maximum Slant Range	1200 m
Pulse Length	15 μ s - 1000 μ s
Pulse Type	Shaped Continuous Wave (CW)
Depth Rating	500 m
Operating Temperature	-10° C to 40° C

Table 4.10: Sonic 2020 multi beam system specifications.

	Conductivity	Temperature
Range	0-80 mS/cm	-4 to +35°C
Accuracy	\pm 0.01 mS/cm	\pm 0.01° C
Resolution	0.001 mS/cm	0.001° C

	Sound Velocity	Pressure
Range	1375-1900 m/s	5, 10, 20, 30, 50, 100 or 600 Bar
Accuracy	\pm 0.017 m/s	\pm 0.05% range
Resolution	0.001 m/s	0.001% range

Table 4.11: MiniCT and MiniSVS technical specifications.

the battery system, all these instruments equipped on Blucy are chosen following a trade off between dimensions, consumption, and price. Before to start the surveying mission a sort of checklist has to be accomplished in order to avoid unpleasantness that could make meaningless the surveying data:

1. Dry calibration for navigation and positioning: initialization of all navigation and positioning systems. The AHRS fixes GNSS position signal and initializes the FOG for position estimation during dead-reckoning.
2. Wet calibration tests for navigation and positioning: deployment of Blucy from the surface vessel. A series of maneuvers, both in ROV and AUV navigation mode, will be performed with the aim to verify the correct functioning of the navigation guidance. During this phase a buoyancy test is performed, since based on the salinity and presence of fresh water it could modify the buoyancy characteristics.
3. Wet calibration for scientific survey: initialization of all surveying instruments. A water column analysis will be performed by monitoring sound speed, pressure, temperature, and conductivity with the purpose to calibrate scientific instruments and to provide data for further scientific post analysis. Initialization of MBES mission profile based on hypothetical navigation altitude.
4. Mission survey: the survey mode depends on the purposes of the mission.
 - MBSE survey: it is possible to survey an huge area in short time and to detect hotspots. The planning survey lines are designed with a trade off among: Blucy navigation depth; seabed morphology in terms of slopes and rocky formations; seabed type in terms of grain size characteristics; swath width. They are planned to be perpendicular to slope directions to maintain a constant swath coverage. During this kind of survey both the navigation mode can be applied.
 - Close seabed survey: it is possible to select previous detected trouble hotspots which need to be further inspected by an optical sensor in order to obtain deeper morphological data. During this kind of survey only the ROV can be applied since Blucy

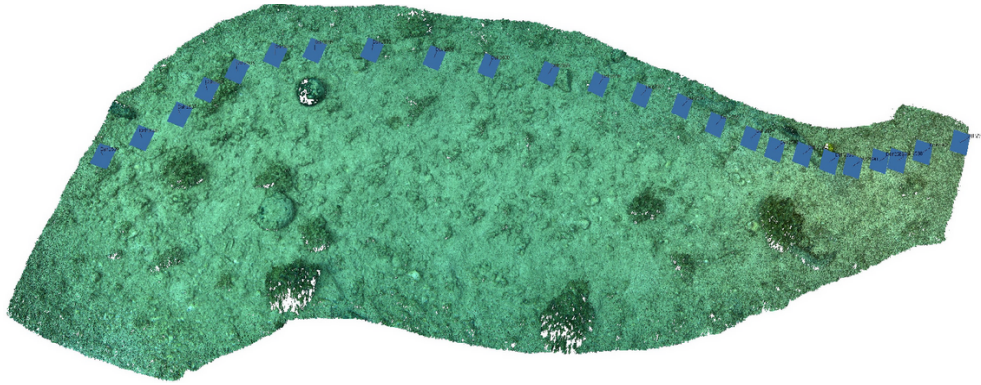


Figure 4.17: Dense 3D point cloud of seabed processed from BottomCAM high-resolution imagery with SFM techniques.

has to move in a range of altitude between 1-5 m, it depends on the turbidity of the surveyed area.

By using telemetry information collected simultaneously by the subsystems on Blucy, it is possible to georeference large portions of the acquired images and produce metric products such as orthophotos of the seabed. Furthermore, SFM from imagery or MBES data processing techniques, it is possible to obtain a three-dimensional reconstruction of the marine environment and an high-resolution DEM of the seabed as showed in Figure [4.17](#), where the pose estimation of the camera is highlighted by the blue squares.

Chapter 5

Structure from Motion technique

5.1 Introduction

Taking a cue from the mechanism by which the human eye observes reality, numerous scholars have wondered how to obtain even the depth of an object from images of it. Some of the earliest scholars to answer this question are J. Koenderink and J. van Doorn, who developed a theorem that is called the Structure from Motion theorem. In 1991 Koenderink and Doorn have given birth with their work [76] to a new line of research that, due to developments over the past decades in the field of computer vision, has and is becoming very successful because of its versatility toward its various applications. Since that time, many are things have changed, but the two scholars laid the groundwork for a method that has gradually been improved, or rather, corrected and refined to high levels of accuracy, thanks in part to recent artificial intelligence techniques. The theorem states that it needs at least 4 points from two views of the scene to be able to appreciate even the depth of an object, thus to obtain from a 2D image a 3D object belonging to the scene or the spatial configuration. By adding other views the reconstruction process improve its efficiency and accuracy. Since that time, the technique has evolved considerably, managing to channel within it concepts of projective geometry as well as topics in numerical computation useful for optimizing computation. So the idea of Structure from Motion that we have today is completely different from the original one. One

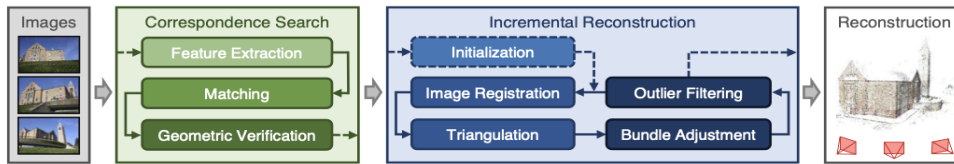


Figure 5.1: Incremental Structure-from-Motion pipeline. Source: [4].

of the most recent Structure from Motion approaches that highlights these differences is, for example, the incremental Structure from Motion technique developed by Schonberger et al. in [4]. They provide a whole open-source framework, called Colmap, in which a strategy for 3D reconstruction from unordered image collections is performed. As you can see from Figure 5.1 the SfM theorem was split in three main phases and each phase contains inherent steps useful to ensure the validity of next step as well as of the entire process. Before addressing the discussion regarding the experiments conducted, since the final products of the experiments are point clouds that can be traced back to the SfM pipeline, it is necessary to introduce the basic geometric concepts, the different possible camera models, projective geometry, and then the principles of 3D reconstruction from multiple views. In addition, a section on challenges in the underwater environment in the context of SfM will be provided, as well as some examples of the application of this technique. As source of information served the work of Hartley and Zisserman 'Multiple View Geometry in Computer Vision'.

5.1.1 Projective Geometry

The instruments provide by Euclidean geometry do not allow to treat a point that is infinitely far apart and to handle transformations as linear ones. These two features are essential when dealing with perspective model and affine transformations. For these reasons, it is said that, Euclidean space \mathbb{R}^n is extended by one dimension to form projective space \mathbb{P}^n . Therefore, projective geometry needs to mapping between 3D world and 2D image. A point in a plane is usually represented as a pair of coordinates $(x, y)^T$ in \mathbb{R}^2 . Therefore, every point that fulfills the equation

$$ax + by + c = 0 \quad (5.1)$$

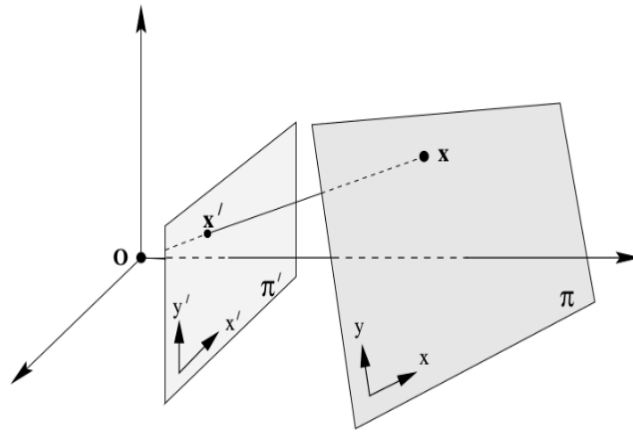


Figure 5.2: Central projection. Source: [5]

lies on the same line within the plane. Considered in vector space, the point $x = (x, y)^T$ can be seen as a vector and the line can be simply defined as

$$(x, y, 1)(a, b, c)^T = 0 \quad (5.2)$$

This representation is also called the homogenous coordinates of point x or the projective space \mathbb{P}^2 , where the additional coordinate dimension allows to treat points as lines and vice versa. Figure 5.2 shows this relationship. By tracing a straight line between a point x in the plane π and the projection Center O , the intersection with an arbitrary plane π' can be computed. The intersection of point x' is the projection of x to plane π' . Actually, in practical issues multiple points from the plane π should be projected to π' . To accomplish this transition, a 3x3 transformation matrix called Homography can be determined from a set of pairs of homologous points. A transformation consisting of a rotation (r_1, r_2, r_3, r_4) and a translation $C = (c_x, c_y)^T$ and it can be expressed by the matrix T :

$$x' = Tx = \begin{bmatrix} r_1 & r_2 & c_x \\ r_3 & r_4 & c_y \\ 0 & 0 & 1 \end{bmatrix} x \quad (5.3)$$

The maximum degrees of freedom (dof) for this linear transformation range from a full projection with 8 dof to a simple Euclidean projection with 3 dof (2 translations, 1 rotation),


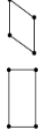


Group	Matrix	Distortion
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$	
Affine 6 dof	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$	

Figure 5.3: Perspective transformations. Source: [5].

depending on the relationship of the planes to be mapped. Therefore transformations are classified depending on their dof; Figure 5.3 gives an idea of all possibilities. Since a projective transformations have 8 dof, the 3x3 transformation matrix is determined up to scale. This means that the bottom row is not of the form $(0, 0, 1)$ anymore and can transfer points onto or away from the plane at infinity. This representation works only when mapping points from one plane to another one, indeed if points do not form an ideal plane, the computed transformation can only be obtained with an error from a best-fit plane, that is the plane that approximately contains the selected points.

5.1.2 Camera Models

In this section how a camera projects a 3D point in space onto a 2D point in the image is explained. For sure, the explanations are valid also for the inverse of such projection, that is, the back-projection: the computation of a 3D ray corresponding to a 2D pixel in the image. There are several possibilities and degrees of accuracy with which to model a camera. Here we consider the pinhole camera model and the axial camera model.

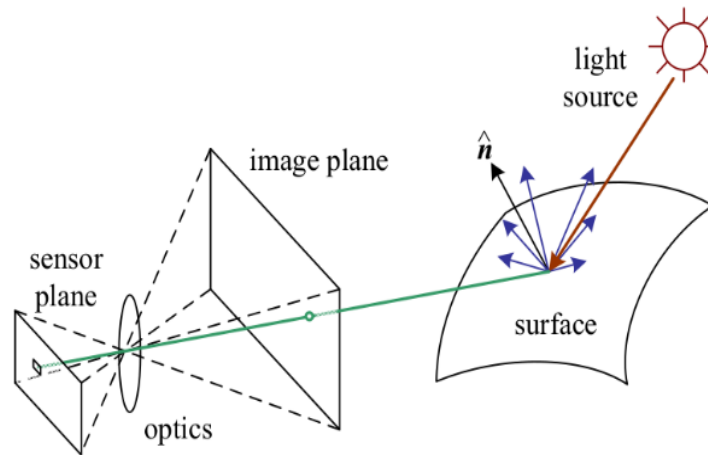


Figure 5.4: Simplified model of the camera's image acquisition process. A ray from the scene enters the camera through the pinhole or center of projection, crossing the optical axis. After that, it is recorded in the image as the point of intersection with the image plane. Source: [6].

Pinhole camera model

The pinhole camera model is the simplest way to model a camera. It is enough to imagine having a camera obscura, a box with a very small hole in one side through which the light enters. It produces an image of the outside world on the side of the box opposite to the hole. Let's replace the backside of the box with a plane able to record the amount of light entered (camera sensor) throughout the hole at different pixel locations. As can be seen in Figure 5.4, a ray connecting a scene point, located on a surface, with the pinhole or center of projection intersects the image plane causing the sensor to record its energy. Note that in Figure 5.4 the image plane is in front of the pinhole, while in a physical camera it is set behind the pinhole. But, from a mathematical point of view it is equivalent. Therefore, in this model all the rays pass through the center of projection contributing to form the final image. Actually, a lens system replaces the pinhole, but the pinhole model is valid for explaining the imaging process. The ray of light can be tracked back from the original world coordinates using the homogeneous representation of the central projection, as long as they are known the internal parameters (intrinsic) of the camera and the translation of the camera in the world frame (extrinsic). This relationship can

be formulated as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = PX \quad (5.4)$$

where P is the camera projection matrix, X the homogeneous world point, $(u, v, 1)^T$ is the 2D projection of X in the image plane and s an arbitrary factor scale. It can also be written as

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} su \\ sv \\ s \end{bmatrix} = x \quad (5.5)$$

and therefore simply as

$$x = PX \quad (5.6)$$

The projection matrix P can also be expressed in terms of the previously mentioned intrinsic and extrinsic camera parameters as

$$P = K(R|t) \quad (5.7)$$

where K is the 3x3 intrinsic camera calibration matrix that projects 3D points given in the camera coordinate system to 2D pixel coordinates and $(R|t)$ the 3x4 extrinsic camera pose matrix that maps 3D points represented in camera coordinates to 2D points in the image plane. The camera pose matrix consisting of a 3x3 rotation matrix R and a 3x1 translation vector t . The camera calibration matrix is a 3x3 matrix which consists of a principal point $c = (c_x = \frac{W-1}{2}, c_y = \frac{H-1}{2})^T$ that is usually close to the image center, an image dimensions $W \times H$ and a focal length f . This matrix does not depend on the scene viewed, so, once estimated, it can be re-used as long as the focal length is fixed. Figure 5.5 represents the schematic composition of the pinhole camera model. The calibration matrix is defined as:

$$K = \begin{bmatrix} f & 0 & c_x \\ 0 & f & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5.8)$$

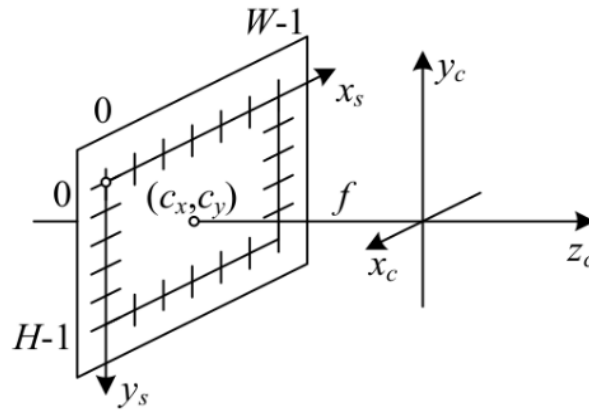


Figure 5.5: Simplified camera intrinsics. Source: [5].

It exists other extended formulation of this matrix that add a skew α parameter for non-rectangular pixel and aspect ratio for non-square image dimensions of the focal length:

$$K = \begin{bmatrix} f_x & \alpha & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5.9)$$

These last parameters will often be computed before complex tasks in a separate calibration step. They are crucial to any reconstruction approach, as small inaccuracies can lead to big errors due to the projective nature of the computation. When considering this model, this is also the reason why the algorithms or camera models take into account for imperfect lens system, which introduce tangential and radial distortion. Radial distortion moves pixel along the radius and is basically caused by the usage of a lens system itself. Tangential distortion is a translation perpendicular to the radius, which is caused by de-centered lenses. Radial distortion is usually classified into cushion and barrel distortion due to the deformation in the image (Figure 5.6). Tangential distortion can often be neglected due to very precisely manufactured lens system. Due to the non-linear behaviour of lens distortion they can not be expressed as a linear transformations. So a common approach is to define a high order polynomial to describe their deflections. Usually, these polynomials provide five coefficients k_1, k_2, k_3 for radial distortion and p_1, p_2 for tangential distortion.

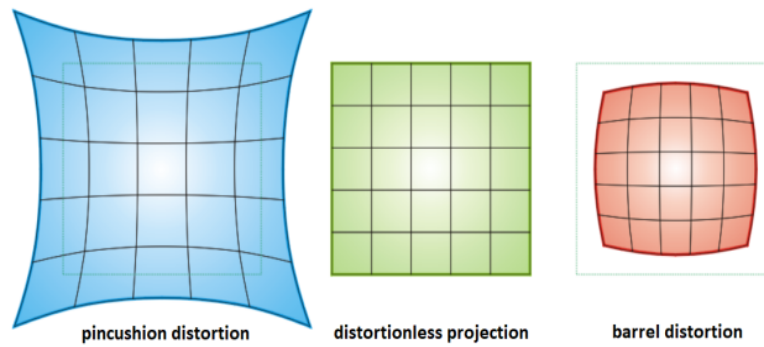


Figure 5.6: Different types of radial lens distortion compared with the case of absence of distortion. Source: [7]

General and Axial camera models

The pinhole camera model fulfills the condition based on the fact that the camera has one center of projection in which all rays of light intersect. This assumption is not hold for fish-eye cameras, catadioptric cameras, and also cameras in underwater housings, just to give some examples. In these cases, a more general camera model need to be used. Instead of describing how a ray is computed for each pixel in the image with a few parameters, a general camera model maps each ray for each pixel independently, considering the point where it starts and its direction. However, determining a calibration for such cameras is difficult to achieve robustly. In [77] cameras are classified based on their distortions:

- Perspective: cameras have a center of projection and do not have any distortion, i. e., the ideal pinhole camera.
- Single-View-Point (SVP): cameras encompass for example wide-angle cameras or perspective cameras with lens distortion. All rays pass through a common center of projection. Image distortions can be compensated if the camera model is known and calibrated.
- Non-Single-View-Point (nSVP): cameras are for example clusters of rigidly coupled cameras, some catadioptric cameras, but also underwater cameras. In order to compensate for distortions, the camera model and the 3D scene structure need to be known.

Grossberg et al. in [78] provide a generic camera model to deal with nSVP camera, where each pixel intercepts one ray of light, called raxel. Each raxel is modeled by a starting point and a direction that do not necessarily coincide with the physical camera. One of the main hypothesis of the Grossberg's work is that the bundle of rays has a singularity and its location is represented by a caustic, a geometric construct to which all rays are tangent, which uniquely describes the camera. Caustics are computed differentiating the mapping from image coordinates $(x, y) \in \mathbb{R}^2$ to rays:

$$X(x, y, k) = \begin{bmatrix} X(x, y, k) \\ Y(x, y, k) \\ Z(x, y, k) \end{bmatrix} = X_S(x, y) + k\tilde{X}(x, y) \quad (5.10)$$

with X_S that represents the starting point and \tilde{X} the direction of the ray. $k \in \mathbb{R}$ is the length of the ray. To find the singularity, the determinant of the Jacobi matrix of the ray function $X(x, y, k)$ has to be compute and set to zero:

$$\det(J(X(x, y, k))) = 0 \quad (5.11)$$

This allows to solve the equation and to determine the caustic point for each pixel (x, y) . The size of the caustic provides an idea of the amount of the deviation from the single-view-point camera.

5.2 Theory

If you look to the Equation [5.4], it is possible to reconstruct 3-space world from 2-space image point only by reversing the camera projection matrix P . Therefore, by knowing the intrinsic and extrinsic parameters would be possible a full reconstruction of the observed scene from a single photo. Actually, it never happens. Indeed, it is not possible to detect the distance or size of unknown object just by looking a photo of it, there is a need for much more information that provides at least a scaling factor. Therefore this fundamental issue can be solved by utilizing additional information: diminishing lines detected by means of deep learning, structured light

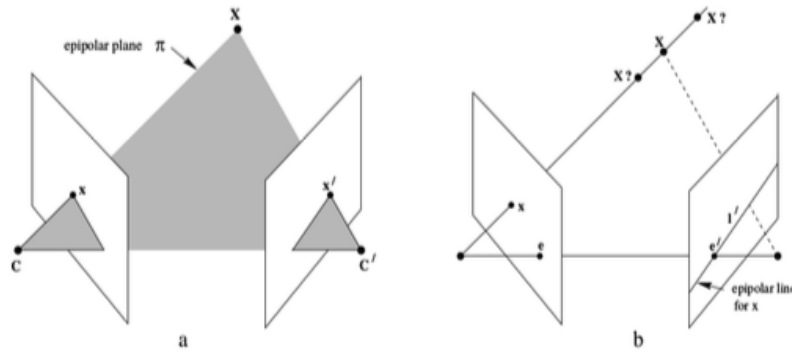


Figure 5.7: Epipolar geometry between two views. (a) The two cameras are indicated by their centres C and C' and image planes. The camera centres, point X and its images x and x' lie in a common plane π . (b) An image point x back-projects to a ray in 3-space defined by the first camera centre C and x . This ray is imaged as a line l' in the second view. Source: [5].

to triangulate between a projector and the camera system, and for aerial photogrammetry in an unknown environment the triangulation between multiple images, also called SfM.

5.2.1 Epipolar Geometry and Triangulation

By triangulation from multiple views, the reconstruction of 3D world points can be carried out. The intersection of rays define the minimum number of views, therefore two is the minimum number. When dealing with two or more views it needs to identify corresponding points between images and extract projection rays. The theory of epipolar geometry represents a useful instrument to handle this issue. Its algebraic representation is described by the fundamental matrix F , which is very similar to the homography transformation describing the mapping between two related planes. By looking Figure 5.7 it is possible to identify the epipolar plane π depicted in gray. It is defined as the plane that pass among three points: an image coordinate x in the left view, and the projection centers C and C' of both cameras. It contains the observed 3D world points and the corresponding image coordinate x' in the right view. One main consequence of this is that for every point in the left view the search for the same observed world point in the right view reduces to a 1D problem along the so called epipolar line l' in the image (Figure 5.7 (b)). The line that connects C with C' is called baseline, and its point of intersection with each image plane is called the epipole e , and respectively e' . Every epipolar line intersects with all other possible epipolar lines within an image in the epipole, due to

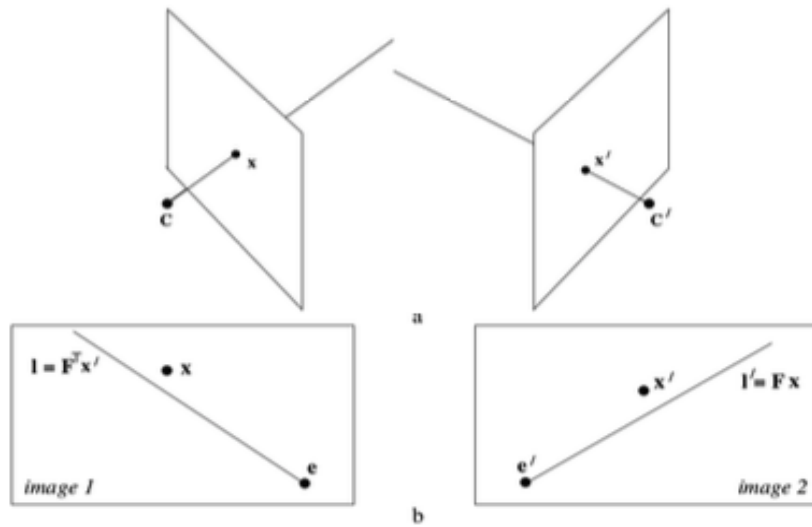


Figure 5.8: Misaligned image pairs. The measured points do not satisfy the epipolar constraint. Since the rays do not intersect, x' does not lie on l' , as well as x does not lie on l . Source: [5].

this fundamental relationship between views. Actually, the pose of two cameras can never perfectly identified. Uncertainties in camera model and acquisition process will always provide errors that lead to misalignments, and the search along the epipolar line will fail in giving the correct corresponding point x' . Figure 5.8 (a) represents the misalignment between two views, where the lines from the projection centers do not intersect the same 3D world point. Figure 5.8 (b) illustrates the resulting epipolar lines in their relative image planes where the image points do not lie on the respectively epipolar line. A straightforward triangulation under these assumptions inherits for sure these errors, because there is not point X which exactly fulfills the projective equations

$$x = PX, x' = P'X \quad (5.12)$$

for the views. Therefore, the choice of how to handle the reconstruction has to take into account this behaviour and it may be arbitrary in its application giving birth to several different approaches. For example in the linear triangulation method the Equation 5.12 is combined into a form $AX = 0$ with

$$A = \begin{bmatrix} xp^{3T} - p^{1T} \\ yp^{3T} - p^{2T} \\ x'p'^{3T} - p'^{1T} \\ y'p'^{3T} - p'^{2T} \end{bmatrix} \quad (5.13)$$

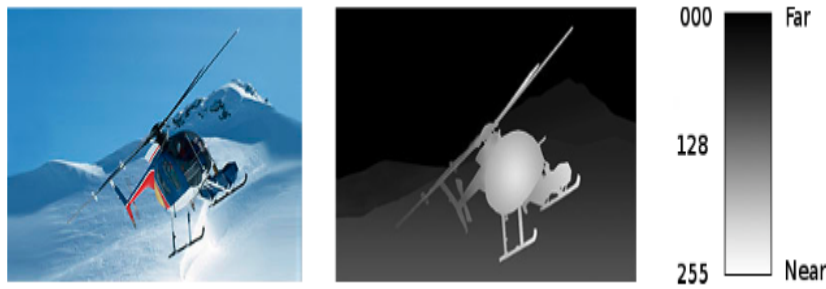


Figure 5.9: Depth map of an observed object. Notice the single channel color scale which gives an idea of the depth. Source: [5].

where p^{iT} are the rows of P and $(x, y), (x', y')$ the coordinates of x and x' . The error between rays can be minimized by applying an optimization to the system of linear equations.

5.2.2 Depth and Depth Maps

A 3D world point can be triangulated using Equation 5.13 from different views. For a dense reconstruction such triangulation is ideally performed once for every pixel of the image. It allows to obtain a very efficient representation of the observed surface. A feature very important, before directly computing all 3D-coordinates, is the distance of each pixel to the scene. This feature provides distance values that are typically referred to as as depth and single channel images with depth values encoded at each pixel as depth maps. Depth maps are more memory efficient and illustrative; Figure 5.9 shows an example of depth map. For the depth map computation, based on Equation 5.6, a world point X can be projected to the ray $x = w(u, v, 1)^T$ using the camera projection matrix with $x = PX$. The last row of this matrix product is

$$(p_{31}, p_{32}, p_{33}, p_{34}) \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = w \quad (5.14)$$

Where w represents the depth if $\det(n) > 0$ and $\|n\| = 1$, with $n = (p_{31}, p_{32}, p_{33})^T$ that is the principal axis direction. The pinhole camera model can provides depth since the rotation

matrix R is orthonormal. Consequently, the depth d can be computed as:

$$(r_{31}, r_{32}, r_{33}, t_z) \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = d \quad (5.15)$$

In order to reproject a depth map to a 3D point cloud, Equation 5.6 can be used again but in the form

$$x = KRX + t \quad (5.16)$$

with $R^{-1} = R^T$ and $x = d(u, v, 1)^T$. Therefore, it follows

$$R^{-T}K^{-1}d \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - R^T t = X \quad (5.17)$$

As the inverse of the euclidean transformation from the camera to the world frame turns to

$$P = K(R^T | -R^T t) \quad (5.18)$$

therefore Equation 5.17 becomes

$$RK^{-1}d \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} + t = X \quad (5.19)$$

5.2.3 Bundle Adjustment

Bundle adjustment is a generic term for optimization in visual reconstruction and photogrammetry. It applies in order to achieve highest accuracies through robust non-linear optimization that try to estimate at the same time 3D structure and viewing parameters. The name of such

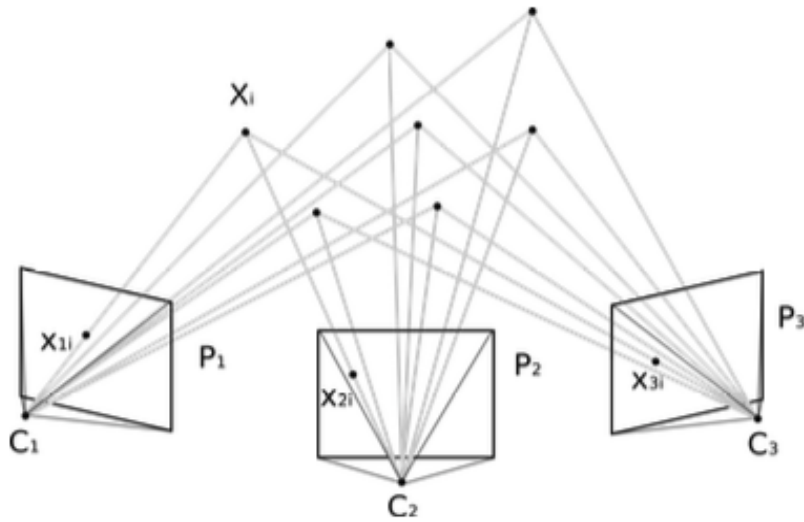


Figure 5.10: Bundles of light rays for multiple cameras observing the same scene. Source: [5].

techniques refers to the bundles of rays, derived from the image acquisition model described in the previous sections, which connects 3D world points and projection centers. The rays are linked among them through the same observed scene, while each camera can be considered a separate system. By looking Figure 5.10 the relationship among rays can be appreciated and it can be formulated as

$$x_{1i} = P_1 X_i, x_{2i} = P_2 X_i, \dots, x_{ni} = P_n X_i \quad (5.20)$$

In order to induce the adjustment, a cost function including optional constraints is minimized. One common measure for deviations is the geometric error, or the reprojection error. A world point X is observed by two cameras with projection centers C and C' at the corresponding point positions x and x' . However, due to the unavoidable errors, the true position \bar{x} and \bar{x}' which exactly fulfills the epipolar constraint $\bar{x}'^T F \bar{x} = 0$ is unknown. By minimizing the euclidean distance $d(*, *)$ in the cost function f with

$$f(x, x') = d(x, \hat{x})^2 + d(x', \hat{x}')^2 \quad (5.21)$$

the points \hat{x} and \hat{x}' can be found. Notice that these constraints can be set up for all cameras and observed points. It can result in a huge over-determined system of non-linear equations. Indeed, when dealing with feature matches that are several millions of measurements and

hundred of thousand of unknown parameters such as 3D point positions, and camera poses, the bundle adjustment can be applied. This can be seen as a complex task, since huge number of operations are needed. But, due to the fact that every feature point $x_{i,j}$ only depends on the one world point position X_i and one camera pose $(R|t)$, the problem is sparse and allows very efficient solutions.

5.2.4 Camera Pose Estimation

Camera pose stands for the 3D transformation of a camera relative to another coordinate frame. If not explicitly defined, it is always a rigid body motion describe as usually by 3x4 matrix

$$M = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \quad (5.22)$$

The estimation of matrix M is crucial. In order to compute it different approaches exist, which depend on the prerequisites given by the scene. The Perspective-n-Point-Algorithm (PnP) is used with 3D world points as prior information allowing to extract M from a single view. In case of multiple views approaches such homography and fundamental matrix can be applied even to unknown scenes.

P-n-P-Algorithm

The PnP Algorithm allow a single view reconstruction of the camera pose. It is applied only when the 3D world points are already known. This is the case during camera calibration with a chessboard whose geometrical characteristics are already known. The coordinate system of the world frame can be chosen as one of the edges of the chessboard. The remaining black and white junctions form a common plane, in which case one coordinate axis can be set to zero. In the next step, it is assumed that the visual angle of any pair of 2D points must be the same angle between their corresponding 3D points. Depending on the number of points n selected

($n \geq 3$) for estimation, different formulations exist to solve this problem. It is important to note that the PnP Algorithm can handle non-planar points, but in case of the calibration they are advantageous to initialize a fixed coordinate system without additional views required.

Homography

The homography H is a 3x3 transformation matrix in the projective space. Its estimation is a multiple view operation which uses corresponding features in the images to detect the camera movement. It follows the equation

$$x = Hx' \quad (5.23)$$

where x and x' are both homogeneous 2D projections of world point X , but from different viewpoints. The underlying assumption implies, that all observed points lie in the same plane. Therefore while PnP-Algorithm can handle non-planar points, homography can not. But, it is able to reconstruct camera movement in unknown environment up to a set of four ambiguous solutions. Subsequently, these can be decomposed with the camera calibration matrix K into pose matrix M . If more than four corresponding points are provided, a best-fit optimization has to be performed.

Fundamental matrix

The fundamental matrix F is a 3x3 matrix similar to homography with the difference that even the intrinsic camera calibration matrix K is unknown. Therefore it is independent of the scene structure and described by

$$x'^T F x = 0 \quad (5.24)$$

where x and x' are the same elements of homography. Despite its independence of the scene, when dealing with planar scene or pure rotational motion, the issue might not be well constrained. Therefore an hybrid approach in which either the homography either the fundamental matrix are computed can yield better solutions.

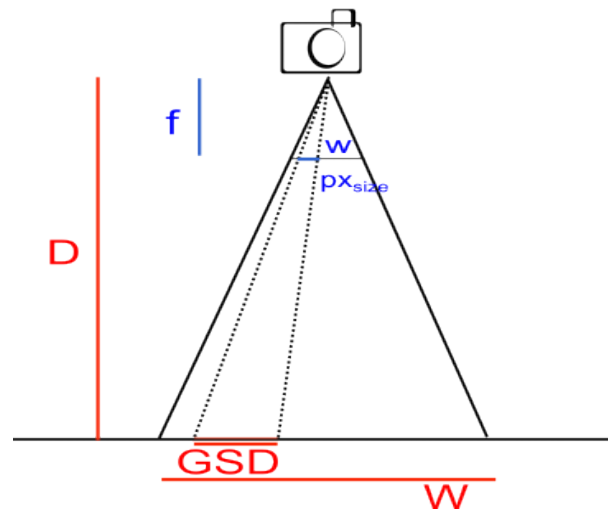


Figure 5.11: GSD relationship illustration.

5.3 Challenges in the underwater environment

SfM technique is used for 3D reconstruction. Most of our sunken heritage is located on the seabed, therefore it is difficult to reach and its reconstruction could contribute to its spreading. At the same time, the interest of natural sciences and biology in studying the dynamics of population that inhabit underwater environment, thanks to the low request of money and resources associated to SfM, is greatly increased. Therefore, to answer to the needs of scientists and cultural heritage management, SfM in underwater environment has to fulfill some good practices in order to obtain a maximum level of accuracy. Before to start an underwater survey it is necessary to define a survey planning, with which checking if the requirements set at the beginning are met. For this reason, it is necessary to face some challenges or to take some precautions during the underwater survey. First of all, attention has to be payed to what resolution and accuracy you want to reach or at most you can reach. You have to decide what will be the smallest detail that must be visible in the image or in general in the products; this is linked to the growth rate of a population or in other cases to the level of detail predetermined. The Ground Sample Distance (GSD) is the main parameter which allows to rule resolution and accuracy. In the computation of GSD for a survey it needs to define the image scale, which is computed as

$$\frac{f}{D}$$

where f is the focal length of the camera system and D the height whence the photos are taken from the object scene. By applying simple rules related to triangles, it is possible to define a relationship among some quantities within the survey

$$\frac{f}{D} = \frac{w}{W} = \frac{px_{size}}{GSD}$$

where px_{size} is the size of the single squared pixel in the photos, W is the field of view or coverage, and w is the length of the field of view inside the camera. Through this relationship it is possible to extract the GSD and W as

$$GSD = \frac{D}{f} px_{size}$$

$$W = \frac{D}{f} w$$

For sure, quantities like f and px_{size} have to be already known and fixed, while for what concerns D it has to decide a proper value based on the conditions with which the survey has to be carried out. Once these geometric quantities have been set, the focus has to be moved towards the choice of the camera system that you want to employ, as for instance, flat or dome ports, use of artificial light or even the use of a UUV. The choice of these system should be justified through underwater calibration tests (see calibration frame or calibration grid) that are able to analyse the image quality and geometric performances. Afterwards, it needs to select a method for datum definition and quality control, which at best it is represented by a reference points of known coordinates that establish a local geodetic network. This is not always possible, therefore it is necessary to define at least one metric reference, such as scale bars. In the phase of underwater site preparation, markers must by size correctly and usually circular coded targets have to be of a diameter bigger than ten times of the GSD . Further, in case it is possible, a good practice is to surrounding the area of survey with colored tapes in order to make it even more visible. A good overlap between subsequent photos has to be ensured (80%). If b is defined as the horizontal distance between two subsequent shots, also

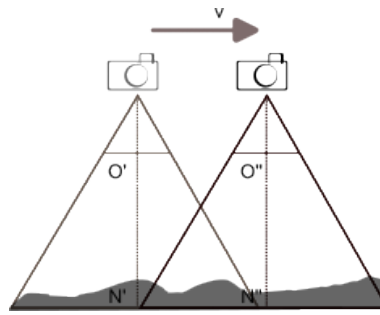


Figure 5.12: Illustration of the relationship between motion speed and shutter time. Inspired by [8].

called baseline, thus it has to be equal to

$$b = \left(1 - \frac{\text{overlap}\%}{100}\right)W$$

in this way the selected overlap is fulfilled. Beyond the choice of the baseline, another main parameter is the relation between motion speed and shutter time. As you can see from Figure 5.12 it needs to maintain a certain swimming speed to avoid motion blur. If the swimming speed is defined as

$$v = \frac{s}{t}$$

where $s = N'\bar{N}''$ is the displacement and t is the exposure time, thus $s < GSD$. Therefore if GSD and v are known, it is possible to extract the exposure time. Otherwise, if v and t are known, it is possible to extract s displacement. Further, it is mandatory to minimize the deformations of photogrammetrically derived 3D models due to accumulation of non compensated systematic errors. For this reason, a good practice, in case of area-based surveys, is to carry out parallel transects with cross strips with an overlap of about 80% and a sidelap of about 60%. While for transect survey is necessary to carry out at least three strips, one nadir and two convergent oblique pointing inward. For fully 3D artifacts, a circular imaging network has to be set. The standard photography parameters have to be set in the right way. They are the aperture, shutter speed and sensor sensitivity and are depicted in Figure 5.13, where based on the direction of the arrows it is represented the effect on the image corresponding to the variation of the current parameter. For sure it has to be found a trade-off among these parameters based on the camera system employed. For example, in low light conditions it

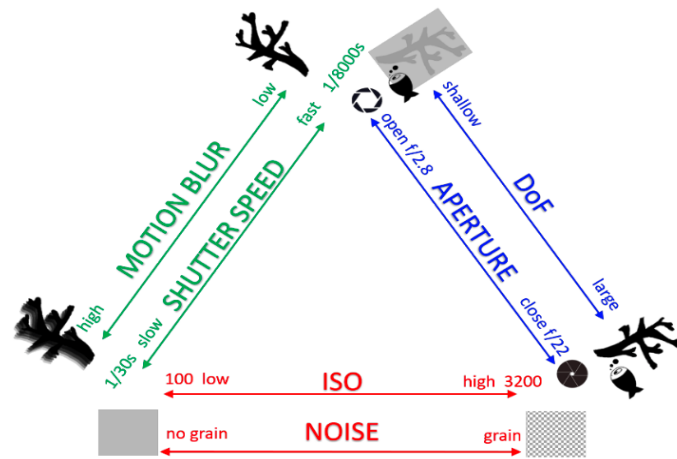


Figure 5.13: The photographic exposure triangle. Source: [8].

is natural to increase ISO, since it amplifies the light signal. It allows to have faster shutter speed but at the cost of increasing the image noise. If you choose a very small aperture (big f number) the depth of field (DoF) increases allowing a large representation of object scene but with a big probability that the image is noisy. Further, when the area of interest is defined, at the beginning of the underwater survey, in order to ensure the stability of the intrinsic camera parameters, the automatic focus option is selected to focus the camera system at the proper acquisition distance (D). Then it is fixed to that value for the whole shooting session. All these good practices are taken from the works conducted by Nocerino, Menna et al., such as [8].

Chapter 6

Underwater experience

6.1 Introduction

SUSHIDROP and TECHERA are the acronyms of two European research projects related to the underwater environment. The former has basically interested the construction from scratch of an UUV prototype, while the latter allowed essentially the employment of the prototype, called Blucy, in the Adriatic Sea. SUSHI DROP SUstainable fiSHeries wIth DROnes data Processing is a project financed by European Union through the Interreg Italy-Croatia Programme. The project aims at enhancing knowledge on accurate and non-invasive methods for mapping the marine ecosystems of Adriatic Sea, in order to assess environmental status of habitats and fish stocks population as reliable and up-to- date information about the state of marine resources are essential to support sound management decisions. The most important goal of SUSHI DROP is to better understand the sensitivity of the habitats to fishing pres-



(a) <https://www.italy-croatia.euwebsushidrop>



(b) <https://www.italy-croatia.euwebtechera>

Figure 6.1: European Interreg research projects.

asures and to design and implement more effective marine management plans. SUSHI DROP evaluates the adoption of drones (UUVs) equipped with sensors to monitor physical, chemical and biological features. In particular, acoustical and optical technologies will be employed as a non-invasive mean to assess fish stocks population. The findings of the opto-acoustic surveys are compared with classical procedures based on fish sampling and to assess the accuracy in deriving single-species abundance indices (in numbers or weight) for direct input into stock assessments. The data gathered during the project will be collected in a Geographical Information System known as GIS. It serves as an open database for collecting, maintaining and sharing the scientific data acquired by the UUVs and as a useful resource in further research and preservation of the biodiversity of the Adriatic. TECHERA is a cross-border cooperation project that aims to capitalize and connect data-driven approaches and practices in the fisheries sector towards a sustainable blue economy in the Adriatic. A sustainable Blue Growth may be undermined by the unavailability of scientific and socio-economic observations related to the status of marine environment. The TECHERA project will tackle this problem by promoting the adoption of emerging technologies such as big-data sharing and analytics to increase the smart specialization of companies operating in the blue economy, and to attract the interest of young people in smart and blue careers. Products of previous projects will be interconnected into study examples that will promote sharing of experience, and provide evidence of potentials and benefits of integration and complementary of approaches. Operational and knowledge gaps will be identified as well. To achieve a consistent, meaningful, and successful communication, TECHERA identifies relevant stakeholders, involving also other Standard and Strategic IT-HR projects operating in the same domain, and adopt different approaches targeting the transferability of results in the Programme Area and beyond. Finally, TECHERA proposes potential inputs for the next programming period in order to contribute to the definition of a strategy for digitization and to the identification of key innovations to improve the competitiveness in the blue economy. Both these projects provided the opportunity to use instruments and equipment related to the marine environment, making possible to carry out underwater surveys, otherwise difficult to accomplish since they require a big effort in organisation as well as an employment of both economic and time resources. Without the support of these research projects, the fol-

lowing experiments would not have been possible. Blucy is the name of the UUV prototype developed in the SUSHIDROP project. At its bottom it is placed a full frame camera useful to perform underwater shooting. The underwater camera system equipped on Blucy will be the crux of the research conducted so far. Therefore the following sections will deal with where the experiments took place, how the experiments are conducted, why it has been chosen a certain approach in order to obtain results, and a finally a description of the obtained results.

6.2 Underwater surveys context

In this section will be described the place where the experiments were conducted, as well as the condition in which the camera or UUV was located in order to provide enough details for the interpretation of the results.

6.2.1 Calibration grid

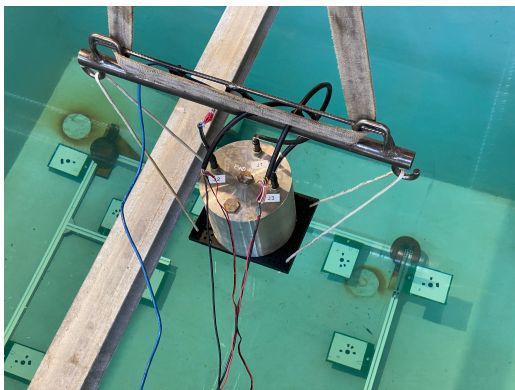
The implementation of the calibration grid was already mentioned in Chapter 2, here the characteristics of the survey and the difficulties encountered are emphasized. First of all, once the calibration grid was built, it was necessary to decide where to carry out the experiment. The requirements are a body of water at least 2.5 m high of fresh water in which to immerse the calibration grid and in which is possible to simply maneuver the camera system in order to catch snapshot from different point of view as well different heights. For the sake of simplicity, the camera system, which consists of camera itself, the underwater housing and a modem useful to connect the camera with a control panel, was separated from the entire structure of Blucy. In this way it was removed the issue related to the movement of the entire UUV, since it weighs about more than two hundred kilograms. However, the only camera system, after the separation from the UUV main body, continues to be quite cumbersome and heavy, about twenty kilograms because of the metal housing. For these reasons it was necessary a support system able to sustain the camera system. After a search of a few weeks, the experiment site was selected within the Department of Hydraulics of the Faculty of Engineering at the

University of Bologna. More precisely, it was a rolling tank useful for regulating the regime of the large adjacent tank used to harness and study wave motion. From here on, we refer to this tank simply as the water tank. It is $2 \times 2.5 \times 3 \text{ m}^3$ large and it is crossed by a traverse used as a support to sustain a metal coverage on which it is possible to transit on foot since the water tank is buried. For this reason it was not possible to remove the traverse which had represented a relevant obstacle during the shooting. Further, since the water tank is buried, when it has been emptied, the calibration grid could not be detected with a TS. The following steps have been carried out with the support of the laboratory technicians:

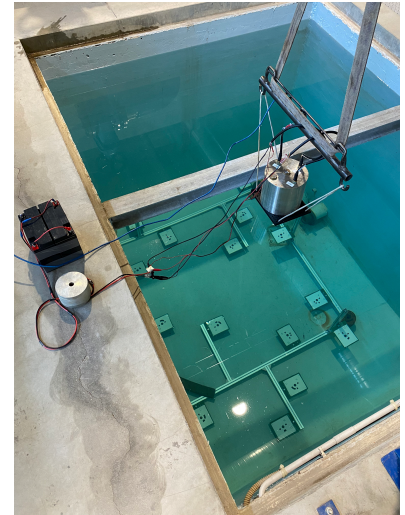
1. water tank cleaning
2. mounting the calibration grid
3. displacement of the calibration grid
4. filling the water tank
5. taking photo from different points of view, with the camera system immersed in the water
6. slow emptying of the tank
7. calibration grid survey with SLS system
8. removal of the grid and restoration of tank function

6.2.2 Calibration frame

Miramare marine protected area (MPA) in Trieste was the place of marine test focused on the survey of the calibration frame under the sea. Thanks to TECHERA project, Blucy was moved to Trieste to carry out a marine survey mission of 10 days in the waters of the Miramare Castle. Miramare MPA was one of oldest MPA in Italy, which has always been an area interposed to traffic, for military reasons, after World War II it continued to preserve this characteristic until it became MPA in the 80s. Several marine population had colonized the beaches near



(a)



(b)

Figure 6.2: Two snapshots taken during the experiment conducted in the water tank of the Department of Hydraulics of the Faculty of Engineering, University of Bologna. Notice the presence of a traverse, which was a fairly significant obstacle during the survey. In (a) there is a zoom of the camera system raised through the use of some ropes tied to an appropriate hook integral with an overhead crane. In (b) a view of the entire system where it is possible to observe the connecting cables with the battery and the control panel.

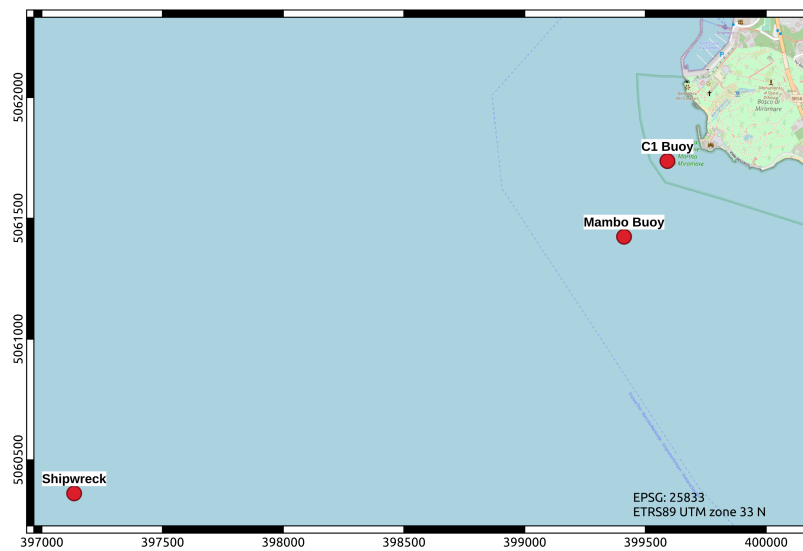


Figure 6.3: Map of the area in front of Miramare marine protected area in the Gulf of Trieste. Red bullets indicate the three points of interest of the survey mission. The calibration frame survey was conducted in the vicinity of 'C1 Buoy', in front of the Miramare castle.



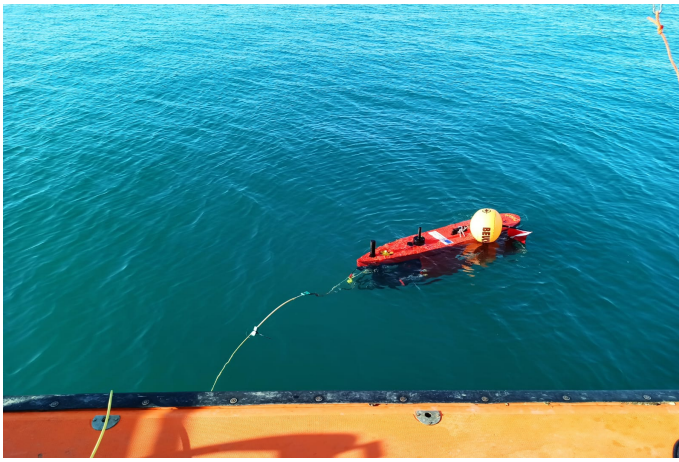
Figure 6.4: View of MPA Miramare, where you can see the vessel used during the mission docked at 'C1 buoy' while Miramare Castle appears in the background.

the MPA giving start to a creation of ecological niche overseen by WWF Italia. In addition to preserving the area from maritime traffic and intensive fishing, WWF Italia has carried out several scientific surveys in the area over time. If we look at Figure [6.3](#), in 'C1 Buoy' an artificial barrier consisting of brick and large tubular material was released on the sea floor to encourage its colonization by the marine fauna and flora present. Here, the sea level height is around 14 m, while the height of the artificial reef is estimated to be within a range of 2 to 5 m. While in 'Mambo Buoy' there is an old mussel breeding device on the seabed that is destined to be removed by the harbor master's office. The third red dot called 'Shipwreck' represents the place where an old vessel sank. Blucy had the task to survey the artificial barrier in acoustic mode since for visibility condition it was impossible to shoot photos of it. Therefore it was used a multi-beam system to survey the area and obtain the profile of the old artificial barrier in such a way useful to determine his shape and volume, without obtaining an estimate of the number of population that inhabits it since optical survey was avoided. The same situation occurred for the survey of the mussel breeding device since its shape and heights were unknown and a more closer survey with Blucy would have mean a risk for her since the presence of some old fishing nets or entangled ropes would have prevented its navigation. It must be remembered that Blucy is a prototype, which is why during the ten-day mission its main functionalities were tested in addition to making calibrations of the instruments on board. When dealing with an open device, such as Blucy, numerous contingencies have to be faced both from the point of view of



Figure 6.5: Vessel used to carry out the survey mission in Miramare MPA moored in the harbor. Also the presence of mechanical arm useful for loading and unloading operations of Blucy can be seen.

its operation and navigation. In fact, the first few days were devoted to setting up Blucy and testing the instruments. The second part of the mission, when the weather permitted, involved surveying the artificial reef and the mussel rearing device. While the last part of the mission involved surveying the shipwreck in addition to validating the previous surveys. Regarding the vessel chosen to transport the UUV and place the control panel, a vessel owned by WWF Italy that had the appropriate characteristics was chosen, see Figure [6.5](#). Indeed, this type of vessel has a crane arm that is very useful in loading, unloading and placing Blucy in the water. In addition to having a dedicated space to hold the large fiber optic coil and allow for basic operations related to the management of the fiber optic cable. The vessel also allowed for three crew members in addition to the two operators who acted only on the proper operation of Blucy. WWF Italia provided a scuba diver and a lifeboat during the mission with the task to support the activities of Blucy, as for instance the transport of the calibration frame and for its placing phase on the seabed. A lifeboat was chosen to avoid collisions of the calibration frame with other objects on the vessel in addition to the fact that it would have been an encumbrance for the operations to be carried out on board. The calibration frame mounted on a rigid metal structure was submerged in water with the help of some ropes attached to the structure so as to slowly promote its sinking to the seabed. There were two attempts to photograph the calibration frame placed on the seafloor. The first one failed, while the second one succeeded.



(a)



(b)

Figure 6.6: Two photos taken during the survey mission conducted in the water of AMP Miramare. In (a) it is showed Blucy during its activity on the surface of water. Pay attention to the fiber optic cable that connects Blucy with the control panel and the fact that there are floats integral to the cable that prevent drift effects. Also present is a buoy to signal the presence of Blucy. In (b) an underwater photo taken by a diver while Blucy takes photos of calibration frame. Due to poor visibility of the water, the rectangular-shaped frame appears highly disturbed but still visible. In addition, it is possible to observe the artificial light system needed to illuminate at great depths, where sunlight is faintly present.

It must be remembered that the success of this task depends on a great many factors, including weather conditions, visibility in the water, and the positioning and navigation system of the UUV. During the first attempt, the survey failed due to poor visibility conditions in the water; suffice it to say that the diver was unable to succeed in locating the UUV, even though he had his lights on. Therefore, the diver once he had an eye on the calibration frame lost contact with Blucy. Whereas, on the other hand, when he was able to locate the UUV he could not reach the frame. To solve this problem then, the idea was to simultaneously sink Blucy, diver and calibration frame using the vessel's anchor chain as a reference. But even in this case, survey operations were not successful, considering a depth of about 20 m. Fortunately, the second attempt was successful. An area near 'C1 buoy' with a depth of about 7 m was chosen. In this case, the visibility in the water was evidently better, and the maneuvering operations between calibration frame and Blucy were carried out strictly by taking advantage of the lighting device. In fact, before taking photos, a method for being able to communicate with the diver via Blucy was determined. So that the diver could move the calibration frame to facilitate taking photos of Blucy, see figure [6.6](#). Piloting Blucy to take pictures over an area of $1 \times 1.5 \text{ m}^2$ was not an

easy matter, as drift effects related to underwater currents had to be counteracted during each maneuver. The distance between the calibration frame and Blucy while taking pictures was kept in a range of 5 to 3 m. This distance value could not be kept constant due to impediments related to navigation and control of Blucy.

6.3 Experiment setup

In this section will be described the set-up of the two experiences conducted to calibrate the camera equipped on board of Blucy in order to reach the maximum photogrammetric accuracy. Indeed, to accomplish this task a calibration algorithm developed by Łuczyński et al. (2017) was applied. For both the experiments the following principle was fulfilled:

- survey of the calibration platforms with proper techniques, used as ground truth
- calibration of the camera system (just one time)
- survey of the calibration platforms by means of the RGB camera equipped on Blucy
- correction of the underwater images through Pinax algorithm
- both corrected and uncorrected photos used to feed the standard pipeline of the incremental SFM [4]
- 3D reconstructions of the calibration platforms
- highlighting the differences between the corrected and uncorrected photos taken by the camera system

In a nutshell, it is intended to show that applying the Pinax algorithm to photos taken by Blucy improves the metrics of 3D reconstructions compared to the ground truth represented by the survey conducted on the ground. As mentioned before the calibration grid experiment has to be seen as a precursor of the experiment related to the calibration frame.

6.3.1 Calibration grid

The instruments and devices used to conduct the following experiment are:

- the camera system: a Nikon Z6 fullframe camera, with a focal length of 24 mm and a resolution of 24 MP; a watertight metal cylinder; a modem required to allow connection between the camera and control panel
- supply power cables required for connection to camera and power supply
- a battery in lead, 12 V
- control panel: a laptop connected to the camera through a fiber optic cable in order to see in real time what the camera observes
- a structured light 3D scanner: ARTEC EVA with a 3D accuracy of 0.1 mm
- Matlab Camera Calibrator for the in air calibration of the camera system
- OpenCV and Octave used to undistort images and apply correction of Pinax algorithm
- Metashape by Agisoft used to obtain 3D reconstruction

The approach followed can be split into three parts and it is sketched in Figure [6.7](#):

1. preliminary actions
2. evaluate section
3. validation process

In preliminary actions stage, since Pinax algorithm requires an in air calibration, a chessboard has to be printed with the aim to calibrate the camera, in such a way that at the end of the process the K calibration matrix of the camera and the distortion coefficients are provided. The mechanism behind this calibration is as described in the section [5.2.4](#). The

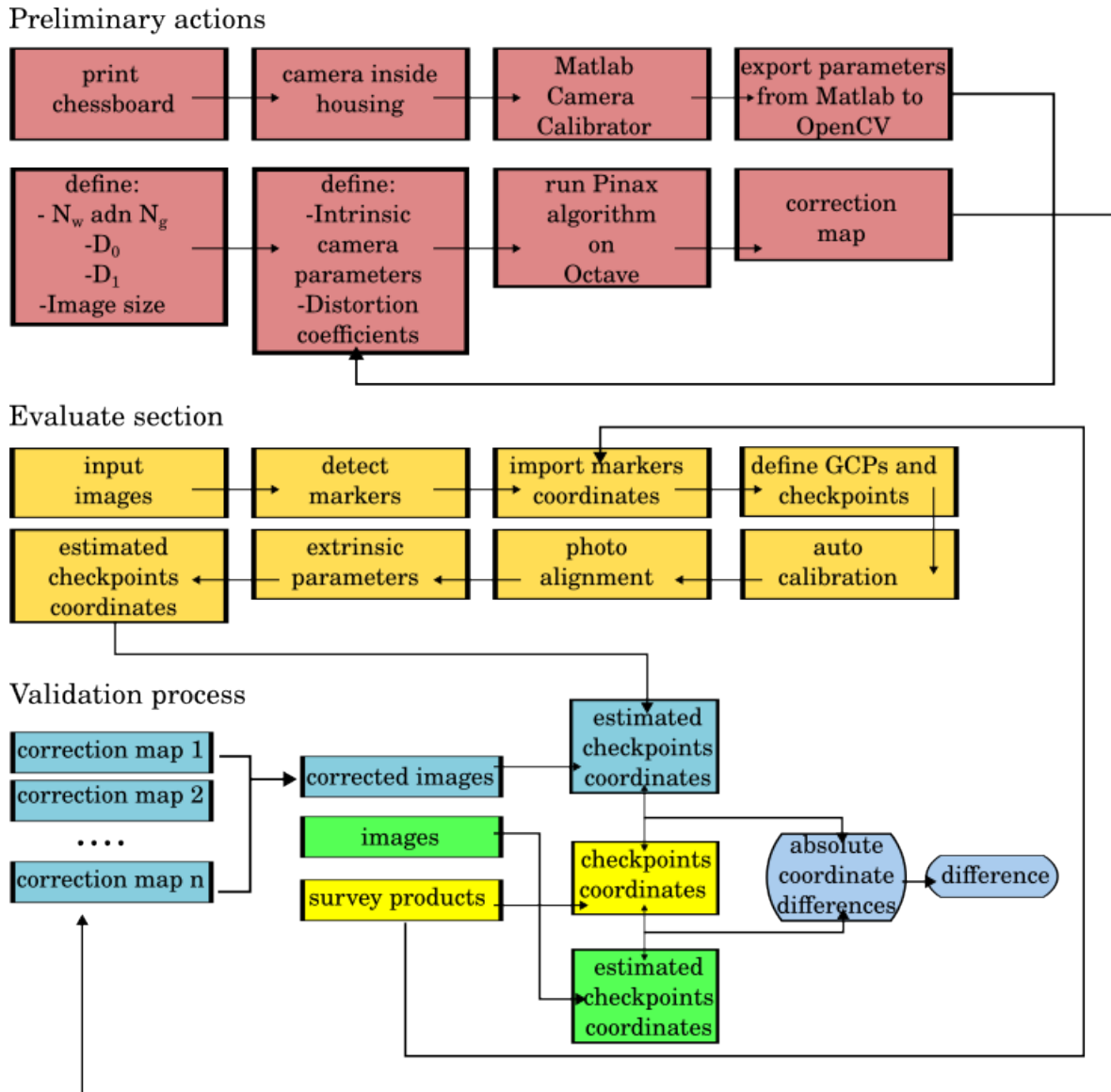


Figure 6.7: Scheme of the steps carried out to validate the calibration and correction flow. The flow is split in three parts, each of which is interconnected. Pay attention to different colors, they are associated to different elements.



Figure 6.8: Camera system, leaning on top of a stand, during the in air calibration. A chessboard was placed on a rigid plane and the size of the alternating squares was given as input for the calibration algorithm.

camera system was separated from Blucy and moved where the chessboard was placed, Figure 6.8 shows the in air calibration phase computed in the Geomatic laboratory of the faculty of engineering at the University of Bologna. Zhang et al. in [79] provide a flexible technique to easily calibrate a camera, it only requires the camera to observe a planar pattern show at least two different orientations. This procedure is embedded in the Matlab Camera Calibrator App. K and the distortion coefficients are exported from Matlab to OpenCV since Pinax algorithm exploit OpenCV routines and the scientific environment such as Octave. Fortunately, a specific Matlab routine is dedicated to transform K and distortion coefficients computed within its environment to be used in OpenCV; the routine declaration is very simple `[k,dist_coeff]=cameraIntrinsicsToOpenCV(intrinsics)`; it provides the objects contained in the left square brackets starting from that within brackets, coming from the Matlab Camera Calibrator App. Once these parameters are computed, they are immediately used as input parameters in the computation of the correction map. In particular, Pinax relies on a special open-source interface which include either the standard routines of Octave either the OpenCV's API. The standard routines carry out tasks related to memory allocation and simple operation as well as allow the creation of computation routine useful to intermediate steps. While for what concerns OpenCV's API, they are useful to perform imaging computation, such as image rectification as well as carry out operations related to images. In order to allow communication between Octave and OpenCV, mexopencv (<https://github.com/kyamagu/mexopencv>, visited 26/01/2023) interface must be implemented. The package, free for charge, provides Octave

MEX functions that interface with hundreds of OpenCV APIs. Also the package contains a C++ class that converts between Octave's native data type and OpenCV data types. The package is suitable for fast prototyping of OpenCV application in Octave, use of OpenCV as an external toolbox in Octave, and development of custom MEX functions. It is necessary to build and install `mexopencv` depending on the operative system. In general the steps necessary to implement `mexopencv` on a Linux machine are:

1. Install Octave 4: `sudo apt-get install octave liboctave-dev`
2. Install some build dependencies: GCC, zlib, JPEG, PNG, TIFF, FFmpeg, Video4Linux, DC1394, Xine, gPhoto, GStreamer, GTK+2, TBB, Eigen, and BLAS/LAPACK/Atlas
3. Download, compile and install OpenCV+`opencv_contrib`:

```
wget -O opencv-3.4.1.zip https://github.com/opencv/opencv/archive/3.4.1.zip
wget -O opencv_contrib-3.4.1.zip https://github.com/opencv/opencv_contrib/archive/3.4.1.zip
mkdir ~/cv/build && cd ~/cv/build
cmake -G "Unix Makefiles" -DOPENCV_EXTRA_MODULES_PATH=~/opencv_contrib-3.4.1/modules ~/opencv-3.4.1
make
sudo make install
```

4. Download and compile the MEX-files for Octave:

```
cd /cv
wget -O mexopencv-master.zip https://github.com/kyamagu/mexopencv/archive/master.zip
cd /cv/mexopencv
make WITH_OCTAVE=true WITH_CONTRIB=true all contrib
```

It may happen that errors arise during the compilation phase, to be able to solve them, it is highly recommended to follow the `mexopencv` guide mentioned above. It exists another procedure to exploit Pinax algorithm, that is through Robotic Operative System (ROS). A more complicated but efficient way to obtain a correction map in very little time. Pinax algorithm is encapsulated in three main routines provides as Octave file:

- RayTrace.m: with this routine it is possible to reconstruct the light path that each ray follows from the image plane to the world points
- RefractedRay.m: with this routine the refractive effects are computed taking into account the underwater housing
- SolveForwardProjectionCase3.m: with this routine the 13th degree polynomial is solved to discover the refractive effects following the results obtained in [25]

While the main file, called FoundMap.m, i.e. the one allocating memory and calling up the functions described above, was modified to feed the validation process and thus implement the iterative process to estimate the optimal d_0 . The structure of FoundMap.m is the following:

- Based on the image size, it allocate enough memory to save image data in homogeneous coordinates
- K matrix calibration and distortion coefficients $(k_1, k_2, p_1, p_2, p_3)$ are passed as input as well as physics parameters, such as refractive indices and other geometric characteristics
- A counter is initialised. It represents d_0 . A value of 1.5 mm is taken as starting value to which will be added/subtracted 1 mm
- A loop statement in which the counter is increased/decreased based on a certain increment is performed. At each cycle a new correction map related to the actual d_0 is produced. Indeed for each image points ($N = width \times height$) the light path is computed taking into account the change of direction of the light
- Inside the loop statement image points are reprojected with the changes provided in the previous step and the corrected map is produced

At the end of the process several correction maps are produced. With an ad-hoc script, called Applymap.m, the different correction maps are applied to the same data set. The one with the least deviation from the reference model, that in this case is represented by the coordinates measured with SLS, is elected as the one that best describes reality, with which a more reliable

estimate of d_0 is consequently associated. `Applymap.m`, first lists the image files in the folder, then for each one it creates a correct image with a similar name. In the processing phase, it was very import to pay attention to metadata included in the images. Since, Pinax creates, starting from a single image, another corrected one but metadata are lost. To solve this problem, it was necessary to transfer image metadata from the original images to the corrected ones exploiting `exiftool` library (<https://exiftool.org/>, visited 26/01/2023) and typing the following bash command:

```
for FILE in IMG_*.JPG;
do b=$(echo "remapped_${FILE}");
exiftool -TagsFromFile $FILE "-all:all all:all" $b;
done
```

Where `FILE` represents a pointer to a current image of JPG format, `b` represents a variable in which is stored the name of the corresponding corrected image, while `exiftool` is the core of the snippet and it allows to transfer from `FILE` to `b` all the metadata. If the metadata were not copied in the corrected images, results were awful. This shows how important metadata is within a photogrammetric computing environment. Pinax algorithm requires as input several parameters, including:

- K calibration matrix and distortion coefficients
- N_w refractive index of water (1.55 salt water, 1.335 fresh water); it is based on the wavelength, so it may vary by temperature and pressure
- N_g refractive index of the glass of which it is made the underwater housing; it is 1.60 a value derived from experimental tests performed in the optics laboratory of the faculty of Physics at University of Bologna
- d_0 distance from the center of projection to the glass window, unfortunately unknown since there is no construction scheme of the underwater camera system. Its value was estimated by means of an iterative method knowing, however, that its value was between 0.5 and 2 mm

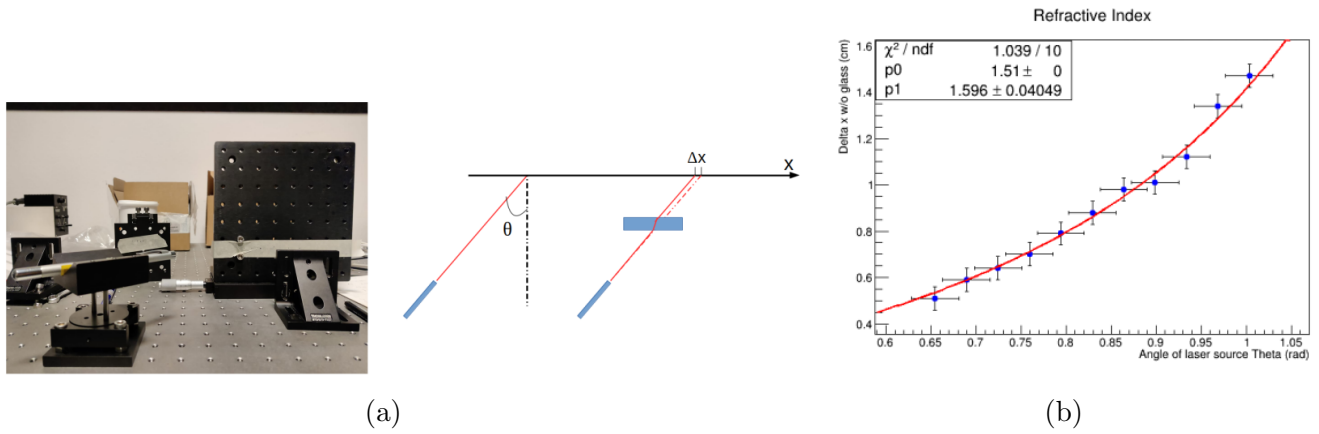


Figure 6.9: Determination of the refractive index of the underwater housing flat glass port. In (a) is showed the experiment setup and a schematic view of the experiment's mechanism. The deflection that the laser beam undergoes due to the abrupt change in refractive index is $\Delta x = d\{tan(\theta) - tan[asin(\frac{sin(\theta)}{N_w})]\}$ where d is the thickness of the glass port. In (b) it is report the experimental data fit, where $p0$ and $p1$ are respectively d and N_w , which is the unknown parameter. Its computed value is $N_w = 1.60 \pm 0.04$. Thanks to Sara Rabaglia which makes available the experiment tools.

- $D_1 = 1.51 \pm 0.01$ cm thickness of the glass
- image size: W=6048 and H=4024

Once the in air calibration was carried out and established the input parameters for Pinax, it is possible starting the computation of the correction map. The time taken to compute it is variable and depends on the performance of the machine on which it takes place as well as the type of computing platform. For what concerns the evaluate section stage, it needs to consider that it occurs within the Metashape Agisoft software. The calibration grid consists of circular target coded mounted on a rigid structure automatically identifiable by the software. Further, the software allows to insert the markers coordinates and to manage them as if they were control points or GCPs and checkpoints. Once defined these features, it is possible to start the photo alignment stage which provides the camera pose and finally the estimated checkpoints coordinates. While for what concerns the validation process, since the d_0 parameter is unknown, an iterative method is applied. Several correction maps as d_0 varies were generated, among these, the one that provided the lowest associated errors was elected as the one actually used. Therefore, to this one corresponded to the value of d_0 that is closest to the actual value. At the end of applying the iterative method, the value of optimal d_0 is estimated to be 1,428 mm. At

this point, since the correction map is known, it can be computed the corrected images via Pinax algorithm. Both for the uncorrected (green color Figure 6.7) and for corrected images (celestial color Figure 6.7) the standard SfM pipeline was carried out in order to estimated in both cases the checkpoint coordinates. In this case it was carried out the survey of the calibration grid by means of a SLS, which provided the checkpoints coordinates. Therefore, absolute differences of the coordinates were computed between estimated checkpoints coordinates and the checkpoints coordinates:

$$|\Delta\bar{X}| = |X_{est} - X_{real}|$$

where \bar{X} indicates a generic spatial coordinate for corrected or uncorrected images. In order to validate the model, the difference of the coordinates

$$\Delta X = |\Delta\bar{X}_{uncorrected}| - |\Delta\bar{X}_{corrected}| > 0$$

has to be positive. Without losing any generality, it can be said that the results obtained have all passed the validation process, so it is possible to say that in the case of the calibration grid, the images treated by the Pinax method are more accurate than the untreated ones. In general the tools that are needed to implement such procedure for a single correction map can be contained in a folder, as depicted in Figure 6.10 where it is represented a standard working directory in which trough a terminal is possible to invoke ApplyMap.m to rectify data set. Coordinate difference calculations, on the other hand, are performed using the basic functions of Octave once the data have been obtained.

Results

In this section results obtained in the underwater experience for the case of the calibration grid are reported. Among the fifteen available targets, seven were chosen as a control points while the remaining as a checkpoints. Once the photogrammetric process described in the previous paragraph is initiated, estimates of target centres are obtained. The difference between the measured value, understood as that generated by the measurements taken with SLS, therefore

```
mc@mcUnibo:~/Desktop/workdir$ tree .
.
├── ApplyMap.m
├── dataset
│   ├── IMG1.JPG
│   ├── IMG2.JPG
│   └── IMGn.JPG
├── FoundMap.m
├── RayTrace.m
├── RefractedRay.m
├── SolveForwardProjectionCase3.m
├── XB3MapCenterX.txt
└── XB3MapCenterY.txt
1 directory, 10 files
```

Figure 6.10: Tree bash command called to display the standard structure of a working directory. For the sake of simplicity, data set consists of only three images. The correction map is split in X and Y spatial coordinates and they are represented by .txt files.

Id	X(cm)	Y(cm)	Z(cm)	Id	X(cm)	Y(cm)	Z(cm)
1	2.03	1.07	2.13	1	1.44	0.78	1.47
2	0.24	-0.65	-3.62	2	0.20	-0.42	-3.17
6	-2.41	-1.47	1.19	6	-2.00	-0.84	1.3
8	-0.36	2.58	-0.05	8	-0.01	2.05	0.07
10	-0.07	0.40	-2.92	10	-0.16	0.41	-2.16
12	0.27	-1.58	2.38	12	-0.10	-1.61	1.58
16	0.31	-0.34	0.88	16	0.71	-0.36	0.91
Total	1.21	1.37	2.21	Total	0.98	1.11	1.77

Table 6.1: Residuals of the errors associated with the control points in the case of uncorrected images(left) and corrected images(right). In the first column there's the target Id. There is target Id 16, however the target number is 15, because target number 7 was affected by a misprint.

considered as the measurement technique that provides higher accuracies than the instrument being tested, and the estimated value represents the residual, that is, the deviation of the estimated value from the actual value. The results that follow are therefore concerned with the residuals associated with the centres of the targets. By looking at the tables in Table [6.1](#) it is possible to observe the errors associated to control points in each of the three spatial coordinates (last three columns of each table). In most cases results carried out by means of Pinax algorithm perform better, as we can see from the Total row. In this case, it shows how the estimates obtained with Pinax are 3-4 mm closer to the true value. While Figure [6.11](#) shows the matrix of relative difference between estimated target coordinates obtained with SLS and those obtained with underwater photogrammetric technique. These matrices are symmetric and characterized by the main diagonal null. At a glance, just by looking the color scale it is possible to appreciate the gain of about 1-2 mm with the use of Pinax algorithm. Further,

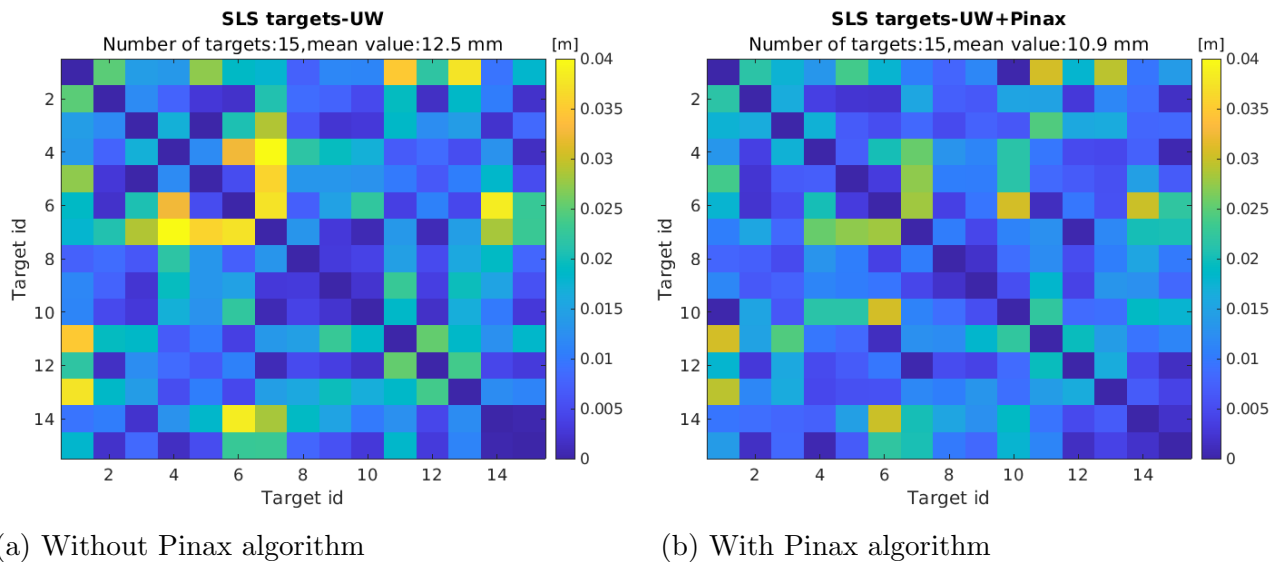


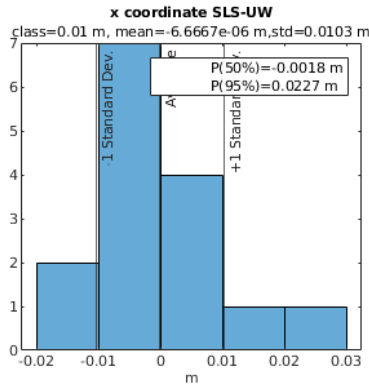
Figure 6.11: Matrices of relative distance differences among SLS target coordinates and underwater photogrammetric extracted target coordinates expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles.

it is possible to observe a pattern on the side diagonal in the left image of Figure [6.11](#) which instead is not visible in the image on the right. It could indicate the lack of a systematic error probably removed with Pinax algorithm. Shown in the Figure [6.12](#) are the trends of the frequency histograms associated with the spatial coordinates of the residuals, both in the case of the photos corrected with Pinax and in the case where the algorithm was not applied. Although the data sample is not large enough to provide reliable statistical indicators, the main statistical features were nevertheless provided. Addressing the description of the histograms in columns, so as to compare for each spatial coordinate both cases, it can be stated that for the x-coordinate the maximum dispersion associated with the results corrected with Pinax is an order of magnitude smaller than those obtained without Pinax. Further the mean value of the Pinax results is more centered around zero than the other ones, as can be seen from the very small values taken by the mean. In both cases, the measurements of differences range between -0.02 m to 0.03 m. For the y-coordinate in both cases there's the same order of magnitude for the standard deviation, however that associated to corrected images is lower. Classes are different from each other by exactly 1 cm, while the average values are comparable. The same considerations can be made for the z-coordinate, except that the classes are of the same magnitude. In this case there are differences associated with the elevations that have

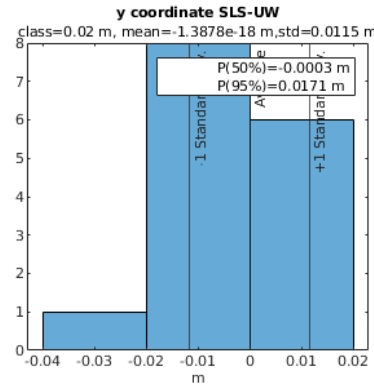
values a few centimeters (2-3cm) larger than the planimetric coordinates. An already expected behavior found in the data is the fact that the highest residuals are those associated with targets that are on the outer edges of the calibration grid. This occurs because the largest deformations in underwater images are mostly seen at the edges of the images. This can be best observed in Figure [6.13](#) where the largest ellipse errors are precisely those located at the outer edges of the calibration grid. To better appreciate this behavior, their amplitude was magnified tenfold. A color legend is also showed and it represents the magnitude of the errors. Below the considerations made before, it can be argued that reconstructing underwater objects from corrected underwater photos using a method that takes into account the path a light ray follows from the moment through water, glass, and air contained in the underwater housing can improve the accuracy of reconstruction by about 2 to 3 cm. This shows that, however the distortion coefficients can compensate for some refractive effects, for high-precision applications, algorithms such as Pinax's are needed. The above applies within a context in which it is not possible or practicable to take advantage of the use of a submerged local geodetic network.

6.3.2 Calibration frame

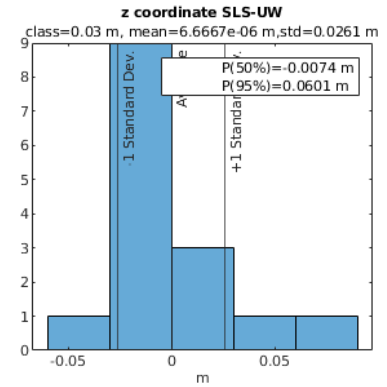
The same camera system used in the previous experiment were used also in case of the experience related to the calibration frame. The configuration of the underwater housing, which contains the camera, did not change, therefore the Pinax input parameters remain the same. Except for the water refractive index, which now is related to salt water instead fresh one. Thanks to the sensors installed on Blucy, precisely the CTD one, it was possible to derive the refractive index linked to the salinity estimate of water by means of tables in [\[80\]](#). Regardless of the setup parameters, a good practice to improve the quality of the Pinax corrections is to carrying out the calibration of the camera at about the same distance from which the underwater scene is observed [\[2\]](#). Indeed, it has been seen that when calibration is performed around this distance at which the object to be detected is located, the the pinhole camera model performs better allowing an optimal rectification of the distorted underwater images. Following this logic, the calibration frame experiment was carried out in an appropriate spatial



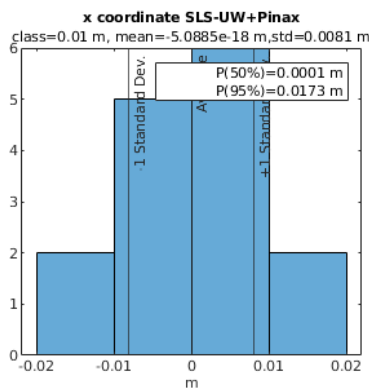
(a) Frequency histogram of the difference of the x coordinates extracted from the standard photogrammetric technique and those obtained with SLS.



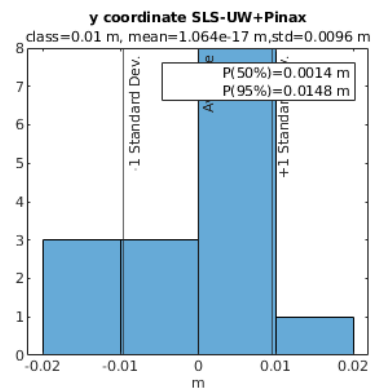
(b) Frequency histogram of the difference of the y coordinates extracted from the standard photogrammetric technique and those obtained with SLS.



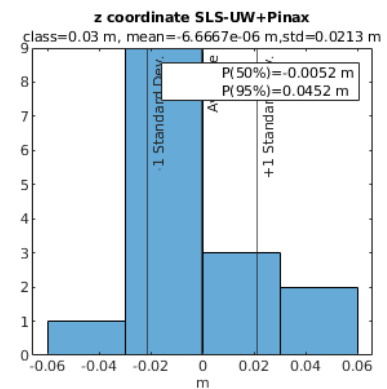
(c) Frequency histogram of the difference of the z coordinates extracted from the standard photogrammetric technique and those obtained with SLS.



(d) Frequency histogram of the difference of the x coordinates extracted from the underwater photogrammetric technique and those obtained with SLS.



(e) Frequency histogram of the difference of the y coordinates extracted from the underwater photogrammetric technique and those obtained with SLS.



(f) Frequency histogram of the difference of the z coordinates extracted from the underwater photogrammetric technique and those obtained with SLS.

Figure 6.12: Frequency histograms of the difference of the center target coordinates between those obtained by means of SLS and those estimated with photogrammetric techniques. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value.

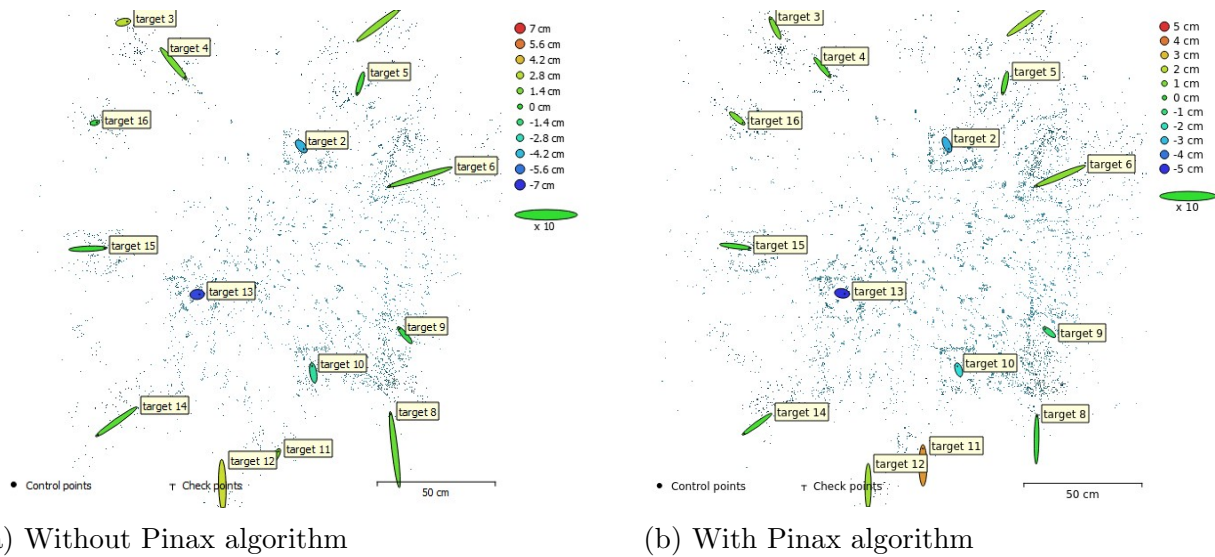


Figure 6.13: Image shows control and check point locations and error estimates. Z error is represented by ellipse color. X,Y errors are represented by ellipse shape. Estimated locations are marked with a dot for control points or T-crossing for check points.

configuration, which is attempted to come as close to reality as possible. For this reason, during the Techera mission it was chosen an area near the 'C1 Buoy', in front of the Miramare Castle, with a depth of about 7 m to carry out the experiment. Having learned the lesson from the first attempt of the experiment, this time a height from the seabed was chosen that was not too high so as to decrease as much as possible the difficulties related to the assisted deposition on the seabed of the platform. This configuration can be compared to when a camera equipped on a ROV device carries out an underwater survey of some cliffs near the shore, where the water height is not too high. It is important to say that, Pinax algorithm does not need for a underwater calibration, but from the point of view to show how 3D reconstructions perform better with its application it is required an underwater survey with at metric reference, which is represented by a calibration platform. In general the same scheme depicted in Figure [6.7](#) was applied to obtain the results, however the iterative process was omitted as well as the preliminary actions stage, except for the creation of the correction map. The mission started on November 30 around 10 am and it last about 3 hours. Unfortunately, the time when the survey mission was organized is not the most appropriate, as it was organized during a winter period with numerous rain events. Isonzo, Timavo and others minor rivers end their course in the Gulf of Trieste, pouring debris and material carried by the current into its waters. As result, the turbidity of the Trieste Gulf increases, by dramatically decreasing visibility under-

water. It must be kept in mind that organizing an underwater survey mission with this kind of equipment, such as a boat and the associated UUV, takes a huge amount of organization and resource consumption for which a long preparation time is needed. That is why the mission was organized more with these factors in mind, and was therefore carried out in the November period. The major difficulty met during the mission is related to the navigation and control of the UUV prototype. Given its nature as a prototype, Blucy is difficult to pilot especially when the target is represented by an object of small extent (1.5 m^2). Stationing above the calibration frame, taking pictures from different vantage points is a very complex operation that required no small amount of effort on the part of those piloting the UUV and the diver supporting its maneuvers. To this end, a communication language was created based on diver gestures and artificial lights installed on Blucy. As mentioned before, a distance of about 3 m has kept from the platform, in such a way that above the UUV there was about 4 m of salt water. One of the first picture taken by the camera and enlightened by the artificial light is showed in Figure [6.14](#). From the hundreds of pictures taken, 76 were selected that met minimum quality standards, such as good focus, lack of obstacles along the line of sight, and no blurring effect. Therefore, the starting dataset for making corrections and starting 3D reconstruction consists of 76 underwater images of the calibration frame. The main instruments and devices used for accomplish the surveying mission were:

- calibration frame
- ropes for slowly deposition of the calibration platform
- UUV equipped the underwater camera system
- a scuba diver necessary to maintain Blucy over the small calibration platform due to low visibility caused by high water turbidity
- a vessel useful for managing the movement of crew and all instruments from the harbour to points of interest, as well as being the control center where the Blucy control panel is located

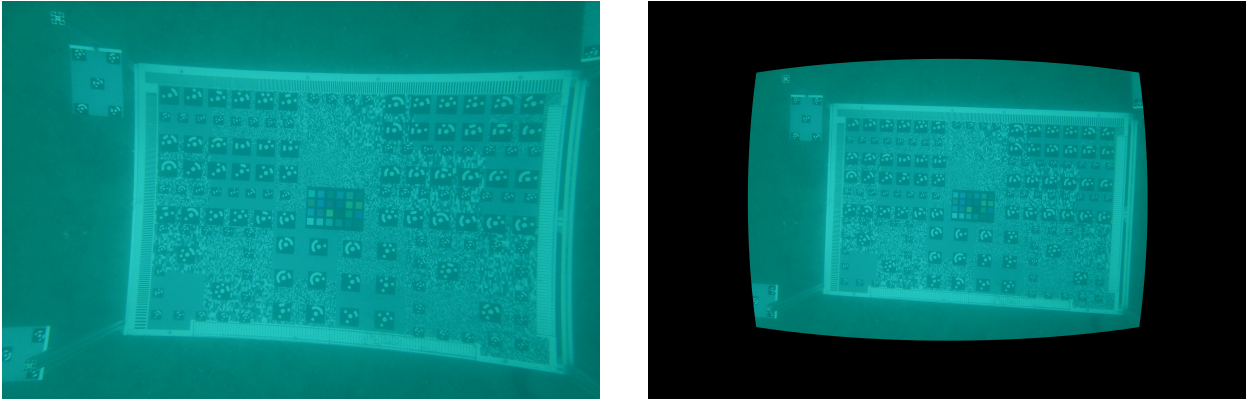


Figure 6.14: First pictures obtained by means of Blucy. In (a) there is the snapshot of the calibration frame. In (b) the same thing except for the corrected distortions that are strongly removed as you can see from the perfect rectangular shape of the platform. It has to be noticed the two resolution bar place on two adjacent sides of the platform. They are designed starting from the smallest GSD considered, please refer to [2.2.2](#) section for more details. During the survey they are used to stabilize the focus of the camera.

- positioning systems without which it would be impossible to determine the position of the vessel and that relative of the UUV

6.4 Analysis of the results

This section describes the post-processing phase subsequent to data collection of calibration frame experience, already described in previous sections. The 76 underwater pictures are imported on Precision 7550 laptop with a total 64 GB of ram and equipped with a graphics processing unit called Quadro RTX 4000 Mobile by NVIDIA (8 GB). These images were rectified using Pinax algorithm exploiting OpenCV library and Octave, as a computation environment. At this point, we had two separated sets of images that show the same object scene, one rectified and the other not, depicted in [Figure 6.16](#) as the primary inputs of the scheme. Both sets of images were subjected to the same incremental SfM pipeline in order to make them comparable, in particular the main workflow parameters applied in Agisoft Metashape are shown in [Table 6.2](#). Attention must be payed to the coded target installed on the calibration frame. They are inverted (white on black) targets of circular 14 bit. This choice was made for two reasons: the marker type can be recognised by different photogrammetric calculation environments, while the fact that they are inverted is due to the fact that underwater the white colour would have

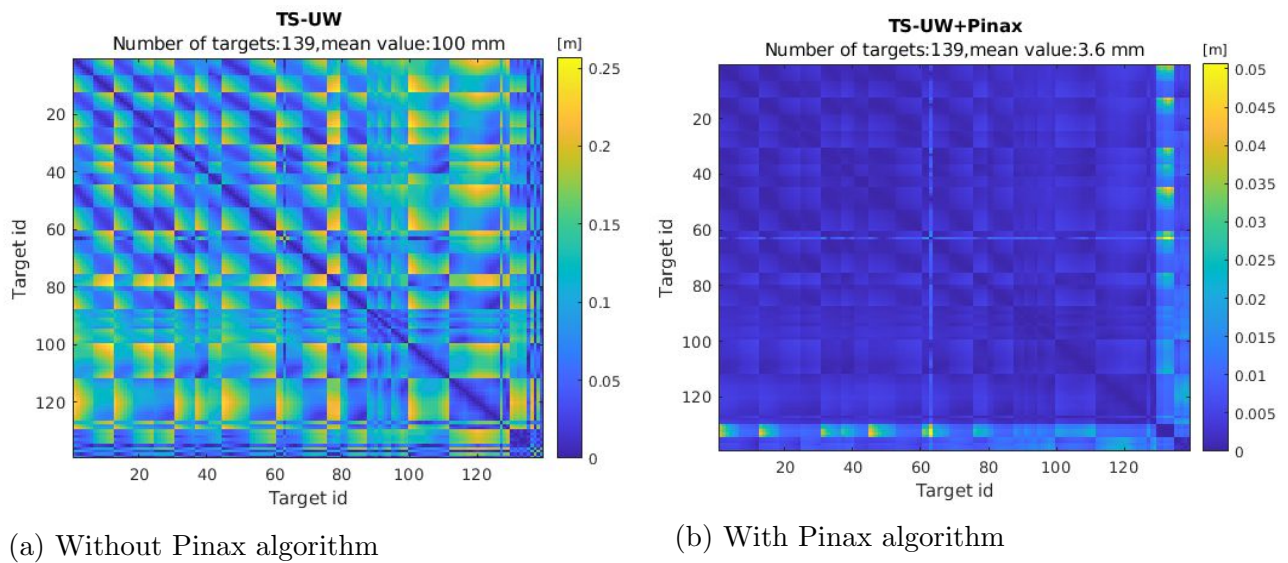


Figure 6.15: Matrices of relative distance differences among TS target coordinates and underwater photogrammetric extracted target coordinates expressed in terms of heatmap. The corresponding target number is shown on the coordinate axes, and global information on measurement performance is provided as subtitles. It has to be noticed that targets at altitude are those with a target id greater than 125, to which correspond higher measurement errors.

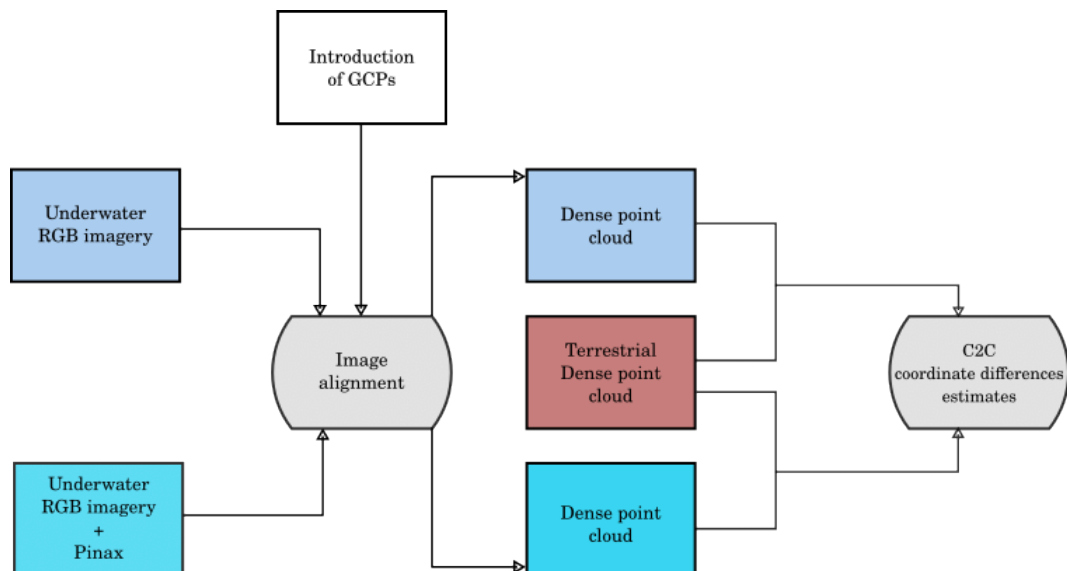


Figure 6.16: Data processing for underwater images. GCPs selected as ground truth are those measured with TS.

Workflow			
Align photo		Camera settings	
Accuracy	Highest	Camera Model	Nikon Z6
Pair preselection	Disabled	Resolution	6048x4024
Point limit	40000	Focal length	24 mm
Aligned cameras	76	Precalibrated	Disabled
Markers	151	Pixel size	$5.96 \times 5.96 \mu m$
Tie point limit	10000	Type	Full frame
Build dense cloud		GCP settings	
Quality	High	Marker accuracy (m)	0.005
Depth filtering	Aggressive	Scale bar accuracy (m)	0.001

Table 6.2: Parameters workflow used in Agisoft Metashape.

created a fogging effect along the contours, degrading the quality of the target centre measurement. In order to invert the target color, as well as to design virtually the frame an open-source graphic environment (InkScape) it is used, and to obtain inverted targets, normal ones were transformed into converted exploiting the invert colors function. Figure 6.18 shows the camera poses with respect to the calibration frame. Blue squares represent the camera orientation at the time the photo was taken. By considering all the photos within the data set, an average flying altitude of 2.33 m is detected as well as an average ground resolution of $0.581 \frac{mm}{pix}$, an amount very close to the maximum GSD that was considered at the design stage. Photos were taken when Blucy was in ROV mode, thus while Blucy was piloted manually and therefore the camera positions distribution was not uniform because of the reasons explained in Section 6.3.2. After the characteristic stage of the image alignment, the target coordinates measured (about 15 measurements) are imported in the photogrammetric software with respect to GCP settings related to Table 6.2. It allowed to estimate the coordinates of the remaining targets, which were then compared with those measured with TS. The procedure of identification of targets and computation of residues is carried out and the results are shown in Figure 6.19, where the 3D model is covered by several labels associated to each target uniformly distributed on the calibration frame and error ellipses show the amount of deviations from the reference model that ranges from -5 mm to 5 mm. Then we started the procedure to obtain the dense cloud of the detected object followed by the generation of a texture to obtain a final product that also had the colors corresponding to reality. Therefore, at the end of the pipeline were provided two colored dense clouds, which are compared with the one provided by the terrestrial



Figure 6.17: Textured point cloud provided from the standard SfM pipeline in Agisoft Metashape that represents the calibration frame placed underwater during the Techera surveying mission.

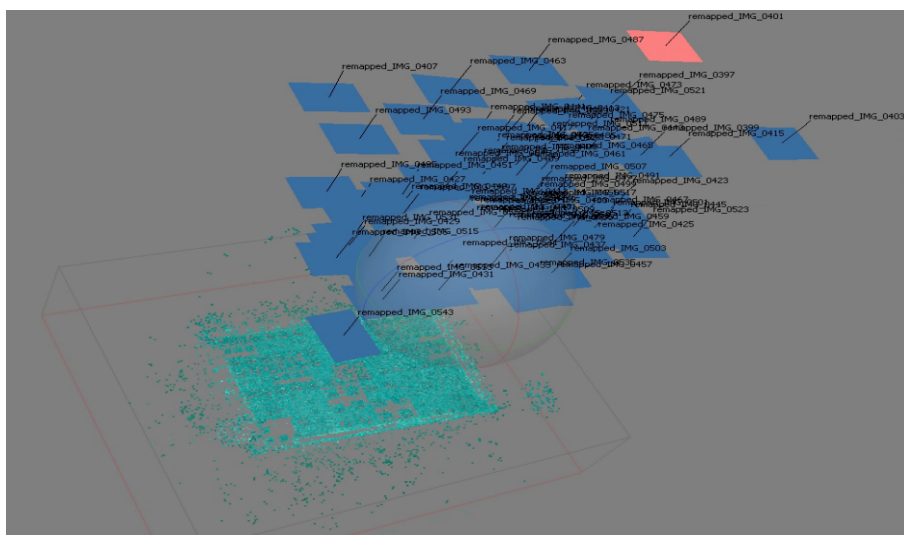


Figure 6.18: Camera positions. An average flying altitude of 2.33 m is computed.

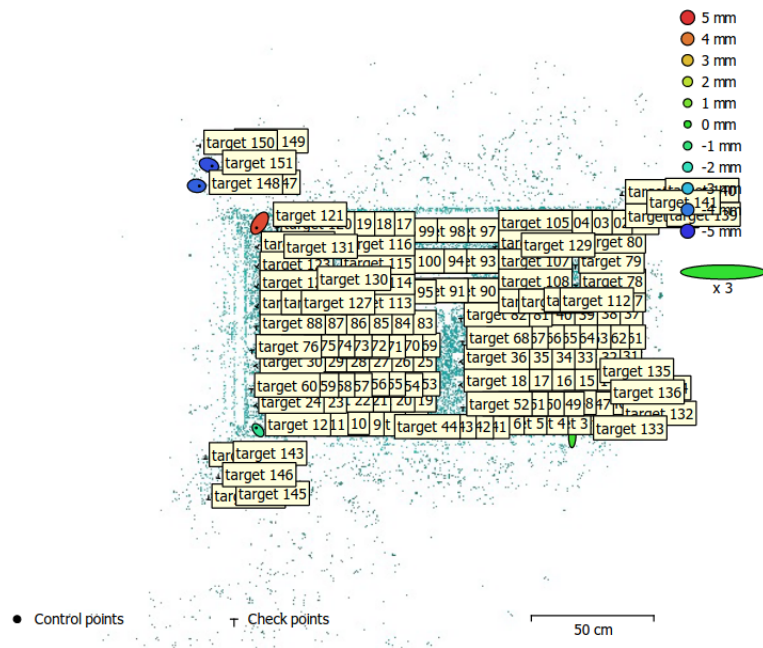
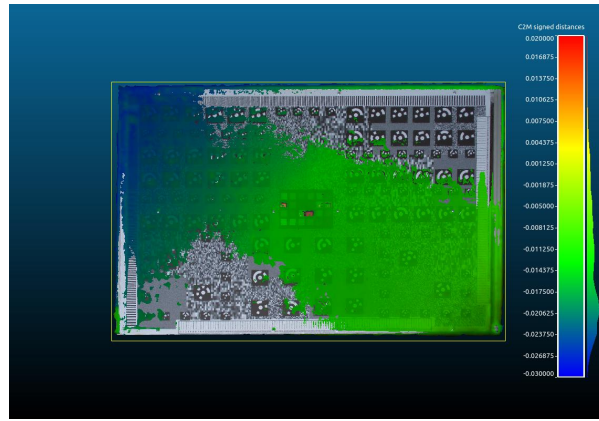


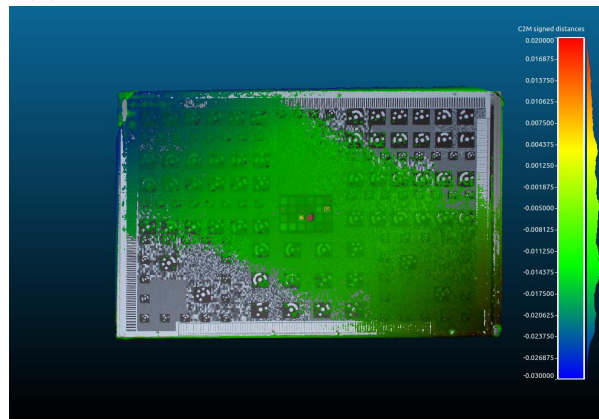
Figure 6.19: GCP locations and error estimates. Z error is represented by ellipse color. X and Y errors are represented by ellipse shape. Estimated GCP locations are marked with a dot or crossing. Case of underwater images corrected with Pinax.

survey in Section 2.3.2. Figure 6.16 tries to summarize this procedure in order to enhance the main steps. The textured point clouds of the calibration frame underwater surveyed is showed in Figure 6.17. It can be noticed that the colors are very different from those of Figure 2.7 because of the presence of water. Green is the color the overlaps the most with the others, as already explained in Section 3.2. The purpose of this section is to demonstrate that the 3D reconstructions of the underwater scene produced from Pinax-corrected photos are more accurate than the case where Pinax was not applied. In this context a series of results will be showed with the intention of validating the previous hypothesis. Figure 6.15 shows the matrices of relative distance difference among targets between TS target coordinates and those extracted with photogrammetric technique. Therefore TS target coordinates were used as ground truth. Since this type of graph has already been used several times, its characteristics are not repeated again. At first glance, what stands out to the eyes is the color scale and its associated values. Qualitatively, the two heatmaps are very different from each other: one almost entirely blue, which therefore indicates very small distances in the coordinates, while the other has all the colors of the color scale, which therefore indicates the fact that the distribution of the distance differences of the coordinates is not uniform but random. Quantitatively, the heatmaps have

the same color scale but with different associated values. Indeed it is possible to notice that to yellow color are associated a distance difference of 25 cm and 5 cm, respectively for distorted images and undistorted images. Those are maximum values of the distance differences, therefore it clearly appears that the 3D reconstruction using Pinax algorithm perform better than without its application. The highest distance differences for image (b) contained in Figure [6.15](#) are for targets positioned at elevation relative to the plane of the calibration frame. Considering the 139 relative distances, an average value of them was derived and it is found that for the photos taken with Pinax it is 3.6 ± 0.2 mm while for the untreated photos, the average value is around 10 ± 1 cm. This would testify to a gain of about two orders of magnitude in the accuracy of the distance difference if one processed the images with Pinax, compared to if one did not. In this context, where the use of a local geodetic network is avoided, this kind of result is very promising. By applying ICP algorithm is possible to compare two point cloud in order to highlight differences between them. Usually one of them is used as reference while the other one represents the term of comparison. The algorithm looks for minimize the difference between two clouds of points and it is used to reconstruct 2D or 3D surfaces from different scans. In this context it is used two overlap the terrestrial cloud point with the submerged ones and detect where there are the main differences. It does not represent an absolute way to display them, but only a way among others. Therefore it is a qualitative method to spatially appreciate the difference between the Pinax treated images and not. Figure [6.20](#) shows the application of ICP algorithm to the point cloud pairs, which are both formed by the terrestrial reference and the terms of comparisons by the reconstructed point cloud starting with the Pinax-corrected images and the incorrect ones. In the case of Pinax application, the distribution of difference values is more centered and covers an oblique band extending from the upper left vertex to the lower right vertex, with values that belong to the green color. While for the case without the application of Pinax algorithm, values are shifted to the negative side and affect a larger portion than the previously identified band. Further, it is possible to notice that values associated to blue color are found mainly near the top left vertex. Hence, the latter statements testify to the fact that the point cloud formed from the Pinax-corrected photos fits better to the terrestrial point cloud. This behaviour is best explained in Figure [6.21](#), where

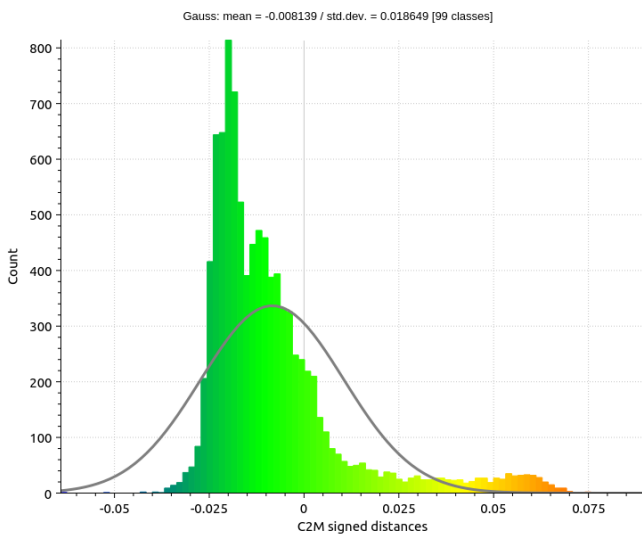


(a) Without Pinax algorithm

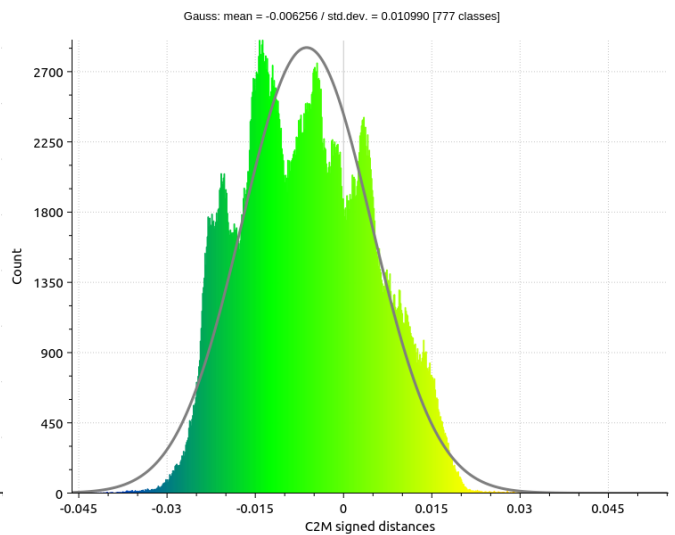


(b) With Pinax algorithm

Figure 6.20: Cloud-to-cloud distance computed with ICP algorithm. We use as reference the terrestrial point cloud computed in the terrestrial surveys obtained by means of the standard photogrammetric technique. Values associated to color scale are reported in metres.



(a) Without Pinax algorithm



(b) With Pinax algorithm

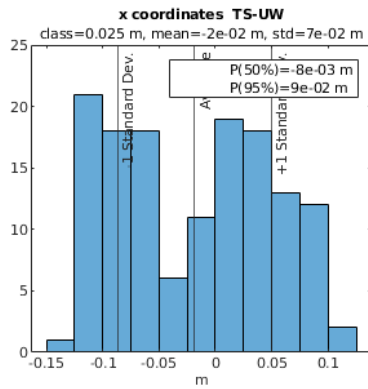
Figure 6.21: Frequency histograms of the cloud to cloud distance with the associated Gauss curve. Mean and standard deviation provided as subtitles.

the frequency histograms of the signed distances are shown. Indeed, Figure (b) [6.21](#) represents better the Gauss curve than (a) image. Further to the values of the tails of the Gauss curve, for the case without Pinax algorithm, are associated classes which have values distributed between -5 cm and 7.5 cm, against those for the with Pinax, which are distribute around - 4cm and 3 cm. Finally by looking to the number of classes obtained for the two cases it can be seen that the number of points forming the point cloud in the Pinax case is much higher than in the case without Pinax, testifying to the fact that in the case where the correction algorithm is applied, the feature detection phase succeeds in finding more homologous points. Better results are obtained even in terms of mean value and standard deviation as you can see from the subtitles of the Figure [6.21](#), where the mean value of cloud to cloud distance of the Pinax case is lower of 2 mm than the case without Pinax. While the standard deviation in the Pinax case is lower by almost an order of magnitude with respect to the case without the application of Pinax algorithm. The standard way to compute distances, and the way that we selected, between two point cloud is the nearest neighbor distance method, that is an algorithm based on k-nearest neighbors algorithm a non-parametric supervised learning method invented in the 1950s and become very famous in recent decades due to developments in the field of computer science[\[81\]](#). In the distance computation, as mentioned before, we use as reference the terrestrial point cloud since an optimal density cloud it required by the algorithm to work good and in our case the point cloud fulfills this requirement. In the case the optimal density had not been acceptable, another algorithm different from k-nearest neighbors could be applied, such as least square plane or Delaunay triangulation. This way to analyse result is not so robust, since the ICP algorithm looks for the best solution as a black box, therefore to have an absolute idea of the analysis it is necessary to pay attention on the frequency histogram of the difference of the coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment. Figure [6.22](#) represents that we are talking about. We proceed to analyze the graphs by columns, so we start from the first column where x spatial coordinate is taken in exam. First of all, it seems that data coming from the 3D reconstruction without Pinax algorithm is distributed into two normal distribution (as two normal distribution that overlap in the average value), and this behaviour repeats also for the other coordinates, while for the

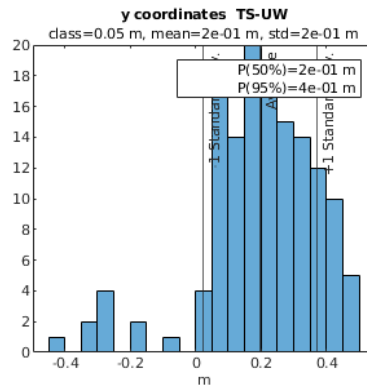
case with Pinax the distribution seems to be unique. Further, either the mean value either the standard deviation is one order of magnitude lower for the case with the application of Pinax algorithm. Actually, it goes from a centimeter level (no Pinax) to millimeter one (Pinax). It is testified also by the percentile values. Attention has to be paid also in the class width of the frequency histograms. The fact that the data coming from the 3D reconstruction with Pinax algorithm, also in this case, lie better on the Gauss curve can be appreciated. In the second column, the same previous assumptions can be made, but actually the mean value of about tens of centimeters is completely decreased at millimeter level in the Pinax reconstruction. While for what concerns standard deviation, Figure (b) and (e) [6.22](#) show a difference of one order of magnitude. The third and final column represent the z coordinates. Even in this case the scenario occurs, where a gain in mean value and standard deviation is shown in Figure (c) and (f) [6.22](#). The main value of these difference for the Pinax case is -1 mm while for the other case is 3 cm, and for the standard deviation respectively 2 mm and 6 cm. In general, looking at the x-axis of all the graphs, it is clear that the range of variation of the coordinate difference is an order of magnitude smaller for the case of 3D reconstructions accomplished with the Pinax algorithm. Therefore if without Pinax it yields a centimeter accuracy, with its application it can be reached a millimeter accuracy. In summary, the contributions that the use of Pinax algorithm brings about in 3D reconstructions are as follows:

- removal of refractive effects induced by underwater housing
- residue abatement from centimeter level to millimeter level
- better distribution of reconstructed points around their mean value
- lower dispersion of reconstructed points around their mean value and thus lower error on measurement

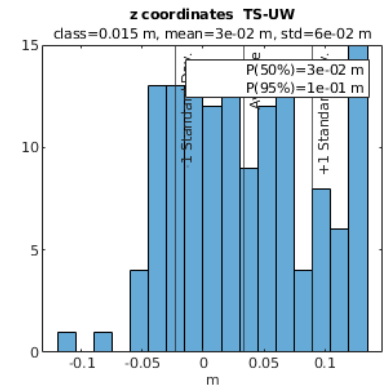
Further, the work done so far demonstrates that if you want to reach high level of accuracy only with the calibration methods that compensate for refractive effects into the distortion coefficients, thus it needs to use other geomatics tools, as for instance a local geodetic network. Indeed, ray tracing camera calibration method are more accurate and their coupling with



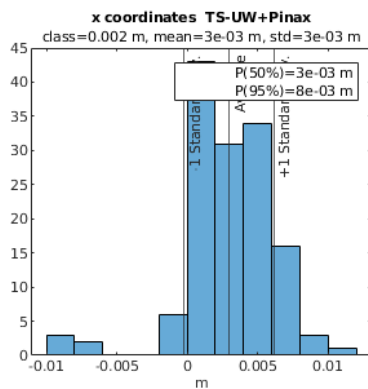
(a) Frequency histogram of the difference of the x coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment without using Pinax algorithm.



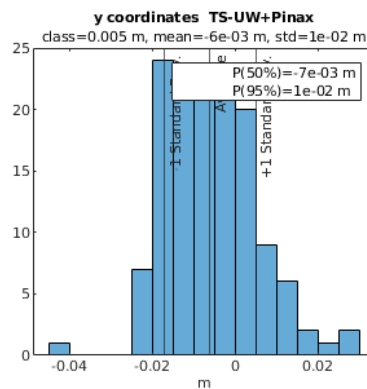
(b) Frequency histogram of the difference of the y coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment without using Pinax algorithm.



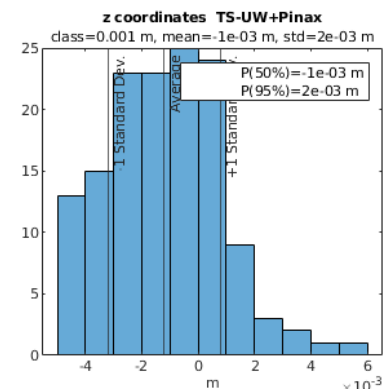
(c) Frequency histogram of the difference of the z coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment without using Pinax algorithm.



(d) Frequency histogram of the difference of the x coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment using Pinax algorithm.



(e) Frequency histogram of the difference of the y coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment using Pinax algorithm.



(f) Frequency histogram of the difference of the z coordinates extracted from the terrestrial survey and those obtained in the calibration frame experiment using Pinax algorithm.

Figure 6.22: Frequency histograms of the difference of the center target coordinates between those obtained by means of SLS and those estimated with underwater photogrammetric techniques. On the x-axis the value of the difference of the coordinates is represented in meters. On the y-axis is shown the frequency of the value associated with a given class. Also shown in the graphs are the mean and \pm standard deviation of the surveyed sample. In the subtitle of each graph it is possible to observe the magnitude of the classes, the value of the mean of the measurements, and the associated standard deviation. While in the legend it is possible to read the 50% and 95% percentile value.

standard methods, as you can see in this work, leads to better results. A comprehensive summary in which global characteristics of the experiment are provided is shown in Table [6.4](#)

Monte Carlo method applied to results

The previous results are represented as spatial coordinates differences, such as

$$r = X - \hat{X}$$

where X is the exact value associated to a general spatial coordinate, \hat{X} is the measured value associated to a general spatial coordinate, and r is the coordinate difference or residual. The coordinates differences are the means through which we can assume that the computations achieved so far are reliable and promising with the contributions listed in Section [6.4](#). The limit of these results is the fact that it was realised only a single test, since organizing additional missions at sea would have wasted many resources. Clearly the underwater photogrammetric measurements are subjected to measurement boundary conditions, such for instance: environmental parameters, instrumental conditions and survey settings. The coordinates coming from TS are used as ground truth (X), therefore they are associated to a null uncertainty ($\sigma_X = 0$), and following the Propagation of uncertainty law, considering the variables as independent, $\sigma = \sqrt{\sum_{i=1}^n (\frac{\partial f}{\partial x_i} \sigma_i)^2}$, we obtain

$$\sigma_r^2 = \sigma_X^2 + \sigma_{\hat{X}}^2 - 2\sigma_{X\hat{X}} = \sigma_{\hat{X}}^2, \quad \sigma_r = \sigma_{\hat{X}}$$

To overcome the limit of a single test, which is valid only for the underwater experience previously described, Monte Carlo method is applied in order to introduce a certain level of randomness. The Monte Carlo method requires as input the standard deviations associated to each spatial coordinate. These choices are made relying on the results achieved in Section [6.4](#) and found in literature ([\[82\]](#)-[\[22\]](#)-[\[83\]](#)-[\[20\]](#)). Therefore, standard deviations were used both for the case of Pinax-corrected and non-Pinax-corrected images, representing average values of the standard deviations obtained in the calibration frame experiment and those instead reported in

case	σ_x	σ_y	σ_z
UW	0.040	0.200	0.040
UW+Pinax	0.005	0.010	0.005

Table 6.3: Standard deviations values used as input of the Monte Carlo method expressed in meters. They are obtained as mean values between those given by the underwater experiment and those found in literature.

the scientific articles cited above. It should be mentioned that the measurements made on the calibration frame can be considered as similar to those reported on an underwater local geodetic network. The standard deviation values selected are thus reported in Table 6.3. The main steps of the application of the Monte Carlo method are the following, for detailed description of the method please refer to [84]:

1. We considered the M (139) scattered coordinate differences on the calibration frame, each with a corresponding mean value and standard deviation $(\mu_x, \mu_y, \mu_z, \sigma_x, \sigma_y, \sigma_z)$
2. These values are assumed to follow a normal distribution of coordinate differences
3. The chosen probability density functions are produced for each target according to the chosen mean and standard deviation values
4. N random simulations (random sampling) of potential coordinate differences within the probability density curves for each target are generated
5. It is verified that the number of simulations conducted is appropriate based on the convergence coefficient of the values obtained (after a certain number of simulations this coefficient must be 1)
6. Interpretation of results

Analysis of the coefficients of variation shows that the chosen number of simulations equals to 1000 is suitable to represent the distributions of values since over 800 the coefficients of variation take value equals to one. In Figure 6.23 are shown trends of convergence coefficients with respect to the number of simulations for some specific cases. Thanks to this result is possible to assume that when the curve becomes flat thus the number of simulation is suitable and the computations

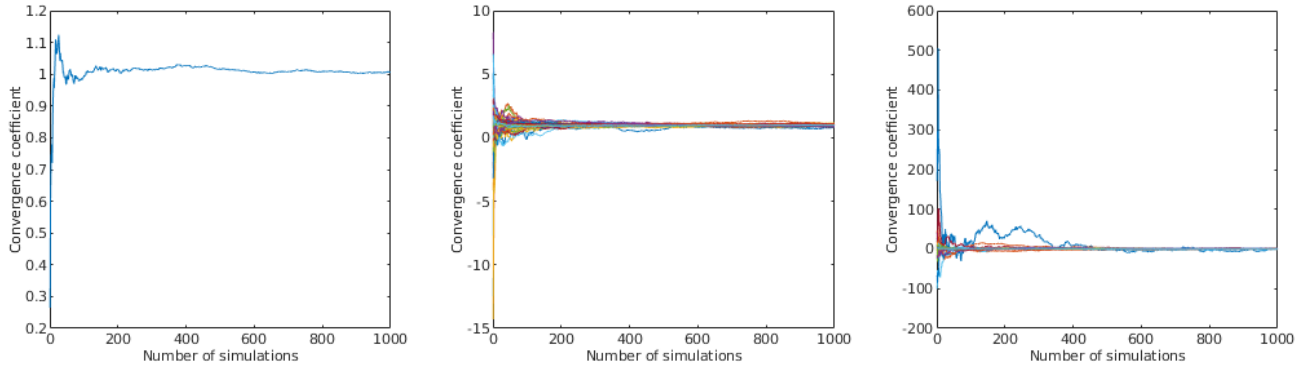


Figure 6.23: Computation of convergence coefficients as the number of simulations increases. (a) is related to a single case extracted randomly within the thousand distributions. (b) is related to all cases of y coordinate without using Pinax algorithm. (c) is related to all cases of z coordinate using Pinax algorithm.

	μ (m)			σ (m)			max (m)			min (m)		
	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
UW	0.020	0.200	0.030	0.070	0.200	0.060	0.105	0.469	0.135	-0.129	-0.420	-0.114
UW+P	0.003	0.006	0.001	0.003	0.010	0.002	0.012	0.029	0.005	-0.009	-0.041	-0.005

Table 6.4: Properties of the underwater calibration frame experiment. It is reported the mean value (μ) for each spatial coordinates, as well as the standard deviation (σ), the maximum value and minimum one. UW stands for underwater. P stands for Pinax.

yield to optimal results. For the sake of simplicity, the others convergence coefficient cases are omitted. Since only a single test was carried out, we can assume that the experiment set-up is reliable with the properties listed in Table 6.4. Thanks to Monte Carlo method it was possible to simulate thousand of potential values of coordinate differences associated to each target, therefore it can be analyzed thousand cases that could include any changes in the boundary conditions of the measurements. In a nutshell, Monte Carlo method allows to add a perturbation in the measurements with a gaussian noise. Therefore, it was possible to define the properties of the synthetic experiment conducted with Monte Carlo method in Table 6.5. By looking to these properties, first of all it notices an higher variation of values, but the same considerations made for the specific case of the calibration frame experiment can also be extended to the case of the synthetic experiment conducted by the Monte Carlo method. In this way, it provides a more robust and more reliable approach to confirm that using Pinax algorithm for correct the refractive effects given by water and underwater housing represent a optimal procedure if you want to reach high level of accuracy without the deployment of an underwater local geodetic

	μ (m)			σ (m)			max (m)			min (m)		
	X	Y	Z	X	Y	Z	X	Y	Z	X	Y	Z
UW	-0.019	0.197	0.034	0.079	0.264	0.069	0.162	0.837	0.203	-0.193	-0.601	-0.136
UW+P	0.003	-0.006	-0.001	0.006	0.011	0.005	0.018	0.029	0.013	-0.014	-0.041	-0.015

Table 6.5: Properties of the Monte Carlo simulations. It is reported the average of the mean values (μ) for each spatial coordinates, as well as the average of the standard deviations (σ), the average of maximum values and minimum ones. UW stands for underwater. P stands for Pinax.

network.

6.5 Several applications of color correction algorithms

As previously mentioned, the calibration frame is equipped with a color palette with which an investigation of color correction can be initiated. Unfortunately, an in-depth investigation could not be conducted here since it was necessary to have a colorimeter to be able to measure the amount of color absorbed by an object or substance. For this reason, color correction algorithm have been used that are not based on any physical theory (see Section 3.4) but only on the characteristics of the individual image. They are called image enhancement methods and use qualitative subjective criteria to produce a more visually pleasing image without considering any physical models. An excellent tool for applying them is the application developed and optimized by Mangeruga M. et al (2019) in [35] within the imareculture european project, who introduced within a single application four different image enhancement methods. The algorithm that were applied are:

- ACE Automatic Color Enhancement: it is a quite complex technique, due to its direct computation on an $N \times N$ image costs $O(N^4)$. It is an unsupervised color equalization algorithm developed by Rizzi et al. [34], it deals with a perceptual approach inspired by some adaptation mechanisms of the human vision system, in particular lightness constancy and color constancy
- SP Screened Poisson Equation for Image Contrast Enhancement: the output of the algorithm is an image which is the result of applying the Screened Poisson equation to

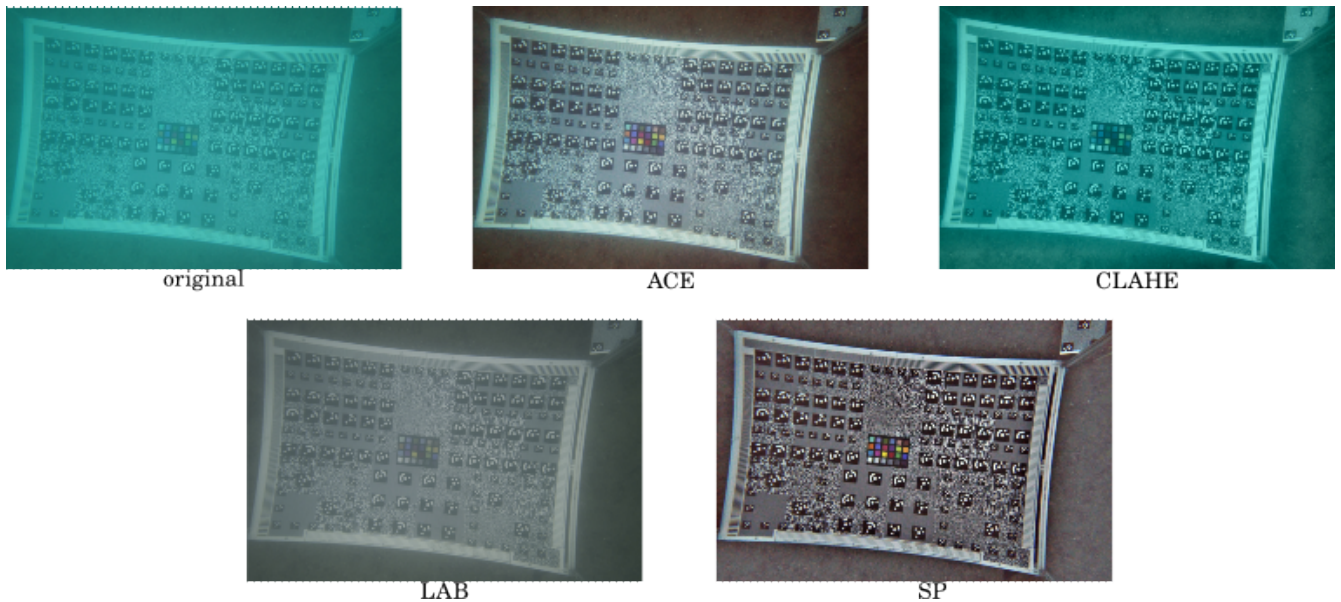


Figure 6.24: Application of ACE, CLAHE, LAB, and SP color correction algorithm obtained with standard parameters.

each colour channel separately, together with a simplest colour balance with a variable percentage of saturation

- LAB Color correction method on $\alpha\beta$ Space: it is based on the assumptions of grey world and uniform illumination of the scene. The idea behind this method is to convert the input image from RGB to LAB space, correct colour casts of an image by adjusting the α and β components
- CLAHE Contrast Limited Adaptive Histogram Equalization: it was originally developed for medical imaging and has proven to be successful for enhancing low-contrast images

In Figure [6.24](#) a single image is taken as example to analyze the application of the four different enhancement methods. The first image is the original one, it is taken by Blucy during its underwater experience. The other ones are the same image but treated with the previous color correction methods. Qualitatively, the algorithm that works worse is the LAB one, indeed, no gain is found even at contrast level. The image appears the same of the original, except for the characteristic grey color of the algorithm. The algorithm that works better is the ACE one. The second image appears as if the water were not there and even color palette and contrast are improved with respect to the original image. Unfortunately, it was not possible to apply

this algorithm to all 76 photos useful for providing 3D reconstruction. Because the photos were taken from different heights, the camera focus was different. Apparently this algorithm is very dependent on the camera focus, and therefore only some photos were improved, namely those that had excellent focus. This drastically reduced the number of photos useful for 3D reconstruction of the calibration frame. As far as CLAHE and SP algorithms, the former improved the contrast but it does not allow a removal of green color, while the latter behaves like the former but the green color is transformed in grey color. Therefore they represent intermediate cases in the qualitative scale. It must be made clear that the statements just made relate only to the realized underwater experience, and are not to be extended to the general case. According with results obtained in [35], no appreciable metric improvements can be appreciated when the images, which are used to create the 3D reconstructions, are pre-treated with this type of algorithms. We can only appreciate an aesthetic quality improvement.

Chapter 7

Conclusion

This thesis presented an accurate scaling and uncertainty estimation of 3D models built with monocular camera system in GPS-denied environments, such as the marine ones. The camera system selected is the one installed on the UUV called Blucy, which is the main subject of this thesis. Since it is characterized by a flat port, Pinax algorithm was exploited in order to take into account the refractive effects induced by the camera system. In addition to the analysis of the error estimates, even the practical issues that arise when an underwater survey with an UUV has to be organised and planned are shown, as well as the needs required when carrying out such measurements. The spirit in which the study was carried out was to build ad-hoc solutions to quantify the errors in point estimates, while also considering the economic aspect, thus finding a trade-off between costs and expected performances. In fact, the same approach has been used in the development of Blucy, a low-cost UUV prototype with the performance of a high-precision commercial product. Within an open modular "low-cost" UUV prototype, two practical underwater tests, properly realised, have been carried out exploiting simple tools found on the common market. The first test served as a training, in order to carry out a more robust test, performed in open sea, however it provided preliminary results that agreed with those obtained later with the second test. The second test represented an opportunity to apply the lessons learned in the first one, as it allowed to perform a more comprehensive error estimated analysis, as well as showcasing all the operational difficulties that

can be encountered when working with a prototype, such as Blucy. The results obtained from these two experiments showed that, for sure the refractive effects due to water and underwater housing can be compensated by the distortion coefficients in the camera model, but if the aim is to achieve high accuracy then a model that takes into account the configuration of the underwater housing, based on ray tracing, must also be coupled. These results were also supported by the use of a statistical method, such as Monte Carlo, which is able to induce a Gaussian perturbation in the measurements and thus strengthen their validity. Further, the phase preceding the implementation of the test, i.e. the terrestrial survey of the calibration platforms, is also described. Two expeditious surveying techniques are thus compared, because for a rigorous camera calibration there is the need to carry out each time the survey of the calibration platforms. The major contributions that this work brought can be summarized as follows:

- An overview of the practical issues when performing surveys exploiting an UUV prototype
- A method to reach a reliable accuracy in the 3D reconstructions without the use of an underwater local geodetic network
- A guide for who addresses underwater photogrammetry topics for the first time
- The use of open-source environments such as OpenCV, Octave and the Pinax algorithm implementation

Among the alternatives available for what concerns the camera calibration algorithms to be applied to underwater photogrammetric data, the selected one is resulted as an efficient solution. For sure, taking into account the continuous evolution on this field, further comparisons in the future may be based on other algorithms, or even take into account further types of corrections. Undoubtedly, many improvements can be made to the technique shown so far, but a necessary approach, which would increase the accuracy of the measurement technique, is to include a laser spot among the instruments available on Blucy so that the model can be scaled up immediately based on the information obtained from the laser itself. Indeed, with the time of flight computed for the laser spot, an estimate of the depth would be provided.

Bibliography

- [1] P. Rossi, M. Ponti, S. Righi, C. Castagnetti, R. Simonini, F. Mancini, P. Agrafiotis, L. Bassani, F. Bruno, C. Cerrano, *et al.*, “Needs and gaps in optical underwater technologies and methods for the investigation of marine animal forest 3d-structural complexity,” *Frontiers in Marine Science*, vol. 8, p. 591292, 2021.
- [2] T. Łuczyński, M. Pfingsthorn, and A. Birk, “The pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings,” *Ocean Engineering*, vol. 133, pp. 9–22, 2017.
- [3] A. Lambertini, M. Menghini, J. Cimini, A. Odetti, G. Bruzzone, M. Bibuli, E. Mandanici, L. Vittuari, P. Castaldi, M. Caccia, and L. De Marchi, “Underwater drone architecture for marine digital twin: Lessons learned from sushi drop project,” *Sensors*, vol. 22, no. 3, 2022.
- [4] J. L. Schonberger and J.-M. Frahm, “Structure-from-motion revisited,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4104–4113, 2016.
- [5] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [6] R. Szeliski, *Computer vision: algorithms and applications*. Springer Nature, 2022.
- [7] “opencv.org.”
- [8] E. Nocerino, F. Menna, A. Gruen, M. Troyer, A. Capra, C. Castagnetti, P. Rossi, A. J. Brooks, R. J. Schmitt, and S. J. Holbrook, “Coral reef monitoring by scuba divers using

- underwater photogrammetry and geodetic surveying,” *Remote Sensing*, vol. 12, no. 18, p. 3036, 2020.
- [9] T. Scambos, R. Bell, R. Alley, S. Anandakrishnan, D. Bromwich, K. Brunt, K. Christianson, T. Creyts, S. Das, R. DeConto, P. Dutrieux, H. Fricker, D. Holland, J. MacGregor, B. Medley, J. Nicolas, D. Pollard, M. Siegfried, A. Smith, E. Steig, L. Trusel, D. Vaughan, and P. Yager, “How much, how fast?: A science review and outlook for research on the instability of antarctica’s thwaites glacier in the 21st century,” *Global and Planetary Change*, vol. 153, pp. 16–34, 2017.
- [10] R. Campos, R. Garcia, P. Alliez, and M. Yvinec, “A surface reconstruction method for in-detail underwater 3d optical mapping,” *The International Journal of Robotics Research*, vol. 34, no. 1, pp. 64–89, 2015.
- [11] P. Martínez-Carricondo, F. Agüera-Vega, F. Carvajal-Ramírez, F.-J. Mesas-Carrascosa, A. García-Ferrer, and F.-J. Pérez-Porras, “Assessment of uav-photogrammetric mapping accuracy based on variation of ground control points,” *International journal of applied earth observation and geoinformation*, vol. 72, pp. 1–10, 2018.
- [12] A. R. Yusoff, M. M. Ariff, K. M. Idris, Z. Majid, and A. K. Chong, “Camera calibration accuracy at different uav flying heights,” *The international archives of photogrammetry, remote sensing and spatial information sciences*, vol. 42, p. 595, 2017.
- [13] J. G. Fryer and C. Fraser, “On the calibration of underwater cameras,” *The Photogrammetric Record*, vol. 12, no. 67, pp. 73–85, 1986.
- [14] J.-M. Lavest, G. Rives, and J.-T. Lapresté, “Underwater camera calibration,” in *European Conference on Computer Vision*, pp. 654–668, Springer, 2000.
- [15] “National snow & ice data center.” <https://nsidc.org/home>. Accessed: 2022-09-25.
- [16] A. Alcocer, P. Oliveira, and A. Pascoal, “Underwater acoustic positioning systems based on buoys with gps,” in *Proceedings of the Eighth European Conference on Underwater Acoustics*, vol. 8, pp. 1–8, 2006.

- [17] A. Capra, C. Castagnetti, M. Dubbini, A. Gruen, T. Guo, F. Mancini, F. Neyer, P. Rossi, and M. Troyer, “High accuracy underwater photogrammetric surveying,” in *3rd IMEKO International Conference on Metrology for Archeology and Cultural Heritage*, 2017.
- [18] A. Calantropio, F. Chiabrando, and R. Auriemma, “Photogrammetric underwater and uas surveys of archaeological sites: The case study of the roman shipwreck of torre santa sabina,” *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 643–650, 2021.
- [19] W.-C. Dullo, “Coral growth and reef growth: a brief review,” *Facies*, vol. 51, no. 1, pp. 33–48, 2005.
- [20] F. Neyer, E. Nocerino, and A. Grün, “Monitoring coral growth—the dichotomy between underwater photogrammetry and geodetic control network,” *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, 2018.
- [21] F. Menna, E. Nocerino, S. Ural, and A. Gruen, “Mitigating image residuals systematic patterns in underwater photogrammetry,” *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 977–984, 2020.
- [22] M. Shortis, “Camera calibration techniques for accurate measurement underwater,” *3D recording and interpretation for maritime archaeology*, pp. 11–27, 2019.
- [23] S. Granshaw, “Bundle adjustment methods in engineering photogrammetry,” *The Photogrammetric Record*, vol. 10, no. 56, pp. 181–207, 1980.
- [24] T. Treibitz, Y. Schechner, C. Kunz, and H. Singh, “Flat refractive geometry,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 1, pp. 51–65, 2011.
- [25] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari, “A theory of multi-layer flat refractive geometry,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3346–3353, IEEE, 2012.

- [26] R. Schettini and S. Corchs, "Underwater image processing: state of the art of restoration and image enhancement methods," *EURASIP journal on advances in signal processing*, vol. 2010, pp. 1–14, 2010.
- [27] B. McGlamery, "A computer model for underwater camera systems," in *Ocean Optics VI*, vol. 208, pp. 221–231, SPIE, 1980.
- [28] J. S. Jaffe, "Computer modeling and the design of optimal underwater imaging systems," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 2, pp. 101–111, 1990.
- [29] N. G. Jerlov, "Optical studies of ocean water," *Rept. Swedish Deep-Sea Exped.*, vol. 3, pp. 1–59, 1951.
- [30] P. J. Werdell, S. Bailey, G. Fargion, C. Pietras, K. Knobelspiesse, G. Feldman, and C. McClain, "Unique data repository facilitates ocean color satellite validation," *Eos, Transactions American Geophysical Union*, vol. 84, no. 38, pp. 377–387, 2003.
- [31] D. Akkaynak, T. Treibitz, T. Shlesinger, Y. Loya, R. Tamir, and D. Iluz, "What is the space of attenuation coefficients in underwater computer vision?," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 568–577, 2017.
- [32] M. Vlachos, A. Calantropio, D. Skarlatos, and F. Chiabrandio, "Modelling colour absorption of underwater images using sfm-mvs generated depth maps," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 959–966, 2022.
- [33] Y. Bekerman, S. Avidan, and T. Treibitz, "Unveiling optical properties in underwater images," in *2020 IEEE International Conference on Computational Photography (ICCP)*, pp. 1–12, IEEE, 2020.
- [34] A. Rizzi, C. Gatta, and D. Marini, "A new algorithm for unsupervised global and local color correction," *Pattern Recognition Letters*, vol. 24, no. 11, pp. 1663–1677, 2003.

- [35] M. Mangeruga, F. Bruno, M. Cozza, P. Agrafiotis, and D. Skarlatos, “Guidelines for underwater image enhancement based on benchmarking of different methods,” *Remote sensing*, vol. 10, no. 10, p. 1652, 2018.
- [36] “Gebco.” https://www.gebco.net/news_and_media/gebco_2022_grid_release.html. Accessed: 2022-07-18.
- [37] “The ocean conference.” <https://www.un.org/sustainabledevelopment/wp-content/uploads/2017/05/Ocean-fact-sheet-package.pdf>. Accessed: 2022-07-18.
- [38] R. Ferretti, M. Bibuli, G. Bruzzone, M. Caccia, A. Odetti, E. Cimenti, M. Demarte, R. Ivaldi, M. Marro, R. Nardini, *et al.*, “Critical marine environment observation: Measurement problems, technological solutions and procedural methods,” *Proceedings of the IMEKO TC-19 Metrology for the Sea, Naples, Italy*, pp. 5–7, 2020.
- [39] T. Liblik, J. Karstensen, P. Testor, P. Alenius, D. Hayes, S. Ruiz, K. Heywood, S. Pouliquen, L. Mortier, and E. Mauri, “Potential for an underwater glider component as part of the global ocean observing system,” *Methods in Oceanography*, vol. 17, pp. 50–82, 2016.
- [40] H. Jung, C. V. Subban, J. D. McTigue, J. J. Martinez, A. E. Copping, J. Osorio, J. Liu, and Z. D. Deng, “Extracting energy from ocean thermal and salinity gradients to power unmanned underwater vehicles: State of the art, current limitations, and future outlook,” *Renewable and Sustainable Energy Reviews*, vol. 160, p. 112283, 2022.
- [41] J. Yuh, G. Marani, and D. R. Blidberg, “Applications of marine robotic vehicles,” *Intelligent service robotics*, vol. 4, no. 4, pp. 221–231, 2011.
- [42] S. Krile, D. Kezić, and F. Dimc, “Nmea communication standard for shipboard data architecture,” *NAŠE MORE: znanstveni časopis za more i pomorstvo*, vol. 60, no. 3-4, pp. 68–81, 2013.
- [43] X. Zhao, Q. Luo, and B. Han, “Survey on robot multi-sensor information fusion technology,” in *2008 7th World Congress on Intelligent Control and Automation*, pp. 5019–5023, IEEE, 2008.

- [44] K. Yuan, H. Wang, and H. Zhang, “Robot position realization based on multi-sensor information fusion algorithm,” in *2011 Fourth International Symposium on Computational Intelligence and Design*, vol. 1, pp. 294–297, 2011.
- [45] K. Yuan, H. Wang, and H. Zhang, “Robot position realization based on multi-sensor information fusion algorithm,” in *2011 Fourth International Symposium on Computational Intelligence and Design*, vol. 1, pp. 294–297, IEEE, 2011.
- [46] F. B. Berger, “The nature of doppler velocity measurement,” *IRE transactions on aeronautical and navigational electronics*, no. 3, pp. 103–112, 1957.
- [47] D. Ribas, P. Ridao, X. Cufí, and A. El-fakdi, “Towards a dvl-based navigation system for an underwater robot,” in *4th Workshop on European Scientific and Industrial Collaboration*, Citeseer, 2003.
- [48] C. J. Vietheer, “Comparison of mems and fog-based inertial navigation systems for hydrographic applications,” 2019.
- [49] A. G. Quinchia, G. Falco, E. Falletti, F. DAVIS, and C. Ferrer, “A comparison between different error modeling of mems applied to gps/ins integrated systems,” *Sensors*, vol. 13, no. 8, pp. 9549–9588, 2013.
- [50] R. M. Eustice, O. Pizarro, and H. Singh, “Visually augmented navigation for autonomous underwater vehicles,” *IEEE Journal of oceanic Engineering*, vol. 33, no. 2, pp. 103–122, 2008.
- [51] M. F. Fallon, M. Kaess, H. Johannsson, and J. J. Leonard, “Efficient auv navigation fusing acoustic ranging and side-scan sonar,” in *2011 IEEE International Conference on Robotics and Automation*, pp. 2398–2405, IEEE, 2011.
- [52] S. Barkby, S. Williams, O. Pizarro, and M. Jakuba, “An efficient approach to bathymetric slam,” in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 219–224, IEEE, 2009.

- [53] Z. Yan, S. Peng, J. Zhou, J. Xu, and H. Jia, "Research on an improved dead reckoning for auv navigation," in *2010 Chinese Control and Decision Conference*, pp. 1793–1797, IEEE, 2010.
- [54] J. Kunhoth, A. Karkar, S. Al-Maadeed, and A. Al-Ali, "Indoor positioning and wayfinding systems: a survey," *Human-centric Computing and Information Sciences*, vol. 10, no. 1, pp. 1–41, 2020.
- [55] Y. Zhang and L. Cheng, "A distributed protocol for multi-hop underwater robot positioning," in *2004 IEEE International Conference on Robotics and Biomimetics*, pp. 480–484, IEEE, 2004.
- [56] P. Batista, C. Silvestre, and P. Oliveira, "A sensor-based long baseline position and velocity navigation filter for underwater vehicles," *IFAC Proceedings Volumes*, vol. 43, no. 14, pp. 302–307, 2010.
- [57] Y. Wu, X. Zhou, L. Yang, *et al.*, "Underwater acoustic positioning system and its application," *Hydrographic Surveying and Charting*, vol. 23, no. 4, pp. 18–21, 2003.
- [58] G. Pilato, A. Augello, M. Missikoff, and F. Taglino, "Integration of ontologies and bayesian networks for maritime situation awareness," in *2012 IEEE Sixth International Conference on Semantic Computing*, pp. 170–177, IEEE, 2012.
- [59] M. T. Sabet, P. Sarhadi, and M. Zarini, "Extended and unscented kalman filters for parameter estimation of an autonomous underwater vehicle," *Ocean Engineering*, vol. 91, pp. 329–339, 2014.
- [60] H. Yao, H. Wang, and Y. Wang, "Uuv autonomous decision-making method based on dynamic influence diagram," *Complexity*, vol. 2020, 2020.
- [61] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part i," *IEEE robotics & automation magazine*, vol. 13, no. 2, pp. 99–110, 2006.

- [62] R. Dubé, A. Gawel, H. Sommer, J. Nieto, R. Siegwart, and C. Cadena, “An online multi-robot slam system for 3d lidars,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1004–1011, IEEE, 2017.
- [63] S. Jia, K. Wang, and X. Li, “Mobile robot simultaneous localization and mapping based on a monocular camera,” *Journal of Robotics*, vol. 2016, 2016.
- [64] P. Drews Jr, S. Botelho, and S. Gomes, “Slam in underwater environment using sift and topologic maps,” in *2008 IEEE Latin American Robotic Symposium*, pp. 91–96, 2008.
- [65] D. Wu, Z. Yan, and T. Chen, “Cooperative current estimation based multi-auvs localization for deep ocean applications,” *Ocean Engineering*, vol. 188, p. 106148, 2019.
- [66] L. Paull, M. Seto, and J. J. Leonard, “Decentralized cooperative trajectory estimation for autonomous underwater vehicles,” in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 184–191, IEEE, 2014.
- [67] “Unmanned underwater vehicles (uuv) market by type (remotely operated vehicle & autonomous underwater vehicle), rovs & auv market by application, product, propulsion system, system, and region - global forecasts to 2025.” [https://www.researchandmarkets.com/reports/5235828/unmanned-underwater-vehicles-uuv-market-by-type?utm_source=GNOM&utm_medium=PressRelease&utm_code=lnplc5&utm_campaign=1488784+-+The+Worldwide+Unmanned+Underwater+Vehicles+\(UUV\)+Industry+is+Expected+to+Reach+%244.4+Billion+by+2025&utm_exec=jamu273prd](https://www.researchandmarkets.com/reports/5235828/unmanned-underwater-vehicles-uuv-market-by-type?utm_source=GNOM&utm_medium=PressRelease&utm_code=lnplc5&utm_campaign=1488784+-+The+Worldwide+Unmanned+Underwater+Vehicles+(UUV)+Industry+is+Expected+to+Reach+%244.4+Billion+by+2025&utm_exec=jamu273prd). Accessed: 2022-07-29.
- [68] T. P. Bean, N. Greenwood, R. Beckett, L. Biermann, J. P. Bignell, J. L. Brant, G. H. Copp, M. J. Devlin, S. Dye, S. W. Feist, *et al.*, “A review of the tools used for marine monitoring in the uk: combining historic and contemporary methods with modeling and socioeconomics to fulfill legislative needs and scientific ambitions,” *Frontiers in Marine Science*, vol. 4, p. 263, 2017.
- [69] G. M. Gavazzi, F. Madricardo, L. Janowski, A. Kruss, P. Blondel, M. Sigovini, and F. Foglini, “Evaluation of seabed mapping methods for fine-scale classification of extremely

- shallow benthic habitats—application to the venice lagoon, italy,” *Estuarine, Coastal and Shelf Science*, vol. 170, pp. 45–60, 2016.
- [70] H. Mohamed, K. Nadaoka, and T. Nakamura, “Assessment of machine learning algorithms for automatic benthic cover monitoring and mapping using towed underwater video camera and high-resolution satellite images,” *Remote Sensing*, vol. 10, no. 5, 2018.
- [71] “Sushidrop link.” <https://www.italy-croatia.eu/web/sushidrop>. Accessed: 2022-07-18.
- [72] A. Lambertini, M. Menghini, J. Cimini, A. Odetti, G. Bruzzone, M. Bibuli, E. Mandanici, L. Vittuari, P. Castaldi, M. Caccia, *et al.*, “Monitoring and surveying from an underwater vehicle in sushi drop project,” in *2021 International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters (MetroSea)*, pp. 189–193, IEEE, 2021.
- [73] V. Raoult, S. Reid-Anderson, A. Ferri, and J. E. Williamson, “How reliable is structure from motion (sfm) over time and between observers? a case study using coral reef bommies,” *Remote Sensing*, vol. 9, no. 7, p. 740, 2017.
- [74] T. I. Fossen and T. Perez, “Kalman filtering for positioning and heading control of ships and offshore rigs,” *IEEE Control Systems Magazine*, vol. 29, no. 6, pp. 32–46, 2009.
- [75] U. Hunkeler, H. L. Truong, and A. Stanford-Clark, “Mqtt-s—a publish/subscribe protocol for wireless sensor networks,” in *2008 3rd International Conference on Communication Systems Software and Middleware and Workshops (COMSWARE’08)*, pp. 791–798, IEEE, 2008.
- [76] J. J. Koenderink and A. J. Van Doorn, “Affine structure from motion,” *JOSA A*, vol. 8, no. 2, pp. 377–385, 1991.
- [77] N. Zheng, G. Loizou, X. Jiang, X. Lan, and X. Li, “Computer vision and pattern recognition,” 2007.
- [78] M. D. Grossberg and S. K. Nayar, “The raxel imaging model and ray-based calibration,” *International Journal of Computer Vision*, vol. 61, no. 2, pp. 119–137, 2005.

- [79] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [80] W. Roswell, A. Halikas, and G. Halikas, “The index of refraction of seawater-technical report,” *University of California, San Diego Visibility Laboratory*, 1976.
- [81] L. E. Peterson, “K-nearest neighbor,” *Scholarpedia*, vol. 4, no. 2, p. 1883, 2009.
- [82] D. Ventura, G. Mancini, E. Casoli, D. S. Pace, G. J. Lasinio, A. Belluscio, and G. Ardizzone, “Seagrass restoration monitoring and shallow-water benthic habitat mapping through a photogrammetry-based protocol,” *Journal of Environmental Management*, vol. 304, p. 114262, 2022.
- [83] M. Ballarin, E. Costa, A. Piemonte, M. Piras, and L. T. Losè, “Underwater photogrammetry: Potentialities and problems results of the benchmark session of the 2019 sifet congress,” *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 925–931, 2020.
- [84] D. J. C. Mackay, “Introduction to monte carlo methods,” in *Learning in graphical models*, pp. 175–204, Springer, 1998.