Low-Cost Vision Based Autonomous Underwater Vehicle for Abyssal Ocean Ecosystem Research

by

Eduardo Alexander Iscar Ruland

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Naval Architecture and Marine Engineering)
in the University of Michigan
2020

Doctoral Committee:

Associate Professor Matthew Johnson-Roberson, Chair
Assistant Professor Alex Shorter
Associate Professor David Singer
Professor Jing Sun

Eduardo Alexander Iscar Ruland
eiscar@umich.edu
ORCID iD: 0000-0003-3052-6743

Für Opa

ACKNOWLEDGMENTS

When I moved to Ann Arbor, I was sure it would only be until I got my Master's degree. I was confident a PhD was not a good idea. Now, five years layer, I would like to thank my advisor, Professor Matthew Johnson-Roberson, for changing my mind on the subject and giving me the opportunity to spend these years working on the fascinating topic of autonomous underwater vehicles. I am very grateful for all the freedom I have had to explore the subject and the constant support and encouragement. I would also like to thank the rest of my committee members, Professors Sun, Singer, and Shorter, for sharing this journey with me.

I could not have finished this project, however, without all the the members of the DROP-Lab: Jie, Katie, Gideon, Nick, Corina, Laura, Liz and Tianyi. Sharing field trips, lunch, beers, and board game nights has been such an important part of my Grad School experience. I would also like to thank Corina and Nick for their special involvement with the DROP-Sphere project: your contributions have been a fundamental driver for what we have achieved. I have been extremely lucky to have such a great team to rely on and a great deal of my successes are due to all of you.

Furthermore, I would like to thank my family for their constant encouragement and patience every time I tried to explain what I was working on all this time. A special thanks to my partner Laura: your help, kindness, and love have pushed me through the hard times and made me enjoy the good ones. I will never forget it.

Last but not least, I would like to thank everyone who makes up the University of Michigan. It took the work of a lot of different people so that I could work on this project for my PhD. I am thankful for everyone who makes the University such a welcoming, enjoyable community.

This work was funded in part by the Ford Motor Company via the Ford-UM Alliance under Award N022884, by the National Oceanic and Atmospheric Administration under Award NA14OAR40110265 and by the National Science Foundation under Award 1452793.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

The oceans have a major impact on the planet: they store 28% of the CO_2 produced by humans, they act as the world's thermal damper for temperature changes, and more than $17,000$ species call the deep oceans their home. Scientific drivers, like climate change, and commercial applications, like deep sea fisheries and underwater mining, are pushing the need to know more about oceans at depths beyond 1000 meters. However, the high cost associated with autonomous underwater vehicles (AUVs) capable of operating beyond the depth of 1000 meters has limited the study of the deep ocean.

Traditional AUVs used for deep-sea navigation are large and typically weigh upwards of 1000-kgs, thus requiring careful planning before deployment and multi-person teams to operate. This thesis proposes the use of a new vehicle design based around a low-cost oceanographic glass sphere as the main pressure enclosure to reduce its size and cost while maintaining the ability for deep-sea operation. This novel housing concept, together with a minimal sensor suite, enables environmental research at depths previously inaccessible at this price point. The key characteristic that enables the cost reduction of this platform is the removal of the Doppler velocity log (DVL) sensor, which is replaced by optical cameras. Cameras allow the vehicle to estimate its motion in the water, but also enable scientific applications such as identification of habitat types or population density estimation of benthic species. After each survey, images can be further processed to produce full, dense 3D models of the survey area.

While underwater optical cameras are frequently placed inside pressure housings behind flat or domed viewports and used for visual navigation or 3D reconstructions, the underlying assumptions for those algorithms do not hold in the underwater domain. Refraction at the housing viewport, together with wavelength-dependent attenuation of light in water, render the ubiquitous pinhole camera model invalid. This thesis presents a quantitative evaluation of the errors introduced by underwater effects for 3D reconstruction applications, comparing low- and high-cost camera systems to quantify the trade-off between equipment cost and performance.

Although the distortion effects created by underwater refraction of light have been extensively studied for more traditional viewports, the novel design proposed necessitates new research into modeling the lensing effect of this off-axis domed viewport. A novel calibration method is presented that explicitly models the effect of the glass interface on image formation based on the characterization of optical distortions. The method is capable of accurately finding the position of the camera within the dome and further enables the use of deconvolution to improve the quality of the taken image.

Finally, this thesis presents the validation of the designed vehicle for optical surveying tasks and introduces a end-to-end ocean mapping pipeline to streamline AUV deployments, enabling efficient use of time and resources.

CHAPTER I

Introduction

1.1 Motivation

Technological advances in the last century have opened the oceans to scientific, industrial, and recreational use. Although the use of the seas for trade, resources, and war goes back thousands of years, it was not until the Challenger Expedition at the end of the 19th century that oceanography emerged as a branch of science. Since then, technological advances have taken many forms to facilitate the scientific work on ocean research: new ship hull materials and designs allowed for extended periods at sea, sonar replaced manual, range-limited bathymetry soundings, and satellite technology allowed the covering of mesoscale ocean processes. Despite technological advances, much is still unknown about the oceans. Only 0.05% of the oceans are mapped at meter resolution [1], while the highest-resolution map covering the ocean's entirety is 5km [2].

In the past, AUVs have appeared as a new tool to allow the collection of detailed oceanographic data. As opposed to remotely operated vehicles (ROVs), AUVs are not tethered to the deployment vessel and do not require human intervention. This has made previously inaccessible areas available and allows researchers to collect oceanographic data at unprecedented scales. Optical cameras are seeing increasing deployment for benthic applications due to their low-cost and low-power characteristics. The 3D photogrammetry tools combine individual images into three dimensional models of the seafloor that show the geometry of the seafloor and its texture. For example, Figure 1.1b shows the reconstruction of a coral reef off of the coast of Hawaii.

Most AUVs have been successfully deployed in shallow coastal waters (depths up to 200m). Some example applications include optical monitoring of coral reefs [3], underwater archaeology [4], and mapping underwater regions [5]. Other projects have used AUVs to control invasive species [6], count animal populations [7], or analyze shark behaviour [8]. Many custom vehicles have been developed for each of these specific tasks, and commercial vehicles exist that can be adapted to mission specific requirements, such as the Kongsberg REMUS 100, Oceanservers IVER (Figure 1.1a), or the Bluefin AUV.

However, there are only a few deep-sea rated AUVs that can go beyond the mesopelagic zone. These vehicles are usually large, heavy pieces of equipment and require teams of multiple operators to deploy and recover. As a result, operation and acquisition of millions of dollars restrict their widespread use. This is of special importance if we consider the resolution of AUV data is orders of magnitude larger than surface-vessel-based technologies. Additionally, AUV area coverage is much greater than the next closest solution by cost for high resolution imaging, a deep sea lander.

The gap in available tools for ocean research, when comparing shallow and deep sea research, is an important constraint on new discoveries regarding oceanic benthos and the target of this thesis.

(a) Shallow water OceanServer IVER3 (Image taken by G.Billings)

(b) 3D reconstruction of a coral reef in Kaneohe Bay, Hawaii

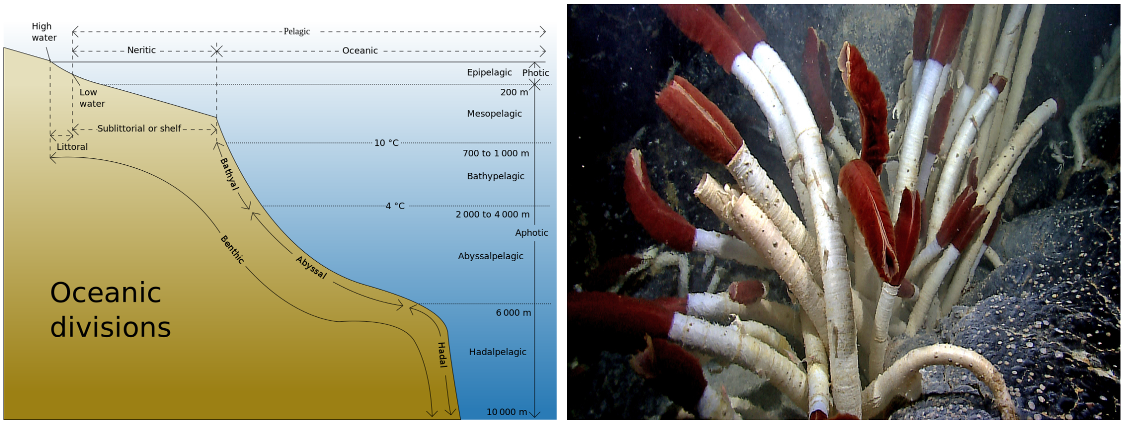
Figure 1.1: AUV and typical 3D reconstruction example

1.2 On the Importance of the Deep Oceans

One of the objectives of the United Nations Sustainable Development Goals is "*to conserve and suitably use the oceans, seas, and marine resources for sustainable development*". Although the general public's understanding of the oceans focuses on pristine beaches and coral reefs, epipelagic coastal waters only account for 5% of the ocean's volume. The rest, beyond 200m depth, is the dark deep sea. Multiple distinct habitats exist in the deep ocean. On a broad-scale, ocean regions can be categorized as soft-sediment planes, volcanic mountain ranges, and ocean trenches. On a smaller scale, hydrothermal vents, methane seeps, cold water coral habitats, and oxygen minimum zones represent some of the most significant areas [9]. These habitats are host to a very large number of marine species. It is estimated that there are approximately two million species [10]. However, the fact that only 9% of the species have been identified [11] highlights the lack of knowledge about our oceans. Life at these depths is highly adapted to its environment. With very little sunlight reaching 1000m, most of the ocean lies in total darkness, requiring deep sea ecosystems to harvest alternative sources of energy. Communities around hydrothermal vents or methane seeps rely on chemosynthetic bacteria to obtain their energy [12]. Others, like Osedax worms, depend on sunken bones from dead fish or mammals to survive [13].

The deep oceans also play a fundamental role in the carbon and nutrient cycles on earth. Cold water near the earth's poles sinks to the bottom of the ocean, dragging dissolved CO_2 with it. These water masses will remain isolated from the atmosphere for ~ 1000 years, creating a buffer that regulates climate and CO_2 exchange. With approximately 25% of the CO_2 emitted by human activity currently locked in the oceans [14], they play a crucial role in reducing the greenhouse effect. However, the consequences, such as ocean acidification, are already being experienced by coral reefs worldwide [15]. In addition to to CO_2 locking, thermohaline circulation is also responsible for cycling and regenerating nutrients that feed shallow water fauna. Some of the worlds most productive fisheries are located in areas of strong upwelling [16].

While we struggle to understand the complicated relationships and processes that govern deep sea ecosystems, studies show that no single ocean area is free of anthropogenic impacts [17]. These impacts can be organized into three main categories [9]: 1. *Disposal*, grouping elements such as sewage, radioactive waste, wrecks, munitions, or chemical contamination; 2. *Exploitation*, which includes fishing, mining for resources like oil, gas, and minerals, as well as underwater cable and pipeline laying; and 3. *Ocean Acidification and Climate Change*, made up of warming temperature, nutrient loading, and hypoxia or deep circulation shutdown. While impacts during the 20th century were mostly of disposal in nature, the current push for fossil fuels, metal mining, and deep sea fisheries presents the highest threat to the deep seas in the short term [9]. The wide range of anthropogenic impacts on the deep ocean and their geographical ubiquity emphasize the need for comprehensive ocean research. As Dr. Sylvia Earle said: "Far and away the biggest threat to the ocean is ignorance".



(a) Ocean zones (Figure by Chris Huh, Public Domain)

(b) Riftia pachyptila tube worms commonly found at hydrothermal vents

Figure 1.2: Ocean zones and deep sea worms

1.3 Problem Statement

In recent decades, optical cameras have been increasingly applied to ocean science. Still images and video data provide scientists with information that enables a large variety of research in deep water biology: the classification and analysis of benthic habitats, validation of ecological niche factor analyses [18], evaluations of anthropogenic impacts such as oil spills on coral reefs [19], benthic habitat classification [20], among other scientific applications. In the case of moving cameras, images can be further joined into photomosaics, or fully textured 3D models, that contextualize data and highlight processes happening at different scales. These examples show the growth of the importance of optical sampling methods for underwater biological research in the abyssal ocean zone. This thesis is focused on enabling abyssal deep sea imaging for biological research and addresses the engineering requirements for reliable, visual data-gathering of the ocean benthos. The main challenges and proposed solutions are discussed in the following sections.

1.3.1 Deep Sea Research Platforms

Optical cameras are mounted onto research equipment that includes mechanical housings, electrical connections, and pertinent software for recording images and parameters. With deployments frequently performed from costly research vessels far from shore, maximizing area coverage and ship-time utilization becomes a critical factor. Figure 1.3 shows a comparison of the different technologies available for performing deep ocean benthic research with respect to their cost and area coverage capabilities. Deploying deep-sea landers for imaging the deep sea is the method with the lowest cost currently available. However, their static position severely limits the area covered. On the other end of the cost spectrum are manned submersibles. Very high costs, complex operation, and reduced availability (less then 10 vehicles worldwide [21]) restrict their use for deep ocean benthic imaging. In between these two technologies, ROVs offer a compromise between cost and area coverage, mostly limited by the tether connecting them to the ship. AUVs significantly increase area coverage but are equipped with additional sensors to enable autonomy that increase their cost. In summary, technologies for deep sea imaging are large pieces of equipment with high associated costs and complex operational requirements. The lack of smaller, lower-cost alternatives reduces a researcher's ability to study benthic ecosystems and are a major obstacle towards understanding the oceans.

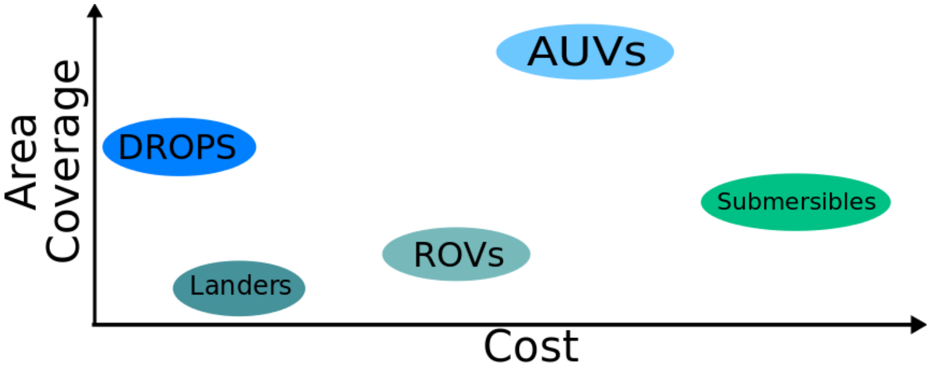


Figure 1.3: Abyssal pptical sensing platforms comparison: Qualitative comparison of cost and area coverage for different technologies capable of performing optical sampling at abyssal ocean depths (1000m to 6000m).

Firstly, this thesis introduces the DROP-Sphere (DROPS), a novel, ultra low-cost autonomous underwater vehicle for abyssal ocean benthic biological research. The vehicle cost of approximately $\$35,000$ is significantly less then AUVs, ROVs, and manned submersibles that can operate up to 6000m depth. Additionally, its area coverage, although smaller then that of AUVs is larger then ROVs, landers, and manned submersibles.

In order to achieve such a significant cost reduction, the designed vehicle replaced the DVL, an acoustic sensor used to measure vehicle speed, with optical cameras for pose estimation. Global acoustic localization methods are also eliminated from the standard vehicle payload. Although the lack of georeferencing prevents the use of the vehicle in applications that require accurate global position information, such as bathymetric mapping or underwater search missions, many biological tasks, such as habitat mapping, species counts, or impact analysis, can still be accomplished with coarse position estimates.

1.3.2 Visual Underwater Navigation

The decision to eliminate the DVL from the sensor suite on board the DROP-Sphere required the use of a reliable, alternative way of estimating vehicle pose underwater. Optical navigation methods have emerged as a complimentary navigation tool to traditional, DVL-based navigation and estimate vehicle movement from image sequences. Cameras are low-power and low-cost sensors. They provide not only data for navigation, but also valuable information about the seafloor they image. Additionally, when used in a simultaneous localization and mapping (SLAM) framework, camera images allow the identification of previously visited areas. These loop closures bound the drift in pose estimation and create consistent trajectories. Visual navigation techniques have been successfully applied underwater in multiple occasions [3], [22]–[24]. Mahon et al. [4] show full 3D reconstructions obtained with a diver-operated camera rig equipped exclusively with cameras, a depth sensor, and an inertial measurement unit (IMU). This proves the feasibility of estimating vehicle pose without requiring the use of a DVL.

However, applying computer vision methods underwater presents additional challenges when compared to terrestrial settings: 1. the lack of ambient light requires the use of artificial light sources, increasing power consumption and potentially creating lightening artifacts; 2. light is strongly attenuated underwater as a function of its wavelength, and this causes underwater images to look predominantly green or blue; 3. backscattering of light by particles in the water; and 4. refraction at the housing glass interface bends the light rays and breaks projective geometry assumptions used in many computer vision methods. These challenges are sources for systematic errors in underwater visual navigation and need to be addressed for vehicles relying on optical sensors as their main navigation instrument. The off-center position of the cameras inside the DROP-Sphere, a consequence of the spherical housing used to enclose the stereo pair, induces large refractive errors that have to be addressed to ensure accurate navigation and benthic reconstructions.

To summarize, underwater effects, such as refraction, render traditional perspective projection models for camera systems inaccurate in the presence of housings. This induces systematic errors in computer vision methods, such as visual odometry, that can compromise the autonomous vehicle navigation capabilities as well as the value of the generated data products.

The second goal was to model underwater image formation effects into a general camera model to improve navigation accuracy and reconstruction fidelity. In order to do so, the impact of underwater effects on the quality of 3D reconstructions with off-the-shelf terrestrial algorithms had to be determined and new calibration techniques developed for cameras positioned in housings with large offsets from the dome center.

1.4 Contributions

The following list enumerates the specific contributions, as well as the corresponding chapter that enable the achievement of this thesis's goals.

- Develop an open source, DVL-less, ultra low-cost deep sea capable robot, the DROP-Sphere, that reduces the barriers of entry to abyssal deep ocean exploration and research through low cost and small size. (Chapter 2)
- Quantify the impact of optical underwater effects on 3D reconstructions with conventional computer vision methods developed for in-air use. (Chapter 3)
- Develop a framework for calibrating off-axis cameras inside domed viewports through the use of the point spread function (PSF). (Chapter 4)
- Develop an end-to-end deep sea surveying and 3D reconstruction pipeline and validate the vehicles optical mapping capabilities. (Chapter 5)

Work presented in this proposal, as well as related research, has been published in the following publications:

Eduardo Iscar, Atulya Shree, Nicholas Goumas, and Matthew Johnson-Roberson. "Low cost underwater acoustic localization." In Proceedings of Meetings on Acoustics 173EAA, vol. 30, no. 1, p. 070006. ASA, 2017.

Eduardo Iscar, Katherine A. Skinner, and Matthew Johnson-Roberson. "Multi-view 3D reconstruction in underwater environments: Evaluation and benchmark." In OCEANS – Anchorage, 2017, pp. 1-8. IEEE, 2017. ©2017 IEEE. Reprinted, with permission, from [25].

Katherine A. Skinner, **Eduardo Iscar Ruland** and Matthew Johnson-Roberson, "Automatic color correction for 3D reconstruction of underwater scenes." In Proceedings of the IEEE International Conference on Robotics and Automation, Singapore, 2017.

Eduardo Iscar, Corina Barbalata, Nicholas Goumas and Matthew Johnson-Roberson, “Towards low cost, deep water AUV optical mapping.” IEEE/MTS OCEANS Conference and Exhibition, Charleston, USA, 2018. ©2018 IEEE. Reprinted, with permission, from [26].

Corina Barbalata, **Eduardo Iscar**, Matthew Johnson-Roberson, “Experimental evaluation of depth controllers for a small-size AUV” 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)

Eduardo Iscar, Matthew Johnson-Roberson, “Autonomous surface vehicle 3D seafloor reconstruction from monocular images and sonar data” IEEE/MTS OCEANS Conference and Exhibition, Washington, USAB, 2015.

Eduardo Iscar, Matthew Johnson-Roberson, “Towards distortion based underwater domed viewport camera calibration”. Under review. Submitted to IEEE Robotics and Automation Letters

CHAPTER II

Low Cost Deep Water Mapping AUV

2.1 Background

The manufacturing cost of electronic goods has decreased exponentially over previous decades. Simultaneously, the appearance of accessible software design tools allow custom solutions to be easily turned into functional parts at a low cost. In parallel, rapid prototyping technologies, such as 3D printing, enable quick iteration over prototypes and proofs of concept for mechanical components. The accessibility of these technologies has fostered the appearance of low-cost robotic platforms in many domains: open source prosthetics [27], aerial drones [28], ground vehicles [29], and manipulators [30]. However, very little of this innovation is making its way to deep sea AUV technology. Some of the reasons include a very aggressive working environment at high pressures, a small user base committed to very complex platforms, and high overhead costs that relativize the impact of vehicle cost reductions. This chapter presents an overview of the state of the art in deep sea AUV technologies as well as a novel, low-cost and open source marine robotic platform for deep ocean mapping.

2.1.1 Deep Sea AUVs

Operating at high depth in the ocean presents AUVs with many challenges, including withstanding the high pressure at abyssal depths, collision risks due to poorly mapped seafloor areas, and frequent far from shore deployments on the open sea that depend on meteorological conditions. As a result, commercial and research deep sea AUVs are large and heavy pieces of equipment, with their weight starting at 630kg, with a typical example being the *Bluefin* 21 AUV [31]. Additionally, commercial AUVs carry a large and varied suite of sensors that includes DVLs, ultra-short-baseline (USBL), or long-baseline (LBL) positioning systems; acoustic modems; inertial navigation systems (INSs); and sonar sensors such as multi-beam or side-scan. Finally, most vehicles offer a reconfigurable payload area where mission specific equipment, such as hydrophones, fluorometers, conductivity-temperature-depths (CTDs), or turbidity sensors, are installed. Due to their large size and complexity, these types of AUVs require the use of large oceanographic support vessels as well as teams of multiple operators [32]. Table 2.1 shows a list of some of the most representative deep sea research AUVs.

Table 2.1: List of deep sea AUVs: For an exhaustive list of sensors please refer to the vehicles data sheets.

Vehicle	**Size**	**Weight**	**Battery**	**Depth**	**Sensors**
Hugin 4500 [33]	⌀:1m L:6m	1900kg	60 kWh	4500m	DVL, INS, Multibeam & Forward Sonar, Iridim, UHF Radio, GPS
Remus 6000 [34]	⌀:0.71m L:3.84m	862kg	11 kWh	6000m	INS, ADCP, LBL, Acoustic modem, Iridium, Wi-Fi, Side-Scan, Camera, Multibeam, CTD, GPS
Bluefin 21 [31]	⌀:0.53m L:4.93m	750kg	13.5 kWh	4500m	INS, DVL, SVS, GPS, USBL, RF, Iridium, Acousitc Modem
Sentry [35]	L:2.9m W:2.2m H:1.8m	1250kg	18 kWh	6000m	DVL, USBL, Acoustic Modem, INS, CTD, Cameras, Multibeam and SideScan Sonar, ADCP, IMU
Nereus [36] [1]	L:3m W:2m H: 2m	2800kg	18 kWh	11000m	CTD, Magnetometer, Forward and Profiling Sonar, DVL

1. Lost at sea during a deployment in 2009.

2.1.2 Low-Cost Underwater Robotics

Multiple low-cost underwater autonomous vehicles have been developed over previous decades for deployment in shallow water depths of up to 100m. The Starbug AUV [37] is a shallow-water reef monitoring AUV with a cost less than $\$10,000$. It navigates by making use of stereo cameras and an IMU. Pirajuba [38] is an AUV designed as a platform for academic research in hydrodynamics and control architectures. Other AUVs are built for very specific applications; one example is Folaga [39], an AUV intended for the collection of vertical oceanographic data in coastal waters. Other low cost vehicles, such as the Stingray AUV [40], imitate the shape of ocean animals in an effort to reduce drag and improve the AUVs characteristics. However, fewer examples of completely open source designs that include hardware and software exist, with quality and completeness of the sources and documentation varying considerably between projects. The HippoCampus [41] is a small vehicle intended for research in underwater swarms, localization, and acrobatics. Other open-source AUVs are the result of student teams participating in robotics competitions [42], [43] but frequently lack design files for mechanical components of the vehicle. While the development of low-cost AUVs has been mostly done at universities, low-cost ROV have seen a rise in popularity as consumer, citizen-science tools/projects. Both the OpenROV and BlueROV are examples of such open source marine platforms for shallow water. They lead the open source marine robotics movement with high quality documentation, a large active user community and continuous development.

2.2 Design Drivers

One of the main reasons for the small number of deep sea AUVs in research is the high cost associated with owning and operating such equipment. This chapter is devoted to the design of the DROP-Sphere (DROPS), a new, low cost underwater vehicle capable of operating at depths of up to 6000m. This vehicle will allow for research operations in more than 50% of the world's oceans surface area and all fresh water bodies. In addition to the depth rating, the DROP-Sphere has been designed to be small enough to be carried and deployed by a single person without requiring specialized equipment. This creates a reduction in operations cost by reducing the number of operators and the required deployment vessel size, and this simplifies equipment transportation by qualifying as regular baggage on commercial flights. These design characteristics allow a switch from traditional monolithic vehicle approaches used for sea exploration to a distributed heterogeneous multi-vehicle approach in which multiple small AUVs deploy different sensing and monitoring capabilities. The proposed design represents the base configuration of such a system. Although the seafloor coverage capabilities of a single DROP-Sphere are not comparable with those of larger AUVs, its low-cost allows researchers and operators to adapt the number of vehicles to the project constraints: survey area size, available support vessels, time frame, and available economic resources. Additionally, offloading the mapping and sampling tasks on multiple smaller vehicles has the advantage of greatly reducing the risk and impact associated with vehicle loss.

2.3 Low Cost for Deep Sea

As mentioned in Section 2.2, one of the main characteristics of the developed vehicle is its low cost to enable access to the deep sea to a broader group of researchers and scientists. The development cost of an AUV can be roughly broken down into three major contributors:

- Mechanical components: Hardware pieces such as pressure housings, underwater buoyancy foam, connectors, and other required hardware to physically construct the vehicle, together with machining of components and raw materials.
- Sensors: All AUVs mount an array of different sensors to allow them to navigate and collect ocean data. Some examples include DVLs, USBL, or CTD probes. These sensors are high accuracy, low volume products and have high price tags attached.
- Development time: Design, assembly, and testing of both the hardware and software required on the vehicle are time consuming operations to be performed by highly qualified operators.

Our chosen design addresses each of these cost blocks in the following manner:

- Mechanical: The DROP-Sphere uses many commercial, off-the-shelf (COTS) components. Most significant is the use of a borosilicate oceanographic glass sphere as the main pressure housing. These devices have been extensively used for oceanographic research, are comparatively low cost, and are rated for extreme water pressures. Additionally, the optical properties of glass allow the placement of imaging cameras inside the housing. Custom mechanical parts or mounts can be manufactured through rapid prototyping technologies or with a small machine shop.

- Sensors: The main navigation sensors on traditional AUVs are the DVL, measuring vehicle speed, and USBL that allows users to acoustically estimate the position of the AUV in the water. Our design replaces both sensors with cameras and relies on state-of-the art computer vision techniques to estimate robot motion.

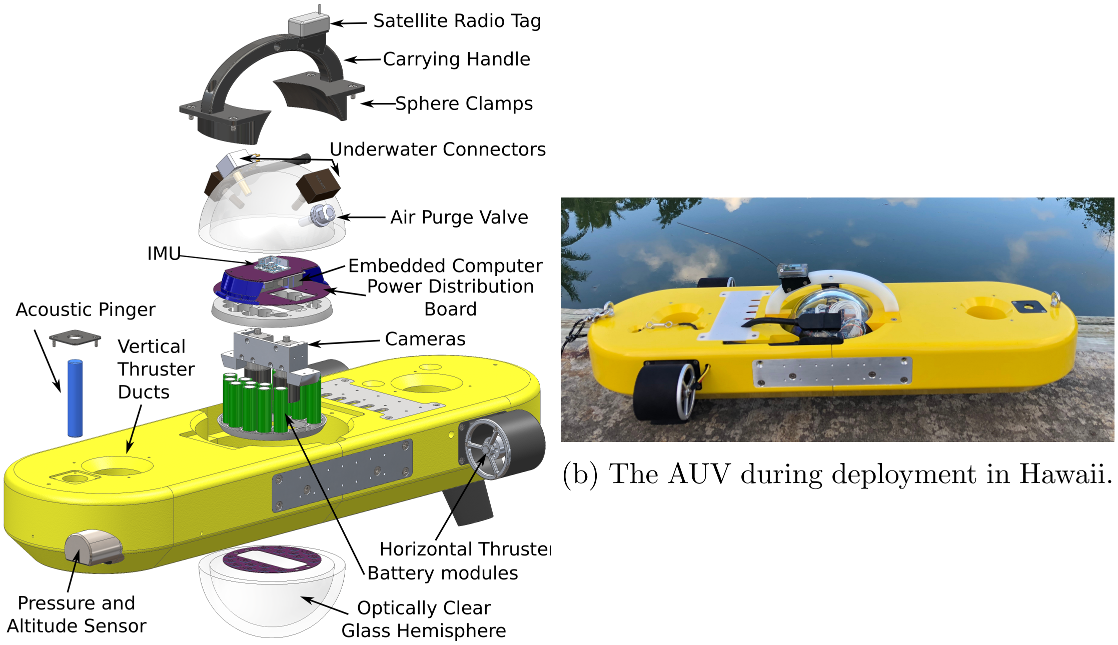
- Development: The complete development of the robot will be released as open source for use by the scientific and engineering community without restrictions. This includes all mechanical design files, electrical schematics, assembly instructions, as well as all associated software and firmware. In this way, reproducing the work presented in this thesis will involve a minimal amount of required time.

2.4 Vehicle Characteristics

As mentioned in Section 2.3, one of the main drivers for deep sea AUV cost is the requirement for custom pressure vessels composed of expensive materials, such as titanium [44] or made with carbon fiber composites [45]. To avoid this, the presented design uses a 0.2m diameter borosilicate glass sphere as its main pressure vessel. Glass spheres have been extensively used as instrument housings for oceanographic research and provide very high pressure ratings. The selected sphere gives our vehicle a theoretical maximum depth of 12000m. Deep sea rated foam provides the required buoyancy to compensate for the sphere weight. It also integrates the ducts for the vertical thrusters, emergency drop weights, and space for ballasting and trim weights. Fig. 2.1a shows an exploded view of the vehicle, illustrating the different components and their relative position.

For propulsion, four low-cost COTS brushless motors were used. These have the advantage of working when flooded, eliminating the need for custom thruster pressure housings. With proper freshwater rinsing and lubrication after each use, they do not suffer excessive wear. Additionally, their low cost makes it possible to simply exchange them when needed. The thruster assembly has been tested to equivalent depths of 7500m in a pressure vessel at Woods Hole Oceanographic Institution. The thruster configuration allows the control of surge, heave, pitch, and yaw degrees-of-freedom. All vehicle components are capable of functioning to at least 6000m of water depth.

The DROP-Sphere main sensing capabilities are enabled by two downward-looking Point Grey BlackFly $U3-23S6C$-C USB3 cameras. These are used to navigate based on visual-SLAM techniques and to produce maps of the surveyed area offline. These



(a) Vehicle exploded view

(b) The AUV during deployment in Hawaii.

Figure 2.1: The developed deep sea AUV

cameras are hardware triggered through a sync signal together with LED strobes for illumination. A Valeport VA500RAP pressure and altitude sensor enables the vehicle to estimate its position in the water column, while a global positioning system (GPS) module provides a geo-referenced position when on the water's surface. Finally, an XSENS IMU is used for vehicle attitude estimation. Other secondary sensors monitor variables such as temperature, current consumption, internal pressure, and humidity. An embedded Linux computer collects the sensor information and runs the vehicle's control, navigation, and localization software stack based on the Robot Operating System (ROS). Power is provided by a battery pack consisting of 18 18650 Li-Ion cells, supplying 210 Wh. The main characteristics of the vehicle have been summarized in Table 2.2, while Fig. 2.1b shows the complete prototype.

Table 2.2: Main characteristics of the DROP-Sphere AUV

Dimensions	0.86 m x 0.43 m x 0.25 m
Weight	20 kg
Max depth	6000 m
Propulsion	4x flooded brushless motors
Power/Autonomy	210Wh / 4-7 hours
Cost	<$35000

2.5 Vehicle Deployment and Operations

Of major importance for our vehicle are its seafloor coverage capabilities. This is dependent both on vehicle characteristics, such as the cameras and lenses mounted, as well as mission specific parameters, such as altitude over the seafloor and vehicle speed. Mission altitude is usually restricted by water clarity. Additionally a minimum of 25% image overlap is required to ensure enough image features can be extracted for a successful seafloor reconstruction. For the given combination of cameras and lenses (resolution, pixel size and focal length), and at a height of 2m from the seafloor a field of view width of 1.5m is expected. At a nominal speed of 1kn, a single AUV will cover 2778 m^2/h. Using the most conservative endurance estimates, this results in 11112 m^2/deployment for a single vehicle. If the same calculations are applied to a *REMUS 6000* AUV with a speed of 5kn and a endurance of 22 hours, 305580 m^2/deployment can be covered.

For descent to the ocean bottom, disposable weights are attached to the nose of the vehicle through galvanic timed releases. Figure 2.2 shows such a corrodible link together with the descent weights. Release times have to be chosen long enough to ensure the AUV has reached the seafloor before dropping the weights. During the descent, the vehicle is kept in a low-power consumption state and only monitors the pitch angle. While the weights are attached to the front of the vehicle the pitch angle remains approximately 90 degrees. Once the link corrodes and the weights are released, the vehicle returns to the horizontal position. At that point the main computer and other sensors are turned on and the mission is started. After the mission is finished, ascent to the surface is achieved through the vehicle's slight positive buoyancy. An ARGOS satellite tracking tag is used to obtain the vehicle location and retrieve it.

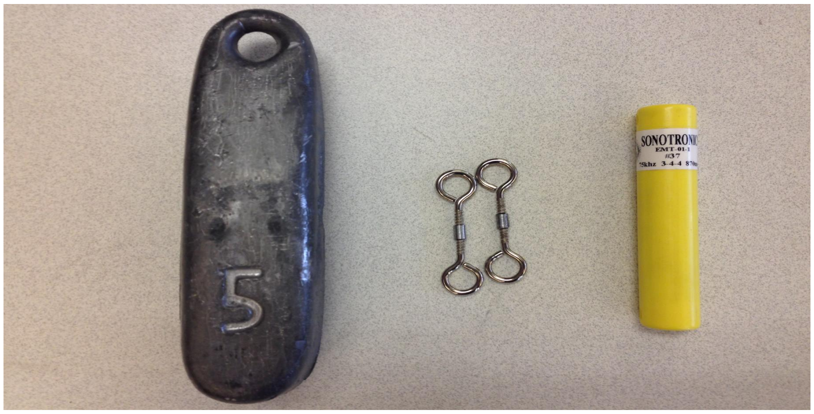


Figure 2.2: Descent weights and acoustic pinger

Finally, deploying multiple robots spreads the risk of catastrophic vehicle failure across the individual agents and mitigates its impact at the cost of map fidelity and the complexity of fusing multiple platforms data. This makes multi-vehicle approaches in its own nature resilient to vehicle loss.

2.6 Cost

One of the main drivers for this project has been to reduce the cost of deep sea capable devices. With the shown configuration, the vehicle cost is very close to $\$35,000$. Table 2.3 summarizes the main cost groups. Additional common hardware shared by multiple vehicles, such as the vacuum pump for closing the glass housings is not included in that list.

Table 2.3: Approximate cost breakdown for one DROP-Sphere.

System	Subsystem	Cost [$]
Mechanical	Glass Sphere	3000
	Buoyancy Foam	12000
	Underwater connectors	1250
	Other hardware	4000
	Thrusters	80
Electronics	Embedded Computer	75
	Printed Circuit Boards & Components	500
	Cameras & Lenses	3000
	Batteries	250
	Altimeter and Pressure Sensor	5500
	Satellite Tag	1500
	Acoustic Pinger	735
	IMU	1400
Misc.		1000
Total		**31890**

CHAPTER III

Underwater 3D reconstruction Benchmarking

The main deployment capability of the DROP-Sphere introduced in Chapter II is high resolution surveying of the seafloor with optical cameras. Recorded images are processed after a mission and assembled in large, textured 3D models. These models allow scientists and researchers to visualize the data at a larger scale and make phenomena that may not be apparent in individual images, visible. Benthic mapping is not exclusive to our vehicle and has become an important function of modern AUVs and ROVs. A wide range of fields from marine biology [46] to underwater archaeology [47] are increasingly relying on accurate 3D models of objects on the seafloor to perform their science. Optical cameras are playing an important role in such mapping due to their ubiquity and low cost. The improvements in open source robust feature matching [48], sparse non-linear bundle adjustment (BA), multi-view stereo [49], and front-end systems [50] have enabled end-users to apply these techniques with great success in a variety of terrestrial applications. Often, these tools designed for terrestrial problems are being used underwater due to cost-effectiveness and convenience; however, a thorough understanding of the accuracy and applicability of these methods to the underwater domain is absent from the literature.

The objective of this chapter is to present a quantitative evaluation of current methods for monocular and stereo 3D reconstruction underwater, using a range of open source and commercially available software and low- to high-cost camera systems. This error analysis motivates the decision to use high quality machine vision stereo cameras instead of monocular approaches in the design of the DROP-Sphere. This is especially important for the vehicle as computer vision techniques are used as the main navigational sensor to replace the acoustic DVL. Although commercial solutions for 3D reconstruction exist, using open-source alternatives presents a series of advantages: open-source code allows the users to customize the code for their specific application and usage scenario, it allows the tracking of changes and gives control over revision history and future development, and reduces the overall system cost. These benefits align with the intended characteristics of the DROP-Sphere as an accessible, low cost, benthic research tool for underwater biology and motivate the decision to develop all navigation and reconstruction software based on open source projects.

Further, the results in accuracy of the different approaches and camera systems motivate the need to explicitly model underwater image formation effects (Chapter IV). Additionally, we are releasing a novel open-source stereo reconstruction implementation to enable the processing of stereo underwater images into 3D models without any other sensor requirements.

3.1 Background

Terrestrial reconstruction techniques operate on the assumption that the brightness constancy constraint (BCC) holds true (i.e., that the intensity of an object viewed from two locations is the same) and that a pinhole camera model calibration describes the mapping of points in the world to points on the camera plane [51]. It has been shown that neither assumption is fully correct underwater [52]. Work has been done to quantify the degree to which the violation of these assumptions induces error in the recovery of 3D information from stereo and structure from motion [53]. Furthermore, the metric scale is an essential attribute of an end-user application (e.g., measuring coral, sizing an amphora in a shipwreck, etc.) and the inherent scale ambiguity in monocular structure from motion confounds the other assumption violations underwater. Jordt-Sedlazeck and Koch developed an underwater camera model to account for refraction through flat viewport housings and incorporated this into a framework for structure from motion [52], but their work focused on the monocular case and did not address stereo reconstruction. Bryson et al. developed a method to correct for several range-dependent effects such as absorption and scattering of light through the water column with automatic selection of attenuation coefficients [54]. In [55], the authors proposed a method that includes the estimation of underwater attenuation coefficients into the bundle adjustment step of 3D reconstruction, estimating a model for water-column effects. Instead of modelling water-column specific effects, this chapter presents a comprehensive evaluation quantifying and comparing the accuracy of methods for 3D reconstruction applied underwater that is absent in the literature.

3.2 Methodology

(a) Scene in air (b) Scene underwater (c) Ground truth surface

Figure 3.1: Artificial rock scene. These figures show the artificial rock scene created to quantitatively evaluate the quality of underwater reconstructions from various camera systems.

The goal of this chapter is to evaluate the relative performance of low- and high-cost imaging systems underwater to establish a relation between equipment cost and performance. We chose a low-cost system (GoPro Hero3+ stereo) that has become common in underwater applications [56] and a high-sensitivity machine-vision stereo pair (Prosilica GT1380) popular with robotics applications [57]. These cameras are representative for low- and high-end quality imaging systems. Despite both rigs being stereo, we emulate a monocular setup by treating each image independently and discarding the stereo calibration.

To allow us to compare the quality of reconstructions from different camera systems using different algorithms, we have developed the processing pipeline shown in Figure 3.3, integrating both a standard monocular and a novel stereo reconstruction pipeline to obtain quantitative results.

3.2.1 Scene and Ground Truth Mesh

In order to evaluate reconstruction accuracy, we created a rigid reference scene consisting of three plastic rocks mounted on a rigid base. The scene can be observed out of the water in Figure 3.1a and submerged during the experiments in Figure 3.1b. The reference artificial rock scene was scanned with a Kinect RGBD sensor, and a dense point-cloud was obtained using ElasticFusion [58]. The point cloud was processed inside Meshlab [59] and a surface mesh obtained through screened poisson surface reconstruction [60]. In Figure 3.3, the lower branch illustrates the ground truth generation steps. Figure 3.1c shows the ground truth mesh.

3.2.2 Reconstruction Methods

Three different reconstruction methods are compared in this chapter. Figure 3.2 displays the open source reconstruction methods processing pipeline.

3.2.2.1 Open Source Monocular Reconstruction

We chose to use the open-source Multi-View Environment implementation by Fuhrmann et al. [61] because it implements several of the current state-of-the-art monocular reconstruction methods. We processed monocular data by extracting all sparse speeded-up robust features (SURF) [62] and scale-invariant feature transform (SIFT) [48] features in all images, followed by matching of features between pairs of images. After matching, camera poses are incrementally computed from a starting pair of images with strong correspondences. A dense point cloud is generated through multi-view stereo [63]. The output surface mesh is obtained by applying Floating Scale Surface Reconstruction (FSSR) [64] to the dense point cloud.

3.2.2.2 Open Source Stereo Reconstruction

Extending [61], we have developed a fixed baseline stereo pipeline to go from underwater images to 3D models. Stereo camera calibration is incorporated into the processing pipeline by matching extracted features between corresponding image pairs first and then imposing the epipolar geometry constraint. This filters out bad features at an early stage and reduces errors when computing relative poses between image pairs. After the features are filtered, the same incremental structure-from-motion (SfM), multi-view stereo, and FSSR steps as in the monocular case are applied. We are releasing our implementation under `https://bitbucket.org/droplabumich/stereo_mve/overview`.

3.2.2.3 Commercial Reconstruction Software

The widespread use of commercial software solutions to perform underwater reconstructions motivates the evaluation of its performance compared to open source solutions to understand trade-offs between ease-of-use, cost, and reconstruction results. We use *Photoscan* because it has become popular in the underwater community despite being designed for terrestrial use [65] [66]. *Photoscan* reconstructions were performed with the highest quality settings available.

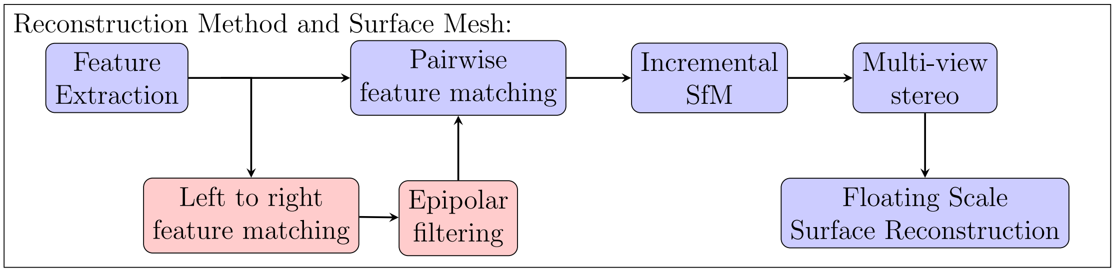


Figure 3.2: 3D reconstruction method pipeline: The reconstruction process for the open source monocular solution is represented by the blue boxes. Two additional steps are added (shown in red) for stereo reconstructions. Although the number of steps for stereo is higher, epipolar filtering reduces the number of features to match. Additionally, only one image per left-right pair is used in the pairwise matching step in case of stereo, further reducing the number of required operations by four.

3.2.3 Mesh Alignment and Error Metrics

The surface meshes generated by the different reconstruction methods were first coarsely aligned by hand with the ground truth reference mesh. We then applied iterative closest point (ICP) [67] to finely register both surfaces. To account for the uncertainty in the scale of the monocular reconstructions, the models were scaled to match the size of the ground truth before alignment, and scaling was added as a free variable to ICP. No scaling was applied to the stereo reconstructions. For the monocular reconstruction with the commercial software *Photoscan*, the same scaling and aligning as for the open source monocular implementation was applied to the obtained surfaces.

The obtained reconstructions were then compared with the ground truth mesh (Figure 3.1c) obtained with a laser-scan in-air by computing the Hausdorff distance [68]. This metric represents the maximum distances between all points on one of the meshes and their closest point on the other mesh. Reconstruction time was also recorded.

3.3 Experiments

To provide a comprehensive evaluation of reconstruction accuracy all three methods presented in Section 3.2.2 have been applied to two different camera systems. The 4ft. by 7ft. (1.22m by 2.14m) artificial scene was placed on the bottom of the University of Michigan's Marine Hydrodynamics Laboratory (MHL) towing tank (Figure 3.1b) at a depth of approximately 3m. Images of it were taken with the camera axis perpendicular to the ground plane. The cameras moved in the direction of

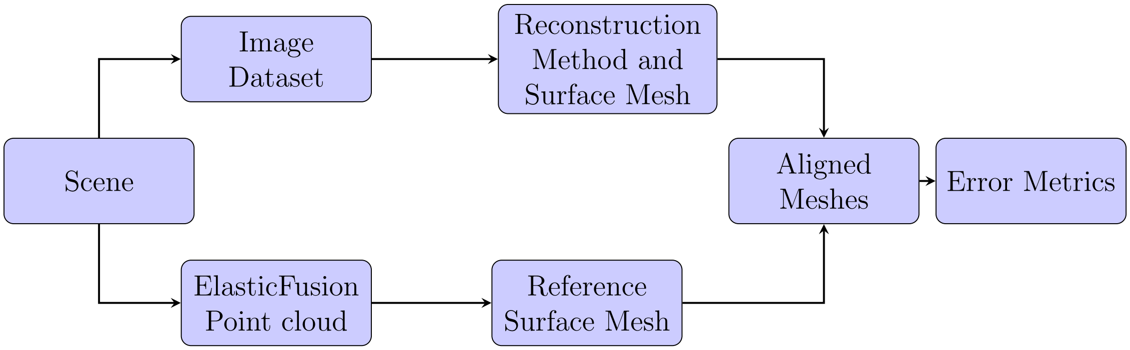


Figure 3.3: 3D reconstruction error evaluation pipeline. The evaluation pipeline used to compare the reconstruction quality of different camera systems with open source monocular and multi-view stereo methods as well as commercial software solutions.

the longer edge of the scene keeping a constant distance from the towing tank floor. Two underwater lights were attached to the camera mounting frame and provided constant illumination to the scene. A total of 160 images (80 image pairs) were used in the reconstructions.

Figure 3.4 shows the error plot color-coded onto the reconstructed mesh for all six combinations of the reconstruction method and camera system. The left column shows the high-end machine vision camera reconstruction, while the low-cost camera reconstructions are shown on the right. The first, second, and third row depict *Photoscan* monocular, open source monocular, and open source stereo reconstruction methods, respectively. A histogram of the error distribution is available on the right of the color scale. The ground truth reference is shown as a green wireframe mesh.

Table 3.1 summarizes the results seen in Figure 3.4, showing the average and standard deviation of the errors on each reconstruction method-camera system pair. The lowest values are highlighted and denote the best reconstruction.

It is important to note that resolution is a significant driver in the quality of reconstruction results. We experimented with different resolutions for both camera systems and found that the distortion and noise in the optics of the low-cost cameras limited the quality of the calibration required for stereo reconstructions. We selected a downsampling factor of four in order to reduce the pixelwise calibration error with respect to image size, resulting in a resolution of 676 by 381 pixels. For the high sensitivity machine vision cameras, we used the full 1360 by 1024 pixel resolution.

3.4 Main Experimental Insights

The completed in-lab experiments and reconstructions using unmodified terrestrial approaches show that the stereo system always outperforms the accuracy of a monocular reconstruction with the trade-off being a more complex setup. Moreover, low-cost action camera stereo systems are able to perform at levels close to those from high quality equipment at a fraction of the cost, making their use highly appealing for the scientific community.

Comparing the mean errors, the high-end cameras performed 38% better than the action cameras in the monocular case, while the improvements increased to 57% for the stereo case. Stereo reconstruction errors are lower: 38% for the high-end cameras and 47% for the low-cost cameras compared to the monocular reconstruction. The results from the commercial software *Photoscan* were very similar to our monocular reconstruction for the action cameras, but had an average error increase of 35% for the high-end imaging system. Both low- and high-end image reconstructions done with *Photoscan* have a large error on the top of the rocks.

The comparison of the histograms (shown to the right of the color scale in Figure 3.4) of the reconstructions of both imaging systems shows that the high-end cameras' errors accumulate in the lower bins, while the low-cost cameras' histogram maximum occurs at bins with higher error. Our hypothesis is that this difference in distribution is caused by higher errors in the calibration of the action cameras, which feature a wide angle lens, as well as higher noise levels in the images as a consequence of lower quality optics and sensors.

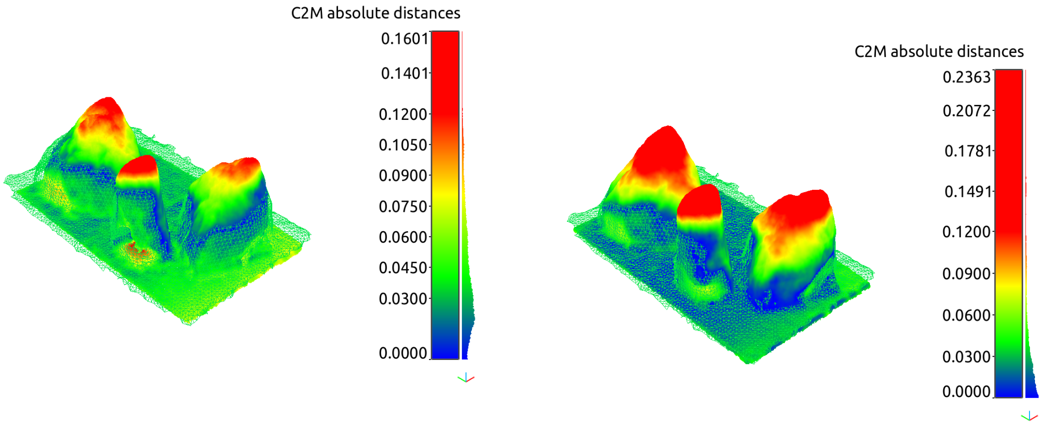
Warping along the axis of camera movement is another effect that can be clearly observed on all monocular reconstructions by noting how error increases along the short edges of the scene base. The higher number of constraints present in stereo pairs reduces this effect.

This comprehensive evaluation of error bounds of each system will enable us to

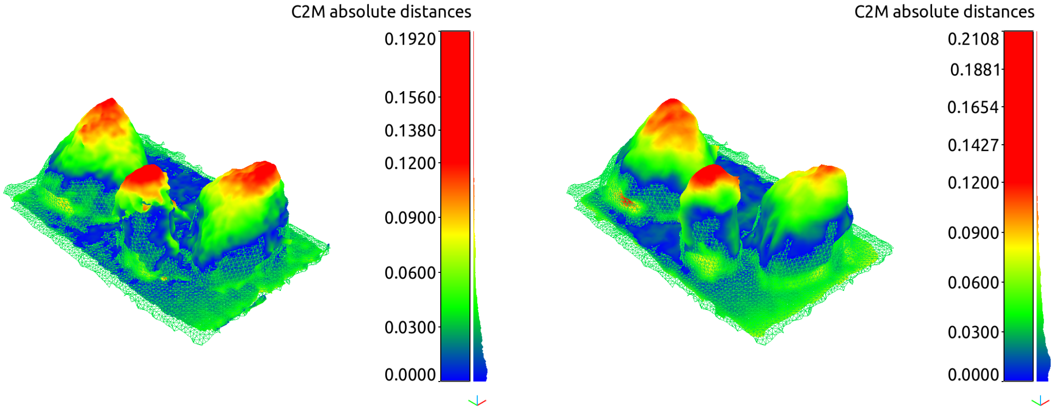
System	Mean [m]	Std Dev [m]
Low-cost Open Source Stereo	0.021	0.020
Low-cost Open Source Mono	0.045	0.040
High-end Open Source Stereo	**0.012**	**0.012**
High-end Open Source Mono	0.032	0.029
High-end PhotoScan	0.044	0.030
Low-cost PhotoScan	0.046	0.049

Table 3.1: Average reconstruction error and standard deviation

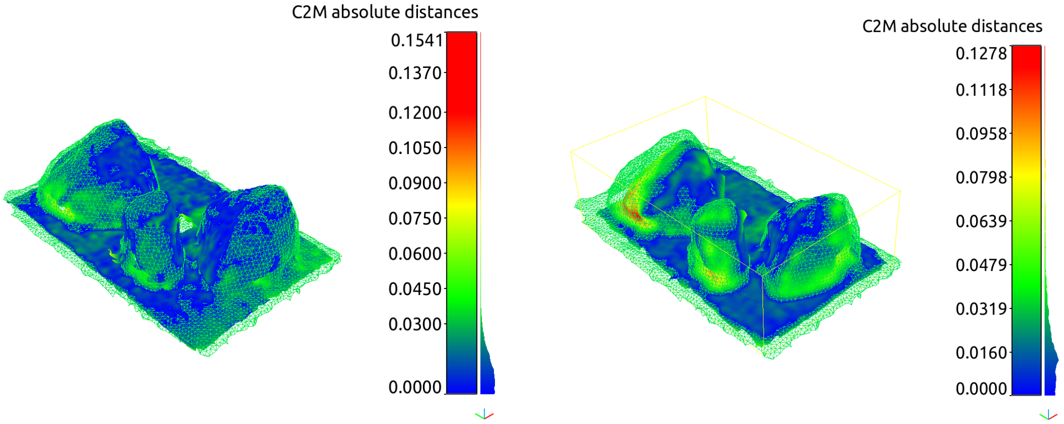
better understand the impact of the variable effects of the underwater domain on the accuracy of reconstructed 3D models in the field. It is important to remark that the absolute value of the error of monocular reconstructions ($\sim$ 4.5cm) is of the same order of magnitude as many biological processes of interest in the underwater domain (e.g., coral growth [69]); thus, single camera systems are not well suited for such applications. This highlights the advantages of stereo machine vision camera setups and drove the decision to use such a system for the DROP-Sphere. Additionally, we have shown the trade-offs between ease-of-use and support of commercial monocular software solutions and performance of stereo reconstructions.



(a) High-end Photoscan reconstruction (b) Low-cost Photoscan reconstruction



(c) High-end open source monocular reconstruction (d) Low-cost open source monocular reconstruction



(e) High-end open source Stereo reconstruction (f) Low-cost open source stereo reconstruction

Figure 3.4: Error plots for both camera systems comparing open source and commercial monocular reconstructions with multi-view stereo techniques. A histogram with error distribution is shown to the right of the color scale. The color scales have been clipped to 0.12m in order to allow direct comparison between the plots. 160 images were used for each of the reconstructions.

CHAPTER IV

Distortion Based Calibrations

4.1 Introduction

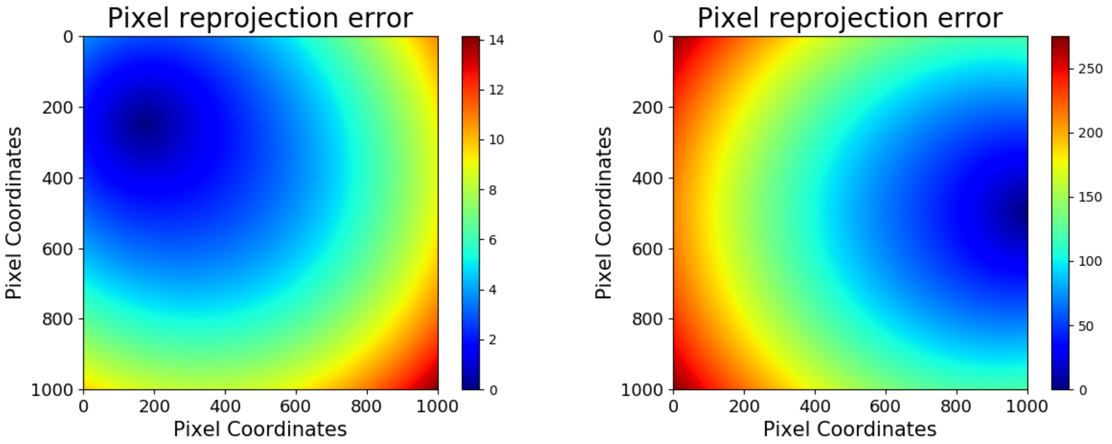
In order to deploy cameras in the aqueous underwater environment , optical systems require an additional housing that protects the lens and camera from water damage and pressure. Light rays travelling through this additional housing interface get refracted when changing mediums. The change in direction is proportional to the incidence angle and the quotient of the material's refractive indices. Because of refraction, camera rays no longer intersect the optical axis at the center of projection [53]. Underwater images are further affected by scattering and absorption, a wavelength dependent phenomenon that modifies the perceived color in the images. As a consequence, the widely-used pinhole camera model [51] is not valid [70].

Chapter III analyzed the errors introduced by the underwater environment on 3D reconstruction accuracy. Specifically, we looked at cameras with flat viewports and domed viewports. Flat viewports are simpler to manufacture and generally smaller. However, they reduce the camera field of view (FOV) by approximately 25% [71]. Inversely, domed viewports are more difficult to manufacture but do not introduce the refraction effects in the image as long as the center of the dome and the camera optical center are located in the same position. Although this is very difficult to achieve in practice, it has been shown [72] that, for a small misalignment, the error is negligible. However, there are situations when ensuring small alignment errors is not physically possible, such as when multiple cameras share a single domed viewport. Figure 4.1 shows the stereo camera pair configuration used by the DROP-Sphere, housed in a borosilicate glass dome. The effect of the large offset from the sphere center on the accuracy of pinhole camera models is illustrated in Figure 4.2. It compares the reprojection errors of small and big (mm vs cm) displacements of the camera inside the domed viewport and highlights the extent a large offset impacts the image formation process. The reprojection error serves as a metric for the quality of camera models and calibration data; this is calculated as the distance between the actual position of a point in the image and the position it would be projected onto using a camera model. More information on its definition and computation can be found in Section 4.4. One of the main characteristics of the DROP-Sphere is that it dispenses with the DVL and relies on the camera feed for navigation while surveying the ocean benthos. As a consequence, it is required to address the errors arising from large offsets in domed underwater cameras to enable accurate vehicle navigation and precise 3D reconstructions of the ocean seafloor. This chapter analyzes the applicability of methods developed for domed housings with small offsets and introduces the usage of the PSF as a means to compute the camera position inside the dome. To do so, the authors collected a dataset of PSFs for multiple camera-dome positions and show how the camera pose can be recovered by analysing the optical distortions.

Figure 4.1: Stereo camera used for vehicle navigation. The position of the cameras is far from the center of the hemisphere, introducing significant distortion in the images.

More specifically, the contributions of this chapter are two-fold. First, we analyze raytracing, a technique frequently used in the literature for underwater refractive models, and show how and why it fails in the case of domed viewports with large camera-dome offsets. Second, we propose the use of the PSF to characterize the complete optical system, consisting of camera, lens, and dome, and use it to obtain the camera pose within the dome.

The following describes how the chapter is organized. Section 4.2 presents a review of related work in the field of underwater refraction modelling. Section 4.3 analyzes the limitations of raytracing methods when applied to underwater domed housings.



(a) Reprojection with small displacement (b) Reprojection with large displacement

Figure 4.2: Pixel reprojection errors due to refraction. Computed for points on a plane at 3m in front of the camera that is located inside a sphere of 8cm radius. The left figure shows the errors for small displacements (∼2.5mm), while the right figure is generated with large (∼2.8cm) offsets between camera and sphere center. Although the errors would be significantly reduced if the system was calibrated with underwater images, refraction errors depend on scene depth and position in the camera frame.

Section 4.4 introduces the use of the PSF as a tool for domed housing camera calibration. Section 4.5 describes the experimental setup developed to measure the PSFs of the system, while Section 4.6 presents the results. Finally, Section 4.7 presents the conclusions and suggestions for future work.

4.2 Related Work

Modelling underwater image formation effects has been an important topic of research in recent years. Extensive work has been done with regard to the effect of flat viewport lenses on underwater cameras. A comprehensive overview of the literature on refractive underwater imaging and camera models is presented by Sedlazeck et. al. [70]. A popular approach has been to use an unmodified pinhole camera model [51] with a conventional calibration based on checkerboard patterns [73] underwater. Errors due to refraction, attenuation, and scattering are absorbed by the estimated camera intrinsic and distortion parameters. Although small reprojection errors can be obtained for a given set of high-quality calibration images, the projection of points and camera raytracing will be affected by systematic errors dependent on scene depth. However, examples of the successful application of this approach are numerous in the literature [74], [75]. Other authors propose to slightly modify the pinhole camera model for underwater use. Lavest et al. [71] analyzed the effect of flat viewports on the camera calibration and established that the effective focal length underwater is 1.33 times the focal length in air. They also adapted the camera distortion parameters to more closely model the refraction. Their derivation, however, only holds for viewports with a constant surface normal and does not generalize to domed housings.

Flat viewports also require the estimation of extra parameters to completely calibrate the imaging system. In addition to the camera extrinsics $[R|t]$, distortion parameters d, and intrinsics K, the distance between the camera and flat viewport, and the viewport normal, need to be estimated. Sedlazeck and Koch [76] presented a calibration method for flat viewport stereo cameras that does not require any special calibration target to be imaged. Additional parameters, such as interface thickness or refractive indices, can also be included in the calibration process. Sedlazeck and Koch [77] developed a method that also includes the estimation of the parameters of a radiometric model of light propagation. Much less work has been done to model, characterize, and correct refraction in domed viewports. One of the main reasons is that, when properly aligned, domed viewports introduce smaller distortions. Kunz and Singh [72] used raytracing to simulate refraction through a spherical interface and evaluate the induced error in 3D point triangulation. However, the authors focused on small displacements between the dome center and the optical center and did not consider situations where the offset is large compared to the hemisphere radius. Additionally, these experiments were limited to simulations. In a practical application, Menna et al. [78] presented methods for characterization of a commercial underwater camera domed housing and analyzed its properties. As opposed to flat viewports, where cameras can be shown to be axial, no simplified model exists for domed optical systems. As a solution, Sturm et al. [79], [80] presented a general imaging model that defined cameras as a collection of correspondences between pixels and imaging rays. Multi-view geometry relations and a calibration procedure are introduced for such a generalized camera model.

While these methods approach refractive image formation from a physics-based perspective, extensive work has been done in the image processing community to recover images that are degraded by a range of different effects, such as motion blur, chromatic aberrations, comma, or defocus. This process generally assumes that the observed image can be expressed as the convolution of the underlying “sharp” image with an unknown kernel, called the PSF. If the PSF for a system is known, the original, un-degraded image can be recovered through deconvolution. The characterization of the optical distortions through the PSF forms the foundation of the method introduced for domed housing camera calibration and is described in detail in Section 4.4.

4.3 Limitations of Raytracing Methods

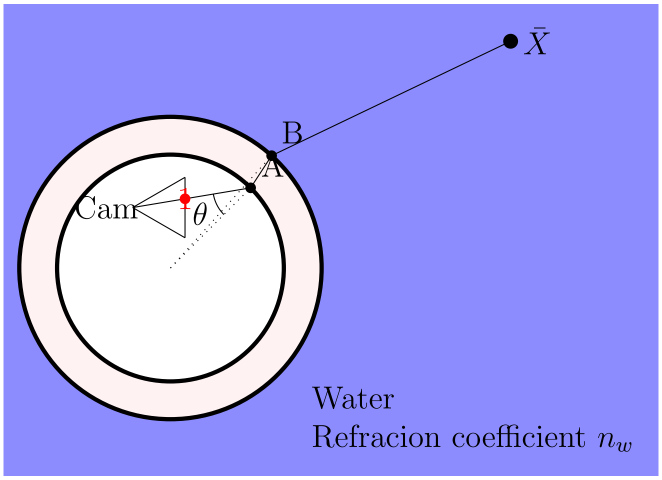


Figure 4.3: Refraction of underwater rays. This figure illustrates the effect of refraction on underwater image formation. Refraction effects have been exaggerated for illustration purposes.

As mentioned in Section 4.2, previous work mostly focuses on the modelling of refraction by tracing individual light rays imaged by the camera and explicitly computing the refraction points at each of the interfaces. Most developed methods rely heavily on a constant interface that is normal in the case of flat viewports, while, to the best of our knowledge, only simulation results have been presented for domed housings. In this Section, we analyze the issues arising when raytracing methods are applied to the calibration of cameras with large offsets in spherical dome housings (Section 4.3.1) and describe the main sources of image degradation (Section 4.3.2). The raytracing process is visualized in Figure 4.3. The figure shows a ray originating from the camera at a given pixel that intersects the inner glass surface at point **A** with angle θ. While going through the interface, the ray is refracted according to Snell's Law and then travels through the glass until it intersects the outer surface at point **B**. At this point, it is again refracted and the ray direction in the water can be obtained. This generates a set of *image pixel - underwater ray* correspondences that fully describe a general camera model and allow to obtain the camera ray representation in water for any given pixel coordinate. The inverse of this problem, to obtain the pixel that corresponds to a given 3D point, is called reprojection and has to be solved as an optimization problem due to the non-linearities introduced by refraction. For more information on general camera models and ray equation derivations, refer to [72], [80] .

4.3.1 Raytracing Based Calibration

In the context of most computer vision applications, the pinhole camera model [51] is used together with a calibration technique based on the imaging of geometric patterns such as checkerboards [73], circle grids, or AprilTags [81] to establish correspondences between pixel coordinates and 3D world coordinates. The raytracing underwater-image formation model requires a total of 24 parameters in the case of spherical dome housings: four camera intrinsics (K) parameters (focal length for each image axis and image center), five lens distortion parameters (three radial and two tangential coefficients), six pose parameters each for the camera and dome poses respectively, as well as the three refraction coefficients of air, glass, and water. The camera's intrinsic and distortion coefficients can be obtained by performing a pinhole camera calibration in air [73]. Furthermore, due to the rotational symmetry of the dome, the camera pose with respect to the dome reduces to the estimation of its translation vector. Finally, refraction coefficients for the glass interface can be obtained from manufacturer data with high accuracy using the Sellmeier Equation [82]. While the refractive index of air depends on temperature and pressure, changes are under 0.008% [83]. The index of sea water refraction is the most variable of the three with up to 3%, but, given pressure, temperature, and salinity, its value can be computed with an accuracy of at least 10^{-4} [84]. With these simplifications, the calibration of a domed underwater camera system only requires the estimation of the relative camera-dome translation $\mathbf{T}_c^{dome}$ and the camera extrinsics ($\mathbf{R}_w^c$ and $\mathbf{T}_w^c$) for each taken image.

Due to the nonlinear effects introduced into the image formation process by refraction, no explicit expression for the model parameters can be derived. Nonlinear optimization methods can be used to search for the correct parameter values through the minimization of a cost function.

Following the procedure outlined in [72], the calibration process is set-up as two nested optimization problems. The outer loop optimizes the relative position between the camera and dome, while the inner loop searches for the optimal camera extrinsics $\mathbf{R}_w^c$

and $\mathbf{T}_w^c$.

$$\mathbf{R}_w^c, \mathbf{T}_w^c, \mathbf{T}_c^{dome} = \arg\min \frac{1}{2} \sum_i^n \rho_i \left\| d_i \right\|^2 \tag{4.1}$$

$$d_i \left(\mathbf{R}_w^c, \mathbf{T}_w^c, \mathbf{T}_c^{dome} \right) = \frac{\| BX \times u \|}{\| u \|} \tag{4.2}$$

where d_i is the distance between the in-water ray, defined by the intersection point, with the outer surface B and its direction u going through pixel i and the corresponding 3D point X. We chose to define the cost function as the L2 distance between the ray and 3D point instead of the reprojection error because the reprojection of world points into the image is a minimization problem itself. Instead, the distance can be computed very efficiently.

The initial values for the camera poses are computed through the Efficient Perspective-n-Point (EPNP) algorithm [85], while initial values for the camera to dome relative position can be either measured on the physical system or CAD model.

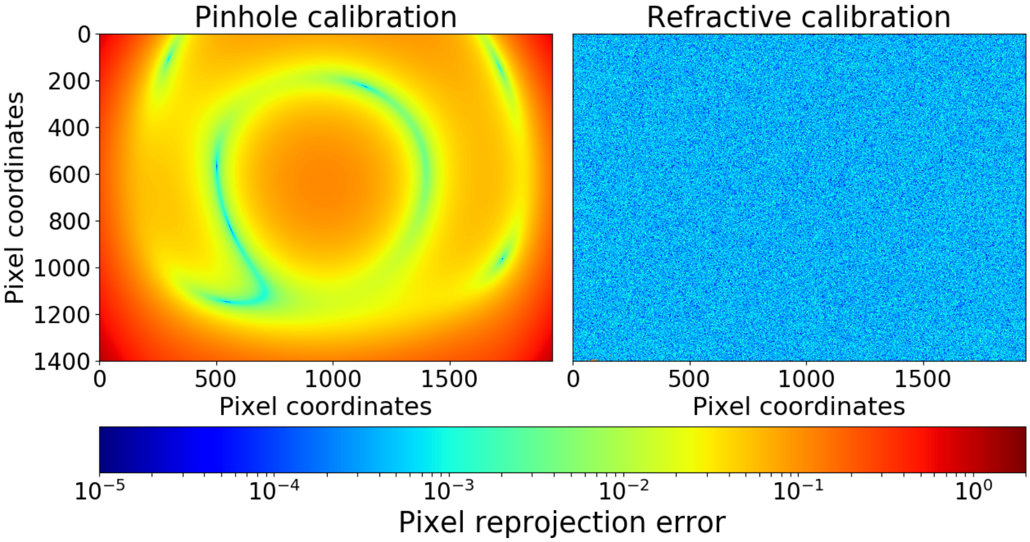


Figure 4.4: Result of camera calibration simulations without noise. Both images show the reprojection error of points on a plane at 3m in front of the camera. The left image represents the reprojection using a conventional pinhole camera model calibrated with underwater images, while the right image incorporates refraction with the proposed improved model.

Figure 4.4 shows the result of applying the proposed algorithm to noiseless pixel-3D point correspondences in simulation. The optimizer successfully converged on all noiseless trials and the resulting reprojection error is orders of magnitude smaller than what would be obtained if a pinhole model was assumed and calibration performed

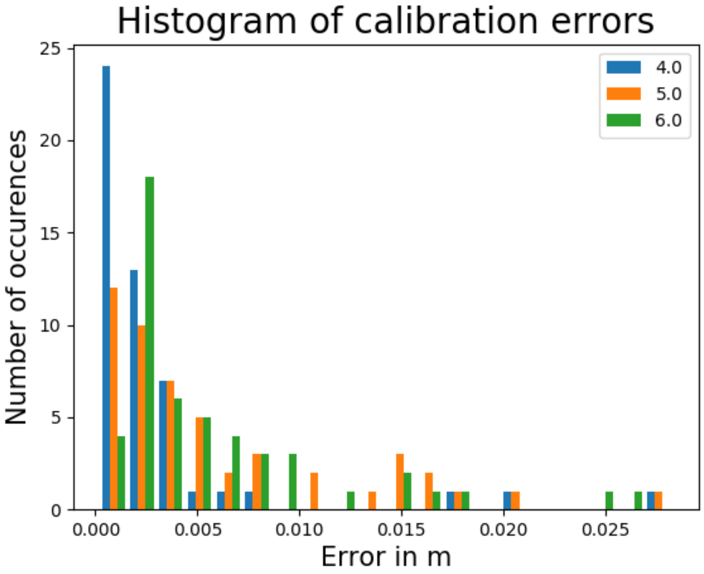


Figure 4.5: Calibration noise sensitivity. As pixel noise increases, the calibration optimization converges to local minima at significant distances from the true camera position and ray tracing will fail. Histogram legend in pixels.

using in-water checkerboard images. After performing experiments with noiseless data, the procedure was repeated introducing increasing levels of noise on the corners extracted from the image. The results are shown in Figure 4.5, showing the histogram of camera pose estimation errors as noise is increased.

4.3.2 Sources of Error

In addition to refraction based geometric distortion, images taken inside glass domed viewports are degraded by additional effects exemplified in Figure 4.6. This figure shows an image of different calibration boards acquired by one of the cameras inside the spherical housing shown in Figure 4.1 and highlights the following effects:

1. Space varying defocus: The spherical glass interface acts as a lens and generates a space varying defocus across the image plane. This effect can be seen when comparing the two zoomed areas in the bottom of Figure 4.6.

2. Chromatic aberration: The different light wavelengths are refracted in different amounts by the interface, generating a change in color.

3. Illumination falloff: Part of the light gets reflected on the glass dome surface in addition to lens vignetting, generating the illumination decay pattern observed.

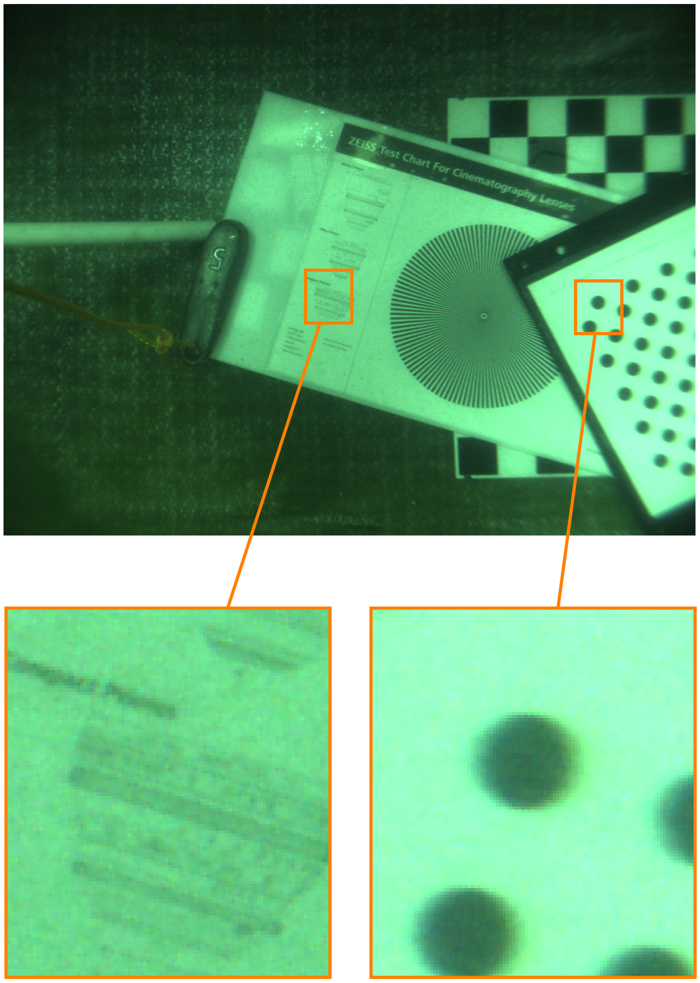


Figure 4.6: Underwater image degradation. Example of image obtained inside the dome with large offset from the center. Note the difference in focus across the image plane as well as the chromatic aberration along the circle edges. Figure details are best seen in digital format

The combination of these effects adds excessive noise to the position of corners extracted from the calibration targets. Multiple attempts were performed to apply the raytracing-based calibration procedure to real datasets with different calibration patterns without success. This section has analyzed the applicability and noise sensitivity of raytracing methods developed for flat viewports, proposed methods for domed viewports [72], and highlights the need for a different approach to the modelling and calibration of optical cameras housed inside spherical housings in the presence of large offsets.

4.4 Distortion Based Calibration

4.4.1 Image Formation

The degradation effects introduced in Section 4.3.2 can be encoded as a convolution kernel K. The observed image O is then the result of convolving the kernel with the original image I in the presence of additive gaussian noise N.

$$O = I \circledast K + N \tag{4.3}$$

The kernel K is known as the PSF. Intuitively, in an ideal optical system a point of light would appear as a single point in the image, with a one to one correspondence between image plane and object plane. However, in a real imaging system the point image is degraded or spread by the kernel K. Deconvolution can be applied to the observed image O to recover the latent image I. Methods can be divided into blind and non-blind methods. While blind methods estimate both the PSF and latent image simultaneously, non-blind methods provide better results when provided with a prior of the PSF. The PSF is a unique property of an optical system and fully describes it, that is, no two distinct optical systems share the same PSF [86]. As a consequence, if the PSF of a system can be measured, it should be possible to recover the parameters of the system that generated it. It is however a very high dimensional function that depends on optical properties such as lens geometry and materials, aperture and imaging distance. While in some cases such as motion blur a static PSF can be used for the whole image the general PSF will also vary along the image plane.

Shih et al. [87] showed that the PSF can be used to refine lens prescription parameters through optimization. Their work focuses however on small deviations from the nominal parameters due to manufacturing inaccuracies, while our work considers the dome-lens-camera system as an optical assembly with large parameter variations.

4.4.2 Distortion Based Calibration

Knowledge about the PSF, and its variations with respect to the optical system variables, can be used to recover the specific set of parameters that generate the image distortions observed. If we consider the relative position between dome and camera as independent variables of the PSF, obtaining the complete lens prescription is equivalent to calibrating the optical system. Such a priori knowledge about the PSF can be obtained in the form of measurement datasets or high-fidelity optical simulations. In this work, we focus on the case where a dataset of the PSF has been previously collected. Details about dataset collection and experimental setup are given in Section 4.5. If such a dataset of PSFs is available for a discrete set of camera-sphere relative positions, it is now possible to formulate the calibration problem as one where, given a measurement of the PSF, we aim to recover the x, y, z ($\mathbf{T}_c^{dome}$) position of the camera that generated such an array of distortions. Given a set of measured PSFs G_{ij}, each centered at pixel i, j, for one system configuration, and a dataset $U_{i,j}(x, y, z)$ of the PSFs for a different camera to sphere position x, y and z at the same i, j positions, we define the cost function as

$$f(x, y, z) = \sum_{R,G,B} \sum_{i,j} \|G_{i,j} - P_{i,j}(x, y, z)\| \tag{4.4}$$

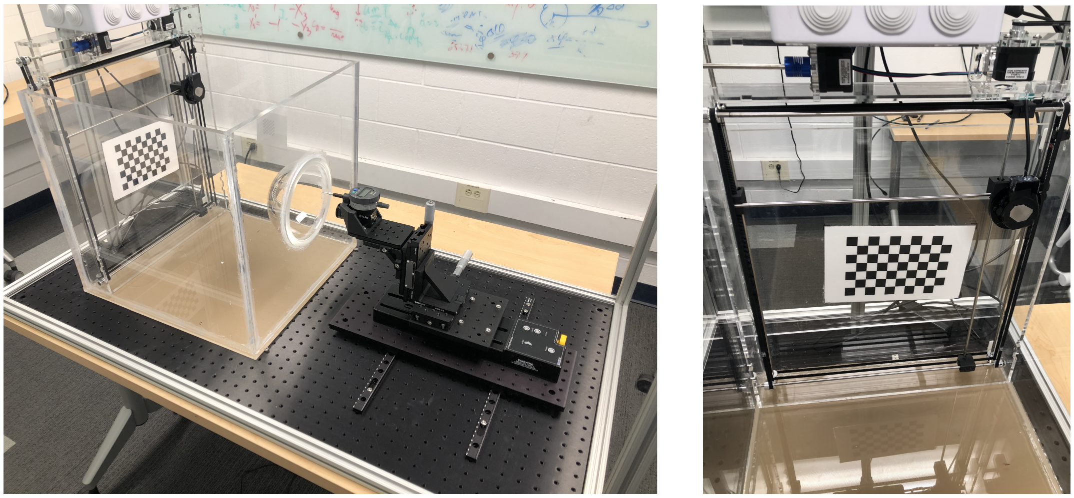
where $P_{i,j}(x, y, z)$ is the estimated response based on the available dataset. Because PSFs can only be measured at a discrete set of coordinates in the image plane, we use linear interpolation to obtain the PSF at the desired pixel coordinates. In the same way, PSFs can be interpolated between different camera poses [88] to obtain $P_{i,j}$ from $U_{i,j}$.

The camera position inside the dome can then be computed as

$$x, y, z = \arg\min f(x, y, z) \tag{4.5}$$

4.5 Experimental PSF Characterization

In order to characterize the changes of the PSF with respect to camera pose inside the spherical dome, we assembled the experimental setup shown in Figure 4.7. The camera is mounted on a three-axis linear-motion stage. The X axis is a electronically controlled, 150mm, travel-range stage; both the Y and Z axes are manual, 50mm, travel stages. The camera is mounted with the optical axis aligned with the system X axis. A precision indicator dial is mounted on top of it, aligned with the X axis as well. A custom acrylic tank was assembled to hold one of the BK-7 glass hemispheres used as the pressure housing by the DROP-Sphere in front of the camera. The hemisphere has an external nominal radius of 187mm and a thickness of 14mm. Finally, in the tank on the opposite side of the spherical dome at a distance of 370mm, a LED light source was mounted on an actuated Y-Z motion platform controlled by stepper motors. A 0.2mm pinhole cut out of 0.1mm thick stainless-steel was placed in front



(a) (b)

Figure 4.7: Experimental PSF measurement setup. The PSF of the complete camera-lens-dome optical system is measured for different camera position within the dome. Subfigure a shows the 3-Axis linear motion stage with camera and indicator dial mounted in front of the water tank with spherical dome. The light source motion system can be seen at the back of the tank. Subfigure b shows the light point source mounted on YZ translation system.

of the LED module to generate a small-point light source. The LED light can be configured to emit red, green, and blue light. Using a movable single light point as opposed to a grid of points ensured repeatable measurements and allowed exposure to be adjusted individually for each position to adequately use the camera's sensor dynamic range.

4.5.1 Measurement Calibrations

Before taking measurements of the PSF, the measurement system had to be calibrated. There were two different calibrations to obtain: a) Dome center calibration and b) Camera pose measurement in the system of coordinates.

4.5.1.1 Dome Center Calibration

The following steps were taken to obtain the center coordinates of the dome. First, the internal diameter of the glass dome was measured with a spherometer. Then, the x-axis of the linear stage assembly was advanced until the dial indicator tip was compressed about 10mm. Subtracting the dial indicator measurement from the linear stage position gives the x coordinate of point of contact between indicator tip and glass dome. The corresponding Y and Z coordinates are the positions of the respective linear stages. This process is then repeated moving the Y and Z positions to obtain a set of sample points of the internal glass surface. Finally, the dome center is computed by fitting the points into the equation of a decenterd sphere.

$$(x - c_x)^2 + (y - c_y)^2 + (z - c_z)^2 = R^2 \tag{4.6}$$

The resulting system of equations can then be solved through Orthogonal Distance Regression [89] that allows the inclusion of measurement errors in both the independent and dependent variables.

4.5.1.2 Camera Pose

Once the position of the sphere dome is known, we needed to compute the position of the camera's center of projection in the measurement coordinate frame. To do so, we first performed a conventional pinhole calibration in air to obtain the camera intrinsics. Next, a small checkerboard was placed on the flat side of the tank and imaged by the camera. Together with the camera intrinsics, this allowed computation of the camera position with respect to the board. Finally, the dial indicator tip was moved onto each of the checkerboard corners to obtain their coordinates in the measurement frame of reference. With the same set of points expressed in the camera and measurement coordinate frames, it was then possible to compute the transformation to express camera pose in the measurement coordinate system.

4.5.2 Experimental Procedure

For each camera pose, the light source was moved to 62 different Y-Z positions, evenly spaced along the back of the tank. Not all positions fall within the field of view of every camera pose, but are designed so that all camera poses have a similar number of measurements taken. For each of the positions, the PSF was measured for red, green, and blue light. The camera was configured to produce images without gamma correction in raw format and set to constant gain. Exposure was adjusted for each measured PSF to ensure the maximum pixel response was between 80% and 98% of the maximum pixel value to avoid over saturated regions. Additionally, the focus of the lens was adjusted for each pose. While this change in the back-focal distance affects the shape and size of the PSF and makes solving Equation 4.5 significantly harder due to the unmodelled focus variable. The collected dataset was also designed with future work in mind that will benefit from this property.

4.5.3 Collected Dataset

A Flir BFS-U3-63S4C camera, with a pixel size of 2.4um and a resolution of 2048x3072, was used together with a Edmund Optics UC-8mm lens. The used IMX178 color sensor uses a color filter array (CFA) in a RGGB Bayer pattern that requires demosaicing the measured PSF response. Schuler et al. [90] showed that the improvement, when explicitly considering demosaicing in the deconvolution process, only improved the resulting image peak signal-to-noise ratio (PSNR) by an average of 0.4dB. Based on these results, we exclusively demosaiced the raw image to obtain the PSF of each channel. The PSFs of a total of 200 camera poses were characterized during the experiments, with samples taken at 3mm intervals. In order to find the sphere center, a total of 180 points lying on the interior sphere surface where sampled, and the diameter measured to be 159.3mm.

While the collected dataset is specific to the camera, lens, and dome used, similar datasets can be collected for different configurations following the outlined procedure. Figure 4.8 shows the normalized PSFs for the three color channels for four different positions in the image plane and highlights the variability of the PSF as a function of wavelength, as well as pixel position, in the image plane.

4.6 Results

4.6.1 Calibration

The proposed distortion-based calibration method (Equation 4.5) was applied to different poses within the dataset described in Section 4.5. The minimization problem was formulated as a non-linear least squares problem and solved using the Levenberg-Marquardt algorithm. Figure 4.9 shows the error between the measured and estimated PSF at multiple different positions in the image plane. The optimizer converged to a position with an error of 0.184, 0.004, and 0.023mm on each axis. The error along the X axis is between one and two orders of magnitude larger than the Y and Z axis. Analysing the dataset showed that variation along the X axis mostly affected the size of the PSFs with a smaller impact on its shape. Blur due to a change in focus has a similar influence, modeled as a solid disk kernel of a varying radius depending on how much the image is out of focus. This similarity, together with the aforementioned fact that the camera focus was readjusted at each pose, is likely the reason for the discrepancy in error magnitude between the axis. Finally, the X axis was sampled very sparsely with only two different values in the dataset, which could

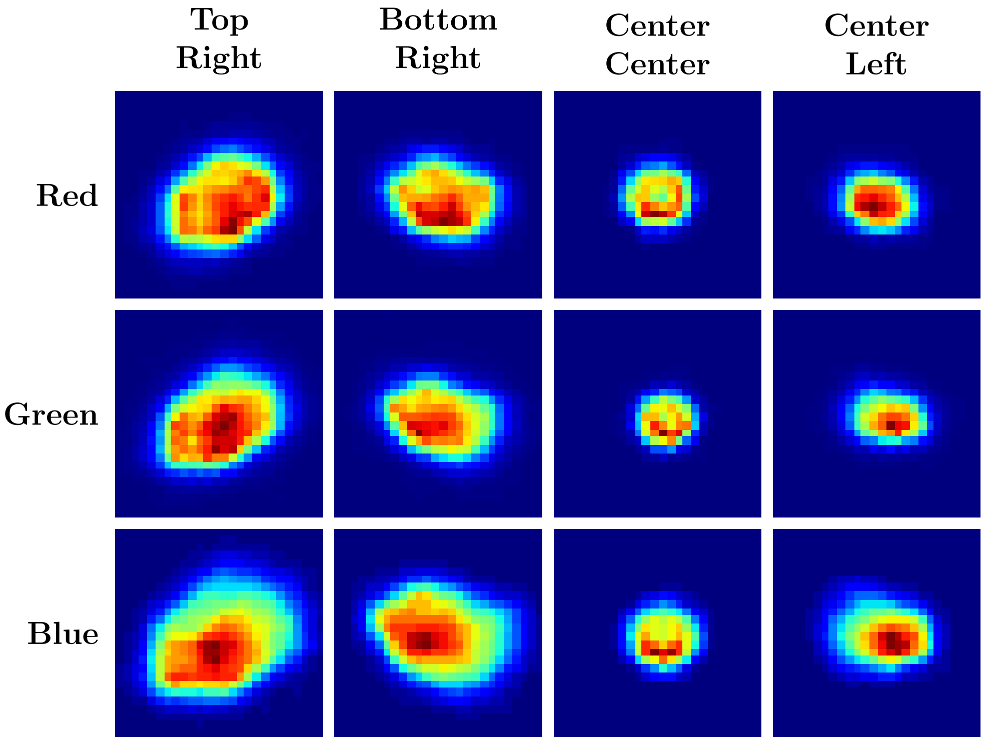


Figure 4.8: Examples of the measured point spread functions for four different positions in the image plane. The PSFs are shown for red, blue and green light. All PSFs are normalized with respect to their maximum response.

further influence the result.

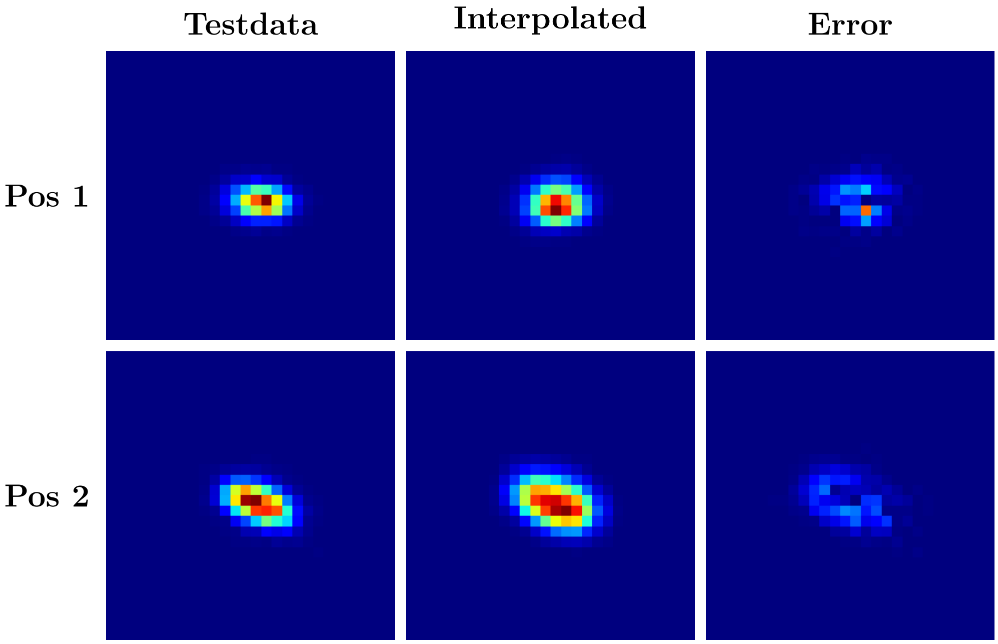


Figure 4.9: Resulting PSFs of calibration routine. Each row shows the measured PSF, as well as the result of the optimization at a different position in the image plane. The error plot shows the difference between the test data and optimization and represents a term in Equation 4.4.

4.6.2 Deconvolution

In addition to camera-to-dome pose calibration, the obtained point spread functions can be used to deconvolve the measured image and recover the underlying image. Figure 4.10 shows an example of a region of an image before and after deconvolution with the obtained PSF. We used the Van Cittert algorithm as implemented in *DeconvolutionLab2* [91]. In order to evaluate image sharpness, we use the variation of Laplacian, $\sigma^2_{\nabla^2}$ [92], a measure that increases for sharper images. Comparing the original and deconvolved image, values of $\sigma^2_{\nabla^2}$ increased from 19 to 62.

4.7 Conclusion and Future Work

In this chapter, we introduced the use of the PSF as an alternative characterization of an underwater optical system and showed it can be used to recover the position of the camera within the dome. This was necessary due to the limitations of state-of-the-art raytracing-based calibration methods when applied to cameras within

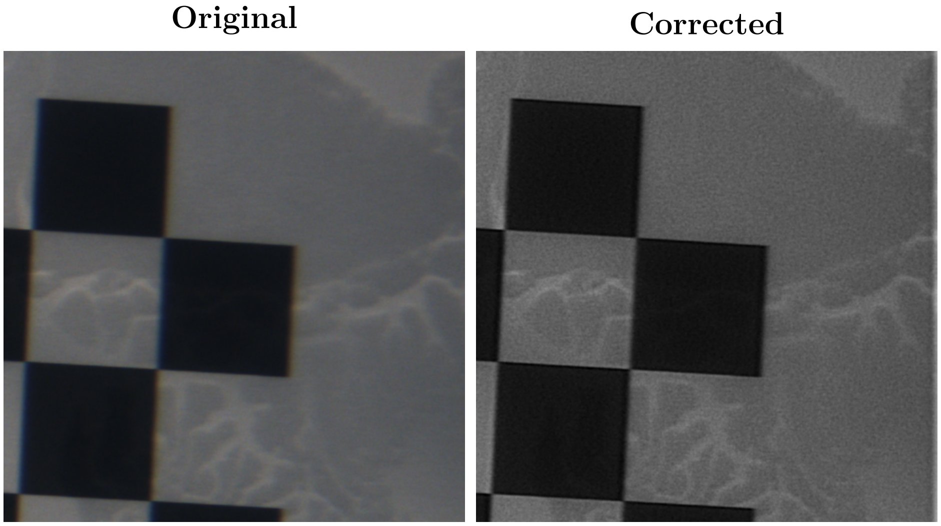


Figure 4.10: Image deconvolved with interpolated PSF. The variance of the Laplacian, a measure of sharpness, increased from 19.2 to 62.9

underwater domed viewports at large offsets.

The presented procedure is very time intensive and requires PSF measurements used for calibration to be performed under the same conditions (such as aperture or working distance) as the original dataset. This severely limits the application of the technique on field-deployed robotic platforms. Future work will address this by replacing the experimental data collection with PSF simulation data and modelling algorithms, enabling the regression dataset to be adapted to the calibration data conditions. Additionally, more advanced interpolation methods, such as methods based on principal component analysis (PCA) or Shapelets, have the potential to increase the calibration performance. We are sharing the collected dataset, under `https://umich.box.com/v/psfdataset-uwdomedcameras` for other researchers to have access to a collection of calibrated PSF measurements.

CHAPTER V

System validation and end-to-end mapping

5.1 Introduction

In addition to the challenges presented by the underwater environment, such as high pressures, corrosion, and low visibility, deep sea mapping requires the combination of many different disciplines and skill sets. Operational constraints, such as restricted working spaces on research vessels, reduced connectivity, and limited deployment time-windows require that AUV operations be highly automated and streamlined. These considerations are especially important for the DROP-Sphere, which is intended to be used by very small groups of researchers and scientists from small vessels.

Research on AUV operations has focused on mission and trajectory planning [93], [94], autonomy architectures [95], [96], and cooperative deployments [97]–[99]. Eichhorn et al. [100] developed a mission planning framework for gliders in dynamic ocean current scenarios. Mission planning for cooperative underwater developments is presented in [101] for adaptive sampling in order to predict ocean processes. In [102], the authors present a framework for the monitoring and quality assurance of AUV sensor data. However, none of the presented works consider the complete pipeline, going from mission specification to processed models and 3D reconstructions of the seafloor.

The main contributions of this chapter are: 1. the development of a complete, end-to-end AUV operation pipeline that goes from desired mission location to complete 3D reconstructions with minimal operator intervention, and 2. the validation of the DROP-Sphere vehicle as a platform for deep ocean optical surveys.

5.2 Autonomous Ocean Mapping Pipeline

The goal of the developed pipeline is threefold: enable vehicle operators to use their time effectively, provide the ability to efficiently process collected data, and provide an overall simplification of AUV operations. This is especially important when operating from small vessels where operator safety and health (e.g. motion sickness) place hard constraints on the tasks that can be performed at sea. This thesis introduces a state-of-the-art pipeline, shown in Figure 5.1, where each deployment of the DROP-Sphere AUV represents a complete run through it. Each step of the pipeline is further described in the following sections.

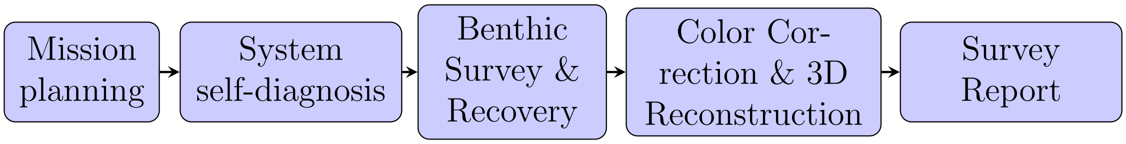


Figure 5.1: Automated end-to-end survey processing pipeline.

5.2.1 Mission Planning Software

DROP-Sphere missions are specified by a set of operator configurable parameters. The first is the survey pattern, which describes the trajectory geometry the vehicle will follow underwater. Frequent mission survey patterns include linear transects as well as rectangular “lawn-mower” grid patterns. The second configuration parameter, target altitude or depth, specifies at what altitude from the seafloor or depth the vehicle will execute the planned survey pattern. The third parameter, duration, specifies the maximum duration for the survey. Last are the safety parameters; these parameters represent limits that force a mission abort if exceeded. The generated mission plan is loaded onto the DROP-Sphere.

5.2.2 Autonomous Systems Diagnosis

Autonomous vehicles consist on multiple interconnected subsystems and it is critical to ensure all are functioning correctly before vehicle deployment. This is a repetitive, error-prone, and time consuming task currently performed by human operators. For swarm based deployments, automatically performing these tests enables timely deployments, builds the operators’ trust of the platform, and reduces the risks associated with AUV operation. On the DROP-Sphere, all sensors are monitored autonomously to ensure both the sensor measurement and measurement frequency are within acceptable ranges. In addition to aborting a survey when values are outside the safe interval, the vehicle uses a set of lights to transmit the vehicle status to the operator by changing the colors and blinking frequency. This allows quick assessment of the readiness of the platform before and after a survey.

5.2.3 Autonomous Benthic Survey and Recovery

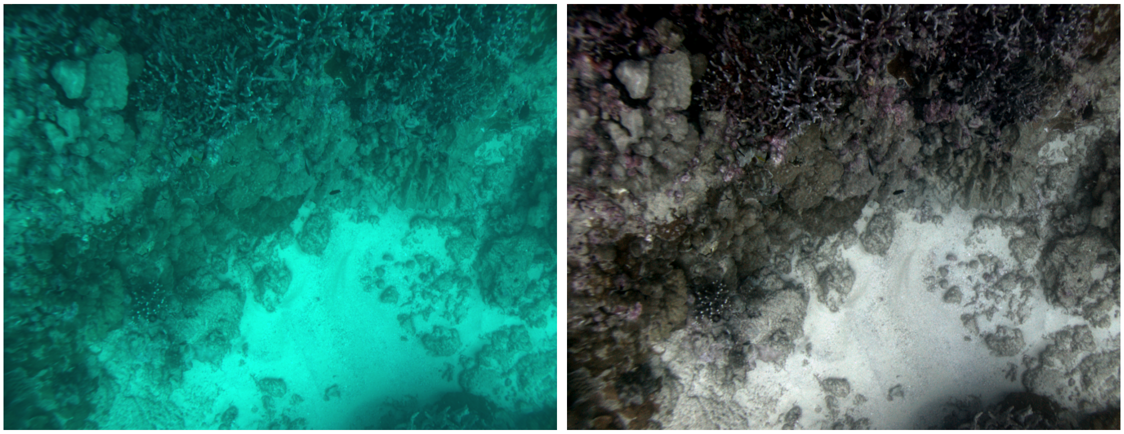
At this step, the vehicle is launched in the ocean and the autonomous navigation software follows the planned survey pattern on the seafloor. Once the survey is finished, the vehicle ascends to the surface through its positive buoyancy. The vehicle is equipped with an ARGOS satellite tracking tag, as well as radio and acoustic pingers, to allow operators to retrieve the vehicle after the mission ends. Argos satellite tags do not use GPS but instead compute their position based on the Doppler shift of the ARGOS satellite signal, with the position accuracy ranging between $150m$ and $1000m$. More details on the performance of the recovery localization devices will be given in Section 5.3.6.

5.2.4 Automated Color Correction and 3D Model Generation

Light attenuates as it travels through the water column. This effect changes the appearance of objects when observed underwater and gives the characteristic green or blue hue to underwater images, as seen in Figure 5.2a. After a survey is completed, images are downloaded from the vehicle and color-corrected to compensate for the wavelength dependent attenuation of light in water. We implement the Greyworld approach [103] to normalize image intensities. This method processes color channels independently and computes the corresponding mean and standard deviation. A gain and offset is then applied to the distribution to increase the image contrast. Figure 5.2 shows an example of an image before and after the correction algorithm was applied. Together, with the navigational data, a 3D textured mesh of the area is constructed. The mesh is generated using the approach detailed in Section 3.2.2.3. These meshes are the main data product generated by the vehicle and can enable scientific research on the ocean benthos, such as geological characterization, population density assessments, species diversity studies, among others.

5.2.5 Survey Report Generation

After the mission has been completed and the data processed, it is important to generate a comprehensive report with all collected information and data products for



(a) Raw Image (b) Color Corrected

Figure 5.2: Example of color correction of underwater imagery

easy human interpretation. The developed reports include all raw sensor readings such as depth, altitude, current and voltage of each power rail, IMU data, thruster set-points etc; mission parameters and vehicle configuration (controller parameters, safety thresholds, camera configuration), as well as processed data outputs such as energy consumption, 3D reconstruction overviews, image summary, and vehicle speed plots. This set of parameters allows operators to quickly identify anomalies in the mission execution, verify controller performance as well as evaluate power consumption. This systematic analysis of survey data allows for quick turnaround times and gives researchers the opportunity to optimize future deployments.

5.3 Autonomous Ocean Mapping Pipeline Validation

5.3.1 Location

In order to validate the vehicle's autonomous survey capabilities with the developed end-to-end processing pipeline, multiple deployments were performed over the course of a week in the coastal waters of the island of Oahu, Hawaii. The waters in Hawaii presented a great deployment opportunity due to the abundance of coral, as well as easy access to the ocean from the harbours along the shore. The shores of Oahu offer areas of steep depth increases in the south shore and more gradual slopes on the east coast of the island. Of special interest are the areas of Makapu'u point, as well as outside of Kane'ohe Bay. These areas have a high diversity of coral ranging from shallow-water corals to mesophotic and deep-sea corals; numerous previous surveys have characterized the area [104], [105]. A 22-foot boat, shown in Figure 5.4a, was

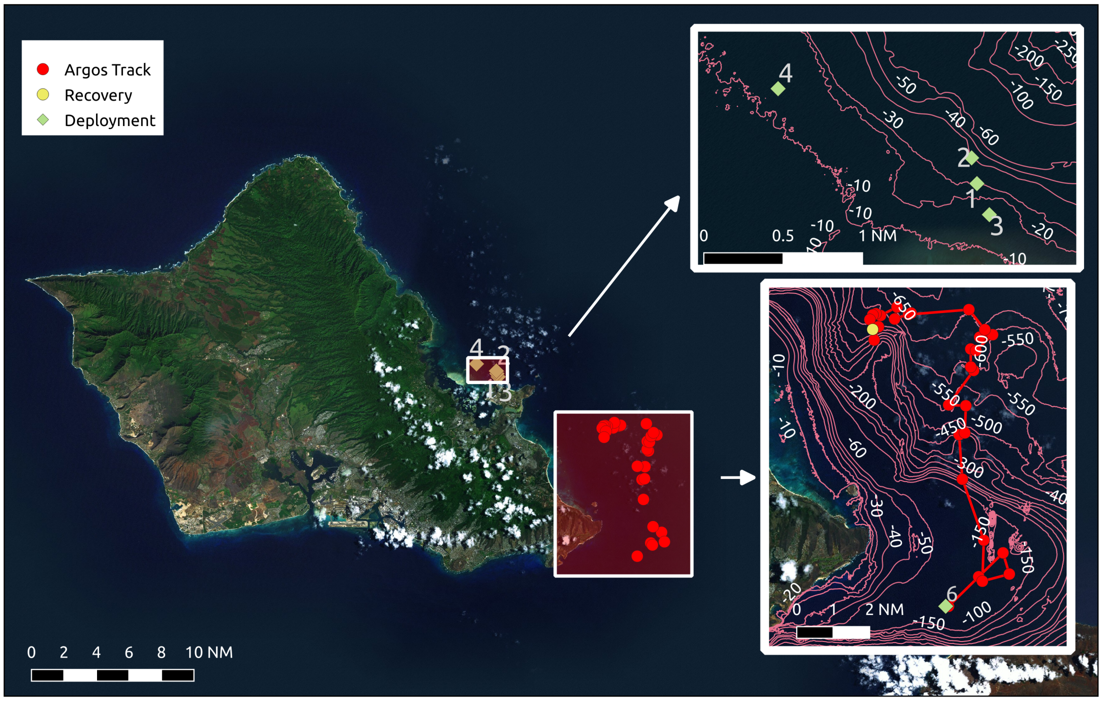


Figure 5.3: Vehicle deployment locations. Deployments were performed outside of Kaneohe bay (top insert) as well as Makapu'u point (bottom insert). The red line shows the trajectory of the vehicle on the ocean surface before retrieval.

chartered with a local captain, and highlights the deployment flexibility of a small, lightweight vehicle like the DROP-Sphere, shown before a deployment in Figure 5.4b.

5.3.2 Deployments Summary

The vehicle was deployed in the waters outside of Kane'ohe Bay (deployments numbered 1 through 5) and Makapu'u point (deployment number 6) at depths ranging from 8 to 80 meters. Missions were planned as linear transects with bottom times between 5 and 20 minutes and conducted at an altitude of $3m$ from the seafloor. Table 5.1 lists the performed deployments and the deployment coordinates. Figure 5.3 highlights deployment locations in both the top and bottom insert over a large-scale map of the island of Oahu. While the exact position coordinates for deployment #5 were not recorded, the vehicle was roughly deployed to the NE of deployment #4. The figure also shows the trajectory followed by the vehicle at the ocean surface after deployment #6. More details about this are given in Section 5.3.6.

Deployment #	Latitude	Longitude	Duration [s]
1	21° 28.32'N	157°45.840'W	1805
2	21°28.494'N	157°45.876'W	1869
3	21°28.110'N	157°45.756'W	644
4	21°28.962'N	157°47.184'W	589
5	-	-	5535
6	21°16.297'N	157°36.681'W	-

Table 5.1: Vehicle Deployments

(a) Vessel used for deployment (b) DROP-Sphere before deployment

Figure 5.4: Deployment Descent

5.3.3 Deployment Phases

Each deployment has four characteristic phases. The first, **Standby** phase, is the phase when the vehicle is turned on and floating on the water's surface. Once the pre-deployments checks have been completed, the vehicle is ready to transition to the second phase, **Descent**. This phase is where the vehicle gains depth either using the vertical thrusters when the target depth was under $60m$, or, for greater depths, using weights to sink to the ocean floor. The weights were connected to the bow of the vehicle through an approximately $2m$ long segment of braided line attached to a corrodible link. The corrodible links had a 2-hour release time, after which the link breaks and the vehicle is released and can start its mission. During the Descent phase, the vehicle achieves vertical speeds of approximately $0.1m/s$ when using the vehicles vertical thrusters; when weighted down with descent weights, the speed increases to

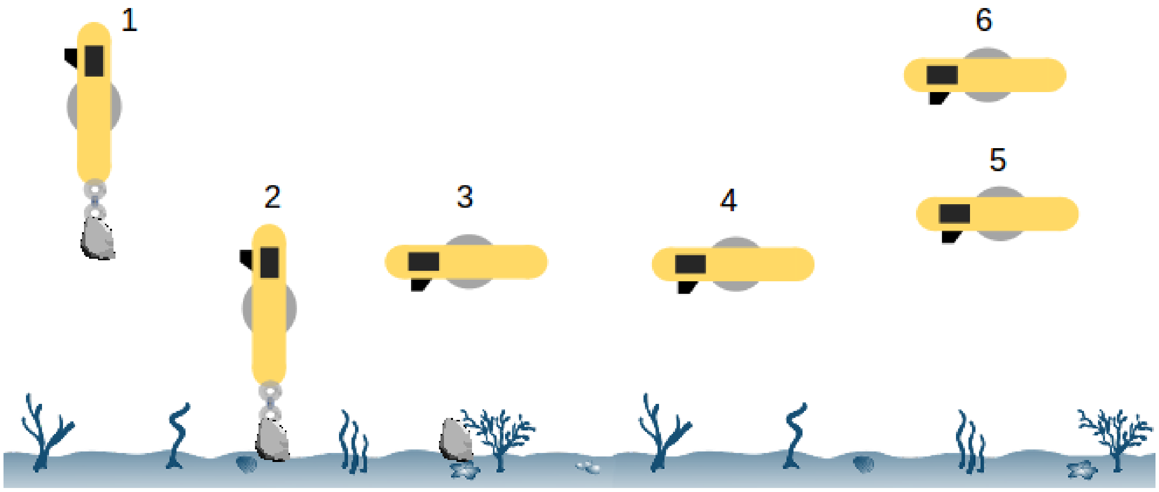


Figure 5.5: Main phases of a DROP-Sphere deployment. 1. The vehicle is weighted down to the seafloor using weights during the **Descent** phase, 2. The vehicle waits for the link to break, 3 & 4. The vehicle is released and the **Survey** phase begins, 5 & 6. The vehicle floats back to the surface during the **Ascent** phase

Deployment #	1	2	3	4	5	6
Descent speed [m/s]	0.10	0.09	0.10	0.11	0.44	0.49
Descent time [s]	284	381	115	112	160	145
Max depth [m]	31.3	43.0	13.9	14	72.5	77.6
Survey speed [m/s]	0.07	0.16	0.08		-	-
Survey time [s]	619	812	121	59	4395	-
Ascent speed [m/s]	0.15	0.14	0.12	-	0.10	-
Ascent time [s]	211	321	104	-	729	-

Table 5.2: Main Deployment Phase Characteristics

$0.45 m/s$. The third phase of the deployment is the **Survey** phase, when the optical survey of the seafloor is started after the vehicle is at the seafloor and free from the weights, if used. Once at the bottom, the speed of these surveys varied greatly between deployments, ranging from 8 to 16 cm/s, due to the deployments in shallower areas being strongly affected by tidal and surface wave-induced currents. Compared to speeds achieved in previous deployments in Traverse City Bay in Lake Michigan, where speed under the same vehicle configuration reached 27 cm/s, the ocean current strength can be estimated between 12 and 20 cm/s. After the survey time has elapsed or an abort condition triggered, the **Ascent** phase is started. Ascent to the surface was achieved using the vehicle's positive buoyancy, with speeds of $0.15 m/s$. Once the vehicle reaches the surface, it goes back into **Standby** phase until it is recovered. The phases are easily recognized in a depth plot such as shown in Figure 5.6. Table 5.2

shows the main statistics corresponding to each of the phases of the deployments.

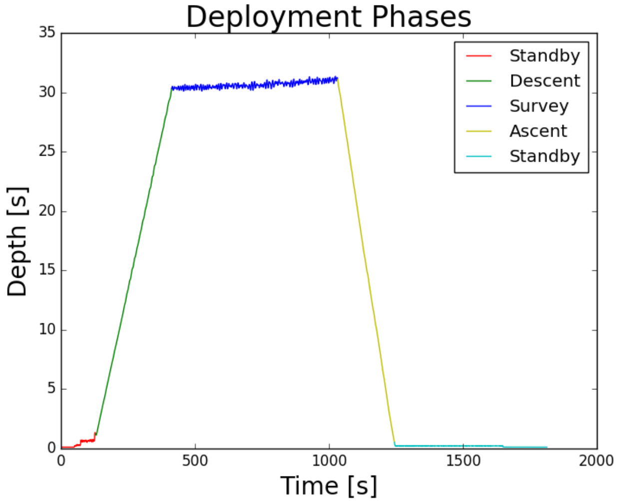
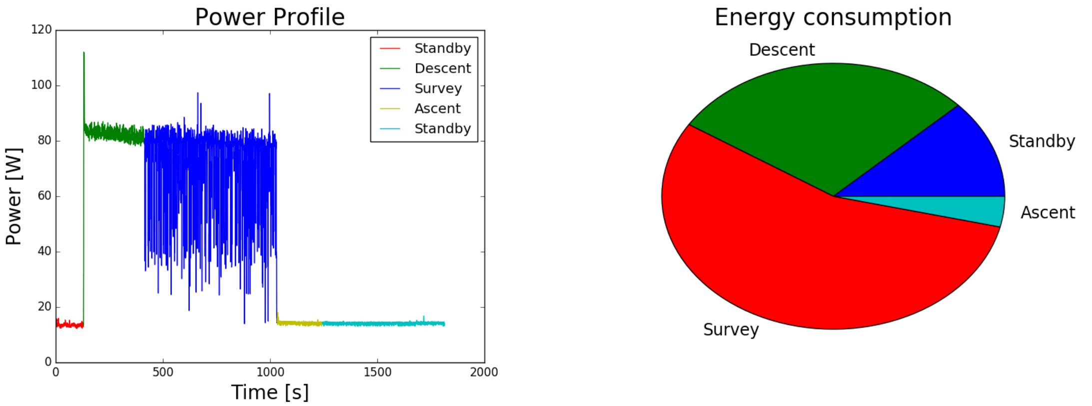


Figure 5.6: Deployment Phases

5.3.4 Energy Balance

Figure 5.7 shows a typical energy consumption profile corresponding to deployment #1. The graph shows how maximum power is demanded during descent due to the use of the vertical thrusters. Once at the bottom, the vehicle intermittently actuates the vertical thrusters to control altitude and the required power strongly oscillates. After the mission is finished, the thrusters are turned off and the vehicle ascends passively to the surface with a great reduction in consumed power. Once at the surface, power demand does not change significantly and remains low until the vehicle is retrieved by the operators. Consumed energy can be computed from the power graph as the integral over time. Figure 5.7b shows a plot of the total energy consumed for each of the mission phases. Table 5.3 summarizes the energy consumption figures for all six deployments.

Furthermore, the energy consumption can be grouped by main consumers, giving insight into not only when but where power is consumed in the AUV. Figure 5.8a shows how most of the energy is consumed by the thrusters. These energy consumption figures indicate that the vehicle is capable of approximately three hour mission surveys with the current $210Wh$ battery pack. Figure 5.8b shows how, during the survey mission phase, the vertical thrusters mostly operate at their maximum power, switching back and forth between push and pull modes. While this controller configuration is capable of satisfactorily controlling the vehicle altitude, it is certainly not



(a) Mission power profile (b) Energy consumption by mission phase

Figure 5.7: Energy Consumption Plots

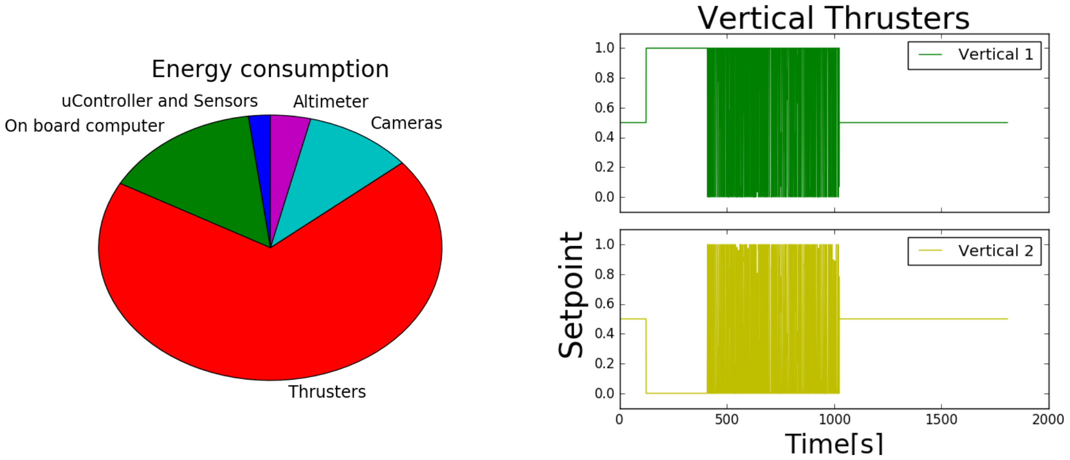
energy efficient and represents a direction for potential future research into energy-optimal control methods. The current configuration was chosen to enable the vehicle to submerge when at the water surface due to the reduced thrust delivered by the motors while at the surface. The energy consumption values obtained from the trials also allow computation of the depth to which it is more energy efficient to descend using the thrusters instead of using descent weights. This is due to the standby power that continues to be consumed while waiting for the corrodible links to release after reaching the seafloor. The critical depth can be computed as:

$$d = \frac{t_{LINK} \cdot v_{DESCENT} \cdot P_{STANDBY}}{P_{DESCENT}} \tag{5.1}$$

where t_{LINK} is the time for the corrodible link to release, $v_{DESCENT}$ and $P_{DESCENT}$ are the descent speed and power when using the vehicles vertical thrusters, and $P_{STANDBY}$ is the standby power. For a nominal corrosion time of two hours, a descent speed of 0.1 m/s and the mean of the powers in Table 5.3, the obtained depth is $123m$.

Deployment #	1	2	3	4	5	6
Mean Descent Power [W]	82.56	76.54	90.27	87.61	14.04	13.59
Mean Survey Power [W]	72.27	66.76	81.59	79.48	30.70	-
Mean Standby Power [W]	14.06	14.14	14.31	14.91	13.86	13.66
Total Energy [Wh]	22.53	25.80	7.33	10.89	48.02	-

Table 5.3: Energy Consumption



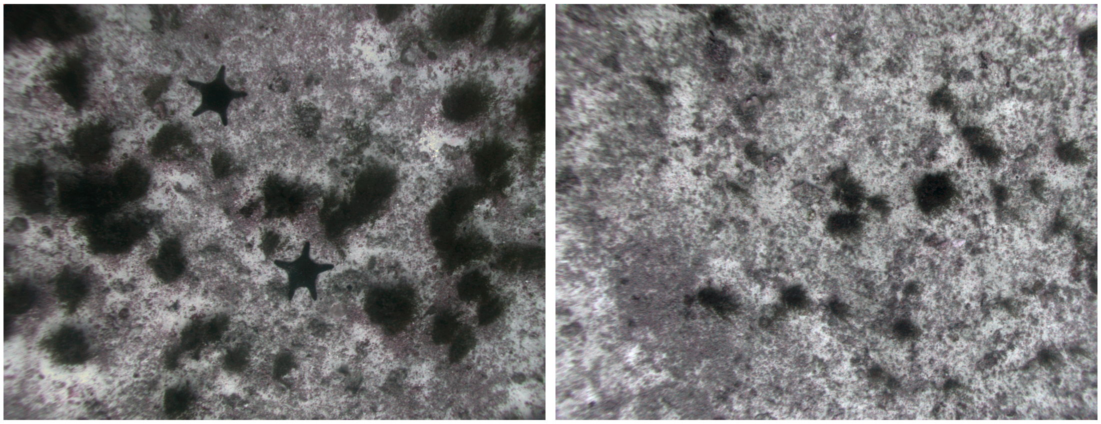
(a) Energy consumption

(b) Thruster usage profile during survey

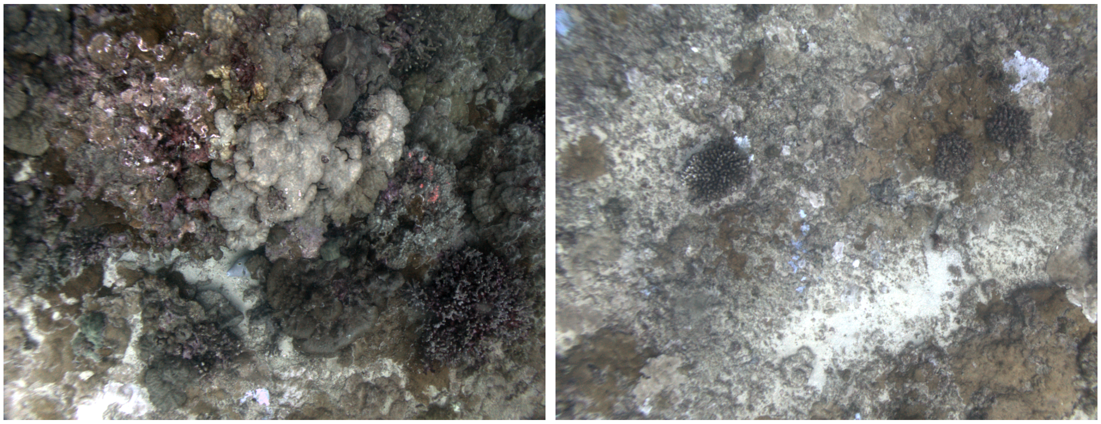
Figure 5.8: Plots of energy consumption by consumer and thruster setpoint profile.

5.3.5 Imagery and 3D Reconstructions

A total of 61505 images were collected during the trials. Images were captured at 10Hz rate during the survey, with a resolution of 1536px width and 1180px height. Images were taken at an altitude of 3m off of the seafloor using 2000lm LED light sources combined with variable intensity natural light, depending on deployment depth. Figure 5.9 shows samples of the obtained images after color correction and highlights how they can be used to count species, e.g., sea stars present in Figure 5.9a or coral coverage and species distribution in Figure 5.9c. After vehicle recovery, the images were color-corrected and processed into full 3D reconstructions. Figure 5.10 shows the resulting models for one of the survey transects and a detail view of the reconstruction of deployment #2. In Figure 5.10c a color coded height map highlights the rugosity of the terrain surveyed during deployment #3. The generated 3D reconstructions cover an approximate area of 500 m^2 of seafloor and provide insight into the terrain composition, rugosity, and biological cover. Software problems prevented the acquisition of images during the last survey over the Makapu'u bed, while a problem with the corrodible link during survey #5 made the vehicle unable to move adequately. As a consequence, no 3D model can be obtained from the images taken during that deployment. Table 5.4 shows the number of images taken, distance traveled, and total area surveyed for each of the deployments.



(a) 40m depth (b) 30m depth



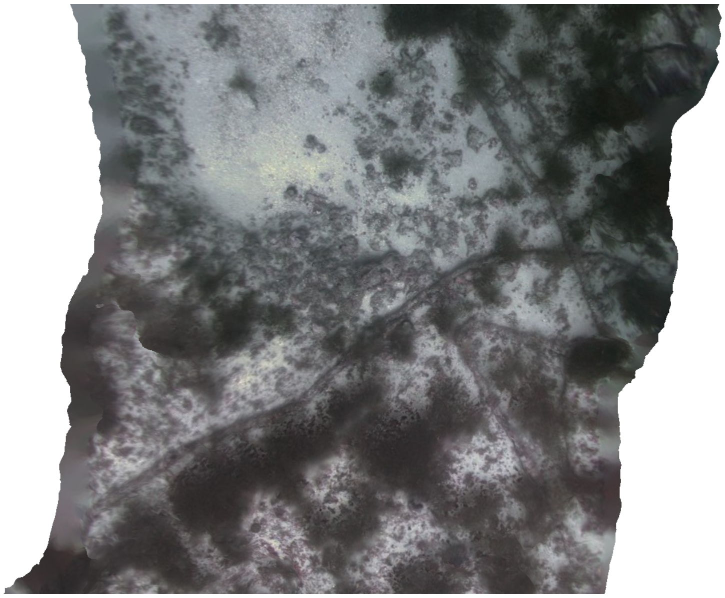
(c) 15m depth (d) 15m depth

Figure 5.9: Survey sample images

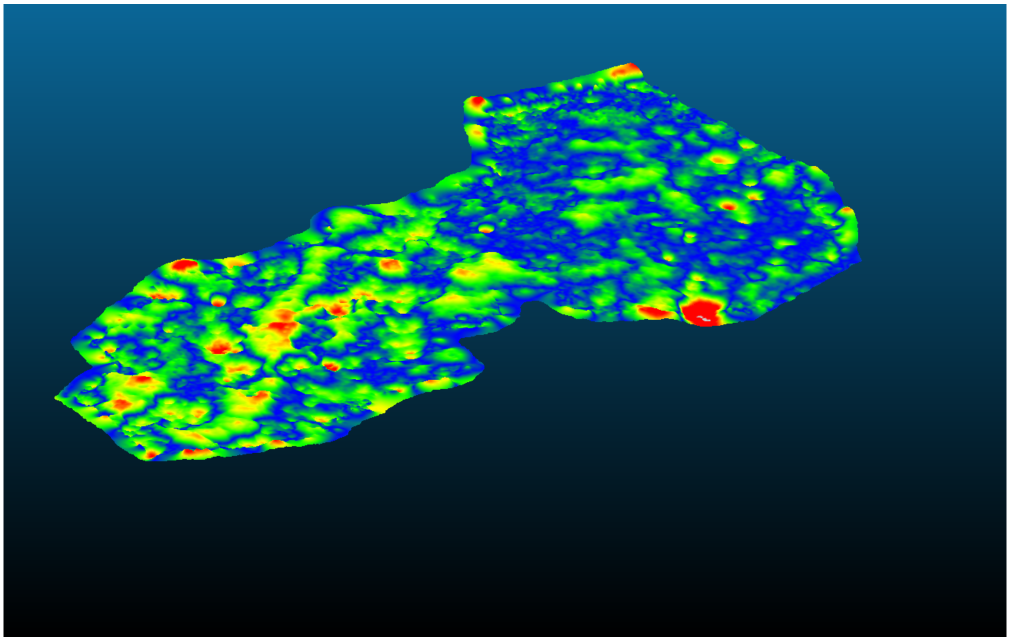
Deployment #	1	2	3	4	5	6
Number of images	14298	18868	3871	7108	17630	-
Transect length	46.8	129.1	10.3	15.6	-	-
Area Covered	126.1	348.6	27.81	42.1	-	-

Table 5.4: Image statistics

(a) Survey #1



(b) Abandoned rope/cables found in the reconstruction of Survey of #2



(c) Color coded relief of Survey #3

Figure 5.10: Survey sample 3D reconstructions

5.3.6 Recovery Issues

The last deployment took place at Makapu'u point, approximately four miles from the coast, and targeted the edge of the Makapu'u bed at a depth of 90m. The vehicle was deployed at 8:00am. While tracking of the underwater acoustic pinger was attempted after deployment with the hydrophone, it was postponed due to rough seas and motion sickness. The first location from the satellite tracking tag was received around 9:45am, indicating the vehicle had returned to the surface. The first fix is shown in Figure 5.3 as the southernmost red dot. Once the deployment boat reached the received position, it was not possible to locate the vehicle. The radio tracking equipment, composed of a directional antenna and radio receiver, was able to listen to the transmitted pings from the satellite tag but it was impossible to identify a direction where the signal was clearly stronger. This can be attributed to the lack of the crew's training in radio tracking, as well as a potential multi-path due to the large waves in the area. Vehicle position updates kept coming in once or twice an hour, but the vehicle could not be found. At 5pm, and after two hours without a satellite position fix, the search was aborted. During the night, the vehicle started sending position messages that indicated it had not been taken offshore by the currents but was rather moving North following the coast profile. The complete track is shown in Figure 5.3. The search was resumed the next morning during more favorable environmental conditions and the vehicle was found at 11:10am. The conclusions from this event are:

- The precision of the ARGOS satellite tag position estimate is not good enough to enable reliable, fast localization of the vehicle in the ocean in the presence of waves that limit line-of-sight.
- The use of radio tracking equipment to triangulate a transmitter requires adequate training and operator expertise. While the devices worked on land, they were of little use on the water. More advanced tracking equipment, such as radio goniometers, could be used to overcome some of the operational difficulties encountered.
- A direct satellite communication modem, such as iridium-based technology, needs to be mounted on the vehicle to transmit accurate, real-time, GPS-obtained position estimates to the operators for retrieval. However, ARGOS tags can still be used as back-up positioning devices.

5.4 Conclusions

This chapter introduced an automated pipeline for the deployment of the DROP-Sphere and the processing of the collected data. We deployed the vehicle in open ocean waters and performed fully autonomous optical surveys of the seafloor. The results highlight the capabilities of the DROP-Sphere to produce dense maps and 3D reconstructions of the seafloor that can be further used by researchers and scientists to characterize the ocean benthos. The recovery issues encountered will be addressed further by hardware changes to incorporate an Iridium-connected satellite modem into the design to enable accurate GPS position transmission.

CHAPTER VI

Conclusions

6.1 Contributions

The main contributions of this thesis are:

1. The development of the DROP-Sphere: a novel, deep-sea-capable AUV with the ability to perform optical surveys at high depth. The vehicle eliminates expensive acoustics-based sensors, like DVL and USBL and custom deep sea enclosures, and replaces them with optical cameras housed inside a commercial, off-the-shelf glass sphere. Its low cost enables smaller research groups to access the ocean benthos and reduces the barriers of entry to deep sea exploration and research.

2. A qualitative evaluation of error metrics for underwater 3D reconstructions, comparing state-of-the-art software implementations and hardware platforms. The results from this analysis were used to influence the design of the optical system of the DROP-Sphere vehicle.

3. The first step towards a calibration method based solely on the analysis of the distortions of underwater images taken through a spherical dome. This analysis also showed that the camera position inside the glass can be recovered by exclusively looking at the shape and size of the distortion kernels.

4. The development of an end-to-end mapping pipeline to streamline AUV operations. Field deployments in Hawaii validated the pipeline, as well as the rest of the DROP-Sphere vehicle, during which we mapped more than 500 m^2 and reached depths of up to 80m.

The presented survey platform, together with the developed software, represents a departure from traditional approaches to deep ocean research and provides an opportunity to explore and further characterize benthic communities. All work presented in this thesis, including the software and hardware, have been shared publicly to further encourage engineers and scientists to adopt the developed technologies and address the lack of knowledge about the deep sea.

6.2 Vehicle Shortcomings and Challenges

Before reaching the state presented in this thesis, the vehicle has gone through multiple iterations of both hardware, electronics and software. Many revisions were necessary to address engineering problems such as noise in the electrical system and to add additional features such as power monitoring or Bluetooth connectivity. Some of the challenges that motivated these revisions, as well as shortcomings of the current design, are listed in this section.

1. Thermal management: The cameras inside the DROP-Sphere generate large amounts of heat. Initial iterations of the vehicle design made extensive used of 3D printed parts that deformed after prolonged use of the cameras. This required a complete redesign of the hardware used to mount the cameras, switching to an aluminium bracket with large contact areas to the glass to transport heat away from the cameras and into the surrounding water. With this change, camera temperatures are kept within the safe operating range.

2. Processing power: The ODROID-XU4 was the most powerful small form factor single board computer available when the design of the DROP-Sphere was first started. While capable of running the basic software stack, more advanced processing algorithms require more processing power. As a consequence, the addition of demanding software modules is limited. More powerful and power efficient alternatives such as Nvidia's Jetson range of single board computers have been introduced to the market since the design of the vehicle and could replace the used ODROID-XU4.

3. Thruster assembly: The thrusters used are simple brushless motors that are low cost and designed to be replaced regularly as they wear out. The current configuration of the vehicle however lacks a way of replacing the thrusters in the field in a fast manner due to the need to cut and solder the thruster wires. This means that when a thruster failure is identified after a deployment, it usually

requires to return to shore to perform the switch. A connector based mechanism should be developed to address this issue and enable better use of deployment windows.

4. Localization system: As mentioned in Section 5.3.6, the localization system based on ARGOS technology failed to work reliably as a means to obtaining the vehicle position on the surface for retrieval. In order to address this, a custom Iridium based modem has to be developed to transmit the vehicle position obtained through GPS. While this will increase the vehicle cost by roughly \$3000 to \$5000, the cost will be offset quickly due to ARGOS tags being single use tags without rechargeable batteries that cost around \$1500.

5. Battery and endurance: The current battery system of 210Wh is, according to the performed field trials in Hawaii, capable of delivering power for approximately three hour missions. The designed endurance however was six hours, and changes to the controller and electronics need to be made to reach this figure.

6.3 Future Work

The development of the DROP-Sphere is a novel AUV concept that relies on optical sensors for navigation. The unique camera configuration inside a glass dome severely impacts image formation and introduces unique challenges not present in other visual navigation application domains, such as terrestrial or aerial robotics. This is critically important as the gathered images are not only used for vehicle navigation but are also the main data product gathered by the platform. While this thesis has proposed a specific design of the optical system and developed methods to help characterize it, there are generally two different ways at two different stages in the research the quality of the images can be improved: at the design stage and through the use of post-processing methods.

At the design stage, an analysis of the camera, lens, and geometric configuration of the system through optical design software allows refining of the camera, lens, and pose selection to minimize the effect of refraction and blur. This type of effort is frequently limited by the availability of high-level fidelity models for the lenses under consideration, as well as the high-cost associated with the required software packages.

Post-processing methods, however, focus on the removal of the optical aberrations after an image is taken. This is a difficult problem because there will be an inevitable loss of information in the image acquisition process. In the underwater domain, post-processing methods have mostly been focused on color correction and the modelling of the refraction at the interface. Domed housings, however, introduce additional blur when the camera is positioned far from the dome's optical center that further degrades the image. This thesis presented initial progress towards the calibration of such an optical system. Including the blur from the images into the analysis of underwater optical systems, and using it as a defining characteristic of the system, offers the potential to better model the effects of the glass interface and the water on image formation.

Additionally, deconvolution methods have shown to be very effective at removing specific blur types from images, including motion-blur, and could be applied to sharpen the underwater images after being captured. Image processing techniques like these can be combined with ray-tracing methods proposed in the literature to create a hybrid approach to image enhancement and geometric feature tracking.

In addition to image formation, localization of a small-sized AUV is a topic of significant importance, not only for navigation but also to localize collected data and provide context for scientists analysing the imagery. This task is further complicated by the lack of a DVL to measure vehicle velocities. Visual inertial implementations have been shown to produce good motion estimates when in close proximity of the seafloor. Localization is especially hard during the descent phase; the seafloor is still too far to enable the tracking of visual features and only the depth sensor and IMU provide navigation information. Research into dynamic modelling of the vehicle characteristics can provide valuable information to estimate the vehicle descent path.

Finally, the development of the DROP-Sphere required a large amount of engineering time to develop the mechanical hardware, electronics, and software that make the vehicle run. As such, it is a one-off prototype and subject to constant refinement and improvement of the design. Furthermore, it is very important to consider the inter-dependencies between the hardware and software that are present in a robotics platform because many software packages, such as visual-inertial navigation, have hardware requirements, such as a camera to IMU synchronization, implemented into hardware. Not considering these early on will force costly hardware design revisions. Many prototypes of each subsystem and module have been built and tested over the last few years until arriving at the current, functional prototype stage, and the cost of doing so was significant. Open sourcing all the software and hardware design files addresses part of the cost, at least for the community as a whole if not for the DROP-Lab. It not only allows others to replicate the design but, more importantly, modify it and use it as a starting point for custom designs and research projects.

The rest of the section discusses some of the mentioned areas of future research in greater detail.

Distortion based camera calibration

The method introduced in Chapter IV represents the first step towards a complete procedure not only capable of calibrating the camera system inside the dome housing with large offsets, but also capable of accounting for refraction and producing improved odometry estimates. In its current state, the proposed method requires large amounts of data captured in a controlled environment. However, optical simulation software has been shown to be able to produce accurate PSF estimations given the optical system prescription. The challenge lies in the high dimensionality of the PSF and the computational power required for its simulation. In parallel, deep learning offers the ability to learn high dimensional functions like the PSF. Future work will aim to combine the two to generate accurate, compact PSF representations to perform regression over and identify the camera model parameters. Furthermore, deconvolution and kernel estimation methods can be combined within the framework to enable the calibration of domed viewports from a single image and without the need of special calibration targets.

Vehicle localization

One of the immediate applications of the DROP-Sphere, enabled by its low-cost, is the simultaneous deployment of large groups of vehicles to cooperatively map an area. Previous work [106] introduced the development of low-cost acoustic transceivers for low-bandwidth acoustic communication and ranging. Multiple vehicles gather simultaneous data about a shared environment and enable communication between AUVs, allowing position refinement estimates for all vehicles involved. Expanding this research, future work will focus on combining coarse bathymetry maps with altimeter-based information through bandwidth-limited acoustical channels to improve the localization estimates of all vehicles.

Vehicle Development

Finally, future research on hardware changes will focus mostly on the development of a satellite communication modem for better retrieval, based around Iridium and GPS technologies, as well as smaller, incremental changes to increase the vehicle's reliability and reduce power consumption.

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] J. Copley. (2014). Just how little do we know about the ocean floor? [Online]. Available: `https://theconversation.com/just-how-little-do-we-know-about-the-ocean-floor-32751`. (accessed: 02.12.2019).

[2] D. T. Sandwell, R. D. Müller, W. H. Smith, E. Garcia, and R. Francis, "New global marine gravity model from cryosat-2 and jason-1 reveals buried tectonic structure," *Science*, vol. 346, no. 6205, pp. 65–67, 2014.

[3] S. B. Williams, O. R. Pizarro, M. V. Jakuba, C. R. Johnson, N. S. Barrett, R. C. Babcock, G. A. Kendrick, P. D. Steinberg, A. J. Heyward, P. J. Doherty, *et al.*, "Monitoring of benthic reference sites: Using an autonomous underwater vehicle," *IEEE Robotics & Automation Magazine*, vol. 19, no. 1, pp. 73–84, 2012.

[4] I. Mahon, O. Pizarro, M. Johnson-Roberson, A. Friedman, S. B. Williams, and J. C. Henderson, "Reconstructing pavlopetri: Mapping the world's oldest submerged town using stereo-vision," in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 2315–2321.

[5] K. Picard, B. P. Brooke, P. T. Harris, P. J. Siwabessy, M. F. Coffin, M. Tran, M. Spinoccia, J. Weales, M. Macmillan-Lawler, and J. Sullivan, "Malaysia airlines flight mh370 search data reveal geomorphology and seafloor processes in the remote southeast indian ocean," *Marine Geology*, vol. 395, pp. 301 –319, 2018.

[6] F. Dayoub, M. Dunbabin, and P. Corke, "Robotic detection and tracking of crown-of-thorns starfish," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 1921–1928.

[7] D. A. Smale, G. A. Kendrick, E. S. Harvey, T. J. Langlois, R. K. Hovey, K. P. Van Niel, K. I. Waddington, L. M. Bellchambers, M. B. Pember, R. C. Babcock, *et al.*, "Regional-scale benthic monitoring for ecosystem-based fisheries management (ebfm) using an autonomous underwater vehicle (auv)," *ICES Journal of Marine Science: Journal du Conseil*, vol. 69, no. 6, pp. 1108–1118, 2012.

[8] C. Forney, E. Manii, M. Farris, M. A. Moline, C. G. Lowe, and C. M. Clark, "Tracking of a tagged leopard shark with an auv: Sensor calibration and state

estimation," in *2012 IEEE International Conference on Robotics and Automation*, May 2012, pp. 5315–5321.

[9] E. Ramirez-Llodra, P. A. Tyler, M. C. Baker, O. A. Bergstad, M. R. Clark, E. Escobar, L. A. Levin, L. Menot, A. A. Rowden, C. R. Smith, and C. L. Van Dover, "Man and the Last Great Wilderness: Human Impact on the Deep Sea," *PLoS ONE*, vol. 6, no. 8, P. Roopnarine, Ed., e22588, Aug. 2011.

[10] M. A. Rex and R. J. Etter, *Deep-sea biodiversity: pattern and scale*. Harvard University Press, 2010.

[11] C. Mora, D. P. Tittensor, S. Adl, A. G. Simpson, and B. Worm, "How many species are there on earth and in the ocean?" *PLoS biology*, vol. 9, no. 8, e1001127, 2011.

[12] L. L. A. and M. R. H., "Isotopic evidence for chemosynthesis-based nutrition of macrobenthos: The lightness of being at pacific methane seeps," *Limnology and Oceanography*, vol. 47, no. 5, pp. 1336–1345,

[13] G. W. Rouse, S. K. Goffredi, S. B. Johnson, and R. C. Vrijenhoek, "Not whale-fall specialists, osedax worms also consume fishbones," *Biology Letters*, vol. 7, no. 5, pp. 736–739, 2011.

[14] C. L. Sabine, R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tilbrook, F. J. Millero, T.-H. Peng, A. Kozyr, T. Ono, and A. F. Rios, "The Oceanic Sink for Anthropogenic CO2," *Science*, vol. 305, no. 5682, pp. 367–371, 2004.

[15] K. R. Anthony, D. I. Kline, G. Diaz-Pulido, S. Dove, and O. Hoegh-Guldberg, "Ocean acidification causes bleaching and productivity loss in coral reef builders," *Proceedings of the National Academy of Sciences*, vol. 105, no. 45, pp. 17 442–17 446, 2008.

[16] A. R. Thurber, A. K. Sweetman, B. E. Narayanaswamy, D. O. B. Jones, J. Ingels, and R. L. Hansman, "Ecosystem function and services provided by the deep sea," *Biogeosciences*, vol. 11, no. 14, pp. 3941–3963, Jul. 2014.

[17] B. S. Halpern, S. Walbridge, K. A. Selkoe, C. V. Kappel, F. Micheli, C. D'agrosa, J. F. Bruno, K. S. Casey, C. Ebert, H. E. Fox, *et al.*, "A global map of human impact on marine ecosystems," *Science*, vol. 319, no. 5865, pp. 948–952, 2008.

[18] M. F. Dolan, A. J. Grehan, J. C. Guinan, and C. Brown, "Modelling the local distribution of cold-water corals in relation to bathymetric variables: Adding spatial context to deep-sea video data," *Deep Sea Research Part I: Oceanographic Research Papers*, vol. 55, no. 11, pp. 1564–1579, 2008.

[19] C. R. Fisher, P.-Y. Hsing, C. L. Kaiser, D. R. Yoerger, H. H. Roberts, W. W. Shedd, E. E. Cordes, T. M. Shank, S. P. Berlet, M. G. Saunders, E. A. Larcom, and J. M. Brooks, "Footprint of deepwater horizon blowout impact to deep-water coral communities," *Proceedings of the National Academy of Sciences*, vol. 111, no. 32, pp. 11 744–11 749, 2014.

[20] P. Rigby, O. Pizarro, and S. Williams, "Toward adaptive benthic habitat mapping using gaussian process classification," *Journal of Field Robotics*, vol. 27, pp. 741–758, Nov. 2010.

[21] MTS. (2017). Active manned submersibles by depth, [Online]. Available: `http://www.mtsmuv.org/active-manned-submersibles-by-depth/`. (accessed: 07.07.2018).

[22] R. M. Eustice, "Large-area visually augmented navigation for autonomous underwater vehicles," PhD thesis, Department of Ocean Engineering, Massachusetts Institute of Technology / Woods Hole Oceanographic Institution Joint Program, Cambridge, MA, USA, Jun. 2005.

[23] I. Mahon, S. B. Williams, O. Pizarro, and M. Johnson-Roberson, "Efficient view-based slam using visual loop closures," *IEEE Transactions on Robotics*, vol. 24, no. 5, pp. 1002–1014, Oct. 2008.

[24] A. Kim and R. M. Eustice, "Real-time visual slam for autonomous underwater hull inspection using visual saliency," *IEEE Transactions on Robotics*, vol. 29, no. 3, pp. 719–733, 2013.

[25] E. Iscar, K. A. Skinner, and M. Johnson-Roberson, "Multi-view 3d reconstruction in underwater environments: Evaluation and benchmark," in *OCEANS 2017 - Anchorage*, Sep. 2017, pp. 1–8.

[26] E. Iscar, C. Barbalata, N. Goumas, and M. Johnson-Roberson, "Towards low cost, deep water auv optical mapping," in *OCEANS 2018 MTS/IEEE Charleston*, IEEE, 2018, pp. 1–6.

[27] G. P. Kontoudis, M. V. Liarokapis, A. G. Zisimatos, C. I. Mavrogiannis, and K. J. Kyriakopoulos, "Open-source, anthropomorphic, underactuated robot hands with a selectively lockable differential mechanism: Towards affordable prostheses," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 5857–5862.

[28] H. Lim, J. Park, D. Lee, and H. J. Kim, "Build your own quadrotor: Open-source projects on unmanned aerial vehicles," *IEEE Robotics Automation Magazine*, vol. 19, no. 3, pp. 33–45, Sep. 2012.

[29] T. Estier, Y Crausaz, B. Merminod, M. Lauria, R. Piguet, and R. Siegwart, "An innovative space rover with extended climbing abilities," in *Robotics 2000*, 2000, pp. 333–339.

[30] M. Quigley, A. Asbeck, and A. Ng, “A low-cost compliant 7-dof robotic manipulator,” in *2011 IEEE International Conference on Robotics and Automation*, May 2011, pp. 6051–6058.

[31] G. Dynamics. (2017). Bluefin 21 datasheet, [Online]. Available: `https://gdmissionsystems.com/bluefinrobotics/vehicles-batteries-and-services/bluefin-21`. (accessed: 09.10.2017).

[32] C. L. Kaiser, D. R. Yoerger, J. C. Kinsey, S. Kelley, A. Billings, J. Fujii, S. Suman, M. Jakuba, Z. Berkowitz, C. R. German, *et al.*, “The design and 200 day per year operation of the autonomous underwater vehicle sentry,” in *Autonomous Underwater Vehicles (AUV), 2016 IEEE/OES*, IEEE, 2016, pp. 251–260.

[33] Kongsberg. (2018). Hughin4500, [Online]. Available: `https://www.kongsberg.com/maritime/products/marine-robotics/autonomous-underwater-vehicles/AUV-hugin/?OpenDocument`. (accessed: 11.07.2018).

[34] Hydroid. (2018). Remus6000, [Online]. Available: `https://www.hydroid.com/remus-6000-commercial-applications`. (accessed: 11.07.2018).

[35] WHOI. (2018). Sentry, [Online]. Available: `http://www.whoi.edu/main/sentry/specifications-sensors`. (accessed: 11.07.2018).

[36] ——, (2018). Nereus, [Online]. Available: `https://www.whoi.edu/main/nereus/specifications`. (accessed: 11.07.2018).

[37] M. Dunbabin, J. Roberts, K. Usher, G. Winstanley, and P. Corke, “A Hybrid AUV Design for Shallow Water Reef Navigation,” in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, Apr. 2005, pp. 2105–2110.

[38] E. A. d. Barros, L. O. Freire, and J. L. Dantas, “Development of the Pirajuba AUV,” *IFAC Proceedings Volumes*, vol. 43, no. 20, pp. 102–107, Sep. 2010.

[39] A. Alvarez, A. Caffaz, A. Caiti, G. Casalino, E. Clerici, F. Giorgi, L. Gualdesi, and A. Turetta, “Desing and realization of a very low cost prototypal autonomous vehicle for coastal oceanographic missions,” en, *IFAC Proceedings Volumes*, vol. 37, no. 10, pp. 471–476, Jul. 2004.

[40] C. Barngrover, R. Kastner, T. Denewiler, and G. Mills, “The stingray auv: A small and cost-effective solution for ecological monitoring,” in *OCEANS’11 MTS/IEEE KONA*, Sep. 2011, pp. 1–8.

[41] A. Hackbarth, E. Kreuzer, and E. Solowjow, “HippoCampus: A micro underwater vehicle for swarm applications,” in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 2258–2263.

[42] S. A. Team. (2018). S.o.n.i.a. auv, [Online]. Available: `http://sonia.etsmtl.ca/`. (accessed: 08.07.2018).

[43] C. Team. (2018). Cuauv, [Online]. Available: `http://cuauv.org/index.php`. (accessed: 08.07.2018).

[44] K. Vestgard, R. Hansen, B. Jalving, O. A. Pedersen, *et al.*, "The hugin 3000 survey auv," in *The Eleventh International Offshore and Polar Engineering Conference*, International Society of Offshore and Polar Engineers, 2001.

[45] T. J. Osse and C. C. Eriksen, "The deepglider: A full ocean depth glider for oceanographic research," in *OCEANS 2007*, IEEE, 2007, pp. 1–12.

[46] S Troisi, S Del Pizzo, S Gaglione, A Miccio, and R. Testa, "3d models comparison of complex shell in underwater and dry environments," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 5, p. 215, 2015.

[47] J. Henderson, O. Pizarro, M. Johnson-Roberson, and I. Mahon, "Mapping submerged archaeological sites using stereo-vision photogrammetry," *International Journal of Nautical Archaeology*, vol. 42, no. 2, pp. 243–256, 2013.

[48] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, Sep. 1999, 1150–1157 vol.2.

[49] Y. Furukawa and J. Ponce, "Accurate, dense, and robust multiview stereopsis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1362–1376, Aug. 2010.

[50] C. Wu, *Visualsfm: A visual structure from motion system*, http://ccwu.me/vsfm/, 2017.

[51] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal on Robotics and Automation*, vol. 3, no. 4, pp. 323–344, Aug. 1987.

[52] A. Jordt-Sedlazeck and R. Koch, "Refractive structure-from-motion on underwater imaowges," in *IEEE International Conference on Computer Vision*, 2013, pp. 57–64.

[53] A. Jordt, *Underwater 3D Reconstruction Based on Physical Models for Refraction and Underwater Light Propagation*, ser. Kiel Computer Science Series 2014/2. Department of Computer Science, CAU Kiel, 2014, Dissertation, Faculty of Engineering,Kiel University.

[54] M. Bryson, M. Johnson-Roberson, O. Pizarro, and S. B. Williams, "True color correction of autonomous underwater vehicle imagery," *Journal of Field Robotics*, vol. 33, no. 6, pp. 853–874, 2016.

[55] K. A. Skinner, E. I. Ruland, and M. Johnson-Roberson, "Automatic color correction for 3d reconstruction of underwater scenes," in *IEEE International Conference on Robotics and Automation*, 2017.

[56] V. E. Schmidt and Y. Rzhanov, "Measurement of micro-bathymetry with a gopro underwater stereo camera pair," in *Oceans, 2012*, IEEE, 2012, pp. 1–6.

[57] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010.

[58] T. Whelan, S. Leutenegger, R. F. Salas-Moreno, B. Glocker, and A. J. Davison, "ElasticFusion: Dense SLAM without a pose graph," in *Robotics: Science and Systems (RSS)*, Rome, Italy, Jul. 2015.

[59] P. Cignoni, M. Corsini, and G. Ranzuglia, "Meshlab: An open-source 3d mesh processing system," *Ercim news*, vol. 73, no. 45-46, p. 6, 2008.

[60] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 3, p. 29, 2013.

[61] S. Fuhrmann, F. Langguth, and M. Goesele, "Mve-a multi-view reconstruction environment," in *Eurographics Workshop on Graphics and Cultural Heritage*, The Eurographics Association, 2014, pp. 11–18.

[62] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (surf)," *Computer vision and image understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[63] M. Goesele, N. Snavely, B. Curless, H. Hoppe, and S. M. Seitz, "Multi-view stereo for community photo collections," in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, IEEE, 2007, pp. 1–8.

[64] S. Fuhrmann and M. Goesele, "Floating scale surface reconstruction," *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, p. 46, 2014.

[65] J. McCarthy and J. Benjamin, "Multi-image photogrammetry for underwater archaeological site recording: An accessible, diver-based approach," *Journal of maritime archaeology*, vol. 9, no. 1, pp. 95–114, 2014.

[66] C Balletti, C Beltrame, E Costa, F Guerra, and P Vernier, "Underwater photogrammetry and 3d reconstruction of marble cargos shipwreck," *The Interna-*

tional Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. 40, no. 5, p. 7, 2015.

[67] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*, International Society for Optics and Photonics, 1992, pp. 586–606.

[68] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the hausdorff distance," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 15, no. 9, pp. 850–863, 1993.

[69] R. Ferrari, W. F. Figueira, M. S. Pratchett, T. Boube, A. Adam, T. Kobelkowsky-Vidrio, S. S. Doo, T. B. Atwood, and M. Byrne, "3d photogrammetry quantifies growth and external erosion of individual coral colonies and skeletons," *Scientific reports*, vol. 7, no. 1, p. 16 737, 2017.

[70] A. Sedlazeck and R. Koch, "Perspective and non-perspective camera models in underwater imaging–overview and error analysis," in *Outdoor and Large-Scale Real-World Scene Analysis*, Springer, 2012, pp. 212–242.

[71] J.-M. Lavest, G. Rives, and J.-T. Lapresté, "Underwater camera calibration," in *European Conference on Computer Vision*, Springer, 2000, pp. 654–668.

[72] C. Kunz and H. Singh, "Hemispherical refraction and camera calibration in underwater vision," in *OCEANS 2008*, IEEE, 2008, pp. 1–7.

[73] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.

[74] M. Bewley, A. Friedman, R. Ferrari, N. Hill, R. Hovey, N. Barrett, E. M. Marzinelli, O. Pizarro, W. Figueira, L. Meyer, R. Babcock, L. Bellchambers, M. Byrne, and S. B. Williams, "Australian sea-floor survey data, with images and expert annotations," vol. 2, 150057 EP –, Oct. 27, 2015.

[75] M. Zhukovsky, V. Kuznetsov, and S. Olkhovsky, "Photogrammetric techniques for 3-d underwater record of the antique time ship from phanagoria," *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci*, vol. 40, pp. 717–721, 2013.

[76] A. Sedlazeck and R. Koch, "Calibration of housing parameters for underwater stereo-camera rigs.," in *BMVC*, Citeseer, 2011, pp. 1–11.

[77] A. Jordt-Sedlazeck and R. Koch, "Refractive calibration of underwater cameras," in *Computer Vision – ECCV 2012*, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 846–859.

[78] F. Menna, E. Nocerino, F. Fassi, and F. Remondino, "Geometric and optic characterization of a hemispherical dome port for underwater photogrammetry," *Sensors*, vol. 16, no. 1, p. 48, 2016.

[79] P. Sturm, S. Ramalingam, and S. Lodha, "On calibration, structure from motion and multi-view geometry for generic camera models," in *Imaging Beyond the Pinhole Camera*, Springer, 2006, pp. 87–105.

[80] P. Sturm and S. Ramalingam, "A generic concept for camera calibration," in *European Conference on Computer Vision*, Springer, 2004, pp. 1–13.

[81] A. Richardson, J. Strom, and E. Olson, "AprilCal: Assisted and repeatable camera calibration," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Nov. 2013.

[82] Sellmeier, "Zur erklärung der abnormen farbenfolge im spectrum einiger substanzen," *Annalen der Physik*, vol. 219, no. 6, pp. 272–282, 1871.

[83] B. Edlén, "The refractive index of air," *Metrologia*, vol. 2, no. 2, p. 71, 1966.

[84] R. W. Austin and G. Halikas, "The index of refraction of seawater," 1976.

[85] V. Lepetit, F.Moreno-Noguer, and P.Fua, "Epnp: An accurate o(n) solution to the pnp problem," *International Journal Computer Vision*, vol. 81, no. 2, 2009.

[86] M. Hirsch and B. Scholkopf, "Self-calibration of optical lenses," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 612–620.

[87] Y. Shih, B. Guenter, and N. Joshi, "Image enhancement using calibrated lens simulations," in *European Conference on Computer Vision*, Springer, 2012, pp. 42–56.

[88] P. Janout, J. Pospíšil, K. Fliegel, M. Klima, and P. Pata, "Interpolation methods for the improvement of the point spread function estimation," in *2018 28th International Conference Radioelektronika (RADIOELEKTRONIKA)*, Apr. 2018, pp. 1–5.

[89] P. T. Boggs and J. E. Rogers, "Orthogonal distance regression," in *Statistical analysis of measurement error models and applications (Arcata, CA, 1989)*, ser. Contemp. Math. Vol. 112, Amer. Math. Soc., Providence, RI, 1990, pp. 183–194.

[90] C. J. Schuler, M. Hirsch, S. Harmeling, and B. Schölkopf, "Non-stationary correction of optical aberrations," in *2011 International Conference on Computer Vision*, Nov. 2011, pp. 659–666.

[91] D. Sage, L. Donati, F. Soulez, D. Fortun, G. Schmit, A. Seitz, R. Guiet, C. Vonesch, and M. Unser, "Deconvolutionlab2: An open-source software for deconvolution microscopy," *Methods*, vol. 115, pp. 28–41, 2017.

[92] J. L. Pech-Pacheco, G. Cristóbal, J. Chamorro-Martinez, and J. Fernández-Valdivia, "Diatom autofocusing in brightfield microscopy: A comparative study," in *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*, IEEE, vol. 3, 2000, pp. 314–317.

[93] F. Repoulias and E. Papadopoulos, "Planar trajectory planning and tracking control design for underactuated auvs," *Ocean Engineering*, vol. 34, no. 11-12, pp. 1650–1667, 2007.

[94] D. Kruger, R. Stolkin, A. Blum, and J. Briganti, "Optimal auv path planning for extended missions in complex, fast-flowing estuarine environments," in *Robotics and Automation, 2007 IEEE International Conference on*, IEEE, 2007, pp. 4265–4270.

[95] D. P. Eickstedt and S. R. Sideleau, "The backseat control architecture for autonomous robotic vehicles: A case study with the iver2 auv," *Marine technology society journal*, vol. 44, no. 4, pp. 42–54, 2010.

[96] C. McGann, F. Py, K. Rajan, H. Thomas, R. Henthorn, and R. McEwen, "A deliberative architecture for auv control," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 1049–1054.

[97] T. R. Cuff and R. W. Wall, "Support platform and communications to manage cooperative auv operations," in *OCEANS 2006 - Asia Pacific*, May 2006, pp. 1–8.

[98] C. C. Sotzing, J. Evans, and D. M. Lane, "A multi-agent architecture to increase coordination efficiency in multi-auv operations," in *OCEANS 2007-Europe*, IEEE, 2007, pp. 1–6.

[99] P Baccou, B Jouvencel, V Creuze, and C Rabaud, "Cooperative positioning and navigation for multiple auv operations," in *OCEANS, 2001. MTS/IEEE conference and exhibition*, IEEE, vol. 3, 2001, pp. 1816–1821.

[100] M. Eichhorn, C. D. Williams, R. Bachmayer, and B. de Young, "A mission planning system for the auv "slocum glider" for the newfoundland and labrador shelf," in *OCEANS 2010 IEEE-Sydney*, IEEE, 2010, pp. 1–9.

[101] E. Fiorelli, N. E. Leonard, P. Bhatta, D. A. Paley, R. Bachmayer, and D. M. Fratantoni, "Multi-auv control and adaptive sampling in monterey bay," *IEEE Journal of Oceanic Engineering*, vol. 31, no. 4, pp. 935–948, Oct. 2006.

[102] C. Murphy, "Data quality monitoring with witness," in *2014 IEEE/OES Autonomous Underwater Vehicles (AUV)*, Sep. 2014, pp. 1–4.

[103] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, “Generation and Visualization of Large-scale Three-dimensional Reconstructions from Underwater Robotic Surveys,” *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010.

[104] D. J. Long and A. R. Baco, “Rapid change with depth in megabenthic structure-forming communities of the makapu’u deep-sea coral bed,” *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 99, pp. 158–168, 2014.

[105] F. Parrish, “Density and habitat of three deep-sea corals in the lower hawaiian chain,” *Bulletin of Marine Science*, vol. 81, pp. 185–194, Nov. 2007.

[106] E. I. Ruland and M. Johnson-Roberson, “Low cost underwater acoustic localization,” in *Acoustics ’17 Boston*, In press, Available on arXiv, 2017.